Published in final edited form as:

J Med Entomol. 2012 May ; 49(3): 563–572.

An Insight Into the Sialotranscriptome of *Triatoma rubida* (Hemiptera: Heteroptera)

JOSÉ M. C. RIBEIRO^{1,2}, TERESA C. F. ASSUMPÇÃO¹, VAN M. PHAM¹, IVO M. B. FRANCISCHETTI¹, and CAROLINA E. REISENMAN³

¹Section of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, 12735 Twinbrook Parkway room 2E32D, Rockville, MD 20852

³Department of Neurosciences, College of Sciences, University of Arizona, PO Box 210077, Tucson, AZ 85721-0077

Abstract

The kissing bug *Triatoma rubida* (Uhler, 1894) is found in southwestern United States and parts of Mexico where it is found infected with *Trypanosoma cruzi*, invades human dwellings and causes allergies from their bites. Although the protein salivary composition of several triatomine species is known, not a single salivary protein sequence is known from *T. rubida*. Furthermore, the salivary diversity of related hematophagous arthropods is very large probably because of the immune pressure from their hosts. Here we report the sialotranscriptome analysis of *T. rubida* based on the assembly of 1,820 high-quality expressed sequence tags, 51% of which code for putative secreted peptides, including lipocalins, members of the antigen five family, apyrase, hemolysin, and trypsin families. Interestingly, *T. rubida* lipocalins are at best 40% identical in primary sequence to those of *T. protracta*, a kissing bug that overlaps its range with *T. rubida*, indicating the diversity of the salivary lipocalins among species of the same hematophagous genus. We additionally found several expressed sequence tags coding for proteins of clear *Trypanosoma* spp. origin. This work contributes to the future development of markers of human and pet exposure to *T. rubida* and to the possible development of desensitization therapies. Supp. Data 1 and 2(online only) of the transcriptome and deduced protein sequences can be obtained from <http://exon.niaid.nih.gov/transcriptome/Trubida/Triru-S1-web.xlsx> and <http://exon.niaid.nih.gov/transcriptome/Trubida/Triru-S2-web.xlsx>.

Keywords

Chagas disease; vector biology; salivary gland; sialome; transcriptome

Triatomine bugs, commonly known as kissing bugs, are prevalent in the Americas where they are the sole vectors of Chagas disease, a disabling and often fatal disease, with 7.7 million people infected and an annual incidence (through vectorial transmission) of 42,500 (Pan–American Health Organization 2006). In southwest United States, kissing bugs are the cause of allergic reactions in humans and animals (Wolf 1969, Pinnas et al. 1986, Moncayo and Ortiz Yanine 2006, Milei et al. 2009, Tanowitz et al. 2009). In the southwestern and western United States infected bugs are found mostly in sylvatic habitats (Wood 1942, 1943; Wood and Wood 1961; Kjos et al. 2008, 2009; Reisenman et al. 2010; Bern et al. 2011), where they are involved in the sylvatic transmission cycle of *Trypanosoma cruzi*, the causative agent of Chagas disease, to mammals such as packrats, mice, armadillos, raccoons,

²Corresponding author: jribeiro@niaid.nih.gov.

opossums, and dogs (Packchanian 1942, Ryckman 1986, Kjos et al. 2008, Beard et al. 2003, Brown et al. 2010). Only seven autochthonous cases of human Chagas Disease have been reported in the United States, all in the southern half of the country (Woody and Woody 1955, Schiffler et al. 1984, Ochs et al. 1996, Herwaldt et al. 2000, Dorn et al. 2007, Bern et al. 2011). The suburbanization of America has led humans to increased contact with nature, including kissing bugs. In Arizona and California this contact is causing triatomines to infest human dwellings and in causing mild to severe allergic reactions, including anaphylaxis, to triatomine bites (Wolf 1969, Marshall et al. 1986, Pinnas et al. 1986, Moffitt et al. 2003, Klotz et al. 2010). Although it has long been assumed that vectorial transmission of Chagas disease is rare in the United States because the native species of triatomines have long defecation delays, a recent study in *Triatoma rubida* (Uhler, 1894) (the most prevalent triatomine species in southern Arizona) showed that defecation during feeding was highly frequent in females (Reisenman et al. 2011). Transmission of *T. cruzi* in wild cycles additionally involves the congenital route and the ingestion of infected bugs by mammalian reservoirs and dogs (Barr 2009, Roellig et al. 2009, Yoshida 2009, Bern et al. 2011).

Two main triatomines, *T. rubida* and *T. protracta* (Uhler, 1894) have been reported as causing human allergies in the southwestern and western United States (Marshall 1982, Marshall and Street 1982, Marshall et al. 1986, Pinnas et al. 1986). *T. rubida* belongs to the *rubrofasciata* species complex, while *T. protracta* belongs to the *protracta* species complex (Ibarra-Cerdena et al. 2009). Both species are found in California, Arizona, New Mexico, Texas, and some northern Mexico states (Pfeiler et al. 2006, Sarkar et al. 2010). Arizona is noteworthy as the state with the highest number of triatomine-human contacts reported in the United States (sources: American Association of Poison Control Centers; Arizona Poison and Drug Information Center, University of Arizona Health Sciences Center) and most contacts with *T. rubida*. While some antigens have been characterized and cloned from *T. protracta* (Paddock et al. 2001), none have been identified from *T. rubida*. Moreover, serum of people allergic to *T. rubida* does not recognize a recombinant antigen, named procalin, which is recognized by the serum of people allergic to *T. protracta* bites (Marshall et al. 1986, Pinnas et al. 1986), emphasizing the antigenic diversity in the saliva of related triatomine insect species.

Saliva of blood feeding insects have a repertoire of antiplatelet, vasodilatory, anticlotting, anti-inflammatory, and immunomodulatory compounds (Ribeiro 1995). Recent studies on salivary gland transcriptomes of hematophagous insects are revealing that saliva of these insects may contain hundreds of proteins, most of which are members of multi gene families. Many of these families are novel, in the sense that they are only found so far in sialomes of particular insect species or genus, and have unknown function. This molecular diversity is proposed to derive from the fast evolution (including gene duplication and divergence of function) of the salivary genes as a result of the immune pressure from their hosts, which can either inhibit protein function, or induce defensive behavior from their hosts that might interrupt feeding or lead to death of the insect (Mans and Neitz 2004, Ribeiro and Arca 2009). In this article, we describe the sialome (from the Greek sialo = saliva) of *T. rubida* aiming to help identification of salivary antigens that might be useful in the diagnosis and treatment of *T. rubida* allergy, and that might be used as epidemiological markers of *T. rubida* exposure in humans and their pets.

Materials and Methods

The methodology used was essentially identical to our previous transcriptomes, such as in *Dipetalogaster maxima* (Assumpção et al. 2011) and is reproduced here with small modifications.

