Universidade de Brasília
Faculdade de Medicina
Programa de Pós-graduação em Ciências Médicas
Laboratório de Interação Parasito-Hospedeiro

Descrição extensiva dos sialomas de *Rhodnius neglectus* e *Triatoma dimidiata*, vetores da doença de Chagas

Paula Beatriz Santiago

Orientador: Prof. Dr. Jaime de Martins Santana

Co-orientadora: Profa. Dra. Carla Nunes de Araújo

Brasília-DF

2016

Universidade de Brasília
Faculdade de Medicina
Programa de Pós-graduação em Ciências Médicas
Laboratório de Interação Parasito-Hospedeiro

Descrição extensiva dos sialomas de *Rhodnius neglectus* e *Triatoma dimidiata*, vetores da doença de Chagas

Paula Beatriz Santiago

Orientador: Prof. Dr. Jaime de Martins Santana

Co-orientadora: Profa. Dra. Carla Nunes de Araújo

Tese apresentada ao Programa de Pós-graduação em Ciências Médicas da Universidade de Brasília como requisito parcial à obtenção ao grau de Doutor em Ciências Médicas.

Trabalho desenvolvido no Laboratório de Interação Parasito-Hospedeiro, Instituto de Biologia, Universidade de Brasília (UnB), com os seguintes apoios financeiros: Programa de apoio a Núcleos de excelência (PRONEX); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF); Financiadora de Estudos e Projetos (FINEP); UnB.

Dedico este trabalho à minha mãe e amiga, Vera Lúcia, por ser a inspiração e o incentivo do meu desenvolvimento acadêmico, me ensinando, através de seu exemplo, que o conhecimento deve ser buscado durante todos os momentos da vida porque nos traz grandeza e alegria.

AGRADECIMENTOS

Ao meu orientador, o professor Jaime Santana, pela oportunidade, orientação e confiança durante todo o tempo de realização deste trabalho.

À minha co-orientadora, a professora Carla Nunes, que foi tão importante em minha vida acadêmica, pelo exemplo, incentivo, paciência, carinho e amizade.

Às professoras Izabela Bastos e Flávia Motta, pelo apoio diário e incentivo.

Aos colegas Rayner Myr e Tainá Raiol pelo apoio durante as atividades desenvolvidas neste trabalho. E em especial à Teresa Assumpção, não só pelo apoio acadêmico, mas também pela amizade.

Aos amigos de laboratório Raquel, Grazi, Clênia, Camila, Cissé, Tati, Hugo, Yanna, Marta, Brina, pela amizade, carinho e momentos de convivência.

À Adriana Xavier pela gentileza e apoio na rotina do biotério.

À toda minha família, em especial ao meu marido Fernando e filhas Maria Eduarda e Maria Fernanda, pela graça do nosso encontro, pelo nosso amor, pela nossa união e por nossos formidáveis momentos. Aos meus pais, Vera e Carlos (*in memoriam*) pelo amor incondicional, pelo exemplo e incentivo. E aos meus irmãos, Patricia e Ricardo, e ao meu padrasto, Henrique, pela amizade e carinho.

SUMÁRIO

LISTA DE ABREVIATURAS

1.	DTN	Doenças Tropicais Negligenciadas
2.	OMS	Organização Mundial de Saúde
3.	DC	Doença de Chagas
4.	ICS	Iniciativa do Cone Sul
5.	OPAS	Organização Pan Americana de Saúde
6.	vWF	Fator de von Willebrand
7.	ADP	Adenosina difosfato
8.	TXA2	Tromboxano A2
9.	cAMP	AMP cíclico
10.	FT	Fator tecidual
11.	ON	Óxido Nítrico
12.	ATP	Adenosina trifosfato
13.	AMP	Adenosina monofosfato
14.	Pi	Fosfato inorgânico
15.	NGS	Sequenciamento de próxima geração
16.	RNA-seq	Sequenciamento de RNA (RNA-*sequencing*)
17.	cDNA	DNA complementar

LISTA DE FIGURAS

RESUMO

A doença de Chagas é causada pelo tripanossomatídeo *Trypanosoma cruzi*, geralmente transmitido ao hospedeiro vertebrado através do contato com as fezes do inseto vetor infectado durante o repasto sanguíneo. A lesão vascular ocasionada pela picada do triatomíneo hematófago desencadeia no hospedeiro manifestações que poderiam interferir na alimentação, tais como resposta imune, inflamação e hemostasia. No entanto, os insetos hematófagos são capazes de neutralizar essas defesas através de um complexo repertório de moléculas salivares, fundamentais para a manutenção do fluxo sanguíneo e o sucesso do repasto. No intuito de conhecer a complexidade molecular da saliva dos triatomíneos *Rhodnius neglectus* e *Triatoma dimidiata*, esta tese descreve os transcritomas de suas glândulas salivares utilizando sequenciamento de próxima geração. Foram obtidos 5.704 transcritos a partir da biblioteca de cDNA das glândulas salivares de *T. dimidiata* e 5.705 da biblioteca de *R. neglectus*, dos quais 662 e 636, respectivamente, codificam proteínas secretadas. A análise revelou a expressão de famílias gênicas já descritas na saliva de outros hematófagos, como lipocalinas, grupo abundante, antígeno-5 e apirases. Além disso, inibidores de proteases do tipo Kazal e proteases também foram encontrados. Adicionalmente, o conteúdo salivar foi submetido à análise proteômica, sendo identificadas 111 e 73 proteínas nos proteomas de *T. dimidiata* e *R. neglectus*, respectivamente. Os resultados obtidos validaram alguns transcritos putativamente secretados. Este estudo evidenciou que a saliva dos triatomíneos investigados tem característica multifuncional diretamente relacionada a diferentes atividades anti-inflamatórias, anti-hemostáticas e imunomoduladoras. O estudo dos compostos salivares auxilia a compreensão acerca da biologia dos vetores e transmissão de doenças, da evolução dos hematófagos, além de estimular a busca por novos fármacos.

Palavras-chave: Doença de Chagas; *Rhodnius neglectus*; *Triatoma dimidiata*; Proteínas Salivares.

SUMMARY

Chagas disease is caused by *Trypanosoma cruzi* trypanosomatid, usually transmitted to vertebrate host through contact with faeces of infected vector insects during a blood meal. The vascular injury caused by the bite triggers symptoms capable to preventing an association with the host, such as immune response, inflammation and haemostasis. However, hematophagous arthropods are able to neutralize these defenses through a complex repertoire of salivary proteins acting as vasodilators, antiplatelets and anticoagulants. These molecules are essential for the blood flow maintenance and the success of the blood meal. In order to understand the salivary molecular complexity of *Rhodnius neglectus* and *Triatoma dimidiata*, this work describes their salivary glands (SG) transcriptome using next-generation sequencing. It was possible to obtain 5,704 transcripts from *T. dimidiata* and 5,705 from *R. neglectus* SGs, of which 662 and 636 encoded putatively secreted proteins respectively. The analysis revealed the expression of gene families already described in the saliva of other hematophagous insects, such as lipocalin, abundant protein family, antigen-5 and apyrases. Besides, Kazal type protease inhibitors and proteases were also found. In addition, the salivary content was submitted to proteomic analysis, being possible to identify 111 and 73 proteins in *T. dimidiata* and *R. neglectus* proteomes respectively. The results validated some putatively secreted transcripts. This work showed that saliva of the triatomines investigated here possess multi-functional feature directly related to different anti-inflammatory, anti-hemostatic and immunomodulatory activities. Salivary research helps the understanding of vector biology and disease transmission, hematophagous evolution, in addition to stimulating the search for new drugs.

Keywords: Chagas disease; *Rhodnius neglectus*; *Triatoma dimidiata*; Salivary Proteins.

INTRODUÇÃO

Visão Global das Doenças Tropicais Negligenciadas

As doenças tropicais negligenciadas (DTN), nomeadas assim pela Organização Mundial de Saúde (OMS), representam um conjunto de infecções que afetam sobretudo a faixa tropical do planeta. A maioria das DTN é antiga e tem atormentado a humanidade há séculos. Textos históricos, como por exemplo Papiros e Bíblia, apresentam registros que fazem descrições de doenças como a lepra e a esquistossomose, destacando algumas como deformantes e/ou incapacitantes (Hotez *et al.*, 2006). Dengue, leishmaniose, esquistossomose, teníase, lepra, filariose, doença de Chagas, e outras, estão entre as DTN. Hoje, essas doenças estão concentradas em regiões de menor progresso sócio-econômico, em áreas rurais remotas ou em favelas urbanas (WHO, 2010). Embora clinicamente diversificadas, formam um grupo porque estão fortemente associadas às áreas pobres, com prevalência nos países em desenvolvimento na zona tropical e subtropical do planeta, onde tendem a coexistir. A falta de saneamento básico e acesso aos serviços de saúde, as moradias inadequadas e a desnutrição são fatores que aumentam a vulnerabilidade de uma população a essas infecções (Lindoso e Lindoso, 2009).

Aproximadamente 1 bilhão de pessoas, ou seja, um sexto da população mundial, sofre de uma ou mais das DTN. Estimativas indicam que essas doenças matam mais de 500 mil pessoas a cada ano no mundo, além de enfraquecer e incapacitar outras tantas. Essas doenças podem manter homens, mulheres e crianças dentro de uma situação de pobreza uma vez que pacientes crônicos têm sua capacidade de trabalhar, estudar ou cuidar de seus familiares diminuídas. Esses fatores que resultam em impactos negativos na qualidade de vida e na produtividade do trabalhador, mantendo baixa a renda de países em situação de pobreza. Um ciclo social e econômico devastador tanto para o indivíduo quanto para sua família, além de sua comunidade (WHO, 2006).

As DTN são consideradas ‘negligenciadas’ porque em geral não possuem investimentos significativos para o desenvolvimento de vacinas e novas drogas, além disso, os programas de controle são insuficientes (Lindoso e Lindoso, 2009). Por muito tempo foi assim, e as DTN receberam pouca ou nenhuma atenção. Essa foi a visão até a última década, quando a OMS tentou convencer o mundo, particularmente os líderes políticos em países endêmicos e não endêmicos, a investir no combate à essas doenças como forma de impulsionar o desenvolvimento humano e econômico (WHO, 2006). Hoje, os manifestos globais estimularam esforços na busca de financiamentos para

programas de controle das DTN, incluindo tanto o financiamento proveniente de países desenvolvidos quanto aquele proveniente de organizações filantrópicas. Existem ainda, parcerias em programas de nutrição e educação em saúde, além de doações responsáveis por permitir às populações o acesso aos medicamentos (WHO, 2010). Programas de administração de medicamento em massa contra as DTN representam um grande esforço de controle em saúde pública.

O primeiro relatório da OMS sobre o assunto, *First WHO report on neglected tropical diseases* (WHO, 2010) apresenta evidências de que as atividades desenvolvidas para prevenir e controlar as DTN estão produzindo resultados. O documento cita que, como estratégia, cerca de 670 milhões de pessoas em 75 países endêmicos receberam quimioterapia específica até o final de 2008, sugerindo uma queda na transmissão das doenças quando a população recebe tratamento preventivo.

O segundo relatório, *Second WHO report on neglected tropical diseases* (WHO, 2013), descreveu que, de fato, a quimioterapia preventiva é uma estratégia chave para desarmar, muitas vezes em conjunto, muitas das DTN. Contudo, uma cobertura universal de prevenção e de controle das DTN dependeria de outros fatores: um forte e eficiente sistema de saúde, com acesso aos medicamentos essenciais com qualidade e preço acessível; envolvimento de outros setores, como economia, educação, agricultura e ação das zoonoses; saneamento básico e gestão ambiental. Soma de fatores nem sempre disponível nos países em desenvolvimento. Outro importante ponto descrito é a transmissão regular de patógenos dependentes de insetos vetores e hospedeiros intermediários. Neste caso, o acesso a medicamentos por si só não controla a incidência destas doenças, destacando-se que medidas para controlar vetores e seus hospedeiros intermediários devem ser eficientes e adequadas para a redução da transmissão, utilizando-se de medidas preventivas sustentáveis guiadas por constante vigilância epidemiológica e entomológica. Neste contexto, destacam-se: dengue, filariose, leishmanioses, oncocercose e doença de Chagas. As doenças transmitidas por vetores representam cerca de 16% da incidência de doenças infecciosas, de acordo com a OMS.

Certamente ainda há muito a ser feito e é necessário que a colaboração e os programas continuem a fim de obter progresso. Atenção à prevenção e ao controle das DTN é um direito humano fundamental. Segundo a OMS, o controle dessas doenças é viável, resultado que poderia quebrar o ciclo de doença e pobreza, melhorando a

qualidade de vida da população e aumentando a produtividade do trabalhador, o que contribui para o crescimento sócio-econômico.

Panorama da doença de Chagas

O Brasil faz parte do grupo de países atingidos pelas DTN. Leishmaniose, dengue, raiva, esquistossomose, teníase e doença de Chagas (DC) estão entre as doenças de ocorrência no país. O cenário do desenvolvimento do presente trabalho concentra-se da DC e seus vetores, trazendo a necessidade de conhecimentos acerca desta doença. No que se refere à sua trajetória, inicialmente a DC era uma doença enzoótica selvagem, e foi posteriormente transformada em uma antropozoonose quando o homem invadiu ecótopos silvestres e se estabeleceu no espaço físico antes ocupado por florestas. Como consequência houve o deslocamento de vertebados selvagens, fonte primária de alimentação para os insetos vetores, gerando a necessidade de busca de novas fontes alimentares. A solução foi evetualmente encontrada na adaptação ao ambiente antrópico, onde os vetores podem facilmente se acomodar, alimentando-se de qualquer vertebrado disponível como galinhas, cães ou até mesmo o homem (Coura, 2007). Atualmente, a domiciliação de vetores está relacionada ao grau de preservação de seus ecótopos naturais, ao tipo de habitação humana e o possível abrigo, da oferta alimentar e do grau de antropofilia de cada espécie (Forattini, 2006).

Foi em 1909 que o jovem cientista Carlos Chagas levou ao conhecimento mundial uma nova doença, seus vetores, os insetos triatomíneos, e seu agente causal, um protozoário do gênero *Trypanosoma*, denominado por ele *Trypanosoma cruzi*, em homenagem a Oswaldo Cruz. Em 1910, a doença recebeu o nome de seu descobridor e passou a se chamar doença de Chagas (Coura, 2013).

Naquele tempo, como não havia medicamentos eficazes contra a doença nem inseticidas para controle dos vetores, a base preventiva era a melhoria das habitações. Na década de 1960, programas nacionais e regionais de combate à DC tiveram início no Brasil, Argentina, Venezuela, Chile e Uruguai (Dias, 2007). Contudo, espacialmente limitados, não surtiram grandes efeitos (OPAS, 2002). Já nas décadas de 1970 a 1990 foram tomadas ações de melhor cobertura, direcionadas à interrupção da transmissão vetorial pelo emprego de inseticidas. O receio da síndrome da imunodeficiência humana

adquirida e da hepatite B ajudou a controlar também a transmissão via transfusão sanguínea (OPAS, 2002; Dias, 2007).

Na década de 1990 os países endêmicos decidiram partilhar esforços contra a doença. A Iniciativa do Cone Sul (ICS), lançada em 1991, é um programa internacional de controle da DC e dele participam: Argentina, Peru, Chile, Uruguai, Paraguai, Bolívia e Brasil. O objetivo do programa é controlar a transmissão da doença, eliminando o principal vetor doméstico, *Triatoma infestans*, em áreas do Cone Sul, além de controlar a transmissão transfusional, mediante triagem sorológica (Dias, 2007). É importante ressaltar que a ICS não foi o início de ações que culminaram com o controle da doença, mas representa a consolidação de ações regulares que o efetivaram.

Com relação à DC, o primeiro relatório da OMS cita que o risco de transmissão encontra-se reduzido devido aos programas de controle vetorial e à triagem sorológica nos bancos de sangue. O número de infectados nas Américas caiu de 20 milhões em 1981 para aproximadamente 7 ou 8 milhões atualmente (WHO, 2013). A estimativa do Ministério da Saúde é que cerca de dois a três milhões dos infectados se encontram no Brasil (disponível em <http://portalsaude.saude.gov.br/index.php/oministerio/principal/secretarias/svs/doenca-de-chagas>, acesso em 10 de novembro de 2015). Aqui no Brasil, todo caso de doença de Chagas aguda é de notificação obrigatória às autoridades locais de saúde. A figura 1 mostra a distribuição mundial dos casos de infecção por *T.cruzi*.

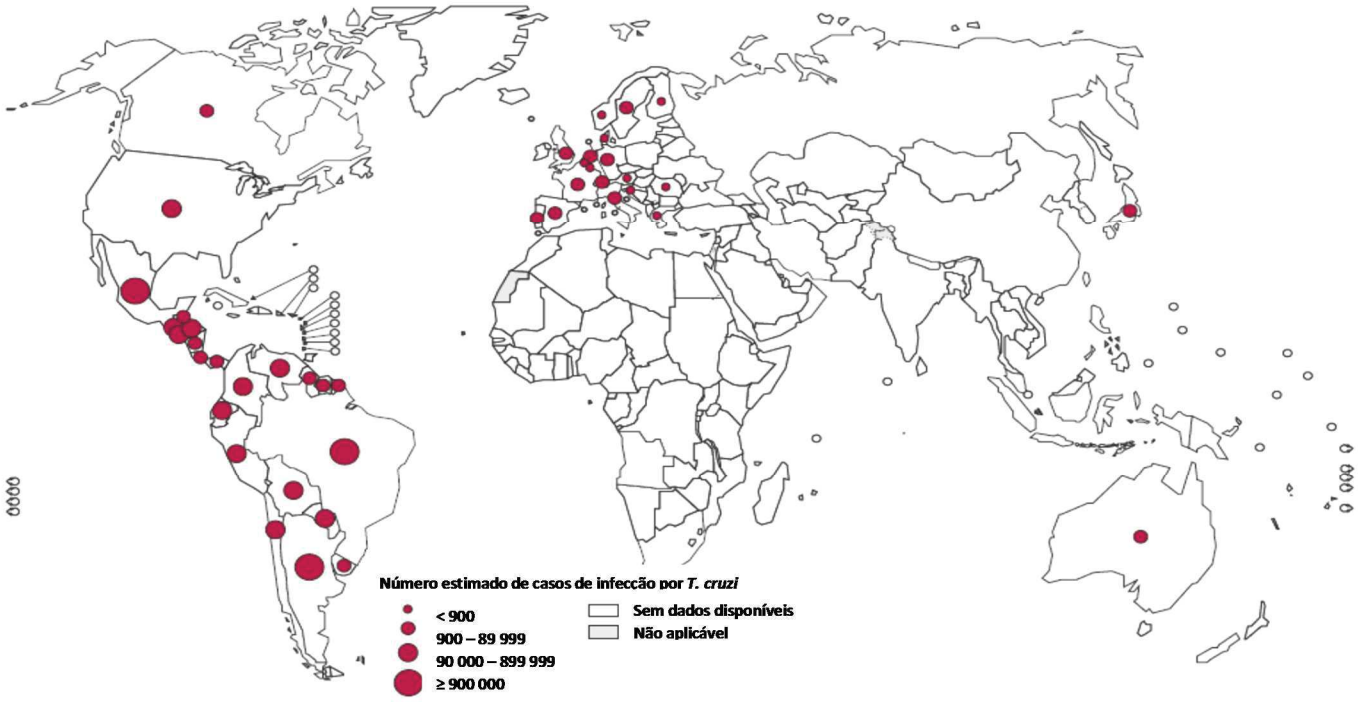


Figura 1: Distribuição de casos de infecção por *Trypanosoma cruzi*, estimativa oficial da OMS entre 2006-2010. Fonte: *Second WHO report on neglected tropical disease* / WHO, 2013 (modificado).

Doença endêmica dos países Latino Americanos, a nação com maior incidência de DC é a Bolívia, onde a soro-prevalência em algumas comunidades chega a 25 % (Salas Clavijo *et al.*, 2012). Contudo, nos últimos anos, a doença tem recebido atenção como um problema emergente na América do Norte e Europa devido às migrações internacionais a partir de áreas endêmicas para as áreas livres de vetores, tornando-se um problema de saúde mundial. A transmissão congênita, aquelas associadas à transfusão sanguínea ou transplantes já foram documentadas em países como EUA, Espanha, Canadá e Suíça (Gascon, Bern e Pinazo, 2010; Schmunis e Yadon, 2010).

Atualmente, no Brasil, a ocorrência de DC aguda tem sido observada principalmente em decorrência da transmissão oral na região Amazônica. Nesta forma de transmissão leva-se em consideração alimentos como sucos de frutas frescas, principalmente açaí e bacaba, contaminados com material infectante proveniente de triatomíneos silvestres (Dias, 2009; Silveira, 2011). O controle da transmissão oral é relativamente difícil dado o caráter aleatório e inesperado de seu acontecimento, assim, medidas gerais de higiene, uma boa seleção alimentar e a pasteurização dos alimentos se

impõem como métodos preventivos nas áreas endêmicas (OPAS, 2009; Dias e Amato Neto, 2011).

Outra possibilidade de transmissão é a invasão esporádica de insetos vetores silvestres no ambiente antrópico, favorecendo a ocorrência humana da DC. A Organização Pan Americana de Saúde (OPAS) faz menção aos limitantes epidemiológicos da doença e cita que como uma enzootia, sua completa erradicação não é possível porque sempre haverá o risco de infecção humana por transmissão silvestre, já que são vastos os reservatórios de animais silvestres, invibializando o esgotamento das fontes de infecção (OPAS, 2002).

Neste contexto, há preocupação com as áreas que demonstram ocorrência de triatomíneos nativos silvestres no peridomicílio, podendo ocorrer esporádica ou progressiva invasão e colonização de moradias humanas, uma situação que exige permanente vigilância epidemiológica. Evidentemente, para que ocorra a propagação da DC em ambiente domiciliar é indispensável que o vetor esteja infectado e colonizando a habitação humana, uma situação condicionada à existência de fatores favoráveis à domiciliação do vetor (Silveira, 2000). Em áreas econômica e socialmente mais pobres, geralmente na zona rural e periurbana, estes insetos se alojam nas frestas de paredes ou nos telhados de palha das casas de má qualidade, onde podem facilmente se esconder durante o dia, e de onde saem à noite para se alimentar. Outras espécies permanecem em ambiente silvestre, nas áreas selvagens remanescentes, e, só ocasionalmente invadem casas, onde podem se alimentar de seres humanos (Lehane, 2005).

Após mais de cem anos de sua descoberta a DC segue incurável. Ainda hoje os medicamentos para o tratamento da fase crônica são limitados, não havendo um tratamento específico ideal contra a doença. Como também não existe uma vacina, o controle do vetor, a triagem sorológica e higiene alimentar ainda são as principais ações preventivas (Dias, 2009). A grande diversidade de reservatórios naturais de parasitos (mamíferos silvestres) e a diversidade de triatomíneos, cujo ambiente natural é continuamente alterado pelo estabelecimento de atividades humanas, devem ser destacadas a fim de desenvolver ações de vigilância em saúde e a sustentabilidade dos programas de controle vetorial para prevenir e controlar a DC (OPAS, 2009; Coura e Borges-Pereira, 2012).

Trypanosoma cruzi e transmissão

O *T. cruzi* é o agente etiológico da doença de Chagas, um protozoário flagelado pertencente à ordem Kinetoplastida e família Trypanosomatidae. Essa espécie apresenta grande diversidade comportamental, sendo encontrada em diferentes hemípteros e animais silvestres e domésticos, além do homem (Brener e Andrade, 2000). O *T. cruzi* é transmitido ao homem através do contato com as fezes infectadas do inseto vetor, um triatomíneo hematófago, quando adaptado ao ambiente doméstico. O inseto defeca no hospedeiro enquanto se alimenta e as fezes contaminadas podem penetrar pela lesão ocasionada pela picada do vetor ou através de mucosas como olhos, nariz e boca (Prata, 2001). Outras vias de transmissão são a transfusional, com sangue contaminado pelo protozoário, ou a congênita, tipos de transmissão que podem ocorrer em áreas não-endêmicas. Acidentes laboratoriais, o manejo de animais infectados e o consumo de alimentos contaminados também podem atuar como vias de transmissão (Prata, 2001).

Na natureza existem três ciclos de transmissão vetorial: o selvagem, o peridoméstico e o doméstico. O ciclo selvagem é enzoótico e é mantido por triatomíneos e animais silvestres contaminados pelo *T. cruzi*. O peridoméstico mantém a ligação entre o ciclo silvestre e o doméstico, e é originado a partir do ciclo selvagem, mantendo a infecção entre os animais domésticos em áreas próximas às habitações humanas. Já o doméstico é mantido na natureza por triatomíneos adaptados ao ambiente humano, transmitindo a infecção aos animais domésticos e eventualmente ao homem (Coura e Dias, 2009).

Ciclo de vida do *T. cruzi*

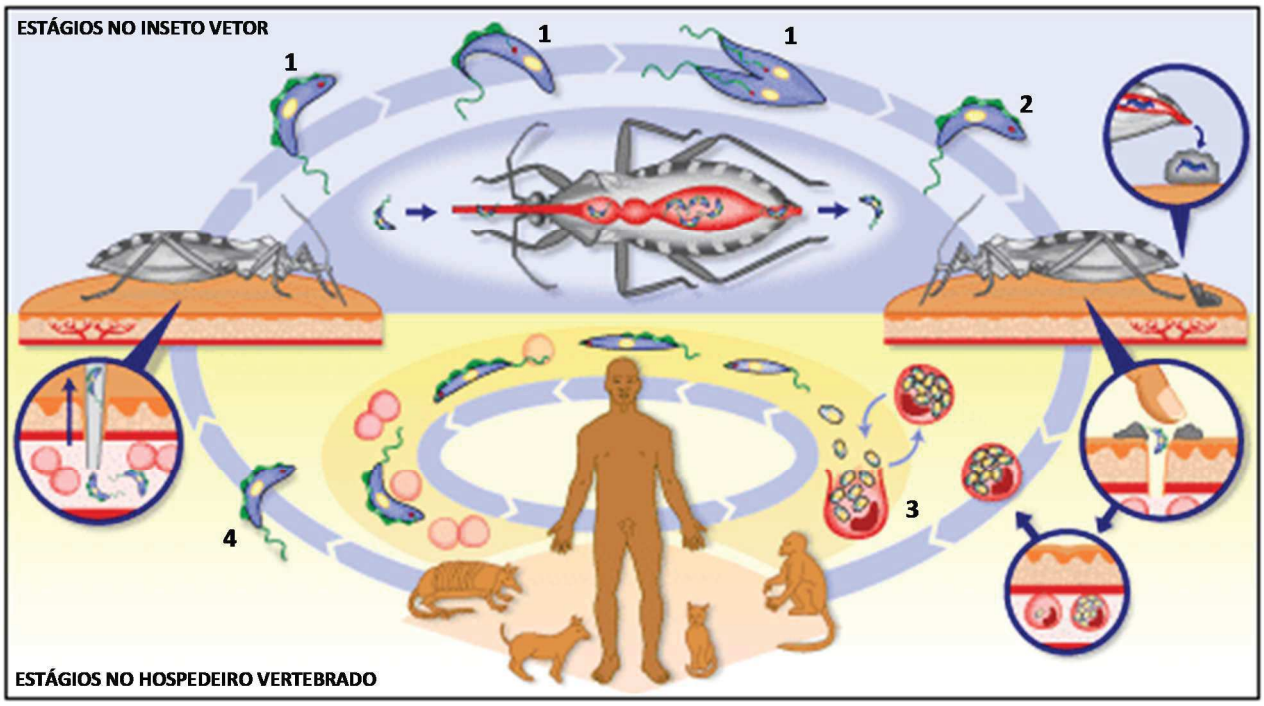


Figura 2. Ciclo de vida do *Trypanosoma cruzi.* Os números indicados na imagem representam as diferentes formas do *T. cruzi*: 1, forma epimastigota proliferativa (presente no intestino médio do vetor). 2, forma tripomastigota metacíclica (presente na porção final do intestino). 3, forma amastigota (presente em tecidos infectados). 4, forma tripomastigota sanguínea. Fonte: *World Health Organization*. Disponível em: http://www.who.int/tdr/diseases-topics/chagas/en/, acesso em 10 de novembro de 2015 (modificado).

O *T. cruzi* apresenta quatro formas de desenvolvimento durante o seu ciclo de vida, o qual se alterna entre o hospedeiro vertebrado e o inseto vetor. A forma epimastigota proliferativa está presente no intestino médio do vetor, e a tripomastigota metacíclica, forma não proliferativa, está presente na porção final de seu intestino. No hospedeiro vertebrado, o parasito apresenta duas formas, tripomastigota e amastigota. A primeira é a forma infectiva presente no sangue e a segunda é a intracelular proliferativa que, ao contrário das outras formas, é arredondada ou oval, apresentando um flagelo curto que não se exterioriza (De Souza, De Carvalho e Barrias, 2010).

O ciclo de vida tem início com o repasto sanguíneo do triatomíneo em um vertebrado contaminado com a forma tripomastigota do *T. cruzi*. No inseto, os parasitos se diferenciam em epimastigotas, multiplicam-se em seu intestino e sofrem diferenciação para a forma infectante tripomastigota metacíclica, a qual é liberada pelas fezes durante

as alimentações sanguíneas subsequentes, gerando o risco de infecção a novos hospedeiros. No hospedeiro vertebrado, os tripomastigotas metacíclicos penetram em suas células, transformam-se em amastigotas, e após crescimento, diferenciam-se em tripomastigotas. Após evasão celular, esta forma entra na circulação sanguínea do hospedeiro vertebrado, e, a partir de novos eventos de repasto sanguíneo, os triatomíneos podem se contaminar e disseminar a infecção (De Souza, De Carvalho e Barrias, 2010) (Figura 2).

Insetos Vetores da Doença de Chagas

Três famílias da ordem hemíptera são representadas por insetos hematófagos. A primeira família é a Cimicidae, a qual possui membros dispersos pelo mundo inteiro, particularmente no hemisfério norte. A maioria parasita morcegos e pássaros, mas duas espécies conhecidas como percevejos, *Cimex lectularius* e *C. hemipterus*, podem se alimentar de sangue humano. Normalmente estes insetos se alimentam à noite, comprometendo o sono do hospedeiro que responde com inflamação e edema. A segunda família é a Polyctenidae, parasitas de morcegos tanto no Novo quanto no Velho mundo, não possuindo importância sócio-econômica (Lehane, 2005).

A terceira família da ordem citada é a Reduviidae e inclui a subfamília Triatominae, a qual é organizada em 5 tribos que abrangem 15 gêneros e mais de 130 espécies, todas caracterizadas pelo hábito hematofágico obrigatório (Schofield e Galvão, 2009). Consomem um grande volume de sangue durante o repasto e, a fim de reduzir o peso, excretam imediatamente grande quantidade de água e sal (Noireau, Diosque e Jansen, 2009).

Esses insetos possuem desenvolvimento hemimetabólico, passando por cinco estágios ninfais antes de atingir a fase adulta. Pelo menos um repasto sanguíneo é necessário para que ocorra mudança de estágio, realizado em uma variedade de vertebrados incluindo o homem. O desenvolvimento do ovo ao inseto adulto dura de 3-6 meses, dependendo da espécie e da freqüência de alimentação, e os insetos podem viver de 6 meses a 2 anos (Beard, 2005). Árvores, troncos ocos e montes de pedras compõem o habitat natural desses insetos, onde podem viver em associação com seus hospedeiros. Contudo, eles pedem eventualmente viver em ambiente humano, onde casas de adobe ou

alvenaria sem reboco possuem fendas que representam o abrigo ideal para o desenvolvimento de uma colônia (Lazzari, Pereira e Lorenzo, 2013). A adaptação ao ambiente humano representa algumas vantagens ao triatomíneo como, por exemplo, refúgio contra predadores, proteção dos eventos climáticos extremos, além de disponibilidade de alimento rico em proteína, fatores que favorecem a reprodução e o aumento da densidade populacional (Noireau, Diosque e Jansen, 2009).

Os triatomíneos são os insetos vetores do *T. cruzi*, o agente etiológico da DC, possuindo diferentes nomes locais, como barbeiro ou chupão. Os triatomíneos com importância para a saúde humana encontram-se na região compreendida desde o sul dos EUA à Argentina e Chile. No total, 22 países são endêmicos para esta enzootia nas Américas. A figura 3, resultado dos estudos de Gourbière e colaboradores, exibe a distribuição geográfica das principais espécies de triatomíneos com relevância epidemiológica (Gourbière *et al.*, 2012).

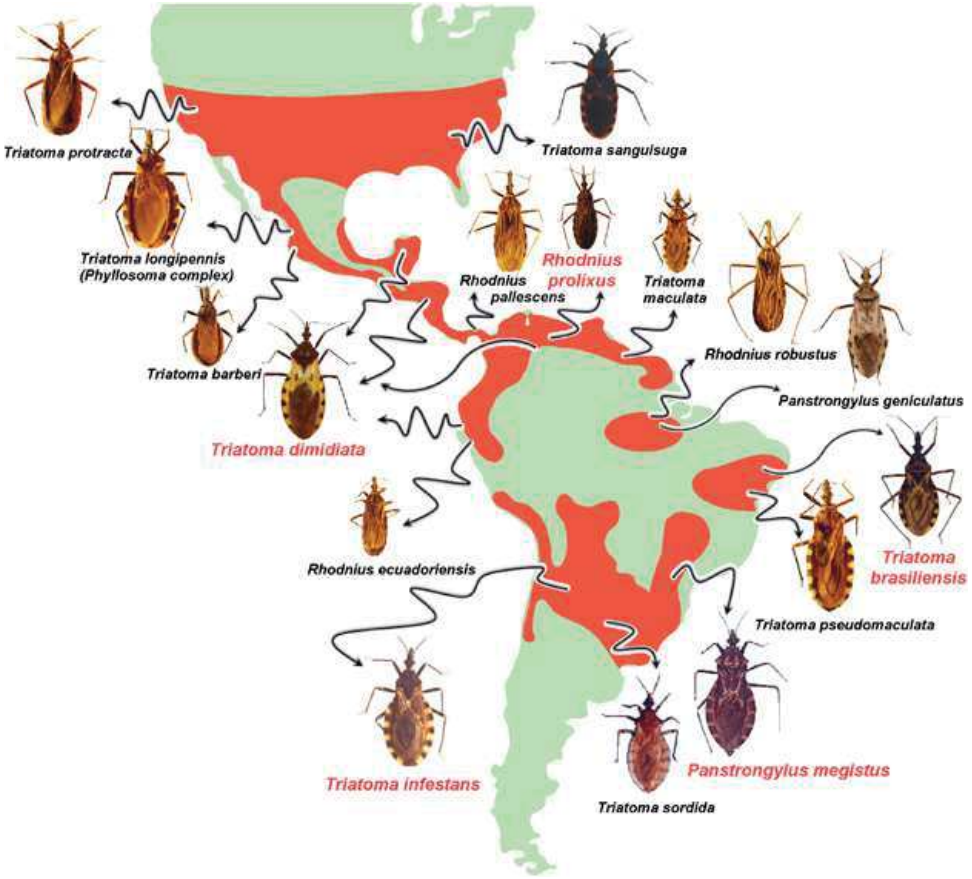


Figura 3: Distribuição dos triatomíneos com importância epidemiológica. Fonte: (Gourbière *et al.*, 2012).

Rhodnius neglectus

Em território brasileiro um estudo investigou a distribuição de 16 espécies de triatomíneos revelando que há associação entre as espécies estudadas e os biomas analisados, destacando-se os gêneros *Panstrongylus, Triatoma* e *Rhodnius*, ocupando o bioma Cerrado, Caatinga e Amazônia respectivamente (Gurgel-Gonçalves *et al.*, 2012).

O triatomíneo *Rhodnius neglectus* (Figura 4) é predominantemente silvestre e está presente no bioma Cerrado. Associa-se a diferentes palmeiras silvestres, sendo a espécie *Mauritia flexuosa* um importante habitat para este e outros triatomíneos (Diotaiuti e Dias, 1984; Gurgel-Gonçalves *et al.*, 2003). Conhecida localmente como 'buriti', cresce nas veredas deste bioma, ambiente caracterizado por pântanos permanentes com palmeiras dispersas e arbustos. Durante a análise da distribuição de vetores descobriu-se que a espécie também pode ser encontrada nos biomas Caatinga e Pantanal, dispersão que pode ocorrer devido à sua associação com as aves (Gurgel-Gonçalves *et al.*, 2012).

Figura 4. **Imagem do triatomíneo *Rhodnius neglectus* e suas glândulas salivares**. Triatomíneo: foto da autora. Glândulas: foto obtida no Laboratório de BioImagem, Embrapa Recursos Genéticos e Biotecnologia.

Estudos revelam que a espécie *R. neglectus* tem um importante papel na manutenção da circulação enzoótica do *T. cruzi* e do *Trypanosoma rangeli* no Cerrado brasileiro ao encontrar indivíduos infectados com os parasitos (Oliveira e da Silva, 2007; Gurgel-Gonçalves *et al.*, 2012). Contudo, apesar de ser um vetor secundário predominantemente silvestre (Barreto-Santana *et al.*, 2011), a espécie foi observada em ambiente intra e peridomiciliar no Estado de Goiás (Oliveira e da Silva, 2007), Mato Grosso do Sul (Almeida e Al, 2008), São Paulo (Rubens, Bonifácio e Wanderley, 1999) e Tocantins (Gurgel-Gonçalves *et al.*, 2008), além do Distrito Federal (Gurgel-Gonçalves *et al.*, 2004). Dada a ocupação eventual de moradias humanas por este triatomíneo, investigações futuras sobre sua competência vetorial podem contribuir para o estabelecimento da importância epidemiológica da espécie.

Triatoma dimidiata

Outro triatomíneo investigado neste projeto é a espécie *Triatoma dimidiata.* Este é um hematófago que apresenta ampla distribuição geográfica: desde o México, passando pela América Central, até Venezuela, Colômbia, Equador e norte do Peru. É atualmente o principal vetor da DC em Guatemala, El Salvador, Nicarágua e Costa Rica, e o segundo vetor mais importante em Honduras e Colômbia (Dorn, Monroy e Curtis, 2007).

A espécie exibe diferentes graus de domesticação ao longo de sua distribuição geográfica, ocupando habitats silvestres, peridomésticos e domésticos. Em alguns locais, como na Guatemala e Nicarágua, o inseto é descrito como um vetor tanto doméstico quanto silvestre (Palma-Guzmán, Rivera e Morales, 1996; Acevedo, Godoy e Schofield, 2000; Nakagawa, Cordón-Rosales*, et al.*, 2003; Nakagawa, Hashimoto*, et al.*, 2003; Weeks *et al.*, 2014).

No México, vetores adaptados ao ambiente doméstico podem ser encontrados, local onde o estado das construções residenciais é o principal fator de risco para a infestação (Dumonteil *et al.*, 2002; Sandoval-Ruiz, Guevara e Ibáñez-Bernal, 2014). Além disso, populações de vetores silvestres foram encontradas na Guatemala e na Costa Rica (Zeledón, Ugalde e Paniagua, 2001; Monroy *et al.*, 2003), enquanto no norte do Peru e no Equador somente triatomíneos domésticos foram descritos (Abad-Franch *et al.*, 2001; Cuba *et al.*, 2002).

Estudos indicam que, embora a população doméstica seja suscetível às ações de controle vetorial, a população silvestre não é atingida, portanto, esses indivíduos são capazes de migrar continuamente ao habitat doméstico e peridoméstico, gerando uma dinâmica de recolonização contínua das moradias humanas, circunstância observada na Guatemala (Monroy *et al.*, 2003) e no México (Dorn, Monroy e Curtis, 2007; Quintal e Polanco, 1977). Este cenário requer a execução de diferentes estratégias para o controle vetorial.

Dorn e colaboradores sugerem que o triatomíneo *T. dimidiata* exibe características distintas em sua abrangência geográfica. Algumas dessas variações são as diferentes taxas de domesticação, de infestação e de infecção pelo *T. cruzi*, que por sua vez são fatores que afetam diretamente a importância epidemiológica do vetor nas diferentes regiões (Dorn, Monroy e Curtis, 2007).

Figura 5. Imagem do triatomíneo *Triatoma dimidiata* e suas glândulas salivares. Triatomíneo: foto da autora. Glândulas: foto obtida no Laboratório de BioImagem, Embrapa Recursos Genéticos e Biotecnologia.

Dinâmica da Interação Vetor-Parasito-Hospedeiro

Uma vez que o vetor localiza o hospedeiro, a pele deve ser penetrada para o acesso ao sangue e, neste intuito, os triatomíneos utilizam uma probóscide longa e flexível para atingir um vaso sanguíneo. A lesão vascular causada pela picada do inseto desencadeia uma rápida resposta fisiológica no hospedeiro contra a invasão externa e a perda sanguínea, sendo determinada por três manifestações: inflamação, resposta imunológica e hemostasia. A inflamação consiste em uma resposta caracterizada pelos cinco sinais cardinais, a saber: edema, calor, rubor, dor e perda de função. A exposição aos antígenos salivares induz o sistema imune a reagir contra essas moléculas. Por fim, a hemostasia visa restaurar a arquitetura vascular e evitar a perda sanguínea. Esses mecanismos são complexos, redundantes e interligados, representando um obstáculo ao sucesso alimentar dos hematófagos (Ribeiro, 1995; Champagne, 2004).

Contudo, apesar da resposta do hospedeiro, um inseto adulto em jejum pode consumir um volume de sangue que corresponde até três vezes o seu peso corporal. Para o sucesso do repasto é fundamental que não cesse o fluxo sanguíneo para o aparelho bucal do hematófago. Dentro desse contexto, o hábito hematofágico evoluiu independentemente entre os artrópodes hematófagos (Francischetti *et al.*, 2009). A resposta do hospedeiro vertebrado é contida através da secreção de diferentes moléculas bioativas presentes na saliva que são injetadas no sítio da picada, estes compostos reconhecem e neutralizam as moléculas envolvidas na inflamação, hemostasia e imunidade, realizando atividades anti-inflamatórias, anti-hemostáticas e imuno-moduladoras (Ribeiro, 1995; Ribeiro e Francischetti, 2003; Champagne, 2005). Essas proteínas são denominadas sialogeninas (do grego *sialo*, saliva; *gen*, origem; *ins*, proteína) (Francischetti, 2010).

O hábito alimentar do vetor proporciona o cenário para a transmissão do *T. cruzi* ao hospedeiro. Já foi demonstrado que, no sítio da picada, a saliva é um poderoso indutor do recrutamento de células mononucleares fagocitárias, e os tripomastigotas metacíclicos eliminados nas fezes por ocasião do repasto podem então interagir com as células do sistema imune e instalar a infecção (Mesquita *et al.*, 2008). Nesse cenário, o patógeno é favorecido pela ação da saliva. Assim, o hábito hematofágico do triatomíneo liga os três elementos da doença: vetor-parasito-hospedeiro.

Hemostasia do hospedeiro

A hemostasia é a resposta fisiológica do hospedeiro contra a perda de sangue. Esse processo começa dentro de segundos após a lesão tecidual e é resultado de uma tríade formada por vasoconstrição, agregação plaquetária e coagulação sanguínea, minimizando a perda sanguínea e promovendo uma restauração da arquitetura vascular normal (Francischetti *et al.*, 2009).

Vasoconstrição

A vasoconstrição é a resposta contrátil da musculatura lisa de um vaso sanguíneo por ocasião de uma lesão física, resultando no estreitamento do vaso. Esse mecanismo pode fechar o lúmen de arteríolas e pequenos vasos, diminuindo o fluxo sanguíneo.

Em condições fisiológicas normais o tônus vascular é mantido, regulando a pressão arterial e a perda de calor a partir da pele através de um equilíbrio endógeno de mediadores vasodilatadores, como o óxido nítrico (ON), e mediadores vasoconstritores, como a norepinefrina e a serotonina. O relaxamento de um vaso sanguíneo pode ser efetuado tanto pela remoção dos vasoconstritores, quanto pela adição de vasodilatadores ao sistema, favorecendo a vasodilatação. A vantagem de usar um equilíbrio entre os vasoconstritores e vasodilatadores para regular o tônus vascular é a velocidade de resposta do tecido: uma pequena mudança em um ou o outro fator pode produzir rapidamente alteração no estado fisiológico do tecido (Kini *et al.*, 2010).

Agregação plaquetária

As plaquetas têm um papel central na hemostasia e, após uma lesão vascular, tem como função primária conter o fluxo sanguíneo através da formação do tampão plaquetário, um evento dinâmico que possui três fases: 1) iniciação: há interação das plaquetas com o endotélio exposto, formando uma monocamada de células plaquetárias ativas; 2) extensão: através da liberação de agonistas, há o recrutamento e ativação adicional de mais plaquetas no sítio da lesão; e 3) estabilização: ocorre a estabilização do tampão plaquetário até que a cicatrização ocorra (Rivera *et al.*, 2009).

A fase de iniciação é marcada pela interação das plaquetas com a matriz extracelular que, uma vez lesionada, expõe fibrinogênio e colágeno. Filamentos de colágeno, por sua vez, possuem afinidade pelo fator de von Willebrand (vWF), e, desta forma, essas duas moléculas encontram-se associadas na matriz extracelular (Rivera *et al.*, 2009).

As plaquetas apresentam diferentes receptores transmembrana que funcionam como receptores de adesão. A interação inicial das plaquetas circulantes com o vaso lesionado é mediada pela atuação do complexo glicoprotéico *(GP) Ib-V-IX*. Individualmente, o receptor plaquetário *GPIb* interage com o vWF imobilizado na superfície do colágeno. Essa interação não promove uma adesão estável, contudo, mantém as plaquetas em contato com a superfície endotelial exposta (Varga-Szabo, Pleines e Nieswandt, 2008). Como resposta inicial também há interação entre o receptor *GPVI* e colágeno. Estes eventos desencadeiam sinais intracelulares plaquetários que

provocam mudança conformacional e ativação da integrina *A2β1* para um estado de alta afinidade, a qual também interage com o colágeno. Contudo, desta vez há uma firme ligação entre as plaquetas e o vaso lesionado (Watson *et al.*, 2005; Varga-Szabo, Pleines e Nieswandt, 2008).

A fase de extensão é marcada pelo recrutamento de plaquetas adicionais que se agregam. A interação entre o receptor plaquetário *GPIb* e o fator vWF promove a liberação de adenosina difosfato (ADP) e tromboxano A2 (TXA2), agonistas que causam a aderência de mais plaquetas no sítio da lesão (Furie e Furie, 2005; Watson *et al.*, 2005). Além disso, a liberação desses agonistas também é responsável por ativar a integrina *α*IIb*β*3, que possui afinidade ao fibrinogênio e ao fator vWF, permitindo a formação de pontes estáveis entre as plaquetas e a constituição de grandes agregados (Furie e Furie, 2005; Watson *et al.*, 2005).

O ADP é um agonista fundamental no processo de extensão, promovendo a ativação de diferentes eventos, como mudança conformacional, secreção de grânulos e agregação através da ativação dos receptores plaquetários $P2Y_1$ e $P2Y_{12}$. O ADP liga-se ao primeiro induzindo a mudança da forma plaquetária, e se liga ao segundo para amplificar a agregação inibindo a produção de AMP cíclico (cAMP) mediado por adenilato ciclase. Já TXA2 é um vasoconstritor e outro importante agonista da ativação e mudança conformacional plaquetária, se ligando ao receptor TP (Varga-Szabo, Pleines e Nieswandt, 2008).

Plaquetas ativas promovem a geração de trombina, o mais potente agonista da agregação, gerando múltiplas respostas mesmo em pequenas doses, como ativação plaquetária, mudança conformacional, secreção de grânulos e agregação. A trombina também ativa os receptores plaquetários ativados por proteases (nomeados PAR-1 e PAR-4 em humanos) desencadeando múltiplas vias de transdução de sinal, e assim modulam a trombose, a coagulação e a inflamação. A resposta plaquetária induzida por trombina é mediada pelo complexo *(GP) Ib-V-IX*. Adicionalmente, plaquetas ativas também secretam moléculas pró-inflamatórias e contribuem para o processo de inflamação local (Coughlin, 2005; Jennings, 2009).

Com a secreção dos grânulos plaquetários há liberação de agonistas considerados mais fracos, como serotonina e epinefrina. A serotonina estimula a mudança da forma plaquetária e intensifica a vasoconstrição, diminuindo o fluxo sanguíneo, enquanto a

epinefrina inibe cAMP sem alterar visivelmente a agregação. Contudo, em conjunto, os dois agonistas potencializam a agregação (Varga-Szabo, Pleines e Nieswandt, 2008).

Na fase de estabilização há um estreito contato entre as plaquetas agregadas que consolidam e estabilizam o tampão, o que, essencialmente, ocorre através da sinalização da integrina αIIbβ3, que uma vez ativa pelo ligante, predominantemente fibrinogênio e vWF como dito anteriormente, desencadeia sinais para a reorganização do citoesqueleto e a estabilização do agregrado plaquetário. Além disso, as plaquetas expressam moléculas de adesão (JAM-A e JAM-C), as quais promovem a interação entre as plaquetas adjacentes, favorecendo a estabilização (Watson *et al.*, 2005).

Assim, a agregação plaquetária é o resultado da ação simultânea de uma série de agonistas, os quais trabalham de maneira redundante em diferentes receptores (Figura 6). A agregação continua até o bloqueio da lesão pela massa de plaquetas agregadas.

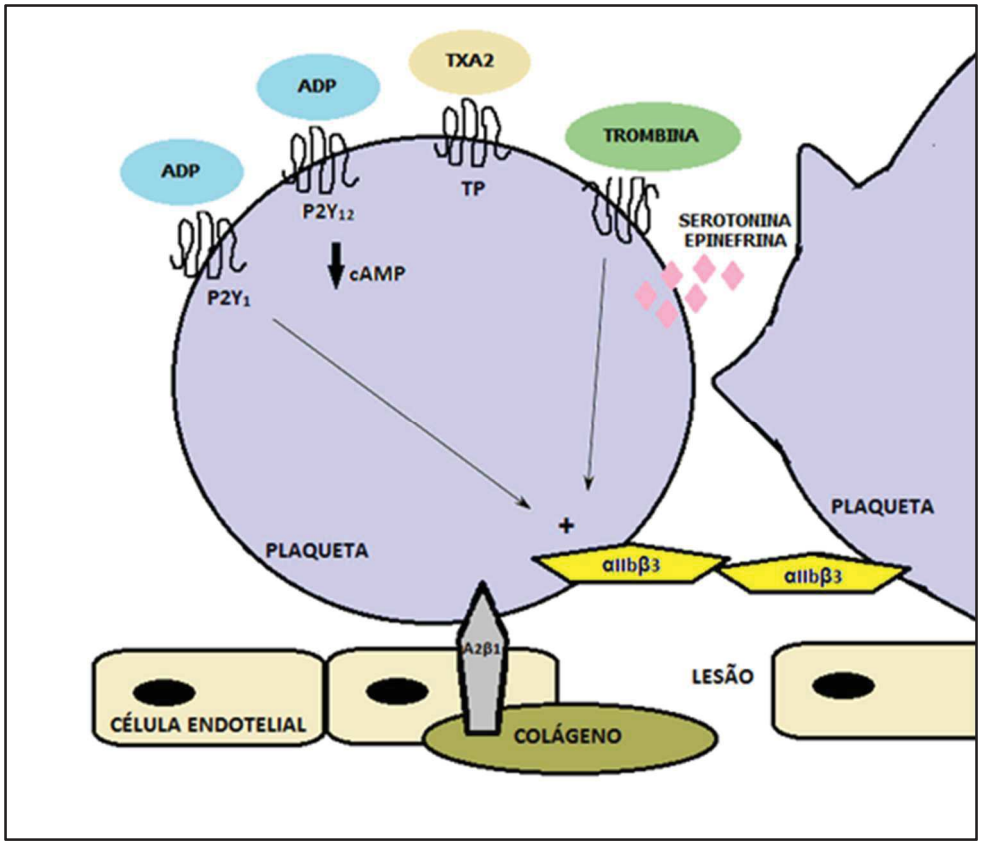


Figura 6. Agonistas da agregação plaquetária. Moléculas agonistas como ADP, colágeno, trombina e TXA2 ativam a agregação plaquetária através de receptores específicos. As plaquetas agregadas obstruem a lesão vascular e evitam o extravasamento sanguíneo. Fonte: elaborado pela autora.

