

RESEARCH

Open Access



Comparative genomics analysis of triatomines reveals common first line and inducible immunity-related genes and the absence of Imd canonical components among hemimetabolous arthropods

Federico Alonso Zumaya-Estrada¹, Jesús Martínez-Barnetche¹, Andrés Lavore², Rolando Rivera-Pomar^{2,3} and Mario Henry Rodríguez^{1*}

Abstract

Background: Insects operate complex humoral and cellular immune strategies to fend against invading microorganisms. The majority of these have been characterized in *Drosophila* and other dipterans. Information on hemipterans, including Triatominae vectors of Chagas disease remains incomplete and fractionated.

Results: We identified putative immune-related homologs of three Triatominae vectors of Chagas disease, *Triatoma pallidipennis*, *T. dimidiata* and *T. infestans* (TTTs), using comparative transcriptomics based on established immune response gene references, in conjunction with the predicted proteomes of *Rhodnius prolixus*, *Cimex lectularis* and *Acyrtosiphon pisum* hemimetabolous. We present a compressive description of the humoral and cellular innate immune components of these TTTs and extend the immune information of other related hemipterans. Key homologs of the constitutive and induced immunity genes were identified in all the studied hemipterans.

Conclusions: Our results in the TTTs extend previous observations in other hemipterans lacking several components of the Imd signaling pathway. Comparison with other hexapods, using published data, revealed that the absence of various Imd canonical components is common in several hemimetabolous species.

Keywords: Innate immunity, Hemipterans, Triatomines, Holometabolous, Hemimetabolous, Imd pathway

Background

Arthropods possess complex innate immune mechanisms to fend against viruses, bacteria, fungi and parasites. When invading microorganisms breach the cuticle and epithelial barriers, they confront humoral and cellular components of the innate immune response [1]. Intruders are recognized as non-self by pattern recognition receptors (PRRs), including peptidoglycan-binding proteins (PGRPs) and Gram-negative bacteria-binding proteins (GNBPs) [1–3]. In *Drosophila*, PRRs bind to conserved pathogen-associated molecular patterns (PAMPs) [1].

These molecular interactions initiate the immune signal transduction through three main pathways, Toll, Jak-STAT and Imd. The immune signaling culminates in the translocation into the nucleus of NF- κ B/Rel transcription factors, which activate humoral responses characterized by the synthesis of antimicrobial peptides (AMPs) with broad activity spectrum against bacteria, and fungi [1, 4].

In addition to AMP production, other effector mechanisms are elicited as first line of defense, which includes coagulation, melanization and the production of nitric oxide (NO) and reactive oxygen species (ROS). Clot formation involves Hemolectin and Fondue proteins, which are critical to immobilize bacteria and initiation of wound healing [1]. Melanization is triggered by injury or recognition of microbial ligands through PRRs [5, 6].

* Correspondence: mhenry@insp.mx

¹Centro de Investigación Sobre Enfermedades Infecciosas (CISEI), Instituto Nacional de Salud Pública, Cuernavaca, México

Full list of author information is available at the end of the article



Pro-phenoloxidase (PPO) is a precursor present in the hemolymph and hemocytes, which is activated by proteolytic cascades to phenoloxidase (PO) for *de novo* synthesis of melanin [7]. NO is a highly toxic for a wide variety of pathogens. This is produced by oxidation of L-arginine to L-citrulline by the nitric oxide synthase (Nos) [8, 9]. ROS are produced by conserved nicotinamide adenine dinucleotide phosphate (NADPH) enzymes; dual oxidase (Duox) generates hydrogen peroxide (H₂O₂) and hypochlorous acid, and a member of the NADPH oxidase family (Nox) generates H₂O₂ [10–12]. In this context, antioxidant enzymes such as catalases, glutathione peroxidases (GPx) and thioredoxin peroxidases (TPx), play important roles in cellular homeostasis [13, 14].

Alongside humoral responses, cellular responses are mediated by hemocytes [15, 16]. The main defense against viruses is RNA interference (RNAi) [17]. RNAi is based on Dicer (Dcr) and Argonaute (Ago) proteins. These mediate the production of short RNAs from double-stranded RNA (dsRNA) to guide the degradation of viral RNA by the small interfering RNA (siRNA) pathway [18, 19].

Hemocytes phagocytose bacteria, encapsulate parasites and produce immune toxic compounds for pathogen lysis [15, 16]. Internalization of invaders involves members of the scavenger receptor (SR) family and the IgSF-domain protein Dscam (Down syndrome cell adhesion molecule) [20–22]. In the hemolymph, thioester-containing proteins (TEPs) act as opsonins to mark pathogens for phagocytosis [23], C-type lectins (CTLs), recognize microbial carbohydrate structures [24, 25], and extracellular serine proteases, including CLIP-domain serine proteases (CLIPs) are activated upon recognition of aberrant tissues or microbial compounds [26].