***T. rubida* and Salivary Glands cDNA Library Construction**

Adult *T. rubida* were collected from the field and fed with bovine blood supplemented with ATP using an artificial feeder as described previously (Reisenman et al. 2011). Insects were starved for 2 wk before salivary glands (SG) were dissected and extracted. SG mRNA was isolated from a total of 22 main and 13 accessory glands obtained from 12 insects using the Micro-FastTrack mRNA isolation kit (Invitrogen, San Diego, CA). The polymerase chain reaction (PCR)-based cDNA library was made following the instructions for the SMART (switching mechanism at 5' end of RNA transcript) cDNA library construction kit (Clontech, Palo Alto, CA). This kit provides a method for producing high-quality, full-length cDNA libraries from nanogram quantities of polyA+ or total RNA. It uses a specially designed oligonucleotide named SMART IV in the first-strand synthesis to generate high yields of full-length, double-stranded cDNA. *T. rubida* SG polyA+ RNA was used for reverse transcription to cDNA using Moloney murine leukemia virus reverse transcriptase (Clontech), SMART IV oligonucleotide, and CDS III/3' primer (Clontech). Trehalose was added to the reaction that was carried out at 60°C for 1 h, then at 42°C for 40 min. Second-strand synthesis was performed according to a long-distance PCR-based protocol using the 5' PCR primer and CDS III/3' primer as sense and antisense primers, respectively. These two primers also create *Sfi*I and *B* restriction enzyme sites at the end of the cDNA. Advantage Taq polymerase mix (Clontech) was used to carry out a long-distance PCR reaction on a GeneAmp PCR system 9700 (Perkin Elmer Corp., Foster City, CA). The PCR conditions were: 95°C for 1 min; 18 cycles of 95°C for 15 s, 68°C for 6 min. An aliquot of the cDNA was analyzed on a 1.1% agarose/EtBr (0.1 µg/ml) gel to check for the quality and range of the synthesized cDNA. Double-stranded cDNA was immediately treated with proteinase K (0.8 µg/µl) at 45°C for 20 min and washed three times with water using Amicon filters with a 100-kDa cutoff (Millipore, Bedford, MA). The purified double-stranded cDNA was then digested with *Sfi*I restriction enzyme at 50°C for 2 h followed by size fractionation on a Chroma-Spin-400 drip column (Clontech). The profiles of the fractions were checked on a 1.1% agarose/EtBr (0.1 µg/ml), and fractions containing cDNA were pooled in three different groups according to size: large, medium, or small sequences. Each group was concentrated and washed three times with water using an Amicon filter with a 100-kDa cutoff. Concentrated cDNA was then ligated into a λ TriplEx2 vector (Clontech), and the resulting ligation mixture was packaged using GigaPack Gold III Plus packaging extract (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. The packaged library was plated by infecting log-phase XL1-Blue *Escherichia coli* cells (Clontech). The percentage of recombinant clones was determined by blue-white selection screening on LB/MgSO₄ plates containing X-gal/IPTG.

Sequencing of the cDNA Library

The SG cDNA library was plated on LB/MgSO₄ plates containing X-gal/IPTG to an average of 250 plaques per 150-mm petri plate. Recombinant (white) plaques were randomly picked up and transferred to 96-well MicroTest U-bottom plates (BD BioScience, San José, CA) containing 75 µl of H₂O per well. The phage suspension was either immediately used for PCR or stored at 4°C until use.

To amplify the cDNA using a PCR reaction, 4 µl of the phage sample were used as a template. The primers were sequences from the λ TriplEx2 vector and named PT2F1 (5'-AAG TACTCTAGCAATTGTGAGC-3') and PT2R1 (5'-CTCTTCGCTATTACGCCAGCTG-3'), positioned at the 5' and 3' end of the cDNA insert, respectively. The reaction was carried out in MicroAmp 96-well PCR plates (Applied Biosystems, Inc., Fullerton, CA) using FastStart PCR Master (Roche Molecular Biochemicals, Indianapolis, IN) on a GeneAmp PCR system 9700 (Perkin Elmer Corp.). The PCR conditions were: one hold of 75°C for 3 min, one hold of 94°C for 4 min, 30

cycles of 94°C for 1 min, 49°C for 1 min, and 72°C for 1 min 20 s. The amplified products were analyzed on a 1.2% agarose/EtBr gel. cDNA library clones (1,880 clones) were PCR amplified, and those showing a single band were selected for sequencing. The PCR products were used as a template for a cycle-sequencing reaction using a DTCS labeling kit (Beckman Coulter, Fullerton, CA). The primer used for sequencing (PT2F3) is upstream of the inserted cDNA and downstream of the PT2F1 primer. The sequencing reaction was performed on a Perkin Elmer 9700 thermocycler. Conditions were one hold of 75°C for 2 min, one hold of 94°C for 4 min, and 30 cycles of 96°C for 20 s, 50°C for 20 s, and 60°C for 4 min. After cycle-sequencing the samples, a cleaning step was performed using the multiscreen 96-well plate cleaning system (Millipore). The 96-well multiscreening plate was prepared by adding a fixed amount (manufacturer's specification) of Sephadex-50 (Amersham Pharmacia Biotech, Piscataway, NJ) and 300 μ l of deionized water. After partially drying the Sephadex in the multiscreen plate, the entire cycle-sequencing reaction was added to the center of each well, centrifuged at $580 \times g$ for 5 min, and the clean sample was collected on a sequencing microtiter plate (Beckman Coulter). The plate was then dried on a Speed-Vac SC110 model with a microtiter plate holder (Savant Instruments Inc., Holbrook, NY). Dried samples were immediately resuspended with 25 μ l of formamide, and one drop of mineral oil was added to the top of each sample. Samples were either sequenced immediately on a CEQ 2000 DNA sequencing instrument (Beckman Coulter) or stored at -30°C . In total, 1,820 cDNA library clones was sequenced.

Bioinformatic Tools and Procedures

Expressed sequence tags (ESTs) were trimmed of primer and vector sequences, clustered, and compared with other databases as previously described (Valenzuela et al. 2003). The BLAST tool (Altschul and Gish 1996), CAP3 assembler (Huang and Madan 1999), ClustalW (Thompson et al. 1994), and TreeView software (Page 1996) were used to compare, assemble, and align sequences and to visualize alignments. For functional annotation of the transcripts, we used blastx (Altschul et al. 1997) to compare the nucleotide sequences with the nonredundant (NR) protein database of the National Center of Biological Information (NCBI) and to the Gene Ontology (GO) database (Ashburner et al. 2000). Rpsblast (Schaffer et al. 2001) was used to search for conserved protein domains in the Pfam (Bateman et al. 2000), Smart (Letunic et al. 2002), Kog (Tatusov et al. 2003), and conserved domains (CDD) databases (Marchler-Bauer et al. 2002). We also compared the transcripts with other subsets of mitochondrial and rRNA nucleotide sequences. Segments of the three-frame translations of the EST (because the libraries were unidirectional, we did not use six-frame translations) starting with a methionine in the first 100 predicted aa—or the predicted protein translation, in the case of complete coding sequences—were submitted to the SignalP server (Nielsen et al. 1997) to help identify translation products that could be secreted. *O*-glycosylation sites on the proteins were predicted with the program NetOGlyc (<http://www.cbs.dtu.dk/services/NetOGlyc/>) (Hansen et al. 1998).