Coagulação sanguínea

A coagulação é um processo complexo e regulado de ativação proteolítica sequencial de diferentes zimogênios presentes no sangue, envolvendo interações entre serino-proteases plasmáticas e seus cofatores, culminando na conversão de fibrinogênio em fibrina e na formação do coágulo sanguíneo.

O início da cascata de coagulação depende da exposição de um componente-chave, o fator tecidual (FT), o qual é expresso por células endoteliais em decorrência de injúrias vasculares ou por monócitos estimulados por infecções e citocinas. O FT é uma glicoproteína de membrana e funciona como receptor para o fator FVII da coagulação, o qual em pequena proporção é encontrado na circulação em sua forma ativa, FVIIa. O FVIIa circulante liga-se ao FT, formando um complexo ativo catalítico, FVIIa/FT, sinal que inicia a cascata de coagulação sanguínea através da ativação dos fatores FIX e FX, os quais devem ser convertidos à FIXa e FXa para continuidade da cascata de coagulação (Bauer *et al.*, 1990; Broze, 1995).

A trombina ativa os fatores FVIIIa e FVa, os quais atuam como importante cofatores para FXa e FIXa nos complexos tenase e protrombinase, respectivamente. O último ativa a conversão de protrombina à trombina na presença do FVa, enquanto a principal função do primeiro complexo é a conversão de FX à FXa na presença de FVIIIa, resultando na amplificação da cascata de coagulação (Monroe, Hoffman e Roberts, 2002; Mann, Butenas e Brummel, 2003; Krishnaswamy, 2005).

Finalmente, a etapa-chave na coagulação sanguínea é a conversão de fibrinogênio em fibrina pela trombina. O coágulo formado consiste em uma malha densa de filamentos de fibrina e engloba os elementos da coagulação no sítio da lesão endotelial, processo crítico para a manutenção da integridade vascular.

Resposta Imune

A resposta do organismo à presença de substâncias estranhas é o que se conhece como resposta imunológica, outro desafio encontrado pelos hematófagos durante o repasto sanguíneo. Tradicionalmente, o sistema imunológico é dividido em inato e adaptativo. O primeiro é a primeira linha de defesa do organismo, sendo composto por componentes solúveis (sistema complemento, citocinas e quimiocinas) e celulares

(neutrófilos, fagócitos e células NK). O segundo, após recorrente exposição ao mesmo antígeno, promove uma resposta imune específica e de memória. É composto por imunidade humoral, mediada por linfócitos B, e imunidade celular, mediada por linfócitos T (Abbas, Lichtman e Pillai, 2008).

O sistema inato está relacionado com a resposta inflamatória, que consiste no recrutamento e ativação de leucócitos e proteínas plasmáticas visando à eliminação do agente infeccioso. Como resultado, ocorre vermelhidão e dor localizados. Durante essa resposta, fagócitos ativos produzem, por exemplo, TNF e IL-1, citocinas mediadoras da inflamação (Abbas, Lichtman e Pillai, 2008).

As principais proteínas plasmáticas da imunidade inata são as moléculas do sistema complemento, as quais reagem umas com as outras através de uma cascata de reações enzimáticas que, além de promover a opsonização do agente infeccioso, induz a uma série de respostas inflamatórias que contribuem para combater a infecção. Existem três vias de ativação do complemento: 1) a via clássica, que é iniciada pela ligação de C1q, a primeira molécula da cascata complemento, a qual detecta o complexo antígeno/anticorpo; 2) a via da lectina, desencadeada pela ligação de lectinas circulantes a polissacarídeos microbianos; 3) a via alternativa, que é iniciada pelo reconhecimento direto de estruturas na superfície microbiana e é regulada pela properdina, molécula que estabiliza a interação (Abbas, Lichtman e Pillai, 2008).

Cada via descrita acima segue uma sequência de reações enzimáticas que culmina com a conversão de C3, a proteína central do sistema complemento, em C3a e C3b. O primeiro peptídeo é uma anafilatoxina responsável por mediar a resposta inflamatória. O segundo é a principal molécula efetora do complemento, que se liga covalentemente aos microorganismos atuando como uma opsonina, assim, marcando-os para a destruição por fagócitos. Além disso, a fração C3b liga-se a outras proteínas do complemento para formar C5 convertase, protease que cliva C5 e inicia os eventos tardios do complemento. Enquanto C5a estimula a entrada de neutrófilos no sítio da infecção, C5b inicia a formação do complexo de ataque à membrana, formando poros nas membranas celulares de patógenos e levando-os à morte (Figura 7) (Abbas, Lichtman e Pillai, 2008).

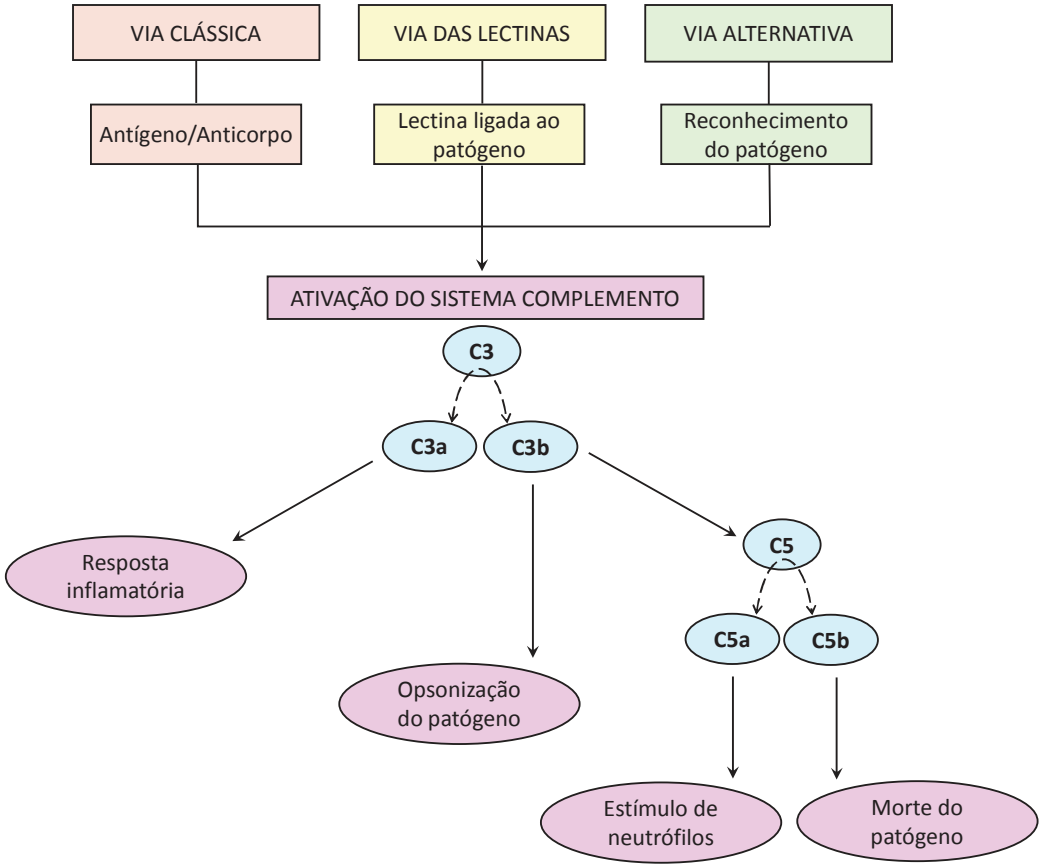


Figura 07. Vias de ativação do sistema complemento. As vias clássica, das lectinas e alternativa são diferentes formas de ativação do sistema complemento, que por meio de uma cascata enzimática, dão origem a mediadores inflamatórios e opsoninas, além da formação de um complexo lítico de membrana que leva o patógeno à morte. Fonte: elaborado pela autora.

Como parte da imunidade adaptativa, linfócitos T e B ativos podem ser observados após uma exposição inicial a determinado antígeno. Neste caso, estas células exibem receptores de antígenos específicos e a indução da resposta imune específica só é possível mediante o reconhecimento do antígeno externo pelo receptor correspondente. Essa interação é consolidada entre células apresentadoras de antígenos, como células dendríticas ou macrófagos do sistema inato, e células T, e entre células B e T (Abbas, Lichtman e Pillai, 2008).

A maturação de células B ocorre na medula óssea, local onde se diferenciam em células que expressam IgM e IgD ligados à membrana, e de onde saem para os órgãos linfóides periféricos para o reconhecimento e resposta aos antígenos estranhos. Já as

células T são produzidas na medula óssea e amadurecem no timo e, em seguida, as células T maduras migram para os tecidos linfóides periféricos ou permanecem na circulação sanguínea. Existem duas populações principais de células T, reconhecidas de acordo com a molécula de superfície que expressam: $CD4^+$ (linfócitos T auxiliares); e $CD8^+$ (linfócitos T citotóxicos), que têm a função de eliminar células infectadas e células tumorais. Ainda, de acordo com o perfil de citocinas secretadas, as células T auxiliares podem ser divididas em subpopulações que apresentam diferentes papéis nas respostas imunes, são as populações Th1 e Th2. A primeira auxilia macrófagos a realizarem fagocitose, enquanto a segunda é associada à resposta humoral, auxiliando a ativação e diferenciação de células B (Abbas, Lichtman e Pillai, 2008).

Propriedades farmacológicas da saliva

Como comentado anteriormente, as respostas fisiológicas do hospedeiro são neutralizadas pelo vetor hematófago, o que garante seu sucesso alimentar através da manutenção do fluxo sanguíneo durante o repasto. Nesse contexto, os triatomíneos apresentam um sofisticado e redundante mecanismo, eles produzem em suas glândulas salivares uma série de moléculas bioativas com função anti-hemostática, incluindo vasodilatadores, anti-agregadores plaquetários e anticoagulantes, além de moléculas imunomoduladoras (Champagne, 2005).

Vasodilatadores

Na saliva dos hematófagos, os vasodilatadores tem função de neutralizar a vasoconstrição local e aumentar o fluxo de sangue, acelerando a aquisição sanguínea e diminuindo o tempo de contato com o hospedeiro no decorrer do repasto.

Dentre as estratégias que culminam com vasodilatação é possível encontrar no triatomíneo *Rodnius prolixus* uma heme proteína denominada *nitroforina*, responsável por estocar e transportar óxido nítrico (ON) nas glândulas salivares. No momento em que o ON é liberado na lesão, liga-se à adenilato ciclase solúvel promovendo o relaxamento da musculatura vascular (Andersen e Montfort, 2000). A coloração avermelhada das glândulas salivares desta espécie é devido à presença abundante dessas moléculas na

saliva. Um mecanismo similar é encontrado na espécie de percevejo *Cimex lectularius* (Valenzuela, Walker e Ribeiro, 1995).

Outra proteína salivar vasodilatadora, denominada *maxadilan*, foi descrita no flebotomíneo *Lutzomia longipalpis* e sugerida como responsável pelo eritema provocado pela picada do mosquito transmissor da leishmaniose (Lerner *et al.*, 1991). Já a espécie *Aedes aegypti* apresenta vasodilatadores pertencentes à família taquiquinina denomindados *sialokinina I* e *II*, os quais mimetizam uma proteína endógena estimulando a produção de ON ao se ligar a um receptor do hospedeiro (Champagne e Ribeiro, 1994).

Com estrutura da família Kazal de inibidores de proteases, um peptídeo denominado *vasotab*, demonstrou atividade vasodilatadora *in vivo* através do bloqueio de canais de Ca^{2+} e consequentemente diminuiu a força contráctil dos vasos (Xu *et al.*, 2008; Ma *et al.*, 2009). Já as espécies de carrapatos *Amblyomma americanum* e *Ixodes scapularis*, da família Ixodidae, durante seu longo tempo de alimentação secretam progressivamente prostaglandinas e prostaciclinas, potentes vasodilatadores que também podem atuar na inibição da agregação plaquetária (Ribeiro *et al.*, 1992; Bowman, Dillwith e Sauer, 1996).

O sequestro de vasoconstritores endógenos é outra estratégia utilizada. Calvo e colaboradores (Calvo *et al.*, 2006) descobriram que a proteína *D7*, presente na saliva de mosquitos, liga-se eficientemente às aminas biogênicas, mais especificamente, à serotonina, norepinefrina e histamina. Como as aminas biogênicas e os leucotrienos são moléculas pró-inflamatórias, a ação da molécula *D7*, além de facilitar o repasto ao conter a inflamação, também inibe a vasoconstrição local ao sequestrar serotonina, atividade que constitui mais uma estratégia para a vasodilatação local. Este mecanismo demonstrou que uma única molécula pode exercer múltiplos efeitos no hospedeiro. Na saliva do triatomíneo *R. prolixus*, a molécula denominada *ABP* (*amine binding protein*) demonstrou atividade semelhante (Andersen *et al.*, 2003).

Esses mecanismos sugerem que a saliva dos insetos hematófagos ocasiona vasodilatação local através da secreção de moléculas que mimetizam vasodilatadores, explorando as vias de sinalização e os receptores endógenos, ou através da remoção de vasoconstrictores endógenos.

Inibidores de agregação plaquetária

Diferentes mecanismos inibidores de agregação plaquetária já foram descritos em vetores hematófagos. Uma importante estratégia é a hidrólise de ATP (adenosina trifosfato) e ADP em AMP (adenosina monofosfato) e Pi (fosfato inorgânico), atividade realizada por uma sialogenina denominada *apirase* (ATP diphosphohydrolase EC 3.6.1.5) (Ribeiro *et al.*, 1984; Francischetti, 2010).

Essa enzima é ubíqua entre os hematófagos, o que é compreensível visto que o ADP é um importante agonista da agregação plaquetária. Já foi descrita na saliva de mosquitos das espécies *Anopheles gambiae* (Lombardo *et al.*, 2000) e *A. aegypti* (Champagne *et al.*, 1995), em carrapatos dos gêneros *Cimex* (Valenzuela *et al.*, 1998), *Ornithodorus* (Stutzer *et al.*, 2009) e *Rhipicephalus* (Liyou et al., 1999) e em triatomíneos como *T. infestans* (Faudry *et al.*, 2004) e *R. prolixus* (Sarkis, Guimarães e Ribeiro, 1986). Além das espécies *Phlebotomus papatasi* (Valenzuela *et al.*, 2001) e *L. longipalpis* (Charlab *et al.*, 1999).

O triatomíneo *R. prolixus* exibe outra estratégia para inibir agregação induzida por ADP. Uma proteína denominada *RPAI* (*Rhodnius platelet aggregation inhibitor*) inibe agregação plaquetária desencadeada por baixas concentrações do agonista (Francischetti *et al.*, 2000). Altas doses de ADP não são afetadas por esses inibidores. Simulação computacional indicou que a coexistência de *apirase* e *RPAI* em uma mesma secreção salivar é responsável por um conjunto complementar na inibição da agregação plaquetária, uma vez que a degradação do ADP pelas *apirases* ocorre mediante alta concentração da adenosina, a qual após ser hidrolisada, permanece presente em baixa concentração, sendo então eliminada pela molécula *RPAI* (Francischetti, Andersen e Ribeiro, 2002).

A *moubatina* é uma sialogenina da saliva do carrapato *O. moubata* e sua ação inibe agregação plaquetária mediada por colágeno humano. A agregação induzida por outros agonistas, por exemplo TXA2, só é afetada em concentração mais elevadas de *moubatina* (Keller *et al.*, 1993; Waxman e Connolly, 1993). A *TSGP3*, clonada de *O. savignyi,* também demonstrou inibir a agregação induzida por colágeno. O estudo constatou que tanto *moubatina* quanto *TSGP3* apresentam alta afinidade à TXA2, mas não à ADP (Mans e Ribeiro, 2008). O TXA2 é não só um importante agonista da

agregação plaquetária, mas também um importante vasoconstritor, o que comprova a importância deste mecanismo de inibição da hemostasia do hospedeiro.

Já a molécula *palidipina*, presente na saliva do triatomíneo *T. pallidipennis*, é relacionada à inibição de agregação plaquetária induzida por colágeno, mas não por ADP, trombina ou TXA2 (Noeske-Jungblut *et al.*, 1994). Clonada a partir das glândulas salivares do triatomíneo *T. infestans*, a *triplatina* possui alta similaridade de sequência com a *palidipina*, também inibindo agregação plaquetária induzida por colágeno. O estudo sugeriu que a *triplatina* pode se ligar ao receptor *GPVI* e assim inibir a transdução de sinal necessária para a ativação plaquetária por colágeno (Morita *et al.*, 2006). Outra sialogenina que afeta a agregação induzida por colágeno foi encontrada na saliva do carrapato *Haemaphysalis longicornis*, a *longicomina* (Cheng, Wu e Li, 1999).

O bloqueio da interação do colágeno com o fator vWF representa outro mecanismo descrito para inibir a agregação e a secreção de grânulos plaquetários, atividade descrita na saliva de *A. aegypti* e executada por uma sialogenina denominada *aegiptina* (Calvo *et al.*, 2007). Mizurini e colaboradores demonstraram que essa proteína possui mecanismo anti-hemostático duplo, além de inibir a agregação, a proteína também inibe a ativação da coagulação mediada por colágeno (Mizurini, Francischetti e Monteiro, 2013). A *anofelina*, molécula presente na saliva do mosquito *Anopheles stephensi*, exibe grande similaridade de sequência com *aegiptina*, e também inibe agregação induzida por colágeno (Yoshida *et al.*, 2008). Investigação posterior revelou que a proteína possui sequência altamente conservada entre outras espécies de mosquitos (Hayashi *et al.*, 2013).

Adicionalmente, a *ABP*, molécula já descrita anteriormente, além de possuir atividade vasodilatadora, também inibe agregação plaquetária induzida por uma combinação de baixa concentração de ADP e serotonina ou epinefrina, ou baixa concentração de colágeno junto à serotonina ou epinefrina (Andersen *et al.*, 2003). A família de proteínas abundantemente expressa na saliva de mosquitos, *D7*, também inibe agregação plaquetária através da ligação à serotonina e norepinefrina (Calvo *et al.*, 2006).

Outro mecanismo comum na saliva de carrapatos é a atividade antagonista às integrinas. O bloqueio da integrina αIIbβ3 realizado pelas desintegrinas previne a ligação ao fibrinogênio e consequente formação do coágulo. A desintegrina *disagregina*, obtida a partir da saliva do carrapato *O. moubata*, inibe a agregação induzida não só por ADP, mas também por colágeno, epinefrina e trombina, ao se ligar ao receptor plaquetário de

fibrinogênio. Além disso, a molécula se liga à integrina αIIbβ3 de plaquetas ativas, confirmando sua atividade desintegrina (Karczewski, Endris e Connolly, 1994; Karczewski *et al.*, 1995). Da mesma maneira, na espécie *Dermacentor variabilis*, a desintegrina *variabilina* (Wang et al., 1996), e na espécie *O. savigni*, a *savignygrina* (Mans, Louw e Neitz, 2002), impede a ligação da intergrina αIIbβ3 ao fibrinogênio. Também na mesma espécie *O. moubata*, a sialogenina TAI (*tick adhesion inhibitor*) inibiu não só a adesão plaquetária ao colágeno, mas também a adesão de células endoteliais ao colágeno, além disso, ensaio com anticorpo sugeriu que a molécula é antagonista da integrina $\alpha 2\beta 1$ (Karczewski *et al.*, 1995).

Enfim, diversos mecanismos de inibição da agregação plaquetária já foram descritos. Agonistas sintetizados pelo hospedeiro durante resposta hemostática, tais como ADP, colágeno, TXA2 e serotonina, compõem o rol de alvos para as sialogeninas (Figura 8). A atividade redundante e sinergística das proteínas salivares é fundamental, visto que a agregação plaquetária é fator chave do processo de hemostasia. Pesquisas futuras podem descobrir outros alvos das sialogeninas inibidoras da agregação, enriquecendo o estudo a respeito da saliva de insetos hematófagos.

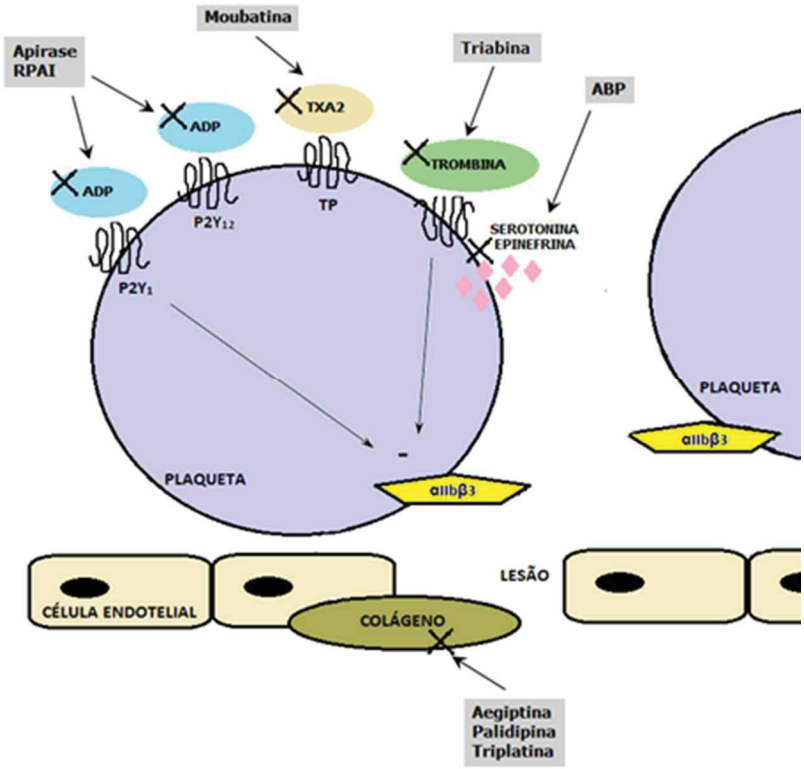


Figura 8. Inibidores da agregação plaquetária. Moléculas salivares, em cinza, evitam a agregação plaquetária inibindo agonistas distintos, o que mantêm o extravasamento sanguíneo na lesão vascular. Fonte: elaborado pela autora.

Inibidores da cascata de coagulação sanguínea

Outro desafio importante para o sucesso do hábito hematofágico é o bloqueio da cascata de coagulação sanguínea. A trombina, além de ser um estímulo para a agregação plaquetária, possui um importante papel na manutenção da hemostasia como um todo, pois cliva o fibrinogênio em fibrina, formando coágulos insolúveis. Consequentemente, a inibição da trombina é uma estratégia atrativa e necessária, uma vez que previne a formação de fibrina e a formação do coágulo.

Inibidores de proteases do tipo Kunitz em carrapatos foram descritos tanto na espécie *Boophilus microplus*, com a molécula *boofilina* (Macedo-Ribeiro *et al.*, 2008), quanto em *O. moubata*, com o exemplo *ornitodorina* (Van De Locht *et al.*, 1996). Ambos inibem o sítio ativo da trombina através de seus domínios Kunitz. Outra classe de inibidores de proteases também merece destaque neste assunto, é a do tipo Kazal. A molécula *rodniina*, encontrada na saliva do triatomíneo *R. prolixus* é um exemplo (Friedrich *et al.*, 1993). Estudo de cristalização revelou que os dois domínios da molécula se ligam em diferentes locais na molécula de trombina, enquanto o N-terminal interage com o sítio ativo, o C-terminal se liga ao exosítio-I (Van De Locht *et al.*, 1995).

Um eficiente inibidor de trombina também foi encontrado na espécie de carrapato *A. variegatum.* Ele é denominado *variegina*, um dos menores inibidores de trombina encontrados na natureza, com somente 32 resíduos de aminoácidos, dos quais 8-14 ligam-se ao sítio ativo da trombina, enquanto os resíduos 15-32 ligam-se ao exosítio-I (Koh *et al.*, 2007).

Outra proteína inibidora da coagulação sanguínea já descrita faz parte da família das lipocalinas, um grande grupo de moléculas que apresenta estrutura tridimensional conservada apesar da diversidade de sequência (Flower, 1996). No triatomíneo *T. pallidipenis* foi descrito um inibidor de trombina denominado *triabina* e, de maneira oposta aos inibidores de trombina descritos acima, a *triabina* liga-se apenas ao exosítio-I (Noeske-Jungblut *et al.*, 1995).

A investigação da composição salivar dos hematófagos tem crescido com o avanço da ciência, e dentro dessa perspectiva, há um bom número de biomoléculas salivares anti-hemostáticas já descritas. O papel individual de algumas dessas moléculas para o hábito hematofágico e para a transmissão de doenças vem sendo cientificamente investigado, contudo, devido à abundância e diversidade, muitas ainda permanecem inexploradas.

Imunomoduladores

Em decorrência da picada de um vetor, o hospedeiro vertebrado responde não só com hemostasia, mas também reage à introdução dos antígenos salivares no sítio da lesão tecidual através de seu sistema imune. Assim, além de apresentar uma variedade de moléculas capazes de neutralizar/modular a resposta hemostática do hospedeiro, a saliva dos hematófagos também possui compostos que apresentam propriedades imunomoduladoras. Neste caso, durante o repasto, diferentes mecanismos são utilizados a fim de neutralizar e/ou modular a reposta imune inata e adaptativa do hospedeiro (Gillespie, Mbow e Titus, 2000; Schoeler e Wikel, 2001). O perfil de resposta imune do hospedeiro é dependente da duração e intensidade de exposição à saliva do hematófago. Por isso, a modulação da resposta imune do hospedeiro pode ser vista como um benefício para o vetor, especialmente, nos casos em que possui uma relação de contato a longo prazo com o hospedeiro, ocorrendo ativação da resposta imune adaptativa, com resposta de células T e B, e consequente produção de anticorpos.

O sistema complemento possui papel fundamental na imunidade inata contra agentes externos (Dunkelberger e Song, 2010). Assim, neutralizar/modular essa resposta é uma estratégia importante para a transmissão de patógenos causadores de doenças transmitidas por vetores hematófagos. Mecanismo de inibição da via alternativa do sistema complemento foi primeiramente observado na saliva de carrapatos Ixodes (Ribeiro, 1987a). E então, posteriormente, vários reguladores do sistema complemento foram descritos principalmente na saliva de carrapatos. *Isac* é uma molécula anti-complemento encontrada na saliva de *I. scapularis* que regula a atividade de C3b, regulando portanto a via alternativa (Valenzuela *et al.*, 2000). Outro regulador do sistema complemento identificado na mesma espécie de carrapato foi a proteína *Salp 20*, a

molécula inibe a properdina, um regulador positivo da via alternativa que se liga à C3 convertase (Tyson *et al.*, 2007 ; Tyson, Elkins e De Silva, 2008). Já na saliva do carrapato mole *O. moubata* foi relatada uma proteína da família das lipocalinas, denominada *OmCI*, que é capaz de inibir tanto a via alternativa, quanto a clássica, através da inibição da C5 convertase. Homólogos à *OmCI*, presentes na saliva de *O. savignyi*, denominados *TSGP2* e *TSGP3*, possuem atividade similar (Mans e Ribeiro, 2008).

Também já foi demonstrado que os extratos salivares dos flebotomíneos *L. longipalpis* e *L. migonei* possuem atividade anti-complemento. Enquanto *L. longipalpis* foi capaz de inibir tanto a via clássica quanto a alternativa, *L. migonei* inibiu apenas a primeira. Similarmente, o mesmo trabalho demonstrou que uma atividade anti-complemento pode estar presente na saliva dos triatomíneos *Panstrongylus megistus*, *T. brasiliensis* e *R. prolixus*, mas não na saliva do mosquito *A. aegyti* e da pulga *Ctenocephalides felis* (Cavalcante, Pereira e Gontijo, 2003).

Mais tarde, a regulação do sistema complemento foi investigada na saliva e nos intestinos dos triatomíneos *T. infestans*, *T. brasiliensis* e *R. prolixus*, observando-se que tanto no extrato salivar quanto no intestino há inibição da deposição de C3b nas vias clássica e alternativa (Barros *et al.*, 2009). Sabendo-se que durante o repasto uma quantidade de saliva é ingerida junto com o sangue, foi sugerido que a função dos inibidores estaria ligada à proteção das células intestinais contra o ataque do sistema complemento após a ingestão de sangue, o que provocaria ruptura das células intestinais.

Outro mecanismo imunomodulador descrito na saliva dos hematófagos é o bloqueio da histamina, um potente indutor da inflamação. Como resultado da reação inflamatória, há aumento da permeabilidade vascular e exsudação de plasma, o que desencadeia a sensação de prurido. Devido à essa sensação incômoda, o hospedeiro pode notar a presença do vetor, o que poderia colocar em risco a vida do mesmo. Moléculas da família das lipocalinas bloqueadoras da histamina vêm sendo descritas na saliva dos carrapatos, e nestes casos, a histamina é sequestrada para dentro da cavidade de ligação da lipocalina (Paesen *et al.*, 1999; Paesen *et al.*, 2000; Mans, 2005). Da mesma forma, o triatomíneo *R. prolixus* possui em sua saliva lipocalinas que apresentam mecanismo de ligação à histamina. Neste caso, as denominadas *nitroforinas*, já discutidas anteriormente,

se ligam tanto ao ON, quanto à histamina, o que otimiza a função da molécula (Ribeiro e Walker, 1994).

Diferente de carrapatos e triatomíneos, em dípteros a atividade anti-histamínica é exercida por proteínas da família *D7*. Essa molécula é abundantemente expressa na saliva e distantemente relacionada com a família OBP, sendo responsável por sequestrar aminas biogênicas, como histamina e serotonina (Calvo *et al.*, 2006).

Assim como ocorre com as moléculas da cascata do sistema complemento, imunoglobulinas também são ingeridas junto com o sangue durante o repasto. Neste contexto, na saliva de diferentes espécies de carrapatos, já foram observadas algumas moléculas capazes de neutralizar a atividade das imunoglobulinas (Wang e Nuttall, 1994). Outros mecanismos imunomoduladores descritos em carrapatos incluem: secreção de prostaglandina E2, atuando na inibição da produção de IL-12 e TNF-α pelas células dendríticas (Sá-Nunes *et al.*, 2007); inibição da proliferação de células T por um inibidor de cisteíno protease, *Sialostatina L* (Sá-Nunes *et al.*, 2009); atividade imunossupressora através da inibição da resposta tanto de linfócitos T, quanto de macrófagos, por uma molécula da família Serpina, *IRIS* (Leboulle *et al.*, 2002); entre outros.

Dentro do contexto de modulação da resposta imune, é importante observar como a infectividade de um parasito pode ser beneficiada a partir das propriedades bioativas das moléculas salivares. Neste cenário, a imunossupressão local em decorrência da atividade de moléculas salivares imunomoduladoras pode maximizar a transmissão e a colonização inicial do parasito no hospedeiro (Ribeiro, 1987b). A esse respeito, alguns estudos sobre os vetores transmissores das leishmanioses foram realizados. Inicialmente foi demonstrado que na presença da saliva de *L. longipalpis* há aumento da infecção por Leishmania *in vivo* (Titus e Ribeiro, 1988). Posteriormente, foi proposto que o aumento da infecção é conduzido por IL-4 (Belkaid *et al.*, 1998) e que o vasodilatador *maxadilan*, encontrado na saliva, estimula a secreção de citocinas Th2 (IL-6 e IL-10) e a redução da secreção de citocinas Th1 (IFN-γ e IL-12) pelas células dendríticas (Brodie *et al.*, 2007; Wheat *et al.*, 2008).

O isolamento e a caracterização de novos imunomoduladores presentes na saliva de vetores transmissores de doenças permanecem como um campo aberto à descoberta de novas moléculas que facilitam a infecção pelo patógeno. Adicionalmente, o estudo dos

imunomoduladores salivares pode ser uma importante ferramenta para o desenvolvimento de fármacos reguladores da resposta imune. Outro aspecto seria o desenvolvimento de vacinas, que, tendo os imunomoduladores como alvos, poderiam atuar tanto contra os vetores hematófagos, quanto contra os patógenos transmitidos pelos mesmos (Gillespie, Mbow e Titus, 2000).

O conjunto das moléculas descritas acima, por toda a seção das propriedades farmacológica da saliva, representam somente uma pequena parte do que já foi descrito na literatura científica, muitas outras moléculas não discutidas aqui fazem parte do conjunto das moléculas salivares de hematófagos. Na Tabela 1 podemos observar alguns exemplos de sialogeninas. Este grupo não para de avançar, e no tópico a seguir entenderemos o que motiva essa crescente descoberta.

Tabela 1: Exemplos de sialogeninas.

HEMATÓFAGO	BIOMOLÉCULA	
Carrapato	Prostaglandina	VASODILATADORES
Triatomíneo	Nitroforina	
Triatomíneo	ABP	
Mosquito	Sialokinina	
Mosquito	D7	
Flebotomíneo	Maxadilan	
Tabanídeo	Vasotab	
Carrapato	Moubatina	INIBIDORES DA AGREGAÇÃO PLAQUETÁRIA
Carrapato	TSGP3	
Carrapato	Longiconina	
Carrapato	Disagregina	
Carrapato	Variabilina	
Carrapato	TAI	
Triatomíneo	Apirase	
Triatomíneo	RPAI	
Triatomíneo	Palidipina	
Triatomíneo	Triplatina	
Triatomíneo	ABP	
Mosquito	Aegiptina	
Mosquito	Anofelina	
Mosquito	D7	
Carrapato	Boofilina	INIBIDORES DA COAGULAÇÃO
Carrapato	Ornitodorina	
Carrapato	Variegina	
Triatomíneo	Rodnina	
Triatomíneo	Triabina	
Carrapato	Isac	IMUNOMODULADORES
Carrapato	Salp20	
Carrapato	OmCI	
Carrapato	TSGP2	
Carrapato	TSGP3	
Carrapato	IRIS	
Carrapato	Sialostatina	
Triatomíneo	Nitroforina	
Mosquito	D7	

Perspectiva da análise de sialomas

Atualmente, a ciência tem disponível o sequenciamento e a espectrometria de massa como principais ferramentas para a descoberta dos novos sialomas. O avanço das técnicas de sequenciamento e o desenvolvimento da bioinformática, orientados para a montagem e anotação de transcritomas, possibilitam um aumento considerável do número de mRNAs sequenciados em investigações científicas. Por sua vez, os recentes progressos

da espectometria de massa alcançam uma maior resolução dos espectros de massa e maior desempenho e eficiência dos equipamentos. Tal progresso científico beneficia as pesquisas referentes à complexidade bioquímica e farmacológica da saliva dos hematófagos, favorecendo as investigações sobre as distintas estratégias anti-hemostáticas das sialogeninas através da obtenção de sequências mais completas e também do aumento do número de moléculas descritas. A descrição do conjunto de mRNAs (transcritoma) e de proteínas (proteoma) presente nas glândulas salivares dos hematófagos forma o **sialoma** (Ribeiro, Andersen*, et al.*, 2004).

A espécie *A. gambiae*, mosquito transmissor da malária teve seu sialoma descrito em 2002 (Francischetti *et al.*, 2002). Outros exemplos de artrópodes com o sialoma já descrito são: *A. aegypti* (Valenzuela, Pham*, et al.*, 2002), *A. stephensi* (Valenzuela *et al.*, 2003), *Culex quinquefasciatus* (Ribeiro, Charlab*, et al.*, 2004), *A. albopictus* (Arcà *et al.*, 2007), *Simulium vittatum* (Andersen *et al*., 2009), os carrapatos *I. scapularis* (Valenzuela, Francischetti*, et al.*, 2002), *O. parkeri* (Francischetti *et al.*, 2008) e *I. ricinus* (Chmelar *et al.*, 2008), além do percevejo *C. lectularius* (Francischetti *et al.*, 2010).

Na literatura científica temos disponível o sialotranscritoma dos triatomíneos *R. Prolixus* (Ribeiro *et al*, 2004) e *Triatoma Rubida* (Ribeiro *et al*, 2012). Além disso, sialomas de diferentes espécies de triatomíneos também já foram investigados, alguns exemplos são: *T. brasiliensis* (Santos *et al.*, 2007), *T. infestans* (Assumpção *et al.*, 2008), *D. maxima* (Assumpção et al., 2011), *T. matogrossensis* (Assumpção *et al.*, 2012) e recentemente, *Panstrongylus megistus* (Ribeiro, Schwarz e Francischetti, 2015). Os resultados destacam a abundância da família de proteínas da família das lipocalinas nos sialomas dos triatomíneos, característica contrastante à ordem dos dípteros, pois esses não possuem lipocalinas salivares descritas. Futuramente, outros sialomas devem ser desvendados com as novas tecnologias, as quais geram um volume grande de dados como resultado, enriquecendo o conhecimento sobre a complexidade salivar dos artrópodes.

Análise de transcritomas utilizando RNA-seq

O método automatizado de Sanger dominou a tecnologia de sequenciamento por quase duas décadas, levando a uma série de realizações no campo científico. Contudo, as limitações desse tipo de sequenciamento e a evidente demanda por tecnologias

revolucionárias que entregassem informações rápidas, baratas e precisas sobre um genoma direcionaram esforços para o desenvolvimento de novos métodos de sequenciamento. O método de Sanger automatizado é então considerado uma tecnologia de "primeira geração", e os novos métodos são conhecidos como *next generation sequencing* (NGS). O grande avanço oferecido pelo NGS é a capacidade de produzir um enorme volume de dados de sequenciamento em uma única execução, com uma resolução muito maior do que aquela disponível com os métodos de Sanger, e com um menor preço. Consequentemente, o advento dos métodos de NGS mudou a maneira como as investigações científicas relacionadas ao sequenciamento são abordadas (Metzker, 2010).

Uma das importantes aplicações do NGS é a técnica conhecida como *RNA-sequencing* (RNA-seq), utilizada para catalogar transcritomas de células e tecidos. Neste procedimento, aos fragmentos de DNA complementares (cDNAs), obtidos a partir do RNA de interesse (total ou poli A+), são adicionados adaptadores a uma ou ambas extremidades (3' e 5'). Cada molécula é diretamente sequenciada utilizando tecnologias de NGS a fim de obter sequências curtas (*reads*) a partir de uma extremidade (sequenciamento *single-end*) ou a partir das duas extremidades (sequenciamento *paired-end*). As *reads* possuem entre 30-400 pb dependendo da plataforma de sequenciamento utilizada. As plataformas 454 (*Roche Life Science*), SOLiD (*Applied Biosystems*) e Illumina (*Illumina*) podem ser aplicadas para obter resultados de sequenciamento utilizando RNA-seq (Wang, Gerstein e Snyder, 2009; Nagalakshmi, Waern e Snyder, 2010).

Após o sequenciamento, com o auxílio da bioinformática, as *reads* obtidas podem ser alinhadas a um genoma de referência ou, se não houver genoma disponível, as mesmas podem ser montadas *de novo* em sequências codificadoras (*contigs*) para a construção em larga escala do transcritoma em estudo, trabalho em que os pesquisadores enfrentam o desafio de montar milhões de *reads* obtidas no sequenciamento (Wang, Gerstein e Snyder, 2009; Nagalakshmi, Waern e Snyder, 2010).

Atualmente o grafo de *De Bruijn* é bastante utilizado como estratégia computacional pelos programas computacionais voltados à montagem de transcritomas. Um grafo é uma rede de vértices conectados por arestas, e neste caso, cada *read* é representada por um vértice, e a sobreposição entre as *reads* é representada por uma seta direcionada, que conecta duas *reads*. No programa de montagem, as *reads* são divididas

em pequenas sequências de comprimento *k*, designadas como *k-mers*. As *k-mers* são então utilizados para formar o grafo (Compeau, Pevzner e Tesler, 2011). *ABySS*, *Velvet*, *SOAPdenovo*, *Trinity* e *Oases* são importantes programas de montagem que utilizam o grafo de *De Bruijn*. A performance da montagem pode ser avaliada por um conjunto de estatísticas, a saber: o número *contigs*, o comprimento N50, a taxa de *reads* mapeadas em *contigs* e a taxa de *contigs* mapeadas em transcritos (Clarke *et al.*, 2013).

Com a montagem dos *contigs*, cada um pode então ser comparado a bancos de dados públicos (tais como NCBI, Swissprot, Gene Ontology; e bancos de dados de proteínas, como CDD, PFAM e SMART) utilizando ferramentas de bioinformática a fim de realizar a classificação funcional de proteínas, identificando sequências motivo e domínios funcionais, além de buscar sinal indicativo da secreção.

Enfim, os recentes avanços das tecnologias de sequenciamento e dos programas computacionais de montagem impulsionaram a construção ou a reconstrução (resequenciamento) de transcritomas inteiros, mesmo na ausência de um genoma de referência. Compreender o transcritoma é essencial para interpretar os elementos funcionais do genoma e revelar os constituintes moleculares de células e tecidos sob determinadas condições de pesquisa. Uma das desvantagens do NGS é a incapacidade de sequenciar fragmentos longos, o que pode afetar o processo de montagem. Assim, um desafio do NGS é o avanço de técnicas de isolamento e de sequenciamento de longas moléculas de DNA (Metzker, 2010).

JUSTIFICATIVA

A diversidade de reservatórios de hospedeiros silvestres e de triatomíneos capazes de transmitir o *T. cruzi* ao homem mantém a doença de Chagas como problema epidemiológico na América Latina. A saliva dos vetores é complexa fonte de moléculas, as quais possibilitam o sucesso do repasto, processo central na transmissão do patógeno ao hospedeiro. Como desafio, fica a determinação da extensão e redundância destes coquetéis salivares. Utilizando RNA-seq é possível obter uma análise completa de um transcritoma salivar, enriquecendo o campo de investigação das GS dos vetores hematófagos.

Conhecer a biologia dos vetores das doenças humanas negligenciadas é essencial para a compreensão do ciclo dessas doenças. Assim, a pesquisa das proteínas salivares enriquece não só o estudo da dinâmica parasito-vetor-hospedeiro e o contexto de transmissão de doenças, mas também estudos acerca da evolução dos hematófagos.

A descrição do sialoma de um vetor hematófago revela a diversidade das proteínas com propriedades farmacológicas presentes na saliva. Desta forma, os estudos relacionados às biomoléculas salivares representam ações promissoras na busca de novos fármacos para o tratamento de distúrbios circulatórios, o que também estimula as atividades realizadas neste trabalho. As tecnologias de sequenciamento por RNA-seq e as análises proteômicas são importantes ferramentas para a investigação dos sialomas desses vetores.

OBJETIVOS

Uma das linhas de pesquisa do Laboratório de Interação Parasito-Hospedeiro, coordenado pelo professor Dr. Jaime Santana, é direcionada à investigação da diversidade molecular e funcional das substâncias farmacologicamente ativas presentes na saliva de triatomíneos vetores da DC.

O presente trabalho visa descrever o sialoma das espécies de triatomíneos *R. neglectus* e *T. dimidiata* com o uso de NGS e espectrometria de massa. Os objetivos específicos do estudo foram:

• Descrever o transcritoma das glândulas salivares do triatomíneo *R. neglectus* utilizando NGS;

• Descrever o transcritoma das glândulas salivares do triatomíneo *T. dimidiata* utilizando NGS;

• Descrever o proteoma do conteúdo salivar do triatomíneo *R. neglectus;*

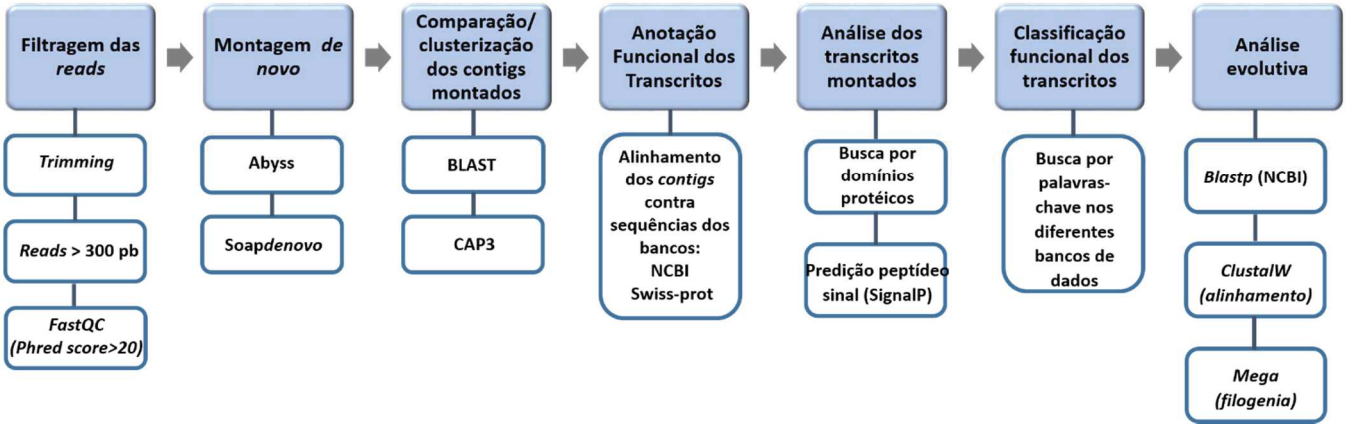
• Descrever o proteoma do conteúdo salivar do triatomíneo *T. dimidiata.*

DESENHO EXPERIMENTAL

Para atingir os objetivos citados acima, as seguintes atividades foram propostas:

❖ Transcritoma:

- Coletar as glândulas salivares;
- Isolar o RNA das glândulas salivares;
- Avaliar a integridade do RNA;
- Sequenciar o cDNA utilizando NGS;
- Submeter os dados obtidos aos programas computacionais, utilizando o seguinte *pipeline*:



❖ Proteoma:

- Coletar o conteúdo das glândulas salivares;
- Realizar protocolo de digestão das proteínas;
- Identificar as proteínas salivares utilizando espectrometria de massa do tipo LC-MS/MS.

RESULTADOS

Os resultados obtidos com as atividades apresentadas neste trabalho proporcionaram a redação de dois manuscritos. O primeiro, "*A Deep Insight into the Sialome of Rhodnius neglectus, a Vector of Chagas Disease*" (manuscrito 1), foi publicado no periódico internacional *PLOS Neglected Tropical Diseases*, fator de impacto 4,446 (2015), classificado como Qualis A1 na área de Medicina I, e analisa o sialoma da espécie *R. neglectus.* O segundo manuscrito, intitulado "*A new insight into the sialome of the salivary glands from Triatoma dimidiata, a vector of Chagas disease*" (manuscrito 2), será submetido para publicação no periódico internacional *Insect Biochemistry and Molecular Biology,* fator de impacto 3,450 (2015), classificado como Qualis A2 na área de Medicina I, e analisa o sialoma da espécie *T. dimidiata.*

MANUSCRITO 1

A Deep Insight into the Sialome of *Rhodnius neglectus*, a Vector of Chagas Disease

Short title: *Rhodnius neglectus* Sialome

Paula Beatriz Santiago[1], Teresa C. F. Assumpção[1,2], Carla Nunes de Araújo[1,3], Izabela Marques Dourado Bastos[1], David Neves[1], Ionizete Garcia da Silva[4], Sébastien Charneau[1], Rayner Myr L. Queiroz[1], Tainá Raiol[1,5], João Victor de Araújo Oliveira[6], Marcelo Valle de Sousa[1], Eric Calvo[2], José M. C. Ribeiro[2] and Jaime M. Santana[1]*

[1] Department of Cell Biology, The University of Brasília, Brasília, Brazil

[2] Vector Biology Section, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Disease, 12735 Twinbrook Parkway, Room 2E32D, Rockville, MD 20852, United States

[3] Ceilândia Faculty, The University of Brasília, Brasília, Brazil

[4] Department of Parasitology, The University of Goiás, Brazil

[5] Instituto Leônidas e Maria Deane - Fiocruz Amazônia, Manaus, AM, Brazil

[6] Department of Computer Science, The University of Brasília, Brasília, Brazil

* Corresponding author: E-mail: jsantana@unb.br

Abstract

Background

Triatomines are hematophagous insects that act as vectors of Chagas disease. *Rhodnius neglectus* is one of these kissing bugs found, contributing to the transmission of this American trypanosomiasis. The saliva of hematophagous arthropods contains bioactive molecules responsible for counteracting host haemostatic, inflammatory, and immune responses.

Methods/Principal Findings

Next Generation Sequencing and Mass Spectrometry-based protein identification were performed to investigate the content of triatomine *R. neglectus* saliva. We deposited 4,230 coding DNA sequences (CDS) in GenBank. A set of 636 CDS of proteins of putative secretory nature was extracted from the assembled reads, 73 of them confirmed by proteomic analysis. The sialome of *R. neglectus* was characterized and serine protease transcripts detected. The presence of ubiquitous protein families was revealed, including lipocalins, serine protease inhibitors, and antigen-5. Metalloproteases, disintegrins, and odorant binding protein families were less abundant.

Conclusions/Significance

The data presented improve our understanding of hematophagous arthropod sialomes, and aid in understanding hematophagy and the complex interplay among vectors and their vertebrate hosts.

Author Summary

Chagas disease is caused by the *Trypanosoma cruzi* protozoan, which is transmitted to vertebrates through the feces of infected triatomines during blood sucking. The vascular injury caused by the bite triggers mechanisms capable of preventing the association with hosts, such as immune response, inflammation and haemostasis. However, hematophagous insects are able to counteract these defenses through a complex repertoire of salivary molecules that have specific targets in the host. Our results show that *R. neglectus* salivary glands express different protein gene families, possessing multi-functional features directly related to different anti-haemostatic activities. For instance, lipocalins are proteins possessing anti-coagulant and vasodilator functions.

Saliva contents have evolved to adapt to blood-feeding habit, ensuring the maintenance of blood flow, the success of the meal, and transmission of diseases.

Introduction

Blood-sucking triatomines (Hemiptera: Reduviidae) feed exclusively on blood in all life stages. They obtain their blood meal from venules or arterioles of their vertebrate hosts. The steps during feeding include piercing of the host skin, followed by a probing period, and finally engorgement [1]. In support of this habit, these arthropods have evolved effective mechanisms to counteract host responses, such as haemostasis, inflammation and immunological reactions. While biting, their salivary glands (SG) release potent pharmacological substances, including vasodilator, anti-inflammatory, antiplatelet, anticlotting and immunomodulatory molecules, to enable the arthropod to obtain a successful blood-meal [2, 3]. These bioactive salivary components represent a promising source of molecules with therapeutic potential for treating circulatory disorders [4, 5].

