**Universidade de Brasília**
**Faculdade de Medicina**
**Laboratório de Interação Parasito-Hospedeiro**

# Diversidade Molecular e Funcional de Proteínas da Saliva de *Triatoma infestans*, um vetor da Doença de Chagas

**Teresa Cristina França de Assumpção**

**Orientador: Prof. Dr. Jaime Martins de Santana**
**UnB - Universidade de Brasília**

**Co-orientador: Dr. José Marcos Ribeiro**
**NIH - National Institutes of Health**

**Tese apresentada ao Programa de Pós-graduação em Patologia Molecular da Universidade de Brasília como requisito parcial à obtenção do Grau de Doutor em Patologia Molecular.**

**Brasília – DF**
**2007**

Trabalho desenvolvido no Laboratório de Interação Parasita-Hospedeiro, Faculdade de Medicina da Universidade de Brasília e no NIH (*National Institutes of Health – USA*), com apoio financeiro do CNPq e CAPES.

*“We must not forget that when radium was discovered no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it. It must be done for itself, for the beauty of science, and then there is always the chance that a scientific discovery may become like the radium, a benefit for humanity.”*

Marie Curie, *Lecture at Vassar College, May 14, 1921*
*French (Polish-born) chemist & physicist (1867 - 1934)*

Dedico este trabalho à minha mãe Abadia e minha irmã Virgínia.

Obrigada pelo amor, carinho, compreensão e incentivo.

# Agradecimentos

Ao Prof. **Jaime Martins de Santana**, pela confiança, orientação, apoio e dedicação na realização deste trabalho.

Ao Dr. **José Marcos Ribeiro**, pela oportunidade, incentivo e exemplo de dedicação ao trabalho científico.

Aos professores da Patologia Molecular que contribuíram para meu crescimento profissional.

Aos colegas de laboratório da UnB: **David, Flávia, Izabela, Danielle, Thiago, Meire, Keyla** e **Hugo**. Em especial, **Flávia e David**, pela amizade e pelos momentos compartilhados.

Aos colegas de laboratório do NIH: **Ivo, John, Eric, Anderson, Michail** e **Ben**. Aos colegas do segundo andar: **Lucinda, Ryan, Van, Clarissa, Régis, Fabiano, Nicolas, Abdoulave, Jennifer, Jesus, Carolina** e **Álvaro**. Obrigada pela acolhida e convivência.

À minha **mãe** e minha **irmã**, pela paciência, compreensão, estímulo e carinho. Ao meu **pai**, pelo apoio e incentivo.

Aos meus parentes que me incentivaram e apoiaram.

Aos meus amigos, em especial **Marcia, Silvia, Eliana, Gloria** e **Kênia**. Obrigada pelo estímulo e amizade.

Aos amigos de EUA: **Theresa, Felipe, Lucinda, Tatiana** e **Marife**. Em especial ao **Andreu**, pelo carinho, apoio e companheirismo.

# Lista de Abreviaturas

| | |
|---|---|
| 2DE | Eletroforese em gel bidimensional |
| AA | Aminoácido |
| BCIP | 5-bromo-4-cloro-3-indolil fosfato |
| DTT | Ditiotreitol |
| EDTA | Ácido etileno bis(oxi-etilenonitrilo) tetraacético |
| FLA$_2$ | Fosfolipase A$_2$ |
| FPLC | Cromatografia Líquida e Rápida de Proteínas |
| HPLC | Cromatografia Líquida de Alta Eficiência |
| IPTG | Isopropil-1-tio-β-D-galactopiranosídeo |
| ITC | Titulação Isotérmica por Calorimetria |
| NBT | p-nitro azul tetrazólio |
| PAF | Fator ativador de plaquetas |
| PAF-AH | PAF-acetil hidrolase |
| SDS-PAGE | Eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio |
| X-gal | 5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo |

# Abreviatura dos Aminoácidos

| | |
|---|---|
| A | Alanina |
| C | Cisteína |
| D | Ácido aspártico |
| E | Ácido glutâmico |
| F | Fenilalanina |
| G | Glicina |
| H | Histamina |
| I | Isoleucina |
| K | Lisina |
| L | Leucina |
| M | Metionina |
| N | Asparagina |
| P | Prolina |
| Q | Glutamina |
| R | Arginina |
| S | Serina |
| T | Treonina |
| V | Valina |
| W | Triptofano |
| Y | Tirosina |

# Índice

# Resumo

O *Triatoma infestans*, um dos vetores mais importantes da Doença de Chagas na América Latina, alimenta-se do sangue de vertebrados em todos os seus estágios – ninfas e adultos. As glândulas salivares de insetos hematófagos produzem compostos farmacologicamente potentes que impedem a hemostasia do hospedeiro, incluindo moléculas anticoagulantes, antiplaquetárias e vasodilatadoras. A saliva de *T. infestans* medeia a hidrólise de NDBC$_6$HPC, um substrato para PAF-AH (platelet-activating factor-acetilhidrolase), em pH neutro. A purificação dessa atividade foi obtida por cromatografia em FLPC utilizando colunas de troca catiônica e interação hidrofóbica. A atividade enzimática ótima foi descrita como independente de Ca$^{+2}$ e associada a uma proteína de 17 kDa (PAF-AH de *T. infestans*; PATi) em SDS-PAGE sob condições redutoras. Experimentos de espectrometria de massa sugerem que PATi é membro da família de fosfolipases A$_2$. Anticorpos específicos localizaram a enzima nas glândulas salivares D2. Esses dados sugerem que a hidrólise de PAF pelo inseto possa interferir nas respostas anti-hemostática e/ou nociceptiva do hospedeiro. Com o objetivo de compreender melhor a complexidade bioquímica e farmacológica deste inseto, uma biblioteca de cDNA de suas glândulas salivares foi seqüenciada. Proteínas salivares também foram submetidas à eletroforese bidimensional seguida de análise por espectrometria de massa. Neste trabalho, nós apresentamos a análise de um grupo de 1534 seqüências de cDNA das glândulas salivares, das quais 645 codificam para proteínas putativamente secretadas. A maioria das proteínas salivares descritas como lipocalinas – 55% da biblioteca de cDNA – coincidiram com seqüências peptídicas dos resultados proteômicos. Esperamos que a obtenção desses novos transcritos salivares possa auxiliar no esclarecimento da função de moléculas salivares nas interações entre o vetor e o hospedeiro e na descoberta de novos agentes farmacológicos.

# Summary

*Triatoma infestans* is one of the most important vectors of Chagas Disease in Latin America, feeding on vertebrate blood in all life stages. Hematophagous insects' salivary glands produce potent pharmacological compounds that counteract host hemostasis, including anti-clotting, anti-platelet, and vasodilatory molecules. The saliva of *T. infestans* mediates hydrolysis of $NDBC_6HPC$, a substrate for Platelet-activating factor-acetylhydrolase (PAF-AH), at neutral pH. Purification of this activity was achieved by a two-step FPLC procedure using cation exchange and hydrophobic columns. Optimal enzyme activity was found to be $Ca^{+2}$-independent and associated with a single 17-kDa protein (PAF-AH of *T. infestans*; PATi) on SDS-PAGE under reducing conditions. Results from mass spectrometry experiments suggest that PATi is a member of the phospholipase $A_2$ family. Specific antibodies localized the enzyme in the luminal content of the salivary glands D2. These findings suggest that hydrolysis of PAF may facilitate the insect to avoid host hemostatic and/or nociceptive responses. To obtain a further insight into the salivary biochemical and pharmacological complexity of this insect, a cDNA library was randomly sequenced. Also, salivary proteins were submitted to 2D gel electrophoresis followed by MS analysis. We present the analysis of a set of 1,534 salivary gland cDNA sequences, 645 of which coding for proteins of a putative secretory nature. Most salivary proteins described as lipocalins – 55% of the cDNA library – matched peptides sequences obtained from proteomic results. We expect this work will contribute with new salivary transcripts that could help the understanding of the role of salivary molecules in host/vector interactions and the discovery of novel pharmacologic agents.

# Introdução

# Introdução

## Doença de Chagas

A doença de Chagas foi descrita por Carlos Chagas em 1909, relatando suas características clínicas, anatomopatológicas e epidemiológicas, seu agente etiológico e vetor como um inseto da ordem Hemiptera (Brener *et alii*, 2000). É uma das patologias de mais ampla distribuição no Continente Americano; estima-se que 18 milhões de indivíduos estejam infectados e cerca de 100 milhões sob risco de contaminação (Dias *et alii*, 2002). A doença de Chagas é ainda hoje, no Brasil e em diversos países da América Latina, um problema grave de saúde pública. Segundo a OMS, constitui uma das principais causas de morte súbita que pode ocorrer com freqüência na fase mais produtiva da vida do indivíduo (Neves *et alii*, 2005). Por isso, a doença de Chagas representa um grande problema social e produz o maior ônus sócio-econômico entre as enfermidades denominadas tropicais. É a doença parasitária mais importante na América Latina em termos de seu impacto na economia regional e no sistema de saúde público (WHO 1991; Banco Mundial, 1993; Miles *et alii*, 2003).

A terapêutica da doença de Chagas continua parcialmente ineficaz, apesar dos grandes esforços que vêm sendo desenvolvidos por vários laboratórios e pesquisadores. Como também não há vacina disponível, a principal estratégia de controle da doença é a prevenção da transmissão do seu agente etiológico, principalmente por meio da eliminação dos insetos vetores domésticos e da infecção por transfusão sangüínea (Schofield *et alii*, 2006; Dias *et alii*, 2002).

## Insetos da Ordem Hemiptera

Compreendem a ordem Hemiptera os insetos com aparelho bucal – probóscida ou tromba – do tipo picador, sugador, que se origina anteriormente aos olhos, constituído por um par de mandíbulas e um de maxilas, envolvidos por um lábio tri ou tetrassegmentado. Hemípteros hematófagos que transmitem o *T. cruzi* pertencem à família Reduviidae e apresentam as seguintes características: cabeça alongada e mais ou menos fusiforme; pescoço nítido unindo a cabeça ao tórax; probóscida reta e trissegmentada. A subfamília Triatominae é constituída por seis tribos organizadas em 19 gêneros, contendo mais de 130 espécies. Das seis tribos, Rhodniini e Triatomini contêm as espécies mais importantes de vetores. As outras não apresentam espécies transmissoras do *T. cruzi* para humanos (Neves *et alii*, 2005).

## *Triatoma infestans*

O *Triatoma infestans*, família Reduviidae e subfamília Triatominae, é o principal inseto transmissor do protozoário *Trypanosoma cruzi* a dezenas de espécies de mamíferos. Sua localização estende-se do sul da Argentina à região nordeste do Brasil (Brener *et alii*, 2000). É espécie predominantemente domiciliar, colonizando-se em grande quantidade nas frestas das cafuas de barro e pau-a-pique. Formas tripomastigotas do parasita no sangue do hospedeiro vertebrado são, eventualmente, sugados por esses triatomíneos durante o repasto. No trato digestivo destes, sofrem diferenciação em epimastigotas replicativos que se diferenciam em formas tripomatigotas metacíclicas infectantes no intestino posterior, sendo eliminados nas fezes ou na urina. Durante o repasto sangüíneo seguinte, os triatomíneos infectados eliminam fezes contendo essas formas que penetram o hospedeiro via mucosa ou lesão da pele infligida ao hospedeiro no sítio da picada, dando continuidade ao ciclo de *T. cruzi* na natureza.

O *T. infestans* possui 3 pares de glândulas salivares bem diferenciadas: D1 – anterior ou principal, D2 – média ou acessória, e D3 – posterior ou suplementar (Barth, 1954; Lacombe, 1999) (Fig. 1). As glândulas salivares localizam-se na cavidade toráxica e contíguas à parte inicial do tubo digestivo, são constituídas por uma camada de células simples e estão conectadas a um hilo comum por meio de ductos (Brener *et alli*, 2000). De forma geral, as glândulas D1 e D2 possuem cor branca leitosa e um pouco amarelada, respectivamente. Já as glândulas D3 mostram-se translúcidas. As colorações das glândulas resultam da presença de secreção no lúmen (Lacombe, 1999). Considerando a hematofagia uma importante característica desses insetos, a presença de componentes como vasodilatadores, anticoagulantes e inibidores da agregação plaquetária em sua saliva é esperada.

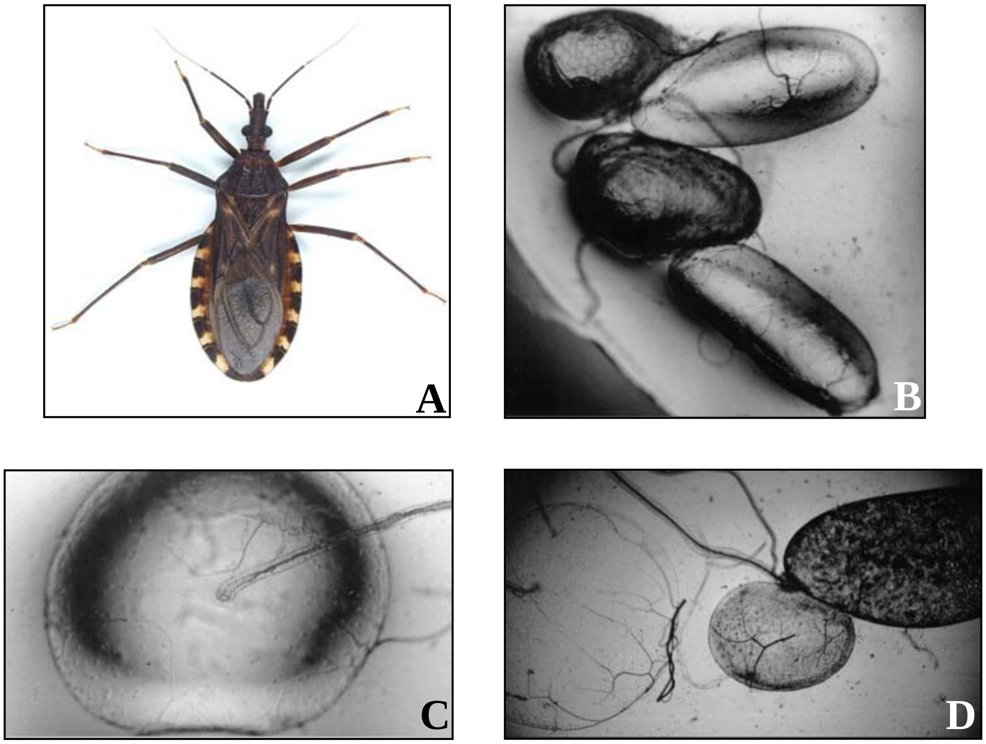


**Figura 1 – *Triatoma infestans* e suas glândulas salivares.** (A) *Triatoma infestans*, adulto. Microscopia, contraste de fase 250X, de glândulas salivares de *T. infestans* adultos: (B) pares D1 e D2, (C) glândula D3, (D) pares D1, D2 e D3. (Foto A: Paulo H. B. Leite; Fotos B, C e D: Teresa Cristina F. de Assumpção).

## Propriedades Anti-hemostáticas da Saliva de Insetos Hematófagos

### Hemostasia

A hemostasia é a resposta fisiológica do hospedeiro e um eficiente mecanismo de defesa que controla a perda de sangue após um dano vascular e é encontrada em todos os organimos que tem um sistema hemostático. Consiste na agregação plaquetária (formação do agregado de plaquetas), cascata de coagulação sangüínea (formação do coágulo sangüíneo) e vasoconstrição (redução do fluxo sangüíneo). Existem vários agonistas independentes para agregação plaquetária (ADP, colágeno, trombina, fator ativador de plaquetas – PAF, etc.) e pelo menos dois vasoconstritores liberados pelas plaquetas (tromboxano $A_2$ e serotonina). A cascata de coagulação é um sistema complexo com vários pontos de amplificação e controle (Ribeiro & Francischetti, 2003).

A elucidação de mecanismos evolutivos dos artrópodes hematófagos podem aumentar o entendimento de sistemas complexos encontrados na interface da hematofagia. As três vias do sistema hemostático são bem interconectadas, fazendo da hemostase um sistema redundante e aumentando o desafio aos insetos hematófagos, pois representam um obstáculo na tentativa de obter sangue do hospedeiro.

### Hematofagia

A hematofagia está presente em diferentes classes e tipos de animais que em sua maioria são invertebrados, incluindo sanguessugas e insetos. Algumas espécies de morcegos, mamíferos vertebrados, também são hematófagas (Basanova *et alii*, 2002). Os primeiros trabalhos sobre substâncias de animais hematófagos capazes de bloquear ou prolongar a coagulação sangüínea de vertebrados datam do século XIX (Moser *et alii*, 1998). Desde então, muitas substâncias de animais hematófagos têm sido descritas.

O hábito da hematofagia evoluiu independentemente entre várias espécies e gêneros de artrópodes hematófagos. A evolução de substâncias anti-hemostáticas que são injetadas junto com a saliva no hospedeiro, no momento do repasto sangüíneo, permitiu antagonizar a hemostase do hospedeiro vertebrado. Essas moléculas em conjunto com adaptações mecânicas do aparelho bucal do inseto, auxiliam na remoção e obtenção do sangue (Ribeiro, 1995).

A saliva de insetos possui substâncias com potentes propriedades farmacológicas que afetam diretamente os sistemas imunológico, inflamatório e hemostático do hospedeiro vertebrado (Ribeiro, 1995). Assim, a saliva afeta a fisiologia do hospedeiro localmente, no sítio da picada, provavelmente resultando em um ambiente favorável aos patógenos transmitidos pelo vetor, transformando a saliva desses insetos hematófagos em um alvo interessante para o controle da transmissão de doenças (Valenzuela, 2002c). Várias substâncias anti-hemostáticas da saliva de insetos hematófagos vetores de doenças têm sido caracterizadas molecular e funcionalmente, incluindo antitrombinas e inibidores do fator Xa da coagulação (Valenzuela *et alii*, 1999).

### Vasodilatadores

São moléculas que aumentam o fluxo sangüíneo mediante antagonismo de substâncias vasoconstritoras produzidas pelo sistema hemostático após injúria tissular provocada pelo aparelho bucal do inseto. Facilitam a alimentação, pois aumentam o calibre das veias sangüíneas, acelerando sua descoberta e o fluxo de sangue para o sítio da picada, logo menos tempo é necessário para a aquisição do sangue. Os vasodilatadores agem direta ou indiretamente em células musculares lisas ativando enzimas intracelulares como adenilato ciclase e guanilato ciclase que levam à formação de AMPc e GMPc, respectivamente (Rang *et alii*, 1997). A sialocinina é um pequeno peptídio vasodilatador, isolado da saliva do *Aedes aegypti*, que age diretamente sobre o endotélio ativando a produção de óxido nítrico (Champagne & Ribeiro, 1994) que ativa a guanilato ciclase em células musculares lisas, resultando na vasodilatação (Valenzuela, 2002c). Esse efeito facilita a localização de vasos e a obtenção de sangue pelo inseto. Em adição, vasodilatadores da saliva também podem facilitar a infecção. Por exemplo, Titus e Ribeiro (1988) demonstraram que a saliva de *Luztomyia longipalpis* aumenta a infecção de mamífero por *Leishmania major* quando o parasita é co-inoculado com saliva da mosca. Este efeito foi associado ao maxadilan, um potente vasodiladador presente na saliva desse inseto (Morris *et alii*, 2001). Isto indica que este parasita utiliza-se das propriedades farmacológicas da saliva para circular na natureza. Outro grupo de moléculas bem estudado é o das nitroforinas de *Rhodnius prolixus*, triatomíneo que também transmite o *T. cruzi* para mamíferos. As nitroforinas consistem em proteínas que contêm grupo heme e possuem cerca de 20 kDa. Sua atividade melhor caracterizada é o armazenamento e transporte de óxido nítrico (NO) que, ao ser liberado, liga-se à guanilato ciclase, resultando em relaxamento muscular e vasodilatação. Essas proteínas também inibem a resposta inflamatória do hospedeiro por interagirem com a histamina (Ribeiro & Walker, 1994; Montfort *et alii*, 2000). Em *T. infestans*, nenhuma molécula apresentando função vasodilatadora foi identificada ainda.

### Inibidores da Coagulação Sangüínea

A cascata de coagulação sangüínea consiste em uma série de serino-proteases que ativam umas às outras de forma seqüencial. A formação do coágulo é a última etapa de uma série de reações proteolíticas que, coordenadas com as plaquetas e células endoteliais, evitam a perda de sangue devido a um dano vascular (Fig. 2). As enzimas proteolíticas, fatores VII, IX, X, XI e trombina são normalmente encontradas na circulação na forma inativa. Sua ativação ocorre pela clivagem de uma ou duas ligações peptídicas (Goodman & Gilman, 1996). A via intrínseca da coagulação começa com a ativação do fator XII, induzida por colágeno, que ativa o fator XI e a calicreína plasmática. Calicreína cliva o cininogênio para formar bradicinina, um peptídio causador de inflamação e sensação de dor (Ribeiro, 1989). A via extrínseca começa com a liberação do fator tecidual (tromboplastina) de células endoteliais danificadas, que ativa o fator VII (Bevers *et alii*, 1993). As duas vias convergem para uma via comum resultando na formação de fator Xa que, por sua vez, ativa protrombina a trombina. Finalmente, o fibrinogênio é clivado pela trombina em fibrina, o principal componente do coágulo sangüíneo junto com as plaquetas e os eritrócitos (Davie *et alii*, 1991; Jackson & Nemerson, 1980).

Componentes anticoagulantes da saliva de artrópodes hematófagos agem especificamente em proteases ou complexos envolvidos na coagulação sangüínea como a trombina e o fator Xa, refletindo o papel central do fator X ou fator Xa nas vias intrínseca e extrínseca, e também a função da trombina na produção de fibrina a partir do fibrinogênio. Substâncias com propriedades antitrombina isoladas da saliva de insetos hematófagos já foram descritas. A anofelina é um peptídio com atividade antitrombina isolado das glândulas salivares do mosquito *Anopheles albimanus*. Seu gene foi clonado e o peptídio sintetizado, confirmando sua especificidade pela trombina, apesar de nenhuma similaridade com outras seqüências em bancos de dados ter sido encontrada (Valenzuela *et alii*, 1999). O mesmo grupo também caracterizou, a partir do homogeneizado das glândulas salivares de *Cimex lectularius*, uma proteína com massa molecular de 17 kDa como inibidor da ativação do fator X em fator Xa (Valenzuela *et alii*, 1996).

Extrato de glândula salivar de *T. infestans* prolongou o tempo de trombina, protrombina e de tromboplastina parcial ativada. Esse efeito anticoagulante da saliva de *T. infestans* observado na via intrínseca da coagulação ocorre principalmente pela interferência no fator VIII (Pereira *et alii*, 1996). Recentemente, Isawa e colaboradores (2007) identificaram duas proteínas nas glândulas salivares de *T. infestans* denominadas triafestina-1 e triafestina-2. Essas proteínas inibem a ativação do sistema calicreína-cinina do hospedeiro, resultando na inibição da liberação de bradicina. Esse sistema participa de respostas inflamatórias mediadas por superfície celular, orginadas após injúria tissular. Triafestina-1 e 2 poderiam atenuar a resposta inflamatória local no sítio da picada, diminuindo os sintomas inflamatórios (vermelhidão, edema e dor) e, provalvemente, facilitando ao inseto a obtenção do repasto sangüíneo (Isawa *et alii*, 2007).

Além dos artrópodes hematófagos, outros animais dependendo de uma alimentação sangüínea desenvolveram mecanismos que interferem com o processo de coagulação. Inibidores de coagulação também foram isolados de animais hematófagos como morcegos (Gardell *et alii*, 1991), sanguessugas (Sawyer, 1986; Sawyer, 1991) e nematódeos (Cappello *et alii*, 1995).

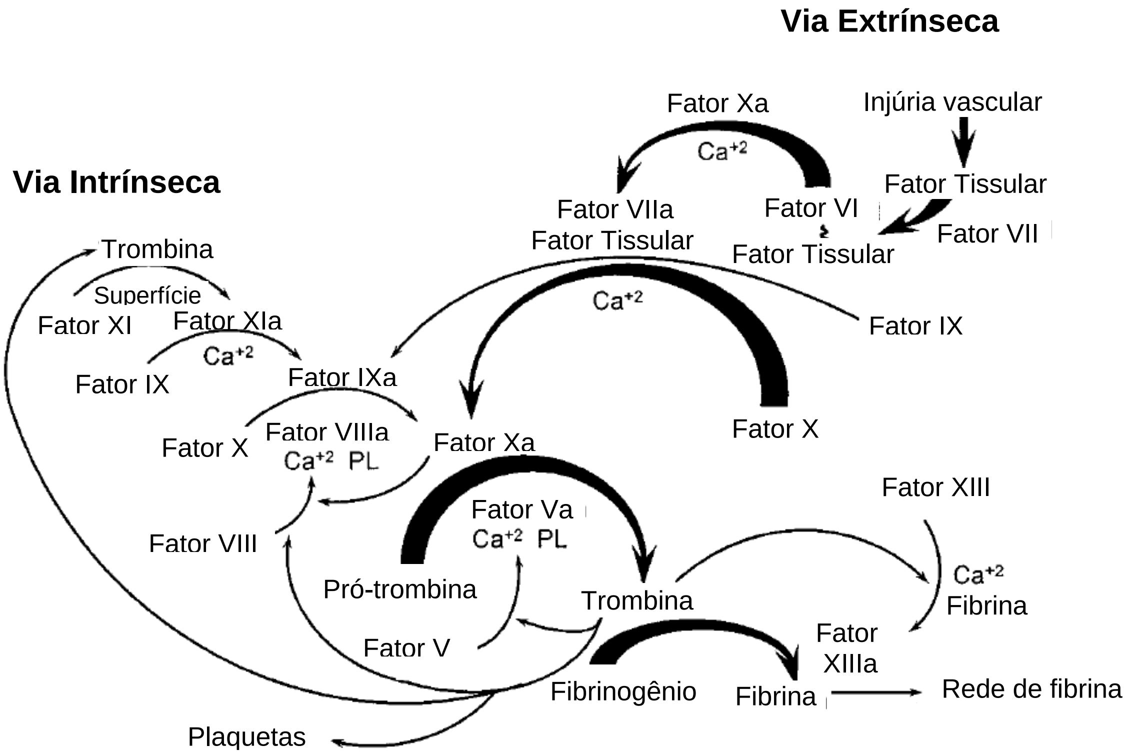


**Figura 2 – Desenho esquemático das vias intrínseca e extrínseca da coagulação sangüínea.** A iniciação da cascata de coagulação ocorre após injúria vascular e exposição do fator tecidual ao sangue. Isto desencadeia a via extrínseca (lado direito), em setas largas. A via intrínseca da coagulação pode ser ativada quando trombina é gerada, levando à ativação do fator XI. As duas vias convergem para a formação do fator Xa. Os fatores de coagulação ativados, exceto a trombina, são designados pela letra **a** minúscula, como IXa, Xa, XIa. PL refere-se a fosfolípidio (adaptado de Davie, 2003).

## Anti-agregadores de Plaquetas

Após dano vascular, as plaquetas são ativadas por vários agonistas como ADP, colágeno, trombina, tromboxano $A_2$, adrenalina, PAF e tromboplastina. Inicialmente, plaquetas ativadas agregam-se no local de injúria, formando um aglomerado celular que reduz ou bloqueia a perda de sangue (Davie *et alii*, 1991). Plaquetas na forma inativa possuem uma superfície lisa e uma forma discóide. A mudança de forma é acompanhada pela extensão de pseudópodes na superfície das plaquetas. A ativação e agregação inicial de plaquetas levam à secreção do conteúdo dos grânulos plaquetários que ativam outras plaquetas e induzem inflamação (Jamaluddin *et alii*, 1991).

Insetos hematófagos inibem a agregação plaquetária por diferentes mecanismos como inibição dos efeitos de trombina e colágeno sobre as plaquetas (Ribeiro, 1987) e hidrólise de PAF (Ribeiro & Francischetti, 2001). No entanto, a estratégia mais utilizada por esses animais para bloquear a agregação plaquetária parece ser por meio da hidrólise de ADP, um importante agonista da agregação, em AMP. Essa reação é catalizada por apirases, enzimas que removem o fosfato inorgânico de ATP e ADP, impedindo a agregação plaquetária induzida pelo ADP (Valenzuela, 2002c). Em invertebrados, a atividade apirásica está associada às glândulas salivares de artrópodes hematófagos. As apirases já foram descritas em *Aedes aegypti* (Champagne *et alii*, 1995b), *Anopheles gambiae* (Arcà *et alii*, 1999), *C. lectularius* (Valenzuela *et alii*, 1998) e outros insetos. A função da apirase em *Rhodnius* está relacionada com sua atividade anti-hemostática, facilitando a obtenção da alimentação sangüínea (Ribeiro & Garcia, 1981). A presença dessa enzima na saliva de *R. prolixus* foi caracterizada por Sarkis e colaboradores (1986). Cinco apirases salivares de *T. infestans* foram purificadas e caracterizadas (Faudry *et alii*, 2004). Charneau e colaboradores (2007) demonstraram que o proteoma da saliva de *T. infestans* contém principalmente inibidores de agregação plaquetária que pertencem às famílias de lipocalinas e apirases. A presença de isoformas de apirase mostra sua diversidade e abundância na saliva de *T. infestans*, diferentemente de outros insetos. Além dos triatomíneos, seqüências putativas codificando para apirases foram encontradas nos sialomas de vários mosquitos como *Ae. aegypti* (Ribeiro *et alii*, 2007), *Ae. albopictus* (Arcà *et alii*, 2007), *An. darlingi* (Calvo *et alii*, 2004), *An. stephensi* (Valenzuela *et alii*, 2003), *An. gambiae* (Francischetti *et alii*, 2002b), *An. funestus* (Calvo *et alii*, 2007a) e *Culex pipiens quinquefasciatus* (Ribeiro *et alii*, 2004b).

O colágeno é uma proteína de matriz extracelular que desempenha uma função importante no processo de hemostase, pois, a sua exposição no local de injúria vascular inicia o recrutamento e estimula a cascata de ativação das plaquetas circulantes, formando o trombo (Farndale *et alii*, 2004). Triplatinas 1 e 2 são duas proteínas salivares de *T. infestans* que inibem a agregação plaquetária induzida por colágeno. Essas proteínas apresentam similaridade à palidipina e acredita-se que sejam antagonistas de GPVI (glicoproteína VI – principal receptor de colágeno) (Morita *et alii*, 2006). Interações entre plaquetas e colágeno são importantes na formação do trombo e GPVI é um receptor de sinalização na superfície de plaquetas que age nessa via de ativação (Nieswandt & Watson, 2003). Recentemente, um membro da família de alérgenos de 30 kDa de *Ae. aegypti*, denominado aegyptina, foi caracterizado como um ligante específico de colágeno. Essa ligação ao colágeno interfere com sua interação com outros ligantes, principalmente o GPVI, inibindo a agregação e adesão plaquetárias (Calvo *et alii*, 2007b).

Lipocalinas constituem um grande grupo de moléculas presentes nas glândulas salivares de triatomíneos. São tipicamente proteínas extracelulares, de baixa massa molecular e compartilham algumas propriedades moleculares como ligação a moléculas pequenas, principalmente hidrofóbicas; ligação a receptores específicos de superfície; formação de complexos covalentes e não-covalentes com outras macromoléculas solúveis. Embora tenham sido classificadas principalmente como proteínas de transporte, está claro que os membros da família de lipocalinas exercem uma grande variedade de funções. Apesar das características e funções comuns, membros da família de lipocalinas têm sido definidos amplamente com base em similaridade estrutural ou de seqüência, compreendendo grande variedade de proteínas (Flower, 2000). Dentre as lipocalinas descritas como inbidoras de agregação plaquetária na saliva de triatomíneos encontramos:

**Palidipina.** É uma lipocalina de 19 kDa purificada da saliva de *T. pallidipennis* que inibe especificamente a agregação plaquetária induzida por colágeno. Nenhum efeito foi observado quando a agregação era induzida por trombina, ADP ou tromboxano $A_2$, mostrando sua especificidade (Noeske-Jungblut *et alii*, 1994).

**Triabina.** A saliva de *T. pallidipennis* inibe não somente a agregação plaquetária induzida por colágeno, mas também a agregação mediada pela trombina. Esse inibidor de trombina foi denominado triabina e forma um complexo com a trombina, causando o prolongamento do tempo de coagulação e do tempo de tromboplastina parcial ativada, além da inibição da agregação plaquetária induzida por trombina (Noeske-Jungblut *et alii*, 1995).

**RPAI-1.** Em *R. prolixus*, uma lipocalina nomeada RPAI-1 (*Rhodnius* platelet aggregation inhibitor 1) impede a agregação plaquetária por inibir a resposta das plaquetas a baixas concentrações de alguns agonistas como ADP, colágeno, trombina, convulxina e tromboxano $A_2$ (Francischetti *et alii*, 2000; Francischetti *et alii*, 2002a).

## Imunidade e Inflamação

Além da necessidade de superar os mecanismos hemostáticos do hospedeiro, os artrópodes hematófagos também precisam impedir suas respostas inflamatória e imune. A inflamação é a reação do hospedeiro à injúria tissular e/ou processo infeccioso e consiste em respostas como dor, hiperemia, calor e edema, resultantes da vasodilatação tissular e liberação de vários fatores com funções farmacológicas específicas. Células polimorfonucleadas e monócitos são importantes mediadores da inflamação. O ATP liberado pelas células ativa os neutrófilos que se acumulam e degranulam no local da inflamação. A trombina da cascata de coagulação sangüínea e outras moléculas pró-inflamatórias, como o fator ativador de plaquetas (PAF), também ativam neutrófilos que produzem prostaglandinas e o próprio PAF, amplificando o sinal (Ribeiro & Francischetti, 2003).

Os invertebrados não possuem uma resposta imune adaptativa e dependem de sistemas de imunidade inata para a sua defesa (Hoffmann *et alii*, 1999). Em insetos, o sistema de ativação de profenoloxidase é parte importante da defesa. Sua função é detectar e eliminar os patógenos invasores, assim como sintetizar melanina para o encapsulamento de patógenos. A forma ativa da enzima fenoloxidase é responsável pela formação de melanina e de intermediários altamente reativos e tóxicos. Com a ativação por proteólise limitada, a fenoloxidase catalisa as primeiras etapas de formação da melanina que encapsula o patógeno e previnindo ou retardando seu crescimento (Ratcliffe *et alii*, 1984; Söderhäll & Cerenius, 1998; Ashida & Brey, 1998). A ativação dessa enzima por meio de uma série de eventos regulados é desempenhada pelo sistema de ativação pró-fenoloxidase que consiste em proteínas capazes de se ligar a polissacarídeos e outros compostos associados a microorganimos. Todas as fenoloxidades de artrópodes já caracterizadas são sintetizadas como precursores inativos que tornam-se enzimaticamente ativos após proteólise por serino-proteases. A ativação de pró-fenoloxidases é mediada por uma cascata enzimática, e esse sistema seria semelhante ao sistema complemento dos vertebrados (Cerenius & Söderhäll, 2004).

Os peptídios antimicrobianos (AMP) são importantes moléculas efetoras no sistema de imunidade inata de insetos (Christophides *et alii*, 2004). Os principais AMPs encontrados em insetos incluem cecropinas, defensinas e peptídios com super-representação de alguns aminoácidos como aqueles ricos em histidina ou prolina. A família de defensinas é o grupo mais amplo de AMPs encontrados em insetos e outros invertebrados. As defensinas são peptídios catiônicos com massa molecular de 4 kDa, ricos em cisteínas e agem contra bactérias gram-positivas (Boman, 1995; Bulet *et alii*, 1999).

## Outras Moléculas da Saliva de Triatomíneos

Outras proteínas também foram descritas nas glândulas salivares de triatomíneos. Algumas dessas proteínas possuem papel importante na hematofagia do inseto, pois podem agir como fatores anti-hemostáticos ou participam de outros processos igualmente importantes como a imunidade inata do inseto.

**Procalina.** Como no momento do repasto sangüíneo os insetos injetam proteínas salivares no hospedeiro, a presença de alérgenos pode resultar em hipersensibilidade do tipo I em indivíduos sensibilizados. As reações anafiláticas mais freqüentes a insetos são atribuídas aos artrópodes da família Reduviidae (Edwards & Lynch, 1984). Paddock e colaboradores (2001) purificaram e identificaram o principal alérgeno das glândulas salivares de *Triatoma protacta*, uma proteína de 20 kDa denominada procalina, membro da família de lipocalinas.

**Sialidase.** Tal atividade enzimática foi identificada e caracterizada nas glândulas salivares de *T. infestans*. É liberada durante o repasto e sua provável função seria a remoção de ácido siálico de moléculas envolvidas na migração celular e na reação inflamatória. Como o ácido siálico participa de alguns processos envolvidos na hemostase, é possível que a sialidase liberada interfira na coagulação ou na agregação plaquetária (Amino *et alii*, 1998).

**Triapsina.** Trata-se de uma protease similar à tripsina liberada com a saliva de *T. infestans*. Está presente na glândula salivar D2 como um precursor inativo (Amino *et alii*, 2001). Esta serino-protease poderia estar envolvida em eventos proteolíticos específicos afetando a coagulação ou a cascata de complemento do hospedeiro. Também poderia estar relacionado com a imunidade, pois as enzimas ativadoras de profenol-oxidase são serino-proteases (Söderhäll & Cerenius, 1998).

**Trialisina.** É uma proteína lítica formadora de poros encontrada na saliva de *T. infestans*. Essa proteína de 22 kDa foi nomeada trialisina por ser capaz de lisar parasitos protozoários e bactérias, indicando uma possível função no controle do crescimento de microorganismos nas glândulas salivares (Amino *et alii*, 2002). A expressão da trialisina recombinante e sua modelagem molecular foram obtidas por Corrêa (2002). A proteína recombinante mostrou efeito citolítico sobre *Escherichia coli*, *T. cruzi*, *Leishmania donovani* e células murinas da linhagem L6. Essas observações sugerem que essa proteína possa fazer parte da imunidade inata do inseto, pois possíveis microorganismos presentes no sangue do hospedeiro seriam lisados, protegendo o inseto de infecção. Concomitantemente, poderia desempenhar uma função anti-hemostática devido à lise de células mamíferas, facilitando a obtenção do repasto.

**Fosfolipases.** A superfamília das fosfolipases ($FLA_2s$) consiste em um amplo espectro de enzimas definidas por sua capacidade de catalisar a hidrólise da ligação éster de fosfolipídios. Os produtos da hidrólise são ácidos graxos livres e lisofosfolipídios. Os lisofosfolipídios são importantes na sinalização celular e no remodelamento de outros fosfolipídios (Six & Dennis, 2000). Algumas fosfolipases têm sido descritas em insetos, mas ainda estão pouco caracterizadas. Uma atividade de $FLA_2$ dependente de cálcio foi identificada na saliva e nas glândulas salivares do carrapato *Amblyomma americanum* (L.) (Bowman *et alii*, 1997). Outro grupo identificou uma fosfolipase C com especificidade por PAF, nomeada PAF-fosforilcolina hidrolase, encontrada na saliva e nas glândulas salivares do mosquito *Culex quinquefasciatus*. A atividade enzimática foi demonstrada pela capacidade de inibir a agregação plaquetária induzida por PAF e pelo consumo do substrato e formação do produto diacil glicerídeo (Ribeiro & Francischetti, 2001).

## Fator Ativador de Plaquetas – PAF

O fator ativador de plaquetas (1-O-alquil-2-acetil-*sn*-glicero-3-fosfocolina, PAF) é um potente mediador biológico que exerce seu efeito em várias células e tecidos. O PAF é um fosfolipídio único em sua função de mediador intracelular e pode ser sintetizado por duas vias distintas: a de remodelamento e a via *de novo*. A primeira está envolvida no remodelamento de fosfolipídios da membrana celular pela hidrólise de um araquidonato a partir da posição *sn*-2 e sua substituição por um acetato. Esta parece ser mais importante em várias respostas inflamatórias e alérgicas. A síntese *de novo* ocorre a partir de 1-O-alquil-*sn*-glicero-3-fosfato através da incorporação de acetato, remoção do fosfato e sua substituição por fosfocolina (Venable *et alii*, 1993).

O PAF não é armazenado nas células, mas é sintetizado em resposta a um estímulo que pode ser reações antígeno-anticorpo ou vários agentes como peptídios quimiotáticos, trombina, colágeno e alguns autacóides, inclusive o próprio PAF. É sintetizado por plaquetas, neutrófilos, monócitos, mastócitos, eosinófilos, células da medula renal e células endoteliais vasculares (Goodman & Gilman, 1996).

As ações intracelulares do PAF são mediadas pela interação com seu receptor (PAFR) que é expresso na superfície de vários tipos celulares. O PAFR pertence à superfamília dos receptores acoplados à proteína G, com sete domínios transmembrânicos (Prescott *et alii*, 2000).

### Propriedades Farmacológicas do PAF

O PAF é um vasodilatador potente que reduz a resistência vascular e a pressão sangüínea sistêmica, quando injetado por via endovenosa. Aumenta a permeabilidade vascular e facilita o movimento dos líquidos para fora deste sistema. (McManus *et alii*, 1981). É um potente estimulador da agregação plaquetária *in vitro*, que é acompanhada da liberação do tromboxano $A_2$ e do conteúdo granular das plaquetas, logo sua ação é independente da presença do tromboxano $A_2$ e de outros agentes agregantes. É um fator quimiotático para eosinófilos, neutrófilos e monócitos, também estimula a adesão dos neutrófilos às células endoteliais e sua diapedese (Goodman & Gilman, 1996). Ainda, o PAF também possui efeitos fisiopatológicos nos sistemas respiratório, gastrointestinal e nervoso (Kuijpers *et alii*, 2001).

### Respostas Inflamatórias e Alérgicas

O PAF exerce efeitos pró-inflamatórios como aumento da permeabilidade vascular, hiperalgesia, edema e infiltração por neutrófilos. A bioatividade potente dessa molécula está associada à sua habilidade de ativar neutrófilos em concentrações picomolares e de induzir quimiotaxia e polimerização de actina em concentrações nanomolares (Kuijpers *et alii*, 2001). A concentração plasmática desse fator está aumentada no choque anafilático experimental (Goodman & Gilman, 1996).

## PAF-Acetil hidrolase

A potência e a natureza de seus efeitos sugerem que tanto a síntese como a degradação do PAF devem ser processos rigorosamente controlados. O PAF produzido por qualquer uma das duas vias é degradado a um produto inativo, o liso-PAF, pelas PAF-acetil hidrolases (PAF-AH; 1-alquil-2-acetil-glicerofosfocolina esterase – EC 3.1.1.47; figura 3) que são enzimas pertencentes à família das fosfolipases $A_2$ (Prescott *et* *alii*, 2000). Esta atividade, encontrada em várias células, e tecidos não requer cálcio e catalisa a hidrólise de análogos acila de PAF bem como fosfolipídios contendo grupos *sn*-2 como ácidos graxos fragmentados e oxidados (Venable *et alii*, 1993).

Os fosfolipídios oxidados também possuem pequenos grupos acila na posição *sn*-2 do glicerol, mas são derivados da oxidação de ácidos graxos poliinsaturados. Aparentemente, esses compostos mimetizam a estrutura do PAF a ponto de se ligarem ao seu receptor e provocarem as mesmas respostas. A principal diferença entre o PAF e esses fosfolipídos oxidados é que a síntese do PAF é altamente regulada, enquanto os fosfolipídios oxidados são produzidos de maneira não regulada (Stafforini *et alii*, 1997).

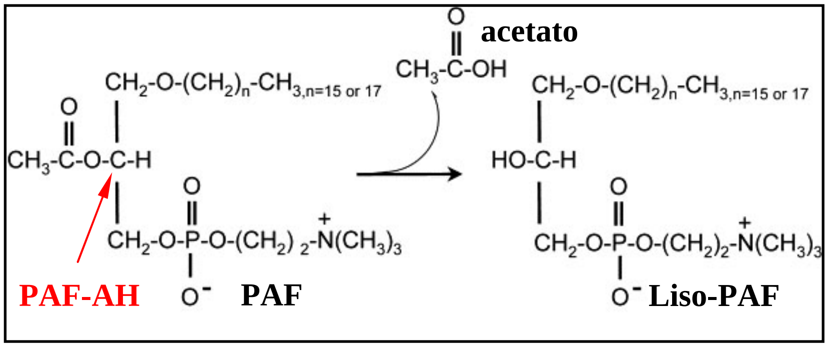


**Figura 3 – Esquema da degradação de PAF pela PAF-acetil hidrolase.** O PAF é degradado a liso-PAF pela PAF-AH através da remoção do grupo acetato na posição *sn*-2, gerando liso-PAF e acetato.

As PAF-AHs humanas representam um grupo único de $FLA_2$s que contém quatro enzimas que exibem especificidade incomum por substratos como PAF e fosfolipídios oxidados. De acordo com a nomenclatura das $FLA_2$s, as PAF-AHs são classificadas como $FLA_2$s dos grupos VII e VIII (Murakami & Kudo, 2002). Essas enzimas ainda podem ser divididas em duas subclasses: intracelulares, encontradas no citossol; e extracelulares, secretadas no plasma sangüíneo ou outros fluidos corporais (Derewenda & Ho, 1999).

A atividade da enzima PAF-AH no plasma e nas células tem idêntica especificidade pelo substrato, mas estudos de massa molecular, inibição química, inativação de protease e reconhecimento por anticorpos têm mostrado que estas enzimas são distintas. A fonte celular da PAF-AH plasmática é provavelmente os macrófagos e hepatócitos, pois ambos sintetizam e secretam uma atividade com propriedades idênticas à enzima plasmática. A secreção da enzima é independente de partículas de lipoproteínas, mas a acetilhidrolase se associa preferencialmente com lipoproteínas do meio (Venable *et alii*, 1993).

# Justificativa

# Justificativa

O conhecimento sobre proteínas da saliva de insetos hematófagos é importante para a melhor compreensão da atuação dessas no processo de alimentação e de sua possível função na transmissão do parasita. A biblioteca de cDNA é uma ferramenta importante para varredura e identificação de genes codificantes de proteínas com atividade relacionada à nossa linha de pesquisa. O seqüenciamento da biblioteca fornece dados suficientes para realização de busca em banco de dados visando encontrar proteínas similares com função conhecida. Também, genes poderão ser clonados a partir da biblioteca de cDNA para estudos funcionais. Além das lipocalinas, proteínas abundantes nas glândulas salivares de triatomíneos, outras classes de proteínas também poderiam desempenhar importante papel como as fosfolipases. Golodne demonstrou a presença de fosfolipídios na saliva de *R. prolixus* assim como propriedades anti-hemostáticas de lisofosfatidilcolina salivar (2003). Uma fosfolipase seria a enzima responsável não só pela geração de fosfolipídios na saliva do inseto mas também pela hidrólise de PAF, um potente mediador da inflamação e estimulador da agregação plaquetária.

Em trabalhos anteriores, a construção de bibliotecas de glândulas salivares foi bem sucedida para muitos artrópodes como *Aedes aegypti* (Valenzuela *et alii*, 2002b), *Anopheles gambiae* (Francischetti *et alii*, 2002b), *Ixodes scapularis* (Valenzuela *et alii*, 2002a), *Anopheles stephensi* (Valenzuela *et alii*, 2003), *Rhodnius prolixus* (Ribeiro *et alii*, 2004a), *Culex quinquefasciatus* (Ribeiro *et alii*, 2004b), *Anopheles darlingi* (Calvo *et alii*, 2004), *Aedes albopictus* (Arcà *et alii*, 2007), *Anopheles funestus* (Calvo *et alii*, 2007a) e *T. brasiliensis* (Santos *et alii,* 2007). Cada transcriptoma origina um banco de dados que pode ser utilizado na busca por seqüências similares. O conhecimento sobre estrutura e função biológica de componentes da saliva de insetos vetores é base para o planejamento de novas estratégias de combate a várias doenças tropicais ainda incuráveis como doença de Chagas, Malária e Leishmaniose.

A análise proteômica, por meio de eletroforese bidimensional das proteínas da saliva e posterior espectrometria de massa, proporciona a identificação de novas proteínas e permite validar os dados obtidos pelo transcriptoma.

A descoberta de novas proteínas que atuam no antagonismo da hemostase também é de grande interesse biotecnológico, pois essas moléculas poderiam ser utilizadas no tratamento de enfermidades como coagulopatias ou diretamente relacionadas à agregação plaquetária.

Esta tese é composta por dois manuscritos e alguns experimentos adicionais. O primeiro manuscrito intitula-se "A PAF-acetylhydrolase activity from the saliva of *Triatoma infestans*". O segundo manuscrito descreve a construção da biblioteca de cDNA de glândulas salivares de *T. infestans* e sua análise transcriptômica, tendo como título: "An insight into the sialome of the blood-sucking bug *Triatoma infestans*, a vector of Chagas' Disease".

# Objetivos

# Objetivos

O objetivo geral desta linha de pesquisa em entomologia molecular é conhecer as características moleculares e funcionais de proteínas da saliva de *T. infestans* relacionadas com o repasto sangüíneo. Esse conhecimento servirá de base para outros estudos visando a inibição de algumas dessas atividades encontradas na saliva do inseto, desfavorecendo o ciclo de vida do inseto e/ou a transmissão do *T. cruzi*.

Os objetivos específicos propostos para este trabalho são:

- Identificar e caracterizar atividade hidrolítica de PAF na saliva de *T. infestans*;
- Obtenção de uma biblioteca de cDNA das glândulas salivares de *T. infestans*;
- Análise transcriptômica das glândulas salivares de *T. infestans* por meio da biblioteca de cDNA;
- Validação da análise transcriptômica por meio de proteoma da saliva de *T. infestans*.

