

Research Article

Variability of defensin genes from a Mexican endemic Triatominae: *Triatoma (Meccus) pallidipennis* (Hemiptera: Reduviidae)

Paulina Díaz-Garrido¹, Omar Sepúlveda-Robles², Ignacio Martínez-Martínez¹ and Bertha Espinoza¹

¹Departamento de Inmunología, Instituto de Investigaciones Biomédicas, UNAM 04510, Ciudad de México, México; ²Catedrático CONACyT – Unidad de Investigación Médica en Epidemiología Clínica UMAE-Hospital de Pediatría, Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México

Correspondence: Bertha Espinoza (besgu@iibiomedicas.unam.mx)



Chagas disease remains a serious health problem for countries where the most common mode of transmission is infection contracted from the feces of a Triatominae insect vector. In México, 32 species of *Triatoma* have been identified; amongst them, *Triatoma (Meccus) pallidipennis* is an endemic species reported to have high percentages of infection with *T. cruzi*. Defensins, cysteine-rich cationic peptides, are a family of antimicrobial peptides (AMPs); the synthesis of these molecules is crucial for insect's immune defense. In the present study, the genes encoding defensins in *T. pallidipennis* were sequenced with the purpose of identifying the variability of these genes in a Mexican vector of *T. cruzi*. We found 12 different genes encoding three mature peptides, all of which had the typical folding of a functional insect defensin. In this work two Defensins type 1 and one type 4 were identified. The pro-peptide domain was highly variable and the mature peptide was not. This is the first report focus on variability of defensins from an epidemiologically important *Triatoma* in Mexico.

Introduction

Chagas disease, one of the neglected tropical diseases, is caused by the protozoan *Trypanosoma cruzi*, the main transmission is through the feces of hematophagous insects of the Triatominae subfamily (Hemiptera: Reduviidae). It is estimated that this disease causes more than 10000 deaths per year and that 8 million people are infected with this parasite [1]. In México, this disease is spread by many species of *Triatoma*. Ramsey et al. (2015) [2] reported 31 different species in the country, including the endemic *Triatoma (Meccus) pallidipennis* (*T. pallidipennis*) that forms part of the Phyllosoma complex and colonizes extensive areas in the center of the country, living in domestic and peridomestic environments and more importantly with a natural infection index of 50% in some regions [3,4].

To control the parasite and pathogen proliferation, insect immune defense includes cellular and humoral responses. Humoral immunity consists of the generation of prophenoloxidase (PPO), hemolymph coagulation cascades, reactive oxygen species (ROS), reactive nitrogen species (RNS), and antimicrobial peptides (AMPs). Most AMPs are synthesized as inactive precursor proteins or propeptides, and active peptides or mature peptides are generated by proteolysis; these molecules have a broad spectrum of targeted organisms ranging from viruses to parasites and are produced by several species other than insects, including species of bacteria, plants, and vertebrates [5,6].

Most insect AMPs are small cationic/basic peptides with activities against bacteria and/or fungi, some AMPs also show activities against several parasites and viruses [7,8]. The fat body of insects is the main organ responsible for the production of AMPs, which can diffuse into the hemolymph and circulates throughout the entire insect body [9]. They can also be secreted by the gut epithelial cells, where parasites

Received: 23 June 2018
Revised: 16 August 2018
Accepted: 30 August 2018

Accepted Manuscript Online:
04 September 2018
Version of Record published:
28 September 2018

such as *T. cruzi* might interact directly with them, in addition to other potential pathogens acquired during feeding [10,11].

Insect defensins are a family of AMPs, distinguished for having six conserved cysteine residues that are linked in three pairs of disulphide bridges (Cys^I–Cys^{IV}, Cys^{II}–Cys^V, and Cys^{III}–Cys^{VI}) and form a consensus folding pattern consisting of the cysteine-stabilized α -helix β -sheet (CS $\alpha\beta$) motif that is the active site of the domain for antimicrobial activity [12–16]. These cationic peptides have a molecular mass of approximately 4 kDa, are typically 33–46 amino acid residues long and show activity mainly against Gram-positive bacteria through a channel-forming mechanism [5,17,18].

Several AMPs have already been found in different species of Triatominae such as *Rhodnius prolixus*, *T. brasiliensis* and *T. infestans*. The only sequenced genome of a hemipteran is that of *R. prolixus*, where 11 genes of defensins were identified [19]. Two transcriptomic analyses of *T. pallidipennis* have been reported, including one of the salivary glands, where no AMPs were found [20]. The other transcriptomic analysis reported eight transcripts of defensins and three transcripts of lysozymes [21].

These facts and their importance for immune response are reasons to explore in detail the variability of the defensin-encoding genes in other important Triatominae vectors of Chagas disease.

Most of the published studies on this topic focus on the expression of mRNA of South American species such as *T. infestans* and some members of the *T. brasiliensis* complex (*T. brasiliensis*, *T. sherlocki*, or *T. juazeirensis*), where the expression of these molecules in different organs has been analyzed [22–26].

Information on other important vectors of Chagas disease from the northern part of the continent is missing, for this reason the aim of the present study was to analyze the variability of defensin genes in the genome of the Mexican endemic bug *T. pallidipennis*. We present a comparative analysis with other defensins previously described and the predicted structures of the mature peptides using two different computational tools for the modeling of the molecules. The presented results offer the possibility of interpreting the phylogenetic relationships of hemipteran defensins, and the predicted mature peptide sequences could provide information for the development of novel drug targets.

Materials and methods

Insect origin and maintenance

For this work, we used adult females of *T. pallidipennis*. The insects were maintained in a colony at the Laboratorio de Estudios sobre la Tripanosomiasis Americana, Instituto de Investigaciones Biomédicas, UNAM, under controlled temperature and humidity. The insects were fed with rabbit blood, and all the procedures were carried out according to the Ethic Code of the Instituto de Investigaciones Biomédicas for animal care.

Genomic DNA isolation

Genomic DNA was isolated from both the fat body and digestive tract (washed several times with saline solution to eliminate intestinal contents) and combined; the organs were macerated in 1 ml of lysis solution (50 mM Tris/HCl, pH 8; 100 mM EDTA, pH 8; 100 mM NaCl; 1% SDS, and 20 μ g/ml RNase) and incubated at 37°C overnight. The phenol chloroform technique was used to extract DNA [27]. Nucleic acid integrity and concentration were quantitated using a NanoDrop 1000 (Thermo Scientific) and checked by gel electrophoresis. The DNA was kept at 4°C until used.