Most information of insect immune responses has been described in *Drosophila melanogaster*, the main model organism in Diptera [1, 27, 28], but knowledge of the immune responses in hemipterans including triatomine bugs is limited [29–31]. The Toll and Jak-STAT signaling components, described in *Drosophila* have also been identified in *Rhodnius prolixus* (RPRO) (Hemiptera: Reduviidae) [32, 33], *Acyrtosiphon pisum* (ACPI) (Hemiptera: Aphididae) [30] and *Cimex lectularis* (CLEC) (Hemiptera: Cimicidae) [31]. However, Imd signaling components, highly conserved in several insect orders [34–42], appear to be absent in RPRO and ACPI [30, 33], and information on the other components of the innate immune response in Triatomine bugs remains fractionated and incomplete [29].

In this study, we used a transcriptomic analysis to describe innate immune response genes of *Triatoma infestans* (TINF), the major vector of Chagas disease in sub-Amazonian endemic regions; *T. dimidiata* (TDIM) a vector in northern South America and Central America, extending into Mexico [43]; and *T. pallidipennis* (TPAL),

an important vector in Mexico [44]. An extended comparative analysis of immune genes of TPAL, TDIM, and TINF (TTTs), and other hemipterans (RPRO, CLEC and ACPI) along with those of other holometabolous and hemimetabolous arthropods revealed the lack of several components of the Imd pathway in the hemimetabolous group.

Methods

Insect rearing

Colonies of *T. pallidipennis* (TPAL) (colony 0230 from Mexico), *T. dimidiata* (TDIM) (colony 0252 from Tegucigalpa, Honduras) and *T. infestans* (TINF) (colony X32 from Santiago del Estero, Argentina) established in the Centro Nacional de Chagas, Córdoba, Argentina were reared in the Centro Regional de Estudios Genómicos (CREG), Universidad Nacional de La Plata (UNLP) and the Centro de Bioinvestigaciones, Universidad Nacional del Noroeste de Buenos Aires (UNNOBA). Insects were reared at 28 °C and a partial humidity of 70% with a 12 h light/dark schedule. Insects were regularly fed using artificial feeders and chicken blood. Insect handling was performed in accordance to the World Health Organization protocol [45].

Transcriptome preparation and sequencing

To maximize the coverage of the gene content for each species, total RNA was isolated from embryos (55) and diverse organs (reproductive and digestive tract, Malpighian tubules, brain, fat body and salivary glands) of fed and starved insects of the five nymphal stages (N1 = 8; N2 = 8; N3 = 8; N4 = 4; N5 = 4), adult mated females (4) and adult males (4) of TPAL, TDIM and TINF using Trizol (Life Technologies, Massachusetts, USA). A pool was made with 2 µg of each total RNA extraction (embryos + insect organs).

A single cDNA library for each species was independently constructed using 1.5 µg of each RNA pool (embryos + insect tissues) using the Mint-2 Kit (Evrogen, Moscow, Russia) according to the manufacturer instructions. To reduce redundancy due to highly expressed transcripts and to increase the representation of poorly represented transcripts, each library was normalized using the Trimer-2 Normalization Kit (Evrogen, Moscow, Russia) according to the manufacturer instructions. The cDNA libraries were barcoded and subjected to the shotgun sequencing protocol using the GS FLX+ (454-Roche, Connecticut, USA). Raw sequence datasets are available at the Sequence Read Archive (SRA) - NCBI: TPAL (SRX2600752), TDIM (SRX2600753) and TINF (SRX2600754).

Data filtering, trimming and assembly

Raw reads from each barcoded library were analyzed with PRINSEQ [46] and filtered according to length,

sequence complexity and quality. Each library was subjected to *de novo* assembly with the GS DeNovo assembler v.2.8 software in cDNA mode using the default parameters, and including the adaptor sequences for trimming. The assembled sequences dataset are available at the NCBI-TSA (GFMK00000000, GFMC00000000 and GFMJ01000000). The non-assembled reads were mapped to the RPRO genome (Rhodnius-prolixus-CDC_SCAF-FOLDS_RproC3.fa) and proteome (Rhodnius-prolixus-CDC_PEPTIDES_RproC3.2.fa) using BLAST (Basic Local Alignment Search Tool) [47] algorithms (BLASTn and BLASTx, respectively). Non-redundant mapped reads to either database were included as singletons into the assembled dataset (full_dataset). A non-redundant database (nr_dataset) was built discarding alternative isotigs belonging to the same isogroup or unigene, by keeping the largest isotig (transcript) per isogroup. The dataset used in this work is available at <http://201.131.57.23:8080/data/triatoma>.