Glycosylphosphatidylinositol anchors were detected with the PredGPI anchor predictor (Pierleoni et al. 2008). Functional annotation of the transcripts was based on all the comparisons above. Following inspection of all results, transcripts were classified as either secretory (S), housekeeping (H), or of unknown (UK) function, with further subdivisions based on function and/or protein families. Sequence alignments were done with the ClustalX software package (Thompson et al. 1997). Phylogenetic analysis and statistical neighbor-joining bootstrap tests (10,000 iterations) of the phylogenies were performed with the Mega package (Kumar et al. 2004). Hyperlinked Excel spreadsheets of the assembled ESTs and of the salivary protein database are supplied as Supp. Data 1 and 2 (online only).

The EST sequences from this work were deposited with DBEST under accession numbers JK627586-JK629435. Coding and protein sequences were deposited in GenBank under

accession numbers JP593052-JP593144 and bioproject ID PRJNA74603 accessible at <http://www.ncbi.nlm.nih.gov/bioproject/74603>.

Statistical Analysis

The Chi-square test was used to compare frequencies of ESTs among different functional classes between different triatomine sialotranscriptomes. The calculations were performed using standard functions from the Excel spreadsheet program (Supp. Data S1 [online only]).

Results and Discussion

Assembly of 1,820 cDNA sequences from the salivary cDNA library of *T. rubida* allowed for identification of 657 contigs including 527 singletons (Supp. Data 1 [online only]). These contigs were classified, based on the several blast and rpsblast comparisons to various databases, as being of S, H, or UK class. Additionally, we have found one sequence matching transposable elements, and 13 ESTs matching proteins very similar to those from trypanosomes (Table 1). Nearly half of the ESTs (51.2%) were from the S class, but it represented only 14% of the contigs because this class had on average 10 ESTs per contig, while the H class had 26% of the ESTs and 28.5% of the contigs, with 2.6 ESTs/contig. The unknown class represented 55% of the contigs, but 22% of the ESTs, with 1.1 ESTs per contig (Table 1; Fig. 1).

Functional classification of the housekeeping contigs indicate that the majority are attached to the protein synthesis machinery, energy metabolism, oxidant metabolism, signal transduction associated with apoptosis and protein modification machinery (Table 2). These results are typical of previous sialotranscriptomes of blood sucking arthropods, except for the relatively high representation of oxidant metabolism ahead of the protein modification/export machineries.

The somewhat surprising increase in the oxidant metabolism/detoxification is associated with a relatively large number of ESTs coding for members of the cytochrome P450 family, which is involved in detoxification, including insecticide detoxification, but also in the genesis of prostaglandins, lipid derivatives of arachidonic acid (Simpson 1997, Spector et al. 2004). Of interest, the contig Triru-contig_83 was assembled from 18 ESTs and produces a best match in the GO database to Cyp6a13 from *Drosophila melanogaster* (Supp. Data 1 [online only]). The additional finding of 12 ESTs associated with members of the 15-hydroxyprostaglandin dehydrogenase family, as indicated by the KOG and Swissprot matches (Supp. Data 1 [online only]; contig Triru-contig_86), further supports an active prostaglandin metabolism in this organ. This family of enzymes is associated with the inactivation of prostaglandins and lipoxins (Ensor and Tai 1995) and could be modulating the levels of these eicosanoids in *T. rubida*. It is possible that prostaglandins are involved in secretion control of *Triatoma* saliva, as they are in ticks (Qian et al. 1997, 1998; Yuan et al. 2000), or that prostaglandins may be actually a mediator found in *Triatoma* saliva, as it is in ticks where they are main salivary vasodilators and host skin immunomodulators (Dickinson et al. 1976, Higgs et al. 1976, Ribeiro et al. 1988, Sa-Nunes et al. 2007, Oliveira et al. 2011).

The analysis of transcripts of the S class indicates that the majority code for members of the lipocalin family ($\approx 90\%$ of ESTs and contigs) with members of the Triabin family representing the majority of the lipocalins. Other represented families, such as antigen-5, Cral-Trio, Trialysin, hemolysin, and apyrase are commonly found in the triatomine transcriptomes done so far (Ribeiro et al. 2004; Santos et al. 2007, Assumpção et al. 2008, 2011; Kato et al. 2010). Coding sequences (CDS) were extracted from the contigs from Supp. Data 1 (online only) to generate Supp. Data 2 (online only), from where protein sequences can be retrieved for comparisons of *T. rubida* salivary proteins with those of other

triatomines. What follows is a more detailed analysis of the deduced proteins of the S class (Table 3).

Lipocalins

The lipocalin family is widespread in prokaryotes and eukaryotes where no primary amino acid sequence alone can identify all family members, but all have a quite conserved barrel structure with hydrophobic amino acids lining its interior. This hydrophobic cavity can carry hydrophobic molecules such as lipids, thus explaining the family's name that means "cup of lipid" (Flower et al. 1993, 2000; Andersen et al. 2005). In blood sucking arthropods, this family is widely represented in ticks and triatomine bugs, each species having dozens of genes. This family has various functions, associated with its binding ability, or by the lateral side chains that might interact with host proteins. For example, in the kissing bug *Rhodnius prolixus* Stahl at least six different lipocalins (named nitrophorins) bind nitric oxide (NO) through a heme moiety and deliver this vasodilator and platelet aggregation inhibitor to their host's skin (Champagne et al. 1995a, Andersen and Montfort 2000, Moreira et al. 2003). These NO carriers are not found in other triatomines (Ribeiro et al. 1998). *Rhodnius* lipocalins are involved in scavenging of biogenic amines such as serotonin and histamine (involved in platelet aggregation, vasoconstriction and itching) (Ribeiro and Walker 1994, Andersen et al. 2003) and (by different lipocalins) of adenosine nucleotides (triggers of platelet and neutrophil aggregation) (Francischetti et al. 2000, 2002). *Dipetalogaster maxima* lipocalin named dipetalodipin sequesters thromboxane A₂ (vasoconstrictor and platelet aggregation agonist) and other lipid agonists (Assumpção et al. 2010) Side chains of these proteins are involved in inhibition of clotting by *R. prolixus* nitrophorin 2 (Ribeiro et al. 1995, Zhang et al. 1998) and triabin from *T. pallidipenis* Stahl (Noeske-Jungblut et al. 1995); inhibition of complement function has been attributed to a tick lipocalin (Nunn et al. 2005).

The sialotranscriptome of *T. rubida* reveals over 50 protein sequences that were assembled with two or more ESTs that may represent expression by different genes as well as alleles of the same gene. Ten of these deduced protein sequences are larger than 100 aa and have at least 10% amino acid difference from their nearest sequence, and thus could result from 10 different genes. Alignment of these 10 sequences with their homologues obtained by blastp matches to the nonredundant protein database from NCBI (NR database) with an e value <1e-4 shows the complexity of this protein family in the triatomines (Supp. Fig. 1 [online only]). Thirteen clades are observed with *T. rubida* lipocalins being represented in clades I, III, VI, and XI, this last clade being the most abundant, where six *T. rubida* lipocalins are found within three sub clades (Supp. Fig. 1 [online only]). The most abundantly expressed lipocalin from *T. rubida*, named Triru-10 on Supp. Data 2 (online only), has only 39% identity to *T. protracta* procalin (Paddock et al. 2001), the main allergen identified for this bug. No other *T. rubida* lipocalin has over 40% identity to procalin (Supp. Data 2 [online only]). Triru-56 is the *T. rubida* lipocalin most similar to triabin, the antithrombin of *T. pallidipenis* but it has only 26% sequence identity over 99% of the triabin sequence. These divergences underscore the proposal that salivary proteins of blood sucking arthropods are under a very fast pace of evolution.