# Manuscrito I

# A PAF-acetylhydrolase activity from the saliva of *Triatoma infestans*

Assumpção, T.C.F.[1], Motta, F.S.N.[1], Lozzi, S.P.[2], Teixeira, A.R.L.[2], Sousa, M.V.[3], Santana, J.M.[1]

Address: [1] Laboratory of Host-Parasite Interface, University of Brasília, Brasília - DF, 70.910-900, Brazil, [2] Laboratory of Chagas Disease, University of Brasília, Brasília - DF, 70.910-900, Brazil, and [3] Brazilian Center for Protein Research, Department of Cell Biology, University of Brasília, Brasília - DF, 70.910-900, Brazil.

E-mail: Teresa C. F. Assumpção - tecris@unb.br; Flávia S.N. Motta - flavianader@gmail.com; Silene P. Lozzi - silene@unb.br; Antonio R. L. Teixeira - ateixeir@unb.br; Marcelo V. Sousa - mvsousa@unb.br; Jaime M. Santana - jsantana@unb.br.

Key words: Saliva, *T. infestans*, phospholipase.

## Abstract

Salivary anti-hemostatic activities are widely distributed in hematophagous arthropods including *Triatoma infestans* (Hemiptera, family Reduviidae, subfamily Triatominae), a vector of Chagas' disease. The saliva of *T. infestans* mediates hydrolysis of NDBC$_6$HPC, a substrate for Platelet-activating factor-acetylhydrolase (PAF-AH), at neutral pH. Purification of the protein responsible for this activity was achieved by a two-step FPLC procedure using cation exchange and hydrophobic columns. Optimal enzyme activity was found to be Ca$^{+2}$-independent and was associated with a single 17-kDa protein (PAF-AH of *T. infestans*; PATi) on SDS-PAGE under reducing conditions. Mass spectrometry experiments suggest that PATi is a member of the phospholipase A$_2$ family. Specific antibodies localized the enzyme in the luminal content of the salivary glands D2. These findings suggest that hydrolysis of PAF may facilitate the insect to avoid host hemostatic and/or nociceptive responses.

## Introduction

The hemiptera *Triatoma infestans*, a vector of Chagas' disease (American trypanosomiasis), feeds exclusively on vertebrate blood in all life stages. Hematophagous insects' salivary glands show a variety of anti-hemostatic compounds, capable to counteract host hemostasis, including anti-clotting, anti-platelet, and vasodilatory molecules, thus helping the bug to obtain its blood meal (Ribeiro and Francischetti, 2003; Ribeiro, 1995). Besides the *T. infestans* salivary apyrases already known for their ability to remove inorganic phosphate from ATP and ADP, preventing platelet aggregation (Faudry et al., 2004), other molecules may mediate inhibition of platelet aggregation through different pathways. Host platelet-activating factor (PAF) could be an interesting target for an arthropod anti-hemostatic enzyme, since PAF is related to inflammation and platelet aggregation.

PAF (1-O-alkyl-2-acetyl-sn-glycero-3-phosphocholine) is a bioactive phospholipid involved in inflammatory reactions (Prescott et al., 2000). It is synthesized by a wide range of inflammatory and non-inflammatory cells (Venable et al., 1993; Snyder, 1995), and has been implicated in both pathological and physiological processes (Venable et al., 1993). PAF is produced by two independent pathways: the remodeling one involves structural modification of a membrane lipid by replacement of the acyl moiety for an acetate group; the other route consists of *de novo* synthesis of PAF from an O-alkyl analogue of a lysophosphatidic acid (Snyder, 1990).

PAF synthesized by any of the two pathways is cleaved to an inactive product – lyso-PAF – by the PAF-acetylhydrolases (PAF-AH; 1-alkyl-2-acetyl-glycerophosphocoline esterase – EC 3.1.1.47), $Ca^{+2}$ independent enzymes belonging to group VII of the phospholipase $A_2$ ($PLA_2$) family (Six and Dennis, 2000). This inactivation of PAF occurs through the hydrolysis of the acetyl group at the *sn*-2 position of the molecule. The $PLA_2$ superfamily consists in a wide range of enzymes defined by their ability to catalyze hydrolysis of the ester bond of phospholipid substrates, resulting in free fatty acids and lysophospholipids. The released fatty acids are important source of energy and act as second messengers and precursors of eicosanoids, which are potent mediators of inflammation and signal transduction. The lysophospholipids are important in cell signaling and in other phospholipids remodeling (Six and Dennis, 2000).

Some phospholipases with activity upon PAF have been described in insects but only a few have been characterized. An activity of calcium-dependent $PLA_2$ was identified in saliva and salivary glands of the tick *Amblyomma americanum* (L.) (Bowman et al., 1997). Another group identified a phospholipase C with specificity for PAF, named PAF-phosphorylcoline hydrolase, found in saliva and salivary glands of the mosquito *Culex quinquefasciatus* (Ribeiro and Francischetti, 2001). Also, a PAF-AH activity was identified in salivary glands of *Ctenocephalides felis* (Cheeseman et al., 2001).

In this study, we examined the PAF hydrolytic activity present in saliva of *T. infestans*, utilizing a specific fluorogenic substrate. We report the identification and partial characterization of such an activity mediated by a 17-kDa Platelet-activating factor-acetylhydrolase of *T. infestans* (PATi). Its activity and molecular properties lead us to consider PATi a member of the phospholipase $A_2$ family. We postulate that PATi may play a role in insect feeding by conteracting host hemostatic and/or nociceptive responses.

## Material and methods

### 2.1. Triatomines and Collection of Saliva and Salivary Glands

*Triatoma infestans* were reared in an insectary room kept at 27 $^{o}$C ± 1.0 °C, with a relative humidity of 70% ± 5.0% and a 16:8 h light:dark photoperiod. The insects were fed on chickens every two weeks. The saliva was obtained from adult insects by spontaneous ejection 1 week after feeding, used immediately or kept at -20 $^{o}$C until use. The luminal content of D1, D2 or D3 salivary gland subunit was carefully collected by syringe puncture and the soluble material obtained upon centrifugation at 4 °C.

### 2.2. Enzymatic Assay

Saliva enzymatic activity was determined by measuring the fluorescence released by hydrolysis of the fluorogenic PAF-acetylhydrolase substrate 2-(6-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amino) hexanoyl-1-hexadecanoyl-*sn*-glycero-3-phosphocholine ($NBDC_6$-HPC; Molecular Probes). Assays were performed by incubating 1.0 μL of saliva or 40 μL of its fractions for 60 min at room temperature in 50 μL of reaction buffer (10 mM Tris-HCl pH 7.5; 10 mM EDTA) in the presence of 10 μM of substrate. The enzymatic activity of PAF-AH from freshly collected saliva, salivary gland extracts, or from purified and partially purified saliva was determined. After incubation at room temperature for 1 h, the emitted fluorescence of free NBD released by the enzymatic reaction was immediately measured at 535 nm on excitation at 475 nm in a fluorescence spectrophotometer (Hitachi F-2000).

### 2.3. Enzyme purification

Twenty-five microlitres of saliva were diluted in 500 μL of 25 mM sodium phosphate ($Na_2HPO_4$) buffer, pH 7.5 and centrifuged for 10 min at 14,000 x g. The supernatant was chromatographed in a Mono S HR 5/5 column (Pharmacia) previously equilibrated with the same buffer. The proteins were eluted with a linear gradient up to 1.0 M NaCl in the equilibration buffer. Fractions were collected and tested for enzymatic activity. Two peaks with activity were obtained. Each activity peak was pooled and applied separately into a hydrophobic interaction Phenyl Superose column (Pharmacia) equilibrated with 25 mM $Na_2HPO_4$ pH 7.5, containing 1.0 M $(NH_4)_2SO_4$. The proteins were eluted with a linear gradient of the above buffer to 25 mM $Na_2HPO_4$, pH 7.5. Fractions were collected and tested for enzymatic activity. The fractions containing activity were concentrated using a Centricon 10 (Amicon) filter to 200 μL. All purification steps were performed using a fast protein liquid chromatography (FPLC) system (Pharmacia) at room temperature. The purity of the preparation was determined by 15% SDS-PAGE (Laemmli, 1970), followed by Coomassie Brilliant Blue or Silver staining. The method of Bradford was used to determine protein concentration (Bradford, 1976). Bovine serum albumin was used as standard.

### 2.4. Gel Electrophoresis and Western Blot Analysis

Samples were boiled in sample buffer in the presence of dithiothreitol for 5 min and electrophoresed on a 15% SDS-polyacrylamide gel according to Laemmli (1970). Protein staining was performed with Coomassie Brilliant Blue or Silver staining.

### 2.5. Antibody Preparation

Purified enzyme (15 μg) was emulsified in Freund's complete adjuvant and injected into a rabbit or mice. Two booster injections, 15 μg in Freund's incomplete adjuvant and without adjuvant, respectively, were given after 3 and 6 weeks. For immunoblotting, the proteins were transferred onto nitrocellulose membrane, probed with polyclonal antibodies raised against native PATi, treated with alkaline phosphatase-linked anti-rabbit IgG, and visualized with NBT/BCIP (Nitro Blue Tetrazolium / Bromo-chloro-indolyl-phosphate).

### 2.7. Mass Spectrometry

Following electrophoresis and Coomassie Blue staining, the 17-kDa band was excised from the gel, washed, and analyzed as described previously (Shevchenko et al., 1996). Briefly, the protein was digested with trypsin (12.5 ng/µL; Promega) in 50 mM $NH_4HCO_3$, 5 mM $CaCl_2$ buffer, for 12 h at 37 $^{o}$C. The resulted peptides were applied to a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry Reflex IV (Bruker Daltonics) in the reflectron mode. The spectra were acquired using the program Flexcontrol 3 (ion source 1 = 20 kV, ion source 2 = 16.35 kV, lens = 9.8 kV, and reflector = 23 kV; pulse ion extraction = 200 ns, and matrix suppression =500 Da) from Bruker Daltonics. The peaks m/z 824.5021 and m/z 2210.0968, originated from trypsin auto-proteolysis, were used for internal calibration. The identification of the protein was performed using the programs BioTool 2.0 (Bruker daltonics) and Mascot (Matrix Science), available at www.matrixscience.com. The database used was SwissProt 41.4x (186207 sequences, 93837565 residues), with mass tolerance of 0.1 Da.

## Results

### Identification of an enzymatic activity in *T. infestans* saliva

Saliva of *T. infestans* readily hydrolyses the synthetic fluorogenic PAF-acetylhydrolase substrate NBDC$_6$-HPC (Fig. 1). To determine whether this activity is differentially expressed by the salivary gland subunits, the content (equal volume) of each one was assayed for enzymatic activity on NBDC$_6$-HPC. Under the conditions of this experiment, D2 salivary gland subunit expresses about 50% of the activity, whereas D1 and D3 express 17 and 33%, respectively. The detected enzymatic activity values were similar whether the assay was performed in the presence or absence of calcium. The enzyme mediating this activity was named PAF-acetylhydrolase of *T. infestans* (PATi) to indicate its activity.

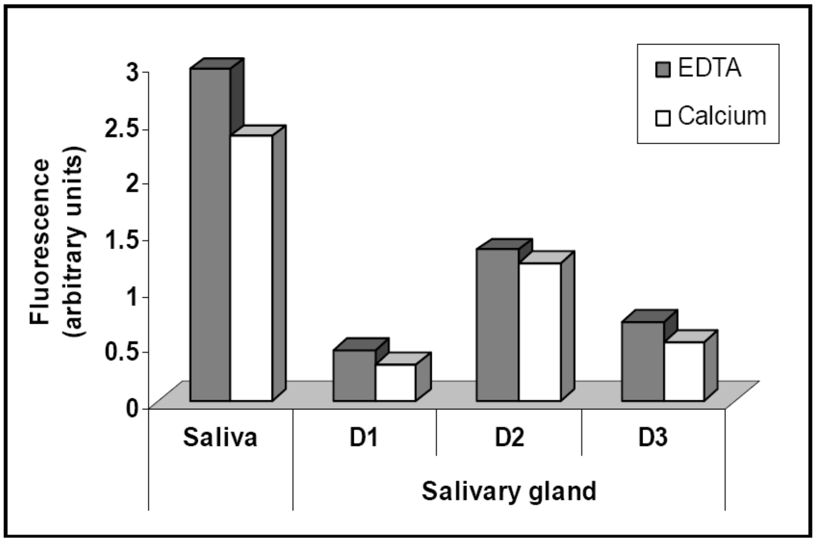


**Figure 1 – *Triatoma infestans* saliva hydrolyses PAF-acetylhydrolase substrate**. Saliva (1.0 μL) freshly collected and 2.0 μL of salivary glands extracts (D1, D2, D3) were incubated for 1 h at room temperature with the fluorogenic substrate NBDC$_6$-HPC, at a final concentration of 10 μM, into the reaction buffer (10 mM Tris-HCl pH 7.4), with 10 mM EDTA (■) or 2 mM CaCl$_2$ (☐). The fluorescence emitted was immediately measured in a fluorescence spectrophotometer at 535 nm, after excitation at 475 nm.

### PATi purification

To further characterize PATi, it was purified from saliva by a combination of ion-exchange and hydrophobic interaction chromatography. Two peaks of enzymatic activity were eluted from the Mono S column: a major peak (A) from 280 to 440 mM NaCl, and a minor one (B) from 800 to 840 mM NaCl (Fig 2A). Each activity peak was pooled and applied separately into a Phenyl Superose column. The elution profile of peak A shows a single peak of activity eluted from 200 to 0 mM of $(NH_4)_2SO_4$ (Fig. 2B). In contrast, we could not detect enzymatic activity when peak B was submitted to this column. After chromatography in both columns, we obtained a purified preparation to visualize in the gel. A single band of 17 kDa was obtained from one of the fractions, and could be visualized in a silver stained gel under reducing conditions (Fig. 3A, lane 5). Polyclonal antibodies raised against purified PATi recognized a single band at the expected size in saliva upon immunobloting (Fig. 3B), indicating that the protein is not degraded or had any modification during the purification process. This result shows that PATi is immunogenic in rabbit.

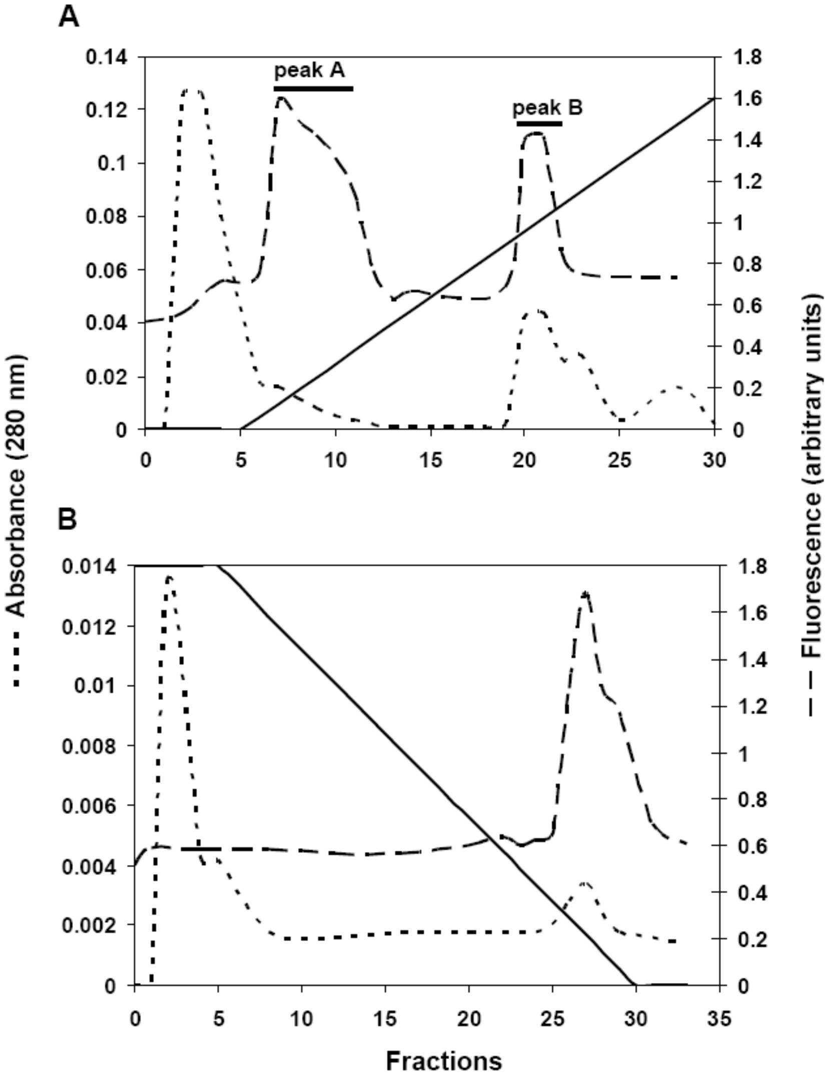


**Figure 2 – PATi purification.** Twenty-five microlitres of saliva were diluted in the buffer and the insoluble material was removed by centrifugation. **(A)** The supernatant was chromatographed in a cation exchange Mono S column and eluted with a linear gradient to NaCl 1.0 M. The first peak of fluorescent activity (pool A) was collected and applied in a Phenyl Superose column **(B)**. The bound material was eluted with a decreasing linear gradient of $(NH_4)_2SO_4$ 1.0 M. The dotted lines represent the absorbance at 280 nm, the dashed lines represent the fluorescent activity using NBDC$_6$-HPC, and the solid lines are the salt gradients.

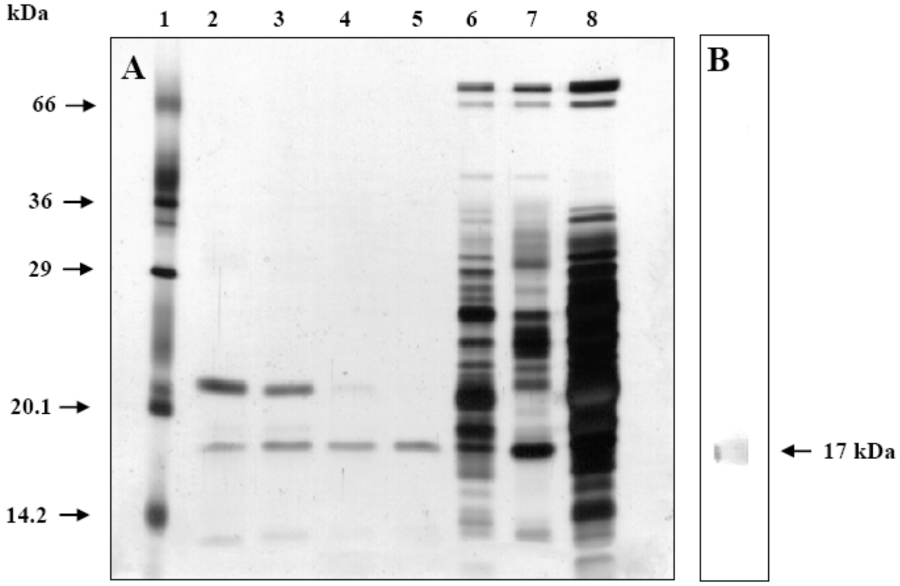


**Figure 3 – (A) SDS-PAGE analysis of the purified enzyme.** Lane 1, molecular mass markers; lanes 2 to 5, fractions with enzymatic activity after purification on both Mono S and Phenyl Superose; lane 6, only Mono S; lane 7, only Phenyl Superose; lane 8, salivary proteins of *Triatoma infestans*. Proteins were visualized by silver staining.

**(B) Western blot.** Purified PATi was submitted to 15% SDS-PAGE, transferred to a nitrocellulose membrane and probed with antibody raised against the purified protein.

### PATi is mainly stored in D2 salivary gland

To determine the localization of the enzyme in salivary glands of the insect, the proteins of each subunit were resolved in a SDS-PAGE (Fig. 4A), transferred to nitrocellulose membranes and probed with anti-PATi antibodies. The antibodies recognized the enzyme in D2 and in a much lesser extent in the D1 and D3 salivary gland subunits (Fig. 4B). No antigen was revealed with pre-immune serum (data not shown). This result correlates well with that obtained upon enzymatic assay shown in Figure 1. In conjunction, these results demonstrated that PATi is mainly stored in the lumen of D2 gland of *T. infestans*.

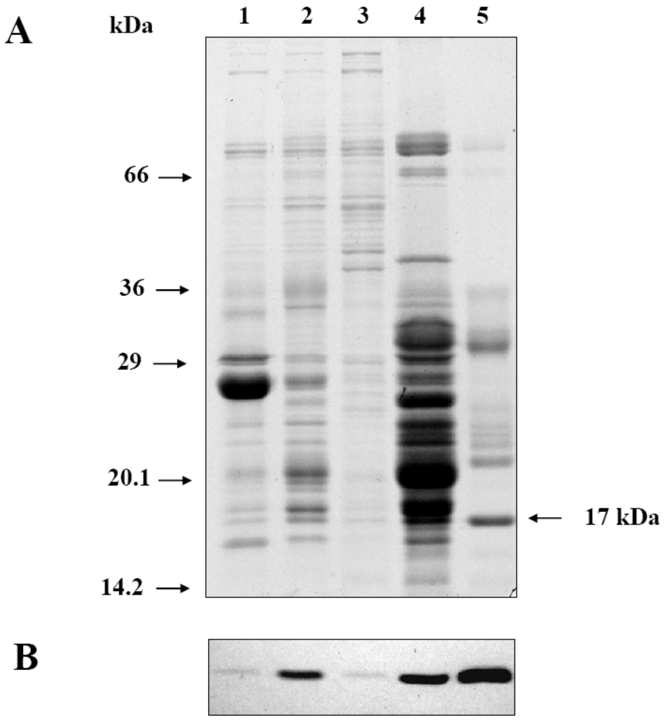


**Figure 4 – (A) SDS-PAGE 15% analysis of the salivary glands and saliva proteins of *T. infestans*.** Lane 1, proteins of salivary gland D1; lane 2, proteins of salivary gland D2; lane 3, proteins of salivary gland D3; lane 4, salivary proteins of *Triatoma infestans*; lane 5, fractions with enzymatic activity after partial purification on Phenyl Superose. The bands of proteins were visualized by Coomassie blue staining.

**(B) *Western blot*.** Salivary glands and saliva proteins (replica of gel A) were submitted to 15% SDS-PAGE, transferred to a nitrocellulose membrane and probed with the antibody raised against the purified protein.

### MS identification

To further identify the purified enzyme mediating $NBDC_6$-HPC hydrolysis, the protein was excised from the gel and digested with trypsin. The peptides obtained were eluted and submitted to mass spectrometry analysis in MALDI-TOF. The identification of the protein was performed using the programs BioTool 2.0 (Bruker daltonics) and Mascot (Matrix Science). The database used was SwissProt 41.4x. The higher score obtained was 78 for the protein Q9PVF4, the phospholipase $A_2$ precursor W6D49 (EC 3.1.1.4 – phosphatidylcholine 2-acylhydrolase) of *Callosellasma rhodostoma*. The purified protein from *T. infestans* saliva was confirmed as an enzyme that shows similarity to members of $PLA_2$ family. The sequences and value of masses of the peptides of phosphatidylcholine 2-acylhydrolase similar to peptides of PATi are present in Table 1. These peptides represent 32% of the enzyme (Fig. 5).

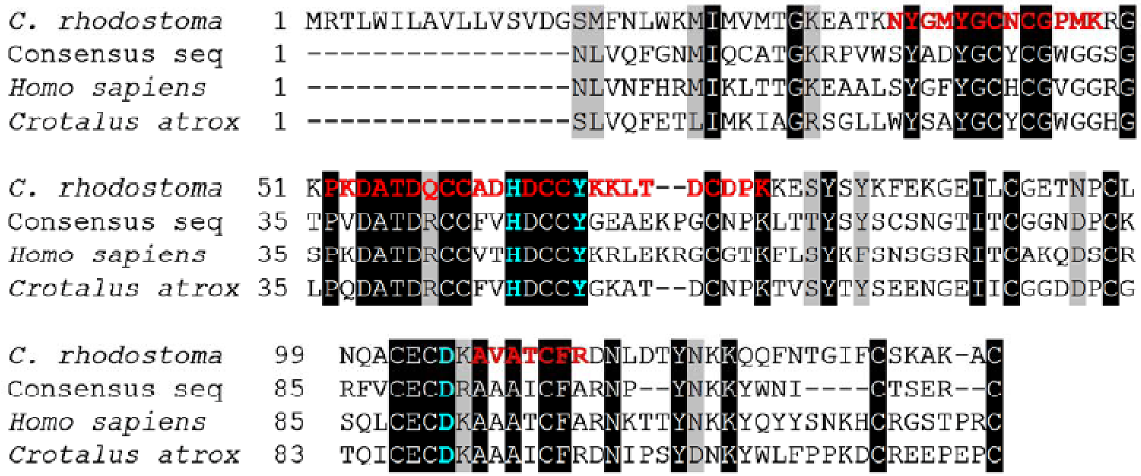


**Figure 5 – Alignment of sequences of different phospholipases $A_2$.** The amino acid sequences of *Callosellasma rhodostoma*, *Homo sapiens*, *Crotalus atrox* and the consensus sequence of a $PLA_2$ domain were aligned with the program CLUSTAL W. The amino acids marked in black show 80% identity and those in gray show 80% similarity. The peptides of the purified protein from *T. infestans* saliva, obtained by mass spectrometry, are in red. The amino acids in blue are in the catalytic site.

## Discussion

The identification of a PAF-AH activity in *T. infestans* saliva was performed with the fluorogenic substrate $NBDC_6$-HPC, which was used in other studies because of its ability to differentiate between PAF-AH and $PLA_2$ activities (Kitsiouli *et al.*, 1999). This feature was our point of start for this work, since we wanted to know if there was an enzyme in *T. infestans* with PAF-AH activity. Thus, PATi was considered a PAF-AH and not a classic $PLA_2$ because of its calcium independent characteristic to hydrolyze the substrate $NBDC_6$-HPC.

**Table 1 – Values of the masses and respective sequences of the peptides obtained after digestion of PAF-AH with trypsin**. The peptides obtained from the digestion with trypsin were submitted to mass spectrometry. The table shows the sequence of the peptides found after search in database SwissProt 41.4x (186207 sequences, 93837565 residues), with the values of the relative masses expected and calculated.

| **Start - End** | **Observed *** | **Mr expected*** | **Mr calculated *** | **Delta** | **Miss** | **Sequence** |
|---|---|---|---|---|---|---|
| 36 – 48 | 1551.57 | 1550.56 | 1550.58 | -0.01 | 0 | NYGMYGCNCGPMK |
| 36 – 48 | 1583.77 | 1582.76 | 1582.57 | 0.19 | 0 | NYGMYGCNCGPMK<br>2 Oxidations (M) |
| 52 – 68 | 2143.83 | 2142.82 | 2142.79 | 0.03 | 1 | PKDATDQCCADHDCCYK |
| 54 – 68 | 1918.77 | 1917.76 | 1917.64 | 0.12 | 0 | DATDQCCADHDCCYK |
| 54 – 69 | 2046.82 | 2045.82 | 2045.73 | 0.08 | 1 | DATDQCCADHDCCYKK |
| 70 – 76 | 848.27 | 847.26 | 847.37 | -0.11 | 0 | LTDCDPK |
| 107 – 113 | 824.21 | 823.20 | 823.40 | -0.20 | 0 | AVATCFR |

(*) Masses expressed in daltons

The size of the carbonic chain in the *sn*-2 position of phospholipids is a parameter for the specificity of enzymes demonstrating affinity for these substrates. PAF-AHs have preference for phospholipids with a short acyl chain at the *sn*-2 position, with no more than 9 carbons, independently of $Ca^{+2}$. Differently, classical phospholipase $A_2$ enzymes cleave carbonic chains up to 20 carbons, like the arachidonic acid in a $Ca^{+2}$-dependent manner (Stafforini *et al.*, 1997). The literature considers that 10 mM EDTA is enough to chelate calcium, and for instance, to inhibit the $PLA_2$ activity (Kitsiouli *et al.*, 1999). Even though EDTA in this concentration is able to chelate other divalent ions besides $Ca^{+2}$, there are no reports of some $PLA_2$ and/or PAF-AH that have any activity dependent on these cofactors. Based on these concepts and taking into account that $NBDC_6$-HPC is a well known substrate for PAF-AH enzymes (Kitsiouli *et al.*, 1999), we named the identified activity as PATi (PAF-AH of *T. infestans*), even though we do not have direct evidence as the activity upon the PAF molecule itself.

PATi was purified by a two-step chromatography procedure using a combination of cationic and hydrophobic columns. The peak B eluted from Mono S column did not show any detectable enzymatic activity after chromatography in Phenyl Superose. The protein could have been altered by chromatography or the ammonium sulfate concentration used induced an irreversible modification of its native structure. The data suggest that *T. infestans* saliva mediates two distinct activities upon the used substrate. This could be due to a different protein or an isoform of that one found in the peak A. There is a possibility that this phenomenon - polymorphism - could also be observed in phospholipases. Despite the low yield of the purification, this protocol was preferred and used because other attempts of purification were not successful. The yield of enzyme purification did not let us perform functional tests such as platelet aggregation assay.

The mass spectrometry analysis after trypsin digestion led to the identification of seven peptides. The search in database SwissProt 41.4x revealed that the masses of the peptides were coincident with peptides of a *Callosellasma rhodostoma* $PLA_2$. *C. rhodostoma* is an asian snake of medical importance. Its venom causes local effects and systemic hemorrhage, and contains moderate levels of $PLA_2$. Ten sequences of different $PLA_2$ were cloned from the cDNA library of the venom gland of *C. rhodostoma*. The major $PLA_2$ from this venom is designed CRV-W6D49 for the presences of Trp6 and Asp49 that can be replaced by other amino acid residues in other $PLA_2$ of the same animal (Tsai *et al.*, 2000). Differently from other $PLA_2$s, CRV-W6D49 does not seem to be active upon known substrate but only, and weakly, upon a pseudo substrate in the presence of $Ca^{+2}$ (Cho *et al.*, 1988); the function would be related with induction of edema through an unknown mechanism.

The snake venom $PLA_2$s have a variety of pharmacologically properties like pre-synaptic neurotoxicity, miotoxicity, induction of edema, hemolytic, anticoagulant and anti-platelet aggregation activities. For example, it was described the presence of three types of $PLA_2$ in other snake venom, *Agkistrodon halys pallas*: acidic, basic and neutral $PLA_2$, according to their predicted isoelectric points. The acidic $PLA_2$s show ability to inhibit platelet aggregation while the basic ones have hemolytic activity (Wang *et al.*, 1996). These functions could be related to the salivary PAF-AH of *T. infestans*, since

PAF hydrolysis would indirectly inhibit the platelet aggregation and also the local inflammation.

The enzyme PAF-AH together with other salivary proteins, probably participate in the insect anti-hemostatic response. Inactivation of PAF by PATi would help the insect to obtain the blood meal and to modulate host immune response through inhibition of local inflammatory reactions. The decrease in local inflammatory response would also facilitate the infection by the parasite, since it would find a more favorable environment at the bite site to invade host cells. $PLA_2$s can also display lytic activity on cells such as platelets, erythrocytes, and cells from the immune system (Hanahan, 1971), thus helping the insect to avoid hemostatic mechanisms and to digest the blood. In fact, hemolytic domain-containing proteins were found in *T. infestans* transcriptome (Assumpção et al., 2007). Another consequence of PAF hydrolysis by insect saliva would be the decrease of the host nociceptive response. The reduction of the pain elicited by the bite would facilitate the insect to obtain blood meal. This possibility is based on the fact that PAF antagonists decrease the hyperalgesy in mice (Teather *et al.*, 2002).

The substances present in the saliva of hematophagous insects that facilitate the acquisition of blood meal and/or affect host immune response, favoring the circulation of parasites in nature, could become a new target for vaccines against vector-borne diseases (Ribeiro & Francischetti, 2003). So, the protein PATi could be used together with other proteins previously described in *T. infestans* saliva as a target for immune prophylaxis of Chagas disease through the interruption of *T. cruzi* transmission to mammalian hosts.

## References

Assumpção, T.C.F., Francischetti, I.M.B., Andersen, J.F., Schwarz, A., Santana, J.M., Ribeiro, J.M.C., 2007. An insight into the sialome of the blood sucking bug *Triatoma infestans*, a vector of Chagas' disease. Insect Biochem. Mol. Biol., in press.

Bowman, A.S., Surdick, M.R., Zhu, K., Essenberg, R.C., Sauer, J.R., Dillwith, J.W., 1997. A novel phospholipase $A_2$ activity in saliva of the lone star tick, *Amblyomma americanum* (L.). Exp. Parasitol. 87, 121-132.

Bradford, M.M., 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254.

Cheeseman, M.T., Bates, P.A., Crampton, J.M., 2001. Preliminary characterization of esterase and platelet-activating factor (PAF)-acetylhydrolase activities from cat flea (*Ctenocephalides felis*) salivary glands. Insect Biochem. Mol. Biol. 31, 157-164.

Cho, W.H., Markowitz, M.A., Ketzdy, F.J., 1988. A new class of phospholipase $A_2$ substrates: kinetics of the phospholipase $A_2$ catalyzed hydrolysis of 3-(acyloxy)-4-nitrobenzoic acids. J. Am. Chem. Soc. 110, 5166-5171.

Faudry, E., Lozzi, S.P., Santana, J.M., D’Souza-Ault, M., Kieffer, S., Felix, C.R., Ricart, C.A., Sousa, M.V., Vernet, T., Teixeira, A.R., 2004. *Triatoma infestans* apyrases belong to the 5’-nucleotidase family. J. Biol. Chem. 279(19), 19607-19613.

Hanahan, D.J., 1971. Phospholipases. The enzymes, V, P.D. Boyer ed. Academic Press, New York, 71-85.

Kitsiouli, E.I., Nakos, G., Lekka, M.E., 1999. Differential determination of phospholipase $A_2$ and PAF-acetylhydrolase in biological fluids using fluorescent substrates. J. Lipid Res. 40, 2346-2356.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Prescott, S.M., Zimmerman, G.A., Stafforini, D.M., McIntyre, T.M., 2000. Platelet-activating factor and related lipid mediators. Annu. Rev. Biochem. 69, 419-445.

Ribeiro, J.M.C., 1995. Blood-feeding arthropods: Live syringes or invertebrate pharmacologists? Infect. Agents Dis. 4, 143-152.

Ribeiro, J.M.C., Francischetti, I.M.B., 2001. Platelet-activating-factor-hydrolyzing phospholipase C in the salivary glands and saliva of the mosquito *Culex quinquefasciatus*. J. Exp. Biol. 204, 3887-3894.

Ribeiro, J.M.C., Francischetti, I.M.B., 2003. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu. Rev. Entomol. 48, 73-88.

Shevchenko, A., Wilm, M., Vorm, O., Mann, M., 1996. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. Anal. Chem. 68, 850-858.

Six, D.A., Dennis, E.A., 2000. The expanding superfamily of phospholipase $A_2$ enzymes: classification and characterization. Biochem. Biophys. Acta 1488, 1-19.

Snyder, F., 1990. Platelet-activating factor and related acetylated lipids as potent biologically active cellular mediators. Am. J. Physiol. 259, C697-708.

Snyder, F., 1995. Platelet-activating factor and its analogs: metabolic pathways and related intracellular processes. Biochim. Biophys. Acta 1254, 231-249.

Stafforini, D.M., McIntyre, T.M., Zimmerman, G.A., Prescott, S.M., 1997. Platelet-activating factor acetylhydrolases. J. Biol. Chem. 272(29), 17895-17898.

Teather, L.A., Magnusson, J.E., Wurtman, R.J., 2002. Platelet-activating factors antagonists decrease the inflammatory nociceptive response in rats. Psychopharmacology 163, 430-433.

Tsai, I.-H., Wang, Y.-M., Au, L.-C., Ko, T.-P., Chen, Y.-H., Chu, Y.-F., 2000. Phospholipases $A_2$ from *Callosellasma rhodostoma* venom gland. Cloning and sequencing of 10 of the cDNAs, three-dimensional modelling and chemical modification of the major isozyme. Eur. J. Biochem. 267, 6684-6691.

Venable, M.E., Zimmerman, G.A., McIntyre, T.M., Prescott, S.M., 1993. Platelet-activating factor: a phospholipid autacoid with diverse actions. J. Lipid Res. 34, 691-702.

Wang, X.-Q., Yang, J., Gui, L.-L., Lin, Z.-J., Chen, Y.-C., Zhou, Y.-C., 1996. Crystal structure of an acidic phospholipase $A_2$ from the venom of *Agkistrodon halys pallas* at 2.0 Å resolution. J. Mol. Biol. 255, 669-676.

# Manuscrito II

**(Aceito para publicação no periódico**
**Insect Biochemistry and Molecular Biology)**

# An insight into the sialome of the blood-sucking bug *Triatoma infestans*, a vector of Chagas' disease

Teresa C. F. Assumpção [a], Ivo M. B. Francischetti [c], John F. Andersen [c], Alexandra Schwarz [b], Jaime M. Santana [a], José M. C. Ribeiro [c,*].

[a] Laboratory of Host-Parasite Interface, University of Brasília, Brasília – DF, 70.910-900, Brazil,

[b] Department of Special Zoology, Ruhr-University, D-44780 Bochum, Germany, and

[c] Vector Biology Section, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Rockville, MD 20852, USA

E-mail: Teresa C F Assumpção – tecris@unb.br; Ivo M B Francischetti – ifrancischetti@niaid.nih.gov; John F Andersen – jandersen@niaid.nih.gov; Alexandra Schwarz - a.schwarz@abdn.ac.uk; Jaime M Santana - jsantana@unb.br; José M C Ribeiro - jribeiro@niaid.nih.gov.

* Corresponding author. Tel.: +1 301496 9389; fax: +1 301 480 2571

*E-mail address*: jribeiro@niaid.nih.gov (J.M.C. Ribeiro).

*Abbreviations*: aa, amino acid; AMP, antimicrobial peptides; EST, expressed sequence tags; OBP, odorant-binding protein; H, putative housekeeping transcripts; S, putative secreted transcripts; U, unknown function transcripts; SG, salivary glands; Ti, *Triatoma infestans*; 2D, two dimensional.

## Abstract

*Triatoma infestans* is a hemiptera, vector of Chagas' disease, that feeds exclusively on vertebrate blood in all life stages. Hematophagous insects' salivary glands (SG) produce potent pharmacological compounds that counteract host hemostasis, including anti-clotting, anti-platelet, and vasodilatory molecules. To obtain a further insight into the salivary biochemical and pharmacological complexity of this insect, a cDNA library from its salivary glands was randomly sequenced. Also, salivary proteins were submitted to two dimensional gel (2D-gel) electrophoresis followed by MS analysis. We present the analysis of a set of 1,534 (SG) cDNA sequences, 645 of which coded for proteins of a putative secretory nature. Most salivary proteins described as lipocalins matched peptide sequences obtained from proteomic results.

*Key words:* Hematophagy, Saliva, Transcriptome, *Triatoma infestans*, Feeding, Sialome.

## 1. Introduction

*Triatoma infestans* (Hemiptera: Reduviidae) is an important vector of *Trypanosoma cruzi*, a protozoan parasite and etiological agent of Chagas' disease (American trypanosomiasis) in Latin America (Dias, 1987). All instar nymphs and adults are hematophagic and need a blood meal to molt and for oviposition. The insect obtains the blood meal by injecting its maxilla into vertebrate's skin searching for a vessel (Lavoipierre, 1965).

The (SG)s of blood-feeding arthropods show a variety of anti-hemostatic compounds that help the bug to obtain its blood meal. Like other blood-sucking arthropods that have been studied (Ribeiro and Francischetti, 2003), *T. infestans* is capable of counteracting host hemostatic responses triggered to prevent blood loss following tissue injury, such as vasoconstriction, blood coagulation and platelet aggregation (Ribeiro, 1995). The molecular diversity of hematophagous insect saliva represents a rich field for the discovery of novel pharmacologically active compounds and for understanding the evolutionary mechanisms leading to the insect's adaptation to this feeding habit. Previous studies describing the sialome (set of RNA message + set of proteins found in (SG)s) of hematophagous insects and ticks (Francischetti et al., 2002; Valenzuela et al., 2002a, b) have revealed that the sialomes of these disease vectors are more complex than expected and contain many proteins to which we can not yet ascribe a function.

Lipocalins are a large and heterogenous group of proteins that play various roles, mainly as carriers of small ligands in vertebrates and invertebrates (Flower et al., 2000). A great array of (SG) proteins belonging to the lipocalin family has generated a large number of different molecules having anti-hemostatic functions while maintaining the fundamental structure of the protein fold (Montfort et al., 2000). Lipocalins were found in the saliva of other blood-sucking insects such as *Rhodnius prolixus* (Ribeiro et al., 2004a) and *Triatoma brasiliensis* (Santos et al., 2007), and also in tick saliva (Paesen et al., 2000). In *R. prolixus*, three types of salivary lipocalins have been characterized: the nitrophorins consisting in a group of lipocalins working as NO carrier and also as anti-clotting; the ADP-binding protein RPAI1, which inhibits platelet activation and aggregation (Francischetti et al., 2000); and a group of lipocalins related to *T. pallidipenis* thrombin inhibitor triabin (Fuentes-Prior et al., 1997). Another lipocalin from the saliva of *T. pallidipenis* denominated pallidipin has been ascribed as a specific inhibitor of collagen-induced platelet aggregation (Noeske-Jungblut et al., 1994).

In this work we present the analysis of a set of 1,534 (SG) cDNA sequences, 645 of which coding for proteins of a putative secretory nature. Most salivary proteins described as lipocalins – 55% of the transcripts coding for putative secreted proteins – matched peptide sequences obtained from proteomic results. We expect this work will contribute with new salivary transcripts that could help the understanding of the role of salivary molecules in host/vector interactions and the discovery of novel pharmacologic agents.

## 2. Materials and methods

### 2.1. Triatomines and Salivary Glands cDNA Library Construction

*Triatoma infestans* were reared in an insectary room kept at $27^{o}$C ± 1.0°C, with a relative humidity ranging from 70 to 75% and a 16 h:8 h light:dark photoperiod. Salivary Glands (SG) were dissected from Vth instar nymphs 2 days after a blood meal and transferred to RNA-Later (Ambion) solution in 1.5 mL polypropylene vials. SG were kept at -20 $^{o}$C for isolating polyA+ RNA.

*T. infestans* SG mRNA was isolated from 15 SG pairs from Vth instar nymphs, using the Micro-FastTrack mRNA isolation kit (Invitrogen). The PCR-based cDNA library was made following the instructions for the SMART (switching mechanism at 5’end of RNA transcript) cDNA library construction kit (Clontech). This kit provides a method for producing high-quality, full-length cDNA libraries from nanogram quantities of polyA+ or total RNA. It utilises a specially designed oligonucleotide named SMART IV™ in the first-strand synthesis to generate high yields of full-length, double-stranded cDNA. *T. infestans* SG polyA+ RNA was used for reverse transcription to cDNA using PowerScript reverse transcriptase (Clontech), the SMART IV oligonucleotide, and the CDS III/3' primer (Clontech). The reaction was carried out at 42°C for 1 h. Second-strand synthesis was performed by a long-distance PCR-based protocol using the 5' PCR primer and the CDS III/3' primer as sense and antisense primers, respectively. These two primers also create *Sfi1A* and *B* restriction enzyme sites at the end of the cDNA. Advantage™ Taq polymerase mix (Clontech) was used to carry out the long-distance PCR reaction on a Perkin Elmer GeneAmp® PCR system 9700 (Perkin Elmer Corp.). The PCR conditions were: 95 °C for 1 min; 14 cycles of 95 °C for 10 s, 68 °C for 6 min. A small portion of the cDNA was analysed on a 1.1% agarose/EtBr (0.1 μg/mL) gel to check for the quality and range of the synthesised cDNA. Double-stranded cDNA was immediately treated with proteinase K (0.8 μg/mL) at 45 °C for 20 min and washed three times with water using Amicon filters with a 100 kDa cutoff (Millipore). The clean double-stranded cDNA was then digested with *SfiI* restriction enzyme at 50 °C for 2 h followed by size fractionation on a ChromaSpin–400 drip column (Clontech). The profiles of the fractions were checked on a 1.1% agarose/EtBr (0.1 μg/mL), and fractions containing cDNA were pooled in three different groups according to their size: large, medium or small sequences. Each group was concentrated and washed three times with water using an Amicon filter with a 100 kDa cutoff. The concentrated cDNA was then ligated into a λ TriplEx2 vector (Clontech), and the resulting ligation mixture was packaged using GigaPack® Gold III Plus packaging extract (Stratagene), according to the manufacturer's instructions. The packaged library was plated by infecting log-phase XL1-Blue *Escherichia coli* cells (Clontech). The percentage of recombinant clones was determined by performing a blue-white selection screening on LB/$MgSO_4$ plates containing X-gal/IPTG.

## 2.2. Sequencing of the *T. infestans* cDNA Library

The *T. infestans* SG cDNA library was plated on LB/$MgSO_4$ plates containing X-gal/IPTG to an average of 250 plaques per 150 mm Petri plate. Recombinant (white) plaques were randomly picked up and transferred to 96-well MICROTEST™ U-bottom plates (BD BioSciences) containing 75 μL of $H_2O$ per well. The phage suspension was either immediately used for PCR or stored at 4 $^{o}$C until use.

To amplify the cDNA using a PCR reaction, 4 μL of the phage sample were used as a template. The primers were sequences from the λ TriplEx2 vector and named PT2F1 (5'-AAGTACTCTAGCAATTGTGAGC-3') and PT2R1 (5'-CTCTTCGCTATTACGCCAGCTG-3'), positioned at the 5' and 3' end of the cDNA insert, respectively. The reaction was carried out in MicroAmp 96-well PCR plates (Applied Biosystems) using Platinum PCR® SuperMix (Invitrogen), on a Perkin Elmer GeneAmp® PCR system 9700 (Perkin Elmer Corp.). The PCR conditions were: 1 hold of 75 $^{o}$C for 3 min, 1 hold of 94 $^{o}$C for 4 min, 33 cycles of 94 $^{o}$C for 1 min, 49 $^{o}$C for 1 min, and 72 $^{o}$C for 1 min and 20 s. The amplified products were analysed on a 1.2% agarose/EtBr gel. cDNA library clones (1800 clones) were PCR amplified, and those showing a single band were selected for sequencing. The PCR products were used as a template for a cycle-sequencing reaction using DTCS labeling kit from Beckman Coulter.

The primer used for sequencing (PT2F3) is upstream from the inserted cDNA and downstream from the PT2F1 primer. The sequencing reaction was performed on a Perkin Elmer 9700 thermocycler. Conditions were 1 hold of 75 $^{o}$C for 2 min, 1 hold of 94 $^{o}$C for 4 min, and 30 cycles of 96 $^{o}$C for 20 s, 50 $^{o}$C for 20 s, and 60 $^{o}$C for 4 min. After cycle-sequencing the samples, a cleaning step was performed using the multiscreen 96-well plate cleaning system (Millipore). The 96-well multiscreening plate was prepared by adding a fixed amount (manufacturer's specification) of Sephadex-50 (Amersham Pharmacia) and 300 μL of deionised water. After partially drying the Sephadex in the multiscreen plate, the whole cycle-sequencing reaction was added to the center of each well, centrifuged at 2,500 rpm for 5 min, and the clean sample was collected on a sequencing microtiter plate (Beckman Coulter). The plate was then dried on a Speed-Vac SC110 model with a microtiter plate holder (Savant Instruments). The dried samples were immediately resuspended with 25 μL of formamide, and one drop of mineral oil was added to the top of each sample. Samples were either sequenced immediately on a CEQ 2000 DNA sequencing instrument (Beckman Coulter) or stored at –30 $^{o}$C. A total of 1,534 cDNA library clones were sequenced.

### 2.3. Reverse Phase Liquid Chromatography (RPLC)

Approximately 20 μL of saliva from adult insects were diluted in 200 μL buffer (50mM Tris-HCl, 120 mM NaCl pH 7.4). Proteins were loaded in a Microcon 10 (YM-10, Millipore) and centrifuged at 10,000 x g for 30 min. The flow through with proteins of low molecular weight (< 10 kDa) was kept. Microbore reversed-phase liquid chromatography (RPLC) was performed using a 0.3 mm C18 column from Phenomenex

(Torrence, CA). After sample injection, the column was washed for 10 min with 95% mobile phase A (0.1% formic acid in water) at 5 µL/min and peptides were eluted using a linear gradient to 90% mobile B (100% acetonitrile and 0.1% formic acid) for 40 min. The column eluate was monitored at 220 nm. A Probot fraction collector (Dionex, Sunnylvale, CA) was used to deliver the fractions to 96 well plates containing 25 µL of water. Fractions of interest were submitted to tryptic digestion followed by mass spectrometry as indicated below.

### 2.4. 2D-Gel Electrophoresis

2D gel electrophoresis was performed using ZOOM IPGRunner System (Invitrogen) under manufacturer's recommended running conditions. Briefly, approximately 500 µg of sample proteins of *T. infestans* saliva were solubilised with 155 µL rehydration buffer (7 M urea, 2 M thiourea, 2% CHAPS, 20 mM DTT, 0.5% carrier ampholytes, pH 3-10). The samples were absorbed by rehydration ZOOM strips (7 cm; pH 3-10 NL) overnight at room temperature and then focused under manufacturer's recommended conditions. The focused IPG strips were reduced/alkylated/equilibrated with reducing and alkylating reagents dissolved in the sample buffer. The strips were then applied onto NuPAGE 4–12% Bis-Tris ZOOM gels (Invitrogen). The gels were run under MOPS buffer and stained with SeeBlue staining solution (Bio-Rad).