Gene amplification cloning and sequencing

The specific forward and reverse primers used to amplify defensin genes of *T. pallidipennis* were designed using the sequences encoding Def1, Def3, and Def4 of *T. brasiliensis* reported by Araújo et al. (2006) [22] and Waniek et al. (2009) [25]: [Def1 – Fwd (5'-GGCGCCATGAAGTGCCTACTCTCTTTG-3'), Def1 – Rev (5'-CATTGCAGGAAGTGAATTGATGCTTGGG-3'), Def3 – Fwd (5'-GGCGCCATGAAGTGTGCACTCTCTTTG-3'), Def3 – Rev (5'-CTATCTGTCACTGCAGGAAGTGAGCTTGGG-3'), Def4 – Fwd (5'-GGCGCCATCATGAAGTGTGCACTCTC-3') and Def4 – Rev (5'-CACTGCAGGAAGTGAAGAGTGCTTGGG-3')]. PCRs were performed in a T100 thermal cycler (Bio-Rad, Inc); after one initial cycle of 95°C for 5 min, PCR was carried out for 30 cycles (1 min at 94°C, 1 min of Tm at 66°C (Def1-oligos) or 70°C (Def3/Def4-oligos), and 1 min at 72°C), followed by a final elongation step of 10 min at 72°C. The putative defensin genes that were amplified (~400 bp) were purified by cutting the gel band and cloning into pJET using a CloneJet PCR Cloning Kit (Thermo Fisher Scientific, Inc) following the manufacturer's instructions. Finally, plasmids were isolated from 15 randomly selected clones and sequenced (five sequences per gene: Def1, Def3, and Def4) using a pJET1.2 forward sequencing primer (5'-CGACTCACTATAGGGAGAGCGGC-3'). The sequencing was done with the Sanger technique, using ABI PRISM 3100 equipment.

Analysis of sequences and identities

T. pallidipennis intron and mRNA sequence identification was performed by alignment using the data from the GenBank Def1 (AY641574.1), Def3 (EU694177.1), and Def4 (EU694178.1) mRNA nucleotide sequences of *T. brasiliensis*. The amino acid sequences were deduced with the translate tool from ExPASy and aligned with the MultAlin server <http://multalin.toulouse.inra.fr/multalin/multalin.html> [28]. Predicted signal peptide cleavage sites and propeptide cleavage sites were calculated using the ProP 1.0 Server of Technical University of Denmark (<http://www.cbs.dtu.dk/services/ProP/>). Molecular masses, isoelectric points, and other protein parameters were determined with the ExPASy Bioinformatics Resource Portal web server.

Phylogenetic analysis

Molecular phylogenetic analyses were conducted with the maximum likelihood (ML) method in MEGA6 software, based on the Tamura 3-parameter model. A discrete γ distribution was used to model evolutionary rate differences amongst sites (five categories (+G, parameter = 2.2566)). The rate variation model allowed some sites to be evolutionarily invariable. The best model was established from a set of 465 nts. There was a total of 465 positions in the final dataset. The analysis involved 32 nucleotide sequences. The codon positions included were 1st+2nd+3rd+noncoding. Ten thousand replicates were used for the analyses. Bayesian inferences (BIs) of *T. pallidipennis* defensin relationships were performed using MrBayes software. Support node values were obtained from 1000000 generations and are shown next to the branches. The evolutionary distances were computed using the GTR substitution model with γ -distributed rate variation across sites and a proportion of invariable sites. The analysis involved a final dataset of 32 nucleotide sequences and a total of 465 positions (gaps included). The analysis was terminated when the S.D. of split frequencies = 0.0091. A matrix of genetic distances was calculated with MEGA6 software with the data used for the ML analysis.

Structure prediction of mature peptides

The three sequences coding for the mature peptides were sent to the I-TASSER server for tertiary structure prediction [29–31], which is available at <http://zhanglab.ccmb.med.umich.edu/I-TASSER/>. The models were refined by energy minimization using YASARA (15097).

Results

Characteristics of the defensin genes from *T. pallidipennis*

Oligonucleotides that comprised the 5' and 3' UTR regions were generated using the previously reported sequences of *T. brasiliensis* Def1, Def3, and Def4. Using genomic DNA as a template and PCR amplification, we obtained amplicons of 400 bp for each gene. The PCR products were cloned into the pJET plasmid to obtain complete gene sequences.

The 15 sequences of *T. pallidipennis* defensins obtained (five for each gene) were aligned with the defensin sequences reported for *T. brasiliensis* to identify the intron position (Figure 1). From these 15 sequences obtained initially, 12 different genes were identified (GenBank accession numbers: MH000324, MH000325, MH000326, MH000327, MH000328, MH000329, MH000330, MH000331, MH000332, MH000333, MH000334, and MH000335).

The sequences obtained with the Def1 oligos present more variability in both the position and the size of the intron as well as in its composition. The introns begin between positions 88 and 93, while the sizes of the introns oscillate between 95 and 112 bp (Figure 1A). The sequences obtained from the amplicons generated with the Def3 oligos are more homogeneous; 3.1 and 3.4 have the same sequence, and in all the sequences, the intron begins at position 89 and ends at position 186. Therefore, the sequences have a size of 97 bp. However, Def3.5 has 106 bp, since it ends at position 195 (Figure 1B). Finally, the sequences obtained with the Def4 oligos, Def4.1, Def4.2, and Def4.4, are the same; the intron begins at position 61 and ends at nucleotide 167, for a length of 106 bp. In the case of Def4.3 and Def4.5, the intron ends at position 162 and consists of 101 bp (Figure 1C).

Characterization of amino acids of defensin genes from *T. pallidipennis*

The genes were edited *in situ*, and bioinformatics analysis was performed to identify the signal peptide, propeptide, and mature peptide (Figure 2). Of the 12 sequences, 10 had 93 aa, and two had 94 aa, with those in the latter category having an extra Glutamine (Q) at position 36 (Def1.1 and Def1.2). The signal peptide is the same for almost all the sequences [MKCALSLVTLFLVAALAYS], however Def1.1 have Alanina (A) on the 8th position unlike the others sequences that have Valine (V) in this position; nevertheless, the propeptide is the most variable region, since eight