Antigen 5

This protein family is also ubiquitous and has been associated with diverse functions, including snake and lizard venoms (Mochca-Morales et al. 1990, Nobile et al. 1996, Yamazaki and Morita 2004), but most have not been characterized functionally. It is found abundantly in diverse gland secretions, such as in vespid venoms, from where the name antigen-5 derives (Hoffman 1993). It is part of the CAP superfamily of proteins that derives its name from the previously defined CRISP (for cysteine rich secretory proteins), antigen-5

and plant pathogen related proteins (Gibbs et al. 2008). The sialotranscriptome of *T. rubida* reveals several members of this family, but all retrieved clones are 5'-truncated. Two fragments with >170 aa were assembled from 39 and 10 ESTs, both covering near 70% of the length of their *T. dimidiata* homologue, with 58% identity.

Apyrase

Apyrases are enzymes that hydrolyze ATP and ADP to AMP thus destroying agonists of platelet and neutrophil aggregation (Francischetti 2010). These enzymes are commonly found in the saliva of blood sucking animals. As is typical of convergent evolution, different protein families with such activity have been recruited to fulfill this role in blood feeding. In the genus *Triatoma*, as in mosquitoes and ticks, members of the 5'-nucleotidase family have been recruited for this purpose (Champagne et al. 1995b, Faudry et al. 2004, Stutzer et al. 2009). In sand flies and bed bugs, and possibly in *Rhodnius*, the *Cimex* type of apyrase is expressed in saliva (Valenzuela et al. 1998, 2001) while fleas have opted for the CD39 family (Andersen et al. 2007). Most 5' nucleotidases are extracellular enzymes typically bound to the cell membrane through a glycosylphosphatidylinositol anchor located in the carboxyterminus domain. Salivary apyrases, though, have a mutated carboxyterminus that do not support the anchor function and these enzymes are thus secreted. We found a single EST in the sialotranscriptome of *T. rubida* coding for a 5'-nucleotidase family member, which matches the final 70 aa carboxyterminus of *T. infestans* apyrase, with 42% identity. When submitted to the PredGPI server (Pierleoni et al. 2008), the *T. infestans* protein is not predicted to have a GPI anchor, as expected. When substituting the last 70 aa of the *T. infestans* protein with the *T. rubida* sequence, the same server does not indicate an anchor, but substituting the last 70 aa of the *T. infestans* protein with the last 70 aa from the rat 5' nucleotidase (NCBI accession gi 148747417) a weak probable anchor is predicted, suggesting the *T. rubida* fragment to be the carboxyterminus of the putative salivary apyrase. The finding of a single apyrase EST in the *T. rubida* sialotranscriptome may indicate a low expression of this enzyme in the saliva of this triatomine species, which was never measured thus far.

Cral-Trio

Proteins with this domain were previously found in the sialotranscriptome of *T. matogrossensis* (gi 307095144) (Assumpção et al. 2012). This domain is typical of secreted proteins that carry lipids, such as the phosphatidylinositol-transfer protein of *Saccharomyces cerevisiae* (Sha et al. 1998). Its function in blood feeding is unknown, but potentially could be associated to binding lipid agonists of inflammation and hemostasis.

Hemolysin Family

This family is uniquely found in sialotranscriptomes of triatomines. Its function is unknown. Triru-412 codes for a truncated protein matching a *T. matogrossensis* of the same family.

Trialsin Family

Trialsin is a salivary antimicrobial protein described from *T. infestans* (Amino et al. 2001b, Martins et al. 2006). Similar proteins are exclusively found on the genus *Triatoma* (Santos et al. 2007, Assumpção et al. 2008). The sequence coded by Triru-377 (Supp. Data 2 [online only]) represents a trialsin member from *T. rubida* having 63% identity to a *T. matogrossensis* homologue.

Deorphanized *T. matogrossensis* Protein Family

Triru-109 codes for a truncated protein having 52% identity over a stretch of 75 aa to an orphan protein of *T. matogrossensis* deposited in GenBank, having no matches to other

known proteins of other organisms. Triru-109 thus characterizes a novel protein family within the *Triatoma* genus.

Comparisons of EST Abundance by Protein Family Among Different Triatomine Species

Saliva of blood sucking arthropods is a complex mixture of compounds that act against a complex set of host reactions including hemostasis, inflammation, and immunity. Because of the fast evolution of the salivary potion of hematophagous arthropods, it can be interesting to compare how each species solves the problem of dosing each available protein family in its sialome. Comparisons between the main salivary secreted proteins families among triatomine sialotranscriptomes (Table 4) shows that *T. rubida* is somewhat unique among three other *Triatoma* transcriptomes in having a relatively high expression of antigen-5 transcripts, accounting for 7.8% of the ESTs associated with secretory products. In this regard the profile is more similar to that of *D. maxima*, where this family represents 4.8% of the transcriptome, with 67 ESTs found. In *T. rubida* 73 ESTs were found coding for Ag-5 proteins, compared with seven in *T. matogrossensis*, one in *T. brasiliensis*, and 13 in *T. infestans*. These differences are significant when compared by a Chi-square test ($P < 0.001$). Most transcriptomes listed in Table 4 have >90% of the transcripts associated with lipocalin expression, except for *T. infestans* (55%) and *T. matogrossensis* (76.4%). These differences are also highly significant when compared by a Chi-square test ($P < 0.001$). These two *Triatoma* species have, however, higher than average expression of transcripts coding for proteins containing the Kazal domain of protease inhibitors (Rimphanitchayakit and Tassanakajon 2010), which are absent in *T. rubida*, *Dipetalogaster*, and *Rhodnius*. These are also highly significant differences ($P < 0.001$). *T. infestans* is also uncommon in having a relatively high expression of the triatomine hemolysin and trialysin families, these two families representing 15% of the ESTs associated with secretory products within the *T. infestans* sialotranscriptome (both comparisons are also highly significantly different, $P < 0.001$). Among the enzymes, *T. brasiliensis* is unique in having 29 ESTs coding for serine proteases, while only two were found in *T. infestans* and none in the remaining sialotranscriptomes ($P < 0.001$, excluding *R. prolixus* and *T. infestans* because their expected numbers were <5). These *Triatoma* serine proteases may be associated with the relatively potent activity, named triapsin, reported in the saliva of *T. infestans* (Amino et al. 2001a). Transcripts coding for inositol phosphatase are also missing in the *T. rubida* sialotranscriptome, while 5, 11, 4, and 34 transcripts were found for *R. prolixus*, *T. infestans*, *T. brasiliensis*, and *T. matogrossensis*, respectively ($P < 0.001$). To the extent that these differences in EST expression determine differences in the amounts of protein expression, the results indicate a quite variable composition of the salivary potion for each individual species.