### 2.5. Protein Identification by Mass Spectrometry

Protein identification of either RPLC or 2D gel-separated proteins was performed on reduced and alkylated trypsin-digested samples prepared by standard mass spectrometry protocols. Tryptic digests were analysed by coupling the Nanomate (Advion BioSciences)—an automated chip-based nano-electrospray interface source—to a quadrupole time-of-flight mass spectrometer, QStarXL MS/MS System (Applied Biosystems/Sciex). Computer-controlled, data-dependent automated switching to MS/MS provided peptide sequence information. AnalystQS software (Applied Biosystems/Sciex) was used for data acquisition. Data processing and databank searching were performed with Mascot software (Matrix Science). The NR protein database from the NCBI, National Library of Medicine, NIH, was used for the search analysis, as was a protein database generated during the course of this work.

## 2.6. Bioinformatic Tools and Procedures

Expressed sequence tags (ESTs) were trimmed of primer and vector sequences, clusterised, and compared with other databases as previously described (Valenzuela et al., 2003). The BLAST tool (Altschul et al., 1996), CAP3 assembler (Huang et al., 1999), ClustalW (Thompson et al., 1994), and Treeview software (Page, 1996) were used to compare, assemble, and align sequences, and to visualise alignments. For functional annotation of the transcripts we used the tool Blastx (Altschul et al., 1997) to compare the nucleotide sequences with the nonredundant (NR) protein database of the NCBI and to the Gene Ontology (GO) database (Ashburner et al., 2000). The tool rpsBlast (Schaffer et al., 2001) was used to search for conserved protein domains in the Pfam (Bateman et al., 2000), Smart (Letunic et al., 2002), Kog (Tatusov et al., 2003), and conserved domains (CDD) databases (Marchler-Bauer et al., 2002). We have also compared the transcripts with other subsets of mitochondrial and rRNA nucleotide sequences downloaded from

NCBI and to several organism proteomes downloaded from NCBI (yeast), Flybase (*D. melanogaster*), or ENSEMBL (*An. gambiœ*). Segments of the three-frame translations of the EST (since the libraries were unidirectional, we did not use six-frame translations) starting with a methionine in the first 100 predicted amino acids (aa)—or the predicted protein translation, in the case of complete coding sequences—were submitted to the SignalP server (Nielsen, 1997) to help identify translation products that could be secreted. *O*-glycosylation sites on the proteins were predicted with the program NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) (Hansen et al., 1998). Functional annotation of the transcripts was based on all the comparisons above. Following inspection of all results, transcripts were classified as either Secretory (S), Housekeeping (H), or of unknown (U) function, with further subdivisions based on function and/or protein families. Sequence alignments were done with the ClustalX software package (Thompson et al., 1997). Phylogenetic analysis and statistical neighbor-joining bootstrap tests of the phylogenies were done with the Mega package (Kumar et al., 2004). Hyperlinked Excel spreadsheets of the assembled EST's and of the salivary protein database are supplied as supplemental tables 1 and 2 at the journal site.

# 3. Results

## General Description of the Salivary Transcriptome Database

### 3.1. Description of the clusterised data set / cDNA library characteristics

1,534 sequences were used to assemble a clusterised database, yielding 657 clusters of related sequences, 500 of which contained only one EST. The consensus sequence of each cluster is named either a contig (deriving from two or more sequences) or a singleton (deriving from a single sequence). In this paper, we will use the denomination contig to address sequences deriving from both consensus sequences and from singletons. The 657 contigs were compared by the program Blastx, Blastn, or rpsBlast (Altschul et al., 1997) to the nonredundant protein database of the National Center of Biological Information (NCBI), to the gene ontology database (Ashburner et al., 2000), to the conserved domains database of the NCBI (Marchler Bauer et al., 2002), and to a custom-prepared subset of the NCBI nucleotide database containing either mitochondrial or rRNA sequences. Because the libraries are unidirectional, the three frame translations of the dataset were also derived, and open reading frames (ORF) starting with methionine and longer than 40 aa residues were submitted to SignalP server (Nielsen et al., 1997) to help identify putative secreted proteins. The EST assembly, BLAST, and signal peptide results were transferred into an Excel spreadsheet for manual annotation.

Five categories of expressed genes derived from the manual annotation of the contigs (Table 1). The putatively secreted (S) category contained 18% of the clusters and 42% of the sequences, with an average number of 5.5 sequences per cluster. The housekeeping (H) category had 36.4% and 35.9% of the cluster and sequences, respectively, and an average of 2.3 sequences per cluster. Forty-four percent of the clusters, containing 21% of all sequences, were classified as unknown (U) because no assignment for their function could be made; most of these consisted of singletons. Possible transposable elements originated 7 clusters, mostly singletons. We have also identified viral transcripts in our dataset. These data can be downloaded as Supplemental Table 1 for the EST data, and Supplemental File 2 for the proteome set.

**Table 1: Types of transcripts found in *Triatoma infestans* salivary glands.**

| Types of transcripts | Clusters | Sequences | Sequences/Cluster |
|---|---|---|---|
| Secreted (S) | 118 | 645 | 5.5 |
| Housekeeping (H) | 239 | 550 | 2.3 |
| Unknown (U) | 292 | 324 | 1.1 |
| Transposable element (TE) | 7 | 11 | 1.6 |
| Viral product | 1 | 4 | 4.0 |
| Total | 657 | 1534 | |

### 3.2. Housekeeping (H) genes

The 239 gene clusters (comprising 550 EST) attributed to H genes expressed in the (SG)s of *T. infestans* were further characterised into 20 groups, according to their possible function (Table 2). According to an organ specialised in secreting polypeptides and as observed in previous sialotranscriptomes (Francischetti et al., 2002; Ribeiro et al., 2004a, b; Calvo et al., 2007), the two larger sets were associated with protein synthesis machinery (298 EST in 51 clusters) and with energy metabolism (17 clusters containing

57 EST). We have also included in this category a group of 66 EST that grouped into 52 clusters and represent conserved proteins of unknown function presumably associated with cellular metabolism. Other sequences with homology to housekeeping protein include those coding for ribosomal, cytochrome, ATP synthase subunit, and NADH-ubiquinone oxidoreductase, among other molecules.

**Table 2: Functional classification of the housekeeping genes expressed in *Triatoma infestans* salivary glands.**

| Types of transcripts | Clusters | % | Sequences | % |
|---|---|---|---|---|
| Conserved, unknown function | 52 | 21.8 | 66 | 11.3 |
| Protein synthesis machinery | 51 | 21.3 | 298 | 54.2 |
| Signal transduction | 18 | 7.5 | 18 | 3.3 |
| Metabolism, energy | 17 | 7.1 | 57 | 10.4 |
| Transport/Storage | 14 | 5.9 | 15 | 2.7 |
| Cytoskeletal | 12 | 5.0 | 13 | 2.4 |
| Nuclear regulation | 12 | 5.0 | 13 | 2.4 |
| Protein export machinery | 10 | 4.2 | 12 | 2.2 |
| Protein modification | 10 | 4.2 | 16 | 2.9 |
| Transcription machinery | 9 | 3.8 | 10 | 1.8 |
| Proteasome machinery | 8 | 3.3 | 10 | 1.8 |
| Metabolism, carbohydrate | 6 | 2.5 | 6 | 1.1 |
| Metabolism, amino acid | 4 | 1.7 | 4 | 0.7 |
| Transcription factor | 3 | 1.3 | 3 | 0.5 |
| Metabolism, intermediate and detoxication | 3 | 1.3 | 3 | 0.5 |
| Metabolism, nucleotide | 3 | 1.3 | 3 | 0.5 |
| Lysosomal products | 2 | 0.8 | 2 | 0.4 |
| Extracellular matrix | 2 | 0.8 | 2 | 0.4 |
| Metabolism, oxidant | 2 | 0.8 | 2 | 0.4 |
| Metabolism, lipid | 1 | 0.4 | 1 | 0.2 |
| Total | 239 | | 550 | |

### 3.3. Viral product

Contig 110 produced a best Blastx match with gi|20451030|, a capsid protein precursor from a *Triatoma* virus (TrV), scoring with other reported capsid proteins. TrV infects triatomines and belongs to a new family of RNA viruses known as Dicistroviridae (Czibener et al., 2000).

### 3.4. Transposable Elements

Seven contigs on our database possibly derive from transposable elements (TE). Their translation products are similar to those of *Drosophila melanogaster* proteins annotated as endonuclease and reverse transcriptase. These transcripts may indicate active ongoing transposition activity in *T. infestans*, or with regulatory transcripts inhibiting TE genomic mobilisation.

### 3.5. Transcripts coding for putative secreted proteins in *Triatoma infestans* salivary glands

#### 3.5.1.1 - Lipocalins

The most abundant group of putative secreted proteins in *T. infestans* (SG)s is the lipocalins, corresponding to 55% of the transcripts in the S class. Lipocalins are widely distributed and heterogeneous proteins occurring in animals, plants and bacteria. Even though they have various molecular masses, the domain, determining the specific properties, is about 18-20 kDa (Flower et al., 1993). The lipocalin primary sequence shows a low percentage of similarity when comparing randomly selected members of the family. By contrast, an interesting feature of the lipocalins is their well conserved three-dimensional structure (Flower, 1995). They are typically small, extracellular proteins sharing several common molecular recognition properties: the binding of small, principally hydrophobic molecules, binding to specific cell-surface receptors; the formation of covalent and non-covalent complexes with other soluble macromolecules. Although they have been classified mainly as transport proteins, it is now clear that members of the lipocalin family fulfill a wide variety of different functions (Flower, 2000). The aim to discover the role of salivary lipocalins in blood-sucking insects through functional genomic and proteomic studies has been able to identify anticoagulants, antiplatelets, and vasodilatory molecules (Andersen et al., 2005).

We found lipocalins similar to salivary proteins of other species of the genus *Triatoma*, such as pallidipin, an inhibitor of collagen-induced platelet aggregation (Noeske-Jungblut et al., 1994); and triabin, a potent and selective thrombin inhibitor, both from the bug *Triatoma pallidipennis* (Noeske-Jungblut et al., 1995); and procalin, a salivary allergen from *T. protacta* (Paddock et al., 2001). The South American *T. brasiliensis* was also shown to contain salivary cDNA sequences similar to these three previously described *Triatoma* lipocalins (Sant'Anna et al., 2002). More recently, the sialotranscriptome of *T. brasiliensis* revealed a high content of lipocalins in its (SG)s, comprising 93.8% of the transcripts coding for putative secreted protein (Santos et al., 2007). Lipocalins have also been found in tick saliva and in *Rhodnius* performing similar functions, such as histamine and serotonin binding (Ribeiro et al., 2004a; Paesen et al., 2000; Sangamnatdej et al., 2002). This scenario contrasts with mosquitoes and sand flies where no salivary lipocalins have been described to date.

**Table 3: Classification of transcripts coding for putative secreted proteins in *Triatoma infestans* salivary glands.**

| Types of transcripts | Clusters | Sequences | % S type sequence |
|---|---|---|---|
| Lipocalins | 69 | 355 | 55.0 |
| Trialysin | 6 | 74 | 11.5 |
| Peptide with trialysin signal peptide | 7 | 64 | 9.9 |
| Hemolysin-domain protein | 7 | 32 | 5.0 |
| Defensin and immunity related peptides | 6 | 32 | 5.0 |
| Enzymes | 7 | 20 | 3.1 |
| Kazal domain-containing peptides | 3 | 17 | 2.6 |
| Antigen 5 protein family | 2 | 13 | 2.0 |
| Other peptides | 5 | 9 | 1.4 |
| Odorant binding protein family | 2 | 9 | 1.4 |
| Triatox | 1 | 8 | 1.2 |
| Similar to Culicoides salivary protein | 1 | 6 | 0.9 |
| Serpin | 1 | 4 | 0.6 |
| Thrombospondin-like | 1 | 2 | 0.3 |
| Total | 118 | 645 | |

Triatomine lipocalin sequences were aligned and a neighbor-joining phylogenetic-tree constructed (Fig. 1). The bootstrap values indicated that these sequences have evolved beyond recognition of a common ancestor (not shown). These results may indicate a long evolutionary history for these proteins, or alternatively, a fast rate of evolution. It is interesting to note here that lipocalins were not found in the sialotranscriptome of *Oncopeltus fasciatus*, a closely related seed feeding bug, indicating that expansion of this gene family expressed in the salivary glands of triatomines occurred as an adaptation to blood feeding (Francischetti et al., 2007)

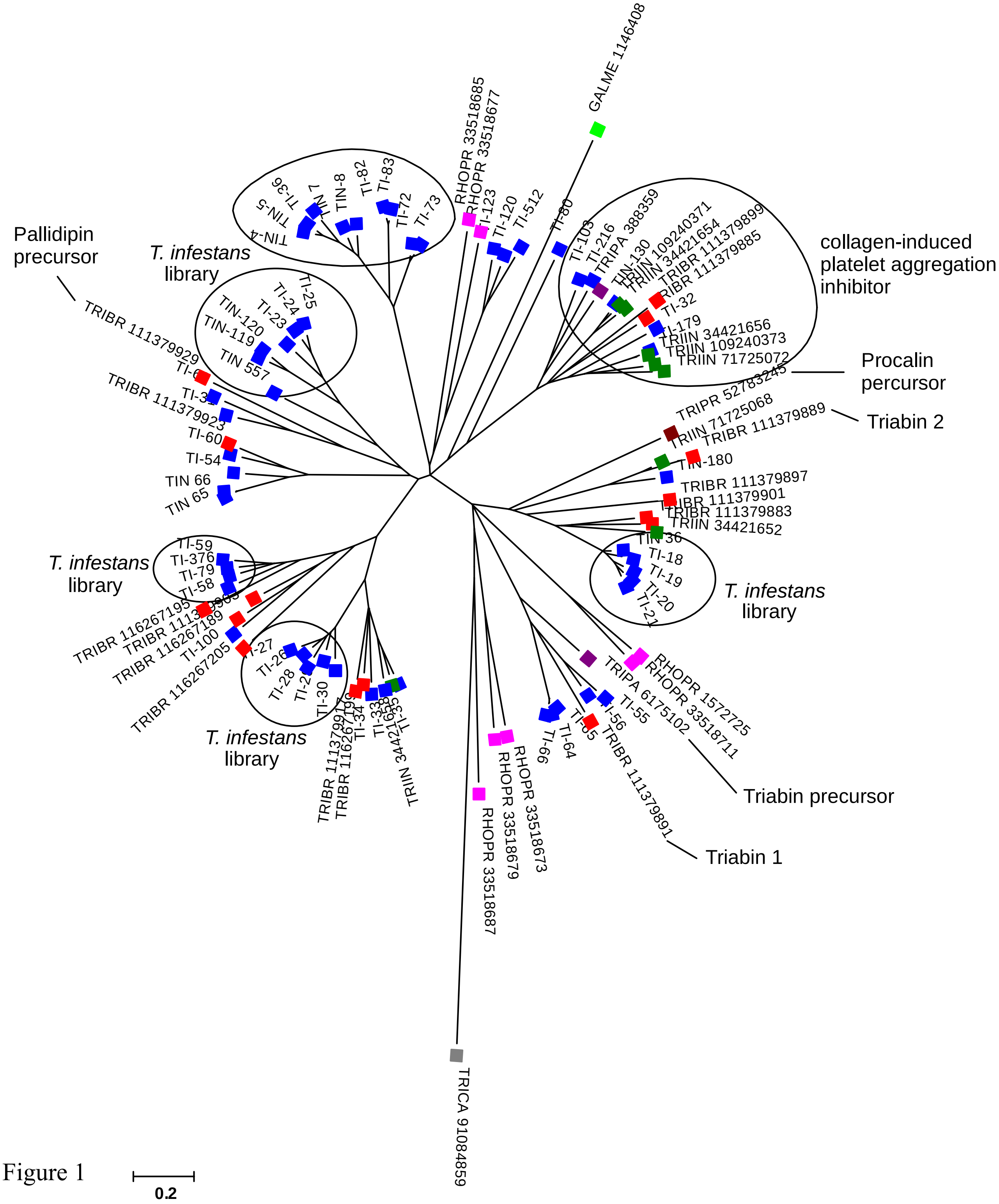


Figure 1

Figure 1. Dendrogram of the *T. infestans* salivary lipocalins with other insect salivary lipocalins. The sequences derived from the nonreduntant (NR) protein database of the National Center for Biotechnology Information (NCBI) are represented by five letters followed by the NCBI gi| accession number. The five letters derive from the first three letters of the genus and the first two letters from the species name. The protein sequences were aligned by the Clustal program (Thompson et al., 1997), and the dendrogram was done with the Mega package (Kumar et al., 2004) after 1,000 bootstraps with the neighbor joining (NJ) algorithm. The bar at the bottom represents 20% amino acid substitution. The colorful squares indicate each insect species whose sequences were used: blue, *T. infestans* sequences from cDNA library; dark green, *T. infestans* proteins described previously; red, *T. brasiliensis*; purple, *T. pallidipennis*; brown, *T. protacta*; magenta, *Rhodnius prolixus*; gray, *Tribolium castaneum*; light green, *Galleria mellonella*.

### 3.5.1.2 - Enzymes

Some transcripts coding for enzymes are identifiable. A serine protease could be involved with specific host proteolytic events that could affect clotting or the complement cascade. It could also be involved in immunity, since prophenoloxidase-activating enzymes are serine proteinases (Söderhäll and Cerenius, 1998). A serine protease with trypsin-like activity was described in *T. infestans* saliva. This salivary proteolytic activity was denominated triapsin and showed to be released with ejected saliva in active form, suggesting a role in blood-feeding (Amino et al., 2001). Since some Hemiptera species utilise cysteine and aspartic proteases for proteolytic digestion (Terra et al., 1996), the presence of a serine-protease could be more related to blood acquisition rather than its digestion. Or it could play a role in the innate immunity of the insect, considering the serine protease as a part of the prophenoloxidase system. This EST has similarities in the alignment with serine proteases from other insects, and the phylogenetic analysis shows a strong bootstrap support for the triatomine clade (Fig. 2).

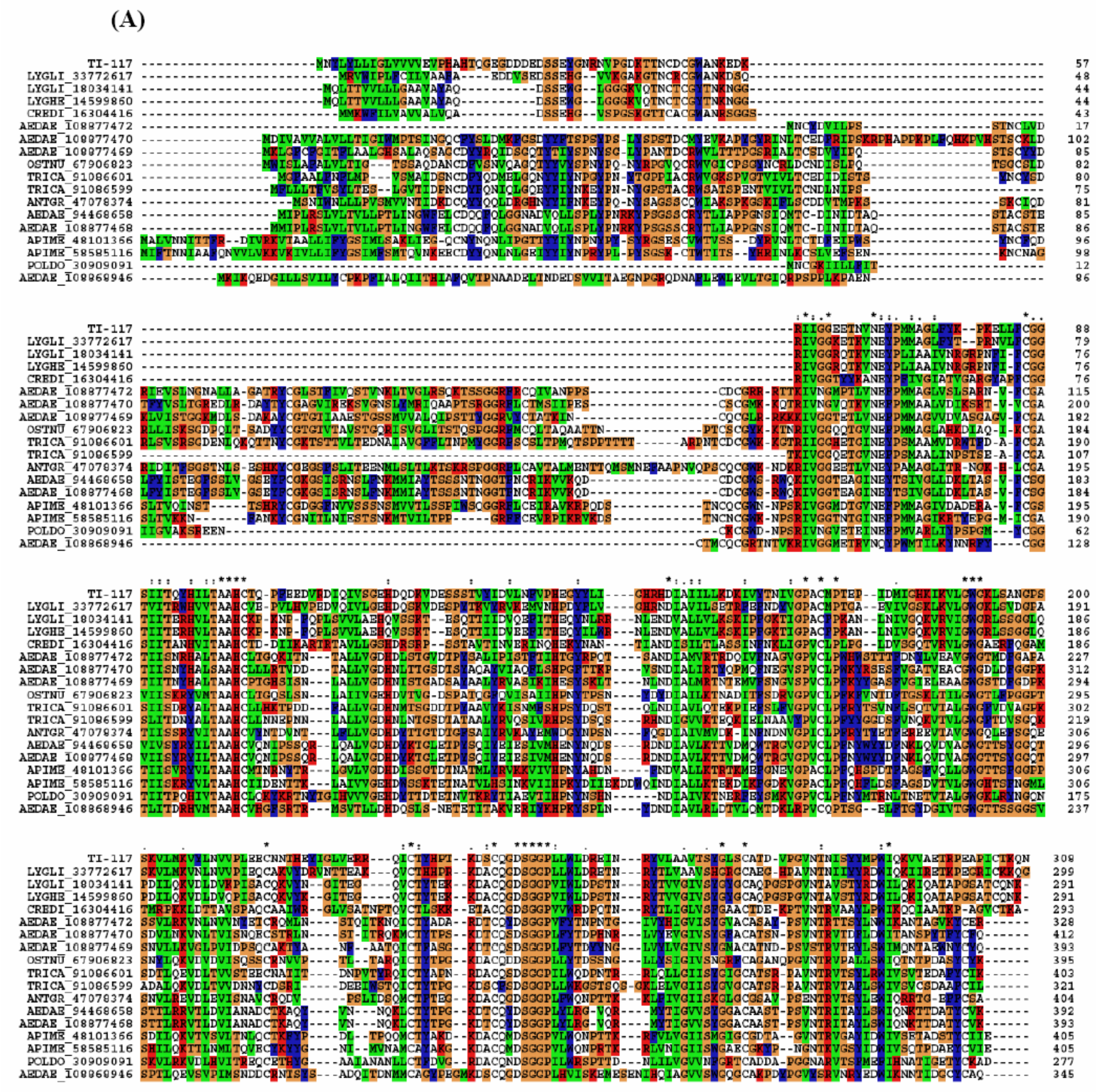


Figura 2A

**(B)**

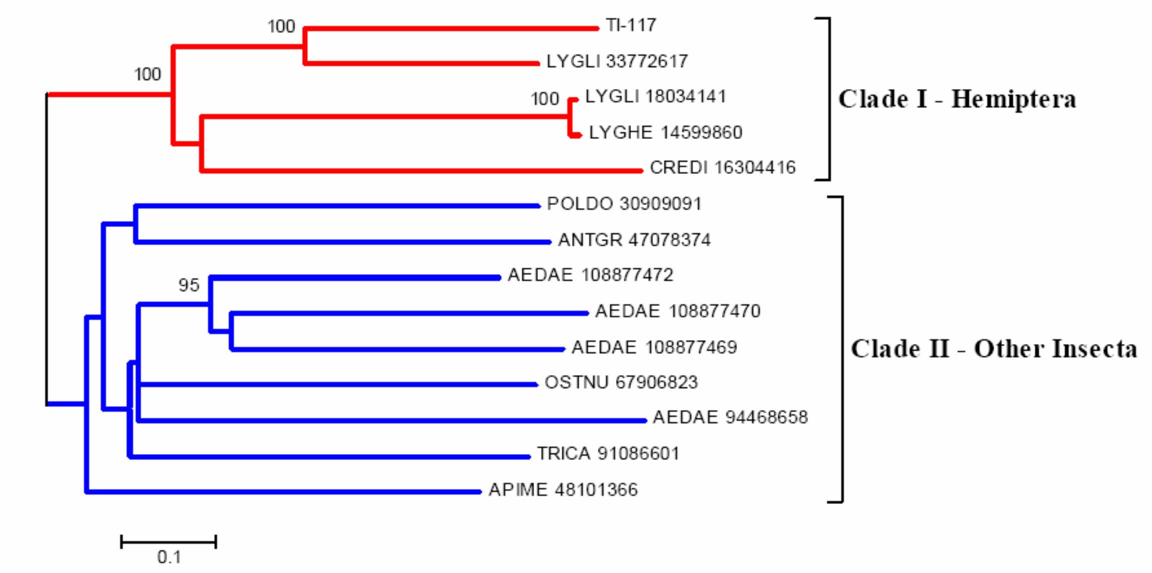


Figure 2. (A) ClustalW alignment of members of the serine protease family deriving from salivary glands of *T. infestans* and from other insects. (B) Neighbor-joining phylogram. The numbers in the phylogram nodes indicate percent bootstrap support for the phylogeny. The bar at the bottom indicates 10% amino acid divergence in the sequences.

### Apyrase

Partial coding sequences (truncated in the 5’ region) for the enzyme apyrase, a member of the 5’ nucleotidase family, were found. This enzyme was described before in *T. infestans* (SG)s (Faudry et al., 2004), where it helps the acquisition of blood meals by the degradation of adenosine diphosphate (ADP), a mediator of platelet aggregation and inflammation (Ribeiro and Francischetti, 2003).

## Inositol phosphatase family

Inositol phosphates are involved in many cellular processes related to signal transduction, secretion and cytoskeletal structure. Numerous enzymes are involved in the metabolism of inositol phosphates and phosphoinositides, including kinases, phosphatases, and phospholipases (Erneux et al., 1998; Guo et al., 1999). Inositol polyphosphate 5-phosphatases (IPPs) share a 5-phosphatase domain and consist of a large family of enzymes acting on different substrates. They regulate the pool of 5-phosphorylated inositol phosphates and phosphoinositides, thereby influencing cellular processes, for which second messenger functions have been reported: Ins(1,4,5)$P_3$, Ins(1,3,4,5)$P_4$, and the lipids PtdIns(4,5)$P_2$ and PtdIns(3,4,5)$P_3$ (Toker et al., 1997; Majerus, et al., 1999).

Alignment of two contigs from the cDNA library with other IPP sequences from vertebrates and insects showed they share similarities and conserved (aa) sequence (Fig. 3A). The phylogram shows a clade with *Triatoma* and *Rhodnius* sequences but apart from other insects, and even the *Strongylocentrotus purpuratus* sequence is more conserved with vertebrates than with insects (Fig. 3B).

*Rhodnius prolixus* also secretes into its saliva an inositol polyphosphate 5-phosphatase during blood feeding (Ribeiro et al., 2004a). It seems to reduce the concentration of some phospholipids in the plasma membrane of cells and platelets; this could have effects such as elimination of substrates for phospholipase C and PI3-kinase, changes in cytoskeletal architecture, and changes in membrane trafficking and vesicle secretion (Andersen et al., 2006). The presence of this phosphatase in *Triatoma* saliva can be also related to antithemostatic response, since changes in platelets membrane could interfere with their aggregation and facilitate the acquisition of blood meal. This protein was also identified in the 2D gel (Fig. 12).

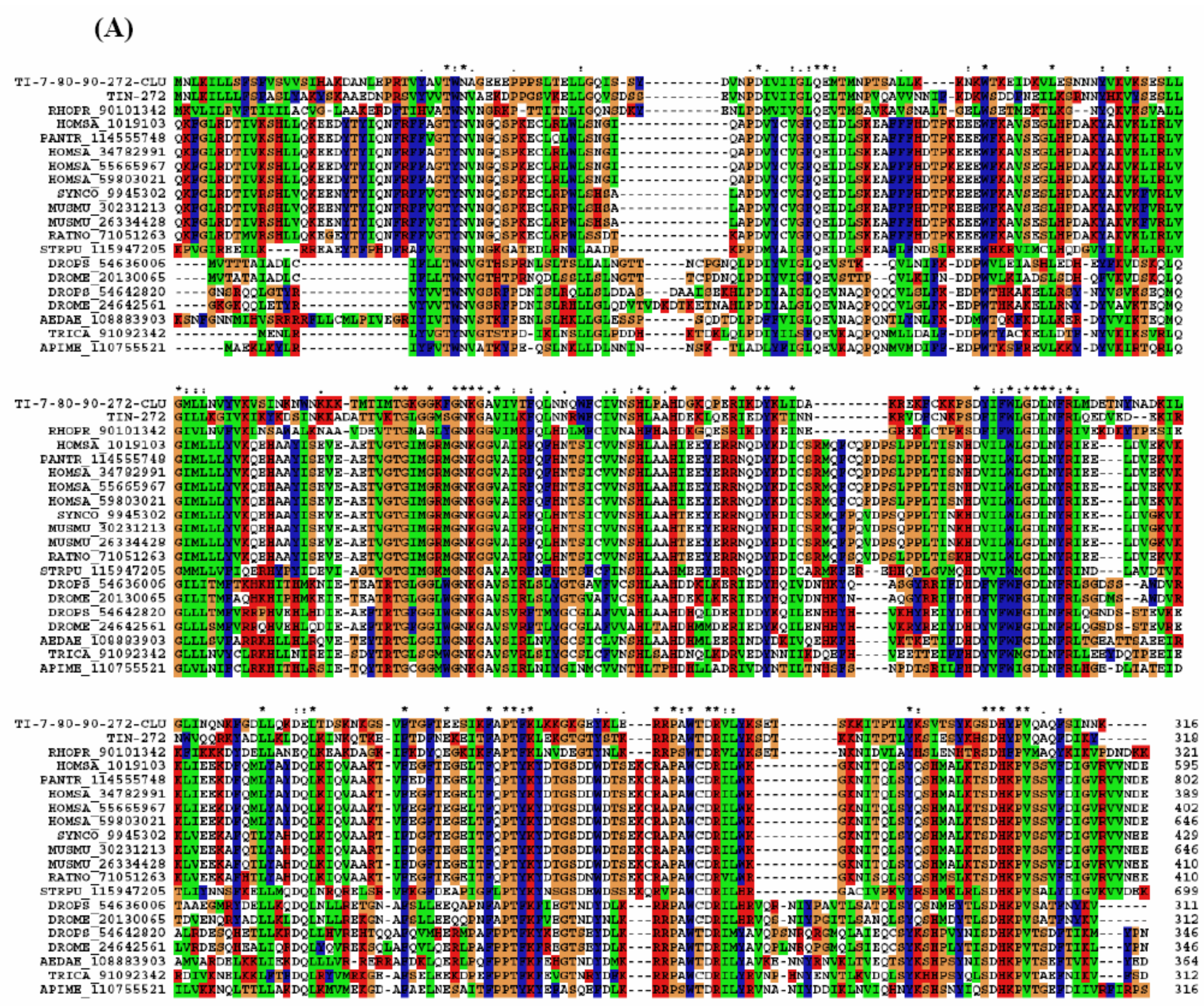


Figura 3A

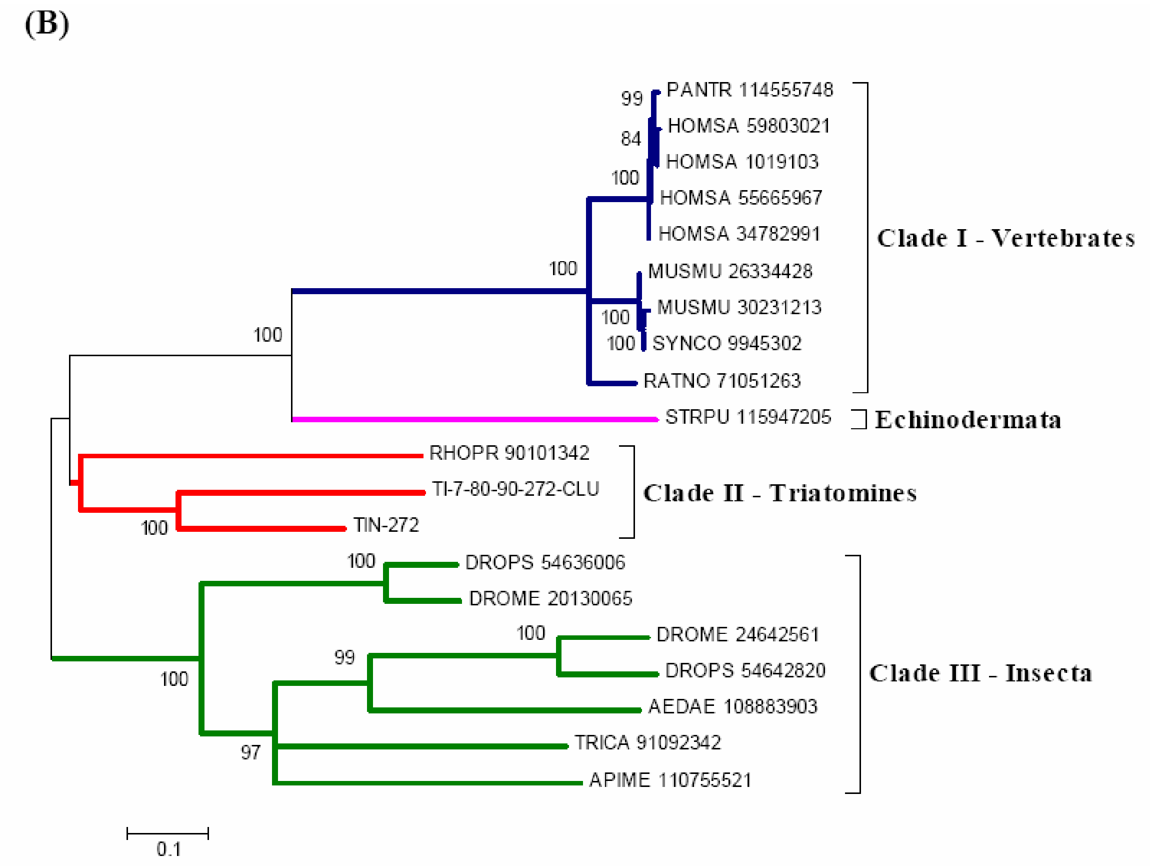


Figure 3. (A) ClustalW alignment of the *T. infestans* inositol polyphosphate 5-phosphatases (IPP) with IPP from other organisms. (B) Neighbor-joining phylogram. The numbers in the phylogram nodes indicate percent bootstrap support for the phylogeny. The bar at the bottom indicates 10% amino acid divergence in the sequences.

### 3.5.1.3 - Serpins

Serpins constitute a family of proteins, most of which function as serine proteinase inhibitors. They are important regulators of serine proteinases involved in inflammation, blood coagulation, fibrinolysis, and complement activation. Other roles for serpins could be protection against microbial proteinases, regulation of both endogenous proteinases and phenoloxidase activation (Kanost, 1999). Among arthropods, serpins have been described from lepidopteran insects, crayfish, ticks, and others (Jiang, et al., 1996; Liang et al., 1995; Prevot, et al, 2006). One of the contigs from the cDNA library matched for a serpin but the sequence was truncated (Table 3).

#### 3.5.1.4 - Kazal domain-containing peptides

Two contigs with transcripts encoding a polypeptide with a mature molecular mass of 8.9 kDa and containing a Kazal domain are predicted. Although this domain is associated with serine protease inhibitors, it is also found in other extracellular proteins without this function and containing antimicrobial activity (Fogaça et al., 2005). Kazal-type serine protease inhibitors are single- or more commonly multi-domain proteins, can be classical or non-classical Kazal-type inhibitors, and share a conserved sequence motif and molecular conformation (one central α-helix and three little antiparallel β- sheets) (Stubbs et al., 1997). Several Kazal-type family members have been previously described to be present in vertebrate and invertebrate animals. The role of Kazal-type inhibitors is not always known. Some invertebrates Kazal-type inhibitors, such as rhodniin, bdellin B-3 and infestin 1-2 (Friedrich et al., 1993; Fink, et al., 1986; Sommerhoff et al., 1994; Campos, et al., 2002) are identified as non-classical Kazal-type inhibitors. Thrombin inhibitors of the Kazal-type family have been found in triatomines. Infestin is the longest Kazal-type proteinase inhibitor precursor described in triatomines. It seems to be similar to rhodniin and dipetalogastin precursors but has more Kazal-type domains (Lovato et al., 2006). Rhodniin, isolated from the blood sucking insect *R. prolixus*, is such a Kazal-type inhibitor composed of two domains and strongly inhibits thrombin, the key enzyme of the blood coagulation cascade (van de Locht et al., 1995).

Alignment with other sequences containing a Kazal domain shows the presence of five conserved cysteines (Fig. 4A). These sequences also showed similarity to the vasodilator named vasotab from the horse fly *Hybomitra bimaculata*. Vasotab is a member of the Kazal-type protease inhibitor family (Takác et al., 2006). The phylogram analysis indicates a strong bootstrap support for the triatomine clade (Fig. 4B). A transcript encoding this polypeptide was also found in the mosquitoes *Ae. aegypti* (Ribeiro et al., 2007) and *Ae. albopictus* (Arcà et al., 2007). And recently, the sialotranscriptome of the triatomine *Triatoma brasiliensis* demonstrated the presence of 9 clusters coding for Kazal-type peptides (Santos et al., 2007).

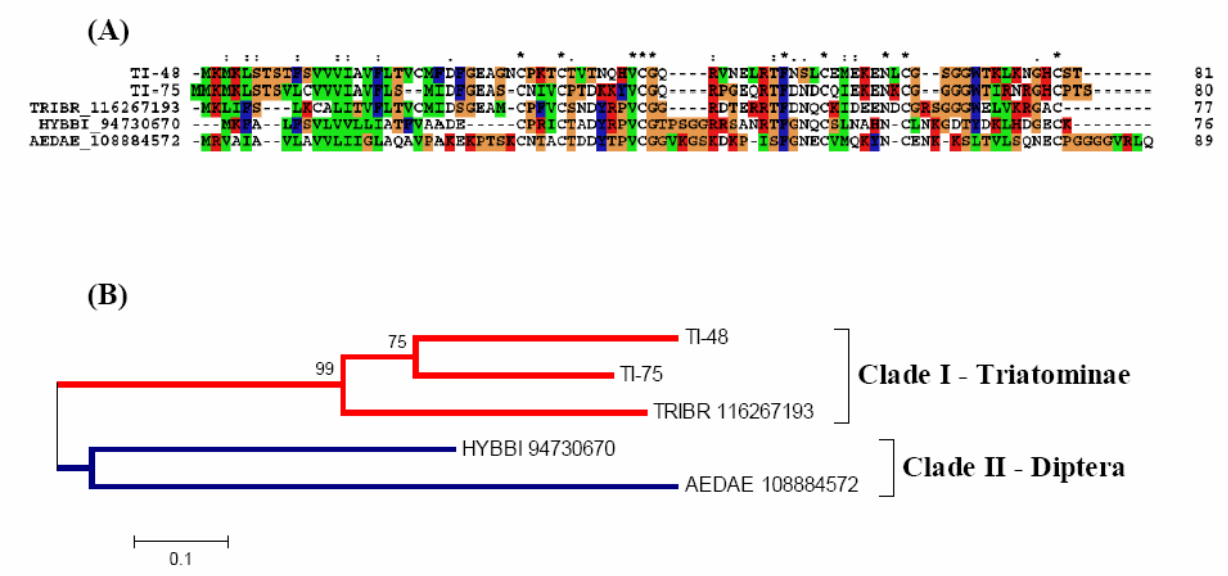
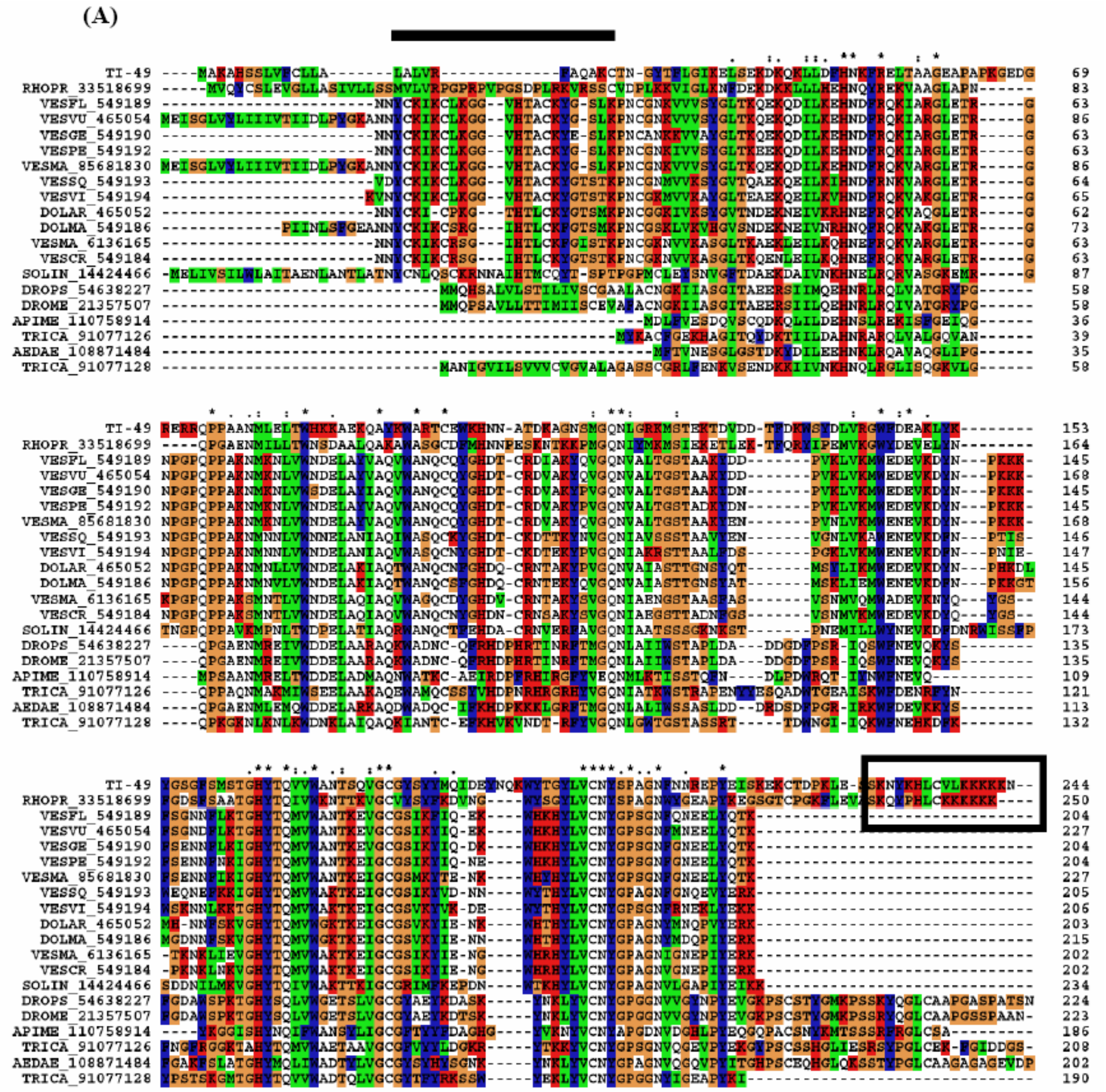


Figure 4. The salivary Kazal-domain containing peptides of *T. infestans*. (A) ClustalW alignment. Six conserved cysteines are shown in black with white background and marked with an asterisk (*). (B) Phylogenetic tree showing the sequence distance relationships between members of the family. The bar shows 10% divergence at the amino acid level. The numbers indicate the bootstrap value from 1000 trials.

### 3.5.1.5 - Antigen 5 protein family

This is a family of secreted proteins that belong to the CAP family (cysteine-rich secretory proteins; AG5 proteins of insects; pathogenesis-related protein 1 of plants) (Megraw et al., 1998). The CAP family is related to venom allergens in social wasps and ants (Hoffman, 1993; King and Spangfort, 2000) and to antifungal proteins in plants (Stintzi et al., 1993; Szyperski et al., 1998). Members of this protein family are found in the (SG)s of many blood-sucking insects (Francischetti et al., 2002; Li et al., 2001; Valenzuela et al., 2002b; Arcà et al., 2005; Calvo et al., 2007). The function of any AG5 protein in the saliva of any blood-sucking arthropod is still unknown. Other animals have shown some insight into their activity: snake venom proteins of the same family have been shown to contain smooth muscle-relaxing activity (Yamazaki et al., 2002; Yamazaki and Morita, 2004), and the salivary neurotoxin of the venomous lizard *Heloderma horridum* is also a member of this protein family (Nobile et al., 1996).

We here report a salivary member of the antigen-5 family found in *T. infestans* (SG)s. Alignment and phylogenetic analysis of insect members of this family indicates that *T. infestans* salivary antigen-5 protein clusters with other reported antigen-5 proteins from *R. prolixus*, vespids and other insects. The hymenoptera proteins, from vespids, have an additional domain after the signal peptide (Fig. 5), indicating that probably their evolution diverged at some point from the others. Both *Rhodnius* and *Triatoma* proteins have a basic tail. This content of a poly-Lys tail in salivary proteins may direct these proteins to the surface of activated platelets, where they could bind and/or release biologically active ligands or interact with other proteins on the platelet surface. They could also block the interaction of coagulation factors with the membrane surface or direct the inhibitor to negatively charged membranes critical for productive blood coagulation complex assembly (Broze Jr., 1995). Demonstration that proteins rich in positively charged residues effectively block the coagulation cascade comes from studies performed with a recombinant *R. prolixus* salivary lipocalin (nitrophorin-7, NP-7). NP-7 contains a cluster of positively charged residues in the N-terminus and specifically binds to anionic phospholipids, preventing thrombin formation by the prothrombinase complex (Andersen et al., 2004).



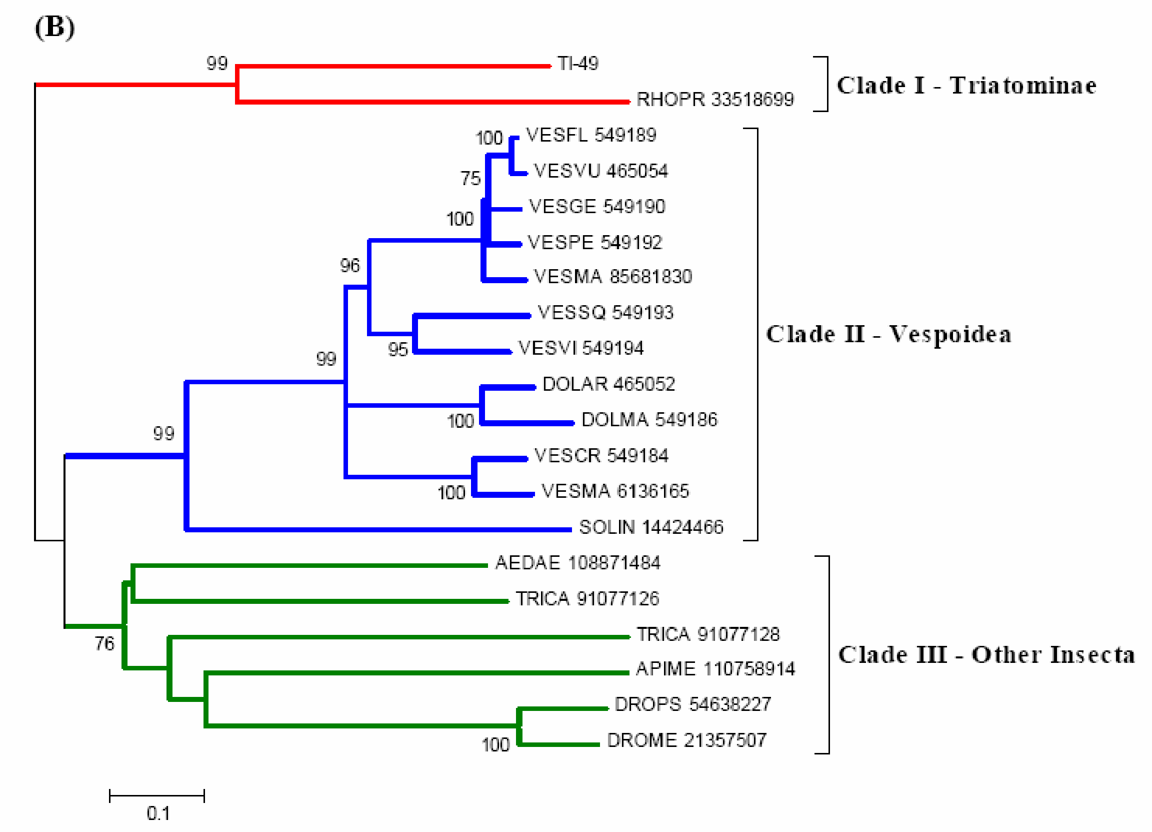


Figure 5. (A) ClustalW alignment of *T. infestans* salivary antigen-5 with members of the same family. (B) Phylogenetic tree of selected members of the antigen-5 family of proteins from Triatominae, Vespoidea and other Insecta, obtained by the NJ method. The numbers in the phylogram nodes indicates percent bootstrap support for the phylogeny. The bar at the bottom indicates 10% amino acid divergence in sequences. For details, see text.

### 3.5.1.6 - Odorant binding protein family

Many airborne molecules, such as hydrophobic odorants and pheromones, must be recognized by a specialised class of proteins that facilitate their delivery to the olfactory receptors (OR) (Forêt and Maleszka, 2006). In both insects and vertebrates this function is provided by odorant binding proteins (OBPs) (Pelosi et al., 1996; Krieger and Breer, 1999; Deyu and Leal, 2002). They are generally thought to solubilise hydrophobic odorants and carry them to the respective receptors (Vogt et al., 1991; Pelosi et al., 1994; Prestwich et al., 1995). Besides contributing to the recognition of odorants in insects, they may also function as carriers in other developmental and physiological processes.

Insect OBPs are small, water soluble molecules expressed in both olfactory and gustatory sensilla, as well as in other specialised tissues (Pelosi et al., 2005). OBPs have also been found in non-olfactory tissue, suggesting that their roles may be related to general carrier capabilities with broad specificity for lipophilic compounds (Forêt and Maleszka, 2006). The heme-binding protein of *R. prolixus* (Paiva-Silva et al., 2002) is one of these OBPs implicated in non-olfactory functions. In fact, there is increasing evidence that OBPs do play an active role in odorant recognition rather than merely serving as passive odorant carriers. One line of evidence is the large number of OBPs present within a variety of insect species. Several proteins of the odorant-binding protein (OBP) family have been described in *Drosophila melanogaster*, *Apis mellifera*, and the hemiptera *R. prolixus*, where a putative OBP with a signal peptide indicative of secretion was identified (Hekmat-Scafe et al., 2000; Forêt and Maleszka, 2006; Ribeiro et al., 2004a). We found transcripts coding for members of the OBP family in *T. infestans*.