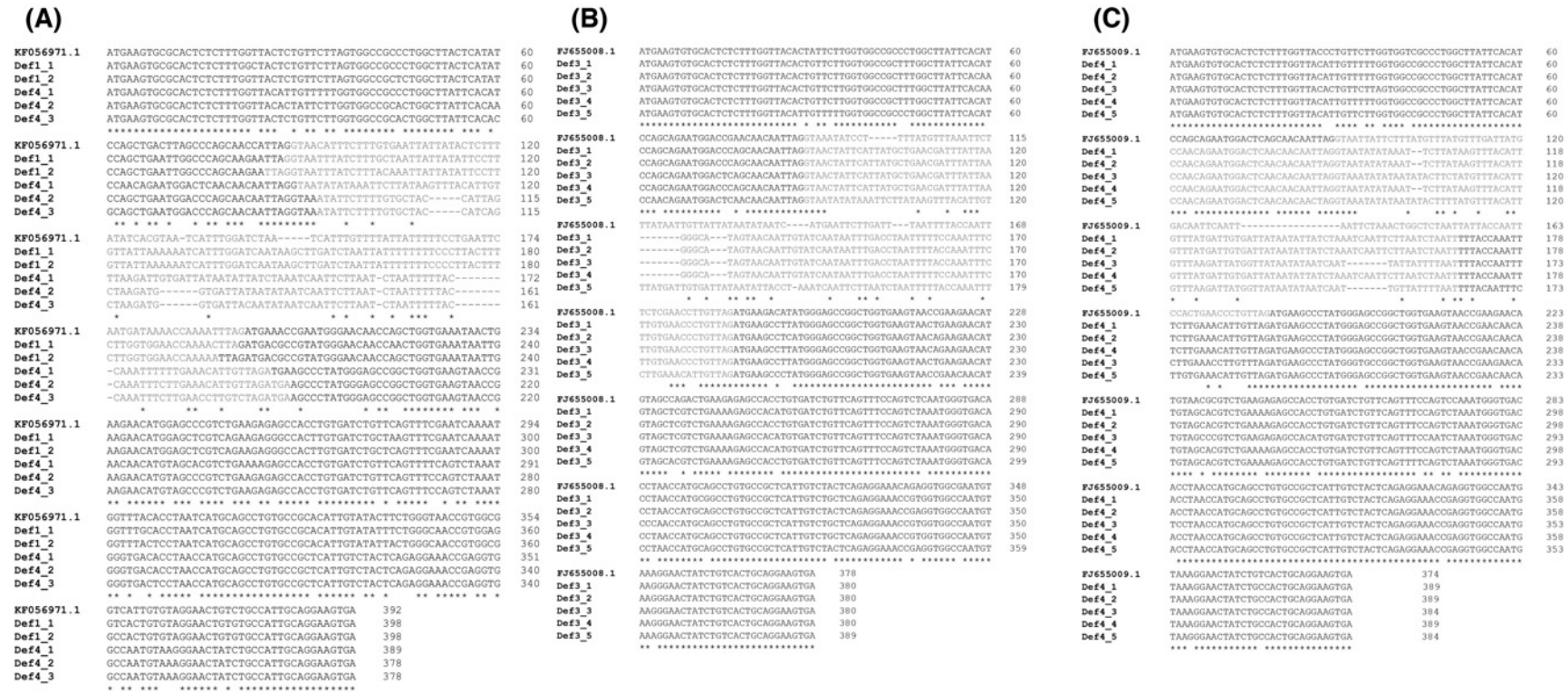


Figure 1. Defensins gene sequences of *Triatoma (Meccus) pallidipennis*

Each alignment consists of five sequences per gen: Def1, Def3, and Def4. The sequences consist of two exons and one intron. Each set of sequences was aligned with Def1 (KF056971.1), Def3 (FJ655008.1), and Def4 (FJ655009.1) from *T. brasiliensis*. **(A)** The set from Def1, varies both in size and position and all the sequences are different. **(B)** From Def3, the position of the intron is the same for all the sequences, also Def3.1 and Def3.4 are identical. **(C)** In the Def4 set, the position of the intron is the same, Def4.1, Def4.2, and Def4.4 are identical. The intron is presented in gray color. Identical nucleotides between the sequences is indicated by asterisks.

		Signal peptide	Propeptide	Mature peptide
Def1.1	(MH000324)	MKCALSLATLFLVAALAYSYPAEALQQELDDAVWEQPTGEIIEEHGARQKRATCDLLSFESKWFAPNHAACAAHCIFLGNRGGHCVGTVCCHCRK		
Def1.2	(MH000325)	MKCALSLVTLFLVAALAYSYPAEALQQELDDAVWEQPAGEIIEEHGARQKRATCDLLSFESKWFPTNHAACAAHCILLGNRGGHCVGTVCCHCRK		
Def4.1	(MH000326)	MKCALSLVTLFLVAALAYSHPTEWTQQQLGEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.2	(MH000327)	MKCALSLVTLFLVAALAYSQPAEWTQQQLGKALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.3	(MH000328)	MKCALSLVTLFLVAALAYSHPTEWTQQQLGKALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.4	(MH000329)	MKCALSLVTLFLVAALAYSHPAEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.5	(MH000330)	MKCALSLVTLFLVAALAYSQPAEWTQQQLDEASWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.6	(MH000331)	MKCALSLVTLFLVAALAYSQPAEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.7	(MH000332)	MKCALSLVTLFLVAALAYSHPTEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.8	(MH000333)	MKCALSLVTLFLVAALAYSHPTEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.9	(MH000334)	MKCALSLVTLFLVAALAYSHPTEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		
Def4.10	(MH000335)	MKCALSLVTLFLVAALAYSHPTEWTQQQLDEALWE-PAGEVTEEHVARLKRATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKGTICHCRK		

Figure 2. Amino acids translated from gene-coding defensins

The 12 different sequences were translated to amino acids, only two correspond to defensin type 1 (Def1.1 and Def1.2), which have an extra aa (Q). The others ten sequences, correspond to defensin type4, in which the propeptide is the more variable region. Only three different mature peptides were obtained. Identical amino acid residues between the sequences are indicated below the alignment sequences by *; (.) indicates conservation between groups of strongly similar properties and (.) indicates conservation between groups of weakly similar properties. The signal peptide, propeptide, and mature peptide are marked.