We have found 13 ESTs in the *T. rubida* sialotranscriptome that produce their best matches to *T. cruzi* proteins. These are mostly singletons, except for a contig assembled from two ESTs coding for a protein with 39% identity to a *T. cruzi* protein annotated as GTPase activating protein of Rab-like GTPase, 34% identity to the *T. vivax* homologue, and 33% to a *Leishmania donovani* protein. This contig clearly derives from a trypanosomatid. The additional 10 singletons produce an average 72% identity at the amino acid level to different *T. cruzi* proteins, except for one contig that produces 92% identity to the ribosomal protein L3 from *T. brucei gambiense*, and 90% identity to the *T. cruzi* homologue. All these putative *Trypanosoma*-derived sequences are clearly from a trypanosomatid, and not from conserved eukaryotic proteins, as can be deduced from the inspection of the blasts against the NR database.

We conclude that the sialotranscriptome herein described shows the diversity of salivary proteins in triatomine bugs, in particular the large differences in sequence identity between *T. rubida* and *T. protracta*, two kissing bugs with overlapping geographical distribution,

explaining the previous finding that sera of people allergic to these bugs are not cross reactive (Marshall et al. 1986, Pinnas et al. 1986). The transcriptome also reveals sequences that are similar, but not identical to *T. cruzi*, representing 0.1% of the transcripts found in the salivary gland, suggesting the presence of a different *T. cruzi* strain in *T. rubida*. The proteins here described provide for a platform for screening human and pet sera immunoreactivity to recombinant *T. rubida* salivary proteins that can be used for desensitization therapies for *T. rubida* allergy, as it was attempted for *T. protracta* (Marshall 1982), as well as for designing recombinant epidemiological markers of *T. rubida* exposure, as has been done for *T. brasiliensis* (Schwarz et al. 2009), *Anopheles gambiae* (Poinssignon et al. 2008), sand flies (Souza et al. 2010), and ticks (Sanders et al. 1998, 1999).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank B. Savary for kindly providing specimens of *T. rubida*. This work was supported by the Intramural Research Program of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

References Cited

- Altschul SF, Gish W. Local alignment statistics. *Methods Enzymol.* 1996; 266:460–480. [PubMed: 8743700]
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–3402. [PubMed: 9254694]
- 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 Biochem Molec Biol.* 2001a; 31:465–472. [PubMed: 11222956]
- Amino R, Martins RM, Procopio J, Hirata IY, Juliano MA, Schenkman S. Trialysin, a novel pore-forming protein from Saliva of haematophagous insects activated by limited proteolysis. *J Biol Chem.* 2001b; 19:19.
- Andersen JF, Montfort WR. The crystal structure of nitrophorin 2. A trifunctional antihemostatic protein from the saliva of *Rhodnius prolixus*. *J Biol Chem.* 2000; 275:30496–30503. [PubMed: 10884386]
- Andersen JF I, Francischetti M, 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; 278:4611–4617. [PubMed: 12464610]
- Andersen JF, Gudderra NP, Francischetti IM, Ribeiro JM. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. *Arch Insect Biochem Physiol.* 2005; 58:97–105. [PubMed: 15660358]
- Andersen JF, Hinnebusch BJ, Lucas DA, Conrads TP, Veenstra TD, Pham VM, Ribeiro JM. An insight into the sialome of the oriental rat flea, *Xenopsylla cheopis* (Rots). *BMC Genomics.* 2007; 8:102. [PubMed: 17437641]
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. *Gene Ontol Consort Nat Genet.* 2000; 25:25–29.
- Assumpção TC I, Francischetti M, Andersen JF, Schwarz A, Santana JM, Ribeiro JM. An insight into the sialome of the blood-sucking bug *Triatoma infestans*, a vector of Chagas' disease. *Insect Biochem Molec Biol.* 2008; 38:213–232. [PubMed: 18207082]
- Assumpção 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: TXA₂, PGF₂α, and 15(S)-HETE. *J Biol Chem.* 2010; 285:39001–39012. [PubMed: 20889972]
- Assumpção TC, Charneau S, Santiago PB, Francischetti IM, Meng Z, Araujo CN, Pham VM, Queiroz RM, de Castro CN, Ricart CA, Santana JM, Ribeiro JM. Insight into the salivary transcriptome and proteome of *Dipetalogaster maxima*. *J Proteome Res.* 2011; 10:669–679. [PubMed: 21058630]
- Assumpção TC, Eaton DP, Pham VM, Francischetti IMB, Aoki V, Hans-Filho G, Rivitti EA, Valenzuela JG, Diaz LA, Ribeiro JMC. An insight into the sialotranscriptome of *Triatoma matogrossensis*, a kissing bug associated with Fogo Selvagem in South America. *Am J Tropical Med Hyg.* 2012 (in press).
- Barr SC. Canine Chagas' disease (American trypanosomiasis) in North America. *Vet Clin North Am Small Anim Pract.* 2009; 39:1055, 1064, v–vi. [PubMed: 19932362]
- Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res.* 2000; 28:263–266. [PubMed: 10592242]
- Beard CB, Pye G, Steurer FJ, Rodriguez R, Campman R, Peterson AT, Ramsey J, Wirtz RA, Robinson LE. Chagas disease in a domestic transmission cycle, southern Texas, USA. *Emerg Infect Dis.* 2003; 9:103–105. [PubMed: 12533289]
- Bern C, Kjos S, Yabsley MJ, Montgomery SP. *Trypanosoma cruzi* and Chagas' disease in the United States. *Clin Microbiol Rev.* 2011; 24:655–681. [PubMed: 21976603]
- Brown EL, Roellig DM, Gompfer ME, Monello RJ, Wenning KM, Gabriel MW, Yabsley MJ. Seroprevalence of *Trypanosoma cruzi* among eleven potential reservoir species from six states across the southern United States. *Vector Borne and Zoonotic Dis.* 2010; 10:757–763.
- Champagne D, Nussenzeig RH, Ribeiro JMC. Purification, characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood sucking insect *Rhodnius prolixus*. *J Biol Chem.* 1995a; 270:8691–8695. [PubMed: 7721773]
- Champagne DE, Smartt CT, Ribeiro JM, James AA. The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a member of the 5'-nucleotidase family. *Proc Natl Acad Sci USA.* 1995b; 92:694–698. [PubMed: 7846038]
- Dickinson RG, O'Hagan JE, Sholtz M, Binnington KC, Hegarty MP. Prostaglandin in the saliva of the cattle tick *Boophilus microplus*. *Aust J Exp Biol Med Sci.* 1976; 54:475–486. [PubMed: 1021092]
- Dorn PL, Perniciaro L, Yabsley MJ, Roellig DM, Balsamo G, Diaz J, Wesson D. Autochthonous transmission of *Trypanosoma cruzi*, Louisiana. *Emerg Infect Dis.* 2007; 13:605–607. [PubMed: 17553277]
- Ensor CM, Tai HH. 15-Hydroxyprostaglandin dehydrogenase. *J Lipid Mediat Cell Signal.* 1995; 12:313–319. [PubMed: 8777575]
- Faudry E, Lozzi SP, Santana JM, D'Souza-Ault M, Kieffer S, Felix CR, Ricart CA, Sousa MV, Vernet T, Teixeira AR. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. *J Biol Chem.* 2004; 279:19607–19613. [PubMed: 14985353]
- Flower DR, North ACT, Attwood TK. Structure and sequence relationships in the lipocalins and related proteins. *Prot Sci.* 1993; 2:753–761.
- Flower DR, North AC, Sansom CE. The lipocalin protein family: structural and sequence overview. *Biochim Biophys Acta.* 2000; 1482:9–24. [PubMed: 11058743]
- Francischetti IM. Platelet aggregation inhibitors from hematophagous animals. *Toxicon.* 2010; 56:1130–1144. [PubMed: 20035779]
- 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; 275:12639–12650. [PubMed: 10777556]
- 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; 41:3810–3818. [PubMed: 11888300]
- 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 Rev.* 2008; 29:865–897. [PubMed: 18824526]

- Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, Brunak S. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj J*. 1998; 15:115–130. [PubMed: 9557871]
- Herwaldt BL, Grijalva MJ, Newsome AL, McGhee CR, Powell MR, Nemec DG, Steurer FJ, Eberhard ML. Use of polymerase chain reaction to diagnose the fifth reported US case of autochthonous transmission of *Trypanosoma cruzi*, in Tennessee, 1998. *J Infect Dis*. 2000; 181:395–399. [PubMed: 10608796]
- Higgs GA, Vane JR, Hart RJ, Porter C, Wilson RG. Prostaglandins in the saliva of the cattle tick, *Boophilus microplus* (Canestrini) (Acarina, Ixodidae). *Bull Entomol Res*. 1976; 66:665–670.
- Hoffman DR. Allergens in Hymenoptera venom. XXV: the amino acid sequences of antigen 5 molecules and the structural basis of antigenic cross-reactivity. *J Allergy Clin Immunol*. 1993; 92:707–716. [PubMed: 8227862]
- Huang X, Madan A. CAP3: a DNA sequence assembly program. *Genome Res*. 1999; 9:868–877. [PubMed: 10508846]
- Ibarra-Cerdena CN, Sanchez-Cordero V, Townsend Peterson A, Ramsey JM. Ecology of North American Triatominae. *Acta Trop*. 2009; 110:178–186. [PubMed: 19084490]
- Kato H, Jochim RC, Gomez EA, Sakoda R, Iwata H, Valenzuela JG, Hashiguchi Y. 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; 10:184–191. [PubMed: 19900580]
- Kjos SA, Snowden KF, Craig TM, Lewis B, Ronald N, Olson JK. Distribution and characterization of canine Chagas disease in Texas. *Vet Parasitol*. 2008; 152:249–256. [PubMed: 18255233]
- Kjos SA, Snowden KF, Olson JK. Biogeography and *Trypanosoma cruzi* infection prevalence of Chagas disease vectors in Texas, USA. *Vector Borne and Zoonotic Dis*. 2009; 9:41–50.
- Klotz JH, Dorn PL, Logan JL, Stevens L, Pinnas JL, Schmidt JO, Klotz SA. “Kissing bugs”: potential disease vectors and cause of anaphylaxis. *Clin Infect Dis*. 2010; 50:1629–1634. [PubMed: 20462351]
- Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform*. 2004; 5:150–163. [PubMed: 15260895]
- Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP, Bork P. Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res*. 2002; 30:242–244. [PubMed: 11752305]
- Mans BJ, Neitz AW. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem Mol Biol*. 2004; 34:1–17. [PubMed: 14723893]
- Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res*. 2002; 30:281–283. [PubMed: 11752315]
- Marshall N. Allergy to *Triatoma protracta* (Heteroptera: Reduviidae). II. Antigen production in vitro. *J Med Entomol*. 1982; 19:253–254. [PubMed: 6750127]
- Marshall NA, Street DH. Allergy to *Triatoma protracta* (Heteroptera: Reduviidae). I. Etiology, antigen preparation, diagnosis and immunotherapy. *J Med Entomol*. 1982; 19:248–252. [PubMed: 6750126]
- Marshall NA, Chapman MD, Saxon A. Species-specific allergens from the salivary glands of Triatominae (Heteroptera:Reduviidae). *J Allergy Clin Immunol*. 1986; 78:430–435. [PubMed: 3531291]
- Martins RM, Sforca ML, Amino R, Juliano MA, Oyama S Jr, Juliano L, Pertinhez TA, Spisni A, Schenkman S. Lytic activity and structural differences of amphipathic peptides derived from trypsin. *Biochemistry*. 2006; 45:1765–1774. [PubMed: 16460023]
- Milei J, Guerri-Guttenberg RA, Grana DR, Storino R. Prognostic impact of Chagas disease in the United States. *Am Heart J*. 2009; 157:22–29. [PubMed: 19081392]
- Mochca-Morales J, Martin BM, Possani LD. Isolation and characterization of helothermine, a novel toxin from *Heloderma horridum horridum* (Mexican beaded lizard) venom. *Toxicon*. 1990; 28:299–309. [PubMed: 1693019]
- Moffitt JE, Venarske D, Goddard J, Yates AB, deShazo RD. Allergic reactions to *Triatoma* bites. *Ann Allergy Asthma Immunol*. 2003; 91:122–128. quiz 128–130, 194. [PubMed: 12952105]

- Moncayo A, Ortiz Yanine MI. An update on Chagas disease (human American trypanosomiasis). *Ann Trop Med Parasitol*. 2006; 100:663–677. [PubMed: 17227647]
- Moreira MF, Coelho HS, Zingali RB, Oliveira PL, Masuda H. Changes in salivary nitrophorin profile during the life cycle of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem Mol Biol*. 2003; 33:23–28. [PubMed: 12459197]
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engl*. 1997; 10:1–6.
- Nobile M, Noceti F, Prestipino G, Possani LD. Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules. *Exp Brain Res*. 1996; 110:15–20. [PubMed: 8817251]
- Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD. Triabin, a highly potent exosite inhibitor of thrombin. *J Biol Chem*. 1995; 270:28629–28634. [PubMed: 7499380]
- Nunn MA, Sharma A, Paesen GC, Adamson S, Lissina O, Willis AC, Nuttall PA. Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. *J Immunol*. 2005; 174:2084–2091. [PubMed: 15699138]
- Ochs DE V, Hnilica S, Moser DR, Smith JH, Kirchhoff LV. Postmortem diagnosis of autochthonous acute chagasic myocarditis by polymerase chain reaction amplification of a species-specific DNA sequence of *Trypanosoma cruzi*. *Am J Trop Med Hyg*. 1996; 54:526–529. [PubMed: 8644910]
- Oliveira CJ, Sa-Nunes A, Francischetti IM, Carregaro V, Anatriello E, Silva JS, de Miranda Santos IK, Ribeiro JM, Ferreira BR. Deconstructing tick saliva: non protein molecules with potent immunomodulatory properties. *J Biol Chem*. 2011; 286:10960–10969. [PubMed: 21270122]
- Packchianian A. Reservoirs hosts of Chagas' disease in the state of Texas. *Am J Trop Med Hyg*. 1942; 22:623–631.
- 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; 167:2694–2699. [PubMed: 11509613]
- Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*. 1996; 12:357–358. [PubMed: 8902363]
- Pan American Health Organization. Quantitative estimation of Chagas' disease in the Americas. Pan American Health Organization; Montevideo, Uruguay: 2006.
- Pfeiler E, Bitler BG, Ramsey JM, Palacios-Cardiel C, Markow TA. Genetic variation, population structure, and phylogenetic relationships of *Triatoma rubida* and *T. recurva* (Hemiptera: Reduviidae: Triatominae) from the Sonoran Desert, insect vectors of the Chagas' disease parasite *Trypanosoma cruzi*. *Mol Phylogenetics Evol*. 2006; 41:209–221.
- Pierleoni A, Martelli PL, Casadio R. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*. 2008; 9:392. [PubMed: 18811934]
- Pinnas JL, Lindberg RE, Chen TM, Meinke GC. Studies of kissing bug-sensitive patients: evidence for the lack of cross-reactivity between *Triatoma protracta* and *Triatoma rubida* salivary gland extracts. *J Allergy Clin Immunol*. 1986; 77:364–370. [PubMed: 2418092]
- Poinsignon A, Cornelié S, Mestres-Simon M, Lanfrancotti A, Rossignol M, Boulanger D, Cisse B, Sokhna C, Arca B, Simondon F, Remoue F. Novel peptide marker corresponding to salivary protein gSG6 potentially identifies exposure to *Anopheles* bites. *PLoS ONE*. 2008; 3:e2472. [PubMed: 18575604]
- Qian Y, Essenberg RC, Dillwith JW, Bowman AS, Sauer JR. A specific prostaglandin E2 receptor and its role in modulating salivary secretion in the female tick, *Amblyomma americanum* (L.). *Insect Biochem Molec Biol*. 1997; 27:387–395. [PubMed: 9219365]
- Qian Y, Yuan J, Essenberg RC, Bowman AS, Shook AL, Dillwith JW, Sauer JR. Prostaglandin E2 in the salivary glands of the female tick, *Amblyomma americanum* (L.): calcium mobilization and exocytosis. *Insect Biochem Molec Biol*. 1998; 28:221–228. [PubMed: 9684330]
- Reisenman CE, Lawrence G, Guerenstein PG, Gregory T, Dotson E, Hildebrand JG. Infection of kissing bugs with *Trypanosoma cruzi*, Tucson, Arizona, USA. *Emerg Infect Dis*. 2010; 16:400–405. [PubMed: 20202413]
- Reisenman CE, Gregory T, Guerenstein PG, Hildebrand JG. Feeding and defecation behavior of *Triatoma rubida* (Uhler, 1894) (Hemiptera: Reduviidae) under laboratory conditions, and its