In the 1990s, multinational control programs against Chagas disease led to a significant reduction of acute cases in many endemic regions of Latin America, mainly through a reduction of domestic vectors [6]. However, factors such as the wide geographical distribution of triatomine species and the availability of different infection reservoirs remain multifactorial obstacles in the control of the disease. Nowadays, there is constant concern regarding the sporadically or progressive (re)invasion and (re)colonization of human dwellings by wild secondary vectors [7, 8]. *Rhodnius neglectus* is found in the Brazilian Savanna (*Cerrado*) in association with different wild palms, playing an important role in the sylvatic maintenance of *T. cruzi* and *Trypanosoma rangeli* [9-11]. In nature, *R. neglectus* feeds mainly on birds and much less on rodents, and rarely on opossum [12]. This species is able to act as a secondary

vector, being observed in both intra and peridomestic environments in five Brazilian states [13-17], a possible result of deforestation and wild ecotope invasion. These anthropogenic environmental changes favor vector dispersion, bridging sylvatic/domestic cycles of the disease.

Sialome studies (from the Greek sialo = saliva) have been developed for many species of bloodsucking insects, which are frequently vectors of human and animal diseases. Sanger automated sequencing technology has been used to investigate the salivary transcriptome for almost two decades. However, Next Generation Sequencing (NGS) is capable of providing much more sequence data in a single run, with a higher resolution than that from the Sanger technique, allowing for deeper analysis of the transcripts. One important application of NGS is RNA sequencing (RNA-seq), used to describe transcriptomes of cells and tissues. Deep sequencing increases the possibilities of finding new biological molecules in the saliva of bloodsucking insects, offering a new array of substances to be further investigated and functionally characterized.

The aim of this report is to catalog the transcripts of *R. neglectus* SGs with probable function in hematophagy using RNAseq and mass spectrometry. This strategy was used to describe the bioactive molecules in triatomine saliva and improve our understanding on the dynamics of the blood-feeding process, vector-host interaction and disease transmission. The data is available at the National Center for Biotechnology Information (NCBI) and can be used in different scientific research projects.

Methods

Insects and Transcriptome Salivary Gland Preparation

R. neglectus triatomines originating from insects collected in 1982 at Itambaracá, in Paraná State, Brazil, were reared in the insectarium at the University of Brasília (Brazil). They were kept at 27±1˚C, a relative humidity of 70-75%, under a 12 h/12 h light/dark cycle. The blood source of these insects was *Gallus gallus domesticus*. The SGs of 5th instar nymphs and adults were dissected at 5, 12, and 24 days post blood meal in cold Trizol reagent (Invitrogen, Carlsbad, CA, USA). A pool of thirty SG pairs was stored at -80˚C prior to RNA extraction.

Salivary gland RNA Isolation, Library Preparation and Sequencing

Total RNA was extracted following the Trizol manufacturer's instructions. RNA integrity and concentration were checked by lab-on-chip analysis using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA). A RNA sample was sent to the Federal District High-Performance Genome Center (DF, Brazil) for Illumina cDNA library construction and next generation sequencing. A Library was prepared with standard protocols using TruSeq RNA kit, v2 (Illumina, San Diego, CA). To generate paired-end reads of 300 nucleotides in length, the sequencing of cDNA libraries was performed on an Illumina MiSeq sequencer (Illumina, USA). One lane of the MiSeq machine was used for sequencing this and another library, distinguished by bar coding. The RNA-seq sequencing generated a total of 12,049,305 reads. The nominal length of the sequences was 301 nt. Following trimming of low quality bases (quality 20 or lower), the average length was 248.07, the median was 301 and L50 was 296 nt. Sequences smaller than 25 nt or with average quality < 20 were rejected.

Bioinformatic analysis

126 Bioinformatic analyses were conducted as previously described [18]. As there
127 was no reference genome to map, the strategy was to perform a *de novo* assembly with
128 Abyss [19] and Soapdenovo Trans [20] assemblers using different kmer (k) values
129 (from 20 to 90). The resulting assemblies were joined by an iterative BLAST and cap3
130 assembler [21]. Sequence contamination between bar-coded libraries were identified
131 and removed when their sequence identities were over 98%. Coding sequences (CDS)
132 were extracted based on the existence of a signal peptide and on similarities to other
133 known proteins [22]. Coding and protein sequences were mapped into a hyperlinked
134 Excel spreadsheet. Reads were mapped into contigs using blastn [23] with a word size
135 of 25, masking homonucleotide decamers and allowing mapping to up to five different
136 CDS if the BLAST results had the same scores. Mapping of the reads was also included
137 in the Excel spreadsheet. CDS were automatically annotated a program written by
138 JMCR that searched a vocabulary of nearly 250 words for matches various databases,
139 including Swissprot, Gene Ontology, KOG, PFAM, and SMART, and a subset of the
140 non-redundant protein database containing proteins from vertebrates (NCBI). Further
141 manual annotation was done as required. Alignment analysis were done with Bioedit
142 software [24] after sequence alignment perfomed using ClustalW [25]. Phylogenetic
143 analysis and statistical neighbor-joining bootstrap tests of the phylogenies were done
144 with Mega package [26]. The sequences used in alignments with *R. neglectus* CDS were
145 obtained from the non-redundant protein database of the NCBI and are represented by
146 six letters followed by the NCBI GI number. The letters derive from the first three
147 letters of the genus and the first three letters of the species name.

148

149 **Data availability**

The raw reads were deposited at the Sequence Read Archive (SRA) in NCBI under bioproject PRJNA292130. A total of 4,230 coding sequences were deposited in DDBJ/EMBL/GenBank through the Transcriptome Shotgun Annotation portal under the accession GDKW00000000.

LC-MS/MS protein identification

The SGs were dissected from 5th instar nymphs and adults at 5, 12 and 24 days post blood meal and carefully punctured at 4ºC. Following centrifugation (16.000 × g, 15 min, 4ºC), the soluble protein fraction from fifteen pairs of SG homogenates was ethanol/acetone precipitated. Resuspended proteins were consecutively alkylated, reduced, digested by trypsin, and subjected to LC-MS/MS analysis as previously described [27]. Briefly, the tryptic peptides were loaded onto a 2 cm fused silica trap column (150 μm inner diameter) packed in-house with reverse phase capillary column ReproSil-Pur C18-AQ 5 μm resin (Dr. Maisch GmbH, Germany) and separated using a DIONEX 3000 nanoUPLC system coupled to an LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific, Waltham, USA). MS1 spectra were recorded in the Orbitrap mass analyzer with 120,000 resolution. After ion fragmentation, MS/MS spectra of the 15 most intense ions were acquired. Raw files were generated and used for protein identification using Proteome Discoverer v.1.3 (Thermo Scientific, Waltham, USA) with in-house SequestHT algorithm for *R. neglectus* SG transcriptome and human keratins, BSA and porcin trypsin. The false discovery rate was less than 1%, with peptide rank of 1 and at least 2 peptides per protein.

Results / Discussion

General description of the sialome of *R. neglectus*

The assembly of *R. neglectus* SG transcriptome enabled the extraction of 5,705 CDS. These CDS mapped a total of over 11 million reads. Following automated and manual annotation, the CDS were classified into putative secreted, housekeeping, unknown, transposable element, and viral product. The CDS of the housekeeping class comprised the largest class (Table 1). They were further characterized into 24 subclasses, according to their possible function, summarized in Table 2.

Table 1: Classification and abundance of coding sequences extracted from the salivary gland transcriptome of *R. neglectus*

Class	No. of CDS	% Total	No. of reads	% Total
Secreted	636	11.15	2,978,414	25.44
Housekeeping	4,739	83.07	8,320,391	71.07
Unknown product	242	4.24	367,035	3.13
Transposable element	86	1.51	41,406	0.36
Viral product	2	0.03	208	0.00
Total	**5,705**	**100**	**11,707,454**	**100**

Table 2: Classification and abundance of coding sequences of putative housekeeping function extracted from the sialotranscriptome of *R. neglectus*

Subclass	No. of CDS	No. of reads	% Total
Unknown conserved	805	1,219,305	14.65
Protein export	271	1,151,192	13.84
Protein synthesis machinery	287	779,314	9.37
Signal transduction	604	725,961	8.73
Protein modification	153	595,681	7.16
Transcription machinery	468	566,268	6.81
Lipid metabolism	217	477,020	5.73
Nucleotide metabolism	87	426,894	5.13
Cytoskeletal protein	190	297,617	3.58
Transporter and Channel	272	293,911	3.53
Protein modification, protease	91	244,360	2.94
Carbohydrate metabolism	150	230,497	2.77
Immunity	78	177,405	2.13
Extracellular matrix	92	175,813	2.11
proteasome	191	157,781	1.90
Energy metabolism	166	143,638	1.73

Amino acid metabolism	79	127,157	1.53
Nuclear Export	26	120,662	1.45
Nuclear Regulation	218	116,946	1.41
Transcription factor	100	99,589	1.20
Detoxification	93	77,717	0.93
Storage	15	56,888	0.68
Intermediary metabolism	57	30,227	0.36
Signal Transduction, apoptosis	28	28,548	0.34
Total	**4739**	**8,320,391**	**100**

Putative secreted proteins

The secreted class was organized in subclasses that include previously known gene families present in hematophagous saliva, such as lipocalin, nitrophorin, antigen-5, as well as gene families not commonly reported in triatomine saliva, such as serine protease and disintegrin (Table 3). The following section describes the putative secreted proteins present in *R. neglectus* sialome, highlighting the remarkable finding of serine proteases in this group.

Table 3: Classification and abundance of coding sequences of putative secretory function extracted from the sialotranscriptome of *R. neglectus*

Subclass	No. of CDS	No. of reads	% Total
Hypothetical secreted protein	198	976,296	32.78
Serine protease	33	820,619	27.55
Conserved secreted protein	89	529,615	17.78
Lipocalin - Triabin	120	471,408	15.83
Disintegrin	2	62,933	2.11
Others	64	52,690	1.77
Lipocalin - Nitrophorin	31	13,737	0.46
Mucin related	11	11,410	0.38
Antigen-5/SCP	8	8,718	0.29
Lipid metabolism	16	6,028	0.20
Major royal jelly protein	2	5,019	0.17
Juvenile hormone related	8	4,328	0.15
Protease inhibitor	14	3,164	0.11
Immunity related	4	2,829	0.09
Insect pheromone-binding	6	2,769	0.09
Protease inhibitor Kazal - type	7	1,993	0.07
OBP	11	1,837	0.06
Toxin	1	1,542	0.05

Nucleotid metabolism	6	735	0.02
5’ nucleotidase	3	477	0.02
Hemolysin-like	1	242	0.01
Metalloprotease	1	35	0.00
Total	**636**	**2,978,424**	**100**

Lipocalins

Lipocalins comprised one of the most abundant groups of transcripts, with 16.29% of putatively secreted reads. These include a large group of extracellular proteins that usually bind to small hydrophobic molecules, cell surface receptors or other proteins. The members of this family have little similarity in peptide sequence, however share a conserved three-dimensional structure, comprised of a single eight-stranded antiparallel β-barrel [28]. In blood-sucking insect and tick saliva the lipocalins are abundantly expressed, but not in Diptera or fleas. In ticks, their function is associated with binding to histamine and serotonin [29]. Triabin and nitrophorin, the two major groups found here, are discussed below.

Lipocalins of the triabin family

First isolated from the saliva of the *Triatoma pallidipennis* kissing bug [30], triabin is a lipocalin-like thrombin inhibitor, which inhibits thrombin-induced platelet aggregation, and prolongs thrombin clotting time through the formation of a noncovalent complex with thrombin at a 1:1 molar ratio. Previous analysis revealed that triabin is a compact one-domain molecule essentially consisting of an eight-stranded β-barrel and inhibits thrombin exclusively via its fibrinogen-recognition exosite [31]. Thrombin is the ultimate serine protease formed during activation of the blood coagulation cascade, which catalyzes the polymerization of fibrinogen to fibrin, the

solid fibrillar component of the blood clot, thereby being a fundamental promoter of blood clotting. Thus, the triabin-like lipocalins may function as thrombin inhibitors in *R. neglectus* saliva. The library analysis shows 120 different CDS from lipocalin family containing the triabin conserved domain, such as triabin, pallidipin, apolipoprotein, procalin and triatin. The alignment of these members with lipocalins already described in triatomines resulted in a phylogram containing different clades (Fig 1). In addition, it is possible to note two divergent clades containing only *R. neglectus* and *Rhodnius prolixus* sequences (RPAI and Apolipoprotein), which may represent additional gene members present in *Rhodnius* spp.. The presence of different clades indicates the expansion of this gene family by gene duplication events, suggesting that, for *R. neglectus*, lipocalins exert a crucial role in success feeding.

Fig 1. Phylogram of lipocalin containing triabin domain from *R. neglectus* SG transcriptome.
Phylogenetic tree derived from the alignment of *R. neglectus* CDS and other triatomine lipocalin sequences as described in Methods section. The bar at the bottom represents 20% amino acid substitution. The colored circles indicate each species whose sequences were used: blue, *R. neglectus* sequences from SG transcriptome; red, *R. prolixus*; yellow, *Triatoma dimidiata*; green, *Triatoma brasiliensis*; dark green, *Triatoma matogrossensis*; dark blue, *T. pallidipennis*; purple, *Triatoma protacta*; magenta, *Triatoma infestans*; gray, *Dipetalogaster maxima*.

Lipocalins of the nitrophorin family

Rhodnius spp. show a characteristic red coloration in their saliva due to the presence of haemoproteins called nitrophorins (NPs). These molecules form a stable complex with nitric oxide (NO), which is sensitive to pH variation, being stabilized by low pH in the lumen of the SGs (pH ~5), and released at neutral pH in the host (pH ~7.5) [32]. The secretion of NO is an efficient way to counteract haemostasis, acting both as a potent vasodilator and as an antagonist of platelet activation. NPs 1-4 can additionally sequester histamine released by host mast cells, reducing inflammation and

immune response [33, 34]. NP 2 inhibits clotting in a mechanism independent of NO or histamine binding, acting as a specific inhibitor of the intrinsic factor X-(FX)-activating complex [35]. As well as reversibly binding to NO or histamine, NP 7 also inhibits prothrombin activation by blocking phospholipid binding sites for the prothrombinase complex on the surfaces of vesicles and activated platelets through binding to phosphatidylserine [36]. The current sialotranscriptome identified 13,737 reads related to the diversity of NPs. The NPs of *R. neglectus* also appear to be a gene family that expanded during evolutionary processes, as inferred by the phylogenetic tree (Fig 2). Notice that there are several sequences homolog to NP1-4 and 7, NPs described in *R. prolixus* saliva.

The mean number of nitrophorins in salivary electrophoretic profiles varies among *Rhodnius* species, with *R. neglectus* showing the fewest. The high polymorphism of NPs may help in the identification of *Rhodnius* species [37]. The lower proportion of nitrophorin content in the saliva compared to those found in the saliva of other *Rhodnius* spp. might not, by itself, explain the reduced feeding performance of *R. neglectus* on mammals. For instance, although *R. neglectus* shows lower amounts of nitrophorins, it feeds more efficiently than *R. robustus* [37]. It is important to note that the exact contribution of each class of saliva molecules on the feeding process is unknown.

Fig 2. Phylogram of lipocalin containing nitrophorin domain from *R. neglectus* SG transcriptome.
Phylogenetic tree derived from the alignment of *R. neglectus* CDS and *R. prolixus* nitrophorin sequences as described in Methods section. The bar at the bottom represents 20% amino acid substitution. The colored circles indicate each species whose sequences were used: blue, *R. neglectus* sequences from SG transcriptome and red, *R. prolixus* sequences from NCBI.

Antigen-5 family

The CAP superfamily members [Cysteine-Rich Secretory Proteins (CRISPS), Antigen 5 (Ag5), and Pathogenesis-Related 1 (Pr-1)] are found in a wide range of organisms, most often as secreted proteins [38]. Ag5, present in the venom of wasps and ants, are considered potent allergens to mammals [39, 40]. This superfamily can also block smooth muscle contraction when present in snake venom [41] and act in the defense response in plants [42]. They have been described in the saliva of some hematophagous, including mosquitoes [43, 44] and sand flies [45]. Among triatomines, Ag5 genes have been reported in the sialotranscriptomes of *R. prolixus* [46], *T. infestans* [47], *D. maxima* [48], *T. matogrossensis* [49] and *Triatoma rubida* [50]. Their functions in blood-feeder saliva remained unexplored for a long time, but a recent report revealed salivary Ag5 of *D. maxima* and *T. infestans* as Cu^{+2}-dependent antioxidant enzymes that inhibit neutrophil oxidative burst and platelet aggregation induced by collagen [51].

The sialotranscriptome analysis revealed eight CDS related to the Ag5 family. The alignment of *R. neglectus* Ag5 with other triatomine Ag5 sequences showed some conserved motifs (S1 Fig). Phylogenetic analysis offers support for the formation of clades I and II comprising triatomine and Diptera sequences, respectively (Fig 3).

Fig 3. Phylogram of Antigen-5 proteins from *R. neglectus* SG transcriptome.
Phylogenetic tree derived from the alignment of *R. neglectus* CDS and other insect antigen-5 sequences as described in Methods section. The bar represents 10% amino acid substitution.

Serine Protease inhibitors

For blood-feeders, targeting components of the coagulation cascade is essential to attenuate the haemostatic response of their hosts. All enzymes participating in this cascade are serine proteases associated with complement activation [52, 53]. The *R. neglectus* sialotranscriptome exhibited a variety of transcripts coding for proteins with

serine protease inhibitory function, comprising 14 CDS and 3,164 reads. Based on their Pfam signature, kazal, pacifastin and serpin families were extracted.

Kazal family

Kazal-type domain-containing proteins are serine protease inhibitors playing important functions in invertebrates, mainly having vasodilation, antimicrobial, and thrombin inbition effects. These protease inhibitors are single or multidomain proteins that share a conserved sequence motif, a distinctive cysteine distribution pattern and highly similar three-dimensional structure [54]. Rhodniin is a kazal-type thrombin inhibitor isolated from *R. prolixus* [55, 56]. Dipetalogastin from *D. maxima* [57], infestin from *T. infestans* [58] and brasiliensin from *T. brasiliensis* [59] are thrombin inhibitors located in the intestines. From the horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) SGs, a vasodilator named vasotab was identified as a member of Kazal-type protease inhibitor family acting through ion channel inhibition and vasodilation [60].

Seven CDS in *R. neglectus* sialotranscriptome possessed the typical sequence of nonclassical Kazal domains characterized by a shorter distance between the first and second cysteine residue, unlike the seven or eight spacer residues found in the classical configuration [55, 57]. The alignment showed a low degree of conserved amino acids, but confirmed the presence of the six cysteine residues responsible for the formation of disulfide bridges (Fig 4). The relative positions of cysteine residues were the same in the compared sequences.

Fig 4. Kazal-type members from *R. neglectus* SG transcriptome.
ClustalW alignment of Kazal-type domain-containing members from *R. neglectus* salivary transcriptome (RN_5563 and RN_549) and other insect kazal-type sequences, identified as described in Methods section. The alignment indicates conserved residues

in black and similar residues in gray background, the six conserved cysteines (boxes) and the blue bar indicates the signal peptide indicative of secretion.

Additionally, one contig was identified as dipetalogastin due to the cysteine residues distribution and the presence of the conserved motif CGXDXXTYXNXC, a distinguishing repeat of Kazal-type inhibitors [57]. This transcript is full length and possesses the signal peptide indicative of secretion. The alignment with other protein sequences with the same features revealed a high degree of conserved amino acids (S2 Fig).

The phylogram of serine protease inhibitor members clearly shows the formation of three clades, with a good bootstrap support, each one representing a different family of serine protease inhibitor discussed above (Fig 5). The CDS RN_21179 is notably distinct from the clades, suggesting the presence of a divergent gene. The different clades may represent sequences differentially expressed sharing the same function regarding haemostasis inhibition.

Fig 5. Phylogram of Serine Protease inhibitors from *R. neglectus* SG transcriptome. Phylogenetic tree derived from the alignment of *R. neglectus* CDS and other insect sequences as described in Methods section. The bar at the bottom represents 20% amino acid substitution.

Pacifastin family

Pacifastin is a family of serine protease inhibitors, mostly multi-domain proteins, first isolated from the plasma of the crayfish *Pacifastacus leniusculus*. The protein is heterodimeric, comprising both a transferrin chain (heavy chain, PHC) and a protease inhibitor chain (light chain, PLC) [61, 62]. Insect pacifastins may have multiple functions, acting as regulators of a wide variety of serine peptidase-dependent processes such as immunity and reproduction [63]. In Hemiptera, two pacifastin-like protease

353 inhibitors from *T. infestans* eggs were functionally characterized, suggesting a role in
354 insect immune response [64]. Here, two CDS are related to pacifastin, RN_17301 and
355 RN_20047, and their alignment with other members of the pacifastin family reveals
356 four conserved domains, containing the cysteine-rich inhibitory pattern of PLC
357 comprised of a triple-stranded antiparallel beta-sheet connected by three disulfide
358 bridges (S3 Fig). This is the first time pacifastin members are identified in triatomine
359 SGs, their function in this organ is still unknown but it might be related to insect
360 immunity.

361

362 Serpin family

363 Serpins are a large family of structurally related proteins found across taxa,
364 showing diverse activities not limited to inhibition of serine proteases [65]. In
365 vertebrates, serpins play crucial control in blood coagulation, fibrinolysis and
366 inflammation. Dysfunction, deficiencies or over-expression of serpins can cause either
367 abnormal bleeding or thrombosis [66]. The function of this protein in saliva of
368 mosquitoes is related to host haemostasis regulation, seeming to act as a potent
369 reversible inhibitor of the host factor Xa [67]. In *Ixodes ricinus* ticks, the molecule was
370 also associated with inhibition of blood coagulation and fibrinolysis of the vertebrate
371 host [68-70]. The consensus three-dimensional fold of serpins is comprised of a bundle
372 of 8-9 α-helices and a β-sandwich composed of three β-sheets [71].

373 Here, four CDS from *R. neglectus* sialotranscriptome were classified as serpins.
374 The phylogram showed four clades with a good bootstrap support (Fig 6). Transcripts
375 RN_9905, RN_20002 and RN_10079 grouped each one in a separate clade while the

fourth CDS, RN_2083, seemed to be a distant divergent gene. In *R. neglectus* saliva this inhibitor might also function in the modulation of coagulation cascade.

Fig 6. Phylogram of serpin proteins from *R. neglectus* SG transcriptome.
Phylogenetic tree derived from the alignment of *R. neglectus* CDS and other insect sequences as described in Methods section. The bar represents 20% amino acid substitution.

Proteases

Metalloprotease

One CDS found in the *R. neglectus* sialotranscriptome is related to the zinc-dependent metalloproteases from the astacin-like metalloproteases, a family of the metzincins superfamily. There are three conserved regions in proteins from this family. The first one is the distinguishing family signature sequence HEXXHXXGXXHE, which is the zinc-binding active site. The second region, RXDRD, is a hydrophilic region, and the third highly conserved region, MXY, is the methionine-containing turn (the Met-turn) [72-74].

This gene family comprises many proteins from diverse species. In the venom of different spider species, there is a common toxin with the ability to hydrolyze fibrinogen and fibronectin [75-77], suggesting a relationship between this proteolytic activity with local hemorrhage, since fibronectin plays a role in platelet aggregation, blood vessel stability and wound healing [78]. Therefore, the biologic function of astacin-like proteases in triatomine saliva could be related to the maintenance of blood flow at the bite site. Astacin domain metalloproteases were already reported in *T. matogrossensis* sialotranscriptome [49]. The CDS RN_21266 is a full-length sequence containing a signal peptide indicative of secretion. Its alignment with other metalloproteases revealed the three conserved motifs of the family (Fig 7, A). The

phylogenetic tree suggests, with a good bootstrap support, that the secreted metalloproteases are closely related proteins (Fig 7, B).

Fig 7. The secreted metalloprotease from *R. neglectus* SG transcriptome.
(A) ClustalW alignment of the secreted metalloprotease from *R. neglectus* salivary transcriptome (RN_21266) and other metalloproteases sequences, identified as described in Methods section. The alignment indicates conserved residues in black and similar residues in gray background. The blue bar indicates the signal peptide indicative of secretion. The boxes limit the family signature sequences showing the determinant residues (black asterisk).
(B) Phylogenetic tree derived from the alignment of *R. neglectus* CDS and other metalloproteases sequences as described in Methods section. The bar represents 10% amino acid substitution.

ADAMTS (ADAM with thrombospondin motifs) / Disintegrins

Two further members of the metzincin metalloprotease superfamily were identified in *R. neglectus* transcriptome and are related to the adamalysin/reprolysin family, which includes ADAM (A Disintegrin And Metalloproteinase domain) and ADAMTS (A Disintegrin And Metalloproteinase with Thrombospondin motifs). ADAMTS is a group of secreted, extracellular and multidomain proteases that have diverse roles in both mammals and invertebrates [79, 80]. They are cysteine-rich molecules that selectively block the function of integrin receptors on the cell membrane surface [81, 82], exhibiting a thrombospondin-like (TS) repeat and a cysteine-rich domain typical of disintegrins [79]. In this family, the third histidine in the family signature sequence containing three zinc ligands is followed by a conserved aspartic acid, HEXXHXXGXXHD. Moreover, it lacks the fifth zinc ligand and the methionine residue of the consensus Met-turn is placed within the sequence V/I-M-A/S [74, 79].

Together with snake venom metalloproteinases (SVMP), ADAM and ADAMTS are a group of versatile molecules in viper venom that affects different elements in haemostasis [83]. The disintegrins can bind to platelets and act as potent inhibitors of

platelet aggregation [84-86]. The molecule can also bind to endothelial cells [87, 88], as well as neutrophils and phagocytes [89]. Rhodostomin is a disintegrin that inhibits activity of LPS-treated monocytes via αvβ3 integrin affecting haemostasis, cell-cell interaction and suppresses tumor growth [90]. In hematophagous organisms, the disintegrins have been described in tick and leech saliva [91]. Here, one CDS related to disintegrins was identified in the *R. neglectu*s SG transcriptome. The alignment exhibited a high degree of homology between *R. neglectus* disintegrin sequence and others of the same family (Fig 8). The identification of ADAMTS is one of the main findings from the sialotranscriptome of *R. neglectus*.

Fig 8. The ADAMTS sequence from *R. neglectus* SG transcriptome. ClustalW alignment of the ADAMTS sequence from *R. neglectus* salivary transcriptome (RN_11351) and other insect ADAMTS sequences, identified as described in Methods section. The alignment indicates conserved residues in black and similar residues in gray background. The bar indicates the signal peptide indicative of secretion. The boxes limit the family signature sequences showing the determinant residues (black asterisk). The symbols (black hash) above indicate the conserved cysteines.

Serine proteases and trypsin-like proteins

The *R. neglectus* SG transcriptome revealed serine proteases reads as the second most abundant group in the secreted class, comprising 820,619 reads. The majority of the sequences identified as serine proteases exhibited the trypsin domain (Tryp-SPc) of the CDD and Smart databases. RN_1189 was assembled from 768,048. Its alignment with serine proteases from other organisms revealed conserved residues located around the cleavage and active sites (Fig 9, A), and the presence of a CUB (complement C1r/C1s, Uegf, Bmp1) domain, a structural motif of approximately 110 residues found almost exclusively in extracellular and plasma membrane-associated proteins. This domain is also present in honeybee allergens Api SI and Api SII, which are probably components of the honeybee defense system [92, 93].

Some serine proteases can function as regulators of coagulation. Thrombin can participate in this regulation by binding to thrombomodulin, a membrane protein present in host endothelial cells. This complex is able to activate Protein C (a serine protease), which acts as a potent anticoagulant enzyme by inactivating factors V and VIII, impairing thrombus progression [94, 95]. In snake venom, blockage of thrombus formation by serine proteases has also been reported. SPSV (Serine Protease Snake Venom) releases a unique fibrinopeptide that produces only instable monomers of fibrin, leading to clots that are rapidly dispersed [96]. Although the specific role in hematophagous saliva is still unknown, an active serine protease was described in *T. infestans* [97], as well as in horse fly *Tabanus yao* saliva, which functions as a fibrinogenolytic enzyme [98]. RN_22226, RN_21634, RN_19989, RN_17969, and RN_10652 were matched by blastp to serine proteases of *T. infestans*, *T. braziliensis,* *Panstrongylus megistus,* and *R. prolixus*. Serine proteases also play important roles in fertilization, embryonic development, and in the processes of molting and metamorphosis of insects [99, 100]. In our sample, the triatomines did not show any sign of larval molting at SG dissection.

Sequences containing CLIP, LDLa and SUSHI domains, which are cysteine-stabilized structures for molecular recognition, were also identified. The CLIP domain is restricted to the Arthropoda and was found N-terminally to the Tryp-SPc domain of RN_16275, RN_18155, and RN_7118. Both domains belong to the serine proteases of the trypsin-like S1 family, that are typically secreted enzymes associated with extracellular proteolysis [101]. CLIP domain has been suggested to be important for dimerization, mediating specific protein-protein interactions involved in the regulation of serine protease activities. The LDLa domain was identified in RN_12992, RN_12776, RN_12432, and RN_21634. The last two sequences also presented the

SUSHI motif of smart database which is known as CCP (Complement Control Protein) module, containing approximately 60 amino acid residues identified in several proteins of the complement system. These *R. neglectus* putative secreted serine proteases may play critical roles in many key biological processes as blood coagulation and immunity. In the vertebrate hosts, allergenicity may reinforce the toxic effect of serine proteases, independently of their catalytic activity, as proposed by Georgieva and colleagues [93].

The phylogram clearly showed four different groups (Fig 9, B), indicating the expression of at least four genes related to serine proteases in the saliva of *R. neglectus*. In regard to the large amount of reads, the results observed here suggests that, to *R. neglectus*, the serine proteases arise as an important salivary secreted gene family, a probably evolutionary adaptation where the protein could gain a new function as a result of selective pressure for the blood-feeding behavior success. These proteases could act in the vertebrate host, as well as in the insect, on pathogens that were ingested with blood. Further experiments are necessary to address the possible roles of those genes on the biology of *R. neglectus*.

Fig 9. Serine proteases from *R. neglectus* SG transcriptome.
(A) ClustalW alignment of a serine protease from *R. neglectus* SG transcriptome (RN_1189) and other serine proteases members, identified as described in Methods section. The alignment indicates conserved residues in black and similar residues in gray background. The blue bar indicates the signal peptide indicative of secretion. The symbols above the residues indicate (black circle) cleavage site and (black triangle) active site showing the HDS triad.
(B) Phylogenetic tree was built from the alignment of *R. neglectus* CDS and other insect sequences as described in Methods section The bar at the bottom represents 20% amino acid substitution. The colored circles identify the sequences used: blue, *R. neglectus* sequences from SG transcriptome; red, Hemiptera order; green, Hymenoptera order; magenta, Dictyoptera order.

OBP

513 The odorant-binding protein (OBP) family is a chemosensory protein ubiquitous
514 in insects commonly associated with solubilizers and carriers of odorants and
515 pheromones. Although associated with chemosensory organs, in recent times this
516 family has been related to other roles such as hydrophobic chemical transportation
517 [102]. The OBPs are characterized by a variable amino acid sequence, but conserve a
518 pattern of six conserved cysteines residues paired to form three disulfide bridges [103].
519 The folding is a typical six α-helices assembled in a compact and stable structure [104,
520 105]. Eleven CDS containing protein sequences related to OBPs were recognized in our
521 transcriptome analysis, all possessing signal peptide prediction. The conserved cysteine
522 residues of *R. neglectus* CDS (Fig 10) were seen during alignment. Phylogenetic
523 analysis with good bootstrap support shows Clade I containing most *R. neglectus*
524 sequences grouped with *R. prolixus* OBP. However, RN_3440 was grouped in Clade
525 III, suggesting this is a more distant OBP (S4 Fig).

526 **Fig 10. The secreted OBP family from *R. neglectus* SG transcriptome.**
527 ClustalW alignment of secreted OBPs from *R. neglectus* SG transcriptome (RN_5213,
528 RN_3440 and RN_18954) and other members from the OBP family, identified as
529 described in Methods section. The alignment indicates conserved residues in black and
530 similar residues in gray background. The blue bar indicates the signal peptide indicative
531 of secretion and the boxes, the six conserved cysteines.
532

533 The proteome of *R. neglectus* saliva

534 *R. neglectus* saliva content was tryptic digested and subjected to mass
535 spectrometry to validate the analysis of the transcripts possibly associated with secreted
536 products. Among the 73 identified secreted proteins groups, 48 were from the lipocalin
537 family, including triabin, pallidipin and nitrophorin proteins, reaffirming their
538 abundance (Table 4). Other soluble proteins, predicted as being secreted by these
539 arthropods, were: secreted metalloprotease, antigen-5, serpin and trypsin-like protease,

each with at least one observation. It is intriguing that only one serine protease was detected by proteomic analysis, regardless the high number of transcripts reads assigned to this subclass of putative secreted proteins (Table 3). There are several possible explanations for this observation. First, the proteins are expressed in a such a small amount not detectable by our proteomic approach; second, the proteins are not secreted; third and most likely, these proteins present in SGs are expressed upon specific physiological conditions, such as during stimulation of salivation (feeding). In addition, it is also possible that those enzymes have both intracellular and extracellular functions as many other proteases do.

Table 4: Classification and abundance of proteins from the salivary proteome of *R. neglectus* based on LC-MS/MS

Class	No. of protein groups	% Total
Lipocalin – Triabin	34	46.58
Lipocalin – Nitrophorin	17	23.29
Others	13	17.81
Conserved secreted protein	2	2.74
Antigen-5/SCP	1	1.37
Inositol polyphosphate phosphatase	1	1.37
Metalloprotease	1	1.37
Protease inhibitor – Serpin	1	1.37
Chitinase-like lectin	1	1.37
Trypsin-like protease	1	1.37
Juvenile hormone related	1	1.37
Total	**73**	**100**

Comparison of protein contents between *R. neglectus* and *R. prolixus*

A comparative blastp analysis was employed to address the similarity of the SGs proteins from *R. neglectus* compared to *R. prolixus*. The two species do not show high evolutionary divergence, presenting at least 80% identity in analyzed sequences (Table 5), suggesting both species share a common ancestral lineage. As described before, this

558 high degree of protein similarity was also seen with *R. brethesi* and *R. robustus* in the
559 Amazon rainforest [106].

560 **Table 5: Identities of *R. neglectus* proteins compared to *R. prolixus* (v 3.0) proteins by**
561 **blastp.**

Class	Average identity	SE*	N
Secreted	86.46	1.47	200
Immunity	91.17	3.76	35
Transporters and channels	92.33	1.22	114
Extracelular matrix	94.07	1.81	43
Transposable elements	94.64	2.50	22
Protein export	95.09	0.84	166
Signal transduction	95.39	0.54	285
Storage	95.50	1.44	8
Cytoskeletal	95.64	0.96	88
Unknown conserved	95.99	0.44	369
Protein modification	95.99	0.78	144
Metabolism	96.28	0.45	405
Transcription machinery	96.50	0.50	250
Proteasome machinery	96.59	0.80	122
Nuclear regulation	96.65	0.75	95
Nuclear export	96.67	1.05	12
Unknown	97.04	1.47	24
Detoxification	97.15	1.16	48
Transcription factor	97.43	0.63	53
Protein synthesis	97.95	0.42	180
Total			**2,663**

562 **SE:* standard error
563

564 Identity among sequences was greater in housekeeping class members, showing
565 that these proteins have a lower evolutionary rate than those of the secreted class. This
566 indicates that antihaemostatic proteins evolve faster after divergence. Different
567 molecular mechanisms may be responsible for the variation between these closely
568 related *Rhodnius* species, expanding their biological diversity patterns. The particularity
569 of each species could be related to their different habitats, including different prey and
570 abiotic factors.

571

Final Considerations

Hematophagy evolved independently at least six times in approximately 15,000 species allowing for adaptation to an existing complex host haemostatic system [5, 107]. Thus, many salivary molecules target different pathways for the insect to achieve a successful blood meal. Here, we described *R. neglectus* sialome in all its complexity to expand our knowledge of the salivary proteins from hematophagous triatomine bugs.

R. neglectus is considered of secondary importance in the transmission of *T. cruzi*, causative agent of Chagas disease. The analysis of salivary secretory products of *R. neglectus* that might be involved in vector-host interactions share similarity with other triatomine species, which can also be infected by and transmit the protozoan.

It is possible that the expression of putative trypsin-like serine proteases in the SGs of *R. neglectus* correlates with blood sources of this species of triatomine. Their role and that of other secreted class, hypothetical and conserved secreted proteins, in hematophagy should be analyzed in future works, and we accentuate that sialome study is still an open field for new discoveries.

Acknowledgments

We thank Ana Cristina Gomes and Fabiano Bastos for technical assistance.

References

1. LAVOIPIERRE MM, DICKERSON G, GORDON RM. Studies on the methods of feeding of blood-sucking arthropods. I. The manner in which triatomine bugs obtain their blood-meal, as observed in the tissues of the living rodent, with some remarks on the effects of the bite on human volunteers. Ann Trop Med Parasitol. 1959;53:235-50. PubMed PMID: 14414675.
2. Fontaine A, Diouf I, Bakkali N, Misse D, Pages F, Fusai T, et al. Implication of haematophagous arthropod salivary proteins in host-vector interactions. Parasit Vectors.

600 2011;4:187. doi: 10.1186/1756-3305-4-187. PubMed PMID: 21951834; PubMed
601 Central PMCID: PMC3197560.
602 3. Champagne DE. Antihemostatic strategies of blood-feeding arthropods. Current
603 drug targets. 2004;4(4):375-96. PubMed PMID: 15578959.
604 4. Champagne DE. Antihemostatic molecules from saliva of blood-feeding
605 arthropods. Pathophysiology of haemostasis and thrombosis. 2005;34(4-5):221-7. doi:
606 10.1159/000092428. PubMed PMID: 16707932.
607 5. Ribeiro JM. Blood-feeding arthropods: live syringes or invertebrate
608 pharmacologists? Infect Agents Dis. 1995;4(3):143-52. PubMed PMID: 8548192.
609 6. Steverding D. The history of Chagas disease. Parasit Vectors. 2014;7:317. doi:
610 10.1186/1756-3305-7-317. PubMed PMID: 25011546; PubMed Central PMCID:
611 PMCPMC4105117.
612 7. Guhl F, Pinto N, Aguilera G. Sylvatic triatominae: a new challenge in vector
613 control transmission. Mem Inst Oswaldo Cruz. 2009;104 Suppl 1:71-5. PubMed PMID:
614 19753461.
615 8. Fitzpatrick S, Feliciangeli MD, Sanchez-Martin MJ, Monteiro FA, Miles MA.
616 Molecular genetics reveal that silvatic Rhodnius prolixus do colonise rural houses.
617 PLoS Negl Trop Dis. 2008;2(4):e210. doi: 10.1371/journal.pntd.0000210. PubMed
618 PMID: 18382605; PubMed Central PMCID: PMC2270345.
619 9. Abad-Franch F, Monteiro FA, Jaramillo O N, Gurgel-Gonçalves R, Dias FB,
620 Diotaiuti L. Ecology, evolution, and the long-term surveillance of vector-borne Chagas
621 disease: a multi-scale appraisal of the tribe Rhodniini (Triatominae). Acta Trop.
622 2009;110(2-3):159-77. doi: 10.1016/j.actatropica.2008.06.005. PubMed PMID:
623 18619938.
624 10. Gurgel-Goncalves R, Galvao C, Costa J, Peterson AT. Geographic distribution
625 of chagas disease vectors in Brazil based on ecological niche modeling. J Trop Med.
626 2012;2012:705326. doi: 10.1155/2012/705326. PubMed PMID: 22523500; PubMed
627 Central PMCID: PMC3317230.
628 11. Gurgel-Goncalves R, Cura C, Schijman AG, Cuba CA. Infestation of Mauritia
629 flexuosa palms by triatomines (Hemiptera: Reduviidae), vectors of Trypanosoma cruzi
630 and Trypanosoma rangeli in the Brazilian savanna. Acta Trop. 2012;121(2):105-11. doi:
631 10.1016/j.actatropica.2011.10.010. PubMed PMID: 22037200.
632 12. Rodrigues VL, Pauliquevis Junior C, da Silva RA, Wanderley DM, Guirardo
633 MM, Rodas LA, et al. Colonization of palm trees by Rhodnius neglectus and household
634 and invasion in an urban area, Araçatuba, São Paulo State, Brazil. Rev Inst Med Trop
635 Sao Paulo. 2014;56(3):213-8. PubMed PMID: 24878999; PubMed Central PMCID:
636 PMCPMC4085872.
637 13. Rubens Antonio Silva PRBeDMVW. Doença de Chagas no Estado de São
638 Paulo: comparaçãoentre pesquisa ativa de triatomíneos em domicílios enotificação de
639 sua presença pela população emárea sob vigilância entomológica. Revista da Sociedade
640 Brasileira de Medicina Tropical1999.
641 14. Gurgel-Gonçalves R, Ramalho ED, Duarte MA, Palma AR, Abad-Franch F,
642 Carranza JC, et al. Enzootic transmission of Trypanosoma cruzi and T. rangeli in the
643 Federal District of Brazil. Rev Inst Med Trop Sao Paulo. 2004;46(6):323-30. doi:
644 /S0036-46652004000600005. PubMed PMID: 15654478.
645 15. W. S. de Oliveira A, G. da Silva I. Distribuição geográfica e indicadores
646 entomológicos de triatomíneos sinantrópicos capturados no Estado de Goiás. Revista da
647 Sociedade Brasileira de Medicina Tropical; 2007. p. 204-8.

16. Almeida PSd, al e. Levantamento da fauna de Triatominae (Hemiptera: Reduviidae) em ambiente domiciliar e infecção natural por Trypanosomatidae no Estado de Mato Grosso do Sul. Uberaba: Rev. Soc. Bras. Med. Trop; 2008.
17. Gurgel-Gonçalves R, Abad-Franch F, Ferreira JB, Santana DB, Cuba CA. Is Rhodnius prolixus (Triatominae) invading houses in central Brazil? Acta Trop. 2008;107(2):90-8. doi: 10.1016/j.actatropica.2008.04.020. PubMed PMID: 18550022.
18. Ribeiro JM, Chagas AC, Pham VM, Lounibos LP, Calvo E. An insight into the sialome of the frog biting fly, *Corethrella appendiculata*. Insect biochemistry and molecular biology. 2013. Epub 2014/02/12. doi: 10.1016/j.ibmb.2013.10.006. PubMed PMID: 24514880.
19. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. Genome Res. 2009;19(6):1117-23. doi: 10.1101/gr.089532.108. PubMed PMID: 19251739; PubMed Central PMCID: PMCPMC2694472.
20. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience. 2012;1(1):18. doi: 10.1186/2047-217X-1-18. PubMed PMID: 23587118; PubMed Central PMCID: PMC3626529.
21. Karim S, Singh P, Ribeiro JM. A deep insight into the sialotranscriptome of the gulf coast tick, Amblyomma maculatum. PLoS ONE. 2011;6(12):e28525. Epub 2012/01/05. doi: 10.1371/journal.pone.0028525. PubMed PMID: 22216098; PubMed Central PMCID: PMC3244413.
22. Nielsen H, Brunak S, von Heijne G. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein engineering. 1999;12(1):3-9. PubMed PMID: 10065704.
23. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-402. PubMed PMID: 9254694; PubMed Central PMCID: PMC146917.
24. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT Oxford University Press: Nucleic Acids Symposium Series; 1999. p. 95-8.
25. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947-8. doi: 10.1093/bioinformatics/btm404. PubMed PMID: 17846036.
26. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform. 2004;5(2):150-63. PubMed PMID: 15260895.
27. Queiroz RM, Charneau S, Motta FN, Santana JM, Roepstorff P, Ricart CA. Comprehensive proteomic analysis of Trypanosoma cruzi epimastigote cell surface proteins by two complementary methods. J Proteome Res. 2013;12(7):3255-63. doi: 10.1021/pr400110h. PubMed PMID: 23682730.
28. Flower DR. The lipocalin protein family: structure and function. Biochem J. 1996;318 (Pt 1):1-14. PubMed PMID: 8761444; PubMed Central PMCID: PMC1217580.
29. Francischetti IM, Sa-Nunes A, Mans BJ, Santos IM, Ribeiro JM. The role of saliva in tick feeding. Front Biosci. 2009;14:2051-88. PubMed PMID: 19273185; PubMed Central PMCID: PMCPMC2785505.

30. Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD. Triabin, a highly potent exosite inhibitor of thrombin. The Journal of biological chemistry. 1995;270(48):28629-34. PubMed PMID: 7499380.
31. Fuentes-Prior P, Noeske-Jungblut C, Donner P, Schleuning WD, Huber R, Bode W. Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. Proc Natl Acad Sci U S A. 1997;94(22):11845-50. PubMed PMID: 9342325; PubMed Central PMCID: PMC23629.
32. Montfort WR, Weichsel A, Andersen JF. Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. Biochimica et biophysica acta. 2000;1482(1-2):110-8. Epub 2000/11/04. doi: S0167-4838(00)00165-5 [pii]. PubMed PMID: 11058753.
33. Ribeiro JM, Walker FA. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorin) of *Rhodnius prolixus*. The Journal of experimental medicine. 1994;180(6):2251-7. Epub 1994/12/01. PubMed PMID: 7964498; PubMed Central PMCID: PMC2191789.
34. Champagne DE, Nussenzveig RH, Ribeiro JM. Purification, partial characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood-sucking insect Rhodnius prolixus. J Biol Chem. 1995;270(15):8691-5. PubMed PMID: 7721773.
35. Zhang Y, Ribeiro JM, Guimaraes JA, Walsh PN. Nitrophorin-2: a novel mixed-type reversible specific inhibitor of the intrinsic factor-X activating complex. Biochemistry. 1998;37(30):10681-90.
36. Andersen JF, Gudderra NP, Francischetti IM, Valenzuela JG, Ribeiro JM. Recognition of anionic phospholipid membranes by an antihemostatic protein from a blood-feeding insect. Biochemistry. 2004;43(22):6987-94. PubMed PMID: 15170336.
37. Soares RP, Sant'Anna MR, Gontijo NF, Romanha AJ, Diotaiuti L, Pereira MH. Identification of morphologically similar Rhodnius species (Hemiptera: Reduviidae: Triatominae) by electrophoresis of salivary heme proteins. Am J Trop Med Hyg. 2000;62(1):157-61. PubMed PMID: 10761743.
38. Gibbs GM, Roelants K, O'Bryan MK. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins--roles in reproduction, cancer, and immune defense. Endocrine reviews. 2008;29(7):865-97. PubMed PMID: 18824526.
39. Fang KS, Vitale M, Fehlner P, King TP. cDNA cloning and primary structure of a white-face hornet venom allergen, antigen 5. Proc Natl Acad Sci U S A. 1988;85(3):895-9. PubMed PMID: 3422469; PubMed Central PMCID: PMC279663.
40. Hoffman DR. Allergens in Hymenoptera venom. XXV: The amino acid sequences of antigen 5 molecules and the structural basis of antigenic cross-reactivity. The Journal of allergy and clinical immunology. 1993;92(5):707-16. PubMed PMID: 8227862.
41. Yamazaki Y, Koike H, Sugiyama Y, Motoyoshi K, Wada T, Hishinuma S, et al. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. Eur J Biochem. 2002;269(11):2708-15. PubMed PMID: 12047379.
42. Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, et al. Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie. 1993;75(8):687-706. PubMed PMID: 8286442.
43. Calvo E, Dao A, Pham VM, Ribeiro JM. An insight into the sialome of Anopheles funestus reveals an emerging pattern in anopheline salivary protein families. Insect biochemistry and molecular biology. 2007;37(2):164-75. Epub 2007/01/25. doi:

745 10.1016/j.ibmb.2006.11.005. PubMed PMID: 17244545; PubMed Central PMCID:
746 PMC1853278.
747 44. Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JM. Toward
748 a description of the sialome of the adult female mosquito Aedes aegypti. Insect
749 biochemistry and molecular biology. 2002;32(9):1101-22. Epub 2002/09/06. PubMed
750 PMID: 12213246.
751 45. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding
752 of the biochemical and pharmacological complexity of the saliva of a hematophagous
753 sand fly Lutzomyia longipalpis. Proc Natl Acad Sci U S A. 1999;96(26):15155-60.
754 PubMed PMID: 10611354; PubMed Central PMCID: PMC24789.
755 46. Ribeiro JM, Andersen J, Silva-Neto MA, Pham VM, Garfield MK, Valenzuela
756 JG. Exploring the sialome of the blood-sucking bug Rhodnius prolixus. Insect Biochem
757 Mol Biol. 2004;34(1):61-79. PubMed PMID: 14976983.
758 47. Assumpcao TC, Francischetti IM, Andersen JF, Schwarz A, Santana JM, Ribeiro
759 JM. An insight into the sialome of the blood-sucking bug *Triatoma infestans,* a vector
760 of Chagas' disease. Insect biochemistry and molecular biology. 2008;38(2):213-32.
761 PubMed PMID: 18207082.
762 48. Assumpcao TC, Charneau S, Santiago PB, Francischetti IM, Meng Z, Araujo
763 CN, et al. Insight into the salivary transcriptome and proteome of Dipetalogaster
764 maxima. Journal of proteome research. 2011;10(2):669-79. doi: 10.1021/pr100866h.
765 PubMed PMID: 21058630; PubMed Central PMCID: PMC3035102.
766 49. Assumpcao TC, Eaton DP, Pham VM, Francischetti IM, Aoki V, Hans-Filho G,
767 et al. An insight into the sialotranscriptome of *Triatoma matogrossensis*, a kissing bug
768 associated with fogo selvagem in South America. The American journal of tropical
769 medicine and hygiene. 2012;86(6):1005-14. Epub 2012/06/06. doi:
770 10.4269/ajtmh.2012.11-0690. PubMed PMID: 22665609; PubMed Central PMCID:
771 PMC3366513.
772 50. Ribeiro JM, Assumpcao TC, Pham VM, Francischetti IM, Reisenman CE. An
773 insight into the sialotranscriptome of *Triatoma rubida* (Hemiptera: Heteroptera).
774 Journal of medical entomology. 2012;49(3):563-72. Epub 2012/06/12. PubMed PMID:
775 22679863; PubMed Central PMCID: PMC3544468.
776 51. Assumpção TC, Ma D, Schwarz A, Reiter K, Santana JM, Andersen JF, et al.
777 Salivary antigen-5/CAP family members are Cu2+-dependent antioxidant enzymes that
778 scavenge O_{2-}. and inhibit collagen-induced platelet aggregation and neutrophil
779 oxidative burst. J Biol Chem. 2013;288(20):14341-61. doi: 10.1074/jbc.M113.466995.
780 PubMed PMID: 23564450; PubMed Central PMCID: PMCPMC3656290.
781 52. Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation,
782 maintenance, and regulation. Biochemistry. 1991;30(43):10363-70. PubMed PMID:
783 1931959.
784 53. Oikonomopoulou K, Ricklin D, Ward PA, Lambris JD. Interactions between
785 coagulation and complement--their role in inflammation. Semin Immunopathol.
786 2012;34(1):151-65. doi: 10.1007/s00281-011-0280-x. PubMed PMID: 21811895;
787 PubMed Central PMCID: PMCPMC3372068.
788 54. Schlott B, Wohnert J, Icke C, Hartmann M, Ramachandran R, Guhrs KH, et al.
789 Interaction of Kazal-type inhibitor domains with serine proteinases: biochemical and
790 structural studies. Journal of molecular biology. 2002;318(2):533-46. PubMed PMID:
791 12051857.
792 55. Friedrich T, Kroger B, Bialojan S, Lemaire HG, Hoffken HW, Reuschenbach P,
793 et al. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. The
794 Journal of biological chemistry. 1993;268(22):16216-22. PubMed PMID: 8344906.