			-Sheet	α -Helix	-Sheet
			C ₁	C ₂ C ₃	C ₄ C ₅ C ₆
TpDef1.1			ATCDLLSFESKWFAPNHAACAAHCIFLGNRGGHCVG-TVCHCRK		
TpDef1.2			ATCDLLSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
TpDef4			ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
ACU11930.1	<i>T. brasiliensis</i>	def3	ATCDLFSFQSQWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02972.1	<i>T. brasiliensis</i>	def1	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02940.1	<i>R. nasutus</i>	defB	ATCDLLSFESKWFPTNHAACAAHCILLGNRGGQCKG-TVCHCRK		
AAO74625.1	<i>R. prolixus</i>	defB	ATCDLLSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AAO74624.1	<i>R. prolixus</i>	defA	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02963.1	<i>T. sherlocki</i>	def3ba	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
ACU11931.1	<i>T. brasiliensis</i>	def4	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02954.1	<i>T. b. macromelasoma</i>	def4	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02953.1	<i>T. juazeirensis</i>	def4	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02955.1	<i>T. sherlocki</i>	def4	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02956.1	<i>T. melanica</i>	def4	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02939.1	<i>R. nasutus</i>	defA	ATCDLFSFQSKWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02962.1	<i>T. sherlocki</i>	def3aa	ATCDLFSFQSQWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
ABR27891.1	<i>T. infestans</i>	def	ATCDLFSFQSQWVTPNHAACAAHCILLGNRGGQCKG-TICHCRK		
AHY02937.1	<i>P. megistus</i>	defB	ATCDLFSFESQWFTPNHSAACAAHCILLGNRGGQCKG-TICHCRK		
ALJ53398.1	<i>R. nasutus</i>	defC	ATCDLLSFTSKWFTPNHAGCAAHCIFLGNRGGHCVG-TVCHCRK		
AAO74626.1	<i>R. prolixus</i>	defC	ATCDLLSLTSKWFTPNHAGCAAHCIFLGNRGGHCVG-TVCHCRK		
AHY02936.1	<i>P. megistus</i>	defA	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02981.1	<i>T. sordida</i>	def1	ATCDLFSFSSKWFTPNHAACAAHCILLGNRGGHCVG-TVCHCRK		
ABD61004.1	<i>T. infestans</i>	def	ATCDLFSFSSKWFTPNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02979.1	<i>T. sherlocki</i>	def1	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02973.1	<i>T. juazeirensis</i>	def1	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02989.1	<i>T. b. macromelasoma</i>	def1a	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AHY02950.1	<i>T. melanica</i>	def1	ATCDLFSFESKWFPTNHAACAAHCILLGNRGGHCVG-TVCHCRK		
AAT41586.1	<i>A. aegypti</i>	defA1	ATCDLLSG----FGVGDSACAAHCIIARGNRGGYCNSSKKVCVCNRN		

Figure 3. Alignment of predicted *T. pallidipennis* defensins with other Triatominae defensins

Twenty-five defensins' sequences were alignment with TpDef1.1, TpDef1.2, and TpDef4. All the sequences showed the six cysteines in the same position and the sequences have the same size. TpDefs from *T. pallidipennis* are in red bold. Identical and similar amino acid residues between the defensins are indicated below the alignment sequences by asterisks and dots respectively; (.) indicate conservative substitution, (.) indicate less conservative substitution. Cysteine residues forming the disulphide bridges are gray shaded. The sequence of *A. aegypti* is shown for comparison.

different sequences were obtained, and all of them present the cleavage site KR (Lysine-Arginine). Interestingly, from the 12 different sequences, only 3 different mature peptides were found. Two correspond to Def1 (named TpDef1.1 and TpDef1.2), with only one different aa Alanine (A) for Threonine (T). The other ten sequences coded for the same mature *T. pallidipennis* defensin, which was named TpDef4 (Figure 3). An alignment was made using the 3

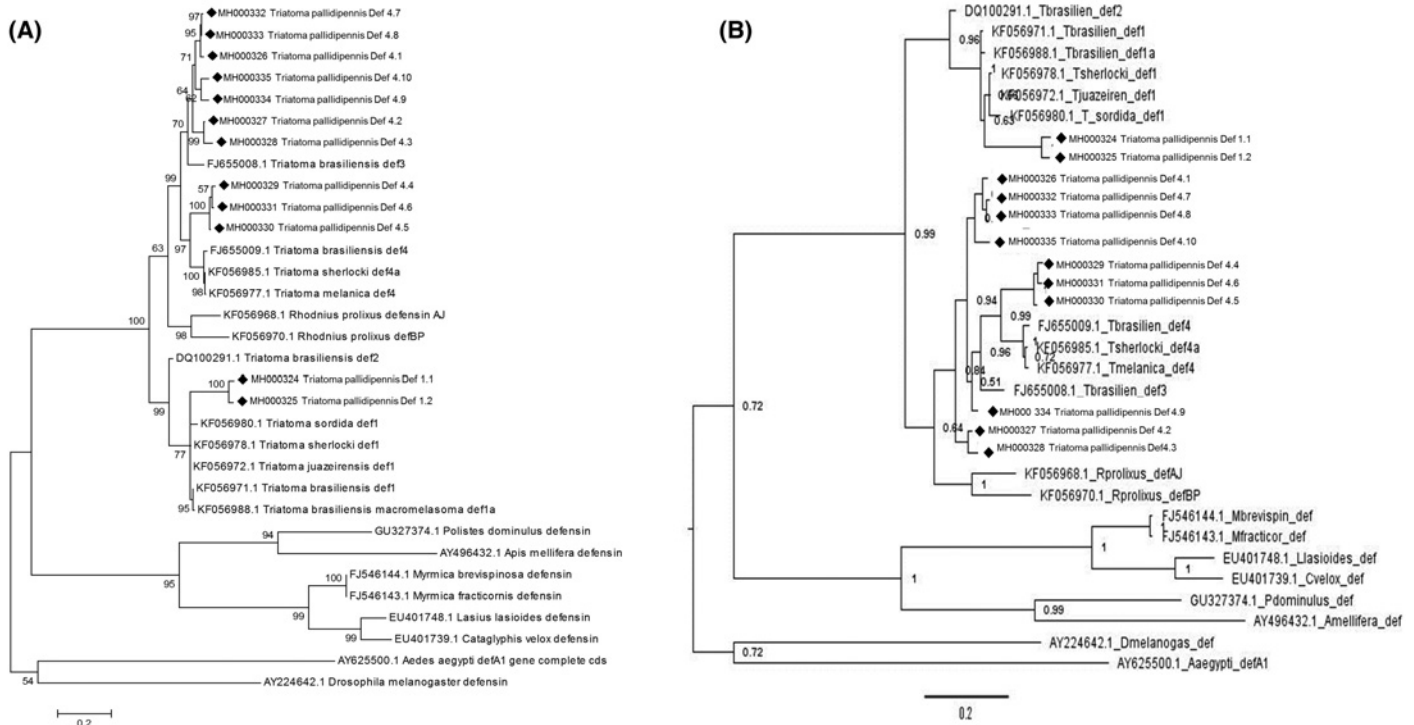


Figure 4. Phylogenetic analysis of defensins

(A) Molecular phylogenetic analysis by ML method. The evolutionary history was inferred by using the ML method based on the Tamura 3-parameter model. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. A discrete γ distribution was used to model evolutionary rate differences amongst sites. The rate variation model allowed for some sites to be evolutionarily invariable. There were a total of 465 positions in the final dataset. Evolutionary analyses were conducted in MEGA6. (B) Bayesian analysis of *T. pallidipennis* defensins. Phylogenetic relationship was established using the MrBayes software. The node values were obtained from 1000000 generations and were showed next to the branches. The evolutionary distances were computed using the GTR substitution model with γ -distributed rate variation across sites and a proportion of invariable sites. The analysis involved 32 nucleotide sequences and a total of 465 positions in the final dataset (gaps including). S.D. deviation of split frequencies = 0.0091. ♦ *T. pallidipennis* defensins.

different mature peptides and 24 reported sequences of defensins in other *Triatoma* species, including one defensin from *Aedes aegypti*. All the sequences have six characteristic cysteines in the same position; furthermore, the size is the same for all the sequences, except for the dipteran defensin.