They share similarities with OBPs from the mosquitoes *An. gambiae* and *Ae. aegypti*. The OBP-like sequences have four conserved cysteines and good bootstrap support for the triatomine clade (Fig. 6). Sequences in Diptera clade do not have a strong bootstrap support, except for some members that probably originate from genetic duplication and/or function conservation. It has been proposed that a moderate number of OBPs could act in a combinatorial manner with a moderate number of OR to greatly increase the discriminating power of an insect's olfactory system (Hekmat-Scafe et al., 2002).

It is possible to speculate whether these salivary OBPs could be acting as partners in gustatory perception in *Triatoma*, as they are related to gustatory perception in *Drosophila* (Galindo and Smith, 2001). If this is the case, saliva would be important not only for its anti-hemostatic function, but also in the chemoreception of phagostimulants (Friend and Smith, 1977). Against this hypothesis is the previous observation that *R. prolixus* surgically deprived of salivary glands were able to feed normally on artificial feeder (Ribeiro and Garcia, 1981)

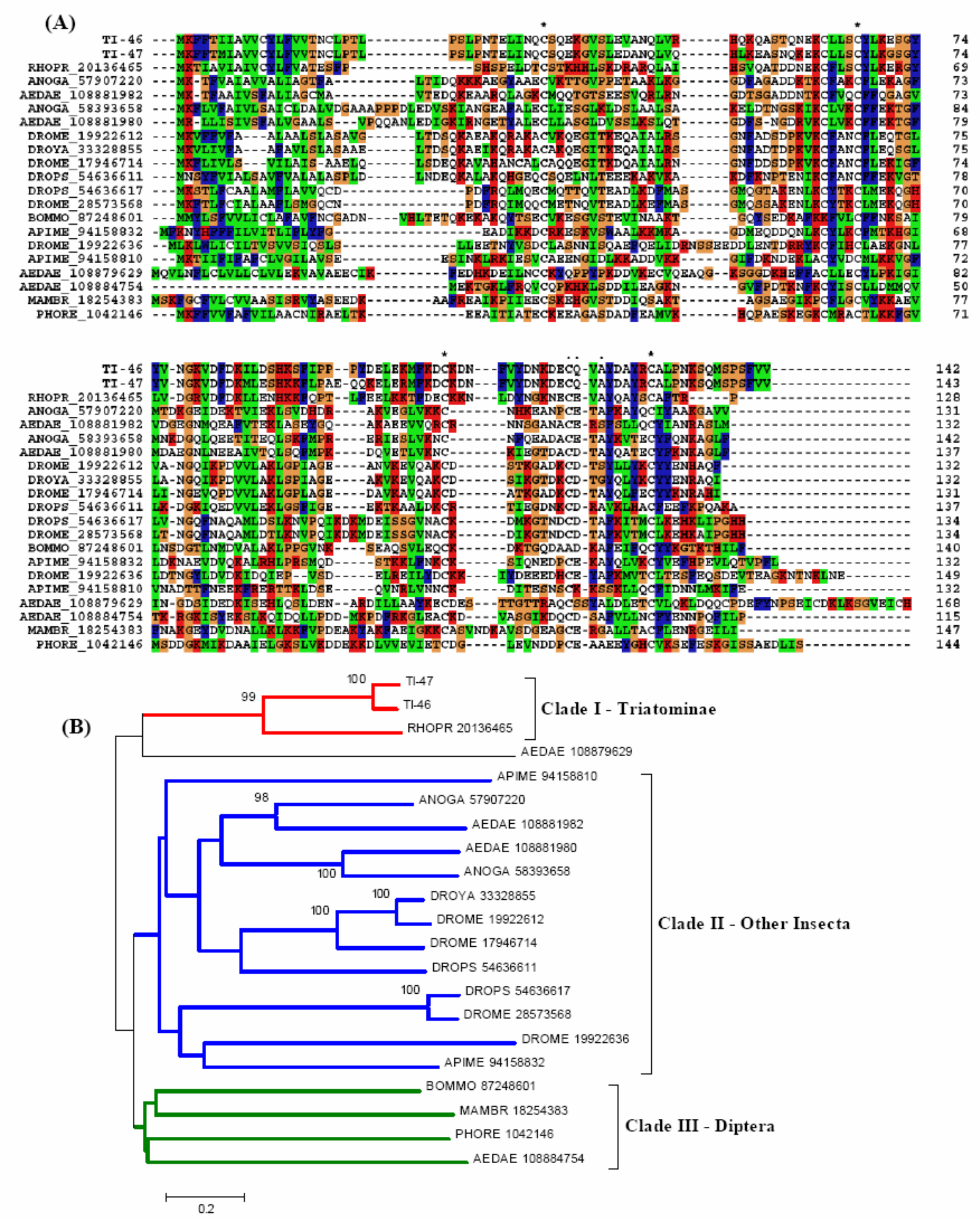


Figure 6. (A) ClustalW alignment of the OBP family of salivary peptides. Four conserved cysteines are shown in black with white background and marked with an asterisk (*). (B) Neighbor-joining phylogram. The numbers in the phylogram nodes indicates percent bootstrap support for the phylogeny. The bar at the bottom indicates 20% amino acid divergence in sequences.

### 3.5.1.7 - Similar to Culicoides salivary protein

This EST sequence found in *T. infestans* shows similarity with Culicoides and Fungi sequences (Fig. 7). Culicoides are competent vectors of a wide range of economically important pathogens that affect both domestic and wild animals (Mellor et al., 2000). A broad range of genes encoding salivary proteins were characterised from a (SG) cDNA library of *C. sonorensis* and revealed many proteins involved in antihemostasis, immunomodulation, and digestion (Campbell et al., 2005). In addition to acting as a vector, Culicoides are the primary cause of an extremely pruritic allergic dermatitis known colloquially as “summer eczema”, or “insect bite hypersensitivity” in horses (Anderson, et al., 1993; Kurotaki, et al., 1994). The saliva of kissing bugs (*Triatominae*) contains several allergens responsible for severe allergic reactions such as urticaria, dyspnea, and anaphylaxis in humans (Moffitt et al., 2003). This transcript, similar to one found in Culicoides, could be related to hypersensitivity or allergic responses by the human host.

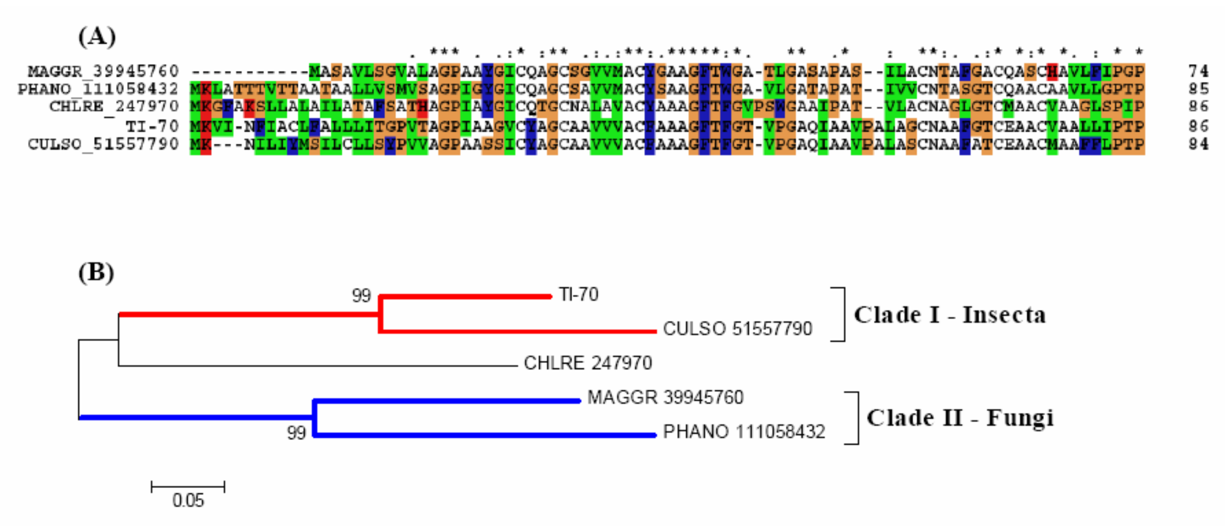
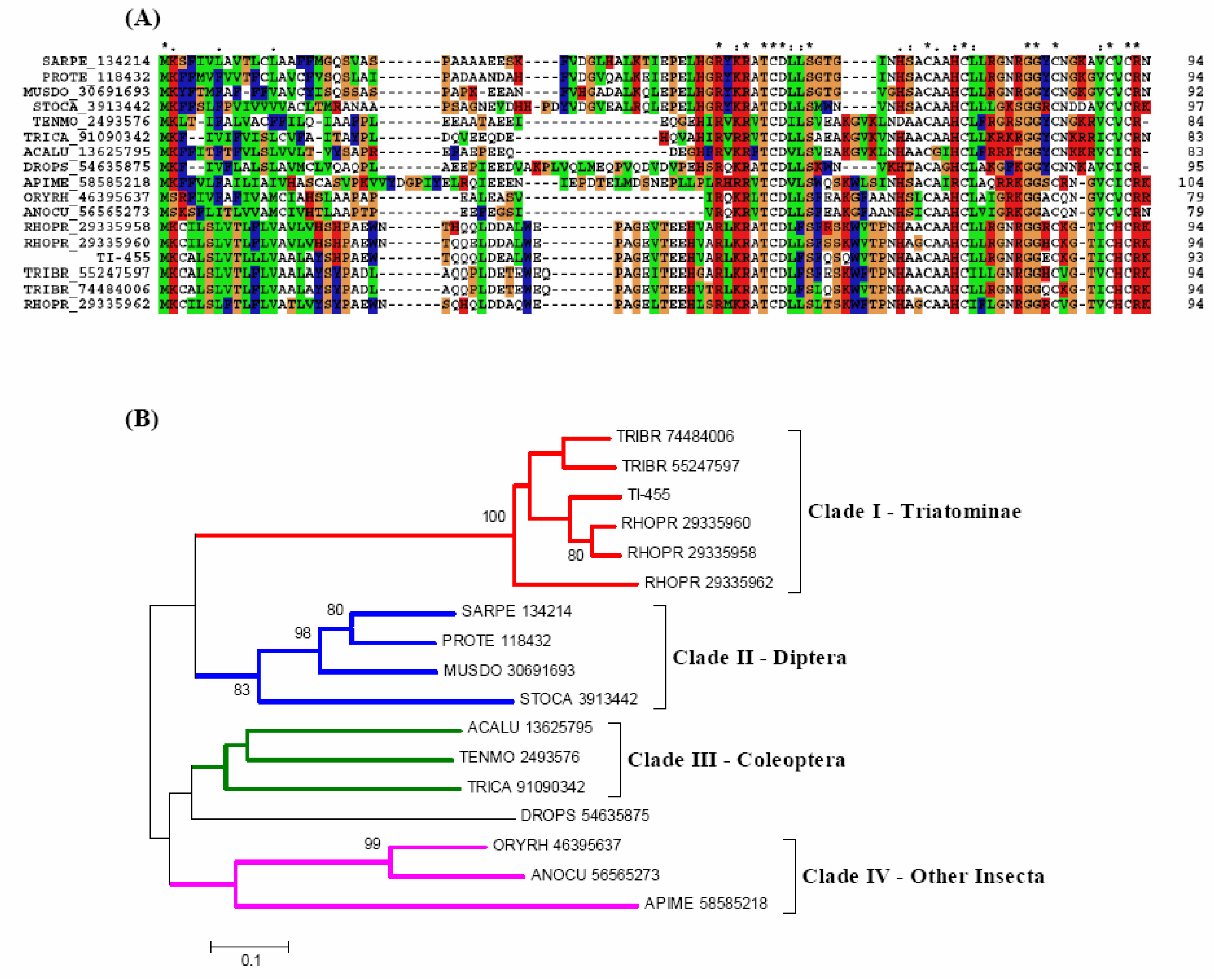


Figure 7. (A) ClustalW alignment of a predicted sequence from a contig that showed similarity with a sequence from Culicoides. (B) Neighbor-joining phylogram. The numbers in the phylogram nodes indicates percent bootstrap support for the phylogeny. The bar at the bottom indicates 5% amino acid divergence in sequences.

### 3.5.1.8 - Defensin and immunity-related peptides

Antimicrobial peptides (AMPs) are widely distributed throughout the animal and plant kingdoms. Despite sharing some common features, such as a small size (often below 10 kDa) and a cationic character, most AMPs differ in their aa sequence and mode of action (Fogaca et al., 2004). Within this AMP family, defensins constitute one of the major families that have been characterised (Bulet et al., 2004). Defensins are cationic peptides with molecular weights of about 4 kDa, three disulphide bridges, formed by six cysteines residues and three characteristic domains, an amino terminal flexible loop, followed by an α-helix and a carboxy-terminal anti-parallel β-sheet (Bonmantin et al., 1992; Bulet et al., 1999). This peptide can be found in different organisms, such as plants, fungi, mollusks, scorpions, insects, and birds, and also in different cells of various mammals. In humans, insects, and plants, defensins contribute significantly to the host defense against invasion by microorganisms (Raj et al., 2002). In addition to the diversity of structure and mode of action, the site and the regulation of the synthesis of AMPs also differs among arthropod groups. In insects, AMPs are mainly synthesised in the fat body and their gene transcription is strongly induced after an injury and/or infection (Bulet et al., 2002). Defensins have been isolated and characterised as part of the innate immune response from the hemolymph of all insect species so far investigated, like Odonata, Diptera, Coleoptera, Lepidoptera, Acari and Hemiptera (Bulet et al., 1992; Ishibashi et al., 1999; Lamberty et al., 1999; Lopez et al., 2003; Ceraul et al., 2003; Johns et al., 2001; Bartholomay et al., 2004).



```
               *.    .     .                                                *  :* ***::*         .: *. :*:     **  *   :* **
  SARPE_134214 MKSFIVLAVTLCLAAFFMGQSVAS-------PAAAAEESK----FVDGLHALKTIEPELHGRYKRATCDLLSGTG----INHSACAAHCLLRGNRGGYCNGKAVCVCRN    94
  PROTE_118432 MKFFMVFVVTFCLAVCFVSQSLAI-------PADAANDAH----FVDGVQALKEIEPELHGRYKRATCDLLSGTG----INHSACAAHCLLRGNRGGYCNGKGVCVCRN    94
MUSDO_30691693 MKYFTMFAF-FFVAVCYISQSSAS-------PAPK-EEAN----FVHGADALKQLEPELHGRYKRATCDLLSGTG----VGHSACAAHCLLRGNRGGYCNGKGVCVCRN    92
  STOCA_3913442 MKFFSLFPVIVVVVACLTMRANAA-------PSAGNEVDHH-PDYVDGVEALRQLEPELHGRYKRATCDLLSMWN----VNHSACAAHCLLLGKSGGRCNDDAVCVCRK    97
  TENMO_2493576 MKLT-IFALVACFFILQ-IAAFPL-------EEAATAEEI---------------EQGEHIRVKRVTCDILSVEAKGVKLNDAACAAHCLFRGRSGGYCNGKRVCVCR-    84
TRICA_91090342 MKF--IVIFVISLCVFA-ITAYPL-------DQVEEQDE----------------HQVAHIRVRRVTCDLLSAEAKGVKVNHAACAAHCLLKRKRGGYCNKRRICVCRN    83
ACALU_13625795 MKFFITFTFVLSLVVLT-VYSAPR-------EFAEPEEQ-----------------DEGHFRVKRFTCDVLSVEAKGVKLNHAACGIHCLFRRRTGGYCNKKRVCICR-    83
DROPS_54635875 MKF--IVFLALSLAVMCLVQAQPL-------AEEPIEEDVAKPLVQLMEQPVQDVDVPEHSRQKRATCDLLSKWN----VKHTACAGHCLAKGFKGGYCNNKAVCICR-    95
APIME_58585218 MKFFVLFAILIAIVHASCASVPKVVYDGPIYELRQIEEEN----IEPDTELMDSNEPLLPLRHRRVTCDVLSWQSKWLSINHSACAIRCLAQRRKGGSCRN-GVCICRK   104
ORYRH_46395637 MSRFIVFAFIVAMCIAHSLAAPAP---------EALEASV--------------------IRQKRLTCDLLSFEAKGFAANHSLCAAHCLAIGRKGGACQN-GVCVCRR    79
ANOCU_56565273 MSKSFLITLVVAMCIVHTLAAPTP---------EEFEGSI--------------------VRQKRVTCDLLSFEAKGFAANHSICAAHCLVIGRKGGACQN-GVCVCRN    79
RHOPR_29335958 MKCILSLVTLFLVAVLVHSHPAEWN------THQQLDDALWE--------PAGEVTEEHVARLKRATCDLFSFRSKWVTPNHAACAAHCLLRGNRGGRCKG-TICHCRK    94
RHOPR_29335960 MKCILSLVTLFLVAVLVHSHPAEWN------TQQELDDALWE--------PAGEVTEEHVARLKRATCDLLSFSSKWVTPNHAGCAAHCLLRGNRGGHCKG-TICHCRK    94
        TI-455 MKCALSLVTLLLVAALAYSHPAEW-------TQQQLDEALWE--------PAGEVTEEHVARLKRATCDLFSFQSQWVTPNHAACAAHCLLRGNRGGECKG-TICHCRK    93
TRIBR_55247597 MKCALSLVTLFLVAALAYSYPADL-------AQQPLDETEWEQ-------PAGEITEEHGARLKRATCDLFSFESKWFTPNHAACAAHCILLGNRGGHCVG-TVCHCRK    94
TRIBR_74484006 MKCALSLVTLFLVAALAYSYPADL-------AQQPLDETEWEQ-------PAGEVTEEHVTRLKRATCDLFSLQSKWVTPNHAACAAHCLLRGNRGGQCKG-TICHCRK    94
RHOPR_29335962 MKCILSLFTLFLVATLVYSYPAEWN------SQHQLDDAQWE--------PAGELTEEHLSRMKRATCDLLSLTSKWFTPNHAGCAAHCIFLGNRGGRCVG-TVCHCRK    94
```

Figure 8. (A) ClustalW alignment of the defensin family of salivary peptides. Six conserved cysteines are shown in black with white background and marked with an asterisk (*). (B) Neighbor-joining phylogram. The numbers in the phylogram nodes indicates percent bootstrap support for the phylogeny. The bar at the bottom indicates 10% amino acid divergence in sequences.

Tick defensins have also been found to exist in (SG)s, an important organ for blood feeding and pathogen transmission (Valenzuela et al., 2002). A defensin-like molecule from *Ixodes ricinus* was recently described and found to be induced following microbial challenge (Rudenko et al., 2005). In triatomines, defensins have been investigated in *R. prolixus*, being isolated, purified and sequenced (Lopez et al., 2003).

From the alignment of the defensin sequence found in *T. infestans* (SG)s with other defensins, we can notice the presence of the six conserved cysteines, characteristic in this family and important for the disulphide bridges conformation (Fig. 8A). The phylogram reveals a triatomine clade sharing sequences from *T. brasiliensis* and *R. prolixus*, and showing a bootstrap support (Fig. 8B).

### Similar to bacterially induced peptide Hdd1 (from *Hyphantria cunea*)

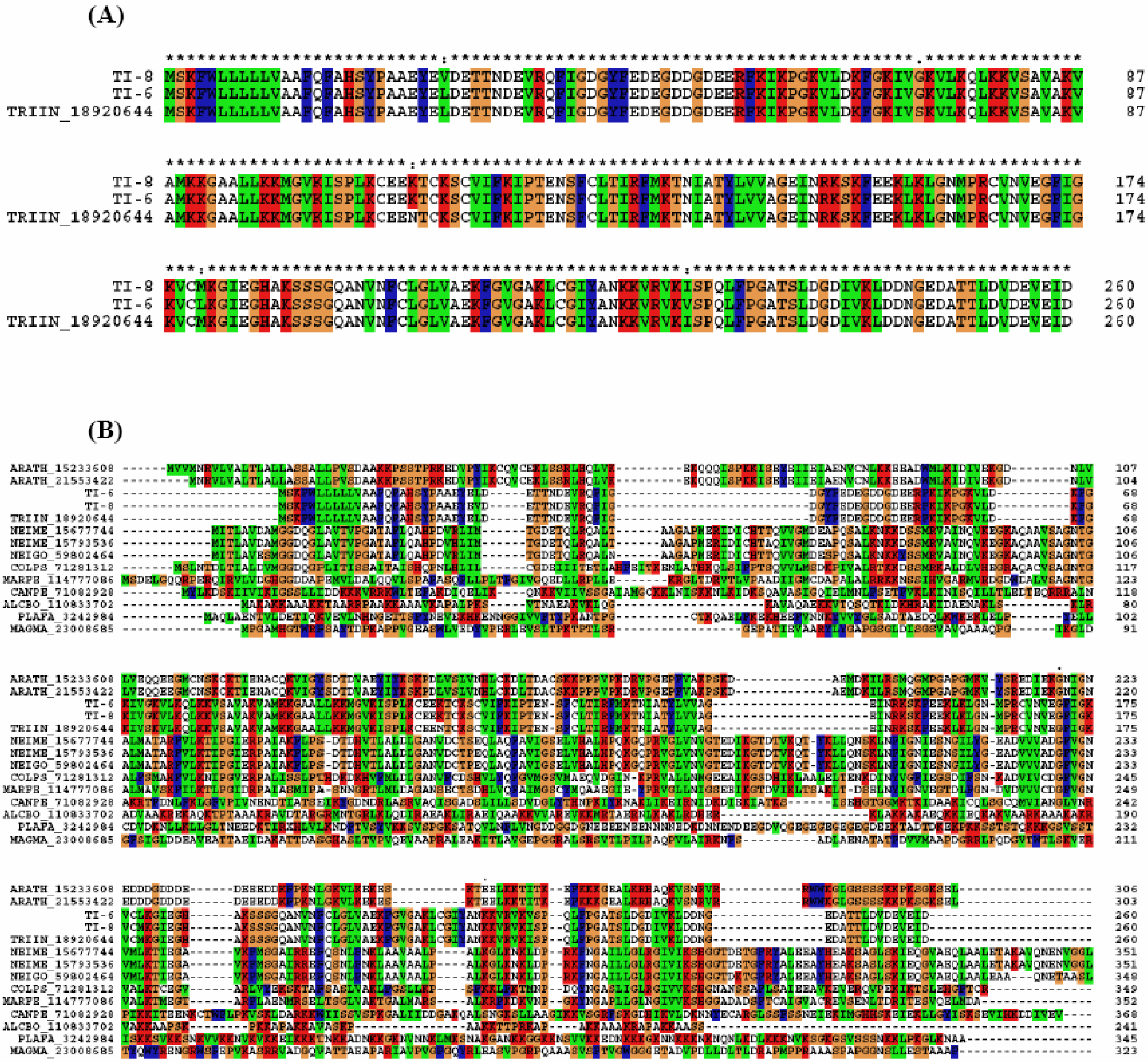
In a previous work, 11 inducible genes were isolated after bacterial challenge in *H. cunea*, one of the most serious insect pests in Korea. Hdd1 (*Hyphantria* differentially displayed clones) was one of this immune-related cDNA but it didn’t produce any significant homology with any known protein (Shin et al., 1998). One contig from the *T. infestans* cDNA library matched this peptide but the function remains unknown.

### 3.5.2 – Coding for proteins found only in triatomines

#### 3.5.2.1 - Trialysin

We identified two contigs that matched the protein trialysin, a pore-forming molecule present in the (SG)s of *T. infestans* (Amino et al., 2002; Martins et al., 2006) (Fig. 9). It is a 22-kDa protein synthesised as a precursor and processed by limited proteolysis. It exerts its potent cytolytic activity on a large variety of cell types, from bacteria to mammalian cells, and it has been hypothesised that it favors the maintenance of the salivary fluid free of microorganisms and parasites. Trialysin possesses a basic amphipathic lytic motif in the N-terminal region containing 27 aa residues, similar to antimicrobial lytic peptides, and a protein portion that increases the lytic specificity toward eukaryotic cells. Studies in which synthetic peptides containing portions of the N-terminus of trialysin have been tested for their lytic properties against the infective stage of *T. cruzi*, *Escherichia coli* and human erythrocytes revealed differences in specificity and activity for the tested targets. The cell membrane is the main target of antimicrobial peptides, thus the diverse lipid composition and distribution or the presence of other components in the cell membrane of the various organisms play an important role in the modulation of peptide/membrane interaction, leading to the differentiation of their action (Amino et al., 2002; Martins et al., 2006).

Further analysis of the near relatives of trialysin found in the nonredundant database indicates related bacterial genes, a proteobacteria clade being strongly associated with the *T. infestans* sequences (Fig. 9b, c).



```
                **************************:********************************************.***************
            TI-8 MSKFWLLLLLVAAFQFAHSYPAAEYEVDETTNDEVRQFIGDGYFEDEGDDGDEERFKIKPGKVLDKFGKIVGKVLKQLKKVSAVAKV    87
            TI-6 MSKFWLLLLLVAAFQFAHSYPAAEYELDETTNDEVRQFIGDGYFEDEGDDGDEERFKIKPGKVLDKFGKIVGKVLKQLKKVSAVAKV    87
TRIIN_18920644 MSKFWLLLLLVAAFQFAHSYPAAEYELDETTNDEVRQFIGDGYFEDEGDDGDEERFKIKPGKVLDKFGKIVSKVLKQLKKVSAVAKV    87

                ***********************:***************************************************************
            TI-8 AMKKGAALLKKMGVKISPLKCEEKTCKSCVIFKIPTENSFCLTIRFMKTNIATYLVVAGEINRKSKFEEKLKLGNMPRCVNVEGFIG   174
            TI-6 AMKKGAALLKKMGVKISPLKCEEKTCKSCVIFKIPTENSFCLTIRFMKTNIATYLVVAGEINRKSKFEEKLKLGNMPRCVNVEGFIG   174
TRIIN_18920644 AMKKGAALLKKMGVKISPLKCEENTCKSCVIFKIPTENSFCLTIRFMKTNIATYLVVAGEINRKSKFEEKLKLGNMPRCVNVEGFIG   174

                ***:*********************************************:************************************
            TI-8 KVCMKGIEGHAKSSSGQANVNFCLGLVAEKFGVGAKLCGIYANKKVRVKISPQLFPGATSLDGDIVKLDDNGEDATTLDVDEVEID   260
            TI-6 KVCLKGIEGHAKSSSGQANVNFCLGLVAEKFGVGAKLCGIYANKKVRVKVSPQLFPGATSLDGDIVKLDDNGEDATTLDVDEVEID   260
TRIIN_18920644 KVCMKGIEGHAKSSSGQANVNFCLGLVAEKFGVGAKLCGIYANKKVRVKISPQLFPGATSLDGDIVKLDDNGEDATTLDVDEVEID   260
```

Figures 9A and 9B.

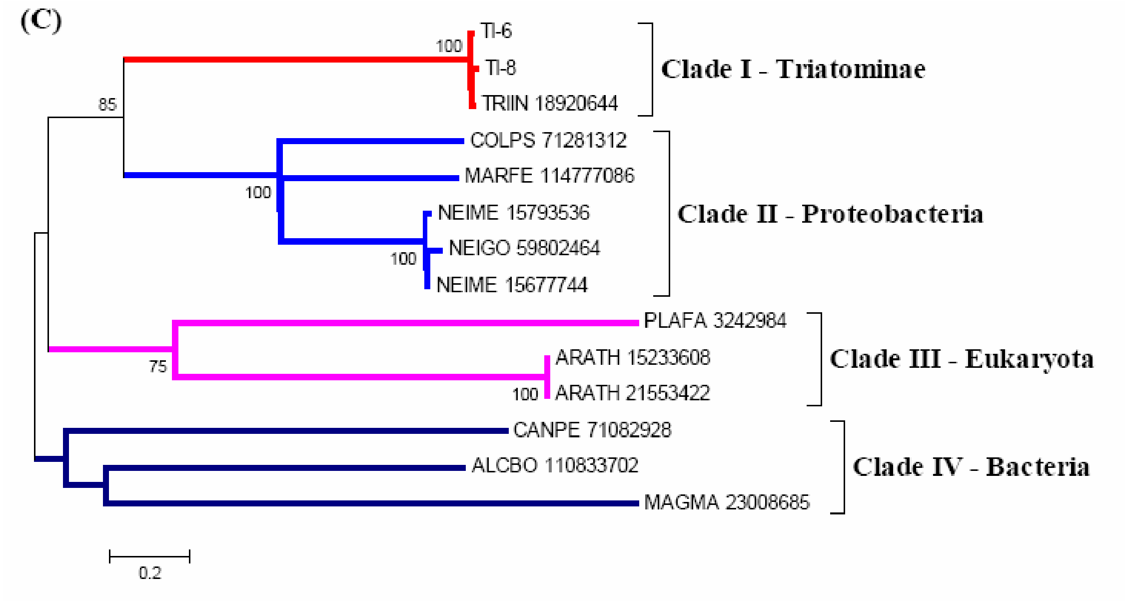


Figure 9. (A) ClustalW alignment of the unique family of trialysin salivary peptides. The sequences shown are from the cDNA library (TI-6, TI-8) and from trialysin from NCBI database (|gi18920644|). (B) ClustalW alignment of the trialysin sequences with other proteins that showed similarity to trialysin. (C) Neighbor-joining phylogram. The numbers in the phylogram nodes indicates percent bootstrap support for the phylogeny. The bar at the bottom indicates 20% amino acid divergence in sequences.

### 3.5.2.2 - Peptides with trialysin signal peptide

This cDNA library also revealed a group of short proteins sharing the signal peptide from trialysin protein. They do not share similarities with other proteins and their function is unknown.

#### 3.5.2.3 - Hemolysin-domain protein

Hemolysins are proteins considered to permeate target cells, forming pores in cytoplasmic membranes of erythrocytes, leukocytes, and other cells, causing the modification of cellular functions and/or lysis of host cells. Hemolysins are more often described in microorganisms, such as *Escherichia coli*, *Staphylococcus aureus*, *Clostridium septicum*, and others (Gouaux, 1998; Melton et al., 2004; Wassenaar, T.M., 2005). The family of hemolysins consists of secreted, water-soluble proteins that form transmembrane, polymeric channels. They are considered a virulence factor from these microorganisms and also a member of the RTX (repeats in toxin) protein family (Welch, 2001). The presence of amphipathic and hydrophilic domains confers to the protein an overall amphiphilic character, which explains its tendency both to aggregate and to interact with membranes (Soloaga et al., 1999; Schindel et al., 2001; Hyland et al., 2001).

We found contigs with proteins containing a hemolysin domain (Fig. 10). These proteins could act as cytolytic proteins, causing erythrocytes lysis in saliva and helping the early steps of the digestion process. Also, it could be related to defense, participating in lysis of microorganisms in insect saliva.

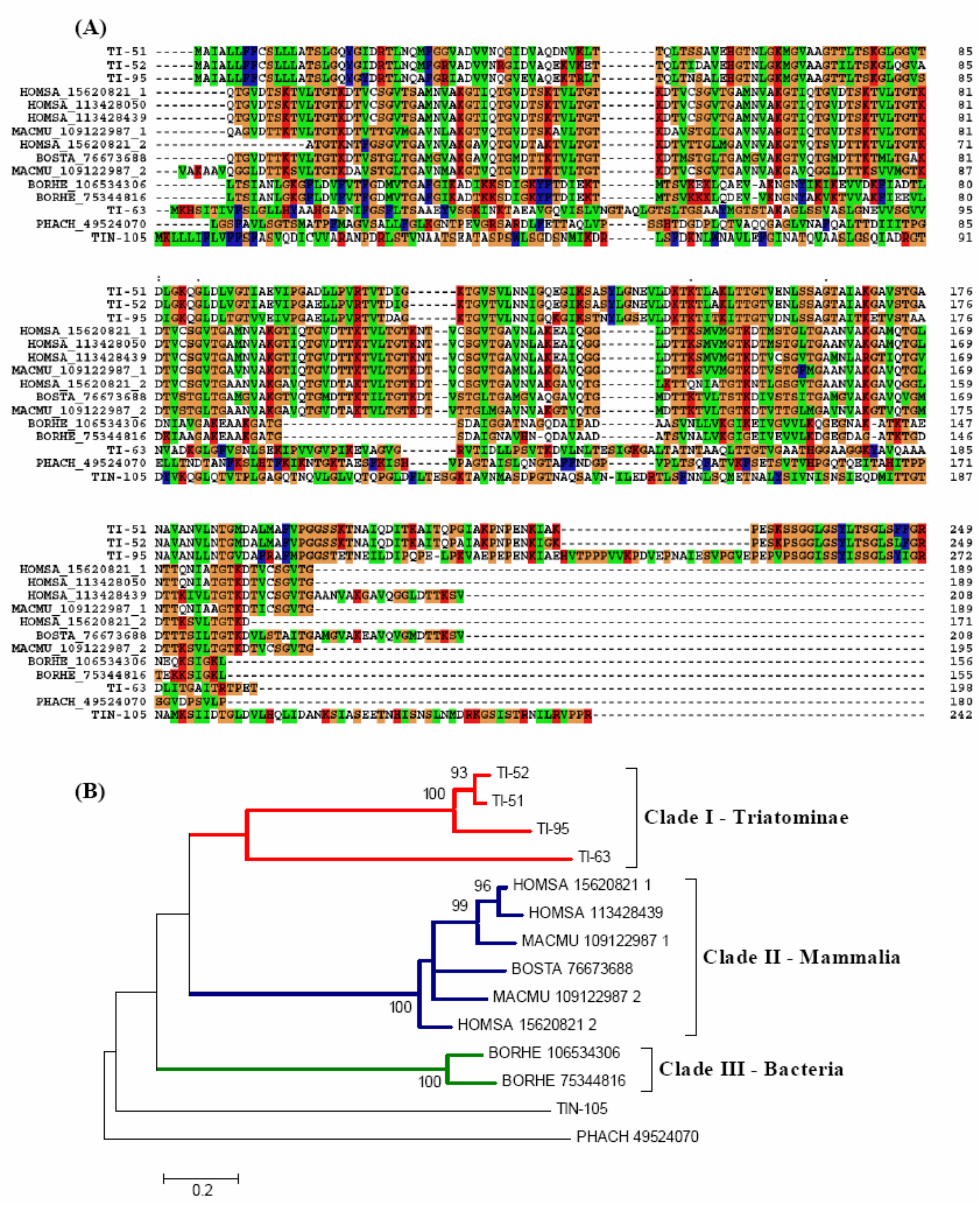
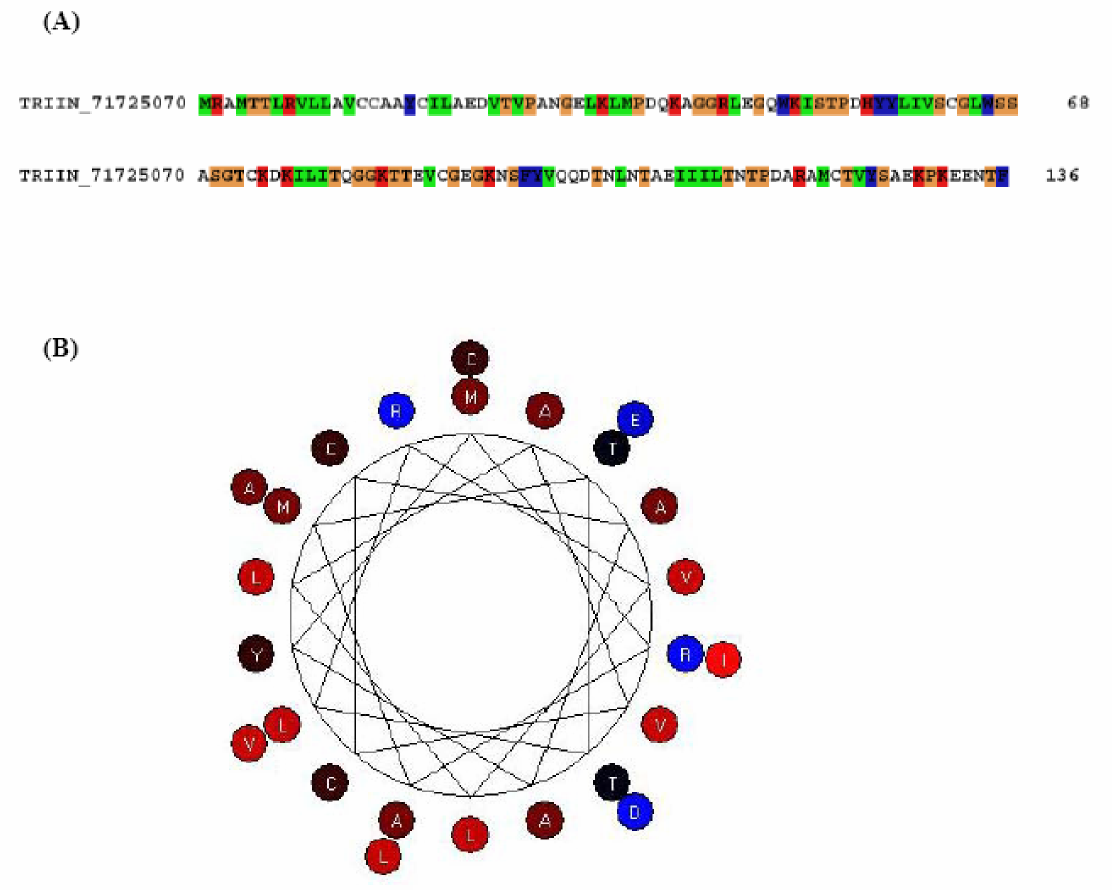


Figure 10. (A) ClustalW alignment of the mature forms of *T. infestans* hemolysin-domain containing proteins with other proteins with hemolysin-domain. The signal peptide region is not shown. (B) Phylogram. Numbers on branches indicate bootstrap value for 1000 trials. Bar indicates 20% distance in amino acid sequence.

### 3.5.2.4 - Triatox

One of the contigs matched a toxin previously described in *T. infestans* saliva as triatox (|gi71725070|) (Fig. 11A). We also found this protein in the 2D gel (Fig. 12). By secondary structure prediction, this protein shows to be an amphipathic α-helix (Fig. 11B). α-helical peptides are linear molecules that exist in aqueous media and become amphipathic helices upon interaction with hydrophobic membranes, like cecropin (Christensen et al., 1988), magainins (Zasloff, 1987), and melittins (Andreu et al., 1992). The amphipathic character of antimicrobial peptides makes them surface-active products, as their biological activity occurs at lipid membrane interfaces (Maget-Dana, 1999). Their amphipathy allows them to be both soluble in an aqueous medium such as the extracellular medium, and to diffuse towards polar/apolar interfaces such as the extracellular medium/cell membrane interface. Linear peptides that can assume an active, amphipathic, α-helical structure are among the most abundant in nature. They have evolved to act against several microbial targets and appear to represent an important role in innate defense (Giangaspero et al., 2001). The structure of triatox suggests it can function as an antimicrobial peptide and be related to the innate defense of the insect in the (SG)s.



(A)

```
TRIIN_71725070 MRAMTTLRVLLAVCCAAYCILAEDVTVPANGELKLMPDQKAGGRLEGQWKISTPDHYYLIVSCGLWSS     68

TRIIN_71725070 ASGTCKDKILITQGGKTTEVCGEGKNSFYVQQDTNLNTAEIIILTNTPDARAMCTVYSAEKPKEENTF   136
```

Figure 11. (A) Triatox sequence |gi71725070|. (B) Prediction of triatox (TI-57) secondary structure using BioEdit program.

### 3.5.2.5 - Thrombospondin-like

This protein shows weak similarity with larger proteins of *Plasmodium* annotated as “thrombospondin-related adhesive protein.” Thrombospondin domains usually bind to sulphated glycoconjugates found on cell surfaces, allowing for host-cell specificity of binding (Dubremetz, et al., 1998). The function is unknown in other genera.

### 3.6. 2D Gel electrophoresis/Mass spectrometry proteomic investigation

The (SG)s of *T. infestans* were shown to contain proteins such as apyrase and the pore-forming protein trialysin (Amino et al., 2002; Faudry et al., 2004). To obtain further information on the salivary proteins of *T. infestans*, electrophoresis of saliva was performed by two-dimensional (2D) SDS-PAGE followed by mass spectrometry. Figure 12 shows the pattern of separation of *T. infestans* salivary proteins by 2D-gel. We were able to identify several salivary lipocalins, and the proteins named “triatox” (|gi71725070|), salivary apyrase precursor (|gi34481604|), salivary inositol polyphosphate 5-phosphate (approximately 37-kDa apparent molecular mass), similar to pallidipin precursor, putative trialysin precursor, similar to lipocalin-like TiLipo77, and similar to pallidipin-2. Supplemental tables S1 and S2 have columns indicating the peptide sequences found in the proteomic experiment.

### 3.7 Salivary peptide identification by RP-HPLC/Mass spectrometry

Because the gel used for the proteomic experiment above does not well identify polypeptides below 10 kDa, we obtained a 10-kDa filtrate from 20 μL of *T. infestans* saliva, submitted it to RP-HPLC and performed tryptic digestion of the fractions to obtain peptide sequences by MS/MS. This allowed identification of four members of the short trialysin-like family, which have predicted mature molecular weights of 6.3 kDa. Several fragments matching the hemolysin-like proteins were also identified, indicating this protein family, whose members have more than 20 kDa, may be proteolytically processed in saliva, possibly by triapsin, the molecularly uncharacterised salivary serine protease of *T. infestans* (Amino et al., 2001). Supplemental tables S1 and S2 have columns indicating the peptides sequences found in this proteomic experiment.

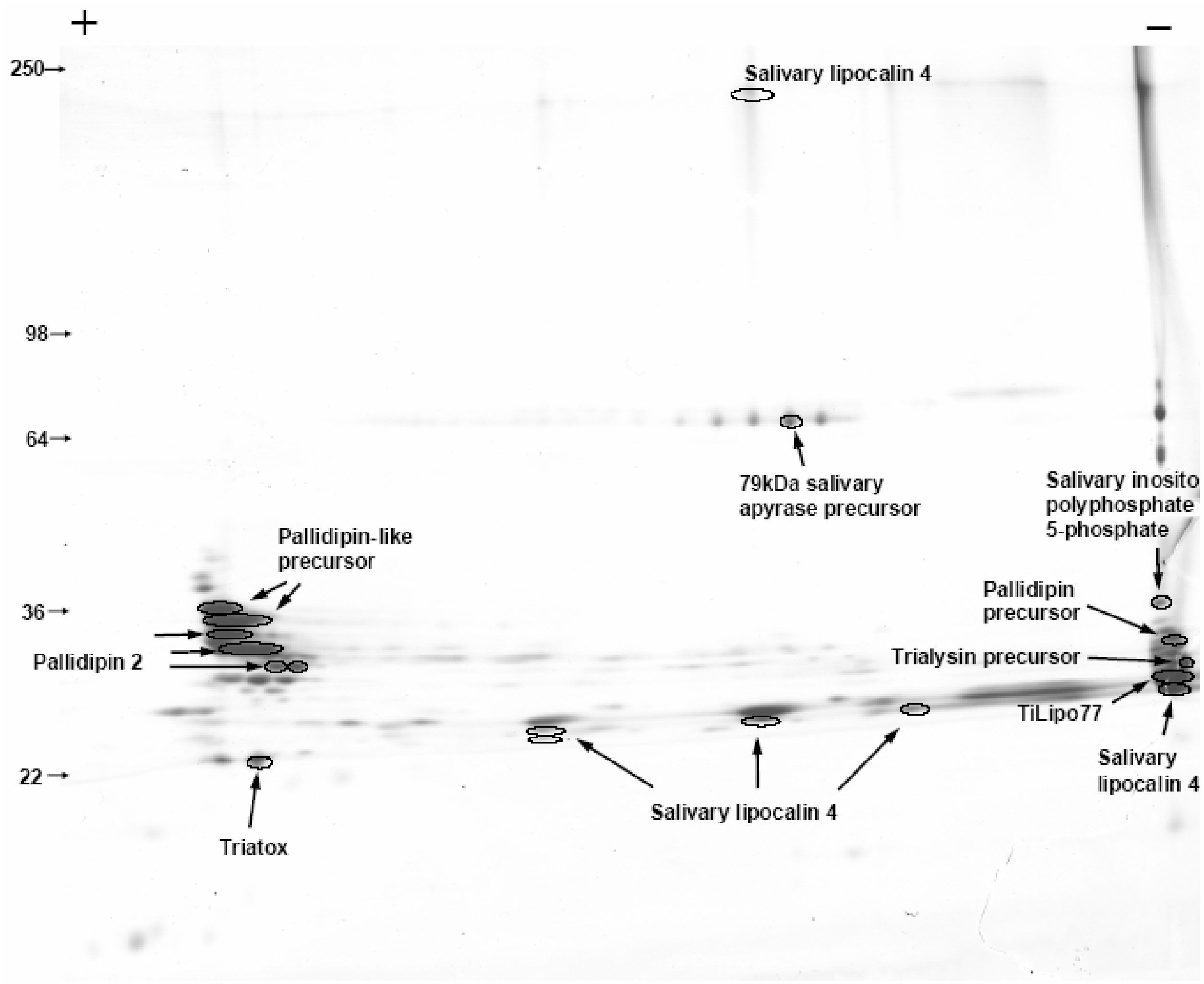


Figure 12. Two-dimensional (2D)-gel electrophoresis of *Triatoma infestans* salivary proteins. Numbers on the left indicate molecular weight marker positions in the gel. The + and – signs indicate the anode or cathode side of the isoelectrophocusing dimension, which ranged from pH 3-10. Gel bands that were identified to a protein (following tryptic digestion and mass spectrometry) are shown in the gel. In some cases, more than one band accounted for the same protein, possibly due to trailing or multiple isoforms. For experimental details, see Materials and Methods.

### 3.8. Concluding Remarks

In an attempt to improve our understanding of the variety of proteins and transcripts expressed in *T. infestans* (SG)s, we have performed a cDNA library and a 2D-gel using, respectively, mRNA and proteins from this same tissue. We described the set of cDNA present in the (SG)s of *Triatoma infestans*. Expression and bioassay of the novel proteins will ultimately characterise the salivary pharmacologic complexity from *T. infestans* evolution to blood feeding.

We believe this cDNA library should help our continuing effort to understand the evolution of blood sucking in vector arthropods and the discovery of novel pharmacologically active compounds.

Acknowledgments: This work was supported in part by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases. TCFA was a recipient of a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Ministry of Education, Brazil) fellowship. We are grateful to the NIAID Research Technologies Brach, directed by Dr. Robert Hohman, for support in the proteomic studies, to Nancy Shulman for editorial assistance, to Chuong Huynh from NCBI for help with posting the data, and Drs. Robert Gwadz and Katherine Zoon for encouragement and support.

## References

Altschul, S.F., Gish, W., 1996. Local alignment statistics. Methods Enzymol. 266, 460-480.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

Amino, R., Martins, R.M., Procopio, J., Hirata, I.Y., Juliano, M.A., Schenkman, S., 2002. Trialysin, a Novel Pore-forming Protein from Saliva of Hematophagous Insects Activated by Limited Proteolysis. J. Biol. Chem. 277, 6207-6213.

Amino, R., Tanaka, A.S., Schenkman, S., 2001. Triapsin, an unusual activatable serine protease from the saliva of the hematophagous vector of Chagas' disease *Triatoma infestans* (Hemiptera: Reduviidae). Insect Biochem. Mol. Biol. 31, 465-472.

Andersen, J.F., Gudderra, N.P., Francischetti, I.M.B., Ribeiro, J.M.C., 2005. The Role of Salivary Lipocalins in Blood Feeding by *Rhodnius prolixus*. Arch. Insect Biochem. Physiol. 58, 97-105.

Andersen, J.F., Guderra, N.P., Francischetti, I.M., Valenzuela, J.G., Ribeiro, J.M. 2004. Recognition of anionic phospholipid membranes by an antihemostatic protein from a blood-feeding insect. Biochemistry 43, 6987-6994.

Andersen, J.F., Ribeiro, J.M.C., 2006. A Secreted Salivary Inositol Polyphosphate 5-Phosphatase from a Blood-Feeding Insect: Allosteric Activation by Soluble Phosphoinositides and Phosphatidylserine. Biochemistry 45, 5450-5457.

Anderson, G.S., Belton, P., Kleider, N., 1993. Hypersensitivity of horses in British Columbia to extracts of native and exotic species of Culicoides (Diptera: Ceratopogonidae). J. Med. Entomol. 30, 657-663.

Andreu, D., Ubach, J., Boman, A., Wahlin, B., Wade, D., Merrifield, R.B., Boman, H.G., 1992. Shortened cecropin A-mellitin hybrids. Significant size reduction retains potent antibiotic activity. FEBS Lett. 296, 190-194.

Arcà, B., Lombardo, F., Francischetti, I.M.B., Pham, V.M., Mestres-Simon, M., Andersen, J.F., Ribeiro, J.M.C., 2007. An insight into the sialome of the adult female mosquito *Aedes albopictus*. Insect Biochemisty and Molecular Biology 37: 107-127.

Arcà, B., Lombardo, F., Valenzuela, J.G., Francischetti, I.M., Marinotti, O., Coluzzi, M., Ribeiro, J.M.C., 2005. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. J. Exp. Biol. 208, 3971-3986.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25-29.