- potential role as a vector of Chagas disease in Arizona, USA. *Am J Trop Med Hyg.* 2011; 85:648–656. [PubMed: 21976567]
- Ribeiro JM, Walker FA. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorin) of *Rhodnius prolixus*. *J Exp Med.* 1994; 180:2251–2257. [PubMed: 7964498]
- Ribeiro JM, Makoul GT, Robinson DR. *Ixodes dammini*: evidence for salivary prostacyclin secretion. *J Parasitol.* 1988; 74:1068–1069. [PubMed: 3057165]
- Ribeiro JM, Schneider M, Guimaraes JA. Purification and characterization of prolixin S (nitrophorin 2), the salivary anticoagulant of the blood-sucking bug *Rhodnius prolixus*. *Biochem J.* 1995; 308(Pt 1):243–249. [PubMed: 7755571]
- Ribeiro JM, Schneider M, Isaias T, Jurberg J, Galvao C, Guimaraes JA. Role of salivary antihemostatic components in blood feeding by triatomine bugs (Heteroptera). *J Med Entomol.* 1998; 35:599–610. [PubMed: 9701952]
- Ribeiro JM, Andersen J, Silva-Neto MA, Pham VM, Garfield MK, Valenzuela JG. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem Mol Biol.* 2004; 34:61–79. [PubMed: 14976983]
- Ribeiro JMC. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? *Infect Agents Dis.* 1995; 4:143–152. [PubMed: 8548192]
- Ribeiro JMC, Arca B. From sialomes to the sialoverse: an insight into the salivary potion of blood feeding insects. *Adv Insect Physiol.* 2009; 37:59–118.
- Rimphanitchayakit V, Tassanakajon A. Structure and function of invertebrate Kazal-type serine proteinase inhibitors. *Dev Comp Immunol.* 2010; 34:377–386. [PubMed: 19995574]
- Roellig DM, Ellis AE, Yabsley MJ. Oral transmission of *Trypanosoma cruzi* with opposing evidence for the theory of carnivory. *J Parasitol.* 2009; 95:360–364. [PubMed: 18763853]
- Ryckman RE. The vertebrate hosts of the Triatominae of North and Central America and the West Indies (Hemiptera: Reduviidae: Triatominae). *Bull Soc Vector Ecol.* 1986; 11:221–241.
- Sa-Nunes A, Bafica A, Lucas DA, Conrads TP, Veenstra TD, Andersen JF, Mather TN, Ribeiro JM, Francischetti IM. Prostaglandin E2 is a major inhibitor of dendritic cell maturation and function in *Ixodes scapularis* saliva. *J Immunol.* 2007; 179:1497–1505. [PubMed: 17641015]
- Sanders ML, Glass GE, Nadelman RB, Wormser GP, Scott AL, Raha S, Ritchie BC, Jaworski DC, Schwartz BS. Antibody levels to recombinant tick calreticulin increase in humans after exposure to *Ixodes scapularis* (Say) and are correlated with tick engorgement indices. *Am J Epidemiol.* 1999; 149:777–784. [PubMed: 10206628]
- Sanders ML, Jaworski DC, Sanchez JL, DeFraités RF, Glass GE, Scott AL, Raha S, Ritchie BC, Needham GR, Schwartz BS. Antibody to a cDNA-derived calreticulin protein from *Amblyomma americanum* as a biomarker of tick exposure in humans. *Am J Trop Med Hyg.* 1998; 59:279–285. [PubMed: 9715947]
- Santos A, Ribeiro JM, Lehane MJ, Gontijo NF, Veloso AB, Sant’Anna MR, Nascimento Araujo R, Grisard EC, Pereira MH. The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). *Insect Biochem Molec Biol.* 2007; 37:702–712. [PubMed: 17550826]
- Sarkar S, Strutz SE, Frank DM, Rivaldi CL, Sissel B, Sanchez-Cordero V. Chagas disease risk in Texas. *PLoS Neglected Trop Dis.* 2010; 4:e836. pii.
- Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res.* 2001; 29:2994–3005. [PubMed: 11452024]
- Schiffler RJ, Mansur GP, Navin TR, Limpakarnjanarat K. Indigenous Chagas’ disease (American trypanosomiasis) in California. *JAMA.* 1984; 251:2983–2984. [PubMed: 6425516]
- Schwarz A, Helling S, Collin N, Teixeira CR, Medrano-Mercado N, Hume JC, Assumpção TC, Marcus K, Stephan C, Meyer HE, et al. Immunogenic salivary proteins of *Triatoma infestans*: development of a recombinant antigen for the detection of low-level infestation of triatomines. *PLoS Neglected Trop Dis.* 2009; 3:e532.
- Sha B, Phillips SE, Bankaitis VA, Luo M. Crystal structure of the *Saccharomyces cerevisiae* phosphatidylinositol-transfer protein. *Nature.* 1998; 391:506–510. [PubMed: 9461221]