Phylogenetic analysis

ML and BI phylogenetic analyses were performed to establish the phylogeny of defensin sequences of *T. pallidipennis* with respect to other sequences of South American Triatominae. The trees from the BI and ML analyses were nearly identical in topology, since all the Def1 sequences are grouped in the same clade, and the Def4 sequences are grouped in a different clade that shows subclades and branch lengths at only a few nodes. Both methods showed Triatominae defensins as a monophyletic group with high support. In the gene tree inferred from ML and BI analyses, the TpDef1.1 and TpDef1.2 sequences are situated in the Triatominae defensin 1 clade with strong statistical support. However, in both analyses, defensin 1 sequences of *T. pallidipennis* show clear differences compared with the defensin 1 sequences of the South American Triatominae species, which form a subgroup with strong support. Additionally, the length of the branch shows that defensin 1 sequences of *T. pallidipennis* have accumulated a greater number of changes than the other sequences of the same subgroup (Figure 4).

In contrast, the sequences of defensins 4.1 to 4.10 show diversity. Some of them are more similar to the defensin 4 sequences reported in South American Triatominae species, while others form clearly separated subgroups. However, amongst them, there is similarity, which is evident when comparing their genetic distances (Table 1).

Table 1 Genetic distances between defensins

	Def 1.1	Def 1.2	Def 4.1	Def 4.2	Def 4.3	Def 4.4	Def 4.5	Def 4.6	Def 4.7	Def 4.8	Def 4.9	Def 4.10
Def 1.1		0.01	0.06	0.06	0.06	0.07	0.07	0.07	0.06	0.06	0.06	0.07
Def 1.2	0.03		0.05	0.05	0.05	0.06	0.06	0.06	0.05	0.05	0.06	0.06
Def 4.1	0.37	0.35		0.02	0.02	0.03	0.03	0.03	0.01	0.01	0.02	0.01
Def 4.2	0.37	0.34	0.07		0.01	0.03	0.03	0.03	0.02	0.02	0.02	0.02
Def 4.3	0.38	0.35	0.09	0.03		0.03	0.03	0.03	0.02	0.02	0.02	0.03
Def 4.4	0.45	0.40	0.17	0.18	0.20		0.01	0.01	0.03	0.03	0.03	0.03
Def 4.5	0.46	0.41	0.16	0.16	0.18	0.03		0.01	0.03	0.03	0.03	0.03
Def 4.6	0.45	0.40	0.17	0.17	0.19	0.01	0.01		0.03	0.03	0.03	0.03
Def 4.7	0.38	0.36	0.02	0.08	0.10	0.16	0.15	0.16		0.00	0.02	0.01
Def 4.8	0.38	0.35	0.01	0.08	0.10	0.17	0.16	0.17	0.00		0.02	0.01
Def 4.9	0.41	0.39	0.07	0.10	0.12	0.15	0.15	0.15	0.07	0.06		0.01
Def 4.10	0.45	0.42	0.04	0.11	0.13	0.17	0.16	0.17	0.04	0.04	0.05	

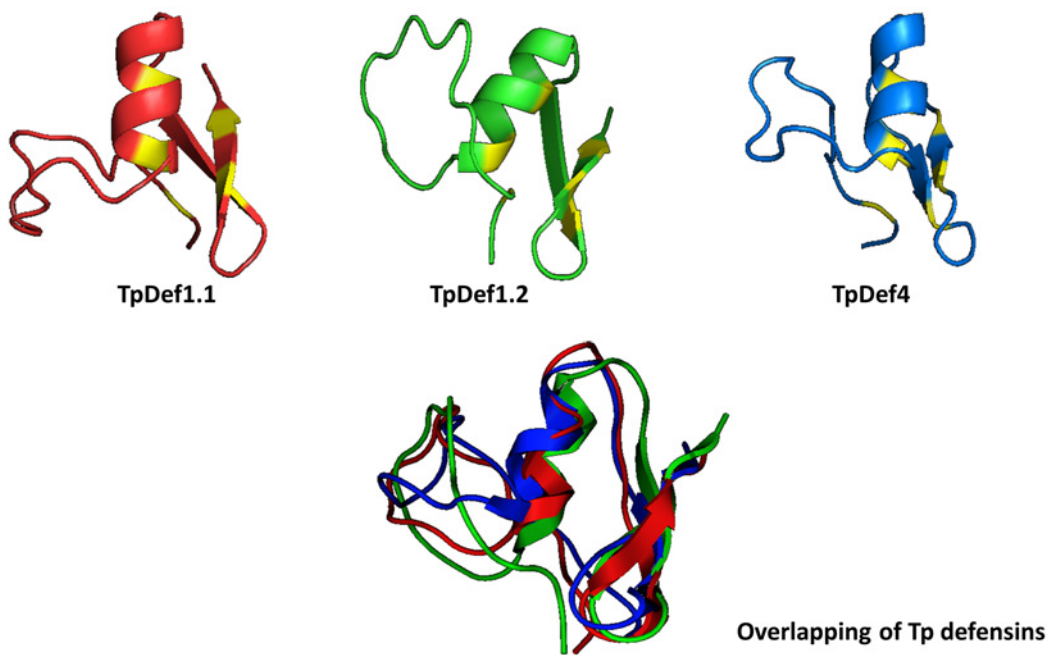


Figure 5. Molecular modeling of Tp defensins

Theoric mature peptides were generated using I-TASSER. All exhibit the same structural topology conformed by three disulphide bonds as well as an α -helix and β -sheets, which are known as cysteine-stabilized ($CS\alpha\beta$) motif. Cysteines are marked in yellow. An overlapping of the three mature peptides is also shown.

Modeling by homology

Molecular modeling of the three theoretical mature peptides was carried out using I-TASSER. All the mature peptides exhibit the same structural topology characterized by three disulphide bonds, an α -helix and β -sheets, which are known as the cysteine-stabilized ($CS\alpha\beta$) motif. TpDef1.1 and TpDef1.2 differ only in two amino acid residues: Alanine replaced by Tryptophan (W), localized in the loop, and Phenylalanine (F) replaced by Leucine (L) in the helix. The TpDef4 sequences have eight different amino acids with respect to TpDef1.1 and TpDef1.2; these changes are located throughout the secondary structure. When overlapping the models of these peptides, it is obvious that the changes in the amino acids do not alter the final conformation of the molecules (Figure 5).

Discussion

A high degree of sequence conservation exists in invertebrate defensins [32,33], which indicates the biological importance of these molecules. In Hemipteran insects, previous studies have demonstrated that the expression of these peptides not only varies in the organ/tissue of origin but also depends on the type of pathogen to which the insect is exposed [34,35].