Bartholomay, L.C., Fuchs, J.F., Cheng, L.-L., Beck, E.T., Vizioli, J., Lowenberger, C., Christensen, B.M., 2004. Reassessing the role of defensin in the innate immune response of the mosquito, *Aedes aegypti*. Insect Mol. Biol. 13, 125-132.

Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., Sonnhammer, E.L., 2000. The Pfam protein families database. Nucleic Acids Res. 28, 263-266.

Bonmantin, J.M., Bonnat, J.L., Gallet, X., Vovelle, F., Ptak, M., Reichhart, J.-M., Hoffmann, J.A., Keppi, E., Legrain, M., Achstetter, T., 1992. Two-dimensional 1H NMR study of recombinant insect defensin A in water: resonance assignments, secondary structure and global folding. J. Biomol. NMR 2, 235-256.

Broze Jr., G.J., 1995. Tissue factor pathway inhibitor and the current concept of blood coagulation. Blood Coagul. Fibrinolysis 6, S7-13.

Bulet, P., Charlet, M., Hetru, C., 2002. Antimicrobial peptides in insect immunity. In: Ezekowitz, A.B., Hoffmann, J.A., editors. Infectious disease: Innate immunity. Totowa, NJ: Humana Press; p. 89-107.

Bulet, P., Cociancich, S., Reuland, M., Sauber, F., Bischoff, R., Hegy, G., Van Dorsselaer, A., Hetru, C., Hoffmann, J.A., 1992. A novel insect defesin mediates the inducible antibacterial activity in larvae of the dragonfly *Aeschna cyanea* (Paleoptera, Odonata). Eur. J. Biochem. 209, 977-984.

Bulet, P., Hetru, C., Dimarcq, J.L., Hoffmann, D., 1999. Antimicrobial peptides in insects; structure and functions. Dev. Comp. Immunol. 23, 329-344.

Bulet, P., Stocklin, R., Menin, L., 2004. Anti-microbial peptides: from invertebrates to vertebrates. Immunol. Rev. 198, 169-184.

Calvo, E., Dao, A., Pham, V.M., Ribeiro, J.M.C., 2007. An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. Insect Biochem. Mol. Biol. 37, 164-175.

Campbell, C.L., Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp, T., Wilson, W.C., 2005. Midgut and salivary gland transcriptomes of the arbovirus vector Culicoides sonorensis (Diptera: Ceratopogonidae). Insect Mol. Biol. 14, 121-136.

Campos, I.T., Amino, R., Sampaio, C.A., Auerswald, E.A., Friedrich, T., Lemaire, H.G., Schenkman, S., Tanaka, A.S., 2002. 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. 32, 991-997.

Ceraul, S.M., Sonenshine, D.E., Ratzlaff, R.E., Hynes, W.L., 2003. An arthropod defensin expressed by the hemocytes of the American dog tick, *Dermacentor variabilis* (Acari:Ixodidae). Insect Biochem. Mol. Biol. 33, 1099-1103.

Christensen, B., Fink, J., Merrifield, R.B., Mauzerall, D., 1988. Channel-forming properties of cecropins and related model compounds incorporated into planar lipid membranes. Proc. Natl. Acad. Sci. USA 85, 5072-5076.

Czibener, C., La Torre, J.L., Muscio, O., Ugalde, R.A., Scodeller, E.A., 2000. Nucleotide sequence analysis of *Triatoma* virus shows that it is a member of a novel group of insect RNA viruses. J. General Virol. 81, 1149-1154.

Deyu, Z., and Leal, W., 2002. Conformational isomers of insect odorant-binding proteins. Arch. Biochem. Biophys. 397, 99-105.

Dias, J.C., 1987. Control of Chagas disease in Brazil. Parasitology Today 3(11), 336-441.

Dubremetz, J.F., Garcia-Réguet, N., Conseil, V., Fourmaux, M.N., 1998. Apical organelles and host-cell invasion by Apicomplexa. Int. J. Parasitol. 28, 1007-1013.

Erneux, C., Govaerts, C., Communi, D., Pesesse, X., 1998. The diversity and possible functions of the inositol polyphosphate 5-phosphatases. Biochim. Biophys. Acta 1436, 185-199.

Faudry, E., Lozzi, S.P., Santana, J.M., D'Souza-Ault, M., Kieffer, S., Felix, C.R., Ricart, C.A., Sousa, M.V., Vernet, T., Teixeira, A.R., 2004. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. J. Biol. Chem. 279(19), 19607-19613.

Fink, E., Rehm, H., Gippner, C., Bode, W., Eulitz, M., Machleidt, W., Fritz, H., 1986. The primary structure of bdellin B-3 from the leech *Hirudo medicinalis*. Bdellin B-3 is a compact proteinase inhibitor of a "non-classical" Kazal type. It is present in the leech in a high molecular mass form. Biol. Chem. Hoppe Seyler 367, 1235-1242.

Flower, D.R., 1995. Multiple molecular recognition properties of the lipocalin protein family. J. Mol. Recogn. 8, 185-195.

Flower, D.R., North, A.C., Attwood, T.K., 1993. Structure and sequence relationships in the lipocalins and related proteins. Protein Sci. 2, 753-761.

Flower, D.R., North, A.C., Sansom, C.E. 2000. The lipocalin protein family: structural and sequence overview. Biochim. Biophys. Acta 1482, 9-24.

Fogaça, A.C., Lorenzini, D.M., Kaku, L.M., Esteves, E., Bulet, P., Daffre, S., 2004. Cysteine-rich antimicrobial peptides of the cattle tick *Boophilus microplus*: isolation, structural characterization and tissue expression profile. Dev. Comp. Immunol. 28, 191-200.

Forêt, S., and Maleszka, R., 2006. Function and evolution of a gene family encoding odorant binding-like protein in a social insect, the honey bee (*Apis mellifera*). Genome Res. 16, 1404-1413.

Francischetti, I.M., Ribeiro, J.M., Champagne, D., Andersen, J., 2000. Purification, cloning, expression, and mechanism of action of a novel platelet aggregation inhibitor from the salivary gland of the blood-sucking bug, Rhodnius prolixus. J. Biol. Chem. 275(17), 12639-12650.

Francischetti, I.M., Valenzuela, J.G., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2002. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. J. Exp. Biol. 205, 2429-2451.

Francischetti, I.M., Lopes, A.H., Dias, F.A., Pham, V.M., Ribeiro, J.M., 2007. An insight into the sialotrascriptome of the seed-feeding bug, *Oncopeltus fasciatus*. Insect Biochem. Mol. Biol. 37, 903-910.

Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H.G., Hoffken, H.W., Reuschenbach, P., Otte, M., Dodt, J., 1993. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. J. Biol. Chem. 268, 16216-16222.

Friend, W.G., Smith, J.J.B., 1977. Factors affecting feeding by bloodsucking insects. Ann. Rev. Entomol. 22, 309-331.

Fuentes-Prior, P., Noeske-Jungblut, C., Donner, P., Schleuning, W.-D., Huber, R., Bode, W., 1997. Structure of thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. Proc. Natl. Acad. Sci. USA 94, 11845-11850.

Galindo, K., Smith, D.P., 2001. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. Genetics 159, 1059-1072.

Giangaspero, A., Sandri, L., Tossi, A., 2001. Amphipathic α helical antimicrobial peptides. A systematic study of the effects of structural and physical properties on biological activity. Eur. J. Biochem. 268, 5589-5600.

Gouaux, E., 1998. α-Hemolysin from *Staphylococcus aureus*: An Archetype of β-Barrel, Channel-Forming Toxins. J. Structural Biol. 121, 110-122.

Guo, S., Stolz, L.E., Lemrow, S.M., York, J.D., 1999. SAC1-like domains of yeast SAC1, INP52, and INP53 and of human synaptojanin encode polyphosphoinositide phosphatases. J. Biol. Chem. 274, 12990-12995.

Hansen, J.E., Lund, O., Tolstrup, N., Gooley, A.A., Williams, K.L., Brunak, S., 1998. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconjugate J. 15, 115-130.

Hekmat-Scafe, D.S., Dorit, R.L., Carlson, J.R., 2000. Molecular Evolution of Odorant-Binding Protein Genes OS-E and OS-F in *Drosophila*. Genetics 155, 117-127.

Hekmat-Scafe, D.S., Scafe, C.R., McKinney, A.J., Tanouye, M.A., 2002. Genome-Wide Analysis of the Odorant-Binding Protein Gene Family in *Drosophila melanogaster*. Genome Res. 12, 1357-1369.

Hoffman, D.R., 1993. Allergens of Hymenoptera venom. XXV: The amino acid sequences of antigen 5 molecule sand the structural basis of antigenic cross-reactivity. J. Allergy Clin. Immunol. 92, 707-716.

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9: 868-877.

Hyland, C., Vuillard, L., Hughes, C., Koronakis, V., 2001. Membrane interaction of *Escherichia coli* hemolysin: flotation and insertion-dependent labeling by phospholipids vesicles. J. Bacteriol. 183, 5364-5370.

Ishibashi, J., Saido-Sakanaka, H., Yang, J., Sagisaka, A., Yamakawa, M., 1999. Purification, cDNA cloning and modification of a defensin from the coconut rhinoceros beetle, *Oryctes rhinoceros*. Eur. J. Biochem. 266, 616-623.

Jiang, H., Wang, Y., Huang, Y., Mulnix, A.B., Kadel, J., Cole, K., Kanost, M.R., 1996. Organization of serpin gene-1 from Manduca sexta: evolution of a family of alternate exons encoding the reactive site loop. J. Biol. Chem. 271, 28017-23.

Johns, R. Sonenshine, D.E., Hynes, W.L., 2001. Identification of a defensin from the hemolymph of the American dog tick, *Dermacentor variabilis*. Insect Biochem. Mol. Biol. 31, 857-865.

Kanost, M.R., 1999. Serine proteinase inhibitors in arthropod immunity. Develop. Comp. Immunol. 23, 291-301.

King, T.P., Spangfort, M.D., 2000. Structure and biology of stinging insect venom allergens. Int. Arch. Allergy Immunol. 123(2), 99-106.

Krieger, J., and Breer, H., 1999. Olfactory reception in invertebrates. Science 286, 720-723.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 5, 150-163.

Kurotaki, T., Narayama, K., Oyamada, T., Yoshikawa, H., Yoshikawa, T., 1994. Immunopathological study on equine insect hypersensitivity ("kasen") in Japan. J. Comp. Pathol. 110, 145-152.

Lamberty, M., Ades, S., Uttenweiler-Joseph, S., Brookharts, G., Bushey, D. Hoffmann, J.A., Bulet, P., 1999. Insect immunity: isolation from the lepidopteran *Heliothis virescens* of a novel insect defensin with potent antifungal activity. J. Biol. Chem. 274, 9320-9326.

Lavoipierre, M.M., 1965. Feeding mechanism of blood-sucking arthropods. Nature (London) 208, 302-303.

Letunic, I., Goodstadt, L., Dickens, N.J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R.R., Ponting, C.P., Bork, P., 2002. Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res. 30, 242-244.

Li, S., Kwon, J., Aksoy, S., 2001. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. Insect Mol. Biol. 10(1), 69-76.

Liang, Z., Söderhäll, K., 1995. Isolation of cDNA encoding a novel serpin of crayfish hemocytes. Comp. Biochem. Physiol. 112B, 385-391.

Lopez, L., Morales, G., Ursic, R., Wolff, M., Lowenberger, C., 2003. Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of Chagas disease. Insect Biochem. Mol. Biol. 33, 439-447.

Lovato, D.V., Campos, I.T.N., Amino, R., Tanaka, A.S., 2006. The full-length cDNA of anticoagulant protein infestin revealed a novel releasable Kazal domain, a neutrophil elastase inhibitor lacking anticoagulant activity. Biochimie 88, 673-681.

Maget-Dana, R., 1999. The monolayer technique: a potent tool for studying the interfacial properties of antimicrobial and membrane-lytic peptides and their interactions with lipid membranes. Biochim. Biophys. Acta 1462, 109-140.

Marchler-Bauer, A., Panchenko, A.R., Shoemaker, B.A., Thiessen, P.A., Geer, L.Y., Bryant, S.H., 2002. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res. 30, 281-283.

Marejus, P.W., Kisseleva, M.V., Norris, F.A., 1999. The role of phosphatases in inositol signaling reactions. J. Biol. Chem. 275, 20110-20116.

Martins, R.M., Sforça, M.L., Amino, R., Juliano, M.A., Oyama, S., Jr., Juliano, L., Pertinhez, T.A., Spisni, A., Schenkman, S., 2006. Lytic Activity and Structural Differences of Amphipathic Peptides Derived from Trialysin. Biochemistry 45, 1765-1774.

Megraw, T., Kaufman, T.C., Kovalick, G.E., 1998. Sequence and expression of Drosophila Antigen 5-related 2, a new member of the CAP gene family. Gene 222(2), 297-304.

Mellor, P.S., Boorman, J., Baylis, M., 2000. *Culicoides* biting midges: their role as arbovirus vectors. Annu. Rev. Entomol. 45, 307-340.

Melton, J.A., Parker, M.W., Rossjohn, J., Buckley, J.T., Tweten, R.K., 2004. The Identification and Structure of the Membrane-spanning Domain of the *Clostridium septicum* Alpha Toxin. J. Biol. Chem. 279, 14315-14322.

Moffitt, J.E., Venarske, D., Goddard, J., Yates, A.B., de Shazo, R.D., 2003. Allergic reactions to *Triatoma* bites. Ann. Allergy Asthma Immunol. 91, 122-128.

Montfort, W.R., Weichsel, A., Andersen, J.F., 2000. Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. Biochim. Biophys. Acta 1482 (1-2), 110-118.

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

Nobile, M., Noceti, F., Prestipino, G., Possani, L.D., 1996. Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules. Exp. Brain Res. 110, 15-20.

Noeske-Jungblut, C., Haendler, B., Donner, P., Alagon, A., Possani, L., Scheleuning, W.-D. 1995. Triabin, a highly potent exosite inhibitor of thrombin. J. Biol. Chem. 270, 28629-28634.

Noeske-Jungblut, C., Krätzschmar, J., Haendler, B., Alagon, A., Possani, L., Verhallen, P., Donner, P., Scheleuning, W.-D., 1994. An inhibitor of collagen-induced platelet aggregation from the saliva of *Triatoma pallidipennis*. J. Biol. Chem. 269, 5050-5053.

O-glycosylation sites on proteins were predicted with NetOGlyc: http://www.cbs.dtu.dk/services/NetOGlyc/

Paddock, C.D., McKerrow, J.H., Hansell, E., Foreman, K.W., Hsieh, I., Marshall, N. 2001. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. J. Immunol. 167, 2694-2699.

Paesen, G.C., Adams, P.L., Nuttall, P.A., Stuart, D.L. 2000. Tick histamine-binding proteins: lipocalins with a second binding cavity. Biochim. Biophys. Acta 1482, 92-101.

Page, R.D., 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357-358.

Paiva-Silva, G.O., Sorgine, M.H.F., Benedetti, C.E, Meneghini, R., Almeida, I.C., Machado, E.A., Dansa-Petretski, M., Yepiz-Plascencia, G., Law, J.H., Oliveira, P.L., Masuda, H., 2002. On the biosynthesis of *Rhodnius prolixus* heme-binding protein. Insect Biochem. Mol. Biol. 32, 1533-1541.

Pelosi, P., 1994. Odorant-binding proteins. Crit. Rev. Biochem. Mol. Biol. 29, 199-228.

Pelosi, P., 1996. Perireceptor events in olfaction. J. Neurobiol. 30, 3-19.

Pelosi, P., Calvello, M., Ban, L., 2005. Diversity of odorant-binding proteins and chemosensory proteins in insects. Chem. Senses 30, i291-i292.

Prestwich, G.D., Du, G., LaForest, S., 1995. How is pheromone specificity encoded in proteins? Chem. Senses 20, 461-469.

Prevot, P.-P., Adam, B., Boudjeltia, K.Z., Brossard, M., Lins, Cauchie, P., Brasseur, R., Vanhaeverbeek, M., Vanhamme, L., Godfroid, E., 2006. Anti-hemostatic Effects of a Serpin from the Saliva of the Tick *Ixodes ricinus*. J. Biol. Chem. 281, 26361-26369.

Raj, P.A., and Dentino, A.R., 2002. Current status of defensins and their role in innate and adaptive immunity. FEMS Microbiol. Lett. 206, 9-18.

Ribeiro, J.M.C., Garcia, E.S., 1981. The role of saliva in feeding in *Rhodnius prolixus*. J. Exp. Biol. 94, 219-230.

Ribeiro, J.M., Francischetti, I.M., 2003. Role of arthropod saliva in blood feeding: Sialome and post-sialome perspectives. Annu. Rev. Entomol. 48, 73-88.

Ribeiro, J.M.C, Charlab, R., Pham, V.M., Garfield, M., Valenzuela, J.G., 2004b. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. Insect Biochem. Mol. Biol. 34, 543-563.

Ribeiro, J.M.C., 1995. Blood-feeding arthropods: Live syringes or invertebrate pharmacologists? Infect. Agents Dis. 4, 143-152.

Ribeiro, J.M.C., Andersen, J., Silva-Neto, M.A.C., Pham, V.M., Garfield, M.K., Valenzuela, J.G., 2004a. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. Insect Biochem. Mol. Biol. 34, 61-79.

Ribeiro, J.M.C., Arcà, B., Lombardo, F., Calvo, E., Pham, V.M., Chandra, P.K., Wikel, S.K., 2007. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. BMC Genomics 8, 6.

Rudenko, N., Golovchenko, M., Edwards, M.J., Grubhoffer, L., 2005. Differential expression of *Ixodes ricinus* tick genes induced by blood feeding or *Borrelia burgdorferi* infection. J. Med. Entomol. 42, 36-41.

Sangamnatdej, S., Paesen, G.C., Slovak, M., Nuttall, P.A. 2002. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol. Biol. 11, 79-86.

Sant'Anna, M.R., Araújo, J.G., Pereira, M.H., Pesquero, J.L. Diotaiuti, L., Lehane, S.M., Lehane, M.J. 2002. Molecular cloning and sequencing of salivary gland-specific cDNAs of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera: Reduviidae). Insect Mol. Biol. 11, 585-593.

Santos, A., Ribeiro, J.M., Lehane, M.J., Gontijo, N.F., Veloso, A.B., Sant'anna, M.R., Nascimento Araújo, R., Grisard, E.C., Pereira, M.H., 2007. The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). Insect Biochem. Mol. Biol. 37(7), 702-12.

Schaffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V., Altschul, S.F., 2001. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Res. 29, 2994-3005.

Schindel, C., Zitzer, A., Schulte, B., Gerhards, A., Stanley, P., Hughes, C., Koronakis, V., Bhakdi, S., Palmer, M., 2001. Interaction of *Escherichia coli* hemolysin with biological membranes. A study using cysteine scanning mutagenesis. Eur. J. Biochem. 268, 800-808.

Shin, S.W., Park, S.-S., Park, D.-S., Kim, M.G., Kim, S.C., Brey, P.T., Park, H.-Y., 1998. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. Insect Biochem. Mol. Biol. 28, 827-837.

Söderhäll, K., Cerenius, L., 1998. Role of the prophenoloxidase-activating system in invertebrate immunity. Curr. Opin. Immunol. 10, 23-28.

Soloaga, A., Veiga, M.P., Garcia-Segura, L.M., Ostolaza, H., Brasseur, R., Goni, F.M., 1999. Insertion of *Escherichia coli* alpha-haemolysin in lipid bilayers as a non-transmembrane integral protein: prediction and experiment. Mol. Microbiol. 31, 1013-1024.

Sommerhoff, C.P., Sollner, C., Mentele, R., Piechottka, G.P., Auerswald, E.A., Fritz, H., 1994. A Kazal-type inhibitor of human mast cell tryptase: isolation from the medical leech *Hirudo medicinalis*, characterization, and sequence analysis. Biol. Chem. Hoppe Seyler 375, 685-694.

Stintzi, A., Heitz, T., Prasard, V., Wiedemann-Merdinoglu, S., Kauffmann, S., Geoffroy, P., Legrand, M., Fritig, B., 1993. Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie 75, 687-706.

Stubbs, M.T., Morenweiser, R., Sturzebecher, J., Bauer, M., Bode, W., Huber, R., Piechottka, G.P., Matschiner, G., Sommerhoff, C.P., Fritz, H., Auerswald, E.A., 1997. The three-dimensional structure of recombinant leech-derived tryptase inhibitor in complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition. J. Biol. Chem. 272, 19931-19937.

Szyperski, T., Fernandez, C., Mumenthaler, C., Wuthrich, K., 1998. Structure comparison of human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system. Proc. Nat. Acad. Sci. USA 95, 2262-2266.

Takác, P., Nunn, M.A., Mészáros, J., Pechánová, O., Vrbjar, N., Vlasáková, P., Kozánek, M., Kazimírová, M., Hart, G., Nuttall, P.A., Labuda, M., 2006. Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) salivary glands. J. Exp. Biol. 209, 343-352.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: an updated version includes eukaryotes. BioMed Central Bioinform. 4, 41.

Terra, W.R., Ferreira, C., Jordão, B.P., Dillon, R.J., 1996. Biology of the Insect Midgut. Chapman & Hall, London.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Toker, A., Cantley, L.C., 1997. Signalling through the lipid products of phosphoinositide-3-OH kinase. Nature 387, 673-676.

Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2003. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. Insect Biochem. Mol. Biol. 33, 717-732.

Valenzuela, J.G., Francischetti, I.M.B., Pham, V.M., Garfield, M.K., Mather, T.N., Ribeiro, J.M.C., 2002a. Exploring the sialome of the tick, *Ixodes scapularis*. J. Exp. Biol. 205, 2843-2864.

Valenzuela, J.G., Pham, V.M., Garfield, M.K., Francischetti, I.M., Ribeiro, J.M.C., 2002b. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. Insect Biochem. Mol. Biol. 32, 1101-1122.

Van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kröger, B., Höffken, W., Bode, W., 1995. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. EMBO J. 14, 5149-5157.

Vogt, R.G., Prestwich, G.D., Lerner, M.R., 1991. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. J. Neurobiol. 22, 74-84.

Wassenaar, T.M., 2005. Use of antimicrobial agents in veterinary medicine and implications for human health. Crit. Rev. Microbiol. 31, 155-169.

Welch, R.A., 2001. RTX toxin structure and function: a story of numerous anomalies and few analogies in toxin biology. Curr. Top. Microbiol. Immunol. 257, 85-111.

Yamazaki, Y., Koike, H., Sugiyama, Y., Motoyoshi, K., Wada, T., Hishinuma, S., Mita, M., Morita, T., 2002. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. Eur. J. Biochem. 269, 2708-2715.

Yamazaki, Y., Morita, T., 2004. Structure and function of snake venom cysteine-rich secretory proteins. Toxicon 44, 227-231.

Zasloff, M., 1987. Magainins, a class of antimicrobial peptides from *Xenopus* skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc. Natl. Acad. Sci. USA 84, 5449-5453.

# Materiais e Métodos

# Materiais e Métodos – Experimentos Adicionais

## Amplificação de genes selecionados da biblioteca de cDNA por PCR

Alguns genes da biblioteca de cDNA de glândula salivar de *T. infestans* foram escolhidos para expressão em sistema heterólogo para melhor caracterização de seus produtos. Os genes escolhidos codificam proteínas homólogas a lipocalinas e são teoricamente secretados de acordo com predição pelo programa SignalP: TINL-P2-B02, TINM-P6-F06, TINM-P7-C08, TINM-P7-D08 e TINM-P10-C04.

O fragmento referente ao peptídio sinal, de acordo com a predição do programa SignalP, foi retirado de cada seqüência de interesse. Os clones escolhidos foram amplificados pela técnica de PCR (reação de polimerização em cadeia), preservando a fase aberta de leitura (ORF). Os iniciadores específicos sintetizados contêm sítios para *NdeI* (senso) ou *XhoI* (antisenso), como apresentado na Tabela A.

O DNA molde utilizado para a amplificação dos clones foi o fago obtido da biblioteca de cDNA, que possue a ORF completa do gene de interesse, em um volume de 4,0 μL. Os iniciadores, magnésio e dNTPs foram utilizados na concentração final de 0,2; 2,0 e 0,2 mM, respectivamente. A enzima usada foi a Taq DNA polimerase *High Fidelity* (Invitrogen) que minimiza a inserção de erros na seqüência sintetizada. As condições da PCR foram: desnaturação inicial a 94 ºC por 2 min seguida por 30 ciclos de desnaturação a 94 ºC por 30 s, anelamento a 55 ºC por 30 s e extensão a 68 ºC por 1 min. Por fim, as reações eram estendidas por mais 5 min a 68 ºC.

Cada produto resultante da PCR, contendo um códon para metionina (ATG) diretamente *upstream* ao primeiro códon da proteína madura, foi clonado no vetor TOPO-TA (Invitrogen) e o plasmídeo gerado utilizado para transformar *E. coli* TOP-10 (Invitrogen). Esse plasmídeo possui o gene *lacZ* que produz a enzima beta-galactosidase. A presença desta enzima faz com que a colônia de bactéria que possui este plasmídeo seja azul quando colocada em presença de X-gal (5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo), pois a quebra do composto químico X-gal pela enzima beta-galactosidase forma um produto de cor azul. Quando adiciona-se um fragmento exógeno de DNA, esse fragmento é inserido no sítio de restrição localizado na região do gene que codifica a enzima beta-galactosidase. Assim, o gene *lacZ* é interrompido e a enzima beta-galactosidase não é produzida. Logo, as colônias de bactéria que possuírem o gene l*acZ* intacto irão produzir colônias azuis, enquanto que as colônias de bactéria com o inserto irão produzir colônias brancas. Os clones recombinantes, obtidos após seleção de colônias azuis/brancas, tiveram os respectivos insertos amplificados por PCR. Uma das colônias crescidas contendo o plasmídeo transformante foi escolhida, o DNA plasmidial isolado e digerido com *NdeI* e *XhoI* para liberação do inserto.

## Clonagem no vetor de expressão

Os genes de interesse foram clonados nos sítios de clivagem das enzimas *NdeI* e *XhoI* do vetor de expressão pET-17b (Novagen). A ligação do fragmento de DNA contendo o gene ao vetor de expressão pET-17b, previamente linearizado com *NdeI* e *XhoI*, foi feita com a T4 DNA ligase (Invitrogen) nas condições indicadas pelo fabricante. Cada produto de ligação foi utilizado para transformar *E. coli* BL21(DE3)pLys-S. Os clones contendo plasmídeos transformantes foram selecionados em meio de cultura sólido Luria-Ágar (Invitrogen) com 100 μg/mL de ampicilina (Sigma) e 35 μg/mL de cloranfenicol. As colônias crescidas em placas que apresentaram insertos do tamanho esperado, após PCR, foram utilizadas para expressão da proteína recombinante em meio líquido, como descrito adiante.

**Tabela A – Seqüência dos Oligonucleotídeos.** As seqüências em azul representam o sítio de restrição para a enzima *NdeI*. As seqüências em vermelho representam o sítio de restrição para a enzima *XhoI*.

| Clone | Tipo de oligonucleotídeo | Seqüência |
|---|---|---|
| TINL-P2-B02 | Senso | 5’GCCATATGCAAAACTCCGGTTGCGAACTGCAG3’ |
|  | Antisenso | 5’GCCTCGAGTCAACATTTCAGGTTTTTGAAATTGC3’ |
| TINM-P6-F06 | Senso | 5’GCCATATGGATTATCCATCTATTGAAAACTGCACTCAC3’ |
|  | Antisenso | 5’GCCTCGAGTTAGGACGCTTTATTTTTCTTTTTGGATGG3’ |
| TINM-P7-C08 | Senso | 5’GCCATATGGAAGAGTGCGTACTCAAGCCAGGT3’ |
|  | Antisenso | 5’GCCTCGAGTTAACAACGAACAGTTTGACGAG3’ |
| TINM-P7-D08 | Senso | 5’GCCATATGGATTATCCGCCAATTGAAAAATGC3’ |
|  | Antisenso | 5’GCCTCGAGTTAGGACGCTTTATTTTTCTTTTTGGATGG3’ |
| TINM-P10-C04 | Senso | 5’GCCATATGCAAAAGAACGGTTGCAACGTGCCG3’ |
|  | Antisenso | 5’GCCTCGAGTCAACACAGGTTTTGGAAATCCG3’ |

## Expressão de Proteínas em Sistema Heterólogo e Purificação

Para a produção da proteína, cada clone foi crescido em meio LB (Luria-Bertani) a 37 $^{o}$C, sob agitação a 250 rpm, até que o valor da densidade óptica a 600 nm ($OD_{600}$) fosse 0,6. Então, isopropil-1-tio-β-D-galactopiranosídeo (IPTG) foi adicionado a uma concentração final de 1 mM e os frascos foram agitados nas mesmas condições anteriores por mais 3 h. Ao final deste período, as células foram coletadas por centrifugação, lavadas com Tris-HCl 20 mM, pH 8,0, ressuspensas em 75 mL do mesmo tampão e lisadas por meio de sonicação a 4 °C.

Sedimento e sobrenadante obtidos após centrifugação do lisado a 9.000 rpm por 20 min foram submetidos a NuPAGE 12% e o gel corado com *Coomassie Blue*. O sedimento insolúvel contendo os corpos de inclusão foi solubilizado com Tris-HCl 20 mM, pH 8,0; Triton X-100 1%, e centrifugado a 9.000 rpm por 15 min. O extrato obtido foi lavado três vezes com Tris-HCl 20 mM pH 8,0, passando por uma centrifugação entre as lavagens. A seguir, a proteína foi solubilizada em 25 mL de Tris-HCl 20 mM pH 8,0, contendo hidrocloreto de guanidina 6,0 M e ditiotreitol (DTT) 10 mM. O material foi diluído em 4 L de Tris-HCl 20 mM, arginina 0,4 M, pH 8,0, mantido sob agitação por 2 horas e, depois, a 4 ºC durante a noite. Após concentração por meio de um equipamento para filtração em fluxo tangencial, a proteína solúvel foi purificada.

## Purificação por HPLC - Exclusão Molecular e Mono S

A proteína de interesse foi purificada a partir do extrato protéico por Cromatografia Líquida de Alta Eficiência (HPLC – *High Performance Liquid Chromatography*). Inicialmente a coluna de exclusão molecular Sephacryl S-100 16/16 HiPrep (Amersham Biotech.) foi utilizada para a purificação. Quando necessário, uma segunda etapa cromatográfica foi realizada em coluna de troca iônica Mono S (Amersham Biotech.). Um volume de 25 mL de extrato protéico era aplicado na coluna

Sephacryl previamente equilibrada com tampão Tris-HCl 20 mM, NaCl 150 mM, pH 8,0. Frações de 4 mL eram coletadas e imediatamente colocadas em banho de gelo. Uma alíquota (20 μL) de cada uma dessas frações foi submetida à eletroforese em gel NuPAGE 12%, em condições desnaturantes e redutoras, para análise de pureza. As frações contendo proteína pura foram concentradas utilizando Centricon 5 (Millipore). A presença de contaminantes indicava a aplicação da amostra em coluna Mono S previamente equilibrada com tampão fosfato de sódio 10 mM, pH 6,0. As proteínas retidas na coluna foram eluídas com gradiente linear de NaCl de zero a 1,0 M por 30 min, com fluxo de 0,5 mL/min. Frações de 1,0 mL eram coletadas e imediatamente colocadas em banho de gelo. Uma alíquota (15 μL) de cada uma dessas frações foi submetida à eletroforese em gel NuPAGE 12%, em condições desnaturantes e redutoras, para análise de pureza. As frações contendo a proteína pura foram agrupadas e concentradas em Centricon 5 (Millipore).

## Titulação Isotérmica por Calorimetria

Medidas de titulação isotérmica por calorimetria foram feitas em um calorímetro Microcal VP-ITC (Northhampton, MA). A proteína em estudo foi dialisada contra Tris-HCl 20mM, NaCl 0,15 M, pH 7,5, o qual também foi usado para preparar as soluções dos ligantes, para diluir a proteína e avaliar a linha de base de calor da diluição. Todas as soluções foram degaseificadas sob vácuo antes do uso. Proteínas na concentração de 2 μM foram inseridas na célula e o ligante na concentração de 25 μM foi usado na seringa. Os ligantes utilizados foram histamina, serotonina, adrenalina, noradrenalina, ADP e AMP. O calor de ligação foi medido à temperatura de 30 $^{o}$C. Após a subtração das medidas do calor de diluição, os dados de entalpia foram analisados com o modelo de ligação em um único sítio, utilizando o software original do Microcal.

# Resultados

# Resultados – Experimentos Adicionais

## Expressão de Proteínas em Sistema Heterólogo

Os clones escolhidos da biblioteca de cDNA de glândulas salivares do inseto – TINL-P2-B02, TINM-P6-F06, TINM-P7-C08, TINM-P7-D08 e TINM-P10-C04 – correspondem a lipocalinas e foram amplificados por PCR utilizando iniciadores específicos e, posteriormente, clonados no vetor de expressão. A lipocalina correspondente a cada clone foi expressa, solubilizada e purificada como descrito em Materiais e Métodos. Cada proteína recebeu o nome original de seu clone.

A proteína TINM-P7-C08 foi obtida em poucas quantidades após o re-enovelamento, provavelmente devido ao baixo rendimento na produção da proteína em corpos de inclusão. Portanto, não procedemos à purificação. A proteína TINL-P2-B02 foi purificada por exclusão molecular, utilizando somente a coluna Sephacryl S-100 no sistema HPLC. A proteína foi eluída no intervalo de 40 a 60 mL de tampão. A purificação das proteínas TINM-P6-F06, TINM-P7-D08 e TINM-P10-C04 foi realizada utilizando as colunas cromatográficas Sephacryl S-100 e Mono S no sistema HPLC. TINM-P7-D08 foi eluída no intervalo de 45 a 70 mL de tampão para a cromatografia de exclusão molecular e com gradiente de NaCl de 650 a 700 mM na cromatografia de troca iônica. Para TINM-P10-C04, a eluição ocorreu de 50 a 80 mL de tampão (exclusão molecular) e com gradiente de NaCl de 170 a 340 mM (troca iônica). O volume de 52 a 70 mL de tampão e o gradiente de 350 a 440 mM de NaCl permitiram a eluição de TINM-P6-F06 nas cromatografias de exclusão molecular e troca iônica, respectivamente. A cromatografia foi realizada como descrito em Materiais e Métodos. As frações coletadas e armazenadas em banho de gelo foram avaliadas em gel NuPAGE 12%, seguido de coloração por *Coomassie Blue* (Fig. 4). Após a concentração das frações puras em centricon, as lipocalinas foram submetidas à analise de titulação isotérmica por calorimetria (ITC).

As lipocalinas, descritas como proteínas carreadoras, são responsáveis pelo transporte de pequenas moléculas como histamina, serotonina e adrenalina, entre outras. Uma característica marcante nas lipocalinas é sua estrutura tridimensional bem conservada (Flower, 1995). A estrutura tridimensional das proteínas é apropriada para produzir sítios de ligação a ligantes, logo o critério universal para o correto dobramento de proteínas seria a medição da estequiometria e afinidade da ligação a ligantes. Uma das maneiras de efetuar essa medição seria pela titulação isotérmica por calorimetria (Brewer, 1999).

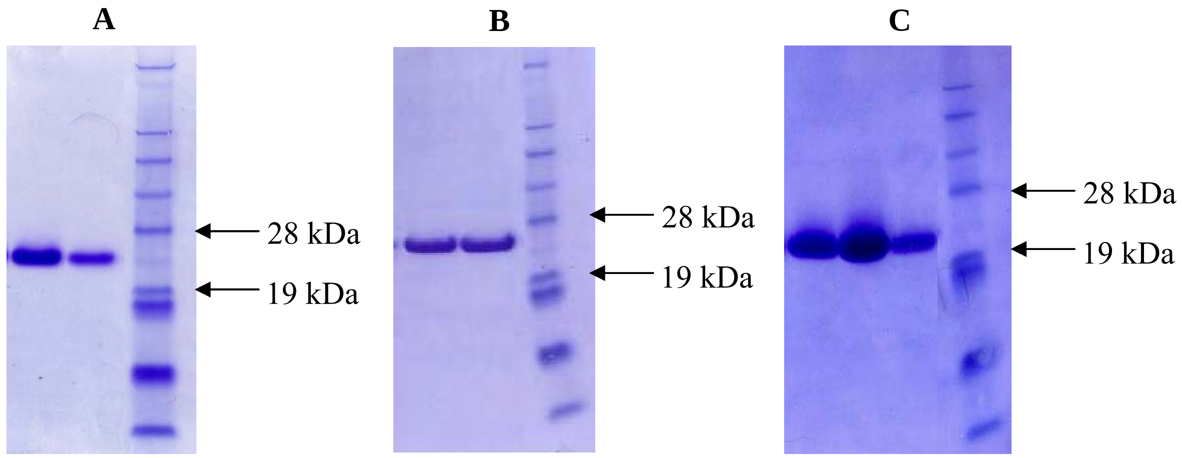


**Figura 4 – Análise em NuPAGE 12% da purificação das lipocalinas em HPLC.** A – TINL-P2-B02. B – TINM-P7-D08. C – TINM-P10-C04. Foram aplicados na coluna 10 mL de extrato protéico e a eluição das proteínas retidas foi realizada como descrito em Materiais e Métodos. As frações coletadas foram armazenadas em banho de gelo e submetidas à análise eletroforética. As bandas de proteína foram visualizadas por coloração por azul de *Coomassie*.

Um instrumento de ITC (titulação isotérmica por calorimetria) consiste em duas células idênticas compostas de material condutor térmico de alta eficiência cercado por uma câmara adiabática. Circuitos sensíveis à temperatura detectam diferença na temperatura entre as duas células e entre as células e a câmara. Em um experimento de ITC, a solução com a macromolécula é colocada na célula. Neste caso, as lipocalinas expressas em sistema heterólogo. O que é medido em um experimento de ITC é a energia tempo-dependente necessária para manter as temperaturas iguais na célula de referência e na célula contendo a amostra. Durante a injeção do titulante na célula com a amostra, calor é absorvido ou liberado, dependendo se a associação das moléculas é endotérmica ou exotérmica. O calor absorvido ou liberado durante a titulação calorimétrica é proporcional à fração de ligante ligado. Assim, é de extrema importância determinar com acurácia as concentrações inicias da macromolécula e do ligante. Nas injeções iniciais, todo ou quase todo o ligante adicionado é ligado à macromolécula, resultando em um amplo sinal exotérmico ou endotérmico, dependendo da natureza da associação. Com o aumento da concentração do ligante, a macromolécula torna-se saturada e, subseqüentemente, menos calor é liberado ou absorvido com a adição do titulante. A vantagem primária do ITC é que o sinal observado é o calor liberado ou absorvido do complexo formado. O único requerimento limitante para estudos por ITC é que a mudança de entalpia da ligação possa ser medida (Pierce *et alii*, 1999).

Após a obtenção da proteína pura, foram realizados testes de titulação isotérmica por calorimetria com cada proteína com os ligantes histamina, serotonina, adrenalina, noradrenalina, ADP e AMP. Nas condições em que os experimentos foram realizados não foi possível detectar a ligação de nenhum desses ligantes com as lipocalinas estudadas (Tabela B).

**Tabela B – Titulação Isotérmica por Calorimetria**

| Reagente<br>Proteína | Histamina | Serotonina | Adrenalina | Noradrenalina | ADP | AMP |
|---|---|---|---|---|---|---|
| **P2B02** | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| **P7D08** | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| **P6F06** | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| **P10C04** | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |

n.d. – Não detectado, sem indicativo de ligação da proteína com o ligante.

# Discussão

# Discussão

Muitos artrópodes estudados até o momento demonstraram estar preparados para desarmar as respostas hemostáticas do hospedeiro, ativadas para prevenir a perda sangüínea após uma injúria tissular (Ribeiro & Francischetti, 2003). Os insetos hematófagos possuem ao menos uma molécula anticoagulante, uma anti-agregadora de plaquetas e um vasodilator em sua saliva, mas a complexidade e redundância das atividades anti-hemostáticas das moléculas salivares garantem um eficiente repasto sangüíneo a estes animais (Ribeiro, 1995). A hidrólise do ADP, um potente nucleotídeo indutor de agregação plaquetária (Sarkis *et alii*, 1986), e a neutralização do vasoconstritor tromboxano $A_2$ liberado pelas plaquetas (Ribeiro & Sarkis, 1982), bem como a presença de lipocalinas carreadoras de óxido nítrico (NO) – as nitroforinas – que causam vasodilatação e inibição da agregação plaquetária (Ribeiro *et alii*, 1993; Champagne *et alii*, 1995a) são diferentes vias abordadas pelo coquetel salivar de insetos hematófagos com o intuito de impedir a hemostase do hospedeiro.

Há dez anos muitas proteínas salivares descritas eram a tradução de ESTs (*expressed sequence tags*) ou de poucos genes completos. Em uma tentativa de aumentar o entendimento da complexidade de proteínas e transcritos expressos nas glândulas salivares de artrópodes hematófagos, Francischetti e colaboradores (2002b) aliaram a construção de uma biblioteca de cDNA com a identificação de proteínas por seqüenciamento N-terminal após separação por eletroforese. Estava definido o sialoma – um conjunto de RNA mensageiros e proteínas expressas nas glândulas salivares (Valenzuela *et alii*, 2002b). Os estudos de sialoma demonstram que a saliva de insetos hematófagos e carrapatos é mais complexa do que o esperado e contém muitas proteínas desconhecidas.

Algumas famílias de proteínas se destacam nos sialomas estudados por sua presença quase que constante entre as espécies investigadas ou pela abundância de transcritos encontrados que codificam para proteínas salivares de uma mesma família. Algumas destas famílias são as lipocalinas, antígeno 5, apirases e defensinas, entre outras. Há ainda *contigs* com grupos de proteínas encontrados somente na espécie *T. infestans* como a trialisina, triatox e proteínas com domínio de hemolisina.

As lipocalinas, proteínas amplamente encontradas em triatomíneos e carrapatos, desempenham funções similares tais como ligação à histamina e serotonina (Paesen *et alii*, 2000; Sangamnatdej *et alii*, 2002). A ampla distribuição de lipocalinas em insetos hematófagos pode ser exemplificada pela presença de proteínas similares à triabina (Noeske-Jungblut *et alii*, 1995), palidipina (Noeske-Jungblut *et alii*, 1994) e procalina (Paddock *et alii*, 2001) nas glândulas salivares de *T. infestans*, *Rhodnius prolixus* (Ribeiro *et alii*, 2004a) e *T. brasiliensis* (Santos *et alii*, 2007). No entanto, lipocalinas ainda não foram descritas em outros insetos hematófagos como os dípteros. Isso poderia ser devido a diferenças na evolução da hematofagia, onde moléculas diferentes exercem função fisiológica idêntica, anti-hemostática. Por exemplo, um grupo de proteínas exclusivo de dípteros, as proteínas D7, são amplamente expressas em suas glândulas salivares e pertencem à superfamília de proteínas que se ligam a odorantes (OBPs – *odorant-binding proteins*) (Hekmat-Scafe *et alii*, 2000). As proteínas D7 poderiam agir como "seqüestradoras" de aminas biogênicas ou agonistas hemostáticos, assim como as lipocalinas salivares de carrapatos e triatomíneos (Calvo *et alii*, 2006). Embora as proteínas D7 não tenham sido encontradas em triatomíneos, o sialoma de *T. infestans* revelou a presença de proteínas similares a OBPs pertencentes à mesma superfamília. Moléculas como odorantes hidrofóbicos e feromônios devem ser reconhecidos por uma classe especializada de proteínas que facilitam sua entrega aos receptores olfativos (Forêt & Maleszka, 2006). Tanto em insetos como em vertebrados, essa função é desempenhada por proteínas ligantes de odorantes (OBPs) (Pelosi, 1996; Krieger & Breer, 1999; Deyu & Leal, 2002). Além de contribuir para o reconhecimento dos odorantes em insetos, as OBPs podem funcionar como carreadores em outros processos fisiológicos, pois têm sido encontradas em tecidos não-olfativos, sugerindo que sua função possa estar relacionada à capacidade de carreador geral com ampla especificidade por compostos lipofílicos (Forêt & Maleszka, 2006).

As apirases, também consideradas nucleotidases, são proteínas responsáveis pela hidrólise de ATP e ADP em AMP. Como o ADP é um importante mediador da agregação plaquetária, sua diminuição no meio causa inibição da agregação, facilitando a ingestão a sangue pelo inseto. Cinco proteínas associadas à atividade apirásica foram descritas na saliva de *T. infestans* anteriormente (Faudry *et alii*, 2004), mas somente uma dessas proteínas possui sua seqüência depositada em banco de dados – o gene que codifica para a apirase de 79 kDa. No sialoma do *T. infestans* encontramos somente um transcrito com similaridade à apirase de 79 kDa, mas a seqüência encontra-se truncada, não permitindo a execução de bioensaios para sua caracterização. Isso pode ocorrer devido à manipulação do mRNA no início da construção da biblioteca, o que poderia gerar transcritos incompletos que não foram totalmente convertidos em cDNA durante as primeiras etapas da técnica. Entre os sialomas já descritos de triatomíneos e dípteros, praticamente todos eles apresentaram seqüências similares à apirase. Uma exceção é o *Rhodnius prolixus*, embora a atividade apirásica tenha sido descrita em suas glândulas salivares (Ribeiro & Garcia, 1981), seu sialoma não identificou transcritos para a enzima apirase, mas mostrou seqüências de inositol fosfatase com similaridade à apirase (Ribeiro *et alii*, 2004a). É preciso ressaltar que a ausência de detecção de genes que codificam as outras apirases de *T. infestans* evidencia uma deficiência da técnica de construção da biblioteca de cDNA. O que também nos leva a pensar que há possibilidade de que a saliva de *R. prolixus* tenha uma apirase. Os transcritos desses possíveis genes podem ter existência apenas em determinado momento da fisiologia do inseto. Neste caso, bibliotecas de cDNA deveriam ser construídas em momentos diversos da vida e tempo pós-repasto do inseto. Esse raciocínio aplica-se também a outros genes cujas funções ainda são desconhecidas.

Membros de um grupo de proteínas denominado Antígeno 5 estão presentes nas glândulas salivares de vários artrópodes e em praticamente todos os sialomas de insetos hematófagos descritos até agora (Francischetti *et alii*, 2002b; Li *et alii*, 2001; Valenzuela *et alii*, 2002b; Valenzuela *et alii*, 2003; Arcà *et alii*, 2005; Calvo *et alii*, 2007a; Arcà *et alii*, 2007; Santos *et alii*, 2007). Essa família de proteínas secretadas pertence à família CAP (proteínas ricas em cisteínas, proteínas antígeno 5 de insetos, proteína 1 relacionada à patogenicidade em plantas) (Megraw *et alii*, 1998). Além desses insetos, proteínas antígeno 5 também foram descritas em vespas (Himenopteras) (Hoffman, 1993). Apesar deste grupo de proteínas ser amplamente encontrado nas glândulas salivares desses insetos, sua função permanece desconhecida.

A biblioteca de *T. infestans* apresentou seqüências similares às defensinas, peptídios relacionados à imunidade encontrados com freqüência em sialomas de insetos hematófagos como *Ae. aegypti*, *An. darlingi* e *An. funestus* (Ribeiro *et alii*, 2007; Calvo *et alii*, 2004; Calvo *et alii*, 2007a). Defensinas também estão presentes nas glândulas salivares de carrapatos, órgão não só importante na alimentação mas também na transmissão de patógenos (Valenzuela *et alii*, 2002a). Em triatomíneos, uma defensina de *Rhodnius prolixus* foi isolada, purificada e seqüenciada, mas estava presente na hemolinfa e não na saliva (Lopez *et alii*, 2003). A presença de defensinas nas glândulas salivares pode estar relacionada à imunidade inata, seriam peptídios com atividade tóxica sobre patógenos que utilizariam as partes bucais do inseto como porta de entrada.

Dentre as proteínas encontradas somente no sialoma do *T. infestans*, algumas já tiveram sua seqüência depositada no *genbank* do NCBI como a triatox (|gi71725070|) e a trialisina (|gi18920644|). A trialisina foi descrita por Amino e colaboradores (2002) como uma proteína formadora de poros, com propriedades citolíticas. Outro grupo de proteínas descrito na biblioteca de *T. infestans* – as que contêm domínio de hemolisina – não apresentou homologia com transcritos encontrados em sialomas de outros artrópodes. Apesar de conter domínio de hemolisina, é necessário realizar ensaios biológicos para testar se esse grupo de proteínas teria mesmo capacidade de lisar células. Um resultado positivo indicaria funcionalidade ligada à nutrição, defesa e mesmo atividade anti-hemostática, dependendo do tipo de célula susceptível.

Algumas proteínas descritas nas glândulas salivares do *T. infestans* em trabalhos anteriores não foram identificadas no seu sialoma como a sialidase (Amino *et alii*, 1998) e proteínas envolvidas no sistema cinina-calicreína (Isawa *et alii*, 2007). Ou ainda, foram encontradas seqüências, mas estas estavam incompletas como no caso da apirase (Faudry *et alii*, 2004). Com a obtenção do sialoma, esperava-se encontrar essas seqüências de apirases e também uma seqüência que codificasse para uma fosfolipase $A_2$, pois essa atividade também foi identificada nas glândulas salivares do *T. infestans* (Assumpção, manuscrito em preparação). Mesmo realizando o seqüenciamento em massa da biblioteca de cDNA, com 1534 clones obtidos; o fato de não encontrarmos essas seqüências seria devido à pouca quantidade de transcritos desses genes, o que tornaria mais difícil selecioná-los para o seqüenciamento. Ainda, o momento de coleta das glândulas pode não ter sido o mais propício, pois cada gene tem uma freqüência de expressão ou responde de maneira diferente após alimentação sangüínea que pode funcionar como indução ou inibição para expressão de genes salivares. Com o gel bidimensional das proteínas da saliva ocorre algo semelhante, pois nem todas as proteínas expressas pelas glândulas salivares se encontravam presentes no momento da coleta da saliva para o estudo em questão. Todos esses fatores influenciam a composição do sialoma descrito. Mesmo assim, o sialoma nos fornece muitos dados sobre as substâncias farmacologicamente ativas encontradas nas glândulas salivares e nos permite compreender melhor os mecanismos que levam à adaptação deste hábito alimentar.