Transcriptome completeness analysis

The assembled dataset for each species was used to identify the proportion of the core eukaryotic genome coverage. We used HMM profiles for 458 core eukaryotic proteins as provided by the Core Eukaryotic Genome Dataset (CEGMA) [48] and HMMER searches with the *hmmsearch* command and the -T 40 and -domT 40 filters, as described in [49]. Following the same approach, a Benchmarking Universal Single-Copy Orthologs (BUSCO) sets for arthropod [50] was used to assess transcriptome datasets completeness.

Comparative genomics datasets

We included three hemipteran species whole-genomes for comparison: the complete predicted-peptide sets of RPRO (Rhodnius-prolixus-CDC_PEPTIDES_RproC3.1.fa) and CLEC (Cimex-lectularius-Harlan_PEPTIDES_ClecH 1.2.fa) available in VectorBase [51] and the predicted-peptide set of ACPI (aphidbase_2.1b_pep.fasta) from AphidBase [52].

A whole set of known immune-related gene sequences (“immunity-genes reference dataset”) of *D. melanogaster*, *Aedes aegypti*, *Anopheles gambiae* and other insects, including triatomines were retrieved from ImmunoDB [36], IID [53], GenBank [54], VectorBase [51], UniProt [55], FlyBase [56] and Ensembl [57] databases. In addition, high-confidence immune-related orthologs of *Bombyx mori*, *Tribolium castaneum* and *Apis mellifera* were retrieved from published literature [35, 38, 40] (Additional file 1: Table S1).

Immune-related homologs search

The “immunity-genes reference dataset” was used as queries to perform BLAST searches against TTTs transcriptomes

and RPRO, CLEC and ACPI predicted-peptide sets. For this purpose, we used multiple BLAST algorithms (tBLASTn, BLASTn, BLASTp) using a cut-off e-value of $1.0e^{-5}$. BLAST outputs were retrieved, listed and compiled in the order of descending sequence identity percentage and score, and ascending e-value. Additionally, BLAST-hits with considerably short alignment lengths compared with the genes of the “immunity-genes reference dataset” were filtered. Then, the best 10 BLAST-hits were selected for detection of conserved protein-domain structures.

To search for immune-related genes of the “immunity-genes reference dataset” that produced no hits through BLAST inquiries, we conducted a tBLASTn search against all contigs and unassembled reads of the TTTs transcriptomes. Further, we performed HMM profile-based searches for those unidentified immune-related genes in TTTs and the other hemipterans (RPRO, CLEC and ACPI). We generated amino acid alignments of the unidentified immune-related genes with MUSCLE [58]. Hidden Markov models of these alignments were built using HMMER [59]. These HMM profiles were used to perform searches (*hmmsearch*) against the six-frame translated sequences of the TTTs transcriptomes and the hemipteran predicted-peptide sets.

To detect conserved protein-domain structures, the immune-related sequences of TTTs and other hemipterans identified were analyzed using InterProScan [60]. The domain signatures recognized were visually inspected and compared against the genes of the “immunity-genes reference dataset” to corroborate their architecture similarities.

The immune-related homologs identified were categorized into the following major immune categories: microbial recognition and activation (GNBPs, PGRPs, CTLs, TEPs, SRs and CLIPs), signaling (Toll, Jak-STAT and Imd signaling pathways), effectors (AMPs, melanization, NO and ROS), regulation (Toll, Jak-STAT and Imd signaling regulators), antioxidant system (catalases and peroxidases), RNA interference (RNA interference machinery) and coagulation (Additional file 1: Tables S2-S8).

Phylogenetic analysis

The amino acid sequences of defensin and lysozyme homologs identified were aligned separately using MUSCLE [58]. To compare the obtained results with previously described triatomine defensins, the sequences of defensin A (GenBank: AY196130), defensin B (GenBank: AY196131) and defensin C (GenBank: AY196132) of RPRO [61], *T. brasiliensis* (TBRA) defensin 1 (GenBank: AAV48636), defensin 2 (GenBank: ABA10770), defensin 3 (GenBank: ACH57150), defensin 4 (GenBank: ACH57151) [62, 63], and defensin (GenBank: ABD61004) of TINF were included in the alignment. Two outgroup sequences were added to the defensin dataset: defensin 1 of *A. gambiae* (VectorBase ID: AGAP01129) and defensin C of *A. aegypti*