795 56. van de Locht A, Lamba D, Bauer M, Huber R, Friedrich T, Kroger B, et al. Two
796 heads are better than one: crystal structure of the insect derived double domain Kazal
797 inhibitor rhodniin in complex with thrombin. EMBO J. 1995;14(21):5149-57. PubMed
798 PMID: 7489704; PubMed Central PMCID: PMC394622.
799 57. Mende K, Petoukhova O, Koulitchkova V, Schaub GA, Lange U, Kaufmann R,
800 et al. Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect.
801 *Dipetalogaster maximus* cDNA cloning, expression and characterization. European
802 journal of biochemistry / FEBS. 1999;266(2):583-90. PubMed PMID: 10561601.
803 58. Campos IT, Amino R, Sampaio CA, Auerswald EA, Friedrich T, Lemaire HG,
804 et al. Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas'
805 disease vector: gene cloning, expression and characterization of the inhibitor. Insect
806 biochemistry and molecular biology. 2002;32(9):991-7. PubMed PMID: 12213235.
807 59. Araujo RN, Campos IT, Tanaka AS, Santos A, Gontijo NF, Lehane MJ, et al.
808 Brasiliensin: A novel intestinal thrombin inhibitor from *Triatoma brasiliensis*
809 (Hemiptera: Reduviidae) with an important role in blood intake. International journal
810 for parasitology. 2007;37(12):1351-8. PubMed PMID: 17575982.
811 60. Takac P, Nunn MA, Meszaros J, Pechanova O, Vrbjar N, Vlasakova P, et al.
812 Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera,
813 Tabanidae) salivary glands. J Exp Biol. 2006;209(Pt 2):343-52. PubMed PMID:
814 16391356.
815 61. Liang Z, Sottrup-Jensen L, Aspan A, Hall M, Soderhall K. Pacifastin, a novel
816 155-kDa heterodimeric proteinase inhibitor containing a unique transferrin chain. Proc
817 Natl Acad Sci U S A. 1997;94(13):6682-7. PubMed PMID: 9192625; PubMed Central
818 PMCID: PMC21218.
819 62. Hergenhahn HG, Aspan A, Soderhall K. Purification and characterization of a
820 high-Mr proteinase inhibitor of pro-phenol oxidase activation from crayfish plasma.
821 Biochem J. 1987;248(1):223-8. PubMed PMID: 3124809; PubMed Central PMCID:
822 PMC1148522.
823 63. Breugelmans B, Simonet G, van Hoef V, Van Soest S, Broeck JV. Identification,
824 distribution and molecular evolution of the pacifastin gene family in Metazoa. BMC
825 Evol Biol. 2009;9:97. doi: 10.1186/1471-2148-9-97. PubMed PMID: 19435517;
826 PubMed Central PMCID: PMC2689174.
827 64. de Marco R, Lovato DV, Torquato RJ, Clara RO, Buarque DS, Tanaka AS. The
828 first pacifastin elastase inhibitor characterized from a blood sucking animal. Peptides.
829 2010;31(7):1280-6. Epub 2010/04/13. doi: S0196-9781(10)00148-8 [pii]
830 10.1016/j.peptides.2010.03.033. PubMed PMID: 20381560.
831 65. Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin
832 superfamily: implications of patterns of amino acid conservation for structure and
833 function. Genome Res. 2000;10(12):1845-64. PubMed PMID: 11116082.
834 66. Rau JC, Beaulieu LM, Huntington JA, Church FC. Serpins in thrombosis,
835 hemostasis and fibrinolysis. J Thromb Haemost. 2007;5 Suppl 1:102-15. PubMed
836 PMID: 17635716.
837 67. Stark KR, James AA. Isolation and characterization of the gene encoding a
838 novel factor Xa-directed anticoagulant from the yellow fever mosquito, Aedes aegypti. J
839 Biol Chem. 1998;273(33):20802-9. PubMed PMID: 9694825.
840 68. Prevot PP, Beschin A, Lins L, Beaufays J, Grosjean A, Bruys L, et al. Exosites
841 mediate the anti-inflammatory effects of a multifunctional serpin from the saliva of the
842 tick Ixodes ricinus. FEBS J. 2009;276(12):3235-46. doi: 10.1111/j.1742-
843 4658.2009.07038.x. PubMed PMID: 19438720.

844 69. Prevot PP, Adam B, Boudjeltia KZ, Brossard M, Lins L, Cauchie P, et al. Anti-
845 hemostatic effects of a serpin from the saliva of the tick Ixodes ricinus. The Journal of
846 biological chemistry. 2006;281(36):26361-9. PubMed PMID: 16672226.
847 70. Leboulle G, Crippa M, Decrem Y, Mejri N, Brossard M, Bollen A, et al.
848 Characterization of a novel salivary immunosuppressive protein from Ixodes ricinus
849 ticks. J Biol Chem. 2002;277(12):10083-9. doi: 10.1074/jbc.M111391200. PubMed
850 PMID: 11792703.
851 71. Gettins PG. Serpin structure, mechanism, and function. Chem Rev.
852 2002;102(12):4751-804. PubMed PMID: 12475206.
853 72. Bode W, Gomis-Ruth FX, Stockler W. Astacins, serralysins, snake venom and
854 matrix metalloproteinases exhibit identical zinc-binding environments
855 (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a
856 common family, the 'metzincins'. FEBS letters. 1993;331(1-2):134-40.
857 73. Jiang W, Bond JS. Families of metalloendopeptidases and their relationships.
858 FEBS letters. 1992;312(2-3):110-4.
859 74. Hooper NM. Families of zinc metalloproteases. FEBS letters. 1994;354(1):1-6.
860 75. da Silveira RB, dos Santos Filho JF, Mangili OC, Veiga SS, Gremski W, Nader
861 HB, et al. Identification of proteases in the extract of venom glands from brown spiders.
862 Toxicon. 2002;40(6):815-22. PubMed PMID: 12175619.
863 76. Feitosa L, Gremski W, Veiga SS, Elias MC, Graner E, Mangili OC, et al.
864 Detection and characterization of metalloproteinases with gelatinolytic,
865 fibronectinolytic and fibrinogenolytic activities in brown spider (Loxosceles intermedia)
866 venom. Toxicon. 1998;36(7):1039-51.
867 77. da Silveira RB, Wille AC, Chaim OM, Appel MH, Silva DT, Franco CR, et al.
868 Identification, cloning, expression and functional characterization of an astacin-like
869 metalloprotease toxin from Loxosceles intermedia (brown spider) venom. Biochem J.
870 2007;406(2):355-63. doi: 10.1042/BJ20070363. PubMed PMID: 17535156; PubMed
871 Central PMCID: PMC1948970.
872 78. Trevisan-Silva D, Gremski LH, Chaim OM, da Silveira RB, Meissner GO,
873 Mangili OC, et al. Astacin-like metalloproteases are a gene family of toxins present in
874 the venom of different species of the brown spider (genus Loxosceles). Biochimie.
875 2010;92(1):21-32. doi: 10.1016/j.biochi.2009.10.003. PubMed PMID: 19879318.
876 79. Porter S, Clark IM, Kevorkian L, Edwards DR. The ADAMTS
877 metalloproteinases. Biochem J. 2005;386(Pt 1):15-27. doi: 10.1042/BJ20040424.
878 PubMed PMID: 15554875; PubMed Central PMCID: PMC1134762.
879 80. Kelwick R, Desanlis I, Wheeler GN, Edwards DR. The ADAMTS (A
880 Disintegrin and Metalloproteinase with Thrombospondin motifs) family. Genome Biol.
881 2015;16:113. doi: 10.1186/s13059-015-0676-3. PubMed PMID: 26025392; PubMed
882 Central PMCID: PMCPMC4448532.
883 81. Gould RJ, Polokoff MA, Friedman PA, Huang TF, Holt JC, Cook JJ, et al.
884 Disintegrins: a family of integrin inhibitory proteins from viper venoms. Proc Soc Exp
885 Biol Med. 1990;195(2):168-71. PubMed PMID: 2236100.
886 82. Calvete JJ, Marcinkiewicz C, Monleon D, Esteve V, Celda B, Juarez P, et al.
887 Snake venom disintegrins: evolution of structure and function. Toxicon.
888 2005;45(8):1063-74. doi: 10.1016/j.toxicon.2005.02.024. PubMed PMID: 15922775.
889 83. Huang TF. What have snakes taught us about integrins? Cell Mol Life Sci.
890 1998;54(6):527-40. PubMed PMID: 9676572.
891 84. Hsu CC, Wu WB, Chang YH, Kuo HL, Huang TF. Antithrombotic effect of a
892 protein-type I class snake venom metalloproteinase, kistomin, is mediated by affecting

893 glycoprotein Ib-von Willebrand factor interaction. Mol Pharmacol. 2007;72(4):984-92.
894 doi: 10.1124/mol.107.038018. PubMed PMID: 17609416.
895 85. Huang TF, Chang MC, Teng CM. Antiplatelet protease, kistomin, selectively
896 cleaves human platelet glycoprotein Ib. Biochim Biophys Acta. 1993;1158(3):293-9.
897 PubMed PMID: 8251530.
898 86. Hsu CC, Wu WB, Huang TF. A snake venom metalloproteinase, kistomin,
899 cleaves platelet glycoprotein VI and impairs platelet functions. J Thromb Haemost.
900 2008;6(9):1578-85. doi: 10.1111/j.1538-7836.2008.03071.x. PubMed PMID: 18624975.
901 87. Wu WB, Chang SC, Liau MY, Huang TF. Purification, molecular cloning and
902 mechanism of action of graminelysin I, a snake-venom-derived metalloproteinase that
903 induces apoptosis of human endothelial cells. Biochem J. 2001;357(Pt 3):719-28.
904 PubMed PMID: 11463342; PubMed Central PMCID: PMC1222001.
905 88. Wu WB, Huang TF. Activation of MMP-2, cleavage of matrix proteins, and
906 adherens junctions during a snake venom metalloproteinase-induced endothelial cell
907 apoptosis. Exp Cell Res. 2003;288(1):143-57. PubMed PMID: 12878166.
908 89. Tseng YL, Lee CJ, Huang TF. Effects of a snake venom metalloproteinase,
909 triflamp, on platelet aggregation, platelet-neutrophil and neutrophil-neutrophil
910 interactions: involvement of platelet GPIbalpha and neutrophil PSGL-1. Thromb
911 Haemost. 2004;91(2):315-24. doi: 10.1160/TH03-07-0426. PubMed PMID: 14961160.
912 90. Yeh CH, Peng HC, Yang RS, Huang TF. Rhodostomin, a snake venom
913 disintegrin, inhibits angiogenesis elicited by basic fibroblast growth factor and
914 suppresses tumor growth by a selective alpha(v)beta(3) blockade of endothelial cells.
915 Mol Pharmacol. 2001;59(5):1333-42. PubMed PMID: 11306719.
916 91. Assumpção TC, Ribeiro JM, Francischetti IM. Disintegrins from hematophagous
917 sources. Toxins (Basel). 2012;4(5):296-322. doi: 10.3390/toxins4050296. PubMed
918 PMID: 22778902; PubMed Central PMCID: PMC3386632.
919 92. Hoffman DR. Hymenoptera venom allergens. Clin Rev Allergy Immunol.
920 2006;30(2):109-28. PubMed PMID: 16645223.
921 93. Georgieva D, Greunke K, Arni RK, Betzel C. Three-dimensional modelling of
922 honeybee venom allergenic proteases: relation to allergenicity. Z Naturforsch C.
923 2011;66(5-6):305-12. PubMed PMID: 21812349.
924 94. Kalafatis M, Rand MD, Mann KG. The mechanism of inactivation of human
925 factor V and human factor Va by activated protein C. J Biol Chem.
926 1994;269(50):31869-80. PubMed PMID: 7989361.
927 95. Kalafatis M, Egan JO, van 't Veer C, Cawthern KM, Mann KG. The regulation
928 of clotting factors. Crit Rev Eukaryot Gene Expr. 1997;7(3):241-80. PubMed PMID:
929 9399073.
930 96. Stocker K, Fischer H, Meier J. Thrombin-like snake venom proteinases.
931 Toxicon. 1982;20(1):265-73. PubMed PMID: 7043783.
932 97. Amino R, Tanaka AS, Schenkman S. Triapsin, an unusual activatable serine
933 protease from the saliva of the hematophagous vector of Chagas' disease Triatoma
934 infestans (Hemiptera: Reduviidae). Insect biochemistry and molecular biology.
935 2001;31(4-5):465-72. Epub 2001/02/27. PubMed PMID: 11222956.
936 98. Xu X, Yang H, Ma D, Wu J, Wang Y, Song Y, et al. Toward an understanding
937 of the molecular mechanism for successful blood feeding by coupling proteomics
938 analysis with pharmacological testing of horsefly salivary glands. Mol Cell Proteomics.
939 2008;7(3):582-90. PubMed PMID: 18087067.
940 99. Jiang H, Kanost MR. The clip-domain family of serine proteinases in
941 arthropods. Insect Biochem Mol Biol. 2000;30(2):95-105. PubMed PMID: 10696585.

100. He W-Y, Zheng Y-P, Tang L, Zheng S-C, Béliveau C, Doucet D, et al. Cloning, expression and localization of a trypsin-like serine protease in the spruce budworm, *Choristoneura fumiferana*. Insect Science. 2009;16(6):455-64.
101. Antalis TM, Buzza MS, Hodge KM, Hooper JD, Netzel-Arnett S. The cutting edge: membrane-anchored serine protease activities in the pericellular microenvironment. Biochem J. 2010;428(3):325-46. doi: 10.1042/BJ20100046. PubMed PMID: 20507279; PubMed Central PMCID: PMCPMC3680374.
102. Pelosi P, Iovinella I, Felicioli A, Dani FR. Soluble proteins of chemical communication: an overview across arthropods. Front Physiol. 2014;5:320. doi: 10.3389/fphys.2014.00320. PubMed PMID: 25221516; PubMed Central PMCID: PMC4145409.

103. Leal WS, Nikonova L, Peng G. Disulfide structure of the pheromone binding protein from the silkworm moth, Bombyx mori. FEBS Lett. 1999;464(1-2):85-90. PubMed PMID: 10611489.

104. Tegoni M, Campanacci V, Cambillau C. Structural aspects of sexual attraction and chemical communication in insects. Trends Biochem Sci. 2004;29(5):257-64. doi: 10.1016/j.tibs.2004.03.003. PubMed PMID: 15130562.

105. Sandler BH, Nikonova L, Leal WS, Clardy J. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem Biol. 2000;7(2):143-51. PubMed PMID: 10662696.

106. Costa CM, Sousa MV, Ricart CA, Santana JM, Teixeira AR, Roepstorff P, et al. 2-DE-based proteomic investigation of the saliva of the Amazonian triatomine vectors of Chagas disease: Rhodnius brethesi and Rhodnius robustus. J Proteomics. 2011;74(9):1652-63. doi: 10.1016/j.jprot.2011.02.022. PubMed PMID: 21362504.

107. Mans BJ, Louw AI, Neitz AW. Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus Ornithodoros. Mol Biol Evol. 2002;19(10):1695-705. PubMed PMID: 12270896.

Supporting Information Legends

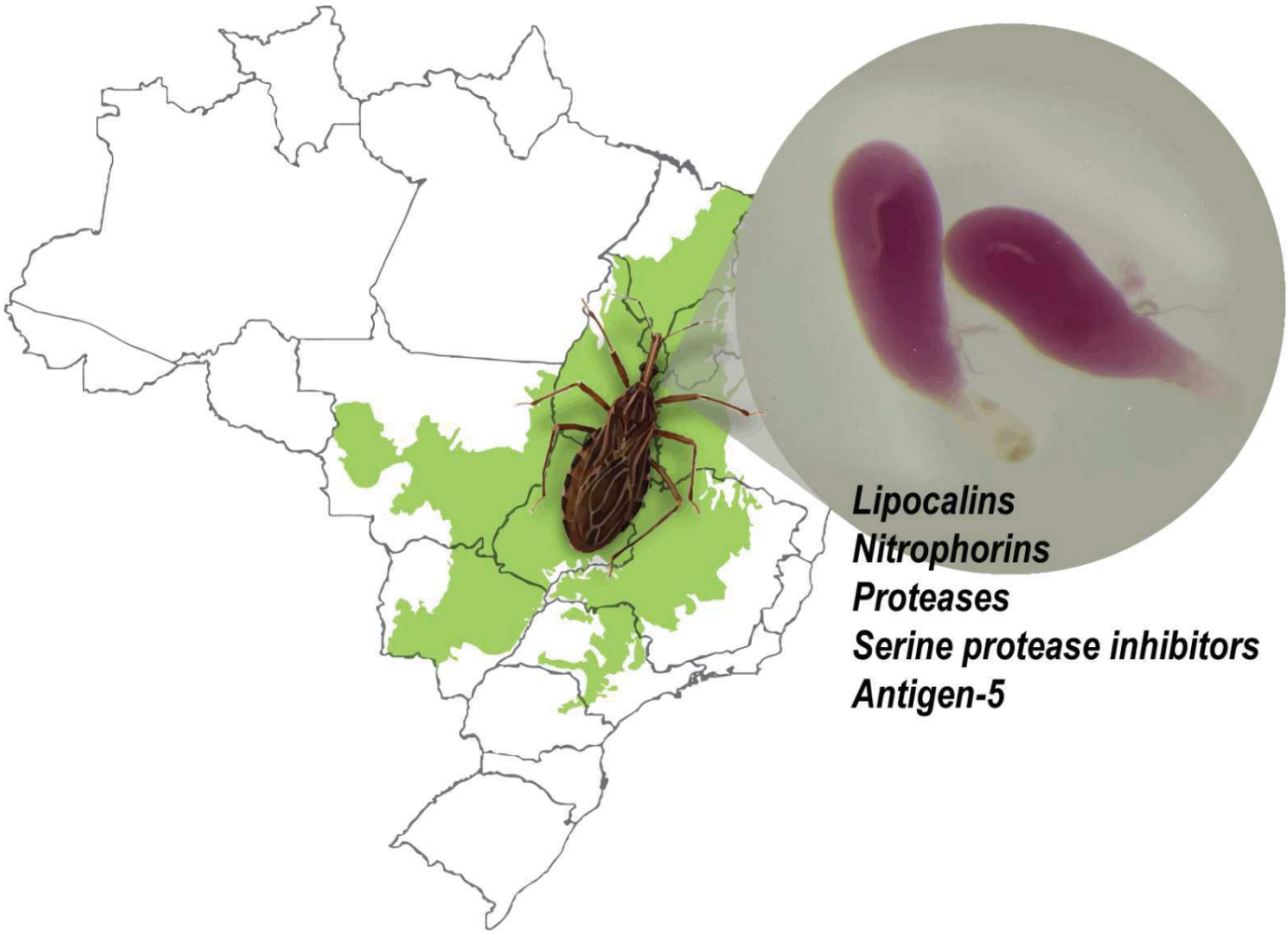
Sup 1. Antigen-5 proteins from *R . neglectus* SG transcriptome.
ClustalW alignment of antigen-5 members from *R. neglectus* salivary transcriptome (RN_21331, RN_21434 and RN_21812) and other hemiptera sequences, identified as described in Methods section. The alignment indicates conserved domains in black and similar domains in gray background.

Sup 2. The dipetalogastin from *R. neglectus* SG transcriptome.
ClustalW alignment of a dipetalogastin member from *R. neglectus* salivary transcriptome (RN_1079) and other sequences from dipetalogastin family, identified as described in Methods section. The alignment indicates conserved residues in black and similar residues in gray background. The boxes indicate conserved motifs, and the blue bar indicates the signal peptide indicative of secretion.

Sup 3. Pacifastin members from *R. neglectus* SG transcriptome.

ClustalW alignment of pacifastin members from *R. neglectus* salivary transcriptome (RN_17301 and RN_20047) and other sequences from the pacifastin family of proteins, identified as described in Methods section. The alignment indicates conserved residues in black and similar residues in gray background. The bars indicate the four conserved pacifastin motifs.

Sup 4. Phylogram of the secreted OBP family from *R. neglectus* SG transcriptome. Phylogenetic tree was built from the alignment of *R. neglectus* CDS and other OBP sequences as described in Methods section. The bar represents 20% amino acid substitution.



Striking still image

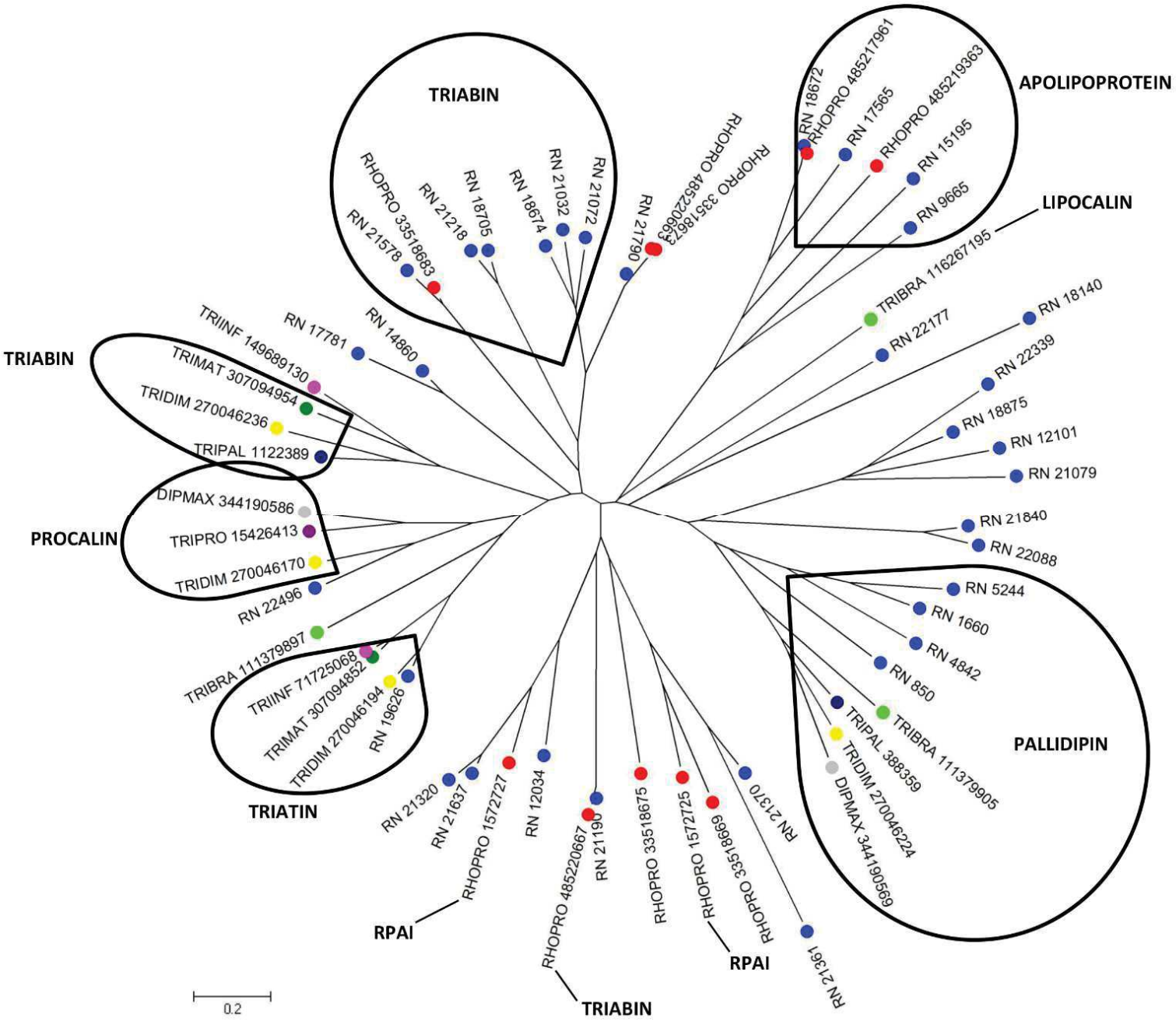

TRIABIN
RHOPRO 33518683
RN 21578
RN 21218
RN 18705
RN 18674
RN 21032
RN 21072
RN 21790
RHOPRO 485220663
APOLIPOPROTEIN
RN 18672
RHOPRO 485217961
RN 17565
RHOPRO 485219363
RN 15195
RN 9665
LIPOCALIN
TRIBRA 116267195
RN 22177
RN 18140
RN 22339
RN 18875
RN 12101
RN 21079
RN 21840
RN 22088
PALLIDIPIN
RN 5244
RN 1660
RN 4842
RN 850
TRIBRA 111379905
TRIPAL 388359
TRIDIM 270046224
DIPMAX 344190569
RN 21370
RHOPRO 33518669
RHOPRO 33518675
RN 21190
RHOPRO 485220667
RN 12034
RHOPRO 1572727
RN 21637
RN 21320
RPAI
TRIABIN
RPAI
TRIATIN
RN 19626
TRIDIM 270046194
TRIMAT 307094852
TRIINF 71725068
TRIBRA 111379897
RN 22496
TRIDIM 270046170
TRIPRO 15426413
DIPMAX 344190586
PROCALIN
TRIPAL 1122389
TRIDIM 270046236
TRIMAT 307094954
TRIINF 149689130
TRIABIN
RN 14860
RN 17781
0.2

Fig 1

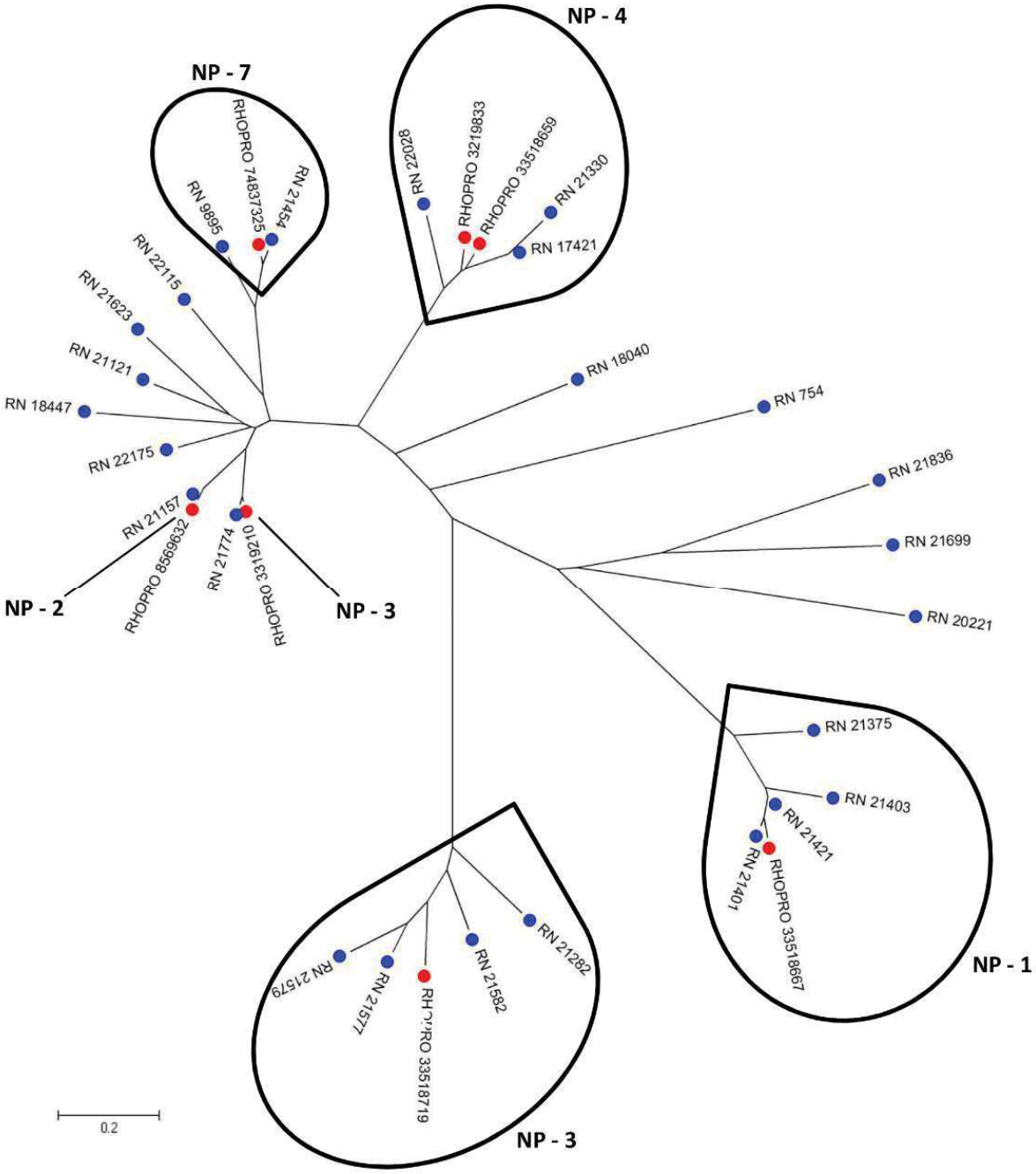

NP - 4
NP - 7
RHOPRO 74837325
RN 2145
RN 9895
RN 22028
RHOPRO 3219833
RHOPRO 33518659
RN 21330
RN 17421
RN 22115
RN 21623
RN 21121
RN 18447
RN 22175
RN 21157
RHOPRO 8569632
RN 21774
RHOPRO 3319210
NP - 2
NP - 3
RN 18040
RN 754
RN 21836
RN 21699
RN 20221
RN 21375
RN 21403
RN 21421
RN 21401
RHOPRO 33518667
NP - 1
RN 21579
RN 21577
RHOPRO 33518719
RN 21582
RN 21282
NP - 3
0.2

Fig 2

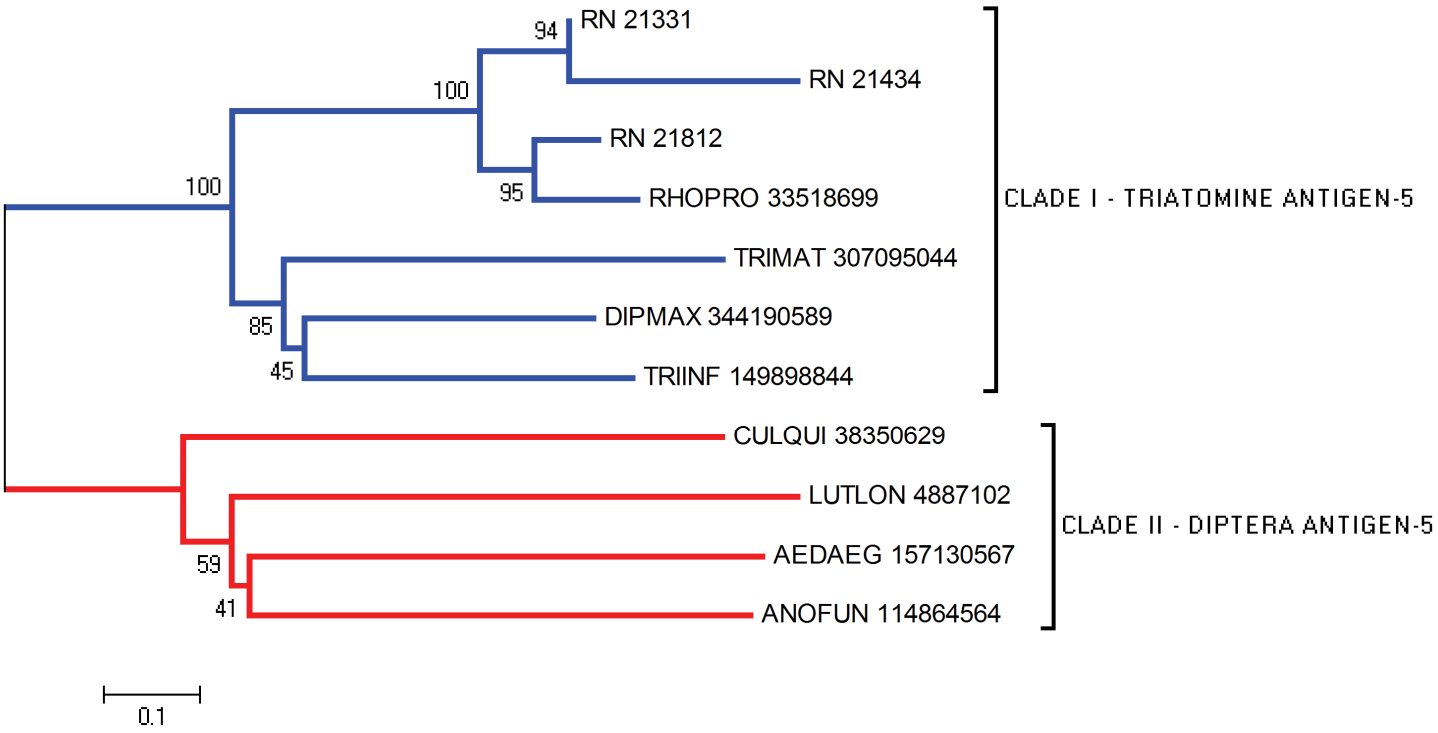

94
RN 21331
RN 21434
100
RN 21812
95
RHOPRO 33518699
100
CLADE I - TRIATOMINE ANTIGEN-5
TRIMAT 307095044
85
DIPMAX 344190589
45
TRIINF 149898844
CULQUI 38350629
LUTLON 4887102
CLADE II - DIPTERA ANTIGEN-5
59
AEDAEG 157130567
41
ANOFUN 114864564
0.1

Fig 3


RN_5563
RN_549
TRIINF_149898876
TRIINF_149898841
PHLPAP_451935058
TRIBRA_116267193
TABYAO_241914365
HIBBIM_94730670

Fig 4

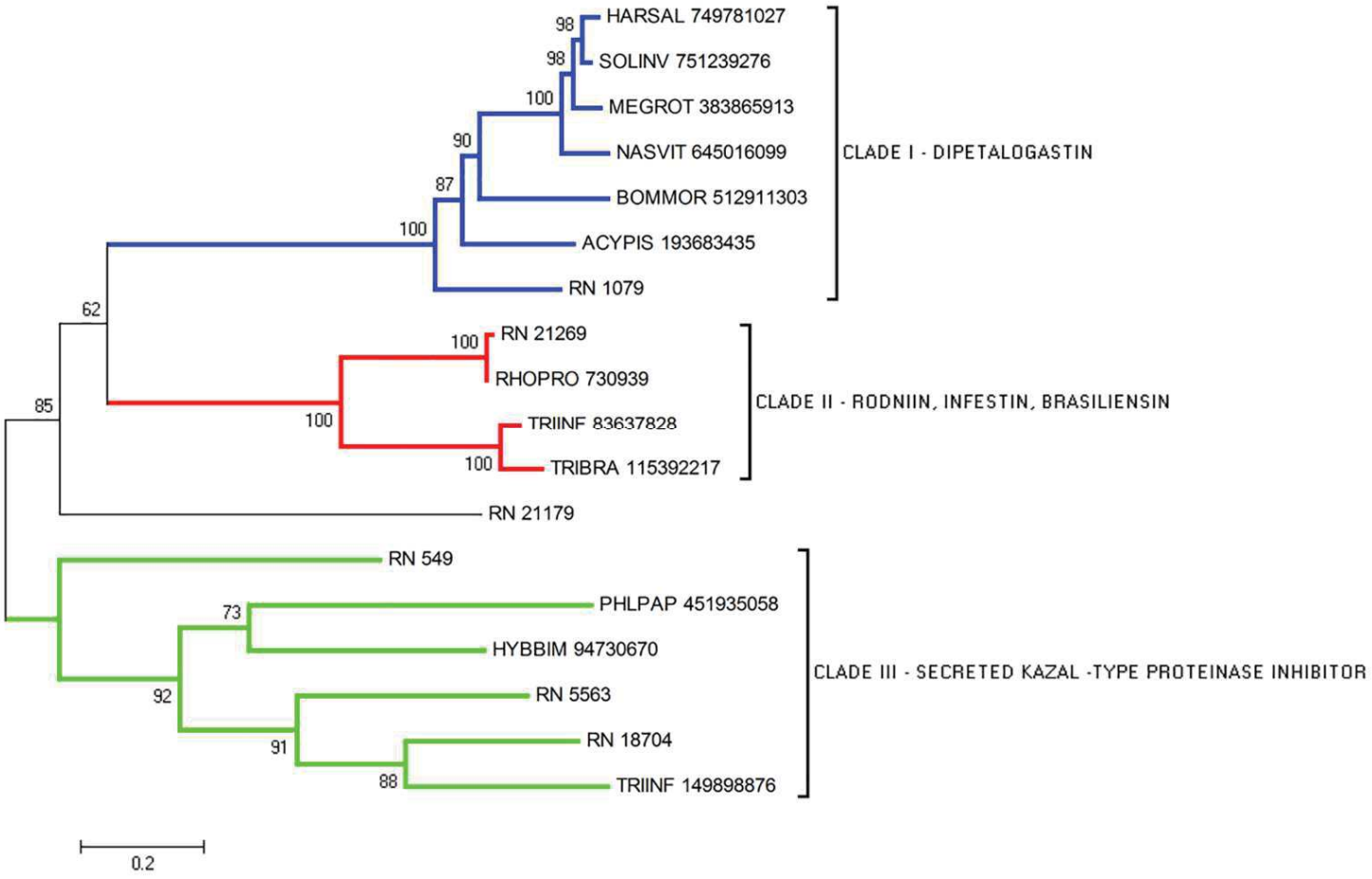

HARSAL 749781027
SOLINV 751239276
MEGROT 383865913
NASVIT 645016099
BOMMOR 512911303
ACYPIS 193683435
RN 1079
CLADE I - DIPETALOGASTIN
RN 21269
RHOPRO 730939
TRIINF 83637828
TRIBRA 115392217
CLADE II - RODNIIN, INFESTIN, BRASILIENSIN
RN 21179
RN 549
PHLPAP 451935058
HYBBIM 94730670
RN 5563
RN 18704
TRIINF 149898876
CLADE III - SECRETED KAZAL -TYPE PROTEINASE INHIBITOR
98
98
100
90
87
100
62
100
100
85
100
73
92
91
88
0.2

Fig 5

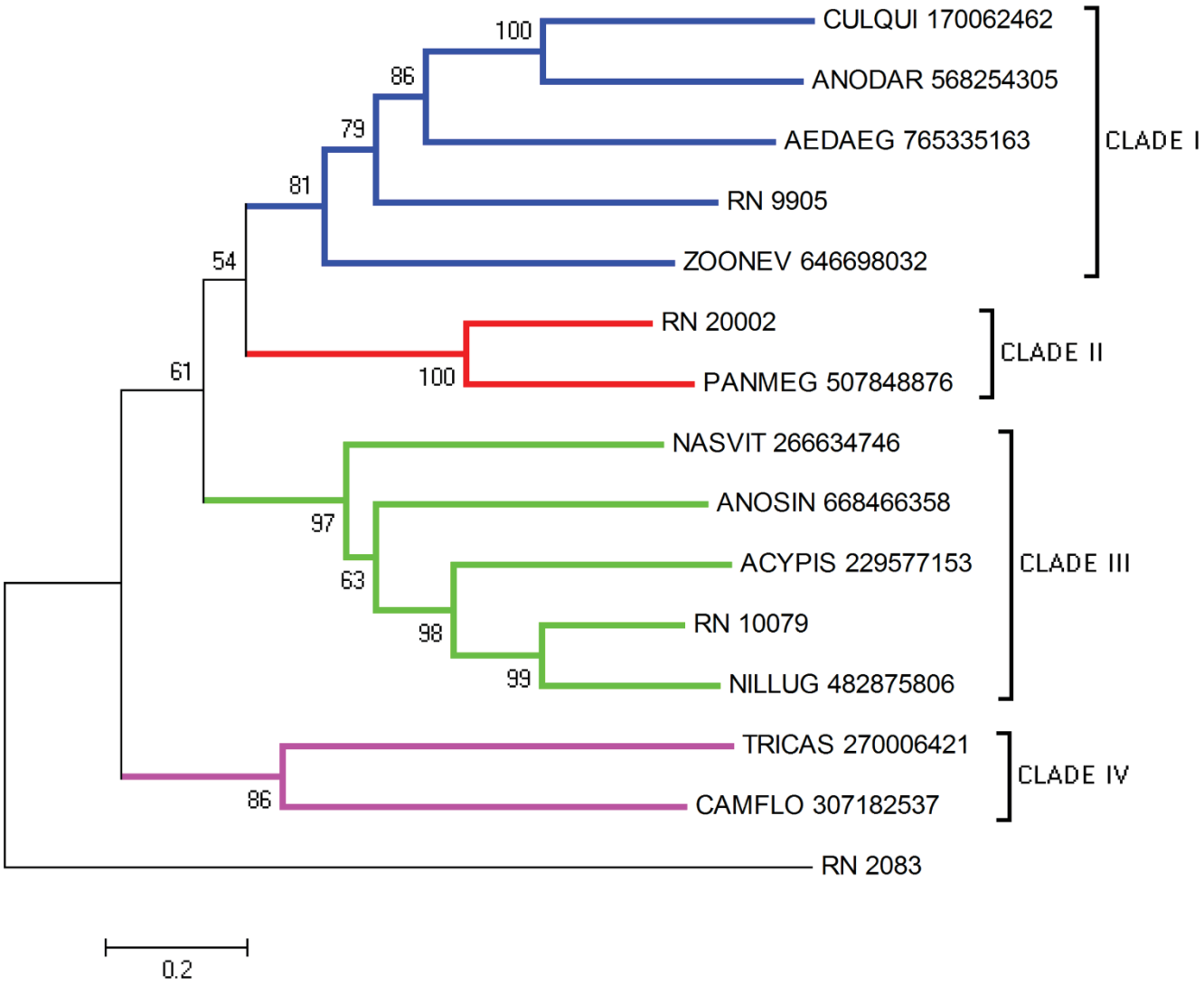

CULQUI 170062462
ANODAR 568254305
AEDAEG 765335163
RN 9905
ZOONEV 646698032
RN 20002
PANMEG 507848876
NASVIT 266634746
ANOSIN 668466358
ACYPIS 229577153
RN 10079
NILLUG 482875806
TRICAS 270006421
CAMFLO 307182537
RN 2083
100
86
79
81
54
61
100
97
63
98
99
86
CLADE I
CLADE II
CLADE III
CLADE IV
0.2

Fig 6

A

```
RN_21266          -------MSLTILIALILATALPRLASGKSIEKPGYELVCPQDENEC---FIKGQEKIDETKLIEGDMVP------------------DEIQQKGNDK 70
ANOSIN_668455453  ------------------------------------MKLA------LFLLAVSSWRSDGVAIERSYFESYLREPPPPTQKDSEK--------HSGGK 47
STEMIM_675388609  -----------MLQKAFILFSLLCISSGKGFVE---------DDD----------SIMENPNLIDGDIIP------------------EDDN---DNR 47
ANODAR_568249911  MPTKMRWKRGWRYYVECYENVLAMHVPGQLDTGYRLDVMRFSSLALGVLLVTCVPVLRTTASSMDRSYFENHYVDPPPAPPKHTE----------DGR 88
ZOONEV_646720098  ------------------------------------------------------------MP-------------------EGR 5
ACYPIS_641648693  ------MKVFLWFVAVCLPLALTQPVNLRAEEEEENEVDNSSDSDLYNGLMHLGAAIFGTPDVKSGEAVSTWKETSQLNPEEMGEYAEGDILHPIGNAR 93
CULQUI_170055406  -------------------------------------------------MFRALVLTTFVVVHCDSSFYESYFNEPVAPNT--------------QTR 35

RN_21266          NALIDPKRKWPDCKIHYAFHHGFTADEKSFIEEQMRFLENKT-CIKFVQSKPEKKANTIVYITNG-DGCVSTVGYAKKQRR--FLNLNQAHCLSKAGRV 166
ANOSIN_668455453  NALRLEYYKWPNCIVPFLFSHGFSLNEQAIILDAINVLQQQT-CVYFIPKTPEQRE-HIRFVKS-DWGCGSSICYRRGQVEPLDVTLDD-FCLGLPGAV 143
STEMIM_675388609  NAIVEESKLWSMCRIPYVISSNLT-DQIPLIKQAMENIENNT-CIQFVPRTIEMDY---INIVTG-DGCSAKVCRGSGGIR--KVTLGDG-CY-KIGRV 137
ANODAR_568249911  NALRLEYYKWPNCVVPYIISEGFSAQEQAAILEAMNVLQQQT-CVYFIPKTADQRE-HIRFVRS-QWGCGSAICFRRGQSEPLDVTLDT-FCLELSGAI 184
ZOONEV_646720098  NALLVKEYLWPSCVVPYVFHSNFTEDEKAKVKAGMKGIQEKT-CVKFVPHTS-EAD-YIEFRKDPQLGCGAMVCRRPGRGFPMAVNYQAPECLQTTGTI 102
ACYPIS_641648693  NGLKAVSSRWPKCVIPYEISPSFGSSDRQLILSSIDEYKKLT-CLKFTPRTSYDTD--YIYFTNGNTGCWSSVC-KIGGRQ--EVNLQSPGCLSKKGTV 187
CULQUI_170055406  NALRLAAYKWPNCIVPFTFDPGCDQRYRTAVLNAISVLEQAASCVHFIPKTDDQTE-HIRFVRS-TSGCGSSICYRSGQREPLDVLLDD-FCLGLAGAV 132

                  ** * * ** * ***                                                                    ***
RN_21266          KHEMLHVIGLLHEQSRPDRDEYVTILEENIEPGKENNFAKADPSEYTTFGVPYNYLSVMHYPALAFSK--------DKKS--PTILPKLPVNPLLLGQRI 255
ANOSIN_668455453  QHELLHVLGLFHEHTRPDRDDYVEVLWENIEPAYRHNFAKGSWDYMETFGLPYDYGSVMHYGNYAYAKAGASVTMVSRINSSAPLGQTDGASFYDLQKIR 243
STEMIM_675388609  IHELTHAIGFYHEQSRSDRDDYINIHWENIKPGRENNFAKLKPYQNALLG-PYDYGSVMHYGPLSFSK--------DKKKGLKTMSPKK--EGIILESP- 224
ANODAR_568249911  QHELLHVLGLFHEHTRPDRDEYVEVFWDNIEPEFQQNFVKGSWEYMETFGLPYDYGSIMHYGSFAFAKAGTSVTMVSRWNSSYPLGQTDGASFYDIQKVR 284
ZOONEV_646720098  QHELLHVLGLFHEQARPDRDNYVTVLWDNIIPEFKNNFVKAPDDVATTYNVPYDYKSIMHYGINTAYSKNGK-NTIVAKNDTSLILGQVEGPTEGDIKKIR 201
ACYPIS_641648693  MHEMLHAAGFMHEQNRPDRDKFVTVNYNNIQSGRENNFEKAQSSMIDSQGIGYDYRSVMHYGANAFSKNGQ--ATIDPKTRGVTMGQREGLSRKDIQKIQ 285
CULQUI_170055406  QHELLHVLGLFHEHTRPDRDEYVEVLWDNIEPEFRRNFARGTEDYMETFGLPYDYGSVMHYPTFAFARPGTSVTMVSRQNRSAELGQTDGASELDLEKVR 232

RN_21266          ------------------------------ 255
ANOSIN_668455453  YMYPCP------------------------ 248
STEMIM_675388609  ------------------------------ 224
ANODAR_568249911  YMYPCP------------------------ 289
ZOONEV_646720098  KLYNCINAQESTLILPWVFKWLSSKKNVKL 230
ACYPIS_641648693  KMYNCKSKRQQSSSSID------------- 301
CULQUI_170055406  RMYDCPRGK--------------------- 240
```