Esses dados sobre a variedade de moléculas anti-hemostáticas fornecem subsídios para uma compreensão mais ampla da função da saliva nos artrópodes hematófagos, podendo contribuir para novas estratégias de combate às doenças que transmitem, interferindo em sua habilidade como vetores de doenças. Além disso, o conhecimento sobre as propriedades farmacológicas da saliva proporciona a descoberta de substâncias para o desenvolvimento biotecnólogico de fármacos atuantes como anticoagulantes ou anti-agregadores de plaquetas.

A análise do transcriptoma é descritiva e revela a diversidade da composição salivar de vários insetos hematófagos, mas a maioria das proteínas não possui função conhecida (Ribeiro & Francischetti, 2003). O uso de ferramentas computacionais para a análise dos transcritos fornece grande quantidade de dados a partir dos quais podemos fazer mais hipóteses sobre as glândulas salivares. A busca por função dentre as proteínas salivares leva a uma pesquisa pós-sialoma. A expressão de proteínas em sistema heterólogo e a posterior realização de testes funcionais permitirão caracterizar melhor essas proteínas em relação à sua função nas glândulas salivares dos insetos.

# Conclusões

# Conclusões

O conjunto de experimentos realizados neste trabalho permitem-nos concluir que:

- A saliva do *T. infestans* cliva o substrato fluorogênico NBDC$_6$-HPC de forma independente de Ca$^{+2}$, indicando presença de enzima com atividade de PAF-AH. O protocolo de purificação foi adequado para obtenção da enzima pura, ainda que em pequena quantidade;

- A proteina PAF-AH purificada tem massa molecular de 17 kDa e é imunogênica pois foi capaz de induzir a produção de anticorpos monoespecíficos em camundongos e coelhos;

- Análise por espectrometria de massa resultou na identificação de massas de peptídios presentes em uma fosfolipase A$_2$ de veneno da serpente *C. rhodostoma*, confirmando que a enzima estudada pertence à família de FLA$_2$;

- O sialotranscriptoma de *T. infestans* demonstra que esse inseto possui um grande repertório de moléculas anti-hemostáticas em sua saliva, sendo as lipocalinas as proteínas mais abundantes. Além das lipocalinas, as seguintes proteínas também foram encontradas: trialisina, inositol fosfatase, defensina, serino-protease, proteínas contendo domínio de hemolisina ou kazal, proteínas da família antígeno 5 e outras que se ligam a odorantes/feromônios, entre outras;

- Observa-se um alto grau de redundância no sialotranscriptoma de *T. infestans*, evidenciado pela presença de transcritos relacionados. Muitas famílias de proteínas foram encontradas previamente nas glândulas salivares de outros triatomíneos, confirmando a natureza ubíqua da composição salivar destes artrópodes hematófagos;

- Análise por espectrometria de massa de proteínas da saliva separadas por gel bidimensional resultou na identificação de massas de peptídios referentes a muitas das proteínas codificadas pelos transcritos obtidos pela biblioteca de cDNA das glândulas salivares, validando o transcriptoma;

- As funções de muitas seqüências descritas neste trabalho são desconhecidas, enfatizando o quanto ainda pode-se aprender sobre moléculas bioativas da saliva de artrópodes hematófagos. A seleção de transcritos para expressão e análise funcional identificaria novas proteínas como agentes antimicrobianos ou anti-hemostáticos, gerando dados que auxiliariam na compreensão de como os triatomíneos se alimentam de maneira bem sucedida. O sucesso na alimentação de sangue resulta também na transmissão de doenças.

- O acúmulo de informações sobre os sialotranscriptomas já descritos permite estabelecer um padrão de proteínas presente na composição salivar desses insetos. Fornece ainda ferramentas para entender a biologia vascular e o sistema imune;

- Espera-se que este trabalho possa contribuir para a melhor compreensão da evolução da hematofagia em artrópodes e na descoberta de novos compostos farmacologicamente ativos.

# Perspectivas

# Perspectivas

As perspectivas deste trabalho consistem em:

- Seleção de genes, especialmente lipocalinas, ainda com função desconhecida para a caracterização molecular e funcional nas glândulas salivares de *T. infestans.*

- Expressão dos genes selecionados em sistema heterólogo para obtenção da forma recombinante em células de insetos.

- Realização de testes funcionais com o intuito de identificar ligantes para as proteínas estudadas, possivelmente alvos enzimáticos ou moléculas transportadas por proteínas carreadoras.

- Elucidação da estrutura tridimensional das proteínas recombinantes, visando a obtenção de informações sobre sítios prováveis para ligação com ligantes ou substratos.

# Referências Bibliográficas

# Referências Bibliográficas

Amino, R., Martins, R.M., Procopio, J., Hirata, I.Y., Juliano, M.A., Schenkman, S. 2002. Trialysin, a Novel Pore-forming Protein from Saliva of Hematophagous Insects Activated by Limited Proteolysis. **J. Biol. Chem.** 277, 6207-6213.

Amino, R., Porto, R.M., Chammas, R., Egami, M.I., Schenkman, S. 1998. Identification and Characterization of a Sialidase Released by the Salivary Gland of the Hematophagous Insect *Triatoma infestans*. **J. Biol. Chem.** 273, 24575-24582.

Amino, R., Tanaka, A.S., Schenkman, S. 2001. Triapsin, an unusual activatable serine protease from the saliva of the hematophagous vector of Chagas' disease *Triatoma infestans* (Hemiptera: Reduviidae). **Insect Biochem. Mol. Biol.** 31, 465-472.

Arcà, B., Lombardo, F., de Lara Capurro, M., della Torre, A., Dimopoulos, G., James A.A., Coluzzi, M. 1999. Trapping cDNAs encoding secreted proteins from the salivary glands of the malaria vector *Anopheles gambiae*. **Proc. Natl. Acad. Sci. USA** 96, 1516-1521.

Arcà, B., Lombardo, F., Francischetti, I.M.B., Pham, V.M., Mestres-Simon, M., Andersen, J.F., Ribeiro, J.M.C. 2007. An insight into the sialome of the adult female mosquito *Aedes albopictus*. **Insect Biochem. Mol. Biol.** 37, 107-127.

Arcà, B., Lombardo, F., Valenzuela, J.G., Francischetti, I.M., Marinotti, O., Coluzzi, M., Ribeiro, J.M.C. 2005. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. **J. Exp. Biol.** 208, 3971-3986.

Ashida, M., Brey, P.T. 1998. Recent advances on the research of the insect phophenoloxidase cascade. **Molecular Mechanisms of Immune Responses in Insects.** London: Chapman & Hall, 135-172.

Banco Mundial. 1993. Capítulo I. La Salud en los países en desarrollo: éxitos e retos a enfrentar. Informe sobre el Desarrollo Mundial 1993 – Invertir en Salud. **Washington DC: Banco Mundial**, 18-37.

Barth, R., 1954. Estudos anatômicos e histológicos sobre a subfamília Triatominae (Heteroptera-Reduviidae). IV Parte: O complexo das glândulas salivares de *Triatoma infestans*. **Mem. Inst. Oswaldo Cruz** 53, 517-585.

Basanova, A.V., Baskova, I.P., Zavalova, L.L. 2002. Vascular-Platelet and Plasma Hemostasis Regulators from Bloodsucking Animals. **Biochemistry (Moscow)** 69, 143-150.

Bevers, E.M., Comfrius, P., Zwaal, R.F.A. 1993. Mechanisms involved in platelet procoagulant response. Mechanisms of platelet activation and control. **Adv. Exp. Med. Biol.** 344, 195-208.

Boman, H.G. 1995. Peptides antibiotics and their role in innate immunity. **Annu. Rev. Immunol.** 13, 61-92.

Bowman, A.S., Surdick, M.R., Zhu, K., Essenberg, R.C., Sauer, J.R., Dillwith, J.W. 1997. A novel phospholipase $A_2$ activity in saliva of the lone star tick, *Amblyomma americanum* (L.). **Exp. Parasitol.** 87, 121-132.

Brener, Z., Andrade, Z.A., Barral-Netto, M. 2000. ***Trypanosoma cruzi* e Doença de Chagas**. Ed. Guanabara Koogan, 2ª ed., Rio de Janeiro.

Brewer, J.M. 1999. The use of differential scanning calorimetry (DSC) to determine the correctness of folding of cloned proteins. **Biotechnol. Appl. Biochem.** 30, 173-175.

Bulet, P., Hetru, C., Dimarcq, J.-L., Hoffmann, D. 1999. Antimicrobial peptides in insects; structure and function. **Dev. Comp. Immunol.** 23, 329-344.

Calvo, E., Andersen, J., Francischetti, I.M., Capurro, M.L., deBianchi, A.G., James, A.A., Ribeiro, J.M.C., Marinotti, O. 2004. The transcriptome of adult female *Anopheles darlingi* salivary glands. **Insect Mol. Biol.** 13, 73-88.

Calvo, E., Dao, A., Pham, V.M., Ribeiro, J.M.C. 2007a. An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. **Insect Biochem. Mol. Biol.** 37, 164-175.

Calvo, E., Mans, B.J., Andersen, J.F., Ribeiro, J.M.C. 2006. Function and Evolution of a Mosquito Salivary Protein Family. **J. Biol. Chem.** 281, 1935-1942.

Calvo, E., Tokumasu, F., Marinotti, O., Villeval, J.-L., Ribeiro, J.M.C., Francischetti, I.M.B. 2007b. Aegyptin, a Novel Mosquito Salivary Gland Protein Specifically

Binds to Collagen and Prevents its Interaction with Glycoprotein VI, Integrin α2β1 and von Willebrand Factor. **J. Biol. Chem.** 282, 26928-26938.

Cappello, M., Vlasuk, G.P., Bergum, P.W., Huang, S., Hotez, P. 1995. *Ancylostoma caninum* Anticoagulant Peptide: A Hookworm-Derived Inhibitor of Human Coagulation Factor Xa. **Proc. Natl. Acad. Sci. USA** 92, 6152-6156.

Cerenius, L., Söderhäll, K. 2004. The prophenoloxidase-activating system in invertebrates. **Immunol. Rev.** 198, 116-126.

Champagne, D., Nussenzveig, R.H., Ribeiro, J.M.C. 1995a. Purification, characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood sucking insect *Rhodnius prolixus*. **J. Biol. Chem.** 270, 8691-8695.

Champagne, D.E., Ribeiro, J.M.C. 1994. Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito *Aedes aegypti*. **Proc. Natl. Acad. Sci. USA**, 91, 138-142.

Champagne, D.E., Smartt, C.T., Ribeiro, J.M.C., James, A.A. 1995b. The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a member of the 5' nucleotidase family. **Proc. Natl. Acad. Sci. USA**, 92, 694-698.

Charneau, S., Junqueira, M., Costa, C.M., Pires, D.L., Fernandes, E.S., Bussacos, A.C., Sousa, M.V., Ricart, C.A.O., Shevchenko, A., Teixeira, A.R.L. 2007. The saliva proteome of the blood-feeding insect *Triatoma infestans* is rich in platelet-aggregation inhibitors. **Int. J. Mass Spectrom.** 268, 265-276.

Christophides, G.K., Vlachou, D., Kafatos, F.C. 2004. Comparative and functional genomics of the innate immune system in malaria vector *Anopheles gambiae*. **Immunol. Rev.** 198, 127-148.

Corrêa, P.S. 2002. Efeito antimicrobiano e citolítico da trialisina recombinante da saliva de *Triatoma infestans*. **Tese de Doutorado**. Faculdade de Medicina, Universidade de Brasília.

Davie, E.W. 2003. A brief historical review of the waterfall/cascade of blood coagulation. **J. Biol. Chem.** 278, 50819-50832.

Davie, E.W., Fujikawa, K., Kisiel, W. 1991. The coagulation cascade: initiation, maintenance, and regulation. **Biochemistry** 30, 10363-10370.

Derewenda, Z.S. & Ho, Y.S. 1999. PAF-acetylhydrolase (Review). **Biochim. Biophys. Acta** 1441, 229-236.

Deyu, Z., Leal, W. 2002. Conformational isomers of insect odorant-binding proteins. **Arch. Biochem. Biophys.** 397, 99-105.

Dias, J.C.P., Silveira, A.C., Schofield, C.J. 2002. The impact of Chagas disease control in Latin America – A Review. **Mem. Inst. Oswaldo Cruz**, Rio de Janeiro, 97(5), 603-612.

Edwards, L., Lynch, P.J. 1984. Anaphylactic reaction to kissing bug bites. **Ariz. Med.** 41(3), 159-161.

Farndale, R.W., Sixma, J.J., Barnes, M.J., Groot, P.G. 2004. The role of collagen in thrombosis and hemostasis. **J. Thromb. Haem.** 2, 561-573.

Faudry, E., Lozzi, S.P., Santana, J.M., D'Souza-Ault, M., Kieffer, S., Felix, C.R., Ricart, C.A., Sousa, M.V., Vernet, T., Teixeira, A.R. 2004. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. **J. Biol. Chem.** 279, 19607-19613.

Flower, D.R. 1995. Multiple molecular recognition properties of the lipocalin protein family. **J. Mol. Recogn.** 8, 185-195.

Flower, D.R., North, A.C., Sansom, C.E. 2000. The lipocalin protein family: structural and sequence overview. **Biochim. Biophys. Acta** 1482, 9-24.

Forêt, S., Maleszka, R. 2006. Function and evolution of a gene family encoding odorant binding-like protein in a social insect, the honey bee (*Apis mellifera*). **Genome Res.** 16, 1404-1413.

Francischetti, I.M., Andersen J.F., Ribeiro, J.M. 2002a. Biochemical and functional characterization of recombinant *Rhodnius prolixus* platelet aggregation inhibitor 1 as a novel lipocalin with high affinity for adenosine diphosphate and other adenine nucleotides. **Biochemistry** 41, 3810-3818.

Francischetti, I.M., Ribeiro, J.M., Champagne, D., Andersen, J. 2000. Purification, cloning, expression, and mechanism of action of a novel platelet aggregation inhibitor from the salivary gland of the blood-sucking bug, *Rhodnius prolixus*. **J. Biol. Chem.** 275, 12639-12650.

Francischetti, I.M., Valenzuela, J.G., Pham, V.M., Garfield, M.K., Ribeiro, J.M. 2002b. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. **J. Exp. Biol.** 205, 2429-2451.

Gardell, S.J., Ramjit, D.R., Stabilito, I.I., Fujita, T., Lynch, J.J., Cuca, G.C., Jain, D., Wang, S.P., Tung, J.S., Mark, G.E. 1991. Effective thrombolysis without marked plasminemia after bolus intravenous administration of vampire bat salivary plasminogen activator in rabbits. **Circulation** 84, 244-253.

Golodne, D.M., Monteiro, R.Q., Graça-Souza, A.V., Silva-Neto, M.A.C., Atella, G.C. 2003. Lysophosphatidylcholine Acts as an Anti-hemostatic Molecule in the Saliva of the Blood-sucking Bug *Rhodnius prolixus*. **J. Biol. Chem.** 278, 27766-27771.

Goodman, L.S. & Gilman, A. 1996. **As Bases Farmacológicas da Terapêutica**. Ed. McGraw-Hill, 9ª ed., Rio de Janeiro.

Hekmat-Scafe, D.S., Dorit, R.L., Carlson, J.R. 2000. Molecular Evolution of Odorant-Binding Protein Genes OS-E and OS-F in *Drosophila*. **Genetics** 155, 117-127.

Hoffman, D.R. 1993. Allergens in Hymenoptera venom XXV: The amino acid sequences of antigen 5 molecules and the structural basis of antigenic cross-reactivity. **J. Allergy Clin. Immunol.** 92, 707-716.

Hoffmann, J.A., Kafatos, F.C., Janeway, C.A., Ezekowitz, R.A.B. 1999. Phylogenetic perspectives in innate immunity. **Science** 284, 1313-1318.

Isawa, H., Orito, Y., Jingushi, N., Iwanaga, S., Morita, A., Chinzei, Y., Yuda, M. 2007. Identification and characterization of plasma kallikrein-kinin system inhibitors from salivary glands of the blood-sucking insect *Triatoma infestans*. **FEBS J.** 274, 4271-4286.

Jackson, C.M., Nemerson, Y. 1980. Blood coagulation. **Ann. Rev. Biochem.** 49, 765-811.

Jamaluddin, M. 1991. New perspectives in blood platelet aggregation. **Current Science** 61, 526-533.

Krieger, J., Breer, H. 1999. Olfactory reception in invertebrates. **Science** 286, 720-723.

Kuijpers, T.W., Van den Berg, J.M., Tool, T.J., Roos, D. 2001. The impact of platelet-activating factor (PAF)-like mediators on the functional activity of neutrophils:

anti-inflamamatory effects of human PAF-acetylhydrolase. **Clin. Exp. Immunol.** 123, 412-420.

Lacombe, D. 1999. Anatomia e Histologia das Glândulas Salivares nos Triatomíneos. **Mem. Inst. Oswaldo Cruz**, Rio de Janeiro, 94(4), 557-564.

Li, S., Kwon, J., Aksoy, S. 2001. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. **Insect Mol. Biol.** 10(1), 69-76.

Lopez, L., Morales, G., Ursic, R., Wolff, M., Lowenberger, C. 2003. Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of Chagas disease. **Insect Biochem. Mol. Biol.** 33, 439-447.

McManus, L.M., Pinckard, R.N., Fitzpatrick, F.A., O'Rourke, R.A., Crawford, M.H., Hanahan, D.J. 1981. Acetyl glyceryl ether phosphorylcholine: intravascular alterations following intravenous infusion into the baboon. **Lab. Invest.** 45, 303-307.

Megraw, T., Kaufman, T.C., Kovalick, G.E. 1998. Sequence and expression of *Drosophila* Antigen 5-related 2, a new member of the CAP gene family. **Gene** 222(2), 297-304.

Miles, M.A., Feliciangeli, M.D., de Arias, A.R. 2003. American trypanosomiases (Chagas' disease) and the role of molecular epidemiology in guiding control strategies. **BMJ** 326, 1444-1448.

Montfort, W.R., Weichsel, A., Andersen, J.F. 2000. Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. **Biochim. Biophys. Acta** 1482, 110-188.

Morita, A., Isawa, H., Orito, Y., Iwanaga, S., Chinzei, Y., Yuda, M. 2006. Identification and characterization of a collagen-induced platelet aggregation inhibitor, triplatin, from salivary glands of the assassin bug, *Triatoma infestans*. **FEBS J.** 273, 2955-2962.

Morris, R. V., Shoemaker, C.B., David, J.R., Lanzaro, G.C., Titus, R.G. 2001. Sandfly maxadilan exacerbates infection with *Leishmania major* and vaccinating against it protects against *L. major* infection. **J. Immunol.** 167, 5226-5230.

Moser, M., Auerswald, E., Mentele, R., Eckerskorn, C., Fritz, H., Fink, E. 1998. Bdellastasin, a serine protease inhibitor of the antistasin family from the medical leech (*Hirudo medicinalis*) – primary structure, expression in yeast, and characterization of native and recombinant inhibitor. **Eur. J. Biochem.** 253, 212-220.

Murakami, M., Kudo, I. 2002. Phospholipases $A_2$. **J. Biochemistry** 131, 285-292.

Neves, D. 2005. **Parasitologia Humana.** Ed. Atheneu, 11ª ed., São Paulo.

Nieswandt, B., Watson, S.P. 2003. Platelet-collagen interactions: is GPVI the central receptor? **Blood** 102, 449-461.

Noeske-Jungblut, C., Haendler, B., Donner, P., Alagon, A., Possani, L., Scheleuning, W.-D. 1995. Triabin, a highly potent exosite inhibitor of thrombin. **J. Biol. Chem.** 270, 28629-28634.

Noeske-Jungblut, C., Krätzschmar, J., Haendler, B., Alagon, A., Possani, L., Verhallen, P., Donner, P., Scheleuning, W.-D. 1994. An inhibitor of collagen-induced platelet aggregation from the saliva of *Triatoma pallidipennis*. **J. Biol. Chem.** 269, 5050-5053.

Paddock, C.D., McKerrow, J.H., Hansell, E., Foreman, K.W., Hsieh, I., Marshall, N. 2001. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. **J. Immunol.** 167, 2694-2699.

Paesen, G.C., Adams, P.L., Nuttall, P.A., Stuart, D.L. 2000. Tick histamine-binding proteins: lipocalins with a second binding cavity. **Biochim. Biophys. Acta** 1482, 92-101.

Pelosi, P., 1996. Perireceptor events in olfaction. **J. Neurobiol.** 30, 3-19.

Pereira, M.H., Souza, M.E.L., Vargas, A.P., Martins, M.S., Penido, C.M., Diotaiuti, L. 1996. Anticoagulant activity of *Triatoma infestans* and *Panstrongylus megistus* saliva (Hemiptera/Triatominae). **Acta Tropica** 61, 225-261.

Pierce, M.M., Raman, C.S., Nall, B.T. 1999. Isothermal Titration Calorimetry of Protein-Protein Interactions. **Methods** 19, 213-221.

Prescott, S.M., Zimmerman, G.A., Stafforini, D.M., McIntyre, T.M. 2000. Platelet-activating factor and related lipid mediators. **Annu. Rev. Biochem.** 69, 419-445.

Rang, H.P., Dale, M.M., Ritter, J.M. 1997. **Farmacologia**, Ed. Guanabara Koogan, 3ª ed., Rio de Janeiro.

Ratcliffe, N.A., Leonard, C., Rowley, A.F. 1984. Prophenoloxidase Activation: Nonself Recognition and Cell Cooperation in Insect Immunity. **Science** 226, 557-559.

Ribeiro, J.M., Francischetti, I.M. 2003. Role of arthropod saliva in blood feeding: Sialome and post-sialome perspectives. **Annu. Rev. Entomol.** 48, 73-88.

Ribeiro, J.M., Garcia, E.S. 1981. Platelet antiaggregating activity in the salivary secretion of the blood sucking bug *Rhodnius prolixus*. **Experientia** 37, 384-386.

Ribeiro, J.M., Walker, F.A. 1994. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorin) of *Rhodnius prolixus*. **J. Exp. Med.** 180, 2251-2257.

Ribeiro, J.M.C. & Francischetti, I.M.B. 2001. Platelet-activating-factor-hydrolyzing phospholipase C in the salivary glands and saliva of the mosquito *Culex quinquefasciatus*. **J. Exp. Biol.** 204, 3887-3894.

Ribeiro, J.M.C. 1987. Role of saliva in blood-feeding by arthropods. **Annu. Rev. Entomol.** 32, 463-78.

Ribeiro, J.M.C. 1989. Role of saliva in tick/host interactions. **Exp. Applied Acarol.** 7, 15-20.

Ribeiro, J.M.C. 1995. Blood-feeding arthropods: live syringes or invertebrate pharmacologist? **Infect Agents Dis.** 4, 143-152.

Ribeiro, J.M.C., Andersen, J., Silva-Neto, M.A.C., Pham, V.M., Garfield, M.K., Valenzuela, J.G. 2004a. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. **Insect Biochem. Mol. Biol.** 34, 61-79.

Ribeiro, J.M.C., Arcà, B., Lombardo, F., Calvo, E., Pham, V.M., Chandra, P.K., Wikel, S.K. 2007. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. **BMC Genomics** 8, 6.

Ribeiro, J.M.C., Charlab, R., Pham, V.M., Garfield, M., Valenzuela, J.G. 2004b. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. **Insect Biochem. Mol. Biol.** 34, 543-563.

Ribeiro, J.M.C., Hazzard, J.M.H., Nussenzveig, R.H., Champagne, D., Walker, F.A. 1993. Reversible binding of nitric oxide by a salivary nitrosylhemeprotein from the blood sucking bug, *Rhodnius prolixus*. **Science** 260, 539-541.

Ribeiro, J.M.C., Sarkis, J.J.F. 1982. Anti-thromboxane activity in *Rhodnius prolixus* salivary secretion. **J. Insect Physiol.** 28, 655-660.

Sangamnatdej, S., Paesen, G.C., Slovak, M., Nuttall, P.A. 2002. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. **Insect Mol. Biol.** 11, 79-86.

Santos, A., Ribeiro, J.M., Lehane, M.J., Gontijo, N.F., Veloso, A.B., Sant'anna, M.R., Nascimento Araújo, R., Grisard, E.C., Pereira, M.H. 2007. The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). **Insect Biochem. Mol. Biol.** 37(7), 702-12.

Sarkis, J.J.F., Guimaraes, J.A., Ribeiro, J.M.C. 1986. Salivary apyrase of *Rhodnius prolixus*: kinetics and purification. **Biochem. J.** 233, 885-891.

Sawyer, R.T. 1986. **Leech Biology and Behavior**, Vol. I, Oxford Science Publications, Clarendon, Oxford.

Sawyer, R.T. 1991. Thrombolytics and anti-coagulants from leeches. **Biotechnology** 9, 513-518.

Schofield, C.J., Jannin, J., Salvatella, R. 2006. The future of Chagas disease control. **Trends Parasitol.** 22(12), 583-588.

Six, D.A. & Dennis, E.A. 2000. The expanding superfamily of phospholipase $A_2$ enzymes: classification and characterization. **Biochim. Biophys. Acta**, 1448, 1-19.

Söderhäll, K., Cerenius, L. 1998. Role of the prophenoloxidase-activating system in invertebrate immunity. **Curr. Opin. Immunol.** 10, 23-28.

Stafforini, D.M., McIntyre, T.M., Zimmerman, G.A., Prescott, S.M. 1997. Platelet-activating factor acetylhydrolases. **J. Biol. Chem.** 272(29), 17895-17898.

Titus, R.G. & Ribeiro, J.M.C. 1988. Salivary gland lysates from the sand fly *Lutzomyia longipalpis* enhance *Leishmania* infectivity. **Science** 239, 1306-1308.

Valenzuela, J.G. 2002c. High-throughput approaches to study salivary proteins and genes from vectors of disease. **Insect Biochem. Mol. Biol.** 32, 1199-1209.

Valenzuela, J.G., Charlab, R., Galperin, M.Y., Ribeiro, J.M.C. 1998. Purification, cloning, and expression of an apyrase from the bed bug *Cimex lectularius*. A new type of nucleotide-binding enzyme. **J. Biol. Chem.** 273, 30583-30590.

Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Ribeiro, J.M. 2003. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. **Insect Biochem. Mol. Biol.** 33, 717-732.

Valenzuela, J.G., Francischetti, I.M.B., Pham, V.M., Garfield, M.K., Mather, T.N., Ribeiro, J.M.C. 2002a. Exploring the sialome of the tick, *Ixodes scapularis*. **J. Exp. Biol.** 205, 2843-2864.

Valenzuela, J.G., Francischetti, I.M.B., Ribeiro, J.M.C. 1999. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*. **Biochemistry** 38, 11209-11215.

Valenzuela, J.G., Guimarães, J.A., Ribeiro, J.M.C. 1996. A novel inhibitor of factor X activation from the salivary glands of the bed bug *Cimex lectularius*. **Exp. Biol.** 83, 184-190.

Valenzuela, J.G., Pham, V.M., Garfield, M.K., Francischetti, I.M., Ribeiro, J.M.C. 2002b. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. **Insect Biochem. Mol. Biol.** 32, 1101-1122.

Venable, M.E., Zimmerman, G.A., McIntyre, T.M., Prescott, S.M. 1993. Platelet-activating factor: a phospholipid autacoid with diverse actions. **J. Lipid Res.** 34, 691-702.

World Health Organization. 1991 Control of Chagas' disease. **WHO, Tech. Rep. Ser.**, 811.

# Anexo I – Tabela Suplementar do Manuscrito II

| Protein link | Description | First residue | Seq size | Nucleotide sequence link | Stop codon? | Counter for number of proteins/class | Best match to TI-ASB database | Number of EST sequences on cluster | HPLC/MS/MS results | 2D-SDS-PAGE/MS/MS results | SigP Result | Predicted cleavage site | MW | pI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Secreted salivary proteins** | | | | | | | | | | | | | | |
| **Lipocalins** | | | | | | | | | | | | | | |
| TI-25 | salivary lipocalin | M | 208 | TI-25 | * | 1 | ti-new-contig_51 | 30 | | G31-12 - | SIG | 18-19 | 24.3 | 9.4 |
| TI-32 | salivary lipocalin | M | 175 | TI-32 | * | 2 | ti-new-contig_61 | 26 | | G37-36 - | SIG | 18-19 | 19.2 | 8.9 |
| TI-31 | salivary lipocalin | M | 202 | TI-31 | * | 3 | ti-new-contig_52 | 24 | | | SIG | 22-23 | 22.7 | 9.3 |
| TI-55 | salivary lipocalin | M | 162 | TI-55 | * | 4 | ti-new-contig_54 | 17 | | | SIG | 18-19 | 18.4 | 4.4 |
| TI-19 | salivary lipocalin | M | 179 | TI-19 | * | 5 | ti-new-contig_29 | 15 | | G104-82 | SIG | 18-19 | 20.0 | 8.2 |
| TI-33 | salivary lipocalin | M | 193 | TI-33 | * | 6 | ti-new-contig_22 | 14 | | G37-37 - | SIG | 18-19 | 21.6 | 8.7 |
| TI-21 | salivary lipocalin | M | 179 | TI-21 | * | 7 | ti-new-contig_31 | 13 | | G103-80 | SIG | 18-19 | 20.0 | 8.7 |
| tin_7 | pallidipin-like lipocalin precursor | M | 269 | tin_7 | * | 8 | ti-new-contig_7 | 13 | | G45-49 - | SIG | 19-20 | 30.7 | 4.4 |
| TI-26 | salivary lipocalin | M | 182 | TI-26 | * | 9 | ti-new-contig_15 | 12 | | G37-33 - | SIG | 18-19 | 20.4 | 9.2 |
| TI-20 | salivary lipocalin | M | 179 | TI-20 | * | 10 | ti-new-contig_28 | 12 | | G104-82 | SIG | 18-19 | 20.0 | 8.2 |
| TI-23 | salivary lipocalin | M | 206 | TI-23 | * | 11 | ti-new-contig_49 | 12 | | G31-15 - | SIG | 18-19 | 23.8 | 8.9 |
| TI-54 | salivary lipocalin | M | 199 | TI-54 | * | 12 | ti-new-contig_67 | 12 | | G36-16 - | SIG | 16-17 | 22.3 | 9.2 |
| TI-18 | salivary lipocalin | M | 179 | TI-18 | * | 13 | ti-new-contig_32 | 9 | | G104-82 | SIG | 18-19 | 19.9 | 6.7 |
| TI-35 | salivary lipocalin | M | 194 | TI-35 | * | 14 | ti-new-contig_20 | 7 | | G37-37 - | SIG | 18-19 | 22.0 | 9.0 |
| TI-24 | salivary lipocalin | M | 208 | TI-24 | * | 15 | ti-new-contig_50 | 7 | | G31-12 - | SIG | 18-19 | 24.1 | 9.2 |
| tin-4 | pallidipin-like lipocalin precursor - | T | 228 | tin-4 | * | 16 | ti-new-contig_4 | 6 | | G45-49 - | CYT | | 26.3 | 4.3 |
| TI-69 | salivary lipocalin | M | 202 | TI-69 | * | 17 | ti-new-contig_53 | 6 | | | SIG | 18-19 | 22.4 | 9.3 |
| TI-36 | truncated pallidipin-like lipocalin p | M | 219 | TI-36 | * | 18 | ti-new-contig_6 | 6 | | G45-49 - | CYT | | 25.5 | 4.4 |
| TI-103 | salivary lipocalin | M | 182 | TI-103 | * | 19 | ti-new-contig_79 | 6 | | | SIG | 18-19 | 19.9 | 5.2 |
| TI-29 | salivary lipocalin | M | 182 | TI-29 | * | 20 | ti-new-contig_18 | 5 | | G37-33 - | SIG | 18-19 | 20.5 | 9.0 |
| TI-59 | salivary lipocalin | M | 194 | TI-59 | * | 21 | ti-new-contig_25 | 5 | | | SIG | 18-19 | 22.2 | 9.0 |
| TI-58 | salivary lipocalin | M | 189 | TI-58 | * | 22 | ti-new-contig_27 | 5 | | | SIG | 18-19 | 21.6 | 8.9 |
| TI-60 | salivary lipocalin | M | 197 | TI-60 | * | 23 | ti-new-contig_90 | 5 | | G36-17 - | SIG | 16-17 | 21.9 | 9.2 |
| tin-5 | pallidipin-like lipocalin precursor - | G | 256 | tin-5 | * | 24 | ti-new-contig_5 | 4 | | G45-49 - | CYT | | 29.5 | 4.4 |
| TI-56 | salivary lipocalin similar to triabin | M | 162 | TI-56 | * | 25 | ti-new-contig_55 | 4 | | | SIG | 18-19 | 18.7 | 6.2 |
| TI-72 | triatin-like salivary lipocalin | M | 197 | TI-72 | * | 26 | ti-new-contig_10 | 3 | | | SIG | 16-17 | 22.7 | 5.4 |
| TI-80 | pallidipin-like salivary lipocalin | M | 250 | TI-80 | * | 27 | ti-new-contig_114 | 3 | | | SIG | 18-19 | 28.8 | 5.5 |
| TI-27 | salivary lipocalin | M | 182 | TI-27 | * | 28 | ti-new-contig_16 | 3 | | G37-33 - | SIG | 18-19 | 20.5 | 9.2 |
| TI-30 | salivary lipocalin | M | 206 | TI-30 | * | 29 | ti-new-contig_19 | 3 | | | SIG | 18-19 | 23.2 | 9.2 |
| TI-79 | salivary lipocalin | M | 193 | TI-79 | * | 30 | ti-new-contig_26 | 3 | | | SIG | 18-19 | 22.3 | 9.1 |
| TI-512 | salivary lipocalin | M | 195 | TI-512 | * | 31 | ti-new-contig_82 | 3 | | | SIG | 18-19 | 22.6 | 9.4 |
| TI-73 | salivary lipocalin | M | 197 | TI-73 | * | 32 | ti-new-contig_11 | 2 | | | SIG | 16-17 | 22.8 | 5.1 |
| tin-119 | salivary lipocalin | M | 208 | tin-119 | * | 33 | ti-new-contig_119 | 2 | | | SIG | 18-19 | 24.1 | 9.8 |
| TI-82 | salivary lipocalin | M | 197 | TI-82 | * | 34 | ti-new-contig_12 | 2 | | | SIG | 16-17 | 23.1 | 5.2 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TI-83 | salivary lipocalin | M | 197 | TI-83 | * | 35 | ti-new-contig_13 | 2 | | | SIG | 16-17 | 23.0 | 5.7 |
| TI-28 | salivary lipocalin | M | 182 | TI-28 | * | 36 | ti-new-contig_17 | 2 | | G37-33 - | SIG | 18-19 | 20.5 | 9.1 |
| ti-179 | lipocalin-like TiLipo39 allele | M | 179 | ti-179 | * | 37 | ti-new-contig_179 | 2 | | G37-36 - | SIG | 22-23 | 19.9 | 8.5 |
| tin-180 | triatin-like salivary lipocalin | M | 195 | tin-180 | * | 38 | ti-new-contig_180 | 2 | | | SIG | 16-17 | 22.0 | 8.7 |
| TI-34 | salivary lipocalin | M | 191 | TI-34 | * | 39 | ti-new-contig_21 | 2 | | G37-37 - | SIG | 18-19 | 21.5 | 6.8 |
| TI-100 | salivary lipocalin | M | 189 | TI-100 | * | 40 | ti-new-contig_23 | 2 | | | SIG | 18-19 | 21.6 | 9.2 |
| TI-376 | salivary lipocalin | M | 193 | TI-376 | * | 41 | ti-new-contig_24 | 2 | | | SIG | 18-19 | 22.4 | 8.6 |
| TI-65 | triabin-like salivary lipocalin | M | 162 | TI-65 | * | 42 | ti-new-contig_56 | 2 | | | SIG | 18-19 | 18.1 | 4.4 |
| TI-64 | triabin-like salivary lipocalin | M | 161 | TI-64 | * | 43 | ti-new-contig_57 | 2 | | | SIG | 18-19 | 18.0 | 4.4 |
| tin_65 | lipocalin-like Ti65 | M | 197 | tin_65 | * | 44 | ti-new-contig_65 | 2 | | G36-18 - | SIG | 15-16 | 21.6 | 9.4 |
| tin_66 | lipocalin-like Tin66 | M | 197 | tin_66 | * | 45 | ti-new-contig_66 | 2 | | G36-18 - | SIG | 15-16 | 21.7 | 9.3 |
| TI-216 | salivary lipocalin | M | 181 | TI-216 | * | 46 | ti-new-contig_78 | 2 | | | SIG | 18-19 | 20.0 | 8.2 |
| tin-8 | pallidipin-like lipocalin precursor | M | 241 | tin-8 | * | 47 | ti-new-contig_8 | 2 | | G45-49 - | SIG | 19-20 | 27.8 | 4.6 |
| TI-120 | salivary lipocalin | M | 195 | TI-120 | * | 48 | ti-new-contig_83 | 2 | | G36-16 - | SIG | 18-19 | 22.7 | 9.2 |
| TI-123 | salivary lipocalin | M | 195 | TI-123 | * | 49 | ti-new-contig_84 | 2 | | G36-16 - | SIG | 18-19 | 22.7 | 9.1 |
| tin-120 | salivary lipocalin | M | 208 | tin-120 | * | 50 | ti-new-contig_120 | 1 | | | SIG | 18-19 | 23.9 | 9.7 |
| tin-130 | lipocalin-like TiLipo33 allele | M | 182 | tin-130 | * | 51 | ti-new-contig_130 | 1 | | G37-36 - | SIG | 18-19 | 20.0 | 8.7 |
| tin_36 | Short salivary lipocalin | M | 116 | tin_36 | * | 52 | ti-new-contig_36 | 1 | | G104-82 | SIG | 18-19 | 12.9 | 6.1 |
| tin_557 | triabin-like lipocalin 4a precursor | M | 199 | tin_557 | | 53 | ti-new-contig_557 | 1 | | | SIG | 18-19 | 23.1 | 9.2 |
| TI-66 | salivary lipocalin | M | 162 | TI-66 | * | 54 | ti-new-contig_59 | 1 | | | SIG | 18-19 | 18.1 | 4.4 |
| **Short Trialysin-like family** | | | | | | | | | | | | | | |
| TI-13 | short trialysin 1 | M | 76 | TI-13 | * | 1 | ti-new-contig_43 | 53 | | | SIG | 19-20 | 8.5 | 4.2 |
| TI-14 | short trialysin 2 | M | 76 | TI-14 | * | 2 | ti-new-contig_44 | 4 | H39 -> A | | SIG | 19-20 | 8.6 | 4.7 |
| TI-15 | short trialysin 3 | M | 76 | TI-15 | * | 3 | ti-new-contig_42 | 3 | H35B -> | | SIG | 19-20 | 8.4 | 4.4 |
| TI-16 | short trialysin 4 | M | 76 | TI-16 | * | 4 | ti-new-contig_48 | 1 | H35B -> | | SIG | 19-20 | 8.3 | 4.6 |
| TI-17 | short trialysin 5 | M | 76 | TI-17 | * | 5 | ti-new-contig_45 | 1 | H48A -> | | SIG | 19-20 | 8.5 | 4.6 |
| **Hemolysin-like family** | | | | | | | | | | | | | | |
| TI-51 | hemolysin-like secreted salivary pr | M | 249 | TI-51 | * | 1 | ti-new-contig_62 | 14 | H35 -> M | | SIG | 18-19 | 25.2 | 7.8 |
| TI-52 | hemolysin-like secreted salivary pr | M | 249 | TI-52 | * | 2 | ti-new-contig_63 | 3 | H35 -> M | | SIG | 18-19 | 25.4 | 8.6 |
| TI-95 | hemolysin-like secreted salivary pr | M | 272 | TI-95 | * | 3 | ti-new-contig_64 | 3 | H35 -> M | | SIG | 18-19 | 28.2 | 4.9 |
| tin-105 | salivary secreted protein similar to | M | 242 | tin-105 | * | 4 | ti-new-contig_105 | 3 | | | SIG | 16-17 | 26.1 | 4.9 |
| TI-63 | salivary secreted protein - possibl | M | 198 | TI-63 | * | 5 | ti-new-contig_86 | 6 | H58A -> | | SIG | 20-21 | 19.6 | 8.0 |
| **Trialysin family** | | | | | | | | | | | | | | |
| TI-6 | trialysin precursor allele | M | 260 | TI-6 | * | 1 | ti-new-contig_37 | 15 | | G36-23 - | SIG | 19-20 | 28.6 | 8.5 |
| TI-8 | trialysin allele | M | 260 | TI-8 | * | 2 | ti-new-contig_39 | 44 | | G36-23 - | SIG | 19-20 | 28.6 | 8.5 |
| **Triatox** | | | | | | | | | | | | | | |
| TI-57 | triatox - salivary lipocalin | M | 136 | TI-57 | * | 1 | ti-new-contig_81 | 8 | | G128-84 | SIG | 22-23 | 14.8 | 5.4 |
| **Antigen 5 family** | | | | | | | | | | | | | | |
| TI-49 | antigen-5-like protein precursor | M | 244 | TI-49 | * | 1 | ti-new-contig_71 | 12 | | | SIG | 23-24 | 28.1 | 9.2 |
| **Kazal containing peptide** | | | | | | | | | | | | | | |
| TI-48 | salivary kazal-type proteinase inhi | M | 81 | TI-48 | * | 1 | ti-new-contig_75 | 10 | | | SIG | 29-30 | 8.9 | 8.2 |
| TI-75 | salivary secreted kazal-type protei | M | 80 | TI-75 | * | 2 | ti-new-contig_88 | 6 | | | SIG | 28-29 | 8.9 | 8.2 |
| **Pheromone binding family** | | | | | | | | | | | | | | |
| TI-46 | heme-binding protein | M | 142 | TI-46 | * | 1 | ti-new-contig_76 | 6 | | | SIG | 19-20 | 16.3 | 5.7 |
| TI-47 | heme-binding protein | M | 143 | TI-47 | * | 2 | ti-new-contig_77 | 3 | | | SIG | 19-20 | 16.4 | 6.1 |
| **Inositol phosphatase family** | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ti-7-80-90-2 | salivary inositol polyphosphate 5-p | M | 316 | ti-7-80-90-2 | * | 1 | ti-new-contig_89 | 5 | G27-10 - | SIG | 19-20 | 36.0 | 9.4 |
| tin-272 | salivary inositol polyphosphate 5-p | M | 318 | tin-272 | * | 2 | ti-new-contig_85 | 6 | G27-5 -> | SIG | 21-22 | 36.9 | 9.1 |
| **Serine protease** | | | | | | | | | | | | | |
| TI-117 | salivary trypsin | M | 308 | TI-117 | * | 1 | ti-new-contig_153 | 2 | | SIG | 21-22 | 34.4 | 5.0 |
| **Similar to bacterially induced Hyphantria cunea peptide Hdd1** | | | | | | | | | | | | | |
| TI-41 | salivary secreted protein | M | 129 | TI-41 | * | 1 | ti-new-contig_69 | 17 | G38-45 - | SIG | 20-21 | 14.7 | 9.9 |
| TI-44 | putative salivary secreted protein | M | 136 | TI-44 | * | 2 | ti-new-contig_73 | 7 | | SIG | 28-29 | 15.3 | 9.8 |
| TI-45 | putative salivary secreted peptide | M | 136 | TI-45 | * | 3 | ti-new-contig_74 | 5 | | SIG | 28-29 | 15.3 | 9.8 |
| **Defensin and other immune-related products** | | | | | | | | | | | | | |
| TI-455 | defensin A | M | 93 | TI-455 | * | 1 | ti-new-contig_402 | 1 | | SIG | 19-20 | 10.3 | 6.9 |
| ti-141 | salivary secreted protein with lipop | M | 310 | ti-141 | * | 2 | ti-new-contig_141 | 2 | | SIG | 17-18 | 33.8 | 6.2 |
| ti-7-80-90-1 | Putative secreted peptide with HH | M | 74 | ti-7-80-90-1 | * | 3 | ti-new-contig_100 | 4 | | BL | 28-29 | 8.5 | 9.5 |
| TI-108 | salivary secreted protein similar to | M | 164 | TI-108 | * | 4 | ti-new-contig_149 | 2 | | SIG | 20-21 | 17.7 | 9.3 |
| **Weak similarity to Plasmodium adhesive protein** | | | | | | | | | | | | | |
| ti-7-80-90-2 | putative secreted protein similar to | M | 142 | ti-7-80-90-2 | * | 1 | ti-new-contig_124 | 2 | | SIG | 21-22 | 16.3 | 4.4 |
| **Secreted protein similar to Culicoides salivary peptide** | | | | | | | | | | | | | |
| TI-70 | putative secreted salivary peptide | M | 86 | TI-70 | * | 1 | ti-new-contig_87 | 6 | | SIG | 22-23 | 8.3 | 5.7 |
| **Other putative secreted proteins** | | | | | | | | | | | | | |
| TI-572 | putative salivary secreted peptide | M | 69 | TI-572 | * | 1 | ti-new-contig_498 | 1 | | SIG | 24-25 | 7.4 | 8.0 |
| ti-new-5-36 | putative salivary secreted protein | M | 175 | ti-new-5-36 | * | 2 | ti-new-contig_638 | 1 | | SIG | 26-27 | 19.9 | 5.6 |
| tin-586 | putative salivary secreted protein | M | 145 | tin-586 | * | 3 | ti-new-contig_586 | 1 | | SIG | 17-18 | 16.7 | 9.3 |
| **Putative housekeeping proteins** | | | | | | | | | | | | | |
| **Nuclear regulation** | | | | | | | | | | | | | |
| tin-606 | similar to mob as tumor suppress | M | 217 | tin-606 | * | 1 | ti-new-contig_606 | 1 | | CYT | | 25.0 | 6.5 |
| TI-118 | H3 histone, family 3B | M | 136 | TI-118 | * | 2 | ti-new-contig_154 | 2 | | CYT | | 15.4 | 11.3 |
| tin-471 | histone H1 - truncated | M | 209 | tin-471 | | 3 | ti-new-contig_471 | 1 | | CYT | | 21.8 | 10.6 |
| **Transcription machinery** | | | | | | | | | | | | | |
| TI-310 | DNA-binding nuclear protein p8 | M | 73 | TI-310 | * | 1 | ti-new-contig_286 | 1 | | CYT | | 8.5 | 8.9 |
| tin-167 | putative elongation factor 1 beta | M | 223 | tin-167 | * | 2 | ti-new-contig_167 | 2 | | CYT | | 24.7 | 4.7 |
| tin-175 | similar to DNA-directed RNA poly | M | 117 | tin-175 | * | 3 | ti-new-contig_175 | 2 | | CYT | | 13.5 | 5.9 |
| TI-604 | small ribonuclear protein | M | 80 | TI-604 | * | 4 | ti-new-contig_516 | 1 | | CYT | | 9.2 | 9.2 |
| **Protein synthesis machinery** | | | | | | | | | | | | | |
| TI-85 | ribosomal protein L41 | M | 25 | TI-85 | * | 1 | ti-contig_85 | 1 | | CYT | | 3.4 | 13.0 |
| TI-287 | 40S ribosomal protein S11 | M | 194 | TI-287 | * | 2 | ti-new-contig_34 | 1 | | CYT | | 22.4 | 10.8 |
| TI-68 | ribosomal protein S15e | M | 147 | TI-68 | * | 3 | ti-new-contig_103 | 4 | | CYT | | 17.0 | 10.5 |
| TI-111 | 40S ribosomal protein S15/S22 | M | 130 | TI-111 | * | 4 | ti-new-contig_148 | 2 | | CYT | | 14.8 | 10.1 |
| TI-257 | 40S ribosomal protein S17 | M | 131 | TI-257 | * | 5 | ti-new-contig_132 | 2 | | CYT | | 15.4 | 9.8 |
| TI-88 | ribosomal protein S20 | M | 117 | TI-88 | * | 6 | ti-new-contig_126 | 2 | | CYT | | 13.4 | 9.9 |
| tin-578 | putative ribosomal protein S26 | M | 116 | tin-578 | * | 7 | ti-new-contig_578 | 1 | | CYT | | 13.4 | 11.0 |
| TI-268 | 40s ribosomal protein S27 | M | 84 | TI-268 | * | 8 | ti-new-contig_255 | 1 | | CYT | | 9.5 | 9.5 |
| TI-440 | 40S ribosomal protein S29 | M | 56 | TI-440 | * | 9 | ti-new-contig_388 | 1 | | CYT | | 6.4 | 9.8 |
| TI-461 | 40S ribosomal protein S3A | M | 262 | TI-461 | * | 10 | ti-new-contig_400 | 1 | | CYT | | 29.7 | 9.8 |
| TI-356 | 40S ribosomal protein S8 | M | 208 | TI-356 | * | 11 | ti-new-contig_319 | 1 | | CYT | | 24.0 | 10.6 |
| TI-77 | ribosomal protein P1 | M | 116 | TI-77 | * | 12 | ti-new-contig_113 | 3 | | BL | | 11.8 | 4.4 |
| tin_177 | ribosomal protein P2 | M | 114 | tin_177 | * | 13 | ti-new-contig_177 | 2 | | BL | | 11.7 | 4.8 |
| tin-445 | similar to ribosomal protein L11 | M | 195 | tin-445 | * | 14 | ti-new-contig_445 | 1 | | CYT | | 22.4 | 10.1 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TI-115 | putative ribosomal protein L15/L27 | M | 148 | TI-115 | * | 15 | ti-new-contig_151 | 2 | CYT |  | 16.7 | 10.7 |
| TI-96 | 60S ribosomal protein L18A | M | 177 | TI-96 | * | 16 | ti-new-contig_104 | 4 | CYT |  | 21.0 | 10.5 |
| TI-337 | 60S ribosomal protein L8 | M | 258 | TI-337 | * | 17 | ti-new-contig_306 | 1 | CYT |  | 28.1 | 11.0 |
| TI-626 | 60S ribosomal protein L26 | M | 149 | TI-626 | * | 18 | ti-new-contig_542 | 1 | CYT |  | 17.4 | 11.0 |
| TI-90 | 60S ribosomal protein L32 | M | 134 | TI-90 | * | 19 | ti-new-contig_128 | 2 | CYT |  | 16.1 | 11.4 |
| TI-498 | 60S ribosomal protein L37 | M | 90 | TI-498 | * | 20 | ti-new-contig_440 | 1 | CYT |  | 10.7 | 11.9 |
| TI-518 | 60S ribosomal protein L37 | M | 92 | TI-518 | * | 21 | ti-new-contig_455 | 1 | CYT |  | 10.2 | 10.7 |
| TI-457 | 60S ribosomal protein L39 | M | 51 | TI-457 | * | 22 | ti-new-contig_404 | 1 | CYT |  | 6.4 | 12.5 |
| TI-274 | Predicted RNA-binding protein cor | M | 300 | TI-274 | * | 23 | ti-new-contig_260 | 1 | CYT |  | 31.3 | 10.1 |
| TI-97 | Translation initiation factor 1 | M | 110 | TI-97 | * | 24 | ti-new-contig_140 | 2 | CYT |  | 12.5 | 6.8 |
| tin-650 | similar to Eukaryotic initiation factc | M | 148 | tin-650 | * | 25 | ti-new-contig_650 | 1 | CYT |  | 17.1 | 5.2 |
| TI-353 | Translation initiation factor 5A | M | 160 | TI-353 | * | 26 | ti-new-contig_316 | 1 | CYT |  | 17.6 | 5.0 |
| TI-336 | eukaryotic translation initiation fact | M | 215 | TI-336 | * | 27 | ti-new-contig_304 | 1 | CYT |  | 24.9 | 6.0 |
| **Protein export machinery** |  |  |  |  |  |  |  |  |  |  |  |  |
| tin-178 | Eclair Golgi protein | M | 217 | tin-178 | * | 1 | ti-new-contig_178 | 2 | SIG | 21-22 | 25.4 | 7.8 |
| tin_299 | putative rab11 | M | 215 | tin_299 | * | 2 | ti-new-contig_299 | 1 | CYT |  | 24.3 | 5.7 |
| TI-420 | Ras-related small GTPase | M | 218 | TI-420 | * | 3 | ti-new-contig_368 | 1 | CYT |  | 24.5 | 6.9 |
| TI-294 | Sec61 protein translocation compl | M | 96 | ti-294 | * | 4 | ti-new-contig_294 | 1 | CYT |  | 20.3 | 5.8 |
| tin-146 | signal recognition particle | M | 148 | tin-146 | * | 5 | ti-new-contig_146 | 2 | CYT |  | 16.5 | 10.0 |
| **Energy metabolism** |  |  |  |  |  |  |  |  |  |  |  |  |
| TI-62 | NADH dehydrogenase subunit 1 - | M | 289 | TI-62 | * | 1 | ti-new-contig_93 | 5 | SIG | 16-17 | 34.2 | 8.9 |
| TI-61 | truncated cytochrome b - mitochor | P | 368 | TI-61 | * | 2 | ti-new-contig_91 | 4 | ANC |  | 42.2 | 9.6 |
| TI-43 | cytochrome c oxidase subunit 2 - r | M | 223 | TI-43 | * | 3 | ti-new-contig_70 | 15 | BL | 40-41 | 25.9 | 6.5 |
| TI-538 | cytochrome c oxidase subunit Va | M | 150 | TI-538 | * | 4 | ti-new-contig_472 | 1 | CYT |  | 17.4 | 6.3 |
| TI-501 | Cytochrome c oxidase | M | 76 | TI-501 | * | 5 | ti-new-contig_439 | 1 | ANC |  | 8.6 | 9.9 |
| TI-606 | putative mitochondrial cytochrome | M | 118 | TI-606 | * | 6 | ti-new-contig_524 | 1 | CYT |  | 13.2 | 9.0 |
| TI-505 | hypothetical conserved protein | M | 100 | TI-505 | * | 7 | ti-new-contig_444 | 1 | ANC |  | 11.1 | 6.0 |
| TI-67 | NADH dehydrogenase subunit 6 - | M | 151 | TI-67 | * | 8 | ti-new-contig_101 | 4 | ANC |  | 17.7 | 10.2 |
| TI-81 | truncated ATPase subunit 6 - mito | M | 222 | TI-81 |  | 9 | ti-new-contig_95 | 5 | CYT |  | 25.4 | 9.7 |
| TI-71 | truncated  mitochondrial ADP/ATP | A | 299 | TI-71 | * | 10 | ti-new-contig_94 | 5 | CYT |  | 32.8 | 9.8 |
| TI-549 | NADH:ubiquinone oxidoreductase | M | 181 | TI-549 | * | 11 | ti-new-contig_483 | 1 | ANC |  | 21.3 | 9.3 |
| **Signal transduction machinery** |  |  |  |  |  |  |  |  |  |  |  |  |
| TI-491 | Ca2+-binding protein, EF-Hand pr | M | 178 | TI-491 | * | 1 | ti-new-contig_431 | 1 | CYT |  | 20.5 | 5.3 |
| TI-513 | G protein gamma subunit | M | 70 | TI-513 | * | 2 | ti-new-contig_452 | 1 | CYT |  | 8.2 | 6.6 |
| TI-408 | Rho GDP-dissociation inhibitor | M | 207 | TI-408 | * | 3 | ti-new-contig_362 | 1 | CYT |  | 23.5 | 5.1 |
| **Transporters and storage proteins** |  |  |  |  |  |  |  |  |  |  |  |  |
| TI-351 | ferritin | M | 172 | TI-351 | * | 1 | ti-new-contig_315 | 1 | CYT |  | 19.5 | 5.4 |
| TI-99 | Vacuolar ATP synthase 16 kDa pr | M | 156 | TI-99 | * | 2 | ti-new-contig_143 | 2 | BL |  | 16.1 | 8.9 |
| TI-375 | putative vacuolar ATP synthase su | M | 241 | TI-375 | * | 3 | ti-new-contig_331 | 1 | CYT |  | 27.1 | 9.7 |
| **Proteasome machinery** |  |  |  |  |  |  |  |  |  |  |  |  |
| TI-138 | E3 ubiquitin ligase interacting with | M | 136 | TI-138 | * | 1 | ti-new-contig_182 | 1 | CYT |  | 15.2 | 8.6 |
| TI-238 | SCF ubiquitin ligase, Rbx1 compo | M | 112 | TI-238 | * | 2 | ti-new-contig_228 | 1 | CYT |  | 12.9 | 5.5 |
| TI-93 | Ubiquitin C-terminal hydrolase UC | M | 228 | TI-93 | * | 3 | ti-new-contig_131 | 2 | CYT |  | 25.3 | 4.8 |
| **Cytoskeletal proteins** |  |  |  |  |  |  |  |  |  |  |  |  |
| tin-652 | actin-related protein Arp2/3 compl | M | 178 | tin-652 | * | 1 | ti-new-contig_652 | 1 | CYT |  | 20.5 | 8.6 |
| TI-283 | Calponin | M | 184 | TI-283 | * | 2 | ti-new-contig_265 | 1 | CYT |  | 20.4 | 7.7 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TI-341 | myosin 2 light chain | M | 151 | TI-341 | * | 3 | ti-new-contig_308 | 1 | CYT | | 16.9 | 4.6 |
| TI-219 | Syntaxin Interacting Protein 1 | M | 76 | TI-219 | * | 4 | ti-new-contig_218 | 1 | CYT | | 8.9 | 7.9 |
| TI-302 | Thymosin beta | M | 169 | TI-302 | * | 5 | ti-new-contig_282 | 1 | CYT | | 18.7 | 5.6 |
| **Detoxicatino and oxidant metabolism** | | | | | | | | | | | | |
| tin-284 | oxidoreductase | M | 252 | tin-284 | * | 1 | ti-new-contig_284 | 1 | BL | | 27.1 | 6.7 |
| tin-641 | superoxide dismutase | M | 154 | tin-641 | * | 2 | ti-new-contig_641 | 1 | CYT | | 16.0 | 5.8 |
| **Nucleic acid metabolism** | | | | | | | | | | | | |
| TI-198 | membrane-bound ribonuclease | M | 94 | TI-198 | * | 1 | ti-new-contig_206 | 1 | ANC | | 10.5 | 6.7 |
| **Possible cuticle protein** | | | | | | | | | | | | |
| ti-546 | conserved protein with chitin bindi | M | 213 | ti-546 | * | 1 | ti-new-contig_546 | 1 | SIG | 17-18 | 23.5 | 6.1 |
| **Conserved proteins of unknown function** | | | | | | | | | | | | |
| tin-589 | similar to testis enhanced gene tra | M | 234 | tin-589 | * | 1 | ti-new-contig_589 | 1 | BL | | 26.1 | 9.4 |
| tin-232 | FYVE finger containing protein | M | 263 | tin-232 | * | 2 | ti-new-contig_232 | 1 | CYT | | 29.6 | 9.0 |
| TI-277 | hypothetical conserved protein | M | 134 | TI-277 | * | 3 | ti-new-contig_138 | 1 | CYT | | 14.6 | 5.7 |
| tin-643 | hypothetical protein | M | 89 | tin-643 | * | 4 | ti-new-contig_643 | 1 | CYT | | 10.2 | 8.9 |
| TI-530 | hypothetical conserved protein | M | 239 | TI-530 | * | 5 | ti-new-contig_467 | 1 | CYT | | 25.6 | 5.2 |
| ti-294 | DJ-1 | M | 194 | ti-294 | * | 6 | ti-new-contig_294 | 1 | CYT | | 20.3 | 5.8 |
| TI-298 | hypothetical conserved insect prot | M | 109 | TI-298 | * | 7 | ti-new-contig_278 | 1 | ANC | | 12.9 | 9.7 |
| TI-248 | hypothetical conserved protein | M | 102 | TI-248 | * | 8 | ti-new-contig_239 | 1 | CYT | | 11.7 | 8.0 |
| TI-409 | hypothetical conserved protein | M | 219 | TI-409 | * | 9 | ti-new-contig_352 | 1 | CYT | | 25.3 | 9.3 |
| **Possible viral protein** | | | | | | | | | | | | |
| ti-590 | hypothetical viral protein found in | M | 152 | ti-590 | * | 1 | ti-new-contig_590 | 1 | CYT | | 17.2 | 7.6 |