The immune system in the midgut of insects is one of the most exposed tissues since it is always in contact with a variety of microorganisms. *T. cruzi* is only found in the digestive tract of the Triatominae, so the production of AMPs in the insect gut is therefore vital to protect against infection and to maintain homeostasis of the intestinal microbiota [36,37]. Yet, few studies focus on the expression of the defensin-encoding genes in the presence of *T. cruzi*. One study in which *R. prolixus* was infected with different *T. cruzi* strains showed that the infection with Dm 28c induced an increase in defensin transcripts in contrast with the infection with Y strain [11,37]. It suggests that depending on the genetic lineage of the *T. cruzi* strain the defending response could be different.

With respect to the percentage of infection by *T. cruzi* and defensin expression there are few studies, nevertheless, *T. pallidipennis* seems to have less infection than *T. dimidiata* or *T. barberi* [3,38], indicating a possible role in the diversity of defensins on the control of parasitemia. But more studies need to be done to clearly elucidate this point.

In the present study female insects were used. It will be interesting to analyze male insects in the future, but until now there are few studies that focus in the parasitemia related to the sex of the insects and they do not show a clear correlation between sex and infection mainly because the number of insects was small and difference in the number of females and males infected was not significant [3,39]. In our personal experience both females and males have similar rates of infections (data not published).

We characterized 12 genes that code for defensins in a Mexican endemic *Triatoma*, and our results indicate that the variability of the genes lies mainly in the introns (position, composition, and size). The 12 ORFs presented a signal peptide with 19 amino acid residues, and the largest difference amongst the amino acid sequences is at the propeptide level. The sequences Def1.1 and Def1.2 have 32 amino acid residues in the propeptide region, while the other sequences have 31 amino acids. Additionally, the KR cleavage site at the end of the propeptide region is conserved in all sequences and in other Triatominae defensins [34].

Three mature peptides from the 12 sequenced genes were obtained, named TpDef1.1, TpDef1.2, and TpDef4. Furthermore, the intermediate region in the signal peptide and the mature peptide is highly variable, as has been reported for *R. prolixus* [34]. The specific biological function of the intermediate region is unknown, but it can participate in the cellular trafficking of the immature peptide; this has been reported for the human BMP protein, for which changes in the sequence of this region slow the secretion of this peptide [40].

The results indicate that the mature peptide is highly conserved and similar to all Hemipteran sequences used in the comparative analysis, especially at the position of the six cysteine residues that permit the formation of three disulphide bridges. These data differ from those reported for mammalian defensins, particularly with β -defensins, where a high variability in the antimicrobial domain has been identified and the functional diversity may depend on the structural variation in the β -defensins. This may be due to the diverse and changing microbial environment in their habitats and for a better deal with of a broader range of pathogens [41].

In the phylogenetic analyses using the genomic sequences, a clear separation of the Triatominae defensins from the sequences of other arthropods was observed, as previously reported [42]. In the Triatominae branch of the phylogenetic trees, a clear separation of two groups of defensins was observed: two type-1 sequences and ten type-4 sequences. The type-1 defensin sequences of *T. pallidipennis* are clearly distant from the South American species sequences, a result of a greater accumulation of changes, reflected in a greater length of branches in the phylogenetic tree (Figure 4). Something similar has been observed when comparing other *T. pallidipennis* proteins such as triafestins and triabins, which, despite their similarity in function and structure, are clearly separated from their South American counterparts [43]. However, despite these differences, Def 1.1 and Def 1.2 of *T. pallidipennis* show a close evolutionary relationship as members of the defensin 1 clade. The sequences of defensin type 4 of *T. pallidipennis* are grouped in the phylogenetic trees and separated from the type-1 defensin sequences.

Additionally, the defensins of the *Triatoma* genus are clearly separated from the *Rhodnius* genus sequences, as previously described [26]. Furthermore, three sequences (Def 4.4, Def 4.5, and Def 4.6) in both the ML and Bayesian analyses are closely related to type-4 defensins of *T. brasiliensis*, *T. sherlocki*, and *T. melanica*, suggesting a close evolutionary relationship amongst them. TpDef4 is coded by ten sequences (Def4.1–Def4.10). This result suggests that this defensin has experienced a duplication phenomenon for retroposition or nonallelic homologous recombination between transposable elements [44], indicating their possible relevance, as has been demonstrated with A and J defensins that are secreted in the intestine and fat body in *R. prolixus* in response to bacterial challenge [34].

This phenomenon should be explored further, since it has been shown that some families of human defensins such as DEFB4 have between 2 and 12 copies per genome [45]. Another possibility is that each of these sequences is regulated by different elements in the flanking 5' region, as has been demonstrated for other sequences of AMPs in arthropods such as *Manduca sexta* [46].

Computational modeling shows that, although the three mature peptides have some different amino acids, this does not affect the final folding of the peptide. All the mature peptides presented three disulphide bridges formed by six cysteine residues and three characteristic domains: an N-terminal flexible loop, followed by an α -helix and two C-terminal antiparallel β -sheets.

Recently, using transcriptome analysis, researchers have reported that *T. pallidipennis* expressed mRNA for eight defensin genes. One of these corresponds to Def4.6; other, to Def1.2 [21], the dataset is available at <http://201.131.57.23:8080/data/triatoma>. That report is in accordance with our finding of divergence in defensin expression on *T. pallidipennis*.

Many questions remain to be answered, such as the following: are the peptides reported here functional? Which are the stimuli that induce the secretion of these peptides? Which organs secreted them? What is the meaning of the sequence diversity in the intron region, if there is any? Furthermore, we are currently analyzing the expression of defensin in infected *T. pallidipennis* and in different stages of the insect.

The present work initiated the genetic characterization of this molecule in an important North American vector of Chagas disease, a human malady that causes thousands of deaths. Knowledge about these molecules can help in designing novel strategies for the control of Chagas disease vectors.

Acknowledgements

We thank Iris N. Serratos, Ph.D., for her help in the generation of computational modeling by homology.

Funding

P.D.G. was supported by a Doctoral Fellowship from the Consejo Nacional de Ciencia y Tecnología (663015).

Competing interests

The authors declare that there are no competing interests associated with the manuscript.

Author contribution

P.D.G. performed most of the experiments, participated in the discussion and in the manuscript writing; O.S.R. helped with the gene cloning and sequencing; I.M.M. performed the genetic analysis and help in the writing of results and discussion; B.E. coordinated the project, designed the experiments, discussed the results and participated in the final writing of the manuscript.

Abbreviations

AMP, antimicrobial peptide; BI, Bayesian inference; CS $\alpha\beta$, cysteine-stabilized α -helix β -sheet; ML, maximum likelihood.