B

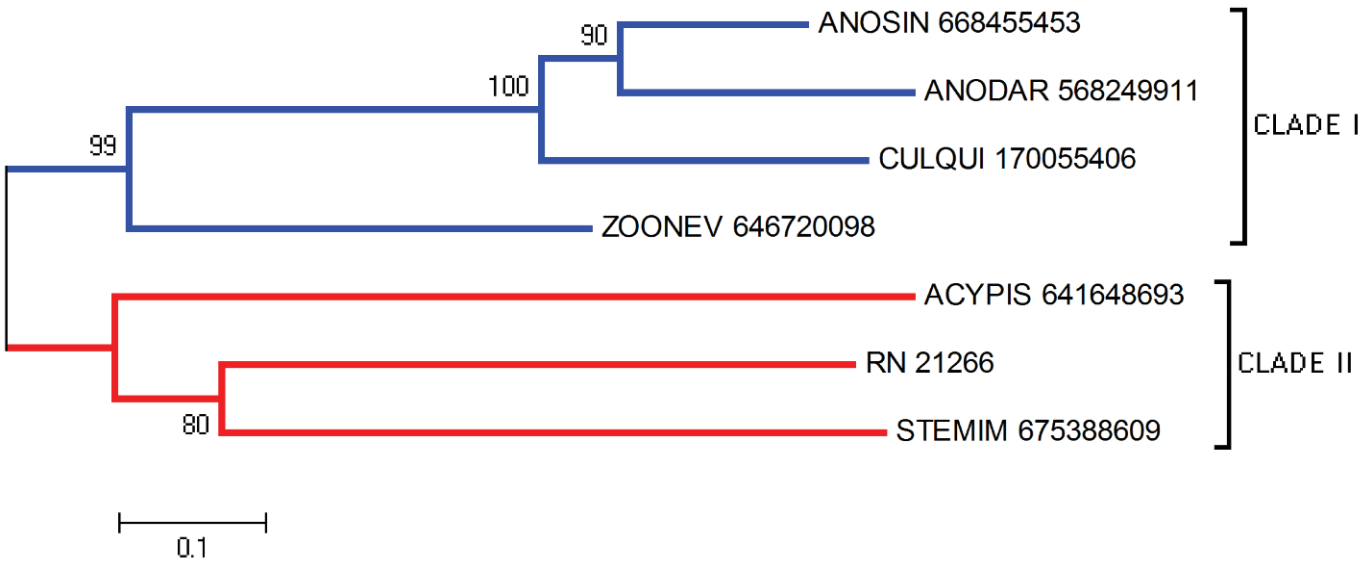


Fig 7

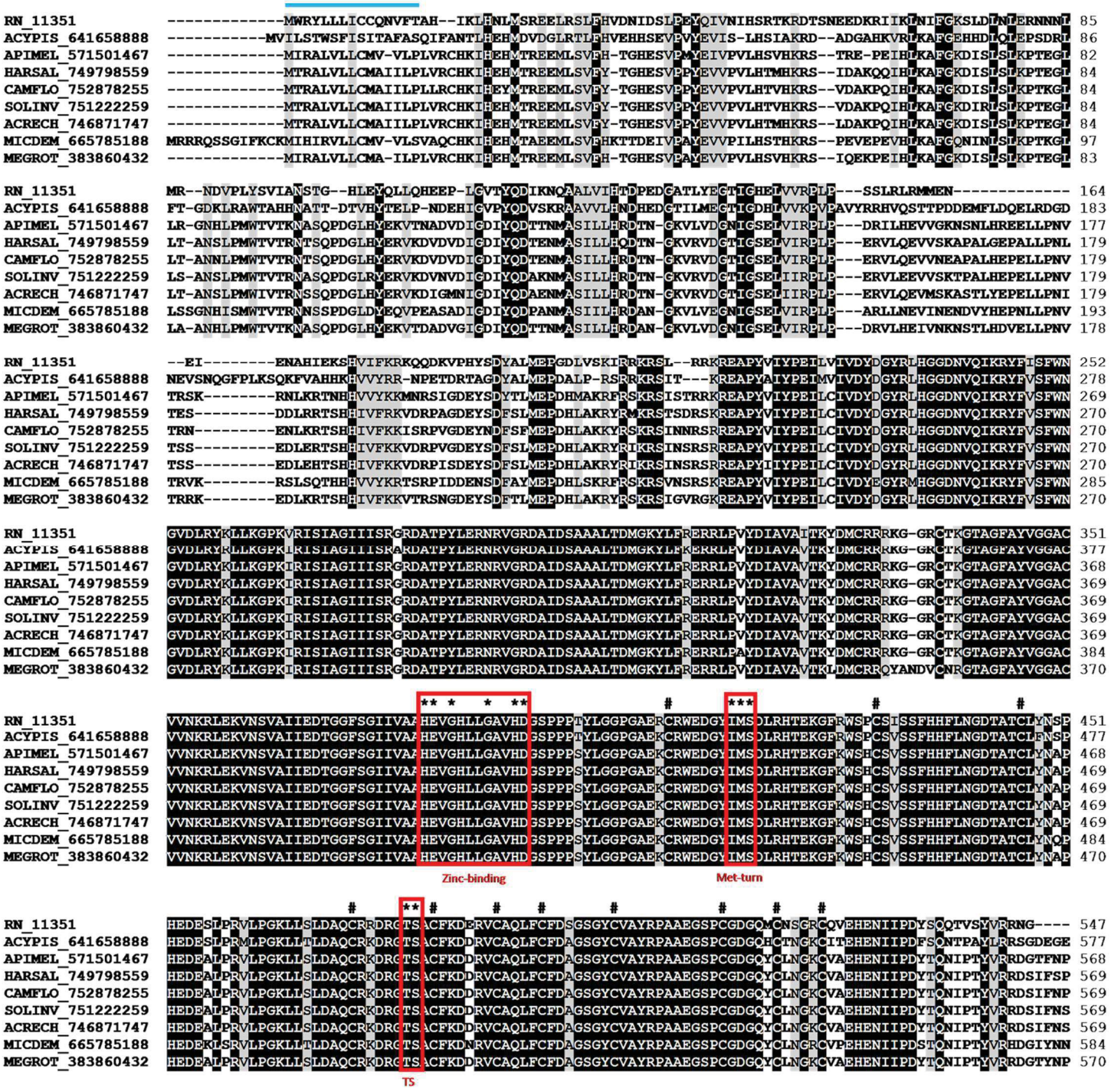


Fig 8

A

```
RN_1189           ----------MDFYHKFIYAATTLLSLWETTRSQNQ-CIYSQYITPCQRYRLNNPSFPNEYPPNLNCSWQMSTRSGYRVILYCQNVKLPVNGNRCLDYLA  89
BOMIMP_350403205  MLSIRTAAFHGFMIQAAVLLFCWSIFSIELSNAANSNCTYYHELTPCSIYFIQNREYPNFYSRSQSCSWIIVS--EYKVDLTCSTFELPWSSNGFQDKVS  98
MEGROT_383850435  ----------------MLSLCWLTPFVKADETTETNCTYVHNLTSEEDYYIYNPEFPKPYRGPQSCSWTINS--DYRVNVVCKLFELPSGFRCLLDNLT  81
CAMFLO_752893391  ----------------------------MSNGLFQN-CNYSQHLKPCNIYYVYNPDYPNVSKGRQFCRWTAES--DYRMKLICNNFNIPWNFKCALDKLT  69
SOLINV_751239412  ----------------------------MSNGLFHQSCNYSQYVKPCSIYYIYNFDYPNVSYGRQHCKWIAES--NYRIRLTCKNFNIPWNFRCALDKLT  70
TRICAS_642928714  -------MWCGNLYRTMGFAALFNFLMP-VSMAIDSNCDFYQDMELCQNYYIYNPGYPNYTGPPIACRWVGKSPVGTVIVLTCEDIDISTSYNCYSDRLS  92
ANODAR_568259905  ------------------------------------------MKFCSTYQLVSPRYPQNYPANIECTYTIRAPYGYRIAVDCPTLQIPSSTNCRLDSLK  57
NASVIT_239050294  -----MAFISKPLKVCCLFGLLLLLFSFEASDAVDANCDFLQKLELCKTYFITNPNYPNNYTASRSCRWHATS--DTRVRITCNEFNIPASPNCEADTFN  93
VESVUL_260670832  ----------MKLDNFFLILYVLLSIIKSNGTIDLNCGYTQKLKSDVNYCVYNPDFPNYYMGEHNCRWSARS--NTRIKLNCTVFDVPPSSNCSLDFMK  87
ANODAR_568259906  ----------MGMKLVFGFVLLIISTIPDGARAQYSGCDYSIDVASCLSYTIYSPNYPGYYQPNSQCRWTLSTTQGSRIALSCSEIAIPSSSSCATDRLV  90

RN_1189           VTTT-------GDLRDSTTYCLPFSATGTEIRLQFVSSVWSWGCIESCDVTSEMIP--VIEANHTKQCGCGWSFSTRIVGGTDACINEFPSMALAVH-VP  179
BOMIMP_350403205  VQVN-----SSTTHRYCCNGNFNIQSESNTMVVTLQSPVWSRGCRFACEVRSVKR------PQDFEDCECGWRNPSRIVGCNDTCVNEFPMMVGIVD-FR  186
MEGROT_383850435  VEAS-----SNSIHRFCCNDAVNVTSDGPTMVVKLPASVWSTGCRFLCDVMAVKK------PSDAEGCECGWRNPNKIVGCMDACVNEFPMMAGLVD-FR  169
CAMFLO_752893391  VYVN-----QSTAHSYCCNKTFSVESTGKGMAIELSTTFWSLGVKFLCELQTIEEE-----PWDDRNCRCGWKNPTKIVGGMETCVNEYPMMAGLVD-PS  158
SOLINV_751239412  VYVN-----DSASQSYCCNETFSIESTGEFMTIELSTSFWSSGVKFLCELQAVDEI-----END--NCRCGWKKPTKIVGGMETCVNEYPMMAGLVD-YF  157
TRICAS_642928714  VSRSGDENLQKQTTNYCCKTSTTVLTEDNAIAVGFFLTNPMYGGRFSCSLTPMQTSPPTTTTARPNTCDCGWKKGTRIHGGHETGINEYPSMAAMVD-RW  191
ANODAR_568259905  VSRSGRLDFSDAYV-YCCVGSLQEKSTGNSLTMRLTSAPTSFGCKFSCTLSIPEQN----------CECGRKKTQRIVNGVETVVNEFPMMAALID-VK  144
NASVIT_239050294  VYTDP----VSGPRRYCCSGDFSRDSTGKEMIITLTSANNSRGGTFRCFFEAIEE----------EECRCGWKKPTKIVGGRETCINEYPMMAGIIN-VP  178
VESVUL_260670832  VKVD-----DDIEYVFCCLNSFAVESTASKMTIKFHSRYNTYGGKFRCNLRSVKE----------KCRCGWKNPSRIVGGVETCVNEYPMMAGIIH-LA  170
ANODAR_568259906  VSKNGQSNLSDGQA-YCCTGVLNIESTYNSLVIALQVAGTSYGGRFYCTANKIE----------CQCCMRRKKKIVNGIETLVNEFPMMAALVDGSG  176

RN_1189           RRVQWCGATIIARTVALTAGHCFITD-SNPHNYGLLVCEHDTSTRAETNATKLLKISKIIVHPYFHPSSDGNDIALMVTETMIEYSSRVGPACLPFNLDK  278
BOMIMP_350403205  RRVVFCGGTIISKRYVLTAGHCVVD--REYQDLGALVGDHDLRTGSDTNATVLHRIIKIVIHPNWGQRDDQNDVAVVKTENEIKFTNEVGPACLPFQHSS  284
MEGROT_383850435  QRNITCGSTIISNRYATTAAHCFKNG-TEPRDYGLLVGDHDLRTGTDTNATVLHRLRSITIHPDVRKNPDVNDIAIVKAELEIKFSNQVGPACLPFQHSP  268
CAMFLO_752893391  QKDVYCGGTIISERYILTAAHCLDN--KNTSNVGILVGDHDLSTGADTNASRLYTVSRFDIHPLVNNESLENDIAIITTNSPISFSEKVGPACLPFQHQS  256
SOLINV_751239412  ERDVFCGATIVSERYVLTAAHCLTD--RNISTMAVVVGDHDLSTGTDTNASRVFTVSRFDIHPFYDSETLENDIAVVMVNSVISFSEEVGPACLPFQHQS  255
TRICAS_642928714  TFDAFCGASIISDRYALTAAHCLLH--KTPDDFALLVGDHNMTSGDDTPYAAVYKISNMFSHPSYDQSTQLNDIAVLQTEKPIEFSLFVGPVCLPFRYTS  289
ANODAR_568259905  TKTVICGATIITDNYALTAAHCLLQ--RTTNDTVPLVGDHNIRSGSDTSFAQIYVVVQFLSHYGFTLKPVANDIALVRTERAIQYNAAVGPACLPWRYTR  242
NASVIT_239050294  IQQVYCGGTIISPKHILTAAHCLNK--LAVNDLGILVGDHDLTTGSETNATKLYRAASYVIHPSYVSNKKDYDIAVITIAGTLTYTNEVGPACLPFQHYL  276
VESVUL_260670832  TRFLYCGATIITPQHVLTAAHCVARYKRILYILGVVVGEHNTWAINDTKATQLYLIDDIIVHPNYR--PKLNDLAVIKLQKRLKYSMRIGPACLPFYYMQ  268
ANODAR_568259906  GEGIFCGATVITNYHAVTAAHCYMG--RSLPNLGLLVGDHNITSGSDTPYSALLLLSSIKLHESYNPTIGTNDIAIVRTRAEIMFNTGVGPVCLPLKYYG  274

RN_1189           NYLLRKYVIVTGWGFLENGCPISNKLQKISLQIIPDSNCTQVYDYKFKPEKHLCAYSPGKDACQFDSGGPLYFYKYLTNLYTVVGIVSYGSGCASADPSL  378
BOMIMP_350403205  DTGGGNYVELLGWGTTDFGGPTSDILQKVILSVLTYLQCSRNYN---VTTDQICTYAENKDSCQMDSGGPVLWQNPTTRRLVLIGIIAKGRSCA-VVGGV  380
MEGROT_383850435  DTGGGSYVELLGWGLTEFGGLPPTVLQKVSLSVLTNLECSKTYQN--VTSSRICTYGQGKDSCQMDSGGPVLWQNPTSRRLVLIGVINQGIDCG-IFPSL  365
CAMFLO_752893391  DSGAGSYVDLLGWGTTEFGCMKSKTLQKVILNVITNRDCRQSDPN--LLYSQLCTSGEKKDACQFDSGGPVLWQNPTTKREVLVGIISYGNGCGSNEPGI  354
SOLINV_751239412  DSGAGSYVDLLGWGTTQFGCMKSQTLQKVTVTVMTYRECRRQYPN--ASYSQLCTYGEGKDACQFDSGGPALWQNPTTKREVLIGIISSGFGCGNDKPGI  353
TRICAS_642928714  VNGLSQTVTALGWGFVDVAGPKSDTLQEVDLTVVSTEECNATITDNPVTYRQICTYAPNRDACQSDSGGPILWQDPNTRRLQLIGIISYGIGCATSRPAV  389
ANODAR_568259905  QSGEGLQLEATGWGDLDFGGPRSDTLQKVQLSVIANSQCSQRLGS-TVPYQQLCTFTPDRDICQADSGGPLFYTNPSTGLLYNVGLVSFGFACATERPSV  341
NASVIT_239050294  DSGGGSFVDVLGWGTTEFACAPSNTLQKVRLSITNFLSCKSYFQN--LEYRQICTYAEGKDACQFDSGGPVLWQNPTTRRLVLVGIISYGSFCASEEPAV  374
VESVUL_260670832  RNGVDTVVTAVGWGLTNFYGVKSEVLRKVDLHVVSMKKCVKYHFL--ATPKQLCTFDMGKDACQFDSGGPILWQNPKSKRIFLLGVINYGRTCADEAPGV  366
ANODAR_568259906  ASGVGIELEAVGWGTLDYGAPKSNVPMKVGLTVVSGAQCSAIYPN-FSVTQQLCTLTANKDSCQSDSGGPLFYTDAYNGRVYLVGIISYGYGCIACASNQPSV  373

RN_1189           NINVHYYLPWIFSQLPDVQFCHG-  401
BOMIMP_350403205  NARVGAYIDWIISATSDASYCIIE  404
MEGROT_383850435  NMRVGAYTDWIVSVTSDASYCISE  389
CAMFLO_752893391  NTRVGTYIDWILHVTSGTSYCKTE  378
SOLINV_751239412  NTRVGTYIDWILYVTPGTNYCMSE  377
TRICAS_642928714  NTRVTSYLRWIVSVTEDAFYCIK-  412
ANODAR_568259905  NTRVTEFLDWIMANSATSFYCFR-  364
NASVIT_239050294  NTRVGSFVDWITSVTGASYCQNE  398
VESVUL_260670832  NLRVTSYLDFITRSTPGEIYCQAY  390
ANODAR_568259906  NTNVLNYMQWITTNTPEWNYCYQ-  396
```

B

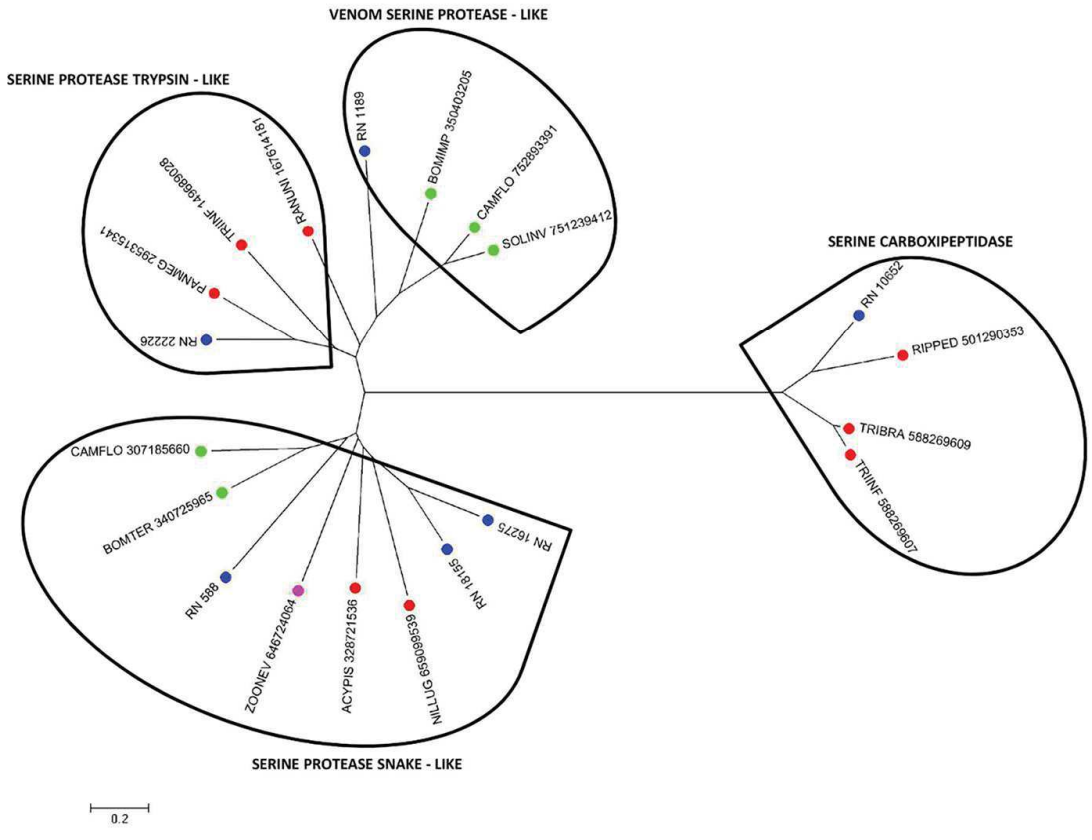


Fig 9

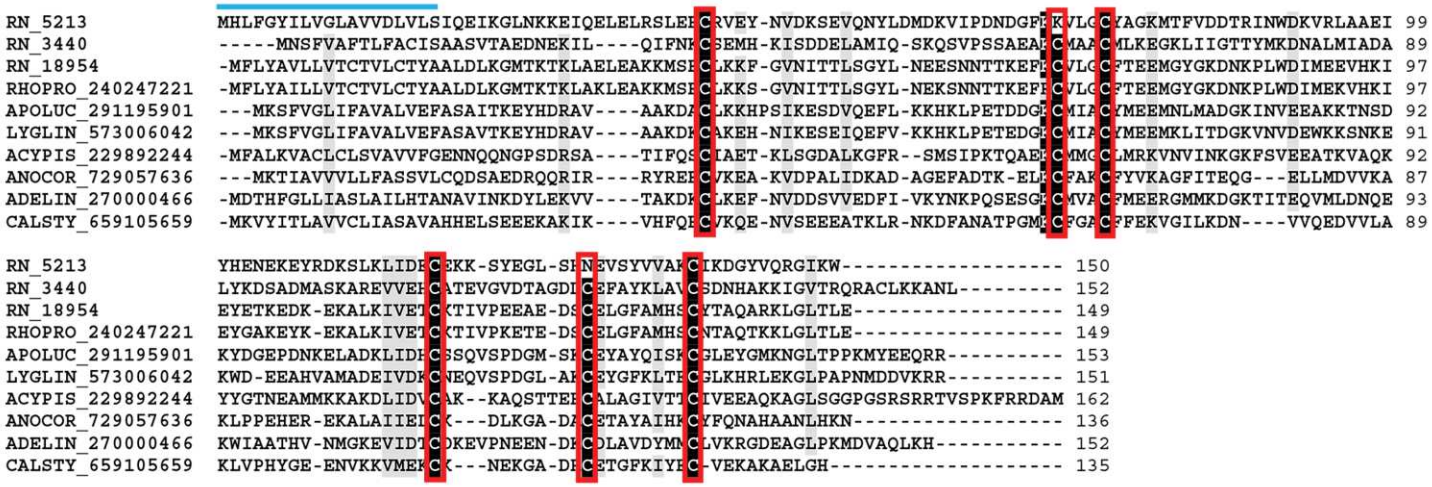
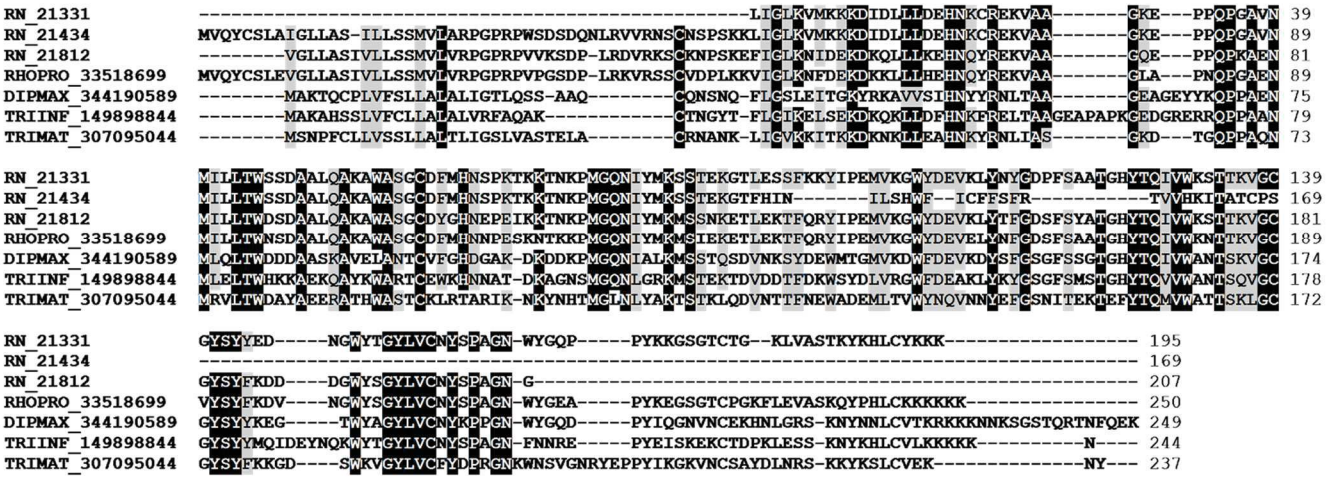
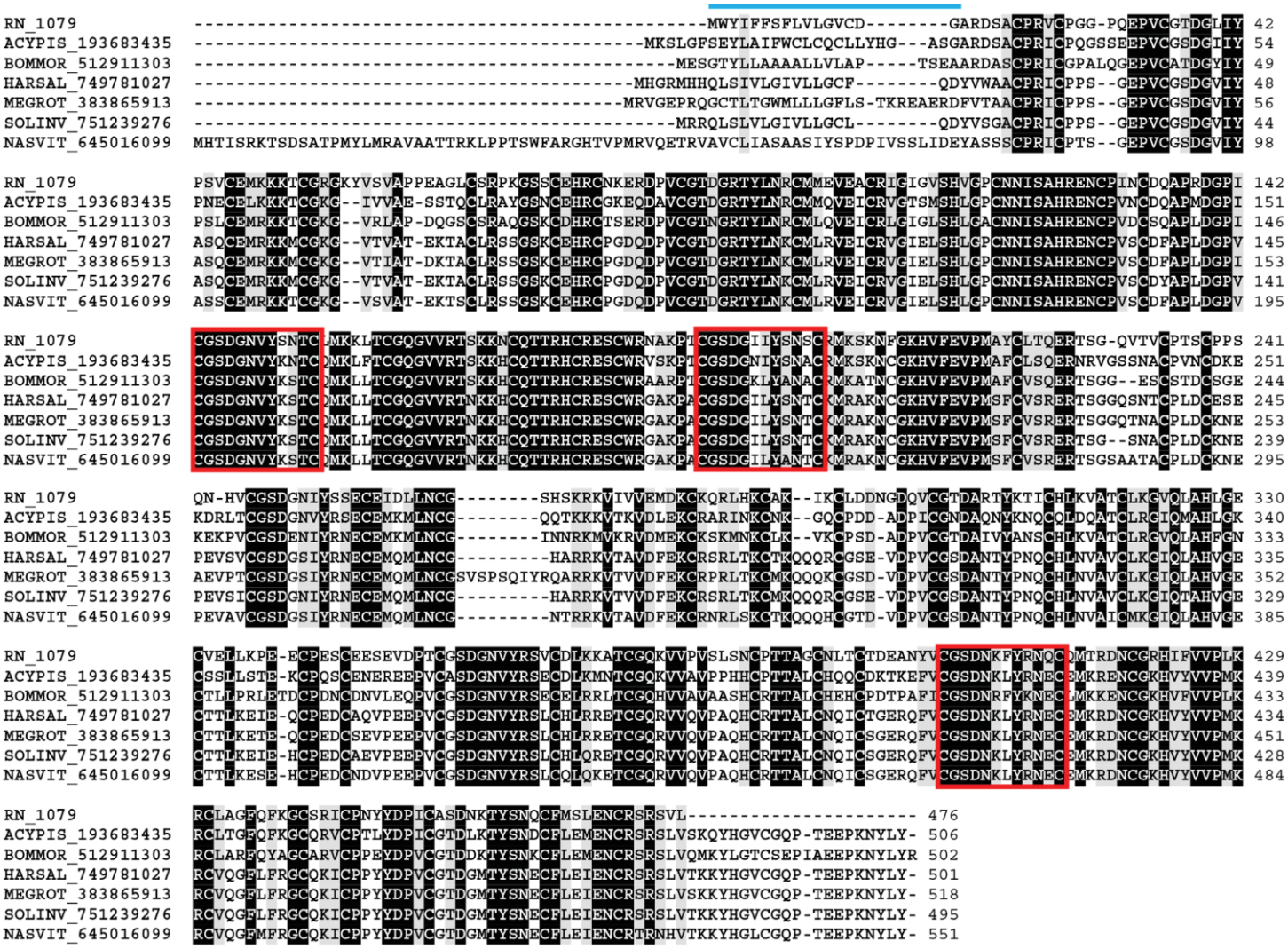

RN_5213 MHLFGYILVGLAVVDLVLSIQEIKGLNKKEIQELELRSLEECRVEY-NVDKSEVQNYLDMDKVIPDNDGFEKVLGCYAGKMTFVDDTRINWDKVRLAAEI 99
RN_3440 -----MNSFVAFTLFACISAASVTAEDNEKIL----QIFNKCSEMH-KISDDELAMIQ-SKQSVPSSAEAECMAACMLKEGKLIIGTTYMKDNALMIADA 89
RN_18954 -MFLYAVLLVTCTVLCTYAALDLKGMTKTKLAELEAKKMSECLKKF-GVNITTLSGYL-NEESNNTTKEFECVLGCFTEEMGYGKDNKPLWDIMEEVHKI 97
RHOPRO_240247221 -MFLYAILLVTCTVLCTYAALDLKGMTKTKLAKLEAKKMSECLKKS-GVNITTLSGYL-NEKSNNTTKEFECVLGCFTEEMGYGKDNKPLWDIMEKVHKI 97
APOLUC_291195901 ---MKSFVGLIFAVALVEFASAITKEYHDRAV----AAKDACLKKHPSIKESDVQEFL-KKHKLPETDDGECMIACYMEEMNLMADGKINVEEAKKTNSD 92
LYGLIN_573006042 ---MKSFVGLIFAVALVEFASAVTKEYHDRAV----AAKDKCAKEH-NIKESEIQEFV-KKHKLPETEDGECMIACYMEEMKLITDGKVNVDEWKKSNKE 91
ACYPIS_229892244 -MFALKVACLCLSVAVVFGENNQQNGPSDRSA----TIFQSCIAET-KLSGDALKGFR--SMSIPKTQAEECMMCCLMRKVNVINKGKFSVEEATKVAQK 92
ANOCOR_729057636 ---MKTIAVVVLLFASSVLCQDSAEDRQQRIR----RYREECVKEA-KVDPALIDKAD-AGEFADTK-ELECFAKCFYVKAGFITEQG---ELLMDVVKA 87
ADELIN_270000466 -MDTHFGLLIASLAILHTANAVINKDYLEKVV----TAKDKCLKEF-NVDDSVVEDFI-VKYNKPQSESGECMVACFMEERGMMKDGKTITEQVMLDNQE 93
CALSTY_659105659 -MKVYITLAVVCLIASAVAHHELSEEEKAKIK----VHFQECVKQE-NVSEEEATKLR-NKDFANATPGMECFGACFFEKVGILKDN----VVQEDVVLA 89
RN_5213 YHENEKEYRDKSLKLIDECEKK-SYEGL-SPNEVSYVVAKCIKDGYVQRGIKW------------------ 150
RN_3440 LYKDSADMASKAREVVEKCATEVGVDTAGDLCEFAYKLAVCSDNHAKKIGVTRQRACLKKANL--------- 152
RN_18954 EYETKEDK-EKALKIVETCKTIVPEEAE-DSCELGFAMHSCYTAQARKLGLTLE----------------- 149
RHOPRO_240247221 EYGAKEYK-EKALKIVETCKTIVPKETE-DSCELGFAMHSCNTAQTKKLGLTLE----------------- 149
APOLUC_291195901 KYDGEPDNKELADKLIDKCSSQVSPDGM-SKCEYAYQISKCGLEYGMKNGLTPPKMYEEQRR---------- 153
LYGLIN_573006042 KWD-EEAHVAMADEIVDKCNEQVSPDGL-AKCEYGFKLTKCGLKHRLEKGLPAPNMDDVKRR---------- 151
ACYPIS_229892244 YYGTNEAMMKKAKDLIDVCAK--KAQSTTEKCALAGIVTTCIVEEAQKAGLSGGPGSRSRRTVSPKFRRDAM 162
ANOCOR_729057636 KLPPEHER-EKALAIIELCK---DLKGA-DACETAYAIHKCYFQNAHAANLHKN----------------- 136
ADELIN_270000466 KWIAATHV-NMGKEVIDTCDKEVPNEEN-DKCDLAVDYMNCLVKRGDEAGLPKMDVAQLKH---------- 152
CALSTY_659105659 KLVPHYGE-ENVKKVMEKCK---NEKGA-DECETGFKIYECLVEKAKAELGH-------------------- 135

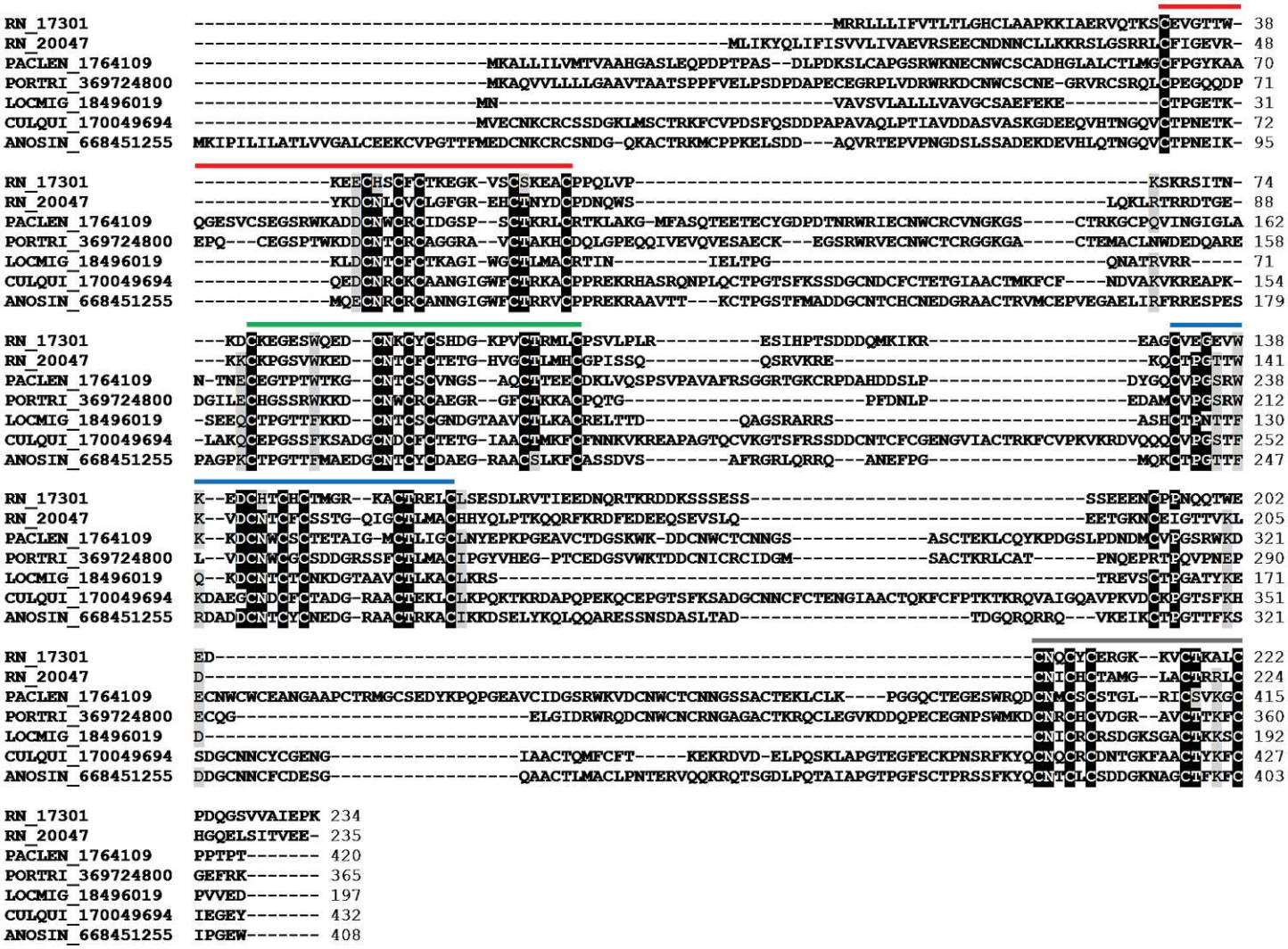
Fig 10


RN_21331
RN_21434
RN_21812
RHOPRO_33518699
DIPMAX_344190589
TRIINF_149898844
TRIMAT_307095044

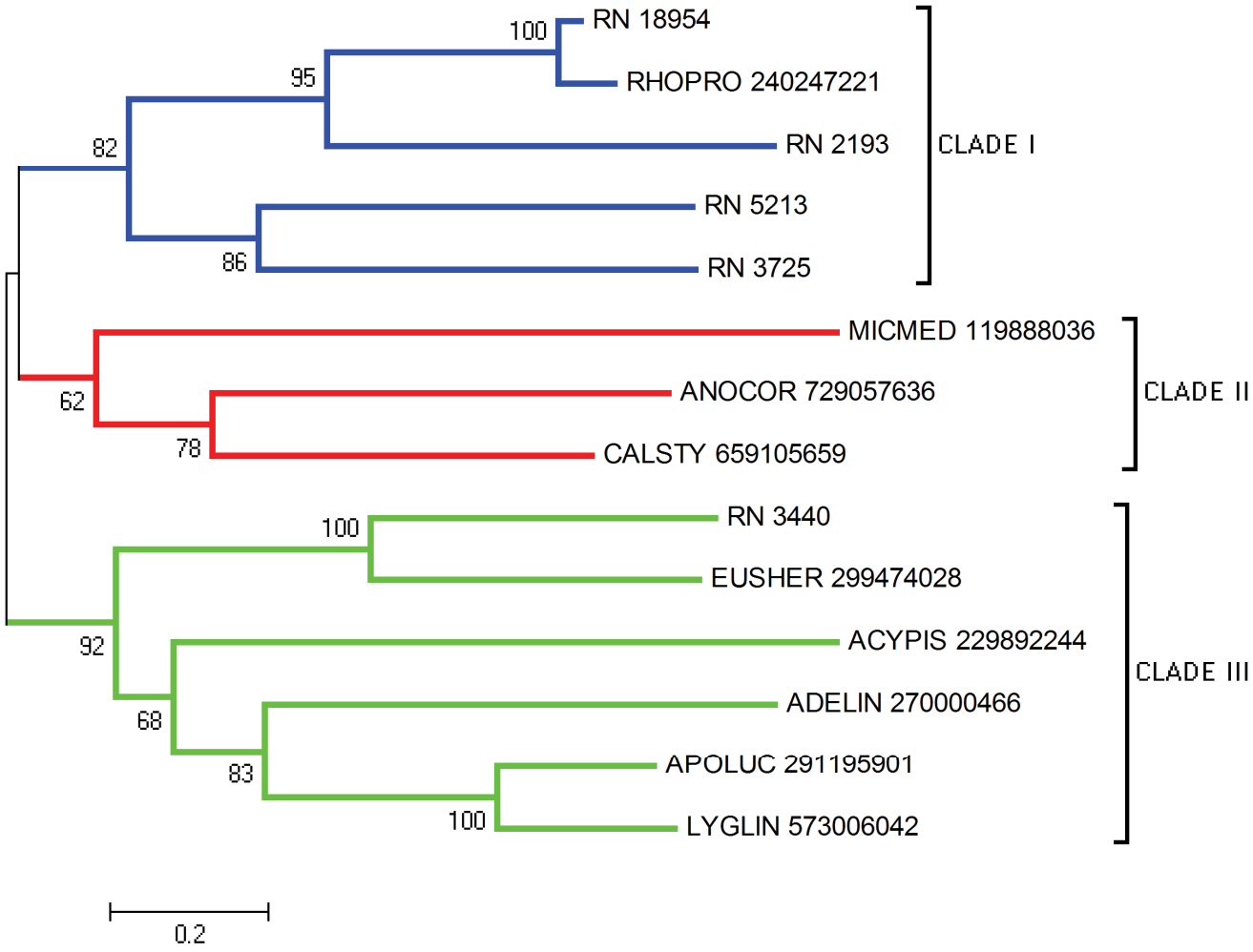
S 1


RN_1079
ACYPIS_193683435
BOMMOR_512911303
HARSAL_749781027
MEGROT_383865913
SOLINV_751239276
NASVIT_645016099

S 2


RN_17301
RN_20047
PACLEN_1764109
PORTRI_369724800
LOCMIG_18496019
CULQUI_170049694
ANOSIN_668451255

S 3


RN 18954
RHOPRO 240247221
RN 2193
RN 5213
RN 3725
CLADE I
MICMED 119888036
ANOCOR 729057636
CALSTY 659105659
CLADE II
RN 3440
EUSHER 299474028
ACYPIS 229892244
ADELIN 270000466
APOLUC 291195901
LYGLIN 573006042
CLADE III
100
95
82
86
62
78
100
92
68
83
100
0.2

S 4

MANUSCRITO 2

A new insight into the sialome of the salivary glands from *Triatoma dimidiata*, a vector of Chagas disease

Paula Beatriz Santiago[a], Teresa C. F. Assumpção[a,b], Carla Nunes de Araújo[a,c], Ionizete Garcia da Silva [d], Rayner M. L. Queiroz[a], Carlos André O. Ricart[a], Tainá Raiol[a,e], João Victor de Araújo Oliveira[f], Eric Calvo[b], José M.C. Ribeiro[b] and Jaime M. Santana[a*]

[a] Department of Cell Biology, The University of Brasília, Brasília, Brazil

[b] Vector Biology Section, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Disease, 12735 Twinbrook Parkway, Room 2E32D, Rockville, MD 20852, United States

[c] Faculty of Ceilândia, The University of Brasília, Brasília, Brazil

[d] Department of Parasitology, The University of Goiás, Brazil

[e] Instituto Leônidas e Maria Deane - Fiocruz Amazônia, Manaus, AM, Brazil

[f] Department of Computer Science

***E-mail addresses:*:** paulabeatriz@hotmail.com; tecris@gmail.com; cnunes@unb.br; profionizete@yahoo.com.br; raynermyr@gmail.com; ricart@unb.br; tainaraiol@gmail.com; joaovicers@outlook.com; ecalvo@niaid.nih.gov; JRIBEIRO@niaid.nih.gov

***Corresponding author:** jsantana@unb.br*

Abbreviations

NGS, next generation sequencing; RNA-seq, RNA sequencing; SG, salivary glands; CDS, coding sequences; S, putative secreted transcripts; H, putative housekeeping transcripts; U, unknown function transcripts; TE, transposable elements; Vr, viral products; Serpin, serine proteinase inhibitors; OBP, odorant binding protein.

ABSTRACT

Chagas disease is a zoonotic infectious disease caused by the parasite *Trypanosoma cruzi* and vector-borne transmitted to mammalian hosts while blood-sucking triatomines take their meal. The salivary glands of hematophagous vectors produce a repertoire of pharmacological molecules that counteract hosts defense responses to assist feeding, including vasodilators, anti-platelet, anti-coagulant and immunomodulators. *Triatoma dimidiata* is an important Chagas disease vector in Latin America, where it is able to invade human dwellings. The dominant transcripts of the salivary glands of this vector were previously described. The aim of the present work was to provide a deeper insight into the salivary glands proteins that contributes to the hematophagic habit of *T. dimidiata.* We used RNA-seq as a tool to sequencing a cDNA library from *T. dimidiata* salivary glands, reporting its expressed transcripts. In addition, to validate the results, the salivary content was submitted to proteomic analysis using LC/MS-MS. More than 51 million reads were obtained and assembled, from where a total of 5,704 coding sequences were extracted. About 87.83 % of total mapped reads were related to putative secreted proteins. Among these, lipocalins were the most abundant gene family products. Other secretory members included antigen-5 family, proteases and protease inhibitors. Lipocalins comprised more than 93.74 % of secreted reads, indicating they are an important expanded family of the salivary repertoire from *T. dimidiata* acting to neutralize host's hemostasis. Accordingly, this work expanded the previous set of annotated sequences from *T. dimidiata* salivary glands available in NCBI, contributing to the pharmacologically active molecules databank and improving our knowledge about the complexity of anti-hemostatic salivary compounds from hematophagous arthropods and its biological interactions.

Keywords: Hematophagy; *Triatoma dimidiata*; Vector biology; Salivary proteins; Transcriptome.

1. Introduction

Hematophagous insects are parasite vectors of debilitating diseases to humans and domestic animals. Blood-sucking triatomine species (Hemiptera: Reduviidae) are able to transmit the protozoan parasite *Trypanosoma cruzi*, the causative agent of Chagas disease, to different mammalian hosts in the wild (sylvatic cycle). Human infection usually occurs through the contact with contaminated feces released after vector feeding. The vector-borne transmission is commonly observed when triatomines adapt to live in human dwellings, possibly in the cracked mud walls in rustic rural houses (domestic cycle). Additional mechanisms of transmission include ingestion of contaminated food, blood transfusion, organ transplantation and congenital transmission (1).

Chagas disease is a neglected tropical disease endemic in Latin America and currently, the control of synanthropic vector populations is the main strategy against transmission. Historically, despite progress in disease control that has been achieved since the 1990s when endemic countries joined in international programs acting against domestic vectors, new scenarios are emerging in some areas, such as domestic re-infestation, resistance to insecticides and sylvatic invasion of urban areas, keeping it a challenge to sustain the disease control (1, 2). In addition, although endemic in rural areas of Latin America, nowadays, due to increased human migration, the disease can be observed in non-endemic areas, affecting about 7 to 8 million people worldwide (1, 3)

Triatoma dimidiata is an important Chagas disease vector widely distributed in Americas, being found from Mexico, through Central America into north of South America (4) (Fig 1). Central America is noteworthy the region where *T. dimidiata* is highlighted as the main vector of *T. cruzi*. The insect shows different phenetic and genetic features throughout its geographical distribution, and can be found in sylvatic,

peridomestic and domestic ecotopes, influencing the effectiveness of vector control activities. Additionally, *T. dimidiata* exhibits variations in domestication rates, in infestation rates and in infection with *T. cruzi* rates, which in turn are factors that directly affect the epidemiological importance of the vector in different regions (4). All these factors make this vector's eradication an extremely difficult process.

Fig. 1: *Triatoma dimidiata* triatomine and its geographical distribution.

All nymph stages and adults of both sexes feed on blood, and at the bite site, in order to achieve a blood vessel, the triatomines insert a long and flexible mouthpart into vertebrate host skin causing tissue injury and bleeding. The host responds with hemostasis (vasoconstriction, platelet aggregation and coagulation), inflammation and immune response, and to obtain a successful meal, these hosts' mechanisms need to be overcome. As a strategy, triatomines release a wide range of salivary molecules acting as platelet aggregation inhibitors, vasodilators, and anticoagulants that make possible a continuous blood flow, counteracting its specific targets in hemostatic response and wounding (5-7). The description of hematophagous saliva composition, the sialomes (from the Greek sialo = saliva), improve the understanding on the dynamics of interaction between the vector and the vertebrate host, as well as how salivary molecules can affect disease transmission. In addition, sialome studies provide a greater insight on the evolution of blood feeding behavior. In this regard, sets of protein families with specialized functions have been found in different hematophagous insect saliva allowing the blood-feeding life style adaptation.

In the last years a number of scientific studies have shown the diversity of hematophagous salivary components through transcriptomic analysis. Sanger made

possible cDNA library sequencing, revealing the potential of salivary molecules to be used as a primary tool for the development of novel drugs with pharmacological applications. However, as this method results in a limited number of sequences, only a portion of the transcriptome could be analyzed, and moreover, isoforms are usually indistinguishable from each other (8). Today, deep-sequencing technologies generate a massive amount of high-resolution data in a faster, cheaper and efficient manner, allowing the collection of more complete transcriptomes, revealing new isoforms and also describing boundaries of incomplete genes (8). The method that uses next generation sequencing (NGS) to sequencing cellular RNA, termed RNA-seq (RNA sequencing), can now be used to re-sequencing hematophagous salivary transcriptomes, increasing our view into its complexity. This finding enables the discovery of complete gene sequences or even novel bioactive salivary molecules, what offers additional data for the protein banks.

The cDNA library of salivary glands (SG) from *T. dimidiata* was previously reported describing a set of 388 sequences (9). Here, aiming to obtain a further insight into *T. dimidiata* salivary repertoire we used RNA-seq to obtain more than 51 million Illumina reads assembled into 5,704 Coding sequence (CDS), expanding the gene families related to secretory nature in *T. dimidiata* SGs transcriptome.

2. Materials and methods

The methodology used here was the same of our previous work reporting the *Rhodnius neglectus* sialome (manuscript in preparation), reproduced with few modifications.

2.1. Insects and Transcriptome SG Preparation

T. dimidiata triatomines used in this work were reared in an insectary room at the University of Brasilia, kept at 27 ˚C ± 1 and relative humidity of 70-75%. Vth instar nymphs and adults starved for 5, 12 and 24 days were dissected to SGs exposure. The pool containing thirty SGs pairs was immediately transferred to cold Trizol reagent (Invitrogen, Carlsbad, CA, USA) and kept at -80 ˚C prior to RNA extraction.

2.2. SG RNA Isolation, Library Preparation and Sequencing

Total RNA from SGs was isolated using Trizol and treated with turbo-DNase (Ambion, Austin, TX, USA) to remove DNA following the manufacturer's protocols. The RNA sample quality was checked by lab-on-chip analysis using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The library construction and the sequencing were executed by *Centro de Genômica de Alto Desempenho do DF* (High Performance Genomics Center of Distrito Federal, Brasília, Brazil). Library construction was performed with standard protocols of TruSeq RNA kit, v2 (Illumina, San Diego, CA) using the mRNA obtained from the pool of SGs. The RNA-seq sequencing step was performed on an Illumina MiSeq sequencer (Illumina, San Diego, CA, USA) and paired-end reads of 300 nucleotides in length were generated. One lane of the MiSeq machine was used for sequencing this and another library, distinguished by bar coding. A total of 51,782,206 reads were checked for data quality using FASTQC (Babraham Institute, Cambridge, UK), following trimming of low quality bases (quality 20 or lower). The nominal length of the sequences was 301 nt. Following trimming of low quality bases and Illumina primers, the average length was 233.6 bases, the median was 293 nt and L50 was 243 nt. Sequences smaller than 25 nt or with average qual < 20 were rejected.

2.3. Bioinformatics Tools for Transcriptome Analysis

Bioinformatic analysis was performed as described before (10). The *de novo* assembly of the reads was done with Abyss (11) as well as Soapdenovo Trans (12) softwares using different kmer (k) values (from 20 to 95 at fivefold intervals). The result data were joined by an iterative BLAST and cap3 assembler as described before (13). Sequence contamination between barcoded libraries were identified and removed when their sequence identities were over 98% but their abundance of reads were > 10 fold between libraries. CDS were extracted based on similarities to known proteins or by containing a signal peptide indicative of secretion as evaluated by version 3.0 of the SignalP program (14). The final assembly of CDS and their protein sequences were mapped into a hyperlinked Excel spreadsheet. These operations were automated by a program written in Visual Basic (VB) by JMCR. To estimate the transcripts abundance, reads were mapped back into the CDS using blastn (15) with a word size of 25, masking homonucleotide decamers and allowing mapping up to five different CDS if the BLAST results had the same score values. Mapping of the reads was also included in the Excel spreadsheet. The automated annotation of proteins was based on a vocabulary of nearly 250 words found in matches to the different databases, including Swissprot, Gene Ontology, KOG, PFAM, SMART and a subset of the non-redundant protein database of the NCBI containing proteins from vertebrates. Further manual annotation was done as required. Alignment analysis was done with Bioedit software (16) after sequence alignment performed by ClustalW (17). The sequences used in the alignments with *T. dimidiata* CDS were obtained from the nonredundant protein database of the National Center for Biotechnology Information (NCBI) and are represented by six letters followed by the NCBI gi accession number. The letters derive from the first three letters of the

genus and the first three letters from the species name. Phylogenetic analysis and statistical neighbor-joining bootstrap test of phylogenies were done with Mega package (18).

2.4. LC - MS/MS Salivary Content Preparation

Vth instar nymphs and adults starved for 5, 12 and 24 days were dissected to SGs exposure. The pool containing the salivary content corresponding to fifteen pairs of SGs was immediately transferred to ice and kept at -80 ˚C until use. Afterwards the protein content from saliva was ethanol/acetone precipitated. Dried samples were suspended in 100 µL of 100 mM TEAB (triethylammonium bicarbonate) diluted with 400 µL of ice-cold ethanol and mixed by vortex. Following, 400 µL of ice-cold acetone was added, mixed vigorously by vortex and incubated overnight at -20 °C. After incubation, the sample was centrifuged at 20,000g at 4 °C for 15 min and the supernatant discarded. The pellet was washed 3 times with ice-cold 40% ethanol/40% acetone solution. The protein pellet was suspended in 100 µL of 8 M urea in 100 mM TEAB, reduced with 20 mM DTT for 60 min, alkylated with 40 mM IAA (iodoacetamide) in the dark for 60 min and digested overnight with 0.01 AU Lys-C at room temperature. After Lys-C digestion, the solution was diluted with 100 mM TEAB to a final urea concentration of 1.0 M and modified trypsin was added (*ca.* 1:100 w/w) and the sample was incubated for more 3 h at room temperature. Finally, the peptide sample was acidified to 0.1% TFA and desalted on homemade microcolumns of Poros Oligo R3 resin packed (1 cm long) into c18 Ziptips™ (adapted from (19)). Desalted samples were dried down and stored at -20 °C until MS analysis.

2.5. LC–MS/MS and data analysis

Desalted sample was suspended in 1% FA (formic acid) and approximately 2 μg were used. Analysis was carried out in a DIONEX 3000 nanoUPLC system (Thermo Scientific, Waltham, USA) coupled online to an LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific, Waltham, USA). Peptides from the sample were loaded onto a 2 cm fused silica trap column (150 μm inner diameter) packed in-house with reverse phase capillary column ReproSil-Pur C18-AQ 5 μm resin (Dr. Maisch GmbH, Germany) and washed for 5 min with 98% phase A (0.1% formic acid) and 2% phase B (0.1% formic acid, 95% acetonitrile). Then, the elution was carried out using a gradient from 2% phase B to 30% phase B for 90 min, 30% to 98% phase B for 5 min, 98% phase B for 10 min, back to 2% phase B in 1 min and re-equilibrated for 15 min in 2 % phase B (a total of 121 min at 300 nl/min) in a 15 cm fused silica emitter (75 μm inner diameter) packed in-house with reverse phase capillary column ReproSil-Pur C18-AQ 3 μm resin (Dr. Maisch GmbH, Germany). Mass spectrums were acquired in positive ion mode applying data-dependent automatic survey MS scan and tandem mass spectra (MS/MS) acquisition. Each MS scan in the orbitrap (mass range of m/z of 400–1300 and resolution 120,000) was followed by MS/MS of the 15 most intense ions in the LTQ. Fragmentation in the LTQ was performed by HCD and selected sequenced ions were dynamically excluded for 15 s. Raw data was viewed in Xcalibur v.2.1 (Thermo Scientific, Waltham, USA). Data processing was performed using Proteome Discoverer v.1.3 (Thermo Scientific, Waltham, USA). Raw files were generated and these were used for protein identification using Proteome Discoverer with in house SequestHT algorithm against *T. dimidiata* SG transcriptome. Contaminant proteins (several types of human keratins, BSA and porcine trypsin) were also added to the database and all contaminant proteins identified were manually removed from the result lists. The searches were performed with the following

parameters: MS accuracy 10 ppm, MS/MS accuracy 0.05 Da, trypsin digestion with up to 2 missed cleavage allowed, fixed carbamidomethyl modification of cysteine and variable modification of oxidized methionine. Number of proteins, protein groups and number of peptides were filtered for false discovery rate (FDR) less than 1%, peptides with rank 1 and proteins with at least 2 peptides.