(VectorBase ID: AAEL003832). In the same way, the sequences of lysozyme A (GenBank: ABX11553), lysozyme B (GenBank: ABX11554) of RPRO [64], TBRA lysozyme 1 (GenBank: AAU04569) [62], and lysozyme 1 (GenBank: AAP83129) and lysozyme 2 (GenBank: ABI94387) [65, 66] of TINF were included in the alignment. Two outgroup sequences were added to the lysozyme dataset: lysozyme 2 of *A. gambiae* (VectorBase ID: AGAP007343) and lysozyme A of *A. aegypti* (VectorBase ID: AAEL003723). Maximum likelihood analyses were carried out using PhyML v3.0 [67] running 1000 bootstrap samples. The output trees were visualized and optimized with FigTree v1.4.3 [68].

Results

Transcriptome datasets metrics

Between 112 and 202 Mbp of filtered raw sequence data were generated for each triatomine species. *De novo* assembly yielded 31,175; 29,024; and 35,629 transcripts for TPAL, TDIM and TINF, respectively. Selection of the largest transcript for each isogroup (unigene) yielded non-redundant datasets with 29,789; 27,652; and 34,646 transcripts for TPAL, TDIM and TINF, respectively. BLASTx matches were identified for 71–74% of RPRO predicted proteome. Transcriptome completeness assessment for the TTTs transcriptomes indicated coverages of 86.9% in TPAL, 85.8% in TDIM and 82.2% in TINF with BUSCO [50], and higher than 90% for CEGMA dataset [48]. The datasets supporting the results of this article are available at <http://201.131.57.23:8080/data/triatoma>.

Microbial recognition

Peptidoglycan and gram-negative bacteria-binding proteins

We identified PGRP-like homologs containing both peptidoglycan recognition protein (IPR015510) and N-acetylmuramoyl-L-alanine amidase (IPR002502) domains in TTTs, and the presence of three PGRP with identical protein signatures in RPRO, which were previously identified [33] (Fig. 1, Additional file 1: Table S2). We also recognized a PGRP-like gene (CLEC005283) in CLEC (PGRP-LF), which was previously related to functions other than microbial recognition [31]. As previously reported, we were unable to detect PGRP-like genes in ACPI [30].

GNBP-like homologs with beta-1,3-glucan-binding N-terminal (IPR031756) and/or concanavalin A-like lectin/glucanase (IPR013320) domains were observed in all hemipterans, except TDIM (Fig. 1, Additional file 1: Table S2).

Lectins

Several CTLs homologs containing C-type lectin fold (IPR016187) domains were identified in all hemipterans (Fig. 1, Additional file 1: Table S2).

Thioester-containing proteins

Notably, we could not identify TEPs in TTTs; however, one TEP encoding gene containing immunoglobulin E-set (IPR001599) and α 2-macroglobulin thiol-ester bond-forming (IPR019565) domains were detected in RPRO and at least two different TEP homologs with similar protein signatures were observed in CLEC and ACPI (Fig. 1, Additional file 1: Table S2).

Cellular receptors

We detected homologs containing a *Dscam*-domain (IPR033027) in RPRO, CLEC and ACPI. Conversely, sequences lacking this domain but containing an immunoglobulin-like fold (IPR013783) and a fibronectin type III (IPR003961) domains were identified in TPAL and TINF.

Scavenger receptors class B (SR-B) homologs, characterized by a CD36 (IPR002159) domain were the most abundant SR class in all hemipterans, except TINF. The SR-B croquemort receptor was not detected in any species, except ACPI (Fig. 1, Additional file 1: Table S2). Scavenger receptors class C (SR-C) encoding genes were identified exclusively in CLEC. However, only one of them (CLEC000453) contains the characteristic extracellular sushi/SCR/CCP (IPR000436) domain (Fig. 1, Additional file 1: Table S2).

CLIP-serine proteases

We identified sequences containing serine protease (IPR001254) and peptidase S1 (IPR009003) domains in all hemipterans. However, homolog genes with a proteinase regulatory CLIP (IPR022700) domain were observed only in RPRO (RPRC003090; RPRC009383) and CLEC (CLEC010998; CLEC006847; CLEC001617) (Fig. 1, Additional file 1: Table S2).

Signaling

Toll signaling pathway

Previously, several Toll signaling proteins were reported in RPRO [32, 33] and we recognized most of the canonical components of the Toll signaling pathway in triatomine bugs (TTTs and RPRO). Of these, we detected homologs of the extracellular cytokine spätzle, the Toll receptor, the death-domain containing adaptor proteins Myd88 and Tube, and the NF- κ B/Rel transcription factor Dorsal (Fig. 2, Additional file 1: Table S3). We also corroborated that canonical components of this signal cascade are conserved in CLEC and ACPI [30, 31] (Fig. 2, Additional file 1: Table S3). In addition, we identified homologs of Cactus in most hemipterans; necrotic homologs were identified only in CLEC and ACPI but wntD was not found in any hemipteran. These molecules are responsible for the negative regulation of the Toll pathway in the absence of pathogenic challenges (Fig. 1, Additional file 1: Table S5).