References

- WHO (2017), [http://www.who.int/es/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](http://www.who.int/es/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis))
- Ramsey, M.J., Peterson, A.C., Carmona-Castro, O., Moo-Llanes, D., Nakasawa, Y., Butrick, M. et al. (2015) Atlas of Mexican Triatominae (Reduviidae: Hemiptera) and vector transmission of Chagas disease. *Mem. Inst. Oswaldo Cruz.* **110**, 1–14, <https://doi.org/10.1590/0074-02760140404>
- Martínez-Ibarra, J.A., Valencia-Navarro, I., León-Saucedo, S., Ibáñez-Cervantes, G., Bustos-Saldaña, R., Montañez-Valdez, O.D. et al. (2011) Distribution and infection of triatomines (Hemiptera: Reduviidae) by *Trypanosoma cruzi* in the state of Michoacán, México. *Mem. Inst. Oswaldo Cruz.* **106**, 445–450, <https://doi.org/10.1590/S0074-02762011000400010>
- Martínez-Hernández, F., Martínez-Ibarra, J.A., Catalá, S., Villalobos, G., de la Torre, P., Laclette, J.P. et al. (2010) Natural crossbreeding between sympatric species of the *Phyllosoma* Complex (Insecta: Hemiptera: Reduviidae) indicate the existence of only one species with morphologic and genetic variations. *Am. J. Trop. Med. Hyg.* **82**, 74–82, <https://doi.org/10.4269/ajtmh.2010.09-0272>
- Ganz, T. (2003) Defensins: antimicrobial peptides of innate immunity. *Nat. Rev. Immunol.* **3**, 710–720, <https://doi.org/10.1038/nri1180>
- Mussabekova, A., Daeflfer, L. and Imler, J.L. (2017) Innate and intrinsic antiviral immunity in *Drosophila*. *Cell. Mol. Life Sci.* **74**, 2039–2054, <https://doi.org/10.1007/s00018-017-2453-9>
- Boulanger, N., Bulet, P. and Lowenberger, C. (2006) Antimicrobial peptides in the interactions between insects and flagellate parasites. *Trends Parasitol.* **22**, 262–268, <https://doi.org/10.1016/j.pt.2006.04.003>
- Hoffman, J.A. and Reichhart, J.M. (2002) *Drosophila* innate immunity: an evolutionary perspective. *Nat. Immunol.* **3**, 121–126, <https://doi.org/10.1038/ni0202-121>
- Hetru, C., Troxler, L. and Hoffman, J.A. (2003) *Drosophila melanogaster* antimicrobial defense. *J. Infect. Dis.* **187**, S327–S334, <https://doi.org/10.1086/374758>

- 10 Ursic-Bedoya, R.J., Buchhop, J., Joy, J.B., Durvasula, R. and Lowenberger, C. (2011) Prolixicin: a novel antimicrobial peptide isolated from *Rhodnius prolixus* with differential activity against bacteria and *Trypanosoma cruzi*. *Insect Mol. Biol.* **20**, 775–786, <https://doi.org/10.1111/j.1365-2583.2011.01107.x>
- 11 Vieira, C.S., Waniek, P.J., Castro, D.P., Mattos, D.P., Moreira, O.C. and Azambuja, P. (2016) Impact of *Trypanosoma cruzi* on antimicrobial peptide gene expression and activity in the fat body and midgut of *Rhodnius prolixus*. *Parasit. Vectors* **9**, 119, <https://doi.org/10.1186/s13071-016-1398-4>
- 12 Ganz, T. and Lehrer, R. (1995) Defensins. *Pharmacol. Ther.* **66**, 191–205, [https://doi.org/10.1016/0163-7258\(94\)00076-F](https://doi.org/10.1016/0163-7258(94)00076-F)
- 13 Bonmatin, J.M., Bonnat, J.L., Gallet, X., Vovelle, F., Ptak, M., Reichhart, J.M. et al. (1992) Two-dimensional ¹H NMR study of recombinant insect defensin A in water: Resonance assignments, secondary structure and global folding. *J. Biomol. NMR* **2**, 235–256, <https://doi.org/10.1007/BF01875319>
- 14 Bulet, P., Hetru, C., Dimarcq, J.-L. and Hoffman, D. (1999) Antimicrobial peptides in insects; structure and function. *Dev. Comp. Immunol.* **23**, 329–344, [https://doi.org/10.1016/S0145-305X\(99\)00015-4](https://doi.org/10.1016/S0145-305X(99)00015-4)
- 15 Thevissen, K., Warnecke, D., François, I.E., Leipelt, M., Heinz, E., Ott, C. et al. (2004) Defensins from insects and plants interact with fungal glucosylceramides. *J. Biol. Chem.* **279**, 3900–3905, <https://doi.org/10.1074/jbc.M311165200>
- 16 Dassanayake, R.S., Gunawardene, Y.I.N. and Tobe, S.S. (2007) Evolutionary selective trends of insect/mosquito antimicrobial defensin peptides containing cysteine-stabilized α/β ; motifs. *Peptides* **28**, 62–75, <https://doi.org/10.1016/j.peptides.2006.09.022>
- 17 Bulet, P. and Stöcklin, R. (2005) Insect antimicrobial peptides: structures, properties and gene regulation. *Protein Pept. Lett.* **12**, 3–11, <https://doi.org/10.2174/0929866053406011>
- 18 Zhang, L.J. and Gallo, R.L. (2016) Antimicrobial peptides. *Curr. Biol.* **26**, 4–9, <https://doi.org/10.1016/j.cub.2015.11.017>
- 19 Mesquita, R.D., Vionette-Amaral, R.J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F.A., Minx, P. et al. (2016) Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1415–E1416, <https://doi.org/10.1073/pnas.1600205113>
- 20 Hernández-Vargas, M.J., Gil, J., Lozano, L., Pedraza-Escalona, M., Ortiz, E., Encarnación-Guevara, S. et al. (2017) Proteomic and transcriptomic analysis of saliva components from the hematophagous reduviid *Triatoma pallidipennis*. *J. Proteomics* **162**, 30–39, <https://doi.org/10.1016/j.jprot.2017.04.022>
- 21 Zumaya-Estrada, F., Martínez-Barnette, J., Lavore, A., Rivera-Pomar, R. and Rodríguez, M.H. (2018) Comparative genomics analysis of triatomines reveals common first line and inducible immunity-related genes and the absence of Imd canonical components among hemimetabolous arthropods. *Parasit. Vectors*, <https://doi.org/10.1186/s13071-017-2561-2>
- 22 Araújo, C.A.C., Waniek, P.J., Stock, P., Mayer, C., Jansen, A.M. and Schaub, G.A. (2006) Sequence characterization and expression patterns of defensin and lysozyme encoding genes from the gut of the reduviid bug, *Triatoma brasiliensis*. *Insect Biochem. Mol. Biol.* **36**, 547–560, <https://doi.org/10.1016/j.ibmb.2006.04.003>
- 23 Kollien, A.H., Fechner, S., Waniek, P.J. and Schaub, G.A. (2003) Isolation and characterization of a cDNA encoding for a lysozyme from the gut of the reduviid bug *Triatoma infestans*. *Arch. Insect. Biochem. Physiol.* **53**, 134–145, <https://doi.org/10.1002/arch.10090>
- 24 Ribeiro, J.M.C., Genta, F.A., Sorgine, M.H.F., Logullo, R., Mesquita, R.D., Paiva-Silva, G.O. et al. (2014) An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *PLoS Negl. Trop. Dis.* **8**, e2594, <https://doi.org/10.1371/journal.pntd.0002594>
- 25 Waniek, P.J., Castro, H.C., Sathler, P.C., Miceli, L., Jansen, A.M. and Araújo, C.A.C. (2009) Two novel defensin-encoding genes of the Chagas disease vector *Triatoma brasiliensis* (Reduviidae, Triatominae): Gene expression and peptide-structure modeling. *J. Insect Physiol.* **55**, 840–848, <https://doi.org/10.1016/j.jinsphys.2009.05.015>
- 26 Araújo, C.A.C., Lima, A.C., Jansen, A.M., Galvão, C., Jurberg, J., Costa, J. et al. (2015) Genes encoding defensins of important Chagas disease vectors used for phylogenetic studies. *Parasitol. Res.* **114**, 4503–4511, <https://doi.org/10.1007/s00436-015-4694-6>
- 27 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York
- 28 Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890, <https://doi.org/10.1093/nar/16.22.10881>
- 29 Roy, A., Kucukural, A. and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Prot.* **5**, 725–738, <https://doi.org/10.1038/nprot.2010.5>
- 30 Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8, <https://doi.org/10.1038/nmeth.3213>
- 31 Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, <https://doi.org/10.1186/1471-2105-9-40>
- 32 Chalk, R., Albuquerque, C.M.R., Ham, P.J. and Townson, H. (1995) Full sequence and characterization of two insect defensins: immune peptides from the mosquito *Aedes aegypti*. *Proc. Biol. Sci.* **261**, 217–221, <https://doi.org/10.1098/rspb.1995.0139>
- 33 Bulet, P., Stöcklin, R. and Menin, L. (2004) Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev.* **198**, 169–184, <https://doi.org/10.1111/j.0105-2896.2004.0124.x>
- 34 Lopez, L., Morales, G., Ursic, R., Wolff, M. and Lowenberger, C. (2003) Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of Chagas disease. *Insect. Biochem. Mol. Biol.* **33**, 439–447, [https://doi.org/10.1016/S0965-1748\(03\)00008-0](https://doi.org/10.1016/S0965-1748(03)00008-0)
- 35 Waniek, P.J., Jansen, A.M. and Araújo, C.A.C. (2011) *Trypanosoma cruzi* infection modulates the expression of *Triatoma brasiliensis* def1 in the midgut. *Vector Borne Zoonotic Dis.* **11**, 845–847, <https://doi.org/10.1089/vbz.2010.0020>
- 36 García, E., Castro, D., Figueiredo, M. and Azambuja, P. (2010) Immune homeostasis to microorganisms in the guts of triatomines (Reduviidae). *Mem. Inst. Oswaldo Cruz* **105**, 605–610, <https://doi.org/10.1590/S0074-02762010000500001>
- 37 Vieira, C., Waniek, P., Mattos, D., Castro, D., Mello, C., Ratcliffe, N. et al. (2014) Humoral responses in *Rhodnius prolixus*: bacterial feeding induces differential patterns of antibacterial activity and enhances mRNA levels of antimicrobial peptides in the midgut. *Parasit. Vectors* **7**, 232, <https://doi.org/10.1186/1756-3305-7-232>