3. Results and discussion

In this work, we present the analysis of a set of 5,704 CDS from the cDNA library of SGs from *T. dimidiata*. The CDS were functionally annotated and then utilized throughout our subsequent analysis to perform a general description of *T. dimidiata* SGs transcriptome. The sequences are available in NCBI databank < http://www.ncbi.nlm.nih.gov/biosample/4240765 >

3.1. General description of transcripts from *T. dimidiata* SG transcriptome

Blood-sucking insects present different strategies to counteract the host defenses mechanisms. To ensure the acquisition of their blood meal, they modulate the host responses through the injection of different biologically active molecules. The biological role of saliva is clear and to catalogue its components in database banks is fundamental to the search of new pharmacologically active compounds and to the understanding of the vector's biology and evolution. Here, in *T. dimidiata* SGs transcriptome analysis, we performed automatic and manual annotation of the CDS to classify them into five broad categories of expressed genes, summarized in Table 1. A total of 662 CDS were probably associated with hematophagy comprising the secreted (S) category, containing the

majority of the reads with 87.83 %. The Housekeeping (H) category comprises 11.77 % of the reads containing 4,565 CDS, followed by the Unknown (U) category, containing 335 CDS and 0.26 % of the reads. Finally, the transposable elements (TE) and viral products (Vr) categories are represented by 138 and 4 CDS respectively, together both classes comprises approximately 0.12% of total reads.

Table 1
Classification and abundance of coding sequences extracted from the salivary gland transcriptome of *Triatoma dimidiata*

Class	No. of CDS	% Total	No. of reads	% Total
Secreted	662	11.60	45,484,416	87.83
Housekeeping	4,565	80.03	6,095,287	11.77
Unknown product	335	5.87	138,678	0.26
Transposable element	138	2.41	62,646	0.12
Viral product	4	0.07	1,179	0.00
Total	**5,704**	**100**	**51,782,206**	**100**

Further, the CDS related to housekeeping function were divided into 24 groups accordingly to their possible biological function (Table 2). The Cytoskeletal protein group comprises the majority of housekeeping reads, with 21.68 %, followed by Protein export 20.75 % and Unknown conserved, with 12.11 % of the reads.

Table 2
Classification and abundance of coding sequences of putative housekeeping function extracted from the sialotranscriptome of *Triatoma dimidiata*

Class	No of CDS	% Total	No of reads	% Total
Cytoskeletal protein	165	3.61	1,321,712	21.68
Protein export	268	5.87	1,264,747	20.75
Unknown conserved	816	17.88	738,030	12.11
Detoxification	84	1.84	440,727	7.23
Protein synthesis machinery	273	5.98	346,627	5.69
Energy metabolism	146	3.20	296,012	4.86
Signal transduction	561	12.29	285,330	4.68
Transcription machinery	484	10.60	204,455	3.35

Lipid metabolism	202	4.42	179,181	2.94
Transporter and Channel	242	5.30	170,425	2.80
Proteasome	197	4.32	146,680	2.41
Protein modification	148	3.24	112,595	1.85
Carbohydrate metabolism	129	2.83	96,720	1.59
Protein modification. protease	86	1.88	90,715	1.49
Immunity	66	1.45	85,992	1.41
Nuclear Regulation	224	4.91	78,366	1.29
Amino acid metabolism	70	1.53	73,860	1.21
Extracellular matrix	101	2.21	40,929	0.67
Transcription factor	102	2.23	34,991	0.57
Nucleotide metabolism	84	1.84	31,533	0.52
Intermediary metabolism	53	1.16	19,572	0.32
Signal Transduction. apoptosis	24	0.53	13,797	0.23
Nuclear Export	32	0.70	13,156	0.22
Storage	8	0.18	9,135	0.15
Total	**4,565**	**100**	**6,095,287**	**100**

3.2. Putative secreted proteins

The CDS of the S category were divided into different groups accordingly to their putative function (Table 3). It is possible to observe previously described gene families in hematophagous saliva. The following section describes the putative secreted proteins from *T. dimidiata* SGs.

Table 3
Classification and abundance of coding sequences of putative secretory function extracted from the sialotranscriptome of *Tritatoma dimidiata*

Class	No of CDS	% Total	No of reads	% Total
Lipocalin	76	11.48	42,637,077	93.74
Protease inhibitor	24	3.63	923,391	2.03
Hypothetical secreted protein	305	46.07	907,707	2.00
Antigen-5	6	0.91	338,467	0.74
(Salivary) secreted protein	61	9.21	221,315	0.49
Nucleotide metabolism	13	1.96	191,812	0.42
Hypothetical hemolysin-like	1	0.15	118,675	0.26
Protease	22	3.32	55,458	0.12
Conserved secreted protein	84	12.69	54,652	0.12

OBP	13	1.96	16,388	0.04
Lipid metabolism	12	1.81	7,674	0.02
Mucin related	3	0.45	4,904	0.01
Hormone related	13	1.96	2,163	0.00
Others	9	1.36	1,986	0.00
ABP	8	1.21	1,087	0.00
5'nucleotidase/Apyrase	2	0.30	976	0.00
Unknow secreted	4	0.60	335	0.00
Immunity related	5	0.76	318	0.00
Reprolysin	1	0.15	31	0.00
Total	**662**	**100**	**45,484,416**	**100**

3.2.1. Lipocalins

The lipocalin proteins are abundantly expressed in *T. dimidiata* SGs, comprising 93.74 % of putative secreted proteins, totalizing 76 CDS and 42,637,077 reads. This data confirms that lipocalins are highly expressed in the SG at the transcriptional level, being the most expressed salivary protein family in *T. dimidiata*. This feature has been found in other Triatoma species, such as *T. infestans* (20), *T. rubida* (21) and *T. matogrossensis* (22) and *Panstrongylus megistus* (23). There is a wide variety of members within this group reflecting the diversity of anti-hemostatic functions assigned to these molecules as a probably result oflineage specific expansion of this protein family during evolution. The salivary repertoire available at the time and the pressure to adapt to a blood feeding habit may have influenced this feature (24, 25).

Present in a variety of organisms, the lipocalins are a widespread functional diverse family of typically small and extracellular members that usually act through the binding of small and mostly hydrophobic molecules, but may also bind to cell-surface receptors or form complexes with other soluble macromolecules (26-28).

The versatility of the lipocalin family can be due to the ability of its members to exhibit a diverse primary structure in contrast to a highly conserved three dimensional

structure: a β-sheet consisting of eight antiparallel β-barrel forming a central hydrophobic cavity which binds to its particular ligand (29). The loops connecting the β-barrels are flexible and tolerate substitutions in the amino acid chain without changing the protein folding. It is exactly the flexibility in the composition of these loops that reflects the different depths and dimensions of the binding site, which are able to accommodate a diverse set of ligands. The changes in the cavity of the barrel are probably the key of lipocalins evolutionary path (26, 28, 29).

It was suggested that in hematophagous arthropods an expansion of the lipocalin family would have occurred twice, independently, in ticks and bugs (30-32). The biological function of lipocalins in tick saliva is diverse, such as histamine and serotonin binding (33, 34), immunomodulation (35), anti-complement activity (36) or toxic properties (31). Salivary lipocalins were also described in triatomine insects. Within *Rhodnius* genus, RPAI (*Rhodnius* platelet aggregation inhibitor) and ABP (*amine-binding protein*) are two of the examples. The first one inhibits platelets by low doses of ADP (37, 38), while the second binds to biogenic amines such as serotonin, norepinephrine and epinephrine, acting as a vasodilator and a platelet aggregation inhibitor (39). Regarding these two molecules, both are commonly reported in Rhodnius but not in Triatoma genus, although one ABP member were found in *T. matogrossensis* sialotranscriptome analysis (22).

Salivary lipocalin components from Triatoma genus acting as thrombin inhibitors that prevent local clotting have also been studied previously. Thrombin is the serine proteinase key to hemostasis and thrombosis, cleaving fibrinogen to fibrin and stimulating platelet aggregation. Additionally, it activates other coagulation factors and interacts with other cells (40). In *Triatoma pallidipennis*, triabin is an anti-coagulant which, in coagulation cascade, inhibits thrombin, forming a noncovalent complex at a molar ratio

of 1:1 (41). From the same species, pallidipin specifically inhibits platelet aggregation induced by collagen (42). Another lipocalin known is procalin, the major allergen protein described in *T. protracta* saliva (43). From *T. infestans*, triafestin exhibits anti-inflammatory activity by inhibiting the kallikrein-kinin system and anti-coagulant activity, interacting with factor XII from coagulation pathway (44). Finally, triplatin shows to inhibit platelet aggregation induced by collagen, showing the molecule interacts with PGF(2α) and inhibits platelet aggregation induced by low doses of collagen and thromboxane A2, but not by ADP (45). From *D. maxima* saliva, a lipocalin termed dipetalodipin presents strong sequence similarity to pallidipin, being suggested it has multifunctional anti-hemostatic properties, reducing platelet aggregation and vasoconstriction, besides blocking angiogenesis (46).

From the 76 CDS related to lipocalin in the transcriptome analysis, the most abundant were those extracted as salivary lipocalins, corresponding to molecules not yet functionally characterized, counting on 52.6 % of these CDS. Regarding lipocalin homologues already studied, the most abundant was a member of the triabin family, represented by 28.9 % of the CDS, followed by pallidipin 9.2 % and triatin 7.9 %. Sequences from *T. dimidiata* SG transcriptome were aligned with other lipocalins obtained from the nonredundant protein database from NCBI (NR database) to construct a phylogenetic tree (Fig 2). The *T. dimidiata* sequences grouped within different monophyletic clades, such as those containing members of triabin, pallidipin, triatin, triplatin, triafestin, procalin and ABP. These results suggest that an expansion of this gene family might have occurred by gene duplication events during this hematophagous insect evolution.

Fig. 2. Phylogram of the lipocalin family from *T. dimidiata* SGs transcriptome.
Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other triatomine

lipocalin sequences showing the distance between members of the family. *T. dimidiata* sequences are named with initials TD. The triatomine sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar represents 20% amino acid substitution. The color circles identify the species whose sequences were used: blue, *T. dimidiata* sequences from SG transcriptome; red, *T. infestans*; green, *T. pallidipennis*; magenta, *T. protracta*; cyan, *T. brasiliensis*; yellow, *R. neglectus.*

3.2.2. Antigen-5

The CAP protein family (cysteine-rich secretory proteins (CRISP), antigen 5, and pathogenesis-related 1 proteins) is a superfamily of secreted proteins found in a wide range of organisms. Together with phospholipase A1, the Antigen 5 is the major protein component of wasps, hornets and fire ants venoms (47, 48). Lately, whereas new sialome studies have been unraveled, Antigen-5 proteins have been shown to be a ubiquitous protein in hematophagous insects' saliva. Members from *T. infestans* and *Dipetalogaster maxima* were reported as molecules that modulate platelet aggregation induced by low doses of collagen (49). In *T. dimidiata* sialotranscriptome, the antigen-5 sequences were extracted from 338,467 reads distributed in 6 CDS. Interestingly, different from the others *T. dimidiata* sequences, which possess only the SCP domain (pfam 00188), the TD_13979 is a longer sequence presenting both a SCP domain and a CRISP domain (pfam 08562). This feature is common in Antigen-5 members from reptile venoms and is characterized by high cysteine content especially in their C-terminal third, which are involved in disulfide bonds (50, 51). The phylogram reveals the sequence TD_13979 grouped in Clade III, together with diptera species and *Nasonia vitripennis*, a wasp

species (Fig 3). The sequence TD_5315 possesses two CAP domains, grouping together with sequences containing this feature in Clade II. Other three sequences from *T. dimidiata* grouped in Clade I, together with *T. infestans* Antigen-5 that contains only one CAP domain. These results revealed that *T. dimidiata* contains three different lineages of Antigen-5 protein. Furthermore, it is possible to suggest that TD_13979 is a multidomain protein, where CAP and CRISP may work in cooperation, strategy that may improve the effective function of this protein increasing binding affinity or targets specificity.

Fig. 3. Phylogram of the Antigen-5 family from *T. dimidiata* SGs transcriptome. Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other triatomine lipocalin sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The triatomine sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar at the bottom represents 20% amino acid substitution.

3.2.3. Proteases

T. dimidiata SGs analysis revealed a variable repertoire of proteases extracted from 22 CDS, including carboxipeptidases, cathepsins L and B, legumain and, in majority, serine proteases. Serine proteases belong to a large family of peptidases and the typical catalytic triad is composed by His/Asp/Ser residues in the polypeptide chain (52). An important activity of this protease is related to the fine regulation of hemostasis and inflammation (53-55). In insects, proteases are responsible for protein catabolism and digestive processes, being related a high protease activity in the midgut. In Hemipterans,

this activity is due to cysteine and aspartic proteases (56). Regarding proteolytic activity in triatomine saliva little is known, although it was already reported that some salivary molecules are stored in the glands as inactive zymogens (57). In *T. infestans* saliva, a trypsin-like activity in saliva was described and the enzyme, termed triapsin, is stored as a proteolytically activatable zymogen, and the compound showed to have a gelatinolytic activity (57). In *Panstrongylus megistus*, a fibrinolytic activity in saliva was also related to a trypsin-like protease (58).

Proteolytic activity has also been detected in SGs from Heteropterans (59-61), suggesting the salivary proteases make part of an ancestral lineage and during the adaptation to a blood-feeding behavior these enzymes may have gained a new function. Here, the alignment of two CDS with other serine proteases shows a catalytic triad containing His and Asp residues, however the third residue, Ser, is replaced by Gly, in a non-classical arrangement (Fig 4, A). Proteases in the same family are related based on the sequence homology of their amino acid sequences and a different active site configuration may allow for activity in a different cellular environment, like pH alterations (62). The phylogram (Fig 4, B.) shows different CDS grouped with other proteases sequences, including carboxipeptidases, legumains and cathepsins. The exactly role of proteases in the saliva remains to be investigated and functionally characterized.

Fig. 4. Secreted proteases from *T. dimidiata* SGs transcriptome. A, ClustalW alignment of protease members from *T. dimidiata* salivary gland transcriptome (TD_4045 and TD_19949) and sequences from other arthropods: *Triatoma infestans* (gi 149689028), *Acyrthosiphon pisum* (gi 240848975), *Apis dorsata* (gi 572262138) and *Athalia rosae* (gi 817058472). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background. The blue bar indicates the signal peptide indicative of

secretion. The # indicates the residues of the active site, while * indicates the substrate binding site. **B,** Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar represents 20% amino acid substitution. Colored circles were used to identify the sequences: blue, *T. dimidiata* sequences from SG transcriptome; red, Hemiptera order; green, Diptera order; magenta, Hymenoptera order; cyan, Dictyoptera order; yellow, Coleoptera order; grey, Arachnid class.

3.2.4. Protease inhibitors

In response to the bite, against blood loss, the host responds with blood coagulation, a complex pathway controlled by serine proteases. In order to maintain a continuous blood flow and success in feeding, the salivary proteins target different enzymes in host's coagulation cascade, eliminating proteolysis and acting as anti-hemostatic proteins. From 923,391 reads, 24 CDS were extracted as serine proteinase inhibitors from Kazal-type, Serpin and Pacifastin families.

3.2.4.1. Kazal-type family

The function of proteinase inhibitors in hematophagous saliva is frequently associated to anticoagulant activity. However, some of these salivary enzymes may play mechanisms non related to proteolysis, working through more than one manner to improve the anti-hemostatic effect. It was proposed that horse fly SGs expresses a

protease inhibitor termed vasotab acting as a vasodilator molecule (63, 64). Moreover, in *Tabanus yao*, it was also proposed that the same molecule inhibits platelet aggregation through a KGD motif, which is a potent inhibitor of platelet aggregation since it blocks the binding of fibrinogen to the platelet receptor GPIIb/IIIa (65).

Homology search suggests vasotab protein as one of best matches for *T. dimidiata* CDS related to kazal-type family. In the alignment, it is possible to observe that the position of six conserved cysteines residues responsible for the intra-chain disulfide bridges is completely conserved (Fig 5, A), but differently from horse flies there is no KGD motif in *T. dimidiata* CDS, suggesting this kazal proteinase inhibitor may function only as a vasodilator protein. If there were also a proteinase inhibitory activity, it would be through another mechanism.

Vasoconstriction is a mechanism used by the host in response to tissue injury, closing the lumen of arterioles and small vessels thereby decreasing blood flow to the blood-sucking mouthparts. Thus, during the blood meal, hematophagous insect secret vasodilators to enlarge the size of the target vessels, making them easier to be located and also speeding the subsequent ingestion of the meal. The relaxation of a blood vessel can be accomplished by removal of vasoconstrictors or by addition of vasodilators, promoting vasodilation (5, 6, 66). In Triatoma genus, vasodilator molecules had not yet been reported, but this important strategy may be represented in these hematophagous saliva.

The phylogram shows three *T. dimidiata* CDS, TD_30859, TD_17576, and TD_21160, grouped in clade II together with other vasotab-like sequences containing only one kazal domain (Fig 5, B). In addition, TD_33228 and TD_28148, that contain three and two domains respectively, grouped in clade I together with proteinase inhibitors isolated from triatomine gut containing kazal multidomains, such as dipetalogastin, from *Dipetalogaster maxima* (67), infestin, from *T. infestans* (68) and brasiliensin, from *T.*

brasiliensis (69). This result suggests different kazal-type proteinases in *T. dimidiata* saliva may work in more than one manner. In sequences homologous to dipetalogastin each domain may bind independently in different surface sites of thrombin, preventing blood coagulation in an efficient mechanism.

Fig. 5. The kazal-type protease inhibitors from *T. dimidiata* SGs transcriptome. A, ClustalW alignment of kazal-type members from *T. dimidiata* salivary gland transcriptome (TD_30859, TD_17576 and TD_21160) and other kazal-type sequences: *Triatoma infestans* (gi 149898876, gi 149898841), *Phlebotomus papatasi* (gi 451935058), *Triatoma brasiliensis* (gi 116267193), *Hybomitra bimaculata* (gi 94730670), and *Tabanus yao* (gi 241914367). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background and the blue bar indicates the signal peptide indicative of secretion. The # indicates the conserved cysteine residues. **B,** Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar at the bottom represents 20% amino acid substitution.

3.2.4.2. Serpin (Serine proteinase inhibitors)

Serine proteinase inhibitors (Serpin) are a superfamily widely distributed among eukaryotes and their members, typically 350–400 amino acids in length, present a consensus tertiary structure consisting of a fold containing a bundle of nine α-helices and

three *β*-sheets (70, 71). In addition, the molecule shows a reactive center loop (RCL) which is cleaved by the target proteinase, and then serpin covalently traps the target disrupting its catalytic site and preventing the release of the protease from the complex (71, 72). Among a diversity of functions, many serpin members regulates the inhibition of serine proteinases in the vertebrate blood coagulation pathway (73).

Common in tick saliva, serpins have been suggested to act as an anticoagulant and immunosuppressive molecule (74-77). Not usually described in triatomine and diptera, a salivary serpin from the Dengue vector, *Aedes aegypti,* was isolated and showed to function as a potent reversible inhibitor of the host's factor Xa (78, 79)

Here, in *T. dimidiata* sialotranscriptome two CDS are related to serpins. The alignment reveals partial conserved sequences but shows the RCL residues, crucial for the ability of the inhibitory conformational mechanism (Fig 6, A). In the phylogram analysis TD_30493 grouped in Clade I with sequences containing just one serpin domain, whereas TD_28621 grouped in Clade II, with sequences containing two serpin domains (Fig 6, B). Serpins may be a part of the remarkable diverse group of salivary anticoagulants targeting factors of blood coagulation cascade.

Fig. 6. The serpin family from *T. dimidiata* SGs transcriptome. A, ClustalW alignment of a serpin member from *T. dimidiata* salivary gland transcriptome (TD_30493) and other arthropod sequences: *Nilaparvata lugens* (gi 482875806), *Zootermopsis nevadensis* (gi 646701027), *Acyrthosiphon pisum* (gi 229577153), *Culex quinquefasciatus* (gi 170053098) and *Aedes aegypti* (gi 157118702). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background, the blue bar indicates the signal peptide indicative of secretion and the red bars indicate the reactive center loop (RCL). **B,** Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the

family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar at the bottom represents 10% amino acid substitution.

3.2.4.3. Pacifastin

Pacifastin is a serine protease inhibitor family firstly described in the plasma of crayfish *Pacifastacus leniusculus* showing to inhibit both trypsin and chymotrypsin in an efficient manner (80). Later on, this crustacean proteinase showed to be a heterodimeric molecule, presenting a transferrin subunit and an inhibitory subunit composed of consensus cysteine-rich motifs, each one containing six cysteines residues, C**1** X9-12 C**2** N X C**3** X C**4** X2-3 G X3-4 C**5** T X3 C**6**, arranged to form a specific three disulfide bond pattern (81, 82). Some studies have reported the presence of this inhibitor in insects and crustaceans (83-87). From blood sucking insects, a pacifastin-like protease inhibitor was characterized showing to be up-regulated in the fat body after a blood meal, suggesting a role in insects immune system as a manner to counteract microorganisms acquired during the meal (88). Although there is some knowledge about pacifastin members characteristics and probable inhibitory mechanisms, little is known about its role and its target protease. In hematophagous insect saliva this protease inhibitor is not commonly described; and in *T. dimidiata* SGs, it comprises 2,057 reads and four CDS. The alignment shows five pacifastin domains in *T. dimidiata* CDS (Fig 7, A). In the phylogram, two CDS grouped in Clade I, together with sequences from other hematophagous arthropods,

including one characterized from *T. infestans* eggs (88), while TD_7118 grouped together with *P. leniusculus* sequence in Clade II (Fig 7, B).

Fig. 7. The pacifastin family from *Triatoma dimidiata* salivary gland. A, ClustalW alignment of pacifastin members from *T. dimidiata* salivary gland transcriptome (TD_7118 and TD_3748) and other sequences from arthropods: *Anopheles sinensis* (gi 668451255), *Aedes aegypti* (gi 157103392) and *Pacifastacus leniusculus* (gi 1764109). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background, the blue bar indicates the signal peptide indicative of secretion and the colored bars indicate the different pacifastin domains. **B,** Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar at the bottom represents 20% amino acid substitution.

3.2.5. Hemolysin

Representing 0.26 % of total reads from the secreted class, a single CDS was extracted as hemolysin. Despite this molecule has been already found in the saliva of some triatomines, its role in the blood meal acquisition is unknown, being suggested the molecule causes erythrocyte lysis to help the early steps of the digestion process (20). Common in microorganisms as *Escherichia coli* (89), *Bacillus cereus* (90) and *Staphylococcus aureus* (91), the hemolysins were firstly named based on its action as red blood cell lytic toxin, being later described its toxicity on epithelial or endothelial cells,

and also on immune cells and platelets. They are water-soluble monomers containing *β*-sheet structure that undergoes conformational changes to form *β*-barrel oligomers organized in the transmembrane space, forming cytotoxic pores in lipid bilayers, what culminates in cellular lysis (89, 91, 92). The biologic function of hemolysins in triatomine saliva remains to be characterized.

3.2.6. OBP (odorant binding proteins)

Odorant binding proteins (OBP) are small soluble polypeptides highly concentrated in olfactory organs. The molecule presents binding activity responsible to detect chemical signals, activating specific chemoreceptors through unknown mechanisms (93, 94). Common both in the lymph of chemosensilla in insects and in the nasal apparatus of vertebrates, the OBPs are not homologous between these taxon, however they present the same physiological function, being crucial at the first step of olfactory signal transmission, transforming an extracellular chemical signal into an intracellular electronic signal (94, 95). Vertebrates OBPs are folded in the typical *β*-barrel structure of lipocalins, belonging to this superfamily (96), while in insects the OBPs are composed by a pattern of six *α*-helical domains arranged to enclose a hydrophobic binding pocket stabilized by three disulfide bridges (97). In *T. dimidiata*, 13 CDS were related to OBP. The alignment shows TD_4743 and other OBPs from Hemipterans containing the six cysteine residues related to disulfide bridges (Fig 8, A). In the phylogram, these sequences were grouped in clade II, a more distant branch (Fig 8, B). It is possible to note *T. dimidiata* CDS grouped in each of the other three branches, together with different arthropod OBPs.

To insects, the olfaction enables the recognition of outer signals ensuring their survival and reproduction. During evolution for blood feeding, the OBP of triatomines were probably adapted to facilitate the meal acquisition process, but the role of OBPs in hematophagous saliva is still unknown.

Fig. 8. The odorant binding protein (OBP) family from *Triatoma dimidiata* salivary gland. A, ClustalW alignment of OBP members from *T. dimidiata* salivary gland transcriptome (TD_4743) and other sequences from Hemiptera: *Adelphocoris lineolatus* (gi 270000360), *Euschistus heros* (gi 299474028), *Lygus lineolaris* (gi 573006010), *Apolygus lucorum* (gi 387158189) and *Nilaparvata lugens* (gi 516299413). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background. The blue bar indicates the signal peptide indicative of secretion, while the # indicates the conserved cysteine residues. **B,** Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar represents 20% amino acid substitution. The color circles identify the sequences used: blue, *T. dimidiata* sequences from SG transcriptome; red, Hemiptera order; green, Diptera order; magenta, Dyctioptera order; cyan, Lepidoptera order.

3.2.7. 5' nucleotidase / Apyrase

During host's hemostasis response, activated platelets release ADP and serotonin, agonists that recruits more platelets, amplifying the activation pathways of platelet aggregation and inducing their aggregation to the damaged vessel wall (98). Among the

inhibitory mechanisms of platelet aggregation already reported in hematophagous, an important strategy is performed by a molecule termed apyrase, which hydrolyse ADP and ATP, to AMP and Pi (99).

The enzyme is ubiquitous in the saliva of hematophagous arthropods, what suggests its fundamental role in blood meal acquisition. Based on substrate specificities and amino acid sequences three genetic families have been reported in hematophagous saliva: (1) the 5′-nucleotidase family, described in mosquitoes (100), triatomines (101, 102) and ticks (103); (2) the Cimex family, founded in bed bugs and flies (104-106); (3) homologues of CD39, a cell-surface apyrase, reported in fleas (107). It has been suggested elsewhere that even distantly related arthropod species can share the same apyrase family (108). 5'nucleotidase CDS from *T. dimidiata* and other species were aligned showing a significant extend of conserved residues (Fig 9, A), even with sequences not derived from hematophagous insects, suggesting a conservation from an ancestral lineage. In addition, TD_31514 and TD_31776 sequence analysis revealed that the CDS are characterized by an N-terminal metallophosphatase domain and a C-terminal 5'nucleotidase domain, as well as in *T. infestans* apyrase (gi 34481604). In the phylogram analysis (Fig 9, B) *T. dimidiata* CDS proved to be closely related to *T. infestans* salivary apyrase, grouped in clade I, which in turn are divergent from those sequences of non hematophagous arthropods, clade II. In addition, TD_31514 and TD_31776 sequence analysis revealed that the CDS are characterized by an N-terminal metallophosphatase domain and a C-terminal 5'nucleotidase domain, as well as in *T. infestans* apyrase (gi 34481604).

Fig. 9. 5'nucleotidase sequences from *Triatoma dimidiata* salivary gland transcriptome. A, ClustalW alignment of 5'nucleotidase members from *T. dimidiata* salivary transcriptome (TD_31776 and TD_31514) and other sequences from arthropods: *Triatoma infestans* (gi 34481604), *Linepithema humile* (gi 815795501), *Wasmannia auropunctata* (gi 780674589), *Atta cephalotes* (gi 801400127), *Apis*

mellifera (gi 66523706) and *Bombus terrestris* (gi 340724470). The sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The alignment indicates conserved residues in black and similar residues in gray background. The blue bar indicates the signal peptide indicative of secretion. The red bar indicates the metallophosphatase domain, while the green bar indicates the 5'nucleotidase domain. **B**, Phylogenetic tree derived from the alignment of *T. dimidiata* CDS and other arthropods sequences showing the distance between members of the family. The *T. dimidiata* sequences are named with initials TD. The arthropods sequences derived from the non-reduntant (NR) protein database of NCBI and are represented by six letters (from the first three letters from the genus and the first three letters from the species name) followed by the NCBI gi accession number. The sequences were ClustalW aligned and the phylogram was done with the Mega package after 1,000 bootstrap iterations using the neighbor-joining (NJ) algorithm. The bar at the bottom represents 10% amino acid substitution.

4. The proteome analysis from *T. dimidiata* SG

The *T. dimidiata* SGs content was tryptic digested and subjected to mass spectrometry as a strategy to obtain an informative view of SGs secreted proteins. The result identified proteins possibly associated with hematophagy. Of to 112 secreted proteins identified, 34 were from the lipocalin family, including triabin, triatin, pallidipin, ABP and trialysin members (Table 4). Other important products identified as soluble proteins in the saliva were: protease, antigen-5, OBP, serpin, hemolysin and 5' nucleotidase, each one with at least one identification.

Table 4: Classification and abundance of proteins from the salivary proteome of *Triatoma dimidiata* based on LC-MS/MS

Class	No. of identifications	% Total
Others	61	55.36
Lipocalin	18	16.07
Triabin	6	5.36

Triatin	3	2.68
Pallidipin	3	2.68
ABP	2	1.79
Trialysin	2	1.79
Conserved secreted protein	6	5.36
Protease Trypsin-like	3	2.68
Antigen-5/SCP	1	0.89
OBP	1	0.89
Protease inhibitor - Serpin	1	0.89
Hemolysin	1	0.89
5'nucleotidase	1	0.89
Protease	1	0.89
Juvenile hormone related	1	0.89
Total	**111**	**100.00**

5. Conclusions

In this report, we found that the major amount of transcripts from *T. dimidiata* SGs related to secretory nature were lipocalins, as reported in the first analysis of its salivary transcriptome (9). The selective pressure during the adaptative process to a blood-feeding life style resulted in different salivary anti-hemostatic strategies, being suggested that some protein families acquired new functions to exert fundamental roles in blood flow maintenance. Here, as the results suggest for *T. dimidiata* species, the lipocalins expansion have emerged as the main gene family acting to counteract the vertebrate hosts' response, and probably have evolved through gene duplication events. Despite the large proportion of lipocalins over other transcripts of secretory nature, observed by reads and CDS percentage, it is evident that these low abundance molecules also have a crucial role in saliva and their amount is sufficient to exert their biological activity at the time of blood meal.

The current analysis allowed a deeper insight into the salivary glands proteins that contributes to the hematophagic habit of *T. dimidiata*. We expect this work will help future investigations regarding vector biology and disease transmission, and also

functional roles of the salivary proteins and their exploitation for the identification and selection of new pharmacologically compounds and/or targets for disease control strategies.

Aknowledgments

We are very grateful to Ana Cristina Gomes and Lucas Barbosa for technical assistance. This work was supported by the Brazilian National Council for Scientific and Technological Development (476252/2012-1), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (PROAP-2013), Financiadora de Estudos e Projetos (CT-Infra 2011), Fundação de Apoio à Pesquisa do Distrito Federal (Pronex 476252/2012-5), and the National Institute of Allergy and Infectious Diseases (Z01 AI000810-18 to Vector-Borne Diseases: Biology of Vector Host Relationship).

References

1. Bern C. Chagas' Disease. N Engl J Med. 2015;373(5):456-66.
2. Andrade DV, Gollob KJ, Dutra WO. Acute chagas disease: new global challenges for an old neglected disease. PLoS Negl Trop Dis. 2014;8(7):e3010.
3. WHO. Sustaining the drive to overcome the global impact of neglected tropical diseases: second WHO report on neglected tropical diseases 2013.
4. Dorn PL, Monroy C, Curtis A. Triatoma dimidiata (Latreille, 1811): a review of its diversity across its geographic range and the relationship among populations. Infect Genet Evol. 2007;7(2):343-52.
5. Ribeiro JM. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? Infect Agents Dis. 1995;4(3):143-52.
6. Champagne DE. Antihemostatic molecules from saliva of blood-feeding arthropods. Pathophysiology of haemostasis and thrombosis. 2005;34(4-5):221-7.
7. Champagne DE. Antihemostatic strategies of blood-feeding arthropods. Current drug targets. 2004;4(4):375-96.
8. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57-63.
9. Kato H, Jochim RC, Gomez EA, Sakoda R, Iwata H, Valenzuela JG, et al. A repertoire of the dominant transcripts from the salivary glands of the blood-sucking bug, *Triatoma dimidiata,* a vector of Chagas disease. Infect Genet Evol. 2010;10(2):184-91.
10. Ribeiro JM, Chagas AC, Pham VM, Lounibos LP, Calvo E. An insight into the sialome of the frog biting fly, *Corethrella appendiculata*. Insect biochemistry and molecular biology. 2014;44:23-32.
11. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. Genome research. 2009;19(6):1117-23.
12. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience. 2012;1(1):18.
13. Karim S, Singh P, Ribeiro JM. A deep insight into the sialotranscriptome of the gulf coast tick, Amblyomma maculatum. PLoS ONE. 2011;6(12):e28525.
14. Nielsen H, Brunak S, von Heijne G. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein engineering. 1999;12(1):3-9.
15. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-402.
16. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT Oxford University Press: Nucleic Acids Symposium Series; 1999. p. 95-8.
17. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics (Oxford, England). 2007;23(21):2947-8.
18. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24(8):1596-9.
19. Gobom J, Nordhoff E, Mirgorodskaya E, Ekman R, Roepstorff P. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom. 1999;34(2):105-16.

20. Assumpcao TC, Francischetti IM, Andersen JF, Schwarz A, Santana JM, Ribeiro JM. An insight into the sialome of the blood-sucking bug *Triatoma infestans,* a vector of Chagas' disease. Insect biochemistry and molecular biology. 2008;38(2):213-32.
21. Ribeiro JM, Assumpcao TC, Pham VM, Francischetti IM, Reisenman CE. An insight into the sialotranscriptome of *Triatoma rubida* (Hemiptera: Heteroptera). Journal of medical entomology. 2012;49(3):563-72.
22. Assumpcao TC, Eaton DP, Pham VM, Francischetti IM, Aoki V, Hans-Filho G, et al. An insight into the sialotranscriptome of *Triatoma matogrossensis*, a kissing bug associated with fogo selvagem in South America. The American journal of tropical medicine and hygiene. 2012;86(6):1005-14.
23. Ribeiro JMC, Schwarz A, Francischetti IMB. A deep insight into the sialotranscriptome of the Chagas disease vector, *Panstrongylus megistus* (Hemiptera: Heteroptera). J Med Entomol. 2015;52(3):351-8.
24. Mans BJ, Neitz AW. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochem Mol Biol. 2004;34(1):1-17.
25. Mans BJ, Ribeiro JM. Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins. Insect biochemistry and molecular biology. 2008;38(9):841-52.
26. Flower DR. Multiple molecular recognition properties of the lipocalin protein family. J Mol Recognit. 1995;8(3):185-95.
27. Flower DR, North AC, Sansom CE. The lipocalin protein family: structural and sequence overview. Biochim Biophys Acta. 2000;1482(1-2):9-24.
28. Grzyb J, Latowski D, Strzałka K. Lipocalins - a family portrait. J Plant Physiol. 2006;163(9):895-915.
29. Flower DR. The lipocalin protein family: structure and function. Biochem J. 1996;318 (Pt 1):1-14.
30. Mans BJ, Neitz AW. Exon-intron structure of outlier tick lipocalins indicate a monophyletic origin within the larger lipocalin family. Insect Biochem Mol Biol. 2004;34(6):585-94.
31. Mans BJ, Louw AI, Neitz AW. The major tick salivary gland proteins and toxins from the soft tick, Ornithodoros savignyi, are part of the tick Lipocalin family: implications for the origins of tick toxicoses. Mol Biol Evol. 2003;20(7):1158-67.
32. Montfort WR, Weichsel A, Andersen JF. Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. Biochimica et biophysica acta. 2000;1482(1-2):110-8.
33. Paesen GC, Adams PL, Harlos K, Nuttall PA, Stuart DI. Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. Mol Cell. 1999;3(5):661-71.
34. Sangamnatdej S, Paesen GC, Slovak M, Nuttall PA. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol Biol. 2002;11(1):79-86.
35. Preston SG, Majtán J, Kouremenou C, Rysnik O, Burger LF, Cabezas Cruz A, et al. Novel immunomodulators from hard ticks selectively reprogramme human dendritic cell responses. PLoS Pathog. 2013;9(6):e1003450.
36. Nunn MA, Sharma A, Paesen GC, Adamson S, Lissina O, Willis AC, et al. Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. J Immunol. 2005;174(4):2084-91.
37. Francischetti IM, Ribeiro JM, Champagne D, Andersen J. 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. 2000;275(17):12639-50.

38. Francischetti IM, Andersen JF, Ribeiro JM. 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. 2002;41(11):3810-8.
39. Andersen JF, Francischetti IM, Valenzuela JG, Schuck P, Ribeiro JM. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. J Biol Chem. 2003;278(7):4611-7.
40. Kalafatis M, Egan JO, van 't Veer C, Cawthern KM, Mann KG. The regulation of clotting factors. Crit Rev Eukaryot Gene Expr. 1997;7(3):241-80.
41. Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD. Triabin, a highly potent exosite inhibitor of thrombin. The Journal of biological chemistry. 1995;270(48):28629-34.
42. Noeske-Jungblut C, Krätzschmar J, Haendler B, Alagon A, Possani L, Verhallen P, et al. An inhibitor of collagen-induced platelet aggregation from the saliva of Triatoma pallidipennis. J Biol Chem. 1994;269(7):5050-3.
43. Paddock CD, McKerrow JH, Hansell E, Foreman KW, Hsieh I, Marshall N. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. J Immunol. 2001;167(5):2694-9.
44. Isawa H, Orito Y, Jingushi N, Iwanaga S, Morita A, Chinzei Y, et al. Identification and characterization of plasma kallikrein-kinin system inhibitors from salivary glands of the blood-sucking insect *Triatoma infestans*. The FEBS journal. 2007;274(16):4271-86.
45. Ma D, Assumpcao TC, Li Y, Andersen JF, Ribeiro J, Francischetti IM. Triplatin, a platelet aggregation inhibitor from the salivary gland of the triatomine vector of Chagas disease, binds to TXA(2) but does not interact with glycoprotein PVI. Thrombosis and haemostasis. 2012;107(1):111-23.
46. Assumpcao TC, Alvarenga PH, Ribeiro JM, Andersen JF, Francischetti IM. Dipetalodipin, a novel multifunctional salivary lipocalin that inhibits platelet aggregation, vasoconstriction, and angiogenesis through unique binding specificity for TXA2, PGF2alpha, and 15(S)-HETE. The Journal of biological chemistry. 2010;285(50):39001-12.
47. Hoffman DR. Allergens in Hymenoptera venom. XXV: The amino acid sequences of antigen 5 molecules and the structural basis of antigenic cross-reactivity. The Journal of allergy and clinical immunology. 1993;92(5):707-16.
48. King TP, Spangfort MD. Structure and biology of stinging insect venom allergens. International archives of allergy and immunology. 2000;123(2):99-106.
49. Assumpção TC, Ma D, Schwarz A, Reiter K, Santana JM, Andersen JF, et al. Salivary antigen-5/CAP family members are Cu2+-dependent antioxidant enzymes that scavenge O_{2-}. and inhibit collagen-induced platelet aggregation and neutrophil oxidative burst. J Biol Chem. 2013;288(20):14341-61.
50. Yamazaki Y, Morita T. Structure and function of snake venom cysteine-rich secretory proteins. Toxicon. 2004;44(3):227-31.
51. Nobile M, Noceti F, Prestipino G, Possani LD. Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules. Exp Brain Res. 1996;110(1):15-20.
52. Polgár L. The catalytic triad of serine peptidases. Cell Mol Life Sci. 2005;62(19-20):2161-72.
53. Mann KG, Butenas S, Brummel K. The dynamics of thrombin formation. Arterioscler Thromb Vasc Biol. 2003;23(1):17-25.

54. Coughlin SR. Protease-activated receptors in hemostasis, thrombosis and vascular biology. J Thromb Haemost. 2005;3(8):1800-14.
55. Cirino G, Napoli C, Bucci M, Cicala C. Inflammation-coagulation network: are serine protease receptors the knot? Trends Pharmacol Sci. 2000;21(5):170-2.
56. Terra WR, Ferreira C. Insect digestive enzymes: properties, compartmentalization and function Comp Biochem Physiol. 1994;109B(1):1-62.
57. Amino R, Tanaka AS, Schenkman S. Triapsin, an unusual activatable serine protease from the saliva of the hematophagous vector of Chagas' disease Triatoma infestans (Hemiptera: Reduviidae). Insect biochemistry and molecular biology. 2001;31(4-5):465-72.
58. Meiser CK, Piechura H, Meyer HE, Warscheid B, Schaub GA, Balczun C. A salivary serine protease of the haematophagous reduviid *Panstrongylus megistus*: sequence characterization, expression pattern and characterization of proteolytic activity. Insect molecular biology. 2010;19(3):409-21.
59. Cohen AC. Organization of Digestion and Preliminary Characterization of Salivary Trypsin-like Enzymes in a Predaceous Heteropteran, *Zelus renardii*. J Insect Physiol. 1993;39(10):823-9.
60. Colebatch G, East P, Cooper P. Preliminary characterisation of digestive proteases of the green mirid, Creontiades dilutus (Hemiptera: Miridae). Insect Biochem Mol Biol. 2001;31(4-5):415-23.
61. Zhu YC, Zeng F, Oppert B. Molecular cloning of trypsin-like cDNAs and comparison of proteinase activities in the salivary glands and gut of the tarnished plant bug Lygus lineolaris (Heteroptera: Miridae). Insect Biochem Mol Biol. 2003;33(9):889-99.
62. Ekici OD, Paetzel M, Dalbey RE. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. Protein Sci. 2008;17(12):2023-37.
63. Takac P, Nunn MA, Meszaros J, Pechanova O, Vrbjar N, Vlasakova P, et al. Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) salivary glands. J Exp Biol. 2006;209(Pt 2):343-52.
64. Ma D, Wang Y, Yang H, Wu J, An S, Gao L, et al. Anti-thrombosis repertoire of blood-feeding horsefly salivary glands. Mol Cell Proteomics. 2009;8(9):2071-9.
65. Zhang Z, Gao L, Shen C, Rong M, Yan X, Lai R. A potent anti-thrombosis peptide (vasotab TY) from horsefly salivary glands. Int J Biochem Cell Biol. 2014;54:83-8.
66. Champagne DE. The role of salivary vasodilators in bloodfeeding and parasite transmission. Parasitol Today. 1994;10(11):430-3.
67. Mende K, Petoukhova O, Koulitchkova V, Schaub GA, Lange U, Kaufmann R, et al. Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect. *Dipetalogaster maximus* cDNA cloning, expression and characterization. European journal of biochemistry / FEBS. 1999;266(2):583-90.
68. Campos IT, Amino R, Sampaio CA, Auerswald EA, Friedrich T, Lemaire HG, et al. Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas' disease vector: gene cloning, expression and characterization of the inhibitor. Insect biochemistry and molecular biology. 2002;32(9):991-7.
69. Araujo RN, Campos IT, Tanaka AS, Santos A, Gontijo NF, Lehane MJ, et al. Brasiliensin: A novel intestinal thrombin inhibitor from *Triatoma brasiliensis* (Hemiptera: Reduviidae) with an important role in blood intake. International journal for parasitology. 2007;37(12):1351-8.

70. Gettins PG. Serpin structure, mechanism, and function. Chem Rev. 2002;102(12):4751-804.
71. Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. Genome Res. 2000;10(12):1845-64.
72. Huntington JA, Read RJ, Carrell RW. Structure of a serpin-protease complex shows inhibition by deformation. Nature. 2000;407(6806):923-6.
73. Rau JC, Beaulieu LM, Huntington JA, Church FC. Serpins in thrombosis, hemostasis and fibrinolysis. J Thromb Haemost. 2007;5 Suppl 1:102-15.
74. Leboulle G, Crippa M, Decrem Y, Mejri N, Brossard M, Bollen A, et al. Characterization of a novel salivary immunosuppressive protein from Ixodes ricinus ticks. J Biol Chem. 2002;277(12):10083-9.
75. Imamura S, da Silva Vaz Junior I, Sugino M, Ohashi K, Onuma M. A serine protease inhibitor (serpin) from Haemaphysalis longicornis as an anti-tick vaccine. Vaccine. 2005;23(10):1301-11.
76. Prevot PP, Adam B, Boudjeltia KZ, Brossard M, Lins L, Cauchie P, et al. Anti-hemostatic effects of a serpin from the saliva of the tick Ixodes ricinus. The Journal of biological chemistry. 2006;281(36):26361-9.
77. Mulenga A, Kim T, Ibelli AM. Amblyomma americanum tick saliva serine protease inhibitor 6 is a cross-class inhibitor of serine proteases and papain-like cysteine proteases that delays plasma clotting and inhibits platelet aggregation. Insect Mol Biol. 2013;22(3):306-19.
78. Stark KR, James AA. Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, Aedes aegypti. J Biol Chem. 1998;273(33):20802-9.
79. Stark KR, James AA. A factor Xa-directed anticoagulant from the salivary glands of the yellow fever mosquito Aedes aegypti. Exp Parasitol. 1995;81(3):321-31.
80. Hergenhahn HG, Aspan A, Söderhäll K. Purification and characterization of a high-Mr proteinase inhibitor of pro-phenol oxidase activation from crayfish plasma. Biochem J. 1987;248(1):223-8.
81. Liang Z, Sottrup-Jensen L, Aspán A, Hall M, Söderhäll K. Pacifastin, a novel 155-kDa heterodimeric proteinase inhibitor containing a unique transferrin chain. Proc Natl Acad Sci U S A. 1997;94(13):6682-7.
82. Mer G, Hietter H, Kellenberger C, Renatus M, Luu B, Lefèvre JF. Solution structure of PMP-C: a new fold in the group of small serine proteinase inhibitors. J Mol Biol. 1996;258(1):158-71.
83. Liu Y, Cui Z, Shi G, Luo D, Wang S, Wang C. PtPLC, a pacifastin-related inhibitor involved in antibacterial defense and prophenoloxidase cascade of the swimming crab Portunus trituberculatus. Fish Shellfish Immunol. 2015;43(1):36-42.
84. Breugelmans B, Simonet G, van Hoef V, Van Soest S, Smagghe G, Vanden Broeck J. A lepidopteran pacifastin member: cloning, gene structure, recombinant production, transcript profiling and in vitro activity. Insect Biochem Mol Biol. 2009;39(7):430-9.
85. Simonet G, Claeys I, Huybrechts J, De Loof A, Vanden Broeck J. Bacterial production and purification of SGPI-1 and SGPI-2, two peptidic serine protease inhibitors from the desert locust, Schistocerca gregaria. Protein Expr Purif. 2003;31(2):188-96.
86. Simonet G, Breugelmans B, Proost P, Claeys I, Van Damme J, De Loof A, et al. Characterization of two novel pacifastin-like peptide precursor isoforms in the desert

locust (Schistocerca gregaria): cDNA cloning, functional analysis and real-time RT-PCR gene expression studies. Biochem J. 2005;388(Pt 1):281-9.
87. Boigegrain RA, Mattras H, Brehélin M, Paroutaud P, Coletti-Previero MA. Insect immunity: two proteinase inhibitors from hemolymph of Locusta migratoria. Biochem Biophys Res Commun. 1992;189(2):790-3.
88. de Marco R, Lovato DV, Torquato RJ, Clara RO, Buarque DS, Tanaka AS. The first pacifastin elastase inhibitor characterized from a blood sucking animal. Peptides. 2010;31(7):1280-6.
89. Bhakdi S, Bayley H, Valeva A, Walev I, Walker B, Kehoe M, et al. Staphylococcal alpha-toxin, streptolysin-O, and Escherichia coli hemolysin: prototypes of pore-forming bacterial cytolysins. Arch Microbiol. 1996;165(2):73-9.
90. Ramarao N, Sanchis V. The pore-forming haemolysins of bacillus cereus: a review. Toxins (Basel). 2013;5(6):1119-39.
91. Berube BJ, Bubeck Wardenburg J. Staphylococcus aureus α-toxin: nearly a century of intrigue. Toxins (Basel). 2013;5(6):1140-66.
92. Montoya M, Gouaux E. Beta-barrel membrane protein folding and structure viewed through the lens of alpha-hemolysin. Biochim Biophys Acta. 2003;1609(1):19-27.
93. Pelosi P. Odorant-binding proteins. Crit Rev Biochem Mol Biol. 1994;29(3):199-228.
94. Pelosi P, Iovinella I, Felicioli A, Dani FR. Soluble proteins of chemical communication: an overview across arthropods. Front Physiol. 2014;5:320.
95. Fan J, Francis F, Liu Y, Chen JL, Cheng DF. An overview of odorant-binding protein functions in insect peripheral olfactory reception. Genet Mol Res. 2011;10(4):3056-69.
96. Tegoni M, Pelosi P, Vincent F, Spinelli S, Campanacci V, Grolli S, et al. Mammalian odorant binding proteins. Biochim Biophys Acta. 2000;1482(1-2):229-40.
97. Pelosi P, Mastrogiacomo R, Iovinella I, Tuccori E, Persaud KC. Structure and biotechnological applications of odorant-binding proteins. Appl Microbiol Biotechnol. 2014;98(1):61-70.
98. Jurk K, Kehrel BE. Platelets: physiology and biochemistry. Semin Thromb Hemost. 2005;31(4):381-92.
99. Ribeiro JM, Sarkis JJ, Rossignol PA, Spielman A. Salivary apyrase of *Aedes aegypti:* characterization and secretory fate. Comparative biochemistry and physiology B, Comparative biochemistry. 1984;79(1):81-6.
100. Champagne DE, Smartt CT, Ribeiro JM, James AA. The salivary gland-specific apyrase of the mosquito Aedes aegypti is a member of the 5'-nucleotidase family. Proc Natl Acad Sci U S A. 1995;92(3):694-8.
101. Faudry E, Rocha PS, Vernet T, Lozzi SP, Teixeira AR. Kinetics of expression of the salivary apyrases in Triatoma infestans. Insect Biochem Mol Biol. 2004;34(10):1051-8.
102. Faudry E, Lozzi SP, Santana JM, D'Souza-Ault M, Kieffer S, Felix CR, et al. Triatoma infestans apyrases belong to the 5'-nucleotidase family. J Biol Chem. 2004;279(19):19607-13.
103. Stutzer C, Mans BJ, Gaspar AR, Neitz AW, Maritz-Olivier C. *Ornithodoros savignyi:* soft tick apyrase belongs to the 5'-nucleotidase family. Experimental parasitology. 2009;122(4):318-27.