| Mature MW | Mature pI | Best match to NR protein database | E value | Match | Extent of match | Length of best match | % identity | % Match length | First residue of match | First residue of sequence | Key words | Species | Best match to GO database | E value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 22.4 | 9.3 | pallidipin precursor | 3E-018 | gi\|1113799 | 172 | 195 | 31 | 88 | 1 | 1 | PALLIDIPI | Triatoma b | | |
| 17.3 | 8.7 | platelet inhibitor tripla | 4E-053 | gi\|1092403 | 180 | 178 | 58 | 101 | 1 | 1 | PLATELET | Triatoma in | | |
| 20.2 | 9.3 | pallidipin precursor | 2E-033 | gi\|1113799 | 192 | 195 | 42 | 98 | 1 | 1 | PALLIDIPI | Triatoma b | retinol binding pr | 0.046 |
| 16.4 | 4.4 | salivary triabin 1 | 2E-043 | gi\|1113798 | 154 | 163 | 55 | 94 | 1 | 1 | SALIVARY | Triatoma b | | |
| 18.0 | 7.7 | salivary lipocalin 4 | 3E-060 | gi\|1113798 | 179 | 179 | 62 | 100 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.0009 |
| 19.7 | 8.5 | lipocalin-like TiLipo77 | 2E-079 | gi\|3442165 | 193 | 193 | 74 | 100 | 1 | 1 | LIPOCALI | Triatoma in | retinol binding pr | 0.006 |
| 18.0 | 8.5 | salivary lipocalin 4 | 3E-058 | gi\|1113798 | 179 | 179 | 60 | 100 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.0004 |
| 28.6 | 4.4 | pallidipin 2 | 2E-013 | gi\|388359 | 199 | 188 | 26 | 106 | 1 | 1 | PALLIDIPI | Triatoma p | Hypothetical prot | 8E-09 |
| 18.6 | 9.1 | lipocalin-like TiLipo77 | 6E-040 | gi\|3442165 | 192 | 193 | 52 | 99 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 18.0 | 7.7 | salivary lipocalin 4 | 2E-059 | gi\|1113798 | 179 | 179 | 61 | 100 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.003 |
| 21.8 | 8.8 | pallidipin precursor | 7E-020 | gi\|1113799 | 179 | 195 | 33 | 92 | 1 | 1 | PALLIDIPI | Triatoma b | | |
| 20.5 | 9.2 | salivary lipocalin | 1E-053 | gi\|1113799 | 153 | 161 | 67 | 95 | 9 | 10 | SALIVARY | Triatoma b | | |
| 17.9 | 6.2 | salivary lipocalin 4 | 3E-060 | gi\|1113798 | 179 | 179 | 62 | 100 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.0001 |
| 20.1 | 9.0 | lipocalin-like TiLipo77 | 4E-093 | gi\|3442165 | 192 | 193 | 83 | 99 | 1 | 1 | LIPOCALI | Triatoma in | retinol binding pr | 0.025 |
| 22.1 | 9.2 | pallidipin precursor | 2E-018 | gi\|1113799 | 172 | 195 | 33 | 88 | 1 | 1 | PALLIDIPI | Triatoma b | | |
| | | pallidipin 2 | 1E-009 | gi\|388359 | 149 | 188 | 27 | 79 | 42 | 17 | PALLIDIPI | Triatoma p | Hypothetical prot | 9E-08 |
| 20.5 | 9.2 | pallidipin precursor | 7E-041 | gi\|1113799 | 184 | 195 | 47 | 94 | 1 | 1 | PALLIDIPI | Triatoma b | retinol binding pr | 0.06 |
| | | pallidipin precursor | 3E-010 | gi\|1113799 | 141 | 195 | 29 | 72 | 20 | 3 | PALLIDIPI | Triatoma b | PTMS: Parathym | 2E-05 |
| 17.9 | 4.9 | pallidipin 2 | 5E-053 | gi\|388359 | 185 | 188 | 57 | 98 | 1 | 1 | PALLIDIPI | Triatoma p | APOD: Apolipop | 0.004 |
| 18.6 | 8.9 | lipocalin-like TiLipo77 | 1E-041 | gi\|3442165 | 192 | 193 | 52 | 99 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 20.1 | 9.1 | salivary lipocalin | 1E-053 | gi\|1162671 | 191 | 190 | 56 | 101 | 3 | 5 | SALIVARY | Triatoma b | | |
| 19.6 | 8.9 | salivary lipocalin 6 | 3E-046 | gi\|1113799 | 190 | 183 | 51 | 104 | 1 | 1 | SALIVARY | Triatoma b | | |
| 20.1 | 9.2 | salivary lipocalin | 1E-062 | gi\|1113799 | 155 | 161 | 73 | 96 | 6 | 7 | SALIVARY | Triatoma b | Glial Lazarillo - e | 0.009 |
| | | lipocalin | 3E-010 | gi\|1113799 | 144 | 164 | 29 | 88 | 8 | 18 | LIPOCALI | Triatoma b | Hypothetical prot | 6E-08 |
| 16.8 | 6.9 | salivary triabin 1 | 1E-049 | gi\|1113798 | 163 | 163 | 57 | 100 | 1 | 1 | SALIVARY | Triatoma b | | |
| 20.9 | 5.3 | salivary lipocalin 5 | 5E-016 | gi\|1113799 | 176 | 178 | 30 | 99 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.02 |
| 26.8 | 5.4 | pallidipin 2 | 5E-013 | gi\|388359 | 229 | 188 | 26 | 122 | 1 | 1 | PALLIDIPI | Triatoma p | | |
| 18.6 | 9.1 | lipocalin-like TiLipo77 | 4E-043 | gi\|3442165 | 192 | 193 | 53 | 99 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 21.3 | 9.1 | lipocalin-like TiLipo77 | 2E-040 | gi\|3442165 | 193 | 193 | 50 | 100 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 20.3 | 9.0 | salivary lipocalin | 8E-057 | gi\|1162671 | 190 | 190 | 56 | 100 | 3 | 5 | SALIVARY | Triatoma b | retinol binding pr | 0.032 |
| 20.5 | 9.4 | salivary lipocalin | 9E-021 | gi\|1113799 | 165 | 179 | 38 | 92 | 3 | 5 | SALIVARY | Triatoma b | | |
| 21.0 | 4.9 | salivary lipocalin 5 | 3E-015 | gi\|1113799 | 176 | 178 | 29 | 99 | 1 | 1 | SALIVARY | Triatoma b | apolipoprotein D | 0.034 |
| 22.1 | 9.8 | pallidipin precursor | 1E-015 | gi\|1113799 | 165 | 195 | 29 | 85 | 1 | 1 | PALLIDIPI | Triatoma b | APOD: Apolipop | 0.0004 |
| 21.3 | 5.1 | pallidipin 2 | 6E-009 | gi\|388359 | 199 | 188 | 23 | 106 | 1 | 1 | PALLIDIPI | Triatoma p | CG33461 - tryps | 0.034 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21.2 | 5.5 | pallidipin 2 | 6E-009 | gi\|388359 | 199 | 188 | 23 | 106 | 1 | 1 | PALLIDIPI | Triatoma p | | |
| 18.6 | 9.0 | lipocalin-like TiLipo77 | 3E-040 | gi\|3442165 | 192 | 193 | 51 | 99 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 17.4 | 8.6 | lipocalin-like TiLipo39 | 4E-097 | gi\|3442165 | 179 | 179 | 96 | 100 | 1 | 1 | LIPOCALI | Triatoma in | | |
| 20.1 | 8.5 | triatin | 2E-074 | gi\|7172506 | 214 | 214 | 63 | 100 | 1 | 1 | TRIATIN T | Triatoma in | APOD: Apolipop | 0.002 |
| 19.5 | 7.0 | salivary lipocalin | 2E-058 | gi\|1162671 | 192 | 186 | 61 | 103 | 1 | 1 | SALIVARY | Triatoma b | | |
| 19.7 | 9.1 | salivary lipocalin 6 | 8E-065 | gi\|1162671 | 189 | 188 | 62 | 101 | 1 | 1 | SALIVARY | Triatoma b | | |
| 20.3 | 8.6 | salivary lipocalin 6 | 4E-046 | gi\|1113799 | 194 | 183 | 51 | 106 | 1 | 1 | SALIVARY | Triatoma b | retinol binding pr | 0.025 |
| 16.1 | 4.5 | salivary triabin 1 | 2E-044 | gi\|1113798 | 154 | 163 | 55 | 94 | 1 | 1 | SALIVARY | Triatoma b | | |
| 16.0 | 4.4 | salivary triabin 1 | 2E-043 | gi\|1113798 | 154 | 163 | 54 | 94 | 1 | 1 | SALIVARY | Triatoma b | | |
| 19.9 | 9.4 | salivary lipocalin | 3E-055 | gi\|1162671 | 168 | 175 | 64 | 96 | 6 | 20 | SALIVARY | Triatoma b | | |
| 20.0 | 9.3 | salivary lipocalin | 1E-049 | gi\|1113799 | 153 | 161 | 60 | 95 | 9 | 10 | SALIVARY | Triatoma b | | |
| 18.0 | 7.7 | pallidipin 2 | 1E-045 | gi\|388359 | 185 | 188 | 50 | 98 | 1 | 1 | PALLIDIPI | Triatoma p | | |
| 25.7 | 4.5 | pallidipin 2 | 3E-013 | gi\|388359 | 197 | 188 | 26 | 105 | 1 | 1 | PALLIDIPI | Triatoma p | bromodomain-co | 0.016 |
| 20.8 | 9.1 | salivary lipocalin 5 | 1E-022 | gi\|1113799 | 154 | 178 | 38 | 87 | 1 | 1 | SALIVARY | Triatoma b | | |
| 20.8 | 9.0 | salivary lipocalin 5 | 2E-022 | gi\|1113799 | 153 | 178 | 39 | 86 | 1 | 1 | SALIVARY | Triatoma b | | |
| 21.9 | 9.7 | pallidipin precursor | 3E-017 | gi\|1113799 | 165 | 195 | 32 | 85 | 1 | 1 | PALLIDIPI | Triatoma b | apolipoprotein D | 0.017 |
| 18.1 | 8.5 | platelet inhibitor tripla | 2E-092 | gi\|1092403 | 182 | 182 | 91 | 100 | 1 | 1 | PLATELET | Triatoma in | | |
| 10.9 | 5.2 | salivary lipocalin 1 | 4E-021 | gi\|1113798 | 73 | 179 | 67 | 41 | 1 | 1 | SALIVARY | Triatoma b | retinol binding pr | 0.015 |
| 21.1 | 9.2 | pallidipin precursor | 2E-023 | gi\|1113799 | 180 | 195 | 35 | 92 | 1 | 1 | PALLIDIPI | Triatoma b | | |
| 16.1 | 4.4 | salivary triabin 1 | 1E-045 | gi\|1113798 | 154 | 163 | 55 | 94 | 1 | 1 | SALIVARY | Triatoma b | | |
| | | | | | | | | | | | | | | |
| 6.3 | 3.9 | trialysin precursor | 3E-004 | gi\|1892064 | 26 | 260 | 80 | 10 | 1 | 1 | TRIALYSI | Triatoma in | OSJNBb0015G0 | 0.058 |
| 6.4 | 4.4 | trialysin precursor | 7E-005 | gi\|1892064 | 64 | 260 | 46 | 25 | 1 | 1 | TRIALYSI | Triatoma in | | |
| 6.2 | 4.1 | trialysin precursor | 1E-004 | gi\|1892064 | 39 | 260 | 66 | 15 | 1 | 1 | TRIALYSI | Triatoma in | | |
| 6.2 | 4.1 | trialysin precursor | 0.091 | gi\|1892064 | 39 | 260 | 58 | 15 | 1 | 1 | TRIALYSI | Triatoma in | | |
| 6.3 | 4.3 | trialysin precursor | 1E-006 | gi\|1892064 | 71 | 260 | 47 | 27 | 1 | 1 | TRIALYSI | Triatoma in | | |
| | | | | | | | | | | | | | | |
| 23.3 | 8.1 | VlpD76silD | 8E-006 | gi\|1065343 | 165 | 345 | 28 | 48 | 33 | 81 | VLPD76SI | Borrelia he | putative glycosid | 0.01 |
| 23.6 | 8.8 | KIAA1881 protein | 7E-005 | gi\|1562082 | 195 | 1348 | 27 | 14 | 6 | 13 | SAPIENS | Homo sapi | histidine kinase, | 0.023 |
| 26.3 | 4.9 | Periplasmic protein T | 1E-005 | gi\|1134762 | 233 | 1197 | 24 | 19 | 208 | 51 | PERIPLAS | Trichodes | MUC5B, MUC5: | 0.012 |
| 24.2 | 4.8 | Uncharacterized pha | 0.024 | gi\|1589515 | 179 | 1819 | 25 | 10 | 1331 | 59 | UNCHARA | Clostridium | sorbin and SH3 | 0.017 |
| 17.4 | 6.8 | lipoprotein, putative [ | 2E-004 | gi\|7073420 | 178 | 539 | 29 | 33 | 135 | 29 | LIPOPROT | Pseudomo | ELN: Elastin pre | 0.002 |
| | | | | | | | | | | | | | | |
| 26.4 | 8.3 | trialysin precursor | 1E-144 | gi\|1892064 | 260 | 260 | 98 | 100 | 1 | 1 | TRIALYSI | Triatoma in | | |
| 26.4 | 8.3 | trialysin precursor | 1E-145 | gi\|1892064 | 260 | 260 | 98 | 100 | 1 | 1 | TRIALYSI | Triatoma in | | |
| | | | | | | | | | | | | | | |
| 12.5 | 4.9 | triatox | 5E-075 | gi\|7172507 | 136 | 136 | 100 | 100 | 1 | 1 | TRIATOX | Triatoma in | | |
| | | | | | | | | | | | | | | |
| 25.7 | 9.1 | antigen-5-like protein | 3E-050 | gi\|3351869 | 213 | 250 | 47 | 85 | 52 | 32 | ANTIGEN- | Rhodnius | CRISP3: Cystein | 2E-20 |
| | | | | | | | | | | | | | | |
| 5.7 | 8.3 | Vasotab precursor | 0.039 | gi\|9473067 | 75 | 76 | 33 | 99 | 6 | 10 | VASOT_H | Hybomitra | | |
| 5.8 | 8.3 | secreted peptide | 6E-014 | gi\|1162671 | 78 | 77 | 53 | 101 | 1 | 4 | SECRETE | Triatoma b | SPINK5: Serine | 0.044 |
| | | | | | | | | | | | | | | |
| 14.1 | 5.3 | heme-binding protein | 4E-031 | gi\|2013646 | 129 | 128 | 49 | 101 | 1 | 1 | HEME-BIN | Rhodnius | Odorant-binding | 4E-05 |
| 14.2 | 5.7 | heme-binding protein | 7E-028 | gi\|2013646 | 130 | 128 | 47 | 102 | 1 | 1 | HEME-BIN | Rhodnius | Odorant-binding | 2E-05 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 34.0 | 9.4 | salivary inositol polyp | 1E-068 | gi\|9010134 | 321 | 321 | 44 | 100 | 1 | 1 | SALIVARY | Rhodnius p | CG9784 - depho | 4E-41 |
| 34.5 | 8.9 | salivary inositol polyp | 6E-072 | gi\|9010134 | 318 | 321 | 46 | 99 | 1 | 3 | SALIVARY | Rhodnius p | CG6805 - depho | 2E-47 |
| | | | | | | | | | | | | | | |
| 32.1 | 4.9 | trypsin precursor LIS | 5E-099 | gi\|3377261 | 284 | 299 | 59 | 95 | 17 | 24 | TRYPSIN | Lygus line | CG14760 - serin | 6E-36 |
| | | | | | | | | | | | | | | |
| 12.5 | 9.9 | conserved hypothetic | 3E-010 | gi\|1088788 | 100 | 117 | 40 | 85 | 21 | 28 | CONSERV | Aedes aeg | CG30413 - biolo | 0.027 |
| 12.3 | 9.8 | CG33998-PA [Droso | 4E-014 | gi\|8572504 | 117 | 119 | 40 | 98 | 5 | 17 | DROSOPH | Drosophila | CG31789 - biolo | 7E-08 |
| 12.3 | 9.8 | CG33998-PA [Droso | 3E-015 | gi\|8572504 | 117 | 119 | 41 | 98 | 5 | 17 | DROSOPH | Drosophila | CG31789 - biolo | 1E-07 |
| | | | | | | | | | | | | | | |
| 8.3 | 6.6 | defensin A | 2E-044 | gi\|2933595 | 94 | 94 | 88 | 100 | 1 | 1 | DEFENSIN | Rhodnius p | Defensin precurs | 2E-07 |
| 32.0 | 6.2 | Large exoprotein invo | 0.003 | gi\|8880982 | 277 | 8129 | 24 | 3 | 380 | 24 | LARGE EX | Synechoco | lipopolysacchari | 0.014 |
| 5.3 | 7.2 | hypothetical protein, | 0.99 | gi\|6812695 | 16 | 627 | 75 | 3 | 172 | 22 | HYPOTHE | Leishmania | | |
| 15.7 | 9.4 | immune-induced prot | 2E-026 | gi\|2773341 | 157 | 166 | 40 | 95 | 10 | 12 | IMMUNE-I | Manduca s | CG7532 - serine | 6E-22 |
| | | | | | | | | | | | | | | |
| 13.8 | 4.3 | hypothetical protein S | 0.031 | gi\|1626502 | 72 | 104 | 33 | 69 | 17 | 76 | HYPOTHE | Sinorhizobi | PRELP: Prolargi | 0.009 |
| | | | | | | | | | | | | | | |
| 5.9 | 4.0 | unknown salivary pro | 7E-029 | gi\|5155779 | 82 | 84 | 73 | 98 | 3 | 5 | UNKNOW | Culicoides | | |
| | | | | | | | | | | | | | | |
| 5.0 | 8.2 | YtxG [Bacillus sp. NR | 2.9 | gi\|8910011 | 53 | 176 | 37 | 30 | 18 | 9 | BACILLUS | Bacillus sp. | | |
| 17.0 | 5.6 | Thiolase [Halorhodos | 0.47 | gi\|8894846 | 73 | 394 | 35 | 19 | 111 | 25 | THIOLASE | Halorhodos | | |
| 14.7 | 8.7 | hypothetical protein T | 0.068 | gi\|8929462 | 55 | 549 | 32 | 10 | 290 | 90 | FUSOBAC | Tetrahyme | lava lamp - actin | 0.085 |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| | | PREDICTED: similar | 1E-121 | gi\|6653002 | 217 | 217 | 96 | 100 | 1 | 1 | SIMILAR T | Apis mellife | MOBKL1A, MOB | 0 |
| | | PREDICTED: H3 hist | 6E-070 | gi\|9439093 | 136 | 138 | 100 | 99 | 1 | 1 | HISTONE I | Mus muscu | H3 histone, famil | 9E-71 |
| | | histone H1 | 7E-041 | gi\|1426941 | 216 | 221 | 48 | 98 | 5 | 9 | HISTONE I | Rhynchoso | His1:CG31617 - | 2E-36 |
| | | | | | | | | | | | | | | |
| | | CG6770-PA [Drosoph | 1E-024 | gi\|2458380 | 62 | 69 | 79 | 90 | 1 | 1 | CG6770-P | Drosophila | NUPR1, COM1: | 9E-08 |
| | | putative elongation fa | 2E-077 | gi\|1106715 | 223 | 214 | 63 | 104 | 1 | 1 | ELONGAT | Diaphorina | Elongation facto | 1E-73 |
| | | PREDICTED: similar | 5E-052 | gi\|9109044 | 117 | 117 | 85 | 100 | 1 | 1 | SIMILAR T | Tribolium c | Rpb11 - DNA-dir | 9E-46 |
| | | PREDICTED: similar | 7E-036 | gi\|1107499 | 78 | 79 | 93 | 99 | 1 | 1 | SIMILAR T | Apis mellife | CG9344 - nuclea | 3E-35 |
| | | | | | | | | | | | | | | |
| | | Ribosomal protein L4 | 3E-006 | gi\|6247179 | 25 | 25 | 100 | 100 | 1 | 1 | RIBOSOM | Aedes aeg | ribosomal protei | 4E-07 |
| | | PREDICTED: similar | 3E-071 | gi\|1107564 | 194 | 155 | 71 | 125 | 1 | 1 | SIMILAR T | Apis mellife | ribosomal protei | 5E-66 |
| | | putative ribosomal pr | 8E-075 | gi\|1106714 | 147 | 147 | 91 | 100 | 1 | 1 | RIBOSOM | Diaphorina | Ribosomal prote | 3E-70 |
| | | PREDICTED: similar | 2E-067 | gi\|9109005 | 130 | 130 | 97 | 100 | 1 | 1 | SIMILAR T | Tribolium c | Ribosomal prote | 1E-66 |
| | | S17e ribosomal prote | 3E-061 | gi\|5034449 | 130 | 131 | 90 | 99 | 1 | 1 | RIBOSOM | Dascillus c | Ribosomal prote | 8E-58 |
| | | ribosomal protein S2 | 3E-049 | gi\|5460932 | 114 | 123 | 82 | 93 | 9 | 5 | RIBOSOM | Bombyx m | ribosomal protei | 2E-47 |
| | | putative ribosomal pr | 5E-055 | gi\|8947370 | 117 | 118 | 90 | 99 | 1 | 1 | RIBOSOM | Acyrthosip | Ribosomal prote | 4E-50 |
| | | putative ribosomal pr | 5E-041 | gi\|1106714 | 83 | 84 | 93 | 99 | 1 | 1 | RIBOSOM | Diaphorina | Ribosomal prote | 1E-41 |
| | | Ribosomal protein S2 | 2E-025 | gi\|4953285 | 56 | 56 | 89 | 100 | 1 | 1 | Q8WQI3 R | Plutella xyl | ribosomal protei | 5E-22 |
| | | Parcxpwex01 | 1E-134 | gi\|5786978 | 263 | 263 | 90 | 100 | 1 | 1 | PARCXPW | Periplaneta | Ribosomal prote | 0 |
| | | ribosomal protein S8 | 1E-100 | gi\|7427182 | 208 | 208 | 84 | 100 | 1 | 1 | RIBOSOM | Apis mellife | RPS8, OK/SW-c | 3E-98 |
| | | ribosomal protein P1 | 2E-045 | gi\|5460918 | 116 | 112 | 84 | 104 | 2 | 1 | RIBOSOM | Bombyx m | Ribosomal prote | 3E-45 |
| | | PREDICTED: similar | 1E-036 | gi\|9108509 | 114 | 112 | 68 | 102 | 1 | 1 | SIMILAR T | Tribolium c | RPLP2, RPP2: 6 | 1E-36 |
| | | PREDICTED: similar | 2E-096 | gi\|1107683 | 191 | 199 | 91 | 96 | 9 | 5 | SIMILAR T | Apis mellife | Ribosomal prote | 1E-90 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | putative ribosomal pr | 3E-078 | gi\|9082001 | 148 | 148 | 90 | 100 | 1 | 1 | RIBOSOM | Graphocep | RPL27A: 60S rib | 3E-70 |
|  |  | putative ribosomal pr | 6E-092 | gi\|9082004 | 176 | 177 | 92 | 99 | 1 | 1 | RIBOSOM | Graphocep | Ribosomal prote | 2E-78 |
|  |  | PREDICTED: similar | 1E-141 | gi\|6651692 | 257 | 257 | 93 | 100 | 1 | 1 | SIMILAR T | Apis mellife | Ribosomal prote | 0 |
|  |  | PREDICTED: similar | 1E-069 | gi\|9109361 | 150 | 150 | 87 | 100 | 1 | 1 | SIMILAR T | Tribolium c | Ribosomal prote | 2E-64 |
|  |  | putative ribosomal pr | 4E-067 | gi\|9082001 | 134 | 134 | 92 | 100 | 1 | 1 | RIBOSOM | Graphocep | ribosomal protein | 2E-60 |
|  |  | ribosomal protein L37 | 4E-042 | gi\|7090988 | 89 | 89 | 88 | 100 | 1 | 1 | RIBOSOM | Timarcha b | Ribosomal prote | 3E-40 |
|  |  | ribosomal protein L37 | 1E-044 | gi\|1162671 | 87 | 91 | 98 | 96 | 5 | 6 | RIBOSOM | Triatoma b | Ribosomal prote | 3E-40 |
|  |  | ribosomal protein L39 | 9E-023 | gi\|6693498 | 51 | 51 | 98 | 100 | 1 | 1 | RIBOSOM | Aedes aeg | Ribosomal prote | 7E-22 |
|  |  | PREDICTED: similar | 1E-055 | gi\|9108196 | 279 | 297 | 49 | 94 | 22 | 48 | SIMILAR T | Tribolium c | CSDA, DBPA: D | 9E-44 |
|  |  | PREDICTED: similar | 5E-058 | gi\|4809987 | 110 | 110 | 98 | 100 | 1 | 1 | SIMILAR T | Apis mellife | CG17737 - prote | 4E-57 |
|  |  | PREDICTED: similar | 4E-073 | gi\|1107560 | 148 | 147 | 90 | 101 | 1 | 1 | SIMILAR T | Apis mellife | Eukaryotic initiati | 1E-71 |
|  |  | Eukaryotic translation | 4E-083 | gi\|5170227 | 160 | 160 | 90 | 100 | 1 | 1 | IF5A_SPO | Spodoptera | eIF-5A - salivary | 5E-73 |
|  |  | eukaryotic translation | 6E-087 | gi\|8957449 | 215 | 215 | 67 | 100 | 1 | 1 | EUKARYO | Acyrthosip | eukaryotic transl | 3E-69 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 23.0 | 7.9 | ENSANGP0000002 | 3E-098 | gi\|5796702 | 205 | 217 | 82 | 94 | 13 | 13 | ANOPHEL | Anopheles | eclair - Golgi app | 3E-93 |
|  |  | putative rab11 | 1E-112 | gi\|4656176 | 215 | 215 | 95 | 100 | 1 | 1 | RAB11 HO | Homalodis | Rab-protein 11 - | 0 |
|  |  | PREDICTED: similar | 1E-083 | gi\|9107823 | 211 | 216 | 71 | 98 | 4 | 7 | SIMILAR T | Tribolium c | Rab GTPase - G | 1E-63 |
|  |  | GA12322-PA | 2E-051 | gi\|5463732 | 184 | 187 | 54 | 98 | 2 | 3 | DROSOPH | Drosophila | zgc:103725 - cel | 6E-51 |
|  |  | PREDICTED: similar | 2E-057 | gi\|9107704 | 148 | 152 | 72 | 97 | 9 | 1 | SIMILAR T | Tribolium c | Signal recognitio | 3E-46 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 32.5 | 9.0 | NADH dehydrogenas | 1E-157 | gi\|1118247 | 286 | 307 | 91 | 93 | 5 | 1 | DEHYDRO | Triatoma d | mitochondrial N | 0 |
|  |  | cytochrome b [Triato | 0.0 | gi\|1118247 | 367 | 377 | 88 | 97 | 10 | 1 | CYTOCHR | Triatoma d | mitochondrial Cy | 0 |
| 21.3 | 7.7 | cytochrome c oxidas | 1E-112 | gi\|1118246 | 221 | 226 | 88 | 98 | 6 | 3 | CYTOCHR | Triatoma d | mitochondrial Cy | 3E-96 |
|  |  | cytochrome c oxidas | 2E-050 | gi\|1630644 | 138 | 150 | 69 | 92 | 15 | 13 | CYTOCHR | Rhyzopert | Cytochrome c ox | 4E-38 |
|  |  | PREDICTED: similar | 4E-016 | gi\|9108557 | 75 | 80 | 52 | 94 | 1 | 1 | SIMILAR T | Tribolium c | cytochrome c ox | 2E-10 |
|  |  | putative mitochondria | 6E-041 | gi\|9082002 | 119 | 119 | 62 | 100 | 1 | 1 | MITOCHO | Graphocep | CG11015 - cyto | 2E-35 |
|  |  | hypothetical protein L | 3E-017 | gi\|4538763 | 102 | 103 | 48 | 99 | 1 | 1 | HYPOTHE | Danio rerio | zgc:77713 - biol | 1E-18 |
|  |  | NADH dehydrogenas | 2E-061 | gi\|1118247 | 151 | 167 | 76 | 90 | 17 | 1 | DEHYDRO | Triatoma d | mitochondrial N | 9E-29 |
|  |  | ATP synthase F0 sub | 1E-094 | gi\|1118246 | 222 | 227 | 74 | 98 | 1 | 1 | SYNTHAS | Triatoma d | mt:ATPase6, AT | 3E-81 |
|  |  | putative mitochondria | 1E-154 | gi\|5383070 | 299 | 309 | 88 | 97 | 11 | 1 | MITOCHO | Oncometo | stress-sensitive | 0 |
|  |  | ENSANGP0000001 | 1E-051 | gi\|5837855 | 188 | 189 | 54 | 99 | 1 | 1 | NADH-UB | Anopheles | lethal (3) neo18 | 4E-47 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | programmed cell dea | 1E-072 | gi\|1088783 | 171 | 174 | 73 | 98 | 4 | 8 | PROGRA | Aedes aeg | CG40410 - calci | 2E-62 |
|  |  | heterotrimeric guanin | 2E-026 | gi\|3346635 | 70 | 70 | 81 | 100 | 1 | 1 | HETEROT | Sitobion a | G protein &ggr; 1 | 2E-21 |
|  |  | PREDICTED: similar | 2E-086 | gi\|4812161 | 199 | 205 | 76 | 97 | 7 | 12 | SIMILAR T | Apis mellife | RhoGDI - Rho G | 5E-74 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | putative ferritin GF2 | 3E-067 | gi\|4656174 | 172 | 172 | 72 | 100 | 1 | 1 | FERRITIN | Homalodis | FTH1, FTH: Fer | 2E-49 |
|  |  | vacuolar H+ ATP syn | 1E-072 | gi\|9510260 | 152 | 155 | 93 | 98 | 4 | 5 | VACUOLA | Bombyx m | Vacuolar H<up> | 6E-73 |
|  |  | putative vacuolar AT | 1E-111 | gi\|4656175 | 244 | 244 | 86 | 100 | 1 | 1 | VACUOLA | Homalodis | Vha36 - hydroge | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | PREDICTED: similar | 1E-048 | gi\|1107593 | 140 | 155 | 65 | 90 | 1 | 1 | SIMILAR T | Apis mellife | CG3800 - nucle | 1E-45 |
|  |  | PREDICTED: similar | 4E-061 | gi\|4809504 | 113 | 113 | 93 | 100 | 1 | 1 | SIMILAR T | Apis mellife | Roc1a - ubiquiti | 2E-59 |
|  |  | PREDICTED: similar | 5E-063 | gi\|9107957 | 230 | 232 | 53 | 99 | 1 | 1 | SIMILAR T | Tribolium c | UCHL3: Ubiquiti | 1E-61 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | actin-related protein | 2E-078 | gi\|9446883 | 177 | 178 | 79 | 99 | 1 | 1 | ACTIN-RE | Aedes aeg | actin related prot | 3E-68 |
|  |  | PREDICTED: similar | 8E-087 | gi\|6651469 | 184 | 184 | 84 | 100 | 1 | 1 | SIMILAR T | Apis mellife | Muscle protein 2 | 8E-80 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | myosin 1 light chain [ | 4E-060 | gi\|5646225 | 150 | 150 | 74 | 100 | 1 | 1 | MYOSIN 1 | Lonomia o | Mlc1: Myosin ligh | 6E-47 |
| | | Syntaxin Interacting | 3E-030 | gi\|2476250 | 76 | 76 | 86 | 100 | 1 | 1 | SYNTAXIN | Drosophila | Syntaxin Interact | 9E-32 |
| | | PREDICTED: similar | 1E-062 | gi\|9107729 | 169 | 169 | 71 | 100 | 1 | 1 | SIMILAR T | Tribolium c | ciboulot - actin m | 2E-41 |
| | | | | | | | | | | | | | | |
| | | PREDICTED: similar | 2E-051 | gi\|9108471 | 250 | 255 | 43 | 98 | 5 | 1 | SIMILAR T | Tribolium c | CG9360 - oxidor | 1E-45 |
| | | superoxide dismutase | 1E-063 | gi\|5611773 | 154 | 154 | 75 | 100 | 1 | 1 | SUPEROX | Gryllotalpa | Superoxide dism | 2E-56 |
| | | | | | | | | | | | | | | |
| | | ribonuclease | 6E-030 | gi\|3213510 | 92 | 95 | 59 | 97 | 1 | 1 | RIBONUCL | Ceratitis ca | | |
| | | | | | | | | | | | | | | |
| 21.7 | 5.9 | PREDICTED: similar | 6E-033 | gi\|9107642 | 207 | 210 | 41 | 99 | 27 | 18 | SIMILAR T | Tribolium c | CG8505 - structu | 6E-08 |
| | | | | | | | | | | | | | | |
| | | PREDICTED: similar | 2E-074 | gi\|6653316 | 234 | 236 | 56 | 99 | 1 | 1 | SIMILAR T | Apis mellife | testis enhanced | 7E-54 |
| | | PREDICTED: similar | 1E-112 | gi\|9108732 | 258 | 282 | 77 | 91 | 1 | 1 | SIMILAR T | Tribolium c | pleckstrin homol | 2E-98 |
| | | PREDICTED: similar | 5E-035 | gi\|1107554 | 145 | 147 | 57 | 99 | 7 | 1 | SIMILAR T | Apis mellife | smell impaired 2 | 3E-32 |
| | | TPA: TPA_inf: HDC0 | 2E-017 | gi\|4161697 | 86 | 88 | 51 | 98 | 1 | 1 | TPA_INF D | Drosophila | zgc:103571 - cel | 1E-08 |
| | | PREDICTED: similar | 8E-038 | gi\|9109430 | 244 | 216 | 41 | 113 | 2 | 10 | SIMILAR T | Tribolium c | zgc:73223 - biolo | 4E-06 |
| | | GA12322-PA | 2E-051 | gi\|5463732 | 184 | 187 | 54 | 98 | 2 | 3 | DROSOPH | Drosophila | zgc:103725 - cel | 6E-51 |
| | | conserved hypothetic | 5E-026 | gi\|1088716 | 108 | 168 | 50 | 64 | 48 | 1 | CONSERV | Aedes aeg | | |
| | | ENSANGP0000001 | 7E-023 | gi\|5838943 | 87 | 109 | 59 | 80 | 23 | 16 | CONSERV | Anopheles | molecular functio | 8E-10 |
| | | PREDICTED: similar | 7E-056 | gi\|9107597 | 203 | 206 | 57 | 99 | 3 | 16 | SIMILAR T | Tribolium c | zgc:77492 - biolo | 2E-49 |
| | | | | | | | | | | | | | | |
| | | putative core protein | 1E-010 | gi\|9964462 | 147 | 156 | 28 | 94 | 6 | 7 | AMSACTA | Amsacta m | tryptophan perm | 0.025 |