- 38 Vidal-Acosta, V., Ibáñez-Bernal, S. and Martínez-Campos, C. (2000) *Infección natural de chinches Triatominae con Trypanosoma cruzi asociadas a la vivienda humana en México*, Salud Pública de México [on line], <http://saludpublica.mx/index.php/spm/article/view/6271/7499> (Accessed 17 September 2018)
- 39 Curtis-Robles, R., Auckland, L., Snowden, K., Hamer, G. and Hamer, S. (2018) Analysis of over 1500 triatomine vectors from across the US, predominantly Texas, for *Trypanosoma cruzi* infection and discrete typing units. *Infect. Genet. Evol.* **58**, 171–180, <https://doi.org/10.1016/j.meegid.2017.12.016>
- 40 Daher, R., Kannengiesser, C., Houamel, D., Lefebvre, T., Bardou-Jacquet, E., Ducrot, N. et al. (2016) Heterozygous mutations in BMP6 pro-peptide lead to inappropriate hepcidin synthesis and moderate iron overload in humans. *Gastroenterology* **150**, 672–683, <https://doi.org/10.1053/j.gastro.2015.10.049>
- 41 Tu, J., Li, D., Li, Q., Zhang, L., Zhu, Q., Gaur, M. et al. (2015) Molecular evolutionary analysis of β -defensin peptides in vertebrates. *Evol. Bioinform. Online* **11**, 105–114, <https://doi.org/10.4137/EBO.S25580>
- 42 Ceraul, S.M., Dreher-Lesnick, S.M., Gillespie, J.J., Rahman, M.S. and Azad, A.F. (2007) New tick defensin isoform and antimicrobial gene expression in response to *Rickettsia montanensis* challenge. *Infect. Immun.* **75**, 1973–1983, <https://doi.org/10.1128/IAI.01815-06>
- 43 Hernández-Vargas, M.J., Santibáñez-López, C.E. and Corzo, G. (2016) An insight into the Triabin protein family of American hematophagous Reduviids: functional, structural and phylogenetic analysis. *Toxins (Basel)* **8**, 44, <https://doi.org/10.3390/toxins8020044>
- 44 Meisel, R.P. (2009) Repeat mediated gene duplication in the *Drosophila pseudoobscura* genome. *Gene* **438**, 1–7, <https://doi.org/10.1016/j.gene.2009.02.019>
- 45 Hollox, E.J., Armour, J.A. and Barber, J.C. (2003) Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am. J. Hum. Genet.* **73**, 591–600, <https://doi.org/10.1086/378157>
- 46 Rayaprolu, S., Wang, Y., Kanost, M.R., Hartson, S. and Jiang, H. (2010) Functional analysis of four processing products from multiple precursors encoded by a leucocin-related gene from *Manduca sexta*. *Dev. Comp. Immunol.* **34**, 63847, <https://doi.org/10.1016/j.dci.2010.01.008>