104. Valenzuela JG, Charlab R, Galperin MY, Ribeiro JM. Purification, cloning, and expression of an apyrase from the bed bug Cimex lectularius. A new type of nucleotide-binding enzyme. J Biol Chem. 1998;273(46):30583-90.
105. Valenzuela JG, Belkaid Y, Rowton E, Ribeiro JM. The salivary apyrase of the blood-sucking sand fly Phlebotomus papatasi belongs to the novel Cimex family of apyrases. J Exp Biol. 2001;204(Pt 2):229-37.
106. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly Lutzomyia longipalpis. Proc Natl Acad Sci U S A. 1999;96(26):15155-60.
107. Andersen JF, Hinnebusch BJ, Lucas DA, Conrads TP, Veenstra TD, Pham VM, et al. An insight into the sialome of the oriental rat flea, Xenopsylla cheopis (Rots). BMC genomics. 2007;8:102.
108. Hughes AL. Evolution of the salivary apyrases of blood-feeding arthropods. Gene. 2013;527(1):123-30.

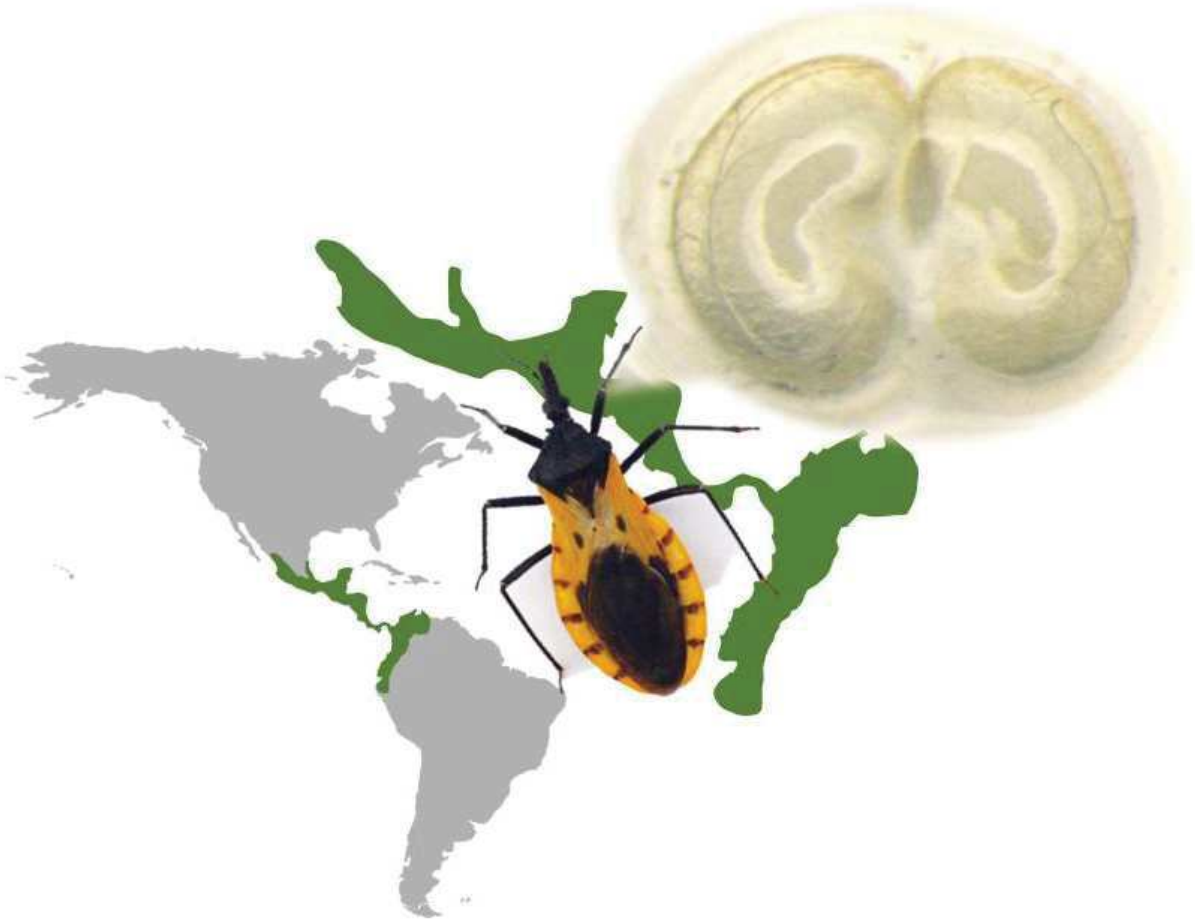


Fig 1

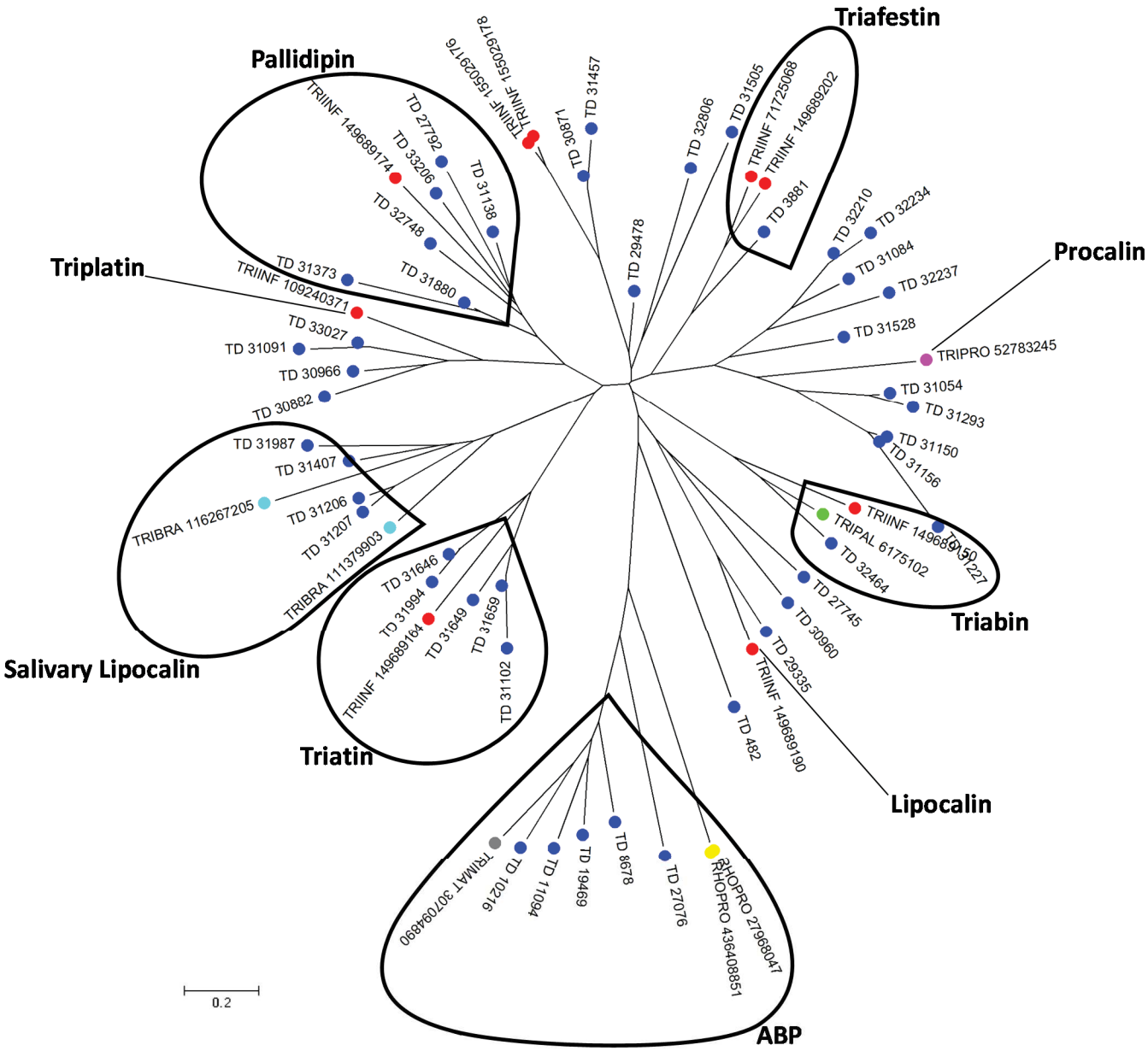

Triafestin
Pallidipin
Triplatin
Procalin
Salivary Lipocalin
Triabin
Triatin
Lipocalin
ABP
TRIINF 149689174
TRIINF 155029178
TRIINF 155029176
TD 27792
TD 33206
TD 31138
TD 32748
TD 31373
TD 31880
TRIINF 109240371
TD 33027
TD 31091
TD 30966
TD 30882
TD 31987
TD 31407
TRIBRA 116267205
TD 31206
TD 31207
TRIBRA 111379903
TD 31646
TD 31994
TRIINF 149689164
TD 31649
TD 31659
TD 31102
TD 30871
TD 31457
TD 29478
TD 32806
TD 31505
TRIINF 71725068
TRIINF 149689202
TD 3881
TD 32210
TD 32234
TD 31084
TD 32237
TD 31528
TRIPRO 52783245
TD 31054
TD 31293
TD 31150
TD 31156
TRIINF 149689191
TRIPAL 6175102
TD 32464
TD 27745
TD 30960
TD 29335
TRIINF 149689190
TD 482
TRIMAT 307094890
TD 10216
TD 11094
TD 19469
TD 8678
TD 27076
RHOPRO 27968047
RHOPRO 436408851
0.2

Fig 2

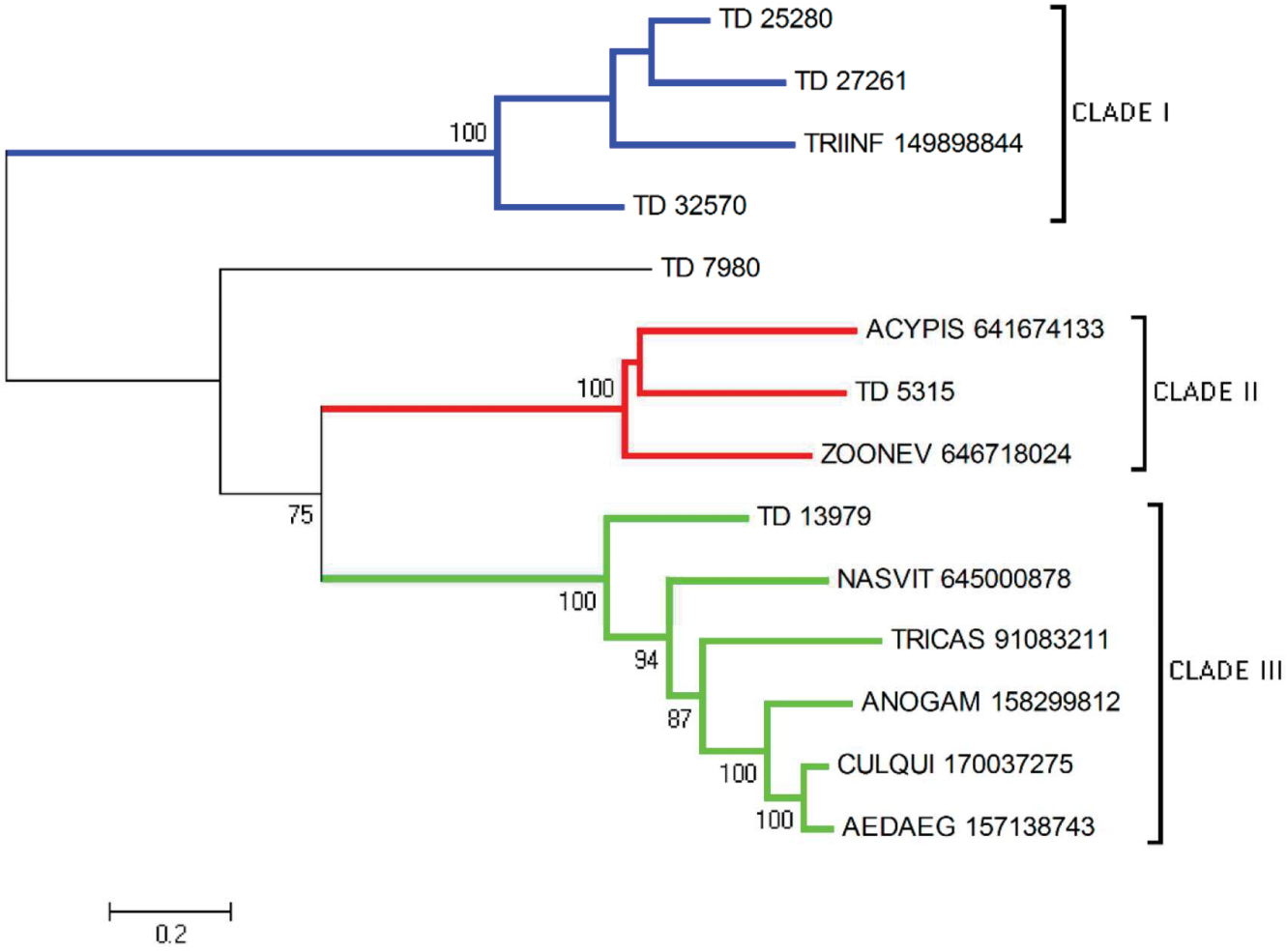

TD 25280
TD 27261
TRIINF 149898844
TD 32570
CLADE I
100
TD 7980
ACYPIS 641674133
TD 5315
ZOONEV 646718024
CLADE II
100
75
TD 13979
NASVIT 645000878
TRICAS 91083211
ANOGAM 158299812
CULQUI 170037275
AEDAEG 157138743
CLADE III
100
94
87
100
100
0.2

Fig 3

```
TD_4045           --MIIKVYHLFIVLFVLYGK---FASSSFLLYEEG--DECDIKRSHRTCQNVKKCPTEAEAVRKGLEP-DICGLRQSQ-LVCCSKTLKKS
TD_19949          ---------MFIQLLLSVCLLSLNYLQCNGQEDQNEDDCVCVPYYQCEEGKVITTGEGLLDLRLKTCT----GIEVCCQLP-DGTDGDST
TRIINF_149689028  --------MNYLYLLIGLVVVVEVPHAHTQGEGDDDEDSSEYGNRNVPGDKTTNCDCGWAN----------------------------
ACYPIS_240848975  MRSSNTNVQLFLTVFLLCFMNLGYNAQILGLYEDEICRKGDGKTPNSICRNIERCPVAKEGLKKNIIP-QLCSFNGSIPIVCCPP--EKN
APIDOR_572262138  ---------MLIIMFLYFLCN-----PIFCNTQDEDEPCITNNNLPGICKLLQQCPSVYENLLKGLSSHKICGYLNFDPVVCCPDIKKIS
ATHROS_817058472  -----MRCSMRLIVLLGLIAVSSGAPQKCGDLDCLIDSIYTPKSPTGSNGGVTPSPINNGGQIPESGG----GEDDCECVPYYQCQNNTI

TD_4045           ISNKKYLLMKEEASNKKKHIIKRAAP---GEKSKRMCQKYSESVYS-FEEPQILLPGATKVKVDNCGIRETPLIVGGELARAKEEPHWAL
TD_19949          SERG-----PLDGGPRTETPVTVTPTREPPTPDTGNVQGPSVPAIS-GGEGCGRRRFAGEGNNIDNSLRITGPAGYPEETFYGEVPWVVR
TRIINF_149689028  ----------------------------------------------------------------KEDKRIIGGEETNVNEVPMWAG
ACYPIS_240848975  TNN-------QITTTTKKPSINPVLK---SYSATEMCSEYSELIYHRIKDPILISDQENYLKVKDC-YDAITLITNGTDAKPKEEPHWAL
APIDOR_572262138  TTT------IRTTTTKKTTAITTFLPFTAKMKTKAKCEEYSRYVYT-TEYPPTLINNKKPVNKSLCYIKDRTLIVGGTKAEAKEEPHWAA
ATHROS_817058472  VTEG-----IGLIDIRVKGACDNYLDICCKKPDVLKPEDKVTPKPP-VRRGCGQRNPDGVG------FRITGHTDN--EAQFGEEPWLVA

                                                  #                                                  #
TD_4045           IGHSEKDDES--AVINWQCGGSLISERWIVSAAHCTATRRG----LAKWATLGDLKVLSEEDEDYTDPVTLRIVERVDHPDYHSGRGYND
TD_19949          LL---FVDNGQINDTFLMCGGALITPSVVLTAAHCINNINN----QMVVRAGDWNIQQNFEPIA--PQDRNVIRTILHPQFNPRNLAND
TRIINF_149689028  LFYK--------PKELLFCGGSIITQYHILTAAHCTQPFEED--VRDIQIVSGEHDQDKVDESSS--TVYIDVLNFVPHEGYYLIGHRHD
ACYPIS_240848975  LGYSEKPD-----INSWACGGSLISKRFVLTAAHCEKLGSKNEPRFASWARLGELDYLSETDEAR--PKDYRIVQRIIHPNYKSTSSYHD
APIDOR_572262138  IGFDTLDG------IVWACGGTLISERFVLTAAHCTFNRNLN--STANWVRLGDLNLERSDDYAK--PENFRVIERIRNPHYKPPSQYHD
ATHROS_817058472  VLREEIIGQNAEKLNVYQCGGALIHPEVVLTAAHCVNGKQPS----QLKVRAGEWDTQTKNEVYP--HQDRNVKSVVVHEGFHSGALYND

TD_4045           IALYKLERDVKFNAYIRPICLN--TDTRPYSGKAIASGWGNTKWK--GSSSQDLLRVVLNLHDMNDCDNAYKNAQKPGLLIKTPQGISPH
TD_19949          IALIGLESPMQLSQTVDIVCTP-DEERVVNRSACLATGWGKNELSRYGTYQDTLRKVSLPLLDKTECQTKLRKTKLG------KYFTLLP
TRIINF_149689028  IAIILLKDKIVYTNIVGPACMP-TEPIDMIGHKIKVLGWGKLSAN--GPSSKVLMKVYLNVVPLEECNNTHEY-----------IGLVER
ACYPIS_240848975  IALFRLERDVDFSAFVRPICLN--TDNTLRPPAIIATGWGKTDVA--SLQSSHLLKVRINTISAEECDKSFLYLVNR--NEKLAQGIVDN
APIDOR_572262138  IALLKLESDVQFNAWIRPSCLPYSLPDSGLDGKATATGWGDVEWH--ESGSNDLLKVTISLVSQPKCNKLFIGNEKN---NKLKFGITED
ATHROS_817058472  VALLILETAVNLAENVDLTCLP-AHGETFDGTRCFASGWGKDIFGKEGHYQVILKRVDLPIVPHDVCQSSLRTTRLG------KYFNLHS

                        *    #                  * *
TD_4045           NLVCAW--ERGKDTCQGDSGGPLQIKLEQPYCMYSIIGITSFGDECA--KDLPGVYTRVSYYLPWIESIVWP------------
TD_19949          GFLCAGG-SQNSDTCEGDGGGPLYCQLQSEPNRFAVVGVISWGIDCG-KG-NPAVFADVNHYMPWIKQQIA------------
TRIINF_149689028  RQICTYH--PTKDSCQGDSGGPLLWLDREIN-RYVLAAVTSYGLSCA--TDVPGVNTNISYYMPWIQKVVAETRPEAPICTKQN
ACYPIS_240848975  LMVCAGNPEGGNDTCGGDSGGPIQIKHNRYTCMYSQIGITSFAGQFCGEKDSPAVYTRVSNYISWIESIVWPKM----------
APIDOR_572262138  SQICAG--ELGKDTCQGDSGGPLVILNRDYDCMYTLIGVTSLGKLCG--TIIPGIYTRVYNYISWIENIVWPDL----------
ATHROS_817058472  SFVCAGG-ESGKDTCKGDGGSPLVCPLKNDPSRYAQVGIVAWGIGCG-ENRVPGVYANVAAARHWIDEQMVYNNLDINTYQA--
```

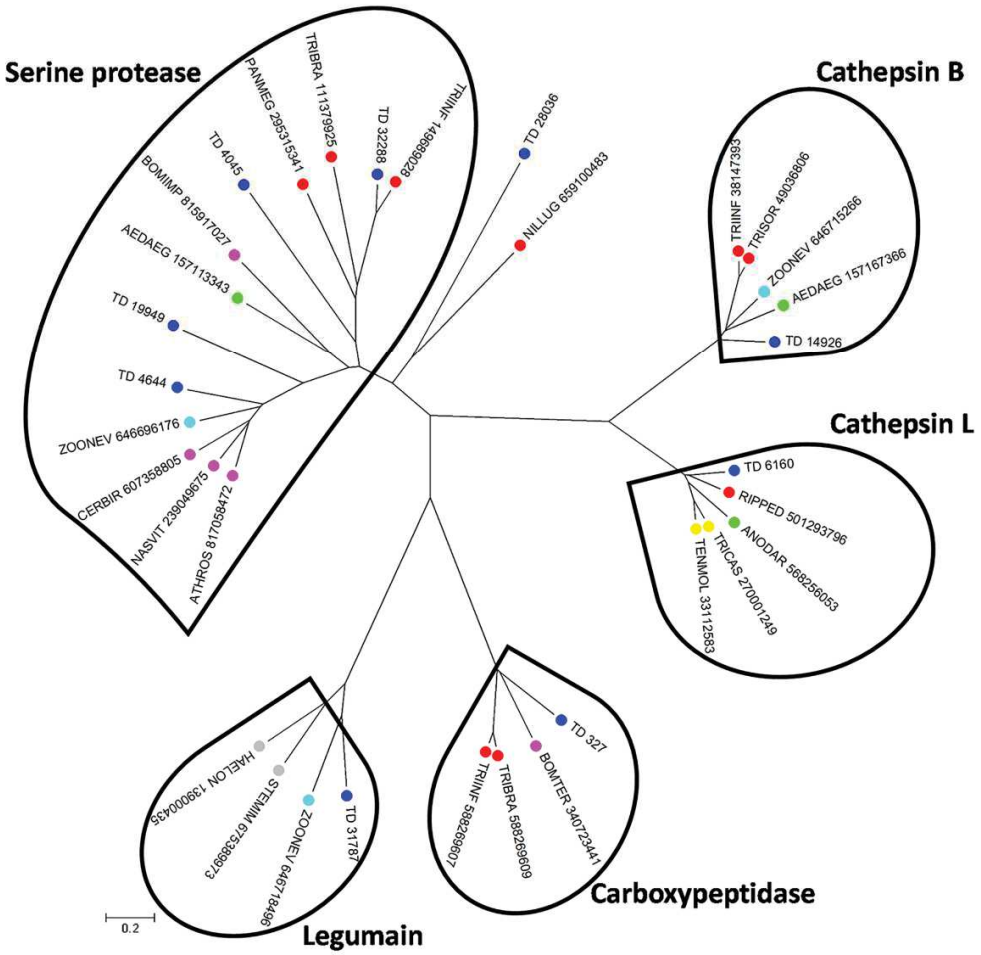


Fig 4

```
                                                                  #   #       #                   #       #               #
TD_30859          ---MKL-SSSVLSVVVIAVFLCMLDKHVFGDEVE-CNTDCGDKYEPVCA--QPSDEEKFEIFKNLCELDKAMKCN-DEKTWKFVSEGECTYRFVFISAISVTFYFITKNMLIKHA
TD_17576          ---MKF-ITSVLSVTVIAMILTVW-MIDFSEAQ--CSIVCTADYRPVCG--QRGQPPRRRTFGNACALSRENQCG-ARPPWRQIHTGECRN----------------------
TD_21160          ---MKL-NIPAFTLFIITTLILVW-SIELTEGQQQCNIACTGEYNPVCG--EQNQ--FFRVFTNRCRLDRENSCG-ANPPWEFVRIGQCP----------------------
TRIINF_149898876  MMKMKL-STSVLCVVVIAVFLS---MIDFGEAS--CNIVCPTDKKYVCG--QRPGEQ--RTFDNDCQIEKENKCG-GG-GWTIRNRGHCPTS--------------------
TRIINF_149898841  -MKMKL-STSTFSVVVIAVFLTVC-MFDFGEAGN-CPKTCTVTNQHVCG--QRVNEL--RTFNSLCEMEKENLCG-SG-GWTKLKNGHCST----------------------
PHLPAP_451935058  ---MKFLCIFALVALFVALSVGAP---AQKTNFNECLKACGYHYSPICAGPKEGAEKP-QTFGNMCALE-TYNCEHK-TEWEVKSQGECPGGGAIRLQ-----------------
TRIBRA_116267193  ---MKL-IFS-LKCALTTVFLTVC-MIDSGEAM--CPFVCSNDYRPVCG--GRDT--ERRTFDNQCKIDEENDCGRSGGGWELVKRGACP----------------------
HYBBIM_94730670   ---MKFALFSVLVVLLIATFVAA----------DECPRICTADYRPVCGTPSGGRRSANRTFGNQCSLN-AHNCLNKGDTYDKLHDGECK----------------------
TABYAO_241914367  ---MKFALFSVLVLMLIATFVAA----------DDCPRICTSDYTPVCGTPSGGRRSANRTFANQCGLD-SHNCLNKGDTYDKLHDGECK----------------------
```

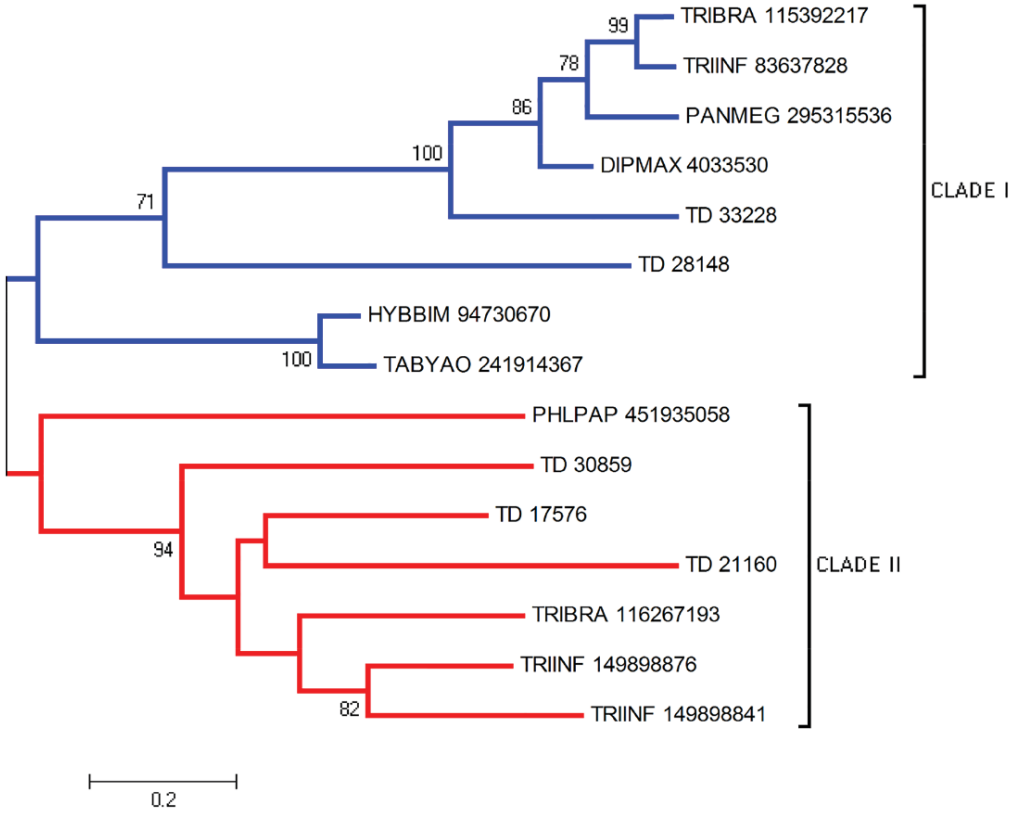


Fig 5

```
TD_30493            --------MMHQLLFMLIVLPAFS--LQQCLTPDD--SKITTD-PNARLALYKGEQQFALSLLQTINGLTPAENIFFSPFSVYNAMLLAY
NILLUG_482875806    -------MMILWMCVVAVTLPVFT--NQQCLTKDD--SKPSTD-PQARLALFRGQQEFSLAMLQTVNHMYPNQNIFFSPHSIYQAMLLSY
ZOONEV_646701027    ------MIIALVVGLTAVLLPGLS--VQQCFSPDD--TKVQNS-PNTRQDLFIGEQAFSLAMLREAVADNPTGNVFFSPVSVYNALLLAY
ACYPIS_229577153    MHWLCKCVVGLTVLATATRSATLPNVAQKCVPPGP--GLSLNDFLSGKFYLFERQQEFSIKLLQTAVTASPKQNLIFSPHSIYTALLITY
CULQUI_170053098    ---MTVMMKFFASIVALLLT-AAVAANAQCLAEDDNVQHTKTDNPLARTRLYKGESIFTLKLLEAINAATPSENVFFSPYSLYHVLLLMY
AEDAEG_157118702    -------MRSLPTVVLLVVALSARTISAQCLAQDDNIQHTKADNPLSRNRLYKGESIFTLKLLEAINTATPTENVFFSPYSLYHVLLLAY

TD_30493            FSASNHTEAALKKSLFIPETQDKLSTMQAYRLEKYFQKMRAYNG-SASYELSSANRLFVSPQQSVRECMKLLFDDEITVTDFAKDPAAAT
NILLUG_482875806    FVAANHTEASIKKAIFLPKQQDKLSTMQAYRLEKFFQSMRLVNG-SDSYELRSANRLFVSQQQKVKECMLELFKDEVQQVDFAKS-AESA
ZOONEV_646701027    FGAANQTELSLKAALRIPESQSKISTMQAYRFEKYFQSMREING-SESYELRSANKLFIAKRLKLRDCIAHLFEDEVEPVDFVANPEAVR
ACYPIS_229577153    FLSNNETEETLKLFLNLPPEQSKLSVMQAYRMEKLFHSMRALNT-SGDYEFSSVDRLFVSQRLPIQKCIADVFKDEVHKMDFVVDPELAR
CULQUI_170053098    FGAKSDTEQTLRKGLELHWTEDKPVVWQAYNVGKKALAMRFGQNKADEIQFVSVDKLFFGKQIPVEECMEDKFDEIEKLDFEKDPESQR
AEDAEG_157118702    FGARAETEKMLRNGLELHWTEDKPVVWQAYNIGKKSLAARFSQ--SGDIQFTSVDKLFFGKQVAIRDCMEDKFEEIEKLDFEKDADGQR

TD_30493            EQINKWVADQTKNQIKDLLHEGQLTSLTQLVLANAAYFKGLWKNKFQPESTRKEVFYISGSQNAFVTMMRQKSTFNHLVSEKLGAHILEL
NILLUG_482875806    QVINQWVANQTKNNIKELIPEGSISETTQLILTNAAYFKGLWKSKFLKSNSRKEIFYINSSKNAFVTMMRQKGTFNHAVSEQLGAHILEL
ZOONEV_646701027    QNINNWVEIQTKRHIQDLIPKDRITADTELVLANAAYFKGLWKSQFLPEHTRKDVFYVSHTKQTLVDMMKQKGTFNHIVSENLGAHVLEL
ACYPIS_229577153    MYINNWVTNQTNQQIKDLLPEGKITYETRLVLANAAYFKGFWSSRFSSESTKDEVFYTSPTENAYVPMMWQSGVFNLLSSDELGAHVLEL
CULQUI_170053098    LYINNWVENTTHGEITDLLIPGSITKNTKLAIANAAYFKGTWQSKFKPEETKKEIFYVSNERQEFVDMMHVEGTFNHAANEKLGCHILEL
AEDAEG_157118702    LYINNWVENVTYGEIKDLLIPGSITKQTKLAIANAAYFKGTWQTEFKPEQTKKEIFYVSSERQEFVDMMHVLGTFSHAANEKLGCHVLEM

TD_30493            PYKGDD----VSMYILLPPFASPNGITNILKRLT--LDKFQEVASEGFMVPRAVDVSVPKFTIEKELE-MTQILESLGVGDVFKPTSDLS
NILLUG_482875806    PYKGDD----VSMFVLLPPFASPSGITNILKRLT--LKTLHEIIDEDSMIPRAVEVSIPKFEVEKSIE-LVQILTSFGI-DMFEDTADLS
ZOONEV_646701027    PYKGSE----VSMFIFLPPYAKSNGIEAVIKQLS--VEGLQDIVEED--SLRAVEVVIPKFTVEQSLE-LTPILERLGIGDLFQFTSDLS
ACYPIS_229577153    PYKGGD----ISLFVILPPFNKQRGISQLSKRLN--TNILQNIVSSGEWTSRPVEVSLPKFNVEQSMDNLVSLLEQMGIGNVFKGKVDLS
CULQUI_170053098    PYTGQDEASRVSMFVFLPP-TESNALDKLLARLTSSPDILSEVVNEG--ISRRVDVKLPKFSIEKTVD-MKTVLGRLGMGKLFENSANFS
AEDAEG_157118702    PYKGQDESTRISMFVFLPP-AVPNSLEQLLARLTSDTGILSEIVNEG--IPRMVDVKFPKFNIEKTVE-LKPVLERLGLGKMFENGADFT

TD_30493            GLTGKPG--LSLDECVHKAKIMVDEDGTTAAAATAVFNFRSSRPLDPARFICNHPFVYLIFDKVSQTILFTGIFNKPPPQQATKQ
NILLUG_482875806    SLTDATGPRVRFTDAVHKARLQVDEDGTTAAAASA----------------------------------------------------
ZOONEV_646701027    GFTEDGD--LHLDKAVHKAKISVDEQGTVAAAASAVISFRSSRPLTPARFIANHPFIYLLFDKTSQTIQFMGVYRSPNN------
ACYPIS_229577153    ALTGNAN-DVTIDDAVHKAKISVDEEGTIAAAATSFLMSRSSRPVEPFKFRCNHPFIYFMFDKVTGSVLFMGVYNSPKSN-----
CULQUI_170053098    GFSKKAQ--IGFDEVLQKSKITVNEEGATAASATVAFSFRSSRPADPAMFHCNHPFIFVIYDYATQAVLFNGVYRQPE-------
AEDAEG_157118702    AFSEKEQ--IGFDEILQKSKIEVNEKGSTAASSTVLFSFRSSRPVEPAMFHCNHPFVFLIYDYGTQAVLFNGVYRQPE-------
```

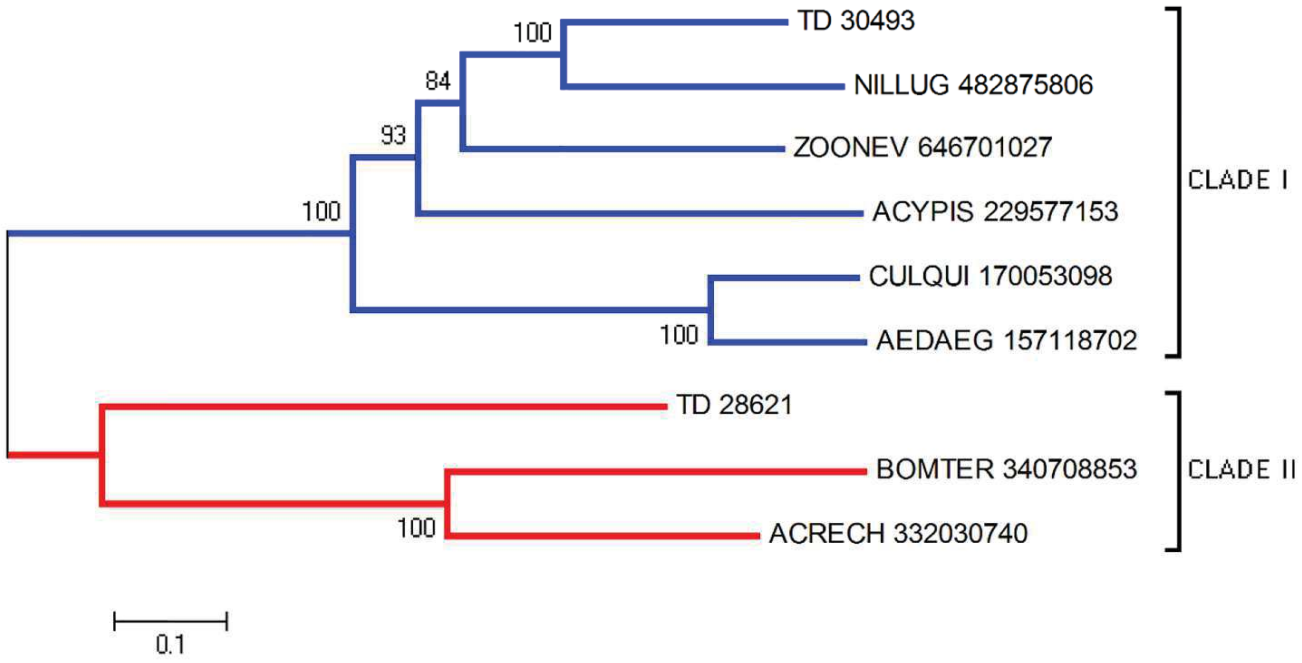


Fig 6

```
TD_7118           MRLLLIFVTLILGYCLAAPERVGDKVPKK----SCEPGTTWKEECHSGVCTSEGK-IDCTKEACPEQLVP----------
TD_3748           MLTKFQLILLPVALVFADPEFE----------TCVPGSTWRKDCNTCTCNLHGR-IVCTQLQCDPDQ------------
ANOSIN_668451255  MKIPILILATLVVGALCEEK-------------CVPGTTFMEDCNKCRCSNDGQ-KACTRKMCPPKELSDDAQVR----
AEDAEG_157103392  MRSLLLIAAVLVAGALSLEAGK----------QCEPGSTFKLDCNTCRCGADGKVMSCTRKFCVPDVQGDDPKAQPMVD
PACLEN_1764109    MKALLILVMTVAAHGASLEQPDPTPASDLPDKSLCAPGSRWKNECNWCSCADHGL-ALCTLMGCFPGYKAAQGESVCSEG

TD_7118           -------------------RSKRSVERQDCK----------------EGETWKEDCHNCFCASGKP--ACTKELCESEAS
TD_3748           ---PNLIYG-----------TISVERRLCS----------------IGELWFEHCNTCICLG------LERAVCTKYDC
ANOSIN_668451255  --TEPVPNGDSLSSADEKDEVHLQTNGQVCT----------------PNEIKMQECNRCRCANNGIGWFCTRRVCPPREK
AEDAEG_157103392  EVSDNLQPDASAAAAAEKEEEEVHTNGQVCT----------------PNEVKMEDCNRCKCAANGIGWFCTRKACPPRER
PACLEN_1764109    SRWKADDCNWCRCIDGSPSCTKRLCRTKLAKGMFASQTEETECYGDPDTNRWRIECNWCRCVNGKG--SCTRKGCPQVIN

TD_7118           HKVVLPSSDDD-------------------------------------------------------------------
TD_3748           HIDNWVHNKIR-------------------------------------------------------------------
ANOSIN_668451255  RAAVTTKCTPGSTFMADDGCNTCHCNEDGRAACTRVMCEPVEGAELIRFRRESPESPAGPKCTPGTTFMAEDGCNTCYCD
AEDAEG_157103392  RSASEFLTKPK-----------------------------------RQVAQPAPPKKECEPGTNFKSADGCNDCFCT
PACLEN_1764109    GIGLANTNECEG-----------------------------------TPTWTKGCNTCSCVNGSAQCTTEECDKLVQS

TD_7118           ------------------------ESNDHQTTKREAECAECAVWK--DDCHTCHCAY-GKKACTRELCPSESD-------
TD_3748           -------------------------KRRETVINAGRKCTPGTTWK--EDCNTCFCSDKGFAGCTLMACPSHLN------R
ANOSIN_668451255  AEGRAACSLKFCASSDVSAFRGRLQRRQANEFPGMQKCTPGTTFRDADDCNTCYCNEDGRAACTRKACIKKDSELYKQLQ
AEDAEG_157103392  ATGIAACTQKFCFNTKT--------KRDTAALAPKKECVPGTTFKSADGCNDCFCSESGHAACTMKFCFPTPP------R
PACLEN_1764109    PSVPAVAFRSGGRTGKCR------PDAHDDSLPDYGQCVPGSRWK--KDCNWCSCTETAIGMCTLIGCLNYEPKPGEAVC

TD_7118           ------------------------------------------------------------LRVTVSEDNERTK
TD_3748           PKRD--------------------------------QGRKCTPGTTWK--EDCNSCFCSEKGFVGCTLMACHSSHLNRPK
ANOSIN_668451255  QARESSNSDA-----------SLTADTDGQRQRRQVKEIKCTPGTTFKSDDGCNNCFCDESGQAACTLMACLP-NTERVQ
AEDAEG_157103392  TKRLASNS----------------------------KDTQCVPGTSFKSADGCNDCFCSANGMAACTLRFCFFEDQPQIK
PACLEN_1764109    TDGSKWKDDCNWCTCNNGSASCTEKLCQYKPDGSLPDNDMCVPGSRWK--DECNWCWCEANGAAPCTRMGCSEDYKPQPG

TD_7118           RDYES--------------SED-----SSSSSEEEKCPPNVKEWKEDCNQCYCDKGVK--VCTKALCPDQGPVVAIKPKQ
TD_3748           RDQAE--------------EEQ----SEESLQELEKNCEIGTTIKLDCNICYCTAMG--LACTKRLCHGHEVSVVVEE--
ANOSIN_668451255  QKRQT--------------SGD-LPQTAIAPGTPGFSCTPRSSFKYQCNTCLCSDDGKNAGCTFKFCIPGEW--------
AEDAEG_157103392  SKRQA--------------NNDELPVSAVAPGAPGFSCSPGKSFKYQCNTCRCDTSGQTAACTFKFCLPGEY--------
PACLEN_1764109    EAVCIDGSRWKVDCNWCTCNNGSSACTEKLCLKPGGQCTEGESWRQDCNMCSCSTGLR--ICSVKGCPPTPT--------
```

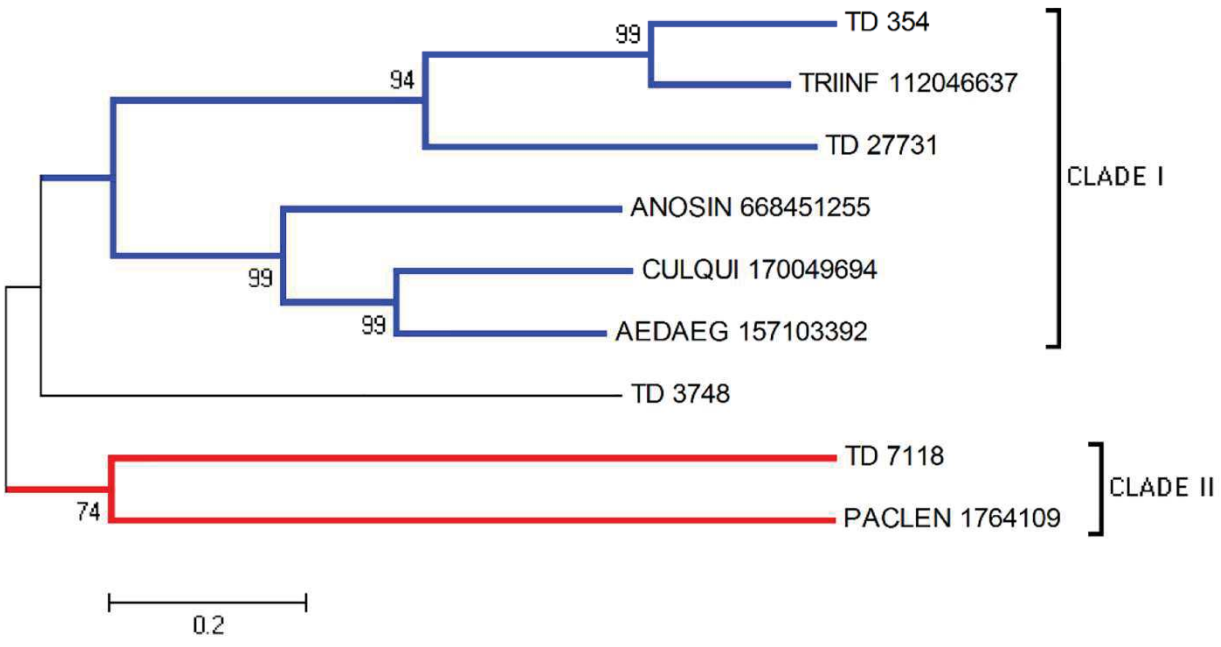


Fig 7

```
                                                       #                            #   #
TD_4743             MKGFTAVCVFTIAITVSLAIDEAAK-KKAINTFNKCKEQHPITDAELEQIKKHE-GLPSSQNAKCLAKCMLTEGNILKDG
ADELIN_270000360    MNSLI--PVLLVVCAAATRADEQTN-AMVAKAFNKCREEFPISDDEIGGVREKT-TIPESHNAKCLMACMLREGKMLRDG
EUSHER_299474028    MKGITTALAFTLLLAT-IEAEENVK-AKIMKVFNTCKEKHPVTDDELAAFRKADIGFGYSHDAKCMLACMLEEGKMLKDG
LYGLIN_573006010    MKFVLSAAVVLLVAAAVKANEKKAN-EKVTEIFNKCKETWPVTDEEIEQVKQKQ-SIPESKNVKCILACMLKEAKILRDG
APOLUC_387158189    MNPLI--LILLVVFAAATRGEEQAN-ALVAKAFNKCFGEFPLGDDEMKEVKDKS-TVPSSHNAKCLMACMLKEGRILRGG
NILLUG_516299413    MLYLLEFVIFSLALTASIPQIMAADSPDMLAVFNKCREEAGATEDDIKNFRAQQ--IPSTTTGKCMLACMFNHSGLMKDD

                                                   #             #        #
TD_4743             KYKKDIAIAFTETLHSDSAEEAEKARQVIEHCASTVGTDVGSEDLCEYAYKMAQCGYSKAKEIGLEKPEWE-
ADELIN_270000360    KYEKENALIMADVLNKDDPASADKAKQLVETCAGKVGTDAGG-DECEFAYKMAVCAAEEAKKLGVRPPDF--
EUSHER_299474028    KYLKENALIMADVFHIDDFDEAAKARKVIENCSVEV-SEVGP-DQCEFAYKMAVCGANESKKLGMKEHDFFE
LYGLIN_573006010    EYNKENAELMADVLYKDEPEHAXKSKQIIEMCSAELGTKTEG-DDCEYAYKMSVCASKHAKELGVKTPEF--
APOLUC_387158189    KYELENAILMADVLNKNDHAATDKAKQLIETCAAQVGTDASA-DECEFAYKMALCASDEAKKLGVRPPDF--
NILLUG_516299413    KYNSEGALKLVSQVFADNPVNLGKAKQLINGCTDEVKNENDK---CEIASKIAHCTVKMTSVVGLS------
```

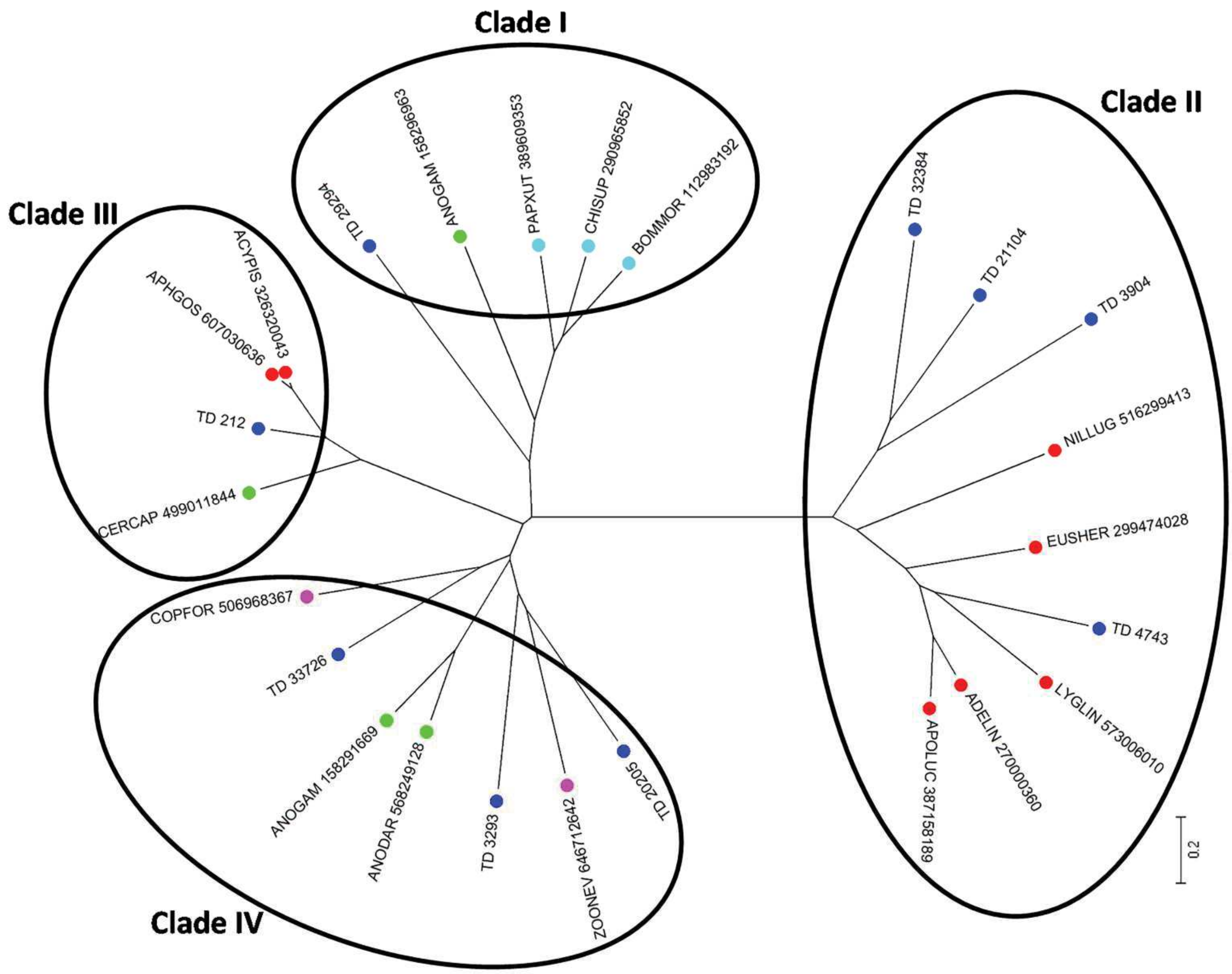
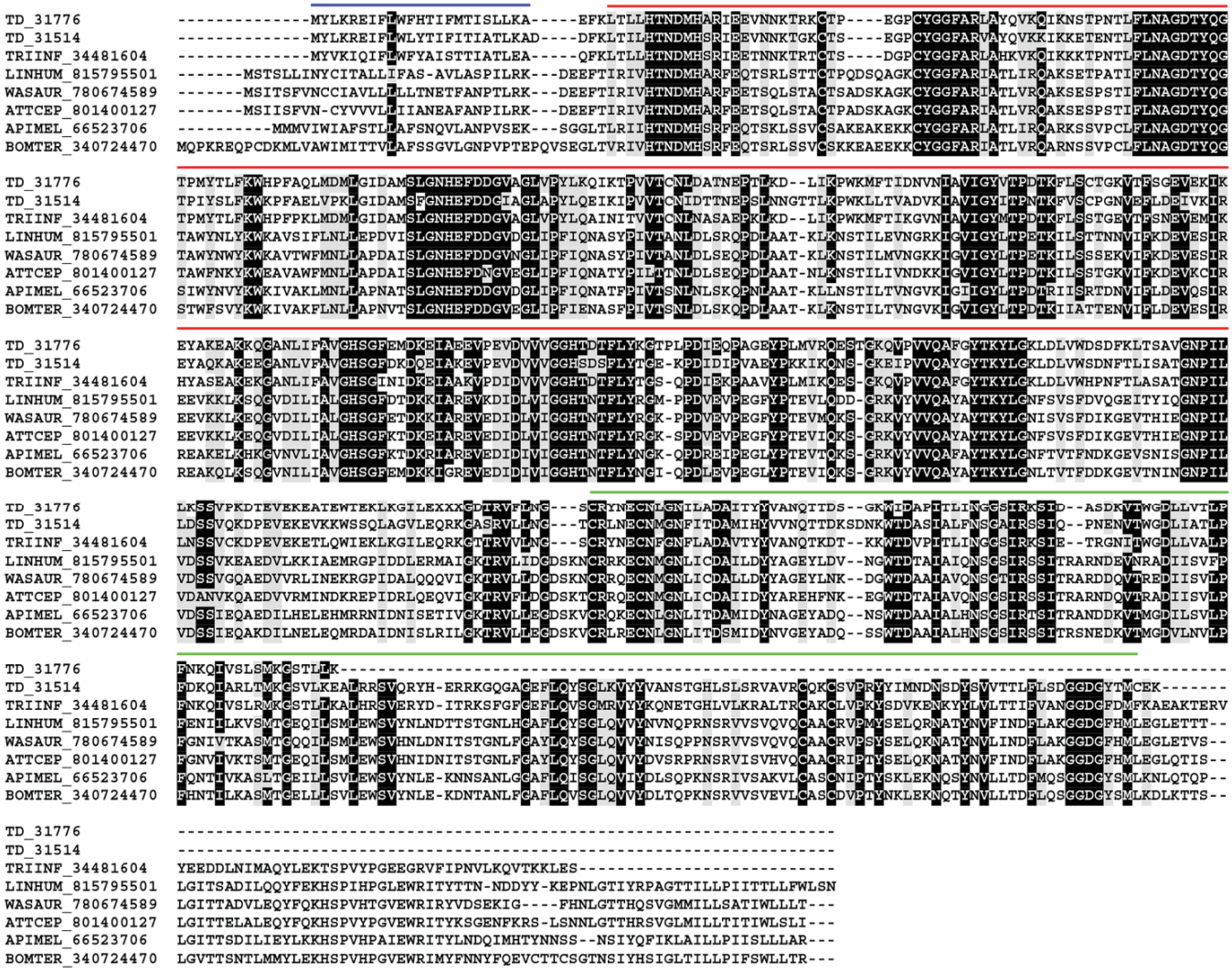


Fig 8


TD_31776
TD_31514
TRIINF_34481604
LINHUM_815795501
WASAUR_780674589
ATTCEP_801400127
APIMEL_66523706
BOMTER_340724470

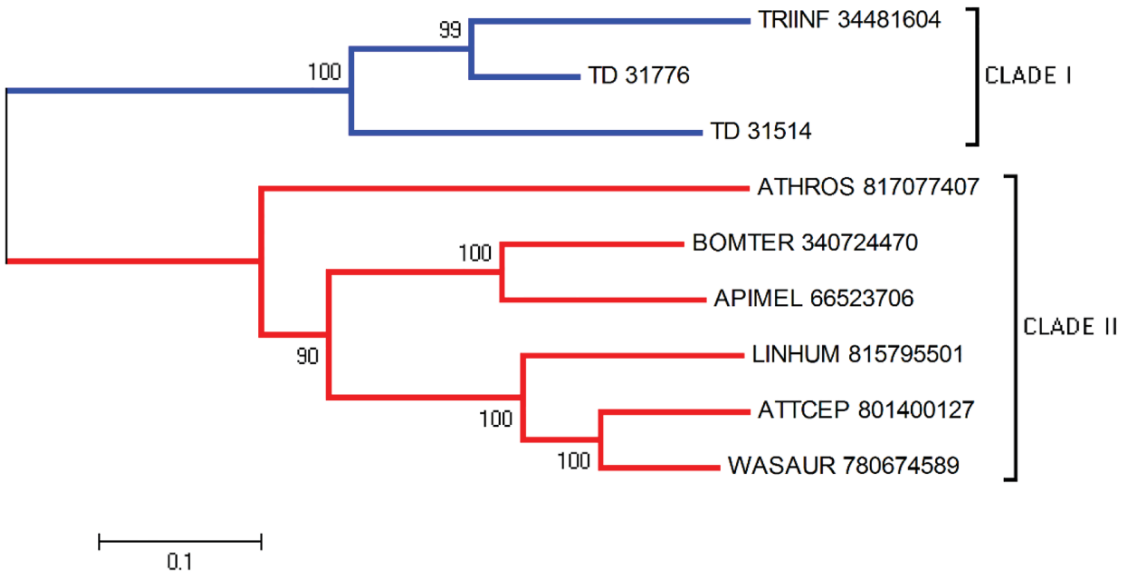

TRIINF 34481604
TD 31776
TD 31514
CLADE I
ATHROS 817077407
BOMTER 340724470
APIMEL 66523706
LINHUM 815795501
ATTCEP 801400127
WASAUR 780674589
CLADE II
99
100
100
90
100
100
0.1

Fig 9

CONCLUSÕES

O conjunto dos resultados obtidos nesta tese gerou um catálogo dos transcritomas das glândulas salivares dos triatomíneos *R. neglectus* e *T. dimidiata,* insetos vetores da DC. No que se refere à análise transcritômica, foram obtidos 5.705 transcritos a partir das GS de *R. neglectus* e 5.704 de *T. dimidiata*. Os dados obtidos estão disponíveis no banco de dados do NCBI (*National Center for Biotechnology Information*), fonte de informação de livre acesso na internet, e agora podem ser utilizados como recurso científico em diferentes projetos de pesquisa. Os endereços de acesso são os seguintes:

http://www.ncbi.nlm.nih.gov/biosample/3975952

http://www.ncbi.nlm.nih.gov/biosample/4240765

O uso de RNA-seq foi fundamental para o desenvolvimento deste trabalho. Como resultado, gerou-se um grande volume de dados, favorecendo as investigações científicas acerca da complexidade bioquímica e farmacológica da saliva dos triatomíneos e as distintas estratégias anti-hemostáticas das sialogeninas.