- Simpson AE. The cytochrome P450 4 (CYP4) family. *Gen Pharmacol.* 1997; 28:351–359. [PubMed: 9068972]
- Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, Alcantara R, Ruiz D, Soto M, Teixeira CR, et al. Using recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector exposure in visceral Leishmaniasis endemic areas. *PLoS Negl Trop Dis.* 2010; 4:e649. [PubMed: 20351785]
- Spector AA, Fang X, Snyder GD, Weintraub NL. Epoxyeicosatrienoic acids (EETs): metabolism and biochemical function. *Prog Lipid Res.* 2004; 43:55–90. [PubMed: 14636671]
- Stutzer C, Mans BJ, Gaspar AR, Neitz AW, Maritz-Olivier C. *Ornithodoros savignyi*: soft tick apyrase belongs to the 5'-nucleotidase family. *Exp Parasitol.* 2009; 122:318–327. [PubMed: 19393241]
- Tanowitz HB, Machado FS, Jelicks LA, Shirani J, de Carvalho AC, Spray DC, Factor SM, Kirchhoff LV, Weiss LM. Perspectives on *Trypanosoma cruzi*-induced heart disease (Chagas disease). *Prog Cardiovasc Dis.* 2009; 51:524–539. [PubMed: 19410685]
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics.* 2003; 4:41. [PubMed: 12969510]
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22:4673–4680. [PubMed: 7984417]
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997; 25:4876–4882. [PubMed: 9396791]
- 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; 273:30583–30590. [PubMed: 9804829]
- 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; 204:229–237. [PubMed: 11136609]
- Valenzuela JG I, Francischetti M, Pham VM, Garfield MK, Ribeiro JM. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochem Mol Biol.* 2003; 33:717–732. [PubMed: 12826099]
- Wolf AF. Sensitivity to triatoma bite. *Ann Allergy.* 1969; 27:271. [PubMed: 4891706]
- Wood S. Observations on vectors of Chagas' disease in the United States. I. California. *Bull South Calif Acad Sci.* 1942; 41:61–69.
- Wood S. Observations on vectors of Chagas' Disease in the United States. II Arizona. *Am J Trop Med Hyg.* 1943; 23:315–320.
- Wood S, Wood F. Observations on vectors of Chagas' Disease in the United States. III. New Mexico. *Am J Trop Med Hyg.* 1961; 10:155–165. [PubMed: 13786516]
- Woody NC, Woody HB. American trypanosomiasis (Chagas's Disease): first indigenous case in the United States. *J Am Med Assoc.* 1955; 159:676–677. [PubMed: 13263111]
- Yamazaki Y, Morita T. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon.* 2004; 44:227–231. [PubMed: 15302528]
- Yoshida N. Molecular mechanisms of *Trypanosoma cruzi* infection by oral route. *Mem Instit Oswaldo Cruz.* 2009; 104(Suppl 1):101–107.
- Yuan J, Bowman AS, Aljamali M, Payne MR, Tucker JS, Dillwith JW, Essenberg RC, Sauer JR. Prostaglandin E(2)-stimulated secretion of protein in the salivary glands of the lone star tick via a phosphoinositide signaling pathway. *Insect Biochem Molec Biol.* 2000; 30:1099–1106. [PubMed: 10989297]
- 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; 37:10681–10690. [PubMed: 9692958]

Fig. 1. Functional classification of transcripts originating from the salivary glands of *T. rubida* as a function of the number of assembled contigs or ESTs.

Table 1
Functional classification of transcripts originating from the salivary gland transcriptome of *T. rubida*

Class	No. contigs	No. ESTs	ESTs/contig	% contigs	% ESTs
Secreted	93	932	10.0	14.2	51.2
Housekeeping	187	477	2.6	28.5	26.2
Unknown	364	397	1.1	55.4	21.8
Protozoal	12	13	1.1	1.8	0.7
Transposable element	1	1	1		
Total	657	1820			

Table 2Functional classification of housekeeping transcripts originating from the salivary gland transcriptome of *T. rubida*

Class	No. contigs	No. ESTs	ESTs/contig	% contigs	% ESTs
Protein synthesis machinery	81	287	3.54	43.32	60.17
Metabolism, energy	16	52	3.25	8.56	10.90
Oxidant metabolism/detoxification	10	28	2.80	5.35	5.87
Signal transduction, apoptosis	13	18	1.38	6.95	3.77
Protein modification machinery	10	16	1.60	5.35	3.35
Unknown, conserved	14	14	1.00	7.49	2.94
Metabolism, lipid	4	13	3.25	2.14	2.73
Signal transduction	9	13	1.44	4.81	2.73
Cytoskeletal	9	12	1.33	4.81	2.52
Protein export machinery	6	8	1.33	3.21	1.68
Proteasome machinery	7	7	1.00	3.74	1.47
Storage	2	3	1.50	1.07	0.63
Metabolism, nucleotide	2	2	1.00	1.07	0.42
Nuclear regulation	1	1	1.00	0.53	0.21
Metabolism, carbohydrate	1	1	1.00	0.53	0.21
Metabolism, amino acid	1	1	1.00	0.53	0.21
Immunity	1	1	1.00	0.53	0.21
Total	187	477			

Table 3
 Classification of putative secreted transcripts originating from the salivary gland transcriptome of *T. rubida*

Class	No. contigs	No. ESTs	ESTs/contig	% contigs	% ESTs
Lipocalins of the triabin family	40	609	15.2	43.0	65.3
Other lipocalins	34	239	7.0	36.6	25.6
Antigen 5	10	73	7.3	10.8	7.8
Cral-Trio family	2	2	1.0	2.2	0.2
Trialsin family	2	2	1.0	2.2	0.2
Triatomine hemolysin family	1	1	1.0	1.1	0.1
Apyrase	1	1	1.0	1.1	0.1
Similar to <i>T. matogrossensis</i> proteins	2	4	2.0	2.2	0.4
Putative secreted peptide	1	1	1.0	1.1	0.1
Total	93	932			

Table 4
Comparisons between the main secreted protein families in Triatomine sialotranscriptomes

Protein family	<i>R. prolixus</i> (782)	<i>D. maxima</i> (1380)	<i>T. infestans</i> (645)	<i>T. brasiliensis</i> (1168)	<i>T. matogrossensis</i> (1182)	<i>T. rubida</i> (932)
Lipocalins	90.53 (708)	93.5 (1,290)	55 (355)	93.8 (1,096)	76.14 (900)	90.1 (848)
Antigen 5	1 (8)	4.8 (67)	2 (13)	0.08 (1)	0.59 (7)	7.8 (73)
Kazal	0	0	2.63 (17)	2.22 (26)	1.3 (16)	0
Hemolysin/Mys family	5.7 (45)	0.07 (1)	4.9 (32)	0	2 (24)	0.1 (1)
Trialsin	0	0	11.5 (74)	0	4.7 (56)	0.2 (2)
Inositol phosphatase	0.34 (5)	0	1.7 (11)	0.34 (4)	2.87 (34)	0
5' nucleotidase/apyrase	0	0.6 (9)	0.93 (6)	0.08 (1)	2.7 (32)	0.1 (1)
Trypsin	0	0	0.3 (2)	2.5 (29)	0	0

The table results represent the percentage (and the actual no. in parenthesis) of ESTs belonging to the indicated protein family. The total no. of ESTs for the secreted class in each sialotranscriptome follows the species name in the table heading.