| Function descriptors | Function second parent | GO # | E value of functional GO | Component descriptors | Component second parent | GO # | E value of component GO | Process descriptors | Process second parent | GO # | E value of process GO | Best match to KOG database | E value | General class | Best match to PFAM database |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | Apolipoprot | 2E-004 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Predicted | 0.63 | Energy pro | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.046 | | | | | Apolipoprot | 0.006 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.036 | | | | | Apolipoprot | 2E-004 | Cell wall/m | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.006 | | | | | CASK-inter | 0.46 | Signal tran | Triabin |
| lipid transp | transporter | GO:000531 | 0.061 | extracellula | extracellula | GO:000557 | 0.061 | lipid metab | transporter | GO:000662 | 0.061 | Apolipoprot | 9E-005 | Cell wall/m | Triabin |
| protein bin | binding | GO:000551 | 2E-07 | nucleus\|\|int | organelle | GO:000563 | 2E-07 | biological p | binding | GO:00000 | 2E-07 | Apolipoprot | 5E-004 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Ubiquitin-s | 0.21 | Posttransla | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.061 | | | | | Apolipoprot | 9E-004 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Apolipoprot | 4E-005 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Glutamine | 0.18 | Amino acid | Triabin |
| lipid transp | transporter | GO:000531 | 0.036 | extracellula | extracellula | GO:000557 | 0.036 | lipid metab | transporter | GO:000662 | 0.036 | Apolipoprot | 4E-004 | Cell wall/m | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.025 | | | | | | | | Triabin |
| | | | | | | | | | | | | Apolipoprot | 0.006 | Cell wall/m | Triabin |
| DNA bindin | binding | GO:000367 | 6E-07 | nucleosom | protein con | GO:000078 | 6E-07 | nucleosom | binding | GO:000633 | 6E-07 | HIV-1 Vpr- | 0.001 | Cell cycle | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.06 | | | | | Transcripti | 0.63 | Transcripti | Triabin |
| molecular f | molecular f | GO:000555 | 0.00002 | cellular con | cellular con | GO:000837 | 0.00002 | biological p | molecular f | GO:00000 | 0.00002 | Apolipoprot | 0.002 | Cell wall/m | Triabin |
| lipid transp | transporter | GO:000531 | 0.004 | extracellula | extracellula | GO:000557 | 0.004 | lipid metab | transporter | GO:000662 | 0.004 | FYVE finge | 0.049 | General fu | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | Apolipoprot | 0.18 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Pyruvate c | 0.092 | Energy pro | Triabin |
| binding\|\|m | binding | GO:000548 | 0.009 | extracellula | extracellula | GO:000561 | 0.009 | | | | | Apolipoprot | 0.004 | Cell wall/m | Triabin |
| DNA bindin | binding | GO:000367 | 7E-07 | nucleosom | protein con | GO:000078 | 7E-07 | nucleosom | binding | GO:000633 | 7E-07 | RNA-bindin | 4E-004 | RNA proce | Triabin |
| | | | | | | | | | | | | Subtilisin k | 0.12 | Posttransla | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.057 | | | | | Apolipoprot | 2E-005 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | Histidine a | 0.73 | General fu | Triabin |
| | | | | | | | | | | | | Thioredoxi | 0.66 | General fu | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.032 | | | | | Apolipoprot | 0.14 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | extracellula | extracellula | GO:000561 | 0.034 | | | | | Apolipoprot | 3E-004 | Cell wall/m | Triabin |
| lipid transp | transporter | GO:000531 | 0.0004 | extracellula | extracellula | GO:000557 | 0.0004 | lipid metab | transporter | GO:000662 | 0.0004 | Apolipoprot | 4E-004 | Cell wall/m | Triabin |
| trypsin acti | catalytic ac | GO:000429 | 0.034 | | | | | proteolysis | catalytic ac | GO:000650 | 0.034 | Apolipoprot | 0.001 | Cell wall/m | Triabin |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | Apolipoprot | 9E-005 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Protein req | 0.69 | Cell cycle c | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| lipid transp | transporter | GO:000531 | 0.002 | extracellula | extracellula | GO:000557 | 0.002 | lipid metab | transporter | GO:000662 | 0.002 | Apolipoprot | 0.50 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.025 | | | | | Apolipoprot | 0.007 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| molecular f | molecular f | GO:000555 | 0.016 | cellular con | cellular con | GO:000837 | 0.016 | biological p | molecular f | GO:000000 | 0.016 | Apolipoprot | 2E-004 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | | | | | | | | | Apolipoprot | 0.095 | Cell wall/m | Triabin |
| | | | | extracellula | extracellula | GO:000561 | 0.017 | | | | | Apolipoprot | 2E-004 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | cellular con | cellular con | GO:000837 | 0.015 | | | | | Apolipoprot | 0.091 | Cell wall/m | Triabin |
| | | | | | | | | | | | | Apolipoprot | 0.007 | Cell wall/m | Triabin |
| | | | | | | | | | | | | | | | Triabin |
| | | | | plastid\|\|intr | organelle | GO:000953 | 0.076 | | | | | | | | |
| hydrolase a | catalytic ac | GO:000455 | 0.01 | cellular con | cellular con | GO:000837 | 0.01 | carbohydra | catalytic ac | GO:000597 | 0.01 | Putative m | 0.11 | Posttransla | |
| myosin bin | binding | GO:001702 | 0.066 | extracellula | extracellula | GO:000557 | 0.066 | transformin | binding | GO:000717 | 0.066 | Putative m | 0.13 | Posttransla | |
| extracellula | structural m | GO:000520 | 0.012 | extracellula | extracellula | GO:000557 | 0.012 | cell adhesi | structural m | GO:000715 | 0.012 | RNA polym | 0.033 | RNA proce | Trypan_P |
| protein kina | catalytic ac | GO:000467 | 0.037 | cellular con | cellular con | GO:000837 | 0.037 | protein ami | catalytic ac | GO:000646 | 0.037 | | | | |
| extracellula | structural m | GO:003002 | 0.002 | extracellula | extracellula | GO:000557 | 0.002 | respiratory | structural m | GO:000758 | 0.002 | | | | |
| | | | | | | | | | | | | | | | DUF1336 |
| | | | | | | | | | | | | | | | DUF1336 |
| | | | | | | | | | | | | Calreticulin | 0.24 | Posttransla | DUF1448 |
| molecular f | molecular f | GO:000555 | 2.00E-20 | extracellula | extracellula | GO:000557 | 2.00E-20 | defense res | molecular f | GO:000695 | 2.00E-20 | Defense-re | 4E-026 | Function ur | SCP |
| | | | | | | | | | | | | | | | DUF1208 |
| serine-type | enzyme reg | GO:000486 | 0.044 | extracellula | extracellula | GO:000557 | 0.044 | negative re | enzyme reg | GO:001652 | 0.044 | | | | |
| odorant bin | binding | GO:000554 | 0.00004 | cellular con | cellular con | GO:000837 | 0.00004 | sensory pe | binding | GO:000760 | 0.00004 | | | | PBP_GOE |
| odorant bin | binding | GO:000554 | 0.00002 | extracellula | extracellula | GO:000557 | 0.00002 | sensory pe | binding | GO:000760 | 0.00002 | | | | PBP_GOE |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| inositol-pol | catalytic ac | GO:000444 | 4.00E-37 | integral to r | cell part | GO:001602 | 4.00E-37 | biological p | catalytic ac | GO:000000 | 4.00E-37 | Inositol-1,4 | 1E-043 | Intracellula | Exo_endo |
| inositol-pol | catalytic ac | GO:000444 | 2.00E-43 | integral to r | cell part | GO:001602 | 2.00E-43 | biological p | catalytic ac | GO:000000 | 2.00E-43 | Inositol-1,4 | 4E-051 | Intracellula | Exo_endo |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| serine-type | catalytic ac | GO:000425 | 3.00E-28 | plasma me | cell part | GO:000588 | 3.00E-28 | cytoskeleto | catalytic ac | GO:000701 | 3.00E-28 | Trypsin | 3E-044 | Amino acid | Trypsin |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| molecular f | molecular f | GO:000555 | 0.027 | cellular con | cellular con | GO:000837 | 0.027 | biological p | molecular f | GO:000000 | 0.027 | Uncharacte | 0.046 | Signal tran | DUF936 |
| molecular f | molecular f | GO:000555 | 7E-08 | cellular con | cellular con | GO:000837 | 7E-08 | biological p | molecular f | GO:000000 | 7E-08 | | | | |
| molecular f | molecular f | GO:000555 | 1E-07 | cellular con | cellular con | GO:000837 | 1E-07 | biological p | molecular f | GO:000000 | 1E-07 | | | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| protease in | enzyme reg | GO:003041 | 0.098 | | | | | defense res | enzyme reg | GO:000695 | 0.098 | Inhibitor of | 0.41 | Signal tran | Defensin_ |
| lipopolysac | binding | GO:000153 | 0.014 | extracellula | extracellula | GO:000561 | 0.014 | | | | | | | | Fibrillarin |
| | | | | | | | | | | | | Chromatin | 0.38 | Chromatin | |
| ferric-chela | catalytic ac | GO:000029 | 1E-13 | | | | | | | | | Rab3 effec | 0.42 | Intracellula | Reeler |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| extracellula | structural m | GO:000520 | 0.009 | extracellula | extracellula | GO:000557 | 0.009 | skeletal de | structural m | GO:000150 | 0.009 | Predicted p | 0.24 | Amino acid | Neisseria_ |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | Mlp |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | DUF569 |
| actin bindin | binding | GO:000377 | 0.085 | Golgi appa | organelle | GO:000579 | 0.085 | cellularizati | binding | GO:000734 | 0.085 | Nuclear pro | 0.065 | General fu | TolA |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| kinase acti | enzyme reg | GO:001920 | 1.00E-115 | nucleus\|\|int | organelle | GO:000563 | 1.00E-115 | protein ami | enzyme reg | GO:004677 | 1.00E-115 | Cell cycle-a | 3E-091 | Cell cycle c | Mob1_pho |
| DNA bindin | binding | GO:000367 | 9.00E-71 | nucleosom | protein con | GO:000078 | 9.00E-71 | nucleosom | binding | GO:000633 | 9.00E-71 | Histones H | 2E-052 | Chromatin | Histone |
| DNA bindin | binding | GO:000367 | 2.00E-36 | nucleus\|\|int | organelle | GO:000563 | 2.00E-36 | chromatin a | binding | GO:000633 | 2.00E-36 | Microtubule | 0.016 | Cell cycle c | Linker_his |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| molecular f | molecular f | GO:000555 | 9E-08 | nucleus\|\|int | organelle | GO:000563 | 9E-08 | induction o | molecular f | GO:000691 | 9E-08 | DNA-bindin | 1E-011 | Transcriptio | |
| translation | translation | GO:000374 | 2.00E-57 | eukaryotic | protein con | GO:000585 | 2.00E-57 | actin filame | translation | GO:003004 | 2.00E-57 | Elongation | 1E-066 | Transcriptio | EF1_GNE |
| DNA-direct | catalytic ac | GO:000389 | 9.00E-46 | DNA-direct | protein con | GO:000566 | 9.00E-46 | transcriptio | catalytic ac | GO:000636 | 9.00E-46 | RNA polym | 3E-041 | Transcriptio | RNA_pol_ |
| protein het | binding | GO:004698 | 2.00E-32 | cytoplasm\|\| | cell part | GO:000573 | 2.00E-32 | mRNA cata | binding | GO:000640 | 2.00E-32 | Small nucle | 2E-030 | RNA proce | LSM |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| molecular f | molecular f | GO:000555 | 4E-07 | cellular con | cellular con | GO:000837 | 4E-07 | biological p | molecular f | GO:000000 | 4E-07 | | | | |
| nucleic aci | binding | GO:000367 | 5.00E-66 | intracellula | cell part | GO:000562 | 5.00E-66 | protein bios | binding | GO:000641 | 5.00E-66 | 40S riboso | 2E-065 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 3.00E-70 | cytosolic sm | protein con | GO:000584 | 3.00E-70 | protein bios | structural m | GO:000641 | 3.00E-70 | 40S riboso | 3E-069 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 1.00E-66 | cytosolic sm | protein con | GO:000584 | 1.00E-66 | protein bios | structural m | GO:000641 | 1.00E-66 | 40S riboso | 3E-062 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 8.00E-58 | cytosolic sm | protein con | GO:000584 | 8.00E-58 | protein bios | structural m | GO:000641 | 8.00E-58 | 40S riboso | 2E-052 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 2.00E-47 | intracellula | cell part | GO:000562 | 2.00E-47 | protein bios | structural m | GO:000641 | 2.00E-47 | 40S riboso | 1E-035 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 4.00E-50 | cytosolic sm | protein con | GO:000584 | 4.00E-50 | protein bios | structural m | GO:000641 | 4.00E-50 | 40s ribosor | 7E-034 | Translation | Ribosoma |
| protein bind | binding | GO:000551 | 4.00E-40 | intracellula | cell part | GO:000562 | 4.00E-40 | signal trans | binding | GO:000716 | 4.00E-40 | 40s ribosor | 2E-034 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 5.00E-22 | intracellula | cell part | GO:000562 | 5.00E-22 | protein bios | structural m | GO:000641 | 5.00E-22 | 40S riboso | 1E-019 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 1.00E-115 | cytosolic sm | protein con | GO:000584 | 1.00E-115 | protein bios | structural m | GO:000641 | 1.00E-115 | 40S riboso | 2E-091 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 3.00E-98 | cytosolic sm | protein con | GO:000584 | 3.00E-98 | protein bios | structural m | GO:000641 | 3.00E-98 | 40S riboso | 4E-079 | Translation | Ribosoma |
| structural c | structural m | GO:000373 | 3.00E-45 | cytosolic la | protein con | GO:000584 | 3.00E-45 | protein bios | structural m | GO:000641 | 3.00E-45 | 60s acidic | 1E-025 | Translation | Ribosoma |
| RNA bindin | binding | GO:000372 | 1.00E-36 | cytosolic la | protein con | GO:000584 | 1.00E-36 | protein bios | binding | GO:000641 | 1.00E-36 | 60S acidic | 9E-027 | Translation | Ribosoma |
| protein bind | binding | GO:000551 | 1.00E-90 | cytosolic la | protein con | GO:000584 | 1.00E-90 | protein bios | binding | GO:000641 | 1.00E-90 | 60S riboso | 7E-087 | Translation | Ribosoma |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RNA bindir | binding | GO:000372 | 3.00E-70 | cytosolic la | protein con | GO:000584 | 3.00E-70 | protein bios | binding | GO:000641 | 3.00E-70 | 60s riboson | 2E-053 | Translation | L15 |
| structural c | structural n | GO:000373 | 2.00E-78 | cytosolic la | protein con | GO:000584 | 2.00E-78 | protein bios | structural n | GO:000641 | 2.00E-78 | 60S ribosoi | 2E-069 | Translation | Ribosoma |
| structural c | structural n | GO:000373 | 1.00E-130 | cytosolic la | protein con | GO:000584 | 1.00E-130 | protein bios | structural n | GO:000641 | 1.00E-130 | 60s riboson | 1E-104 | Translation | Ribosoma |
| structural c | structural n | GO:000373 | 2.00E-64 | cytosolic la | protein con | GO:000584 | 2.00E-64 | protein bios | structural n | GO:000641 | 2.00E-64 | 60S ribosoi | 3E-044 | Translation | KOW |
| structural c | structural n | GO:000373 | 2.00E-60 | intracellula | cell part | GO:000562 | 2.00E-60 | protein bios | structural n | GO:000641 | 2.00E-60 | 60S ribosoi | 3E-046 | Translation | Ribosoma |
| structural c | structural n | GO:000373 | 3.00E-40 | cytosolic la | protein con | GO:000584 | 3.00E-40 | protein bios | structural n | GO:000641 | 3.00E-40 | 60S ribosoi | 4E-024 | Translation | Ribosoma |
| structural c | structural n | GO:000373 | 3.00E-40 | cytosolic la | protein con | GO:000584 | 3.00E-40 | protein bios | structural n | GO:000641 | 3.00E-40 | 60S ribosoi | 5E-029 | Translation | Ribosoma |
| structural c | structural n | GO:000373 | 7.00E-22 | cytosolic la | protein con | GO:000584 | 7.00E-22 | protein bios | structural n | GO:000641 | 7.00E-22 | 60s riboson | 1E-006 | Translation | Ribosoma |
| double-stra | binding | GO:000369 | 9.00E-44 | cytoplasm\|\| | cell part | GO:000573 | 9.00E-44 | negative re | binding | GO:000012 | 9.00E-44 | Predicted F | 2E-018 | Translation | CSD |
| translation | translation | GO:000374 | 4.00E-41 | cytoplasm\|\| | cell part | GO:000573 | 4.00E-41 | regulation c | translation | GO:000644 | 4.00E-41 | Translation | 1E-041 | Translation | SUI1 |
| translation | translation | GO:000374 | 1.00E-71 | cytosol\|\|cyt | cell part | GO:000582 | 1.00E-71 | smoothene | translation | GO:000722 | 1.00E-71 | Translation | 7E-054 | Translation | eIF-1a |
| translation | translation | GO:004518 | 5.00E-73 | cytosol\|\|cyt | cell part | GO:000582 | 5.00E-73 | <salivary g | translation | GO:003507 | 5.00E-73 | Translation | 2E-061 | Translation | eIF-5a |
| translation | translation | GO:000374 | 1.00E-61 | eukaryotic | protein con | GO:000585 | 1.00E-61 | | | | | Uncharacte | 7E-073 | Function ur | eIF-3_p25 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| protein carr | transporter | GO:000832 | 6.00E-77 | membrane | cell part | GO:001602 | 6.00E-77 | intracellula | transporter | GO:000688 | 6.00E-77 | emp24/gp2 | 2E-074 | Intracellula | EMP24_G |
| transporter | transporter | GO:000521 | 1.00E-96 | Golgi trans | organelle p | GO:000580 | 1.00E-96 | two-compo | transporter | GO:000016 | 1.00E-96 | GTPase Ra | 5E-093 | Intracellula | Ras |
| GTP bindir | binding | GO:000552 | 1.00E-63 | intracellula | cell part | GO:000562 | 1.00E-63 | small GTPa | binding | GO:000726 | 1.00E-63 | GTPase Ra | 2E-081 | General fur | Ras |
| molecular f | molecular f | GO:000555 | 6.00E-51 | cellular con | cellular con | GO:000837 | 6.00E-51 | biological p | molecular f | GO:000000 | 6.00E-51 | Putative tra | 1E-050 | General fur | DJ-1_PfpI |
| 7S RNA bir | binding | GO:000831 | 3.00E-46 | signal reco | protein con | GO:000578 | 3.00E-46 | SRP-deper | binding | GO:000661 | 3.00E-46 | Signal reco | 5E-033 | Intracellula | SRP19 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NADH deh | catalytic ac | GO:000813 | 3.00E-35 | plastid\|\|intr | organelle | GO:000953 | 3.00E-35 | electron tra | catalytic ac | GO:000611 | 3.00E-35 | NADH deh | 1E-080 | Energy pro | NADHdh |
| ubiquinol-c | transporter | GO:000812 | 1.00E-115 | respiratory | protein con | GO:000575 | 1.00E-115 | aerobic res | transporter | GO:000906 | 1.00E-115 | Cytochrom | 4E-069 | Energy pro | Cytochron |
| cytochrome | transporter | GO:000412 | 2.00E-60 | respiratory | protein con | GO:000575 | 2.00E-60 | aerobic res | transporter | GO:000906 | 2.00E-60 | Cytochrom | 8E-093 | Energy pro | COX2 |
| cytochrome | transporter | GO:000412 | 4.00E-38 | respiratory | protein con | GO:000575 | 4.00E-38 | electron tra | transporter | GO:000611 | 4.00E-38 | Cytochrom | 2E-043 | Energy pro | COX5A |
| oxidoreduc | catalytic ac | GO:001649 | 2E-10 | mitochondr | organelle | GO:000573 | 2E-10 | electron tra | catalytic ac | GO:000611 | 2E-10 | | | | COX6C |
| cytochrome | transporter | GO:000412 | 2E-07 | mitochondr | organelle | GO:000573 | 2E-07 | electron tra | transporter | GO:000611 | 2E-07 | Cytochrom | 6E-029 | Energy pro | COX5B |
| molecular f | molecular f | GO:000555 | 1.00E-18 | cellular con | cellular con | GO:000837 | 1.00E-18 | biological p | molecular f | GO:000000 | 1.00E-18 | | | | Ion_trans |
| NADH deh | catalytic ac | GO:000813 | 9.00E-29 | mitochondr | organelle | GO:000573 | 9.00E-29 | | | | | | | | Herpes_L |
| hydrogen-e | transporter | GO:000855 | 3.00E-81 | mitochondr | organelle | GO:000573 | 3.00E-81 | proton tran | transporter | GO:001599 | 3.00E-81 | ATP syntha | 4E-034 | Energy pro | ATP-synt_ |
| ATP\:ADP | transporter | GO:000547 | 1.00E-142 | mitochondr | organelle p | GO:000574 | 1.00E-142 | flight behav | transporter | GO:000762 | 1.00E-142 | Mitochondr | 1E-132 | Energy pro | Mito_carr |
| molecular f | molecular f | GO:000555 | 7.00E-25 | cellular con | cellular con | GO:000837 | 7.00E-25 | biological p | molecular f | GO:000000 | 7.00E-25 | NADH:ubiq | 4E-043 | Energy pro | Octopine_ |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| protein binc | binding | GO:000551 | 2.00E-60 | soluble frac | cell part | GO:000562 | 2.00E-60 | caspase ac | binding | GO:000691 | 2.00E-60 | Ca2+-bindi | 5E-060 | Signal tran | GD_AH_C |
| GTPase ac | catalytic ac | GO:000392 | 2.00E-21 | heterotrime | protein con | GO:000583 | 2.00E-21 | actin filame | catalytic ac | GO:000701 | 2.00E-21 | G protein g | 1E-009 | Signal tran | G-gamma |
| GTPase ac | enzyme reg | GO:000509 | 5.00E-54 | cytoplasm\|\| | cell part | GO:000573 | 5.00E-54 | Rho protein | enzyme reg | GO:000726 | 5.00E-54 | Rho GDP-c | 1E-061 | Signal tran | Rho_GDI |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| kinase binc | binding | GO:001990 | 2.00E-49 | plasma me | cell part | GO:000588 | 2.00E-49 | immune res | binding | GO:000695 | 2.00E-49 | Ferritin | 9E-053 | Inorganic ic | Ferritin |
| hydrogen-e | transporter | GO:000855 | 6.00E-73 | hydrogen-tı | protein con | GO:000022 | 6.00E-73 | proton tran | transporter | GO:001599 | 6.00E-73 | Vacuolar H | 1E-046 | Energy pro | ATP-synt_ |
| hydrogen-e | transporter | GO:000855 | 1.00E-108 | hydrogen-tı | protein con | GO:000022 | 1.00E-108 | proton tran | transporter | GO:001599 | 1.00E-108 | Vacuolar H | 1E-083 | Energy pro | ATP-synt_ |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| transcriptio | transcriptio | GO:000370 | 3.00E-27 | nucleus\|\|int | organelle | GO:000563 | 3.00E-27 | positive reg | transcriptio | GO:000828 | 3.00E-27 | E3 ubiquitir | 7E-011 | Posttransla | zf-CCHC |
| ubiquitin-pr | catalytic ac | GO:000484 | 2.00E-59 | nucleus\|\|int | organelle | GO:000563 | 2.00E-59 | protein ubic | catalytic ac | GO:001656 | 2.00E-59 | SCF ubiqui | 3E-042 | Posttransla | zf-C3HC4 |
| ubiquitin th | catalytic ac | GO:000422 | 1.00E-61 | cytoplasm\|\| | cell part | GO:000573 | 1.00E-61 | ubiquitin-de | catalytic ac | GO:000651 | 1.00E-61 | Ubiquitin C | 6E-069 | Posttransla | Peptidase |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| structural c | structural n | GO:000520 | 3.00E-68 | <lamellipodium | | GO:003002 | 3.00E-68 | cell motility | structural n | GO:000692 | 3.00E-68 | Actin-relate | 1E-073 | Cytoskeleto | P21-Arc |
| actin bindir | binding | GO:000377 | 8.00E-80 | contractile | cell part | GO:004329 | 8.00E-80 | cell adhesi | binding | GO:000715 | 8.00E-80 | Calponin | 9E-057 | Cytoskeleto | CH |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| microfilame | motor activ | GO:000014 | 6.00E-47 | muscle my | protein con | GO:000585 | 6.00E-47 | muscle cor | motor activ | GO:000693 | 6.00E-47 | Myosin ess | 1E-029 | Cytoskeleto | Caleosin |
| protein bin | binding | GO:000551 | 5E-12 | SCAR com | protein con | GO:003120 | 5E-12 | regulation | binding | GO:000806 | 5E-12 | Predicted | 0.024 | Function u | PV-1 |
| actin mono | binding | GO:000378 | 2.00E-41 | cytosol\|\|cyt | cell part | GO:000582 | 2.00E-41 | <brain deve | binding | GO:000742 | 2.00E-41 | Thymosin | 3E-012 | Cell motility | Thymosin |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| oxidoreduc | catalytic ac | GO:001662 | 1.00E-21 | nucleus\|\|int | organelle | GO:000563 | 1.00E-21 | metabolism | catalytic ac | GO:000815 | 1.00E-21 | Predicted | 5E-047 | Secondary | adh_short |
| copper\, zi | catalytic ac | GO:000478 | 2.00E-56 | cytoplasm\|\| | cell part | GO:000573 | 2.00E-56 | determinati | catalytic ac | GO:000834 | 2.00E-56 | Cu2+/Zn2+ | 4E-049 | Inorganic i | Sod_Cu |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | Uncharacte | 0.16 | Function u | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| molecular f | molecular f | GO:000555 | 0.0002 | cellular con | cellular con | GO:000837 | 0.0002 | biological p | molecular f | GO:000000 | 0.0002 | RNA polym | 0.063 | Transcriptio | Chitin_bin |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| molecular f | molecular f | GO:000555 | 2E-08 | cellular con | cellular con | GO:000837 | 2E-08 | biological p | molecular f | GO:000000 | 2E-08 | Bax-mediat | 2E-050 | Defense m | UPF0005 |
| protein bin | binding | GO:000551 | 3.00E-27 | <ruffle | | GO:000172 | 3.00E-27 | cytoskeleto | binding | GO:000701 | 3.00E-27 | FYVE finge | 9E-068 | General fu | FYVE |
| | | | | | | | | olfactory be | binding | GO:004204 | 3.00E-32 | Glucose-re | 0.32 | General fu | HypA |
| molecular f | molecular f | GO:000555 | 1E-08 | cellular con | cellular con | GO:000837 | 1E-08 | biological p | molecular f | GO:000000 | 1E-08 | Glycine del | 0.12 | Amino acid | GDC-P |
| molecular f | molecular f | GO:000555 | 0.000004 | cellular con | cellular con | GO:000837 | 0.000004 | biological p | molecular f | GO:000000 | 0.000004 | Nidogen a | 0.42 | Cell wall/m | Spot_14 |
| molecular f | molecular f | GO:000555 | 6.00E-51 | cellular con | cellular con | GO:000837 | 6.00E-51 | biological p | molecular f | GO:000000 | 6.00E-51 | Putative tra | 1E-050 | General fu | DJ-1_PfpI |
| | | | | | | | | | | | | Ribosomal | 0.97 | General fu | |
| molecular f | molecular f | GO:000555 | 6E-07 | mitochondr | organelle | GO:000573 | 6E-07 | biological p | molecular f | GO:000000 | 6E-07 | Uncharacte | 4E-015 | Function u | WTF |
| molecular f | molecular f | GO:000555 | 4.00E-25 | cellular con | cellular con | GO:000837 | 4.00E-25 | biological p | molecular f | GO:000000 | 4.00E-25 | Uncharacte | 7E-025 | Function u | DUF1014 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| aromatic a | transporter | GO:001517 | 0.025 | plasma me | cell part | GO:000588 | 0.025 | telomere m | transporter | GO:000072 | 0.025 | Ca2+-perm | 0.32 | Inorganic i | Herpes_L |

| E value | Best match to SMART database | E value | Clustered at 30% Sim- on 50% of length - Cluster# | # seqs | Clustered at 50% Sim- on 50% of length - Cluster# | # seqs | Clustered at 70% Sim- on 50% of length - Cluster# | # seqs | Clustered at 90% Sim- on 50% of length - Cluster# | # seqs | Clustered at 95% Id- on 70% of length - Cluster# | # seqs | Psi-blast clustering - no signal peptide | # seqs |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3E-012 |  |  | 1 | 55 | 1 | 48 | 3 | 5 | 7 | 2 | 34 | 1 | 1 | 52 |
| 6E-019 | TOP2c | 0.70 | 1 | 55 | 2 | 5 | 41 | 1 | 45 | 1 | 51 | 1 | 1 | 52 |
| 6E-015 |  |  | 1 | 55 | 1 | 48 | 16 | 2 | 43 | 1 | 49 | 1 | 1 | 52 |
| 3E-017 | LIGANc | 0.51 | 1 | 55 | 1 | 48 | 6 | 5 | 80 | 1 | 90 | 1 | 1 | 52 |
| 1E-020 | FN3 | 0.077 | 1 | 55 | 1 | 48 | 7 | 4 | 1 | 4 | 24 | 1 | 1 | 52 |
| 8E-009 | POL3Bc | 0.37 | 1 | 55 | 1 | 48 | 11 | 3 | 46 | 1 | 52 | 1 | 1 | 52 |
| 7E-019 | FN3 | 0.19 | 1 | 55 | 1 | 48 | 7 | 4 | 1 | 4 | 27 | 1 | 1 | 52 |
| 2E-011 | PTPc | 0.44 | 1 | 55 | 1 | 48 | 5 | 5 | 147 | 1 | 159 | 1 | 1 | 52 |
| 4E-018 |  |  | 1 | 55 | 1 | 48 | 4 | 5 | 2 | 4 | 36 | 1 | 1 | 52 |
| 3E-019 | FN3 | 0.068 | 1 | 55 | 1 | 48 | 7 | 4 | 1 | 4 | 26 | 1 | 1 | 52 |
| 1E-013 | RL11 | 0.61 | 1 | 55 | 1 | 48 | 3 | 5 | 30 | 1 | 30 | 1 | 1 | 52 |
| 6E-015 | S4 | 0.98 | 1 | 55 | 1 | 48 | 8 | 4 | 78 | 1 | 88 | 1 | 1 | 52 |
| 2E-019 | FN3 | 0.084 | 1 | 55 | 1 | 48 | 7 | 4 | 1 | 4 | 23 | 1 | 1 | 52 |
| 1E-009 | ZnF_TTF | 0.25 | 1 | 55 | 1 | 48 | 11 | 3 | 51 | 1 | 57 | 1 | 1 | 52 |
| 6E-013 |  |  | 1 | 55 | 1 | 48 | 3 | 5 | 7 | 2 | 32 | 1 | 1 | 52 |
| 3E-010 | PTB | 0.60 | 1 | 55 | 1 | 48 | 5 | 5 | 130 | 1 | 142 | 1 |  |  |
| 3E-018 |  |  | 1 | 55 | 1 | 48 | 16 | 2 | 95 | 1 | 105 | 1 | 1 | 52 |
| 4E-011 | DWB | 0.80 | 1 | 55 | 1 | 48 | 5 | 5 | 55 | 1 | 61 | 1 |  |  |
| 1E-021 |  |  | 1 | 55 | 2 | 5 | 14 | 2 | 17 | 1 | 10 | 1 | 1 | 52 |
| 2E-017 |  |  | 1 | 55 | 1 | 48 | 4 | 5 | 2 | 4 | 44 | 1 | 1 | 52 |
| 2E-012 |  |  | 1 | 55 | 1 | 48 | 1 | 5 | 85 | 1 | 95 | 1 | 1 | 52 |
| 9E-012 |  |  | 1 | 55 | 1 | 48 | 1 | 5 | 84 | 1 | 94 | 1 | 1 | 52 |
| 3E-013 | Aamy_C | 0.69 | 1 | 55 | 1 | 48 | 8 | 4 | 86 | 1 | 96 | 1 | 1 | 52 |
| 6E-011 | PTPc | 0.66 | 1 | 55 | 1 | 48 | 5 | 5 | 133 | 1 | 145 | 1 |  |  |
| 5E-020 | SERPIN | 0.64 | 1 | 55 | 1 | 48 | 6 | 5 | 81 | 1 | 91 | 1 | 1 | 52 |
| 7E-016 | SR | 0.54 | 1 | 55 | 1 | 48 | 9 | 4 | 12 | 2 | 6 | 2 | 1 | 52 |
| 6E-007 | SapB | 0.18 | 1 | 55 | 71 | 1 | 83 | 1 | 104 | 1 | 114 | 1 | 1 | 52 |
| 3E-016 |  |  | 1 | 55 | 1 | 48 | 4 | 5 | 2 | 4 | 38 | 1 | 1 | 52 |
| 3E-018 | B_lectin | 0.099 | 1 | 55 | 1 | 48 | 4 | 5 | 41 | 1 | 47 | 1 | 1 | 52 |
| 4E-013 | RPOLD | 0.33 | 1 | 55 | 1 | 48 | 1 | 5 | 103 | 1 | 113 | 1 | 1 | 52 |
| 3E-015 |  |  | 1 | 55 | 1 | 48 | 10 | 3 | 73 | 1 | 82 | 1 | 1 | 52 |
| 2E-014 | SR | 0.74 | 1 | 55 | 1 | 48 | 9 | 4 | 12 | 2 | 6 | 2 | 1 | 52 |
| 4E-015 | DUF1 | 0.61 | 1 | 55 | 1 | 48 | 3 | 5 | 14 | 2 | 131 | 1 | 1 | 52 |
| 9E-012 | TLDc | 0.44 | 1 | 55 | 1 | 48 | 9 | 4 | 13 | 2 | 7 | 2 | 1 | 52 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7E-012 | TLDc | 0.61 | 1 | 55 | 1 | 48 | 9 | 4 | 13 | 2 | 7 | 2 | 1 | 52 |
| 4E-016 | | | 1 | 55 | 1 | 48 | 4 | 5 | 2 | 4 | 41 | 1 | 1 | 52 |
| 2E-019 | Zn_pept | 0.40 | 1 | 55 | 2 | 5 | 93 | 1 | 115 | 1 | 125 | 1 | 1 | 52 |
| 1E-020 | | | 1 | 55 | 1 | 48 | 104 | 1 | 126 | 1 | 138 | 1 | 1 | 52 |
| 5E-011 | PINT | 0.19 | 1 | 55 | 1 | 48 | 11 | 3 | 49 | 1 | 55 | 1 | 1 | 52 |
| 5E-009 | | | 1 | 55 | 1 | 48 | 1 | 5 | 16 | 1 | 9 | 1 | 1 | 52 |
| 7E-013 | LMWPc | 0.098 | 1 | 55 | 1 | 48 | 1 | 5 | 57 | 1 | 63 | 1 | 1 | 52 |
| 1E-014 | | | 1 | 55 | 1 | 48 | 6 | 5 | 5 | 3 | 2 | 3 | 1 | 52 |
| 6E-014 | | | 1 | 55 | 1 | 48 | 6 | 5 | 5 | 3 | 2 | 3 | 1 | 52 |
| 6E-015 | | | 1 | 55 | 1 | 48 | 8 | 4 | 15 | 2 | 8 | 2 | 1 | 52 |
| 6E-015 | | | 1 | 55 | 1 | 48 | 8 | 4 | 15 | 2 | 8 | 2 | 1 | 52 |
| 3E-018 | | | 1 | 55 | 2 | 5 | 14 | 2 | 28 | 1 | 28 | 1 | 1 | 52 |
| 2E-012 | | | 1 | 55 | 1 | 48 | 5 | 5 | 142 | 1 | 154 | 1 | 1 | 52 |
| 6E-017 | | | 1 | 55 | 1 | 48 | 10 | 3 | 22 | 1 | 15 | 1 | 1 | 52 |
| 3E-019 | WNT1 | 0.45 | 1 | 55 | 1 | 48 | 10 | 3 | 23 | 1 | 16 | 1 | 1 | 52 |
| 1E-013 | | | 1 | 55 | 1 | 48 | 3 | 5 | 14 | 2 | 132 | 1 | 1 | 52 |
| 4E-019 | | | 1 | 55 | 2 | 5 | 99 | 1 | 121 | 1 | 133 | 1 | 1 | 52 |
| 3E-008 | | | 1 | 55 | 104 | 1 | 119 | 1 | 145 | 1 | 157 | 1 | 1 | 52 |
| 2E-017 | | | 1 | 55 | 1 | 48 | 120 | 1 | 146 | 1 | 158 | 1 | 1 | 52 |
| 5E-014 | | | 1 | 55 | 1 | 48 | 6 | 5 | 5 | 3 | 2 | 3 | 1 | 52 |
| | | | | | | | | | | | | | | |
| | | | 2 | 5 | 3 | 5 | 2 | 5 | 3 | 3 | 17 | 1 | 2 | 5 |
| | | | 2 | 5 | 3 | 5 | 2 | 5 | 25 | 1 | 19 | 1 | 2 | 5 |
| | CNX | 0.46 | 2 | 5 | 3 | 5 | 2 | 5 | 3 | 3 | 20 | 1 | 2 | 5 |
| | CNX | 0.36 | 2 | 5 | 3 | 5 | 2 | 5 | 3 | 3 | 21 | 1 | 2 | 5 |
| | | | 2 | 5 | 3 | 5 | 2 | 5 | 26 | 1 | 22 | 1 | 2 | 5 |
| | | | | | | | | | | | | | | |
| | | | 3 | 3 | 4 | 3 | 12 | 3 | 11 | 2 | 81 | 1 | 3 | 5 |
| | | | 3 | 3 | 4 | 3 | 12 | 3 | 11 | 2 | 85 | 1 | 3 | 5 |
| 0.17 | | | 3 | 3 | 4 | 3 | 12 | 3 | 110 | 1 | 120 | 1 | 3 | 5 |
| | | | 83 | 1 | 85 | 1 | 98 | 1 | 120 | 1 | 130 | 1 | 3 | 5 |
| | | | 63 | 1 | 63 | 1 | 75 | 1 | 92 | 1 | 102 | 1 | 3 | 5 |
| | | | | | | | | | | | | | | |
| 0.21 | | | 4 | 3 | 5 | 3 | 13 | 3 | 4 | 3 | 1 | 3 | 5 | 3 |
| 0.24 | | | 4 | 3 | 5 | 3 | 13 | 3 | 4 | 3 | 1 | 3 | 5 | 3 |
| | | | | | | | | | | | | | | |
| 0.26 | | | 56 | 1 | 56 | 1 | 68 | 1 | 82 | 1 | 92 | 1 | 14 | 1 |
| | | | | | | | | | | | | | | |
| 2E-023 | SCP | 2E-027 | 46 | 1 | 46 | 1 | 58 | 1 | 68 | 1 | 76 | 1 | 13 | 1 |
| | | | | | | | | | | | | | | |
| 0.42 | KAZAL | 0.003 | 10 | 2 | 10 | 2 | 20 | 2 | 67 | 1 | 75 | 1 | 7 | 2 |
| | KAZAL | 0.009 | 10 | 2 | 10 | 2 | 20 | 2 | 101 | 1 | 111 | 1 | 7 | 2 |
| | | | | | | | | | | | | | | |
| 6E-009 | PhBP | 2E-007 | 9 | 2 | 9 | 2 | 19 | 2 | 10 | 2 | 72 | 1 | 6 | 2 |
| 1E-007 | PhBP | 4E-007 | 9 | 2 | 9 | 2 | 19 | 2 | 10 | 2 | 74 | 1 | 6 | 2 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1E-026 | IPPc | 4E-059 | 11 | 2 | 11 | 2 | 21 | 2 | 97 | 1 | 107 | 1 | 8 | 2 |
| 2E-026 | IPPc | 7E-066 | 11 | 2 | 11 | 2 | 21 | 2 | 128 | 1 | 140 | 1 | 8 | 2 |
| | | | | | | | | | | | | | | |
| 1E-037 | Tryp_SPc | 3E-050 | 15 | 1 | 14 | 1 | 24 | 1 | 20 | 1 | 13 | 1 | 10 | 1 |
| | | | | | | | | | | | | | | |
| 0.053 | | | 6 | 2 | 7 | 2 | 17 | 2 | 8 | 2 | 4 | 2 | 4 | 4 |
| | B_lectin | 0.018 | 8 | 2 | 8 | 2 | 18 | 2 | 9 | 2 | 5 | 2 | 4 | 4 |
| | B_lectin | 0.10 | 8 | 2 | 8 | 2 | 18 | 2 | 9 | 2 | 5 | 2 | 4 | 4 |
| | | | | | | | | | | | | | | |
| 6E-006 | Knot1 | 0.003 | 43 | 1 | 43 | 1 | 55 | 1 | 64 | 1 | 70 | 1 | 12 | 1 |
| 0.11 | LIGANc | 0.35 | 78 | 1 | 80 | 1 | 92 | 1 | 114 | 1 | 124 | 1 | 20 | 1 |
| | | | 66 | 1 | 66 | 1 | 78 | 1 | 96 | 1 | 106 | 1 | 17 | 1 |
| 8E-011 | FN3 | 0.073 | 13 | 1 | 12 | 1 | 22 | 1 | 18 | 1 | 11 | 1 | 9 | 1 |
| | | | | | | | | | | | | | | |
| 0.014 | ZnF_C4 | 0.66 | 67 | 1 | 67 | 1 | 79 | 1 | 98 | 1 | 108 | 1 | 18 | 1 |
| | | | | | | | | | | | | | | |
| 0.020 | | | 68 | 1 | 68 | 1 | 80 | 1 | 99 | 1 | 109 | 1 | 19 | 1 |
| | | | | | | | | | | | | | | |
| | | | 57 | 1 | 57 | 1 | 69 | 1 | 83 | 1 | 93 | 1 | 15 | 1 |
| 0.43 | | | 82 | 1 | 84 | 1 | 97 | 1 | 119 | 1 | 129 | 1 | 22 | 1 |
| 0.22 | SEA | 0.16 | 93 | 1 | 95 | 1 | 110 | 1 | 135 | 1 | 147 | 1 | 24 | 1 |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| 6E-078 | | | 95 | 1 | 97 | 1 | 112 | 1 | 137 | 1 | 149 | 1 | | |
| 5E-016 | H3 | 4E-040 | 16 | 1 | 15 | 1 | 25 | 1 | 21 | 1 | 14 | 1 | | |
| 1E-011 | H15 | 1E-011 | 91 | 1 | 93 | 1 | 108 | 1 | 132 | 1 | 144 | 1 | | |
| | | | | | | | | | | | | | | |
| | | | 31 | 1 | 30 | 1 | 40 | 1 | 44 | 1 | 50 | 1 | | |
| 4E-030 | | | 85 | 1 | 87 | 1 | 101 | 1 | 123 | 1 | 135 | 1 | | |
| 3E-010 | RPOLD | 0.060 | 86 | 1 | 88 | 1 | 102 | 1 | 124 | 1 | 136 | 1 | | |
| 6E-012 | Sm | 7E-012 | 58 | 1 | 58 | 1 | 70 | 1 | 87 | 1 | 97 | 1 | | |
| | | | | | | | | | | | | | | |
| | | | 71 | 1 | 73 | 1 | 85 | 1 | 106 | 1 | 116 | 1 | | |
| 3E-020 | | | 27 | 1 | 26 | 1 | 36 | 1 | 38 | 1 | 43 | 1 | | |
| 1E-027 | Glyco_18 | 0.21 | 65 | 1 | 65 | 1 | 77 | 1 | 94 | 1 | 104 | 1 | | |
| 1E-035 | | | 5 | 2 | 6 | 2 | 15 | 2 | 6 | 2 | 3 | 2 | | |
| 2E-045 | 53EXOc | 0.96 | 22 | 1 | 21 | 1 | 31 | 1 | 33 | 1 | 35 | 1 | | |
| 4E-024 | | | 72 | 1 | 74 | 1 | 86 | 1 | 107 | 1 | 117 | 1 | | |
| 5E-040 | | | 92 | 1 | 94 | 1 | 109 | 1 | 134 | 1 | 146 | 1 | | |
| 1E-018 | ZP | 0.13 | 23 | 1 | 22 | 1 | 32 | 1 | 34 | 1 | 37 | 1 | | |
| 0.011 | | | 42 | 1 | 42 | 1 | 54 | 1 | 63 | 1 | 69 | 1 | | |
| 9E-088 | | | 45 | 1 | 45 | 1 | 57 | 1 | 66 | 1 | 73 | 1 | | |
| 2E-037 | | | 37 | 1 | 36 | 1 | 47 | 1 | 54 | 1 | 60 | 1 | | |
| 2E-009 | | | 12 | 2 | 70 | 1 | 82 | 1 | 102 | 1 | 112 | 1 | | |
| 3E-006 | | | 12 | 2 | 102 | 1 | 117 | 1 | 143 | 1 | 155 | 1 | | |
| 3E-037 | | | 90 | 1 | 92 | 1 | 107 | 1 | 131 | 1 | 143 | 1 | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4E-006 | | | 14 | 1 | 13 | 1 | 23 | 1 | 19 | 1 | 12 | 1 | | |
| 6E-070 | | | 75 | 1 | 77 | 1 | 89 | 1 | 111 | 1 | 121 | 1 | | |
| 4E-036 | | | 33 | 1 | 32 | 1 | 43 | 1 | 48 | 1 | 54 | 1 | | |
| 0.010 | KOW | 0.011 | 62 | 1 | 62 | 1 | 74 | 1 | 91 | 1 | 101 | 1 | | |
| 3E-033 | POL3Bc | 0.69 | 73 | 1 | 75 | 1 | 87 | 1 | 108 | 1 | 118 | 1 | | |
| 1E-011 | POLAc | 0.94 | 48 | 1 | 48 | 1 | 60 | 1 | 70 | 1 | 78 | 1 | | |
| 4E-026 | | | 52 | 1 | 52 | 1 | 64 | 1 | 75 | 1 | 84 | 1 | | |
| 8E-006 | | | 44 | 1 | 44 | 1 | 56 | 1 | 65 | 1 | 71 | 1 | | |
| 1E-018 | CSP | 5E-010 | 24 | 1 | 23 | 1 | 33 | 1 | 35 | 1 | 39 | 1 | | |
| 1E-024 | | | 76 | 1 | 78 | 1 | 90 | 1 | 112 | 1 | 122 | 1 | | |
| 1E-037 | eIF1a | 6E-027 | 98 | 1 | 100 | 1 | 115 | 1 | 140 | 1 | 152 | 1 | | |
| 6E-015 | MIF4G | 0.88 | 36 | 1 | 35 | 1 | 46 | 1 | 53 | 1 | 59 | 1 | | |
| 7E-064 | | | 32 | 1 | 31 | 1 | 42 | 1 | 47 | 1 | 53 | 1 | | |
| | | | | | | | | | | | | | | |
| 3E-048 | | | 87 | 1 | 89 | 1 | 103 | 1 | 125 | 1 | 137 | 1 | 23 | 1 |
| 2E-064 | RAB | 6E-075 | 7 | 2 | 103 | 1 | 118 | 1 | 144 | 1 | 156 | 1 | | |
| 9E-053 | RAB | 8E-052 | 7 | 2 | 40 | 1 | 52 | 1 | 61 | 1 | 67 | 1 | | |
| 2E-025 | | | 28 | 1 | 27 | 1 | 37 | 1 | 39 | 1 | 45 | 1 | | |
| 3E-014 | | | 84 | 1 | 86 | 1 | 100 | 1 | 122 | 1 | 134 | 1 | | |
| | | | | | | | | | | | | | | |
| 4E-085 | | | 61 | 1 | 61 | 1 | 73 | 1 | 90 | 1 | 100 | 1 | 16 | 1 |
| 1E-064 | SH3 | 0.11 | 60 | 1 | 60 | 1 | 72 | 1 | 89 | 1 | 99 | 1 | | |
| 7E-065 | eIF5C | 0.46 | 41 | 1 | 41 | 1 | 53 | 1 | 62 | 1 | 68 | 1 | 11 | 1 |
| 7E-040 | SapB | 0.20 | 54 | 1 | 54 | 1 | 66 | 1 | 77 | 1 | 87 | 1 | | |
| 2E-014 | | | 49 | 1 | 49 | 1 | 61 | 1 | 71 | 1 | 79 | 1 | | |
| 6E-017 | | | 59 | 1 | 59 | 1 | 71 | 1 | 88 | 1 | 98 | 1 | | |
| 0.22 | | | 50 | 1 | 50 | 1 | 62 | 1 | 72 | 1 | 80 | 1 | | |
| 0.33 | | | 64 | 1 | 64 | 1 | 76 | 1 | 93 | 1 | 103 | 1 | | |
| 9E-024 | | | 70 | 1 | 72 | 1 | 84 | 1 | 105 | 1 | 115 | 1 | | |
| 4E-020 | FH | 0.56 | 69 | 1 | 69 | 1 | 81 | 1 | 100 | 1 | 110 | 1 | | |
| 0.45 | TOP4c | 0.22 | 55 | 1 | 55 | 1 | 67 | 1 | 79 | 1 | 89 | 1 | | |
| | | | | | | | | | | | | | | |
| 0.044 | EFh | 0.005 | 47 | 1 | 47 | 1 | 59 | 1 | 69 | 1 | 77 | 1 | | |
| 4E-009 | GGL | 1E-011 | 51 | 1 | 51 | 1 | 63 | 1 | 74 | 1 | 83 | 1 | | |
| 3E-055 | ML | 0.041 | 39 | 1 | 38 | 1 | 50 | 1 | 59 | 1 | 65 | 1 | | |
| | | | | | | | | | | | | | | |
| 5E-037 | WNT1 | 0.014 | 35 | 1 | 34 | 1 | 45 | 1 | 52 | 1 | 58 | 1 | | |
| 2E-005 | | | 77 | 1 | 79 | 1 | 91 | 1 | 113 | 1 | 123 | 1 | | |
| 4E-045 | NUC | 0.11 | 38 | 1 | 37 | 1 | 48 | 1 | 56 | 1 | 62 | 1 | | |
| | | | | | | | | | | | | | | |
| 0.19 | ZnF_C2H | 0.015 | 17 | 1 | 16 | 1 | 26 | 1 | 24 | 1 | 18 | 1 | | |
| 0.84 | RING | 0.030 | 20 | 1 | 19 | 1 | 29 | 1 | 31 | 1 | 31 | 1 | | |
| 1E-057 | | | 74 | 1 | 76 | 1 | 88 | 1 | 109 | 1 | 119 | 1 | | |
| | | | | | | | | | | | | | | |
| 2E-076 | DM4_12 | 0.63 | 99 | 1 | 101 | 1 | 116 | 1 | 141 | 1 | 153 | 1 | | |
| 1E-015 | CH | 1E-018 | 26 | 1 | 25 | 1 | 35 | 1 | 37 | 1 | 42 | 1 | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.16 | | | 34 | 1 | 33 | 1 | 44 | 1 | 50 | 1 | 56 | 1 | | |
| 0.20 | | | 19 | 1 | 18 | 1 | 28 | 1 | 29 | 1 | 29 | 1 | | |
| 0.024 | THY | 2E-004 | 30 | 1 | 29 | 1 | 39 | 1 | 42 | 1 | 48 | 1 | | |
| | | | | | | | | | | | | | | |
| 6E-041 | HhH1 | 0.93 | 89 | 1 | 91 | 1 | 106 | 1 | 129 | 1 | 141 | 1 | | |
| 2E-056 | | | 96 | 1 | 98 | 1 | 113 | 1 | 138 | 1 | 150 | 1 | | |
| | | | | | | | | | | | | | | |
| | | | 18 | 1 | 17 | 1 | 27 | 1 | 27 | 1 | 25 | 1 | | |
| | | | | | | | | | | | | | | |
| 3E-007 | | | 80 | 1 | 82 | 1 | 95 | 1 | 117 | 1 | 127 | 1 | 21 | 1 |
| | | | | | | | | | | | | | | |
| 1E-018 | | | 94 | 1 | 96 | 1 | 111 | 1 | 136 | 1 | 148 | 1 | | |
| 9E-019 | FYVE | 2E-018 | 88 | 1 | 90 | 1 | 105 | 1 | 127 | 1 | 139 | 1 | | |
| 0.11 | | | 25 | 1 | 24 | 1 | 34 | 1 | 36 | 1 | 40 | 1 | | |
| 0.071 | | | 97 | 1 | 99 | 1 | 114 | 1 | 139 | 1 | 151 | 1 | | |
| 1E-011 | | | 53 | 1 | 53 | 1 | 65 | 1 | 76 | 1 | 86 | 1 | | |
| 2E-025 | | | 28 | 1 | 27 | 1 | 37 | 1 | 39 | 1 | 45 | 1 | | |
| | S_TKc | 0.37 | 29 | 1 | 28 | 1 | 38 | 1 | 40 | 1 | 46 | 1 | | |
| 0.19 | CheW | 0.45 | 21 | 1 | 20 | 1 | 30 | 1 | 32 | 1 | 33 | 1 | | |
| 6E-024 | | | 40 | 1 | 39 | 1 | 51 | 1 | 60 | 1 | 66 | 1 | | |
| | | | | | | | | | | | | | | |
| 0.056 | PSN | 0.20 | 81 | 1 | 83 | 1 | 96 | 1 | 118 | 1 | 128 | 1 | | |