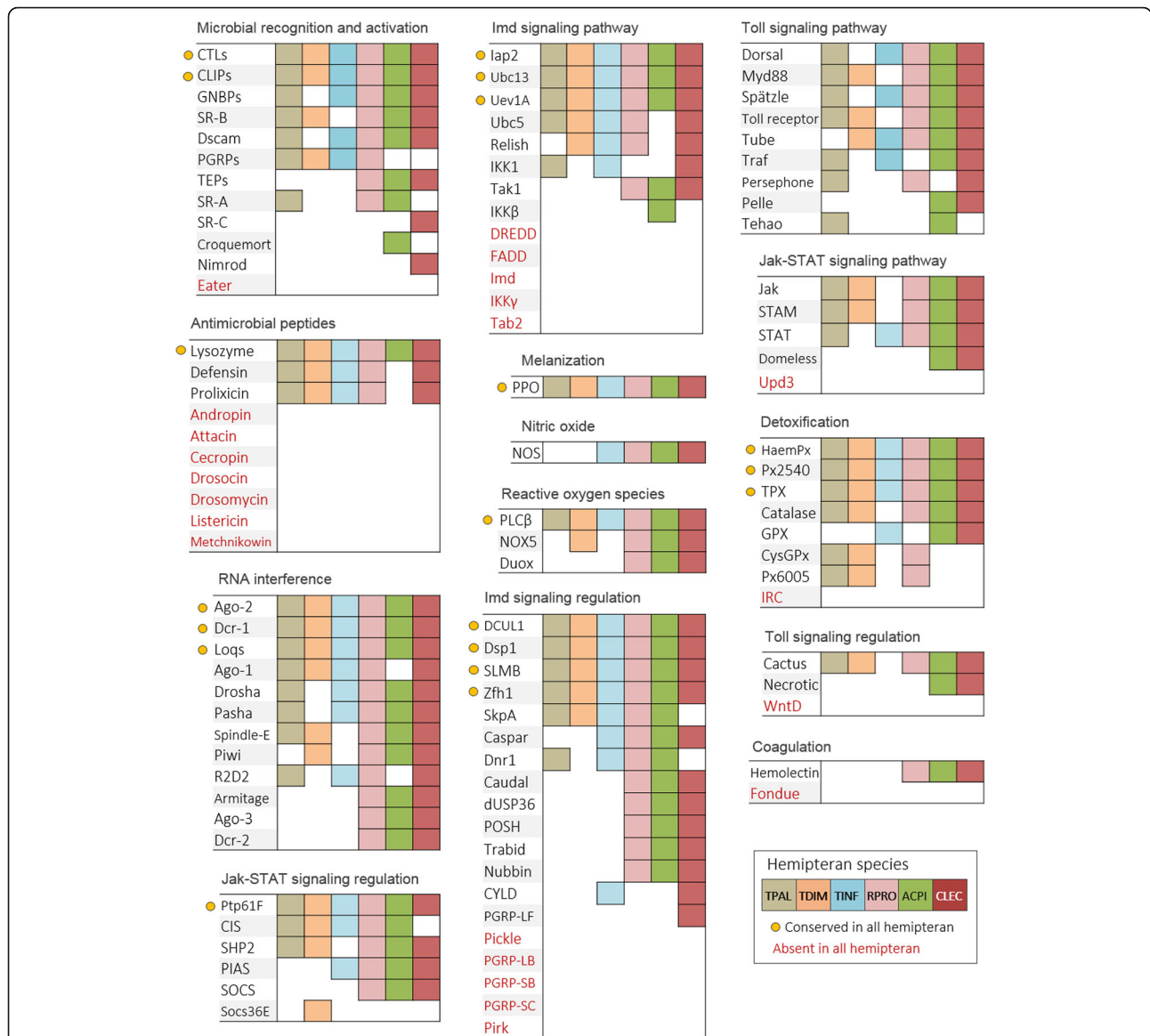


Fig. 1 Presence/absence of immune homologs in hemipterans. Homolog tables presented according to the immune categories in which they participate. Homologs are presented in each first column (corresponding names could be seen in the abbreviations list). The presence of homologs in each species is represented by colored squares in subsequent columns. Brown squares represent *T. pallidipennis* (TPAL), orange *T. dimidiata* (TDIM), blue *T. infestans* (TINF), pink *R. prolixus* (RPRO), green *A. pisum* (ACPI) and red *C. lectularis* (CLEC). The absence of homologs corresponds to blank squares/spaces. The conserved immune-related homologs in all the hemipterans are indicated with yellow circles. Names in red font correspond to immune-related homologs absent in all hemipterans

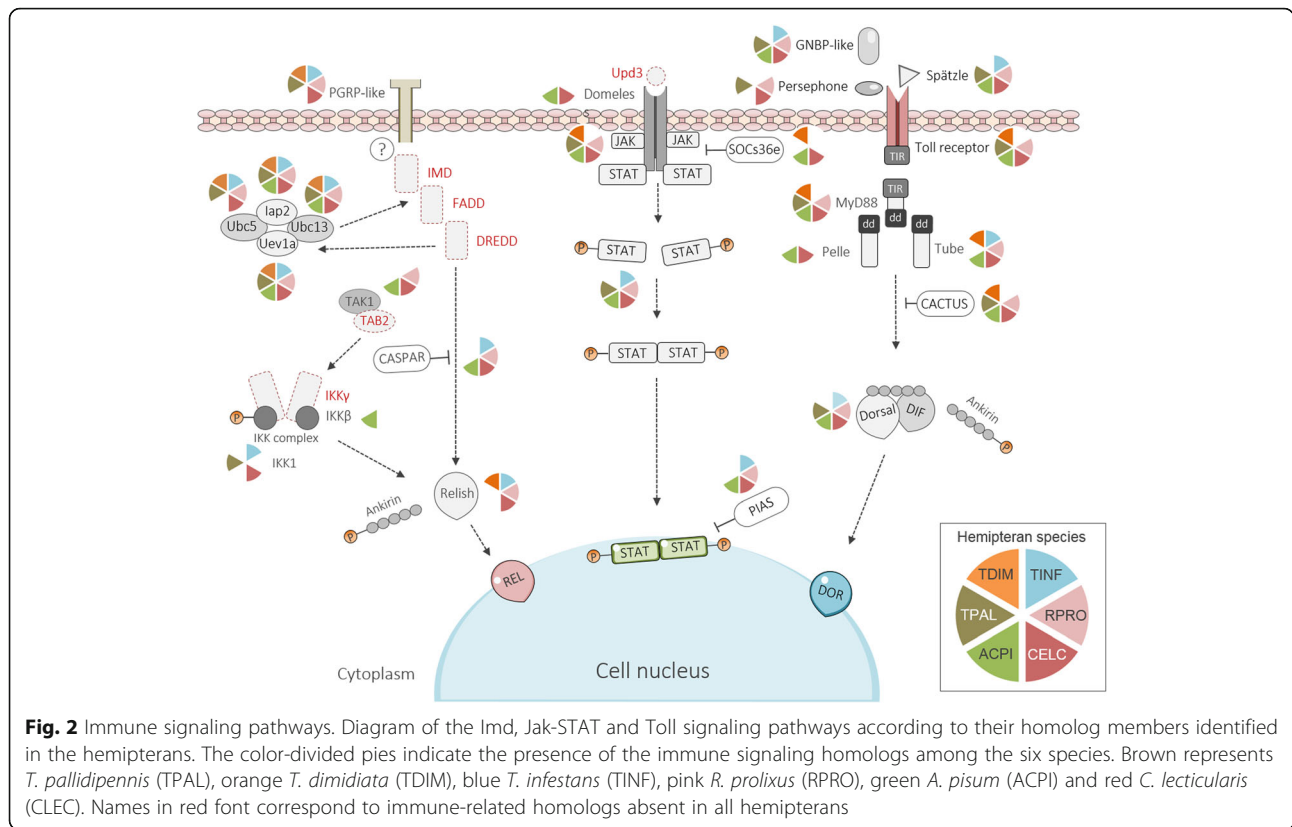
Jak-STAT signaling pathway

Domeless receptor homologs with both fibronectin type III and immunoglobulin-like fold domains were detected only in CLEC and ACPI (Fig. 2, Additional file 1: Table S3). We did not detect homologs of the hemocyte-released cytokine Unpaired-3 (Upd-3) in any of the hemipterans. However, homologs of Janus kinase (Jak), containing a serine-threonine/tyrosine-protein kinase catalytic domain (IPR001245) and homologs of STAT with transcription factor STAT domain (IPR001217) were observed in all hemipterans, except TINF and TDIM, respectively.

We also identified several homologs of negative regulators, as SOCS36E, PIAS and Ptp61F. Of these, homologs of Ptp61F with a PTPase (IPR000242) domain were present in all hemipterans (Fig. 1, Additional file 1: Table S5).

Imd signaling pathway

We were unable to detect key components of the Imd pathway in all hemipteran species. Of these, as described for ACPI, RPRO and CLEC, the death-domain protein Imd, FADD (Fas-Associated protein with death-domain) and the caspase DREDD (death-related



ced-3/Nedd2-like protein) were not detected in any TTT species. However, homologs the three E2 ubiquitin conjugating enzymes: Uev1a, Ubc13 (Bendless), Ubc5 (Effete) and the E3 ubiquitin ligase Inhibitor of Apoptosis Protein (IAP2) were present in all hemipterans, except for Effete in ACPI (Fig. 2, Additional file 1: Table S3). Homologs of the NF-κB/Rel transcription factor Relish, with Rel homology, DNA-binding domain (IPR011539) were observed in CLEC and all triatomine bugs, except TPAL. As previously reported, we were unable to detect Relish genes in ACPI [30] (Fig. 2, Additional file 1: Table S3).

We detected several negative regulators of the Imd signaling cascade. Homologs of Caspar with UBX (IPR001012) and UAS (IPR006577) domains were identified in all hemipterans, except TPAL and TDIM (Fig. 1, Additional file 1: Table S5). Homologs that alter the stability of the Imd pathway members such as dUSP36 and POSH were observed in ACPI, CLEC and RPRO, but not in TTTs. In addition, all components of the SKPA/SLMB/DSL1 complex were identified in all hemipterans, except for SKPA in CLEC (Fig. 1, Additional file 1: Table S5).

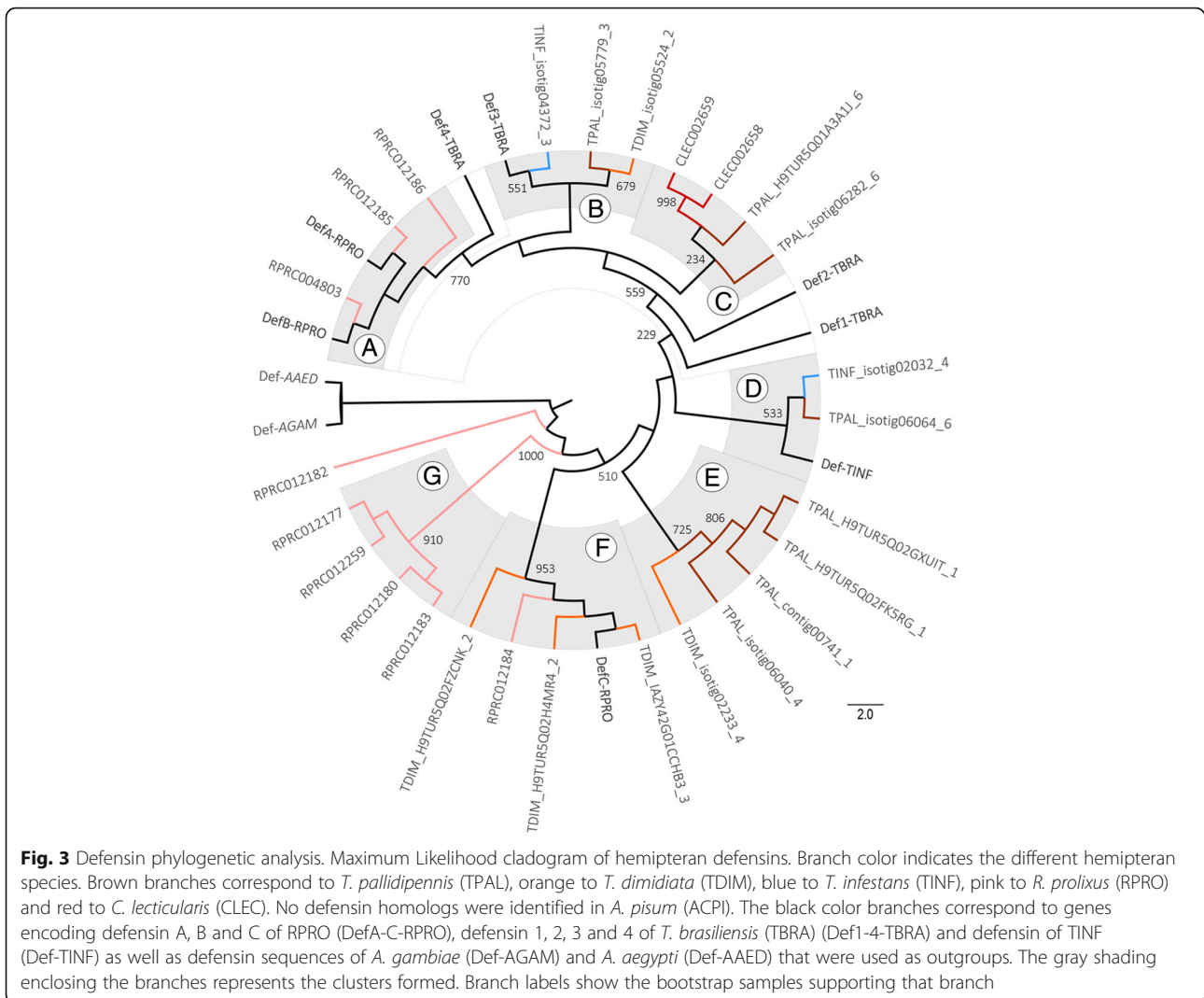
Homologs of the transcriptional repressor Zfh1 (Zn finger homeodomain 1) with zinc finger C2H2-type (IPR013087) and homeobox (IPR001356) domains were identified in all hemipterans. Similarly, homologs of the caspase inhibitor

Dnr1 (Defense repressor 1) were not identified only in TDIM and CLEC (Fig. 1, Additional file 1: Table S5). While homologs of the transcription factor caudal with a helix-turn-helix (IPR000047) motif and multiple homeobox (IPR001356; IPR020479; IPR009057; IPR017970) domains were observed in all hemipterans, but not in TTTs (Fig. 1, Additional file 1: Table S5). Likewise, homologs of the deubiquitinase Trabid [69] with zinc finger, RanBP2-type (IPR001876) and OTU (IPR003323) domains were present in RPRO and ACPI. A Trabid homolog sequence containing an OTU domain, but not zinc finger motifs, was detected in CLEC. No Trabid homologs were detected in TTTs. While homologs of Pickle, a *Drosophila* negative regulator that selectively inhibit Relish [70] were not detected in any hemipteran (Fig. 1, Additional file 1: Table S5).

Immune effectors

Antimicrobial peptides

A number of defensin homologs, containing both knottin scorpion toxin-like (IPR003614) and defensin invertebrate/fungal (IPR001542) domains, were identified in all hemipterans, except ACPI (Fig. 1, Additional file 1: Table S4). ACPI genome apparently has no genes encoding for defensins [30]. The maximum likelihood tree, based on defensin homolog sequences, showed the presence of seven clades (A-G) (Fig. 3). In clade A, genes encoding defensin A and B of RPRO (DefA-RPRO,



DefB-RPRO) were grouped with three sequences of RPRO (RPRC004803, RPRC012185 and RPRC012186) with high bootstrap support (770). A gene encoding defensin 4 of TBRA (Def4-TBRA) was located at the basal position of this clade. In clade B, a gene encoding defensin 3 of TBRA (Def3-TBRA) was grouped with defensin sequences of TINF (TINF_isotig04372_3), TPAL (TPAL_isotig05779_3) and TDIM (TDIM_isotig05524_2) with medium bootstrap support values (551–679). In clade C, both defensin homolog sequences of CLEC (CLEC002659 and CLEC002658) were grouped with two TPAL sequences (TPAL_H9TUR5Q01A3A1J_6 and TPAL_isotig06282_6) with variable bootstrap values (234–998). Two genes encoding defensins of TBRA (Def1-TBRA and Def2-TBRA) were located at the basal position of the A, B and C clades, with variable bootstrap values (229–559). In clade D, a defensin from the gut of TINF (Def-TINF) was grouped with a TINF (TINF_isotig02032_4) and TPAL (TPAL_isotig06064_6) sequences with medium bootstrap

support (533). Clade E grouped four TPAL sequences (TPAL_H9TUR5Q02GXUIT_1, TPAL_H9TUR5Q02FK5RG_1, TPAL_contig00741_1, and TPAL_isotig06040_4) and a sequence of TDIM (TDIM_isotig02233_4) with a bootstrap support value of 725. Clade F grouped a gene encoding defensin C of RPRO (DefC-RPRO) with another sequence of RPRO (RPRC012184), as well as three TDIM sequences (TDIM_IAZY42G01CCHB3_3, TDIM_H9TUR5Q02H4MR4_2 and TDIM_H9TUR5Q02FZCNK_2) with high bootstrap support (953). In clade G, four sequences of RPRO (RPRC012183, RPRC012180, RPRC012259 and RPRC012177) grouped separately from other defensin sequences of RPRO and were supported by high bootstrap value (1000). A RPRO sequence (RPRC012182) was located separate from the other defensin clades. Dipteran defensin 1 of *A. gambiae* (Def-