Dentre as proteínas com classificação funcional extraída durante as análises encontram-se as lipocalinas, antígeno-5 e apirases, moléculas previamente encontradas em outros sialomas. Outras proteínas, como proteases (metaloproteases, desintegrinas e serino proteases), OBP, e inibidores de proteases, não tão comumente descritas na saliva de insetos hematófagos, também puderam ser observadas. Os resultados obtidos evidenciaram a característica multifuncional do repertório salivar dos triatomíneos estudados, composto predominantemente por moléculas com funções (algumas hipotéticas) diretamente relacionadas a diferentes atividades anti-hemostáticas. Adicionalmente, os resultados relativos ao transcritoma das GS de *T. dimidiata* ampliaram o conjunto de sequências depositadas no banco de dados do NCBI, de 388 para 3.815.

O hábito hematofágico evoluiu de forma independente várias vezes entre os hematófagos, o que leva à diversidade das moléculas encontradas na saliva. Assim, a partir de uma combinação de evolução convergente e seleção positiva, é possível observar uma diversificada composição salivar em insetos hematófagos (Ribeiro, Schwarz e Francischetti, 2015). Esta observação reflete a existência de diferentes estratégias anti-hemostáticas entre as espécies. Neste trabalho, a análise de espécies pertencentes a diferentes tribos, Rhodniini e Triatomini revelou uma distinta distribuição dos transcritos putativamente secretados. Enquanto *R. neglectus* revelou uma composição salivar

abundante tanto de serino proteases quanto de lipocalinas, *T. dimidiata* revelou uma abundância significativa de lipocalinas, característica comum da tribo Triatomini.

A análise proteômica do conteúdo salivar resultou na identificação de peptídeos de massa relacionados aos transcritos encontrados nas glândulas salivares, validando o transcritoma nos dois estudos. Os resultados mostraram que, em geral, aqueles transcritos observados em maior número durante as análises do transcritoma, como as lipocalinas, foram os mais facilmente identificados.

Durante as investigações desenvolvidas neste trabalho explorou-se uma abordagem evolutiva através da busca por homologia e construção de árvores filogenéticas. Muitas vezes, as sequências analisadas foram homólogas a proteínas de diferentes linhagens filogenéticas de artrópodes, inclusive não hematófagos. Esses achados suportam a ideia de que o comportamento hematofágico pode ter evoluído a partir de artrópodes não hematófagos, como hemípteros predadores ou sugadores por exemplo. A pressão evolutiva para o hábito hematofágico teria selecionado indivíduos que apresentavam proteínas salivares adaptadas para novas funções, no caso, funções anti-hemostáticas. O que impulsionou esses eventos evolutivos ainda não está claro, mas, possivelmente, a forte associação entre o inseto e o hospedeiro vertebrado é um cenário que pode ter influenciado o sucesso desta relação de parasitismo.

Por fim, este trabalho amplia o conhecimento não só para a compreensão da biologia dos triatomíneos e do papel da saliva durante a atividade vetorial, mas auxilia no entendimento da dinâmica estabelecida entre vetor-parasito-hospedeiro e o processo de transmissão de doenças.

PERSPECTIVAS

A análise do sialoma dos hematófagos identifica as moléculas presentes nas glândulas salivares providenciando um banco de dados contendo suas sequências, e o uso das tecnologias científicas modernas ampliaram o potencial para a descoberta de novas moléculas e funções biológicas. Quando nos referimos às sialogeninas, um grande conteúdo científico está disponível, contudo, ainda há muito a ser descoberto. Diversas moléculas ainda possuem função desconhecida na saliva, assim, diferentes projetos que visam a caracterização funcional de moléculas individuais ainda devem ser realizados. Um bom exemplo seria a investigação das serino proteases de *R. neglectus*, respondendo se, de fato, as mesmas são abundantemente secretadas durante a salivação; e ainda, qual seria seu papel biológico para o sucesso do repasto. Sem dúvida, outras proteínas, como os inibidores de serino proteases ou as OBPs, também precisariam ter seu papel funcional e seus alvos específicos investigados.

Uma abordagem adicional seria o estudo funcional das desintegrinas, explorando seu possível potencial para aplicação terapêutica anti-angiogênica. Moléculas vasodilatadoras e anti-coagulantes, como as nitroforinas e as triabinas, também poderiam ser avaliadas para uso farmacológico em desordens vasculares. Outros temas que se abrem com a descrição do sialoma de um triatomíneo seriam: assistência ao estudo do processo evolutivo dos insetos hematófagos; papel da saliva (moléculas individuais) dos hematófagos na transmissão de microorganismos infecciosos durante o processo de repasto; ou na dinâmica da interação vetor-parasito-hospedeiro.

REFERÊNCIAS

ABAD-FRANCH, F. et al. Biogeography of Triatominae (Hemiptera: Reduviidae) in Ecuador: implications for the design of control strategies. **Mem Inst Oswaldo Cruz,** v. 96, n. 5, p. 611-20, Jul 2001. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11500757 >.

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. **Imunologia celular e molecular**. 6ª Edição. ELSEVIER editora LTDA, 2008. 576 p.

ACEVEDO, F.; GODOY, E.; SCHOFIELD, C. J. Comparison of intervention strategies for control of Triatoma dimidiata in Nicaragua. **Mem Inst Oswaldo Cruz,** v. 95, n. 6, p. 867-71, 2000 Nov-Dec 2000. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11080777 >.

ALMEIDA, P. S. D.; AL, E. Levantamento da fauna de Triatominae (Hemiptera: Reduviidae) em ambiente domiciliar e infecção natural por Trypanosomatidae no Estado de Mato Grosso do Sul. Uberaba: **Rev. Soc. Bras. Med. Trop**. v. 41, n. 4, p. 374-380 2008.

ANDERSEN, J. F. et al. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. **J Biol Chem,** v. 278, n. 7, p. 4611-7, Feb 2003. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12464610 >.

ANDERSEN, J. F.; MONTFORT, W. R. The crystal structure of nitrophorin 2. A trifunctional antihemostatic protein from the saliva of Rhodnius prolixus. **J Biol Chem,** v. 275, n. 39, p. 30496-503, Sep 2000. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10884386 >.

ANDERSEN, J. F. et al. Insight into the sialome of the Black Fly, Simulium vittatum. **J Proteome Res,** v. 8, n. 3, p. 1474-88, Mar 2009. ISSN 1535-3893. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19166301 >.

ARCÀ, B. et al. An insight into the sialome of the adult female mosquito Aedes albopictus. **Insect Biochem Mol Biol,** v. 37, n. 2, p. 107-27, Feb 2007. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17244540 >.

ASSUMPÇÃO, T. C. et al. Insight into the salivary transcriptome and proteome of Dipetalogaster maxima. **J Proteome Res,** v. 10, n. 2, p. 669-79, Feb 2011. ISSN 1535-3907. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21058630 >.

ASSUMPÇÃO, T. C. et al. An insight into the sialotranscriptome of Triatoma matogrossensis, a kissing bug associated with fogo selvagem in South America. **Am J**

Trop Med Hyg, v. 86, n. 6, p. 1005-14, Jun 2012. ISSN 1476-1645. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22665609 >.

ASSUMPÇÃO, T. C. et al. An insight into the sialome of the blood-sucking bug Triatoma infestans, a vector of Chagas' disease. **Insect Biochem Mol Biol,** v. 38, n. 2, p. 213-32, Feb 2008. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18207082 >.

BARRETO-SANTANA, D. et al. [Comparative biology and feeding behavior of Rhodnius neglectus and Rhodnius robustus (Triatominae) under laboratory conditions]. **Rev Soc Bras Med Trop,** v. 44, n. 4, p. 490-5, 2011 Jul-Aug 2011. ISSN 1678-9849. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21860897 >.

BARROS, V. C. et al. The role of salivary and intestinal complement system inhibitors in the midgut protection of triatomines and mosquitoes. **PLoS One,** v. 4, n. 6, p. e6047, 2009. ISSN 1932-6203. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19557176 >.

BAUER, K. A. et al. Factor IX is activated in vivo by the tissue factor mechanism. **Blood,** v. 76, n. 4, p. 731-6, Aug 1990. ISSN 0006-4971. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/2383653 >.

BEARD. Kissing bugs and bedbugs, the heteroptera. In: Marquardt WC (ed). Biology of Disease Vector. 2ª edição. **Elsevier Academic Press**, 2005.

BELKAID, Y. et al. Development of a natural model of cutaneous leishmaniasis: powerful effects of vector saliva and saliva preexposure on the long-term outcome of Leishmania major infection in the mouse ear dermis. **J Exp Med,** v. 188, n. 10, p. 1941-53, Nov 1998. ISSN 0022-1007. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/9815271 >.

BOWMAN, A. S.; DILLWITH, J. W.; SAUER, J. R. Tick salivary prostaglandins: Presence, origin and significance. **Parasitol Today,** v. 12, n. 10, p. 388-96, Oct 1996. ISSN 0169-4758. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15275289 >.

BRENER Z.; ANDRADE Z. Trypanosoma Cruzi e Doença de Chagas. **Nova Guanabara:** 440 p. 2000.

BRODIE, T. M. et al. Immunomodulatory effects of the Lutzomyia longipalpis salivary gland protein maxadilan on mouse macrophages. **Infect Immun,** v. 75, n. 5, p. 2359-65, May 2007. ISSN 0019-9567. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17339357 >.

BROZE, G. J. Tissue factor pathway inhibitor and the revised theory of coagulation. **Annu Rev Med,** v. 46, p. 103-12, 1995. ISSN 0066-4219. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7598447 >.

CALVO, E. et al. Function and evolution of a mosquito salivary protein family. **J Biol Chem,** v. 281, n. 4, p. 1935-42, Jan 2006. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16301315 >.

CALVO, E. et al. Aegyptin, a novel mosquito salivary gland protein, specifically binds to collagen and prevents its interaction with platelet glycoprotein VI, integrin alpha2beta1, and von Willebrand factor. **J Biol Chem,** v. 282, n. 37, p. 26928-38, Sep 2007. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17650501 >.

CAVALCANTE, R. R.; PEREIRA, M. H.; GONTIJO, N. F. Anti-complement activity in the saliva of phlebotomine sand flies and other haematophagous insects. **Parasitology,** v. 127, n. Pt 1, p. 87-93, Jul 2003. ISSN 0031-1820. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12885192 >.

CHAMPAGNE, D. E. Antihemostatic strategies of blood-feeding arthropods. **Curr Drug Targets Cardiovasc Haematol Disord,** v. 4, n. 4, p. 375-96, Dec 2004. ISSN 1568-0061. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15578959 >.

CHAMPAGNE, D. E. Antihemostatic molecules from saliva of blood-feeding arthropods. **Pathophysiol Haemost Thromb,** v. 34, n. 4-5, p. 221-7, 2005. ISSN 1424-8832. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16707932 >.

CHAMPAGNE, D. E.; RIBEIRO, J. M. Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito Aedes aegypti. **Proc Natl Acad Sci U S A,** v. 91, n. 1, p. 138-42, Jan 1994. ISSN 0027-8424. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8278354 >.

CHAMPAGNE, D. E. et al. The salivary gland-specific apyrase of the mosquito Aedes aegypti is a member of the 5'-nucleotidase family. **Proc Natl Acad Sci U S A,** v. 92, n. 3, p. 694-8, Jan 1995. ISSN 0027-8424. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7846038 >.

CHARLAB, R. et al. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly Lutzomyia longipalpis. **Proc Natl Acad Sci U S A,** v. 96, n. 26, p. 15155-60, Dec 1999. ISSN 0027-8424. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10611354 >.

CHENG, Y.; WU, H.; LI, D. An inhibitor selective for collagen-stimulated platelet aggregation from the salivary glands of hard tickHaemaphysalis longicornis and its mechanism of action. **Sci China C Life Sci,** v. 42, n. 5, p. 457-64, Oct 1999. ISSN 1006-9305. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18726508 >.

CHMELAR, J. et al. Insight into the sialome of the castor bean tick, Ixodes ricinus. **BMC Genomics,** v. 9, p. 233, 2008. ISSN 1471-2164. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18489795 >.

CLARKE, K. et al. Comparative analysis of de novo transcriptome assembly. **Sci China Life Sci,** v. 56, n. 2, p. 156-62, Feb 2013. ISSN 1869-1889. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23393031 >.

COMPEAU, P. E.; PEVZNER, P. A.; TESLER, G. How to apply de Bruijn graphs to genome assembly. **Nat Biotechnol,** v. 29, n. 11, p. 987-91, Nov 2011. ISSN 1546-1696. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22068540 >.

COUGHLIN, S. R. Protease-activated receptors in hemostasis, thrombosis and vascular biology. **J Thromb Haemost,** v. 3, n. 8, p. 1800-14, Aug 2005. ISSN 1538-7933. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16102047 >.

COURA, J. R. Chagas disease: what is known and what is needed--a background article. **Mem Inst Oswaldo Cruz,** v. 102 Suppl 1, p. 113-22, Oct 2007. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17992371 >.

COURA, J. R. The discovery of chagas disease (1908-1909): great successes and certain misunderstandings and challenges. **Rev Soc Bras Med Trop,** v. 46, n. 4, p. 389-90, 2013 Jul-Aug 2013. ISSN 1678-9849. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23982093 >.

COURA, J. R.; BORGES-PEREIRA, J. Chagas disease. What is known and what should be improved: a systemic review. **Rev Soc Bras Med Trop,** v. 45, n. 3, p. 286-96, Jun 2012. ISSN 1678-9849. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22760123 >.

COURA, J. R.; DIAS, J. C. Epidemiology, control and surveillance of Chagas disease: 100 years after its discovery. **Mem Inst Oswaldo Cruz,** v. 104 Suppl 1, p. 31-40, Jul 2009. ISSN 1678-8060. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19753455 >.

CUBA, C. A. et al. The triatomines of northern Peru, with emphasis on the ecology and infection by trypanosomes of Rhodnius ecuadoriensis (Triatominae). **Mem Inst**

Oswaldo Cruz, v. 97, n. 2, p. 175-83, Mar 2002. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12016438 >.

DE SOUZA, W.; DE CARVALHO, T. M.; BARRIAS, E. S. Review on Trypanosoma cruzi: Host Cell Interaction. **Int J Cell Biol,** v. 2010, 2010. ISSN 1687-8884. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20811486 >.

DIAS, J. C. Southern Cone Initiative for the elimination of domestic populations of Triatoma infestans and the interruption of transfusional Chagas disease. Historical aspects, present situation, and perspectives. **Mem Inst Oswaldo Cruz,** v. 102 Suppl 1, p. 11-8, Oct 2007. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17891281 >.

DIAS, J. C. Elimination of Chagas disease transmission: perspectives. **Mem Inst Oswaldo Cruz,** v. 104 Suppl 1, p. 41-5, Jul 2009. ISSN 1678-8060. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19753456 >.

DIAS, J. C.; AMATO NETO, V. [Prevention concerning the different alternative routes for transmission of Trypanosoma cruzi in Brazil]. **Rev Soc Bras Med Trop,** v. 44 Suppl 2, p. 68-72, 2011. ISSN 1678-9849. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21584360 >.

DIOTAIUTI, L.; DIAS, J. C. [Occurrence and biology of Rhodnius neglectus Lent, 1954 in palm trees of suburban areas of Belo Horizonte, Minas Gerais]. **Mem Inst Oswaldo Cruz,** v. 79, n. 3, p. 293-301, 1984 Jul-Sep 1984. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/6443015 >.

DORN, P. L.; MONROY, C.; CURTIS, A. Triatoma dimidiata (Latreille, 1811): a review of its diversity across its geographic range and the relationship among populations. **Infect Genet Evol,** v. 7, n. 2, p. 343-52, Mar 2007. ISSN 1567-1348. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17097928 >.

DUMONTEIL, E. et al. Geographic distribution of Triatoma dimidiata and transmission dynamics of Trypanosoma cruzi in the Yucatan peninsula of Mexico. **Am J Trop Med Hyg,** v. 67, n. 2, p. 176-83, Aug 2002. ISSN 0002-9637. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12389944 >.

DUNKELBERGER, J. R.; SONG, W. C. Complement and its role in innate and adaptive immune responses. **Cell Res,** v. 20, n. 1, p. 34-50, Jan 2010. ISSN 1748-7838. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20010915 >.

FAUDRY, E. et al. Triatoma infestans apyrases belong to the 5'-nucleotidase family. **J Biol Chem,** v. 279, n. 19, p. 19607-13, May 2004. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/14985353 >.

FLOWER, D. R. The lipocalin protein family: structure and function. **Biochem J,** v. 318 (Pt 1), p. 1-14, Aug 1996. ISSN 0264-6021. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8761444 >.

FORATTINI, O. P. Biogeografia, origem e distribuição da domiciliação de triatomíneos no Brasil. **Rev. Saúde Pública**. São Paulo, 2006. v. 40, n. 6**:** p. 964-998.

FRANCISCHETTI, I. M. Platelet aggregation inhibitors from hematophagous animals. **Toxicon,** v. 56, n. 7, p. 1130-44, Dec 2010. ISSN 1879-3150. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20035779 >.

FRANCISCHETTI, I. M.; ANDERSEN, J. F.; RIBEIRO, J. M. 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,** v. 41, n. 11, p. 3810-8, Mar 2002. ISSN 0006-2960. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11888300 >.

FRANCISCHETTI, I. M. et al. Insight into the Sialome of the Bed Bug, Cimex lectularius. **J Proteome Res,** v. 9, n. 8, p. 3820-31, Aug 2010. ISSN 1535-3907. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20441151 >.

FRANCISCHETTI, I. M. et al. An insight into the sialome of the soft tick, Ornithodorus parkeri. **Insect Biochem Mol Biol,** v. 38, n. 1, p. 1-21, Jan 2008. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18070662 >.

FRANCISCHETTI, I. M. et al. 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,** v. 275, n. 17, p. 12639-50, Apr 2000. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10777556 >.

FRANCISCHETTI, I. M. et al. The role of saliva in tick feeding. **Front Biosci,** v. 14, p. 2051-88, 2009. ISSN 1093-4715. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19273185 >.

FRANCISCHETTI, I. M. et al. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector Anopheles gambiae. **J Exp Biol,** v. 205, n. Pt 16, p. 2429-51, Aug 2002. ISSN 0022-0949. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12124367 >.

FRIEDRICH, T. et al. A Kazal-type inhibitor with thrombin specificity from Rhodnius prolixus. **J Biol Chem,** v. 268, n. 22, p. 16216-22, Aug 1993. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8344906 >.

FURIE, B.; FURIE, B. C. Thrombus formation in vivo. **J Clin Invest,** v. 115, n. 12, p. 3355-62, Dec 2005. ISSN 0021-9738. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16322780 >.

GASCON, J.; BERN, C.; PINAZO, M. J. Chagas disease in Spain, the United States and other non-endemic countries. **Acta Trop,** v. 115, n. 1-2, p. 22-7, 2010 Jul-Aug 2010. ISSN 1873-6254. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19646412 >.

GILLESPIE, R. D.; MBOW, M. L.; TITUS, R. G. The immunomodulatory factors of bloodfeeding arthropod saliva. **Parasite Immunol,** v. 22, n. 7, p. 319-31, Jul 2000. ISSN 0141-9838. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10886716 >.

GOURBIÈRE, S. et al. Genetics and evolution of triatomines: from phylogeny to vector control. **Heredity (Edinb),** v. 108, n. 3, p. 190-202, Mar 2012. ISSN 1365-2540. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21897436 >.

GURGEL-GONÇALVES, R. et al. Is Rhodnius prolixus (Triatominae) invading houses in central Brazil? **Acta Trop,** v. 107, n. 2, p. 90-8, Aug 2008. ISSN 0001-706X. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18550022 >.

GURGEL-GONÇALVES, R. et al. Infestation of Mauritia flexuosa palms by triatomines (Hemiptera: Reduviidae), vectors of Trypanosoma cruzi and Trypanosoma rangeli in the Brazilian savanna. **Acta Trop,** v. 121, n. 2, p. 105-11, Feb 2012. ISSN 1873-6254. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22037200 >.

GURGEL-GONÇALVES, R. et al. Geographic distribution of chagas disease vectors in Brazil based on ecological niche modeling. **J Trop Med,** v. 2012, p. 705326, 2012. ISSN 1687-9694. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22523500 >.

GURGEL-GONÇALVES, R. et al. Sampling Rhodnius neglectus in Mauritia flexuosa palm trees: a field study in the Brazilian savanna. **Med Vet Entomol,** v. 17, n. 3, p. 347-50, Sep 2003. ISSN 0269-283X. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12941022 >.

GURGEL-GONÇALVES, R. et al. Enzootic transmission of Trypanosoma cruzi and T. rangeli in the Federal District of Brazil. **Rev Inst Med Trop Sao Paulo,** v. 46, n. 6, p.

323-30, 2004 Nov-Dec 2004. ISSN 0036-4665. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15654478 >.

HAYASHI, H. et al. Identification of the active region responsible for the anti-thrombotic activity of anopheline anti-platelet protein from a malaria vector mosquito. **Platelets,** v. 24, n. 4, p. 324-32, 2013. ISSN 1369-1635. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22738392 >.

HOTEZ, P. et al. The neglected tropical diseases: the ancient afflictions of stigma and poverty and the prospects for their control and elimination. **Adv Exp Med Biol,** v. 582, p. 23-33, 2006. ISSN 0065-2598. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16802616 >.

JENNINGS, L. K. Role of platelets in atherothrombosis. **Am J Cardiol,** v. 103, n. 3 Suppl, p. 4A-10A, Feb 2009. ISSN 1879-1913. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19166707 >.

KARCZEWSKI, J.; ENDRIS, R.; CONNOLLY, T. M. Disagregin is a fibrinogen receptor antagonist lacking the Arg-Gly-Asp sequence from the tick, Ornithodoros moubata. **J Biol Chem,** v. 269, n. 9, p. 6702-8, Mar 1994. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8120028 >.

KARCZEWSKI, J. et al. An inhibitor from the argasid tick Ornithodoros moubata of cell adhesion to collagen. **Biochem Biophys Res Commun,** v. 208, n. 2, p. 532-41, Mar 1995. ISSN 0006-291X. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7695604 >.

KELLER, P. M. et al. Cloning of the cDNA and expression of moubatin, an inhibitor of platelet aggregation. **J Biol Chem,** v. 268, n. 8, p. 5450-6, Mar 1993. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8449907 >.

KINI, R.M.; CLEMETSON, K.J.; MARKLAND, F.S.; MCLANE, M.A.; MORITA, T. Toxins and Hemostasis. From Bench to Bedside. **Springer,** 2010.

KOH, C. Y. et al. Variegin, a novel fast and tight binding thrombin inhibitor from the tropical bont tick. **J Biol Chem,** v. 282, n. 40, p. 29101-13, Oct 2007. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17684009 >.

KRISHNASWAMY, S. Exosite-driven substrate specificity and function in coagulation. **J Thromb Haemost,** v. 3, n. 1, p. 54-67, Jan 2005. ISSN 1538-7933. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15634266 >.

LAZZARI, C. R.; PEREIRA, M. H.; LORENZO, M. G. Behavioural biology of Chagas disease vectors. **Mem Inst Oswaldo Cruz,** v. 108 Suppl 1, p. 34-47, 2013. ISSN 1678-8060. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/24473801 >.

LEBOULLE, G. et al. Characterization of a novel salivary immunosuppressive protein from Ixodes ricinus ticks. **J Biol Chem,** v. 277, n. 12, p. 10083-9, Mar 2002. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11792703 >.

LEHANE. The Biology of Blood-Sucking in Insects. Second Edition. **Cambridge University Press**, 2005.

LERNER, E. A. et al. Isolation of maxadilan, a potent vasodilatory peptide from the salivary glands of the sand fly Lutzomyia longipalpis. **J Biol Chem,** v. 266, n. 17, p. 11234-6, Jun 1991. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/2040631 >.

LINDOSO, J. A.; LINDOSO, A. A. Neglected tropical diseases in Brazil. **Rev Inst Med Trop Sao Paulo,** v. 51, n. 5, p. 247-53, 2009 Sep-Oct 2009. ISSN 1678-9946. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19893976 >.

LIYOU, N. et al. Cloning and expression of ecto 5-nucleotidase from the cattle tick Boophilus microplus. **Insect Mol Biol,** v. 8, n. 2, p. 257-66, May 1999. ISSN 0962-1075. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10380109 >.

LOMBARDO, F. et al. Promoter sequences of the putative Anopheles gambiae apyrase confer salivary gland expression in Drosophila melanogaster. **J Biol Chem,** v. 275, n. 31, p. 23861-8, Aug 2000. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10801886 >.

MA, D. et al. Anti-thrombosis repertoire of blood-feeding horsefly salivary glands. **Mol Cell Proteomics,** v. 8, n. 9, p. 2071-9, Sep 2009. ISSN 1535-9484. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19531497 >.

MACEDO-RIBEIRO, S. et al. Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. **PLoS One,** v. 3, n. 2, p. e1624, 2008. ISSN 1932-6203. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18286181 >.

MANN, K. G.; BUTENAS, S.; BRUMMEL, K. The dynamics of thrombin formation. **Arterioscler Thromb Vasc Biol,** v. 23, n. 1, p. 17-25, Jan 2003. ISSN 1524-4636. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12524220 >.

MANS, B. J. Tick histamine-binding proteins and related lipocalins: potential as therapeutic agents. **Curr Opin Investig Drugs,** v. 6, n. 11, p. 1131-5, Nov 2005. ISSN 1472-4472. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16312134 >.

MANS, B. J.; LOUW, A. I.; NEITZ, A. W. Savignygrin, a platelet aggregation inhibitor from the soft tick Ornithodoros savignyi, presents the RGD integrin recognition motif on the Kunitz-BPTI fold. **J Biol Chem,** v. 277, n. 24, p. 21371-8, Jun 2002. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11932256 >.

MANS, B. J.; RIBEIRO, J. M. Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins. **Insect Biochem Mol Biol,** v. 38, n. 9, p. 841-52, Sep 2008. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18694828 >.

MESQUITA, R. D. et al. Trypanosoma cruzi infection is enhanced by vector saliva through immunosuppressant mechanisms mediated by lysophosphatidylcholine. **Infect Immun,** v. 76, n. 12, p. 5543-52, Dec 2008. ISSN 1098-5522. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18794282 >.

METZKER, M. L. Sequencing technologies - the next generation. **Nat Rev Genet,** v. 11, n. 1, p. 31-46, Jan 2010. ISSN 1471-0064. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19997069 >.

MIZURINI, D. M.; FRANCISCHETTI, I. M.; MONTEIRO, R. Q. Aegyptin inhibits collagen-induced coagulation activation in vitro and thromboembolism in vivo. **Biochem Biophys Res Commun,** v. 436, n. 2, p. 235-9, Jun 2013. ISSN 1090-2104. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/23726920 >.

MONROE, D. M.; HOFFMAN, M.; ROBERTS, H. R. Platelets and thrombin generation. **Arterioscler Thromb Vasc Biol,** v. 22, n. 9, p. 1381-9, Sep 2002. ISSN 1524-4636. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12231555 >.

MONROY, M. C. et al. Habitats, dispersion and invasion of sylvatic Triatoma dimidiata (Hemiptera: Reduviidae: Triatominae) in Petén, Guatemala. **J Med Entomol,** v. 40, n. 6, p. 800-6, Nov 2003. ISSN 0022-2585. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/14765656 >.

MORITA, A. et al. Identification and characterization of a collagen-induced platelet aggregation inhibitor, triplatin, from salivary glands of the assassin bug, Triatoma infestans. **FEBS J,** v. 273, n. 13, p. 2955-62, Jul 2006. ISSN 1742-464X. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16759235 >.

NAGALAKSHMI, U.; WAERN, K.; SNYDER, M. RNA-Seq: a method for comprehensive transcriptome analysis. **Curr Protoc Mol Biol,** v. Chapter 4, p. Unit

4.11.1-13, Jan 2010. ISSN 1934-3647. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/20069539 >.

NAKAGAWA, J.; CÓRDON-ROSALES C.; JUÁREZ J.; ITZEP C.; NONAMI T. Impact of residual spraying on Rhodnius prolixus and Triatoma dimidiata in the department of Zacapa in Guatemala. **Mem Inst Oswaldo Cruz,** v. 98, n. 2, p. 277-81, Mar 2003. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12764447 >.

NAKAGAWA, J.; HASHIMOTO K.; CÓRDON-ROSALES C.; ABRAHAM J.; TRAMPE R.; MARROQUÍN MARROQUÍN L. The impact of vector control on Triatoma dimidiata in the Guatemalan department of Jutiapa. **Ann Trop Med Parasitol,** v. 97, n. 3, p. 288-97, Apr 2003. ISSN 0003-4983. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12803860 >.

NOESKE-JUNGBLUT, C. et al. Triabin, a highly potent exosite inhibitor of thrombin. **J Biol Chem,** v. 270, n. 48, p. 28629-34, Dec 1995. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7499380 >.

NOESKE-JUNGBLUT, C. et al. An inhibitor of collagen-induced platelet aggregation from the saliva of Triatoma pallidipennis. **J Biol Chem,** v. 269, n. 7, p. 5050-3, Feb 1994. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8106481 >.

NOIREAU, F.; DIOSQUE, P.; JANSEN, A. M. Trypanosoma cruzi: adaptation to its vectors and its hosts. **Vet Res,** v. 40, n. 2, p. 26, 2009 Mar-Apr 2009. ISSN 0928-4249. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19250627 >.

OLIVEIRA, A. W. S.; da SILVA I.G. Distribuição geográfica e indicadores entomológicos de triatomíneos sinantrópicos capturados no Estado de Goiás. **Revista da Sociedade Brasileira de Medicina Tropical**. 40**:** 204-208 p. 2007.

OPAS. O controle da doença de Chagas nos países do Cone Sul da América. **Organização Panamericana de Saúde**, 2002. Disponível em: < http://www.paho.org/hq/index.php?gid=20193&option=com_docman&task=doc_view>

OPAS. Guia para vigilância, prevenção, controle e manejo clínico da doença de Chagas agudatransmitida por alimentos. **Organização Panamericana de Saúde**. Rio de Janeiro, 2009.

PAESEN, G. C. et al. Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. **Mol Cell,** v. 3, n. 5, p. 661-71, May 1999. ISSN 1097-2765. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10360182 >.

PAESEN, G. C. et al. Tick histamine-binding proteins: lipocalins with a second binding cavity. **Biochim Biophys Acta,** v. 1482, n. 1-2, p. 92-101, Oct 2000. ISSN 0006-3002. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11058751 >.

PALMA-GUZMÁN, R.; RIVERA, B.; MORALES, G. Domestic vectors of Chagas' disease in three rural communities of Nicaragua. **Rev Inst Med Trop Sao Paulo,** v. 38, n. 2, p. 133-40, 1996 Mar-Apr 1996. ISSN 0036-4665. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/9071033 >.

PRATA, A. Clinical and epidemiological aspects of Chagas disease. **Lancet Infect Dis,** v. 1, n. 2, p. 92-100, Sep 2001. ISSN 1473-3099. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11871482 >.

QUINTAL, R. E.; POLANCO, G. G. Feeding preferences of Triatoma dimidiata maculipennis in Yucatan, Mexico. **Am J Trop Med Hyg,** v. 26, n. 1, p. 176-8, Jan 1977. ISSN 0002-9637. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/320892 >.

RIBEIRO, J. M. Ixodes dammini: salivary anti-complement activity. **Exp Parasitol,** v. 64, n. 3, p. 347-53, Dec 1987a. ISSN 0014-4894. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/3119364 >.

RIBEIRO, J. M. Vector salivation and parasite transmission. **Mem Inst Oswaldo Cruz,** v. 82 Suppl 3, p. 1-3, 1987b. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/3334398 >.

RIBEIRO, J. M. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? **Infect Agents Dis,** v. 4, n. 3, p. 143-52, Sep 1995. ISSN 1056-2044. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8548192 >.

RIBEIRO, J. M. et al. Exploring the sialome of the blood-sucking bug Rhodnius prolixus. **Insect Biochem Mol Biol,** v. 34, n. 1, p. 61-79, Jan 2004. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/14976983 >.

RIBEIRO, J. M. et al. An insight into the salivary transcriptome and proteome of the adult female mosquito Culex pipiens quinquefasciatus. **Insect Biochem Mol Biol,** v. 34, n. 6, p. 543-63, Jun 2004. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15147756 >.

RIBEIRO, J. M. et al. Amblyomma americanum: characterization of salivary prostaglandins E2 and F2 alpha by RP-HPLC/bioassay and gas chromatography-mass spectrometry. **Exp Parasitol,** v. 74, n. 1, p. 112-6, Feb 1992. ISSN 0014-4894. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/1730268 >.

RIBEIRO, J. M.; FRANCISCHETTI, I. M. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. **Annu Rev Entomol,** v. 48, p. 73-88, 2003. ISSN 0066-4170. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12194906 >.

RIBEIRO, J. M. et al. Salivary apyrase of Aedes aegypti: characterization and secretory fate. **Comp Biochem Physiol B,** v. 79, n. 1, p. 81-6, 1984. ISSN 0305-0491. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/6094095 >.

RIBEIRO, J. M.; SCHWARZ, A.; FRANCISCHETTI, I. M. A Deep Insight Into the Sialotranscriptome of the Chagas Disease Vector, Panstrongylus megistus (Hemiptera: Heteroptera). **J Med Entomol,** v. 52, n. 3, p. 351-8, May 2015. ISSN 0022-2585. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/26334808 >.

RIBEIRO, J. M.; WALKER, F. A. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorin) of Rhodnius prolixus. **J Exp Med,** v. 180, n. 6, p. 2251-7, Dec 1994. ISSN 0022-1007. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7964498 >.

RIBEIRO, J. M. et al. Exploring the sialome of the blood-sucking bug Rhodnius prolixus. **Insect Biochem Mol Biol**, v. 34, n. 1, p. 61-79, Jan 2004. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/14976983 >.

RIBEIRO, J. M. An insight into the sialotranscriptome of Triatoma rubida (Hemiptera: Heteroptera). **J Med Entomol**, v. 49, n. 3, p. 563-72, May 2012. ISSN 0022-2585. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22679863 >.

RIVERA, J. et al. Platelet receptors and signaling in the dynamics of thrombus formation. **Haematologica,** v. 94, n. 5, p. 700-11, May 2009. ISSN 1592-8721. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19286885 >.

RUBENS A. S.; BONIFÁCIO P. R.; WANDERLEY D. M. V. Doença de Chagas no Estado de São Paulo: comparaçãoentre pesquisa ativa de triatomíneos em domicílios enotificação de sua presença pela população emárea sob vigilância entomológica. **Revista da Sociedade Brasileira de Medicina Tropical.** 32(6):653-659 1999.

SALAS CLAVIJO, N. A. et al. Prevalence of Chagas disease in pregnant women and incidence of congenital transmission in Santa Cruz de la Sierra, Bolivia. **Acta Trop,** v. 124, n. 1, p. 87-91, Oct 2012. ISSN 1873-6254. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/22772023 >.

SANDOVAL-RUIZ, C. A.; GUEVARA, R.; IBÁÑEZ-BERNAL, S. Household risk factors associated to infestation of Triatoma dimidiata, the Chagas disease vector in Central Region of Veracruz, Mexico. **Salud Publica Mex,** v. 56, n. 2, p. 213-20, Apr 2014. ISSN 1606-7916. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25014428 >.

SANTOS, A. et al. The sialotranscriptome of the blood-sucking bug Triatoma brasiliensis (Hemiptera, Triatominae). **Insect Biochem Mol Biol,** v. 37, n. 7, p. 702-12, Jul 2007. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17550826 >.

SARKIS, J. J.; GUIMARÃES, J. A.; RIBEIRO, J. M. Salivary apyrase of Rhodnius prolixus. Kinetics and purification. **Biochem J,** v. 233, n. 3, p. 885-91, Feb 1986. ISSN 0264-6021. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/3010945 >.

SCHMUNIS, G. A.; YADON, Z. E. Chagas disease: a Latin American health problem becoming a world health problem. **Acta Trop,** v. 115, n. 1-2, p. 14-21, 2010 Jul-Aug 2010. ISSN 1873-6254. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19932071 >.

SCHOELER, G. B.; WIKEL, S. K. Modulation of host immunity by haematophagous arthropods. **Ann Trop Med Parasitol,** v. 95, n. 8, p. 755-71, Dec 2001. ISSN 0003-4983. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11784430 >.

SCHOFIELD, C. J.; GALVÃO, C. Classification, evolution, and species groups within the Triatominae. **Acta Trop,** v. 110, n. 2-3, p. 88-100, 2009 May-Jun 2009. ISSN 1873-6254. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19385053 >.

SILVEIRA A. C. Situação do controle da transmissão vetorial da doença de Chagas nas Américas. Rio de Janeiro: **Cad. Saúde Pública**. 16(Sup. 2):35-42 2000.

SILVEIRA, A. C. [New challenges and the future of control]. **Rev Soc Bras Med Trop,** v. 44 Suppl 2, p. 122-4, 2011. ISSN 1678-9849. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21584365 >.

STUTZER, C. et al. Ornithodoros savignyi: soft tick apyrase belongs to the 5'-nucleotidase family. **Exp Parasitol,** v. 122, n. 4, p. 318-27, Aug 2009. ISSN 1090-2449. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19393241 >.

SÁ-NUNES, A. et al. The immunomodulatory action of sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. **J Immunol,** v. 182, n. 12, p. 7422-9, Jun 2009. ISSN 1550-6606. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19494265 >.

SÁ-NUNES, A. et al. Prostaglandin E2 is a major inhibitor of dendritic cell maturation and function in Ixodes scapularis saliva. **J Immunol,** v. 179, n. 3, p. 1497-505, Aug 2007. ISSN 0022-1767. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17641015 >.

TITUS, R. G.; RIBEIRO, J. M. Salivary gland lysates from the sand fly Lutzomyia longipalpis enhance Leishmania infectivity. **Science,** v. 239, n. 4845, p. 1306-8, Mar 1988. ISSN 0036-8075. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/3344436 >.

TYSON, K. et al. Biochemical and functional characterization of Salp20, an Ixodes scapularis tick salivary protein that inhibits the complement pathway. **Insect Mol Biol,** v. 16, n. 4, p. 469-79, Aug 2007. ISSN 0962-1075. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17651236 >.

TYSON, K. R.; ELKINS, C.; DE SILVA, A. M. A novel mechanism of complement inhibition unmasked by a tick salivary protein that binds to properdin. **J Immunol,** v. 180, n. 6, p. 3964-8, Mar 2008. ISSN 0022-1767. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18322205 >.

VALENZUELA, J. G. et al. The salivary apyrase of the blood-sucking sand fly Phlebotomus papatasi belongs to the novel Cimex family of apyrases. **J Exp Biol,** v. 204, n. Pt 2, p. 229-37, Jan 2001. ISSN 0022-0949. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11136609 >.

VALENZUELA, J. G. et al. Purification, cloning, and expression of an apyrase from the bed bug Cimex lectularius. A new type of nucleotide-binding enzyme. **J Biol Chem,** v. 273, n. 46, p. 30583-90, Nov 1998. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/9804829 >.

VALENZUELA, J. G. et al. Purification, cloning, and expression of a novel salivary anticomplement protein from the tick, Ixodes scapularis. **J Biol Chem,** v. 275, n. 25, p. 18717-23, Jun 2000. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/10749868 >.

VALENZUELA, J. G. et al Exploring the sialome of the tick Ixodes scapularis. **J Exp Biol,** v. 205, n. Pt 18, p. 2843-64, Sep 2002. ISSN 0022-0949. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12177149 >.

VALENZUELA, J. G. et al. Exploring the salivary gland transcriptome and proteome of the Anopheles stephensi mosquito. **Insect Biochem Mol Biol,** v. 33, n. 7, p. 717-32, Jul 2003. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12826099 >.

VALENZUELA, J. G. et al. Toward a description of the sialome of the adult female mosquito Aedes aegypti. **Insect Biochem Mol Biol,** v. 32, n. 9, p. 1101-22, Sep 2002. ISSN 0965-1748. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/12213246 >.

VALENZUELA, J. G.; WALKER, F. A.; RIBEIRO, J. M. A salivary nitrophorin (nitric-oxide-carrying hemoprotein) in the bedbug Cimex lectularius. **J Exp Biol,** v. 198, n. Pt 7, p. 1519-26, Jul 1995. ISSN 0022-0949. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7658188 >.

VAN DE LOCHT, A. et al. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. **EMBO J,** v. 14, n. 21, p. 5149-57, Nov 1995. ISSN 0261-4189. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7489704 >.

VAN DE LOCHT, A. et al. The ornithodorin-thrombin crystal structure, a key to the TAP enigma? **EMBO J,** v. 15, n. 22, p. 6011-7, Nov 1996. ISSN 0261-4189. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8947023 >.

VARGA-SZABO, D.; PLEINES, I.; NIESWANDT, B. Cell adhesion mechanisms in platelets. **Arterioscler Thromb Vasc Biol,** v. 28, n. 3, p. 403-12, Mar 2008. ISSN 1524-4636. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18174460 >.

WANG, H.; NUTTALL, P. A. Excretion of host immunoglobulin in tick saliva and detection of IgG-binding proteins in tick haemolymph and salivary glands. **Parasitology,** v. 109 (Pt 4), p. 525-30, Nov 1994. ISSN 0031-1820. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7794319 >.

WANG, H.; NUTTALL, P. A. Immunoglobulin-G binding proteins in the ixodid ticks, Rhipicephalus appendiculatus, Amblyomma variegatum and Ixodes hexagonus. **Parasitology,** v. 111 (Pt 2), p. 161-5, Aug 1995. ISSN 0031-1820. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/7675530 >.

WANG, X. et al. Variabilin, a novel RGD-containing antagonist of glycoprotein IIb-IIIa and platelet aggregation inhibitor from the hard tick Dermacentor variabilis. **J Biol**

Chem, v. 271, n. 30, p. 17785-90, Jul 1996. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8663513 >.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nat Rev Genet,** v. 10, n. 1, p. 57-63, Jan 2009. ISSN 1471-0064. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/19015660 >.

WATSON, S. P. et al. GPVI and integrin alphaIIb beta3 signaling in platelets. **J Thromb Haemost,** v. 3, n. 8, p. 1752-62, Aug 2005. ISSN 1538-7933. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16102042 >.

WAXMAN, L.; CONNOLLY, T. M. Isolation of an inhibitor selective for collagen-stimulated platelet aggregation from the soft tick Ornithodoros moubata. **J Biol Chem,** v. 268, n. 8, p. 5445-9, Mar 1993. ISSN 0021-9258. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8449906 >.

WEEKS, E. N. et al. Community-based sampling methods for surveillance of the Chagas disease vector, Triatoma dimidiata (Hemiptera: Reduviidae: Triatominae). **J Med Entomol,** v. 51, n. 5, p. 1035-42, Sep 2014. ISSN 0022-2585. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25276934 >.

WHEAT, W. H. et al. Lutzomyia longipalpis salivary peptide maxadilan alters murine dendritic cell expression of CD80/86, CCR7, and cytokine secretion and reprograms dendritic cell-mediated cytokine release from cultures containing allogeneic T cells. **J Immunol,** v. 180, n. 12, p. 8286-98, Jun 2008. ISSN 0022-1767. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18523295 >.

WHO. Hidden successes, Emerging opportunities. **World Health Organization**, 2006. Disponível em: < http://apps.who.int/iris/handle/10665/44214 >

WHO. First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases. **World Health Organization**, 2010. Disponível em: < http://www.who.int/neglected_diseases/2010report/en/ >

WHO. Second WHO report on neglected tropical diseases: sustanining the drive to overcome the global impact of neglected tropical diseases. **World Health Organization**, 2013. Disponível em: < http://www.who.int/neglected_diseases/9789241564540/en/ >

XU, X. et al. Toward an understanding of the molecular mechanism for successful blood feeding by coupling proteomics analysis with pharmacological testing of horsefly salivary glands. **Mol Cell Proteomics,** v. 7, n. 3, p. 582-90, Mar 2008. ISSN 1535-9484. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18087067 >.

YOSHIDA, S. et al. Inhibition of collagen-induced platelet aggregation by anopheline antiplatelet protein, a saliva protein from a malaria vector mosquito. **Blood,** v. 111, n. 4, p. 2007-14, Feb 2008. ISSN 0006-4971. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18056842 >.

ZELEDÓN, R.; UGALDE, J. A.; PANIAGUA, L. A. Entomological and ecological aspects of six sylvatic species of triatomines (Hemiptera, Reduviidae) from the collection of the National Biodiversity Institute of Costa Rica, Central America. **Mem Inst Oswaldo Cruz,** v. 96, n. 6, p. 757-64, Aug 2001. ISSN 0074-0276. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/11562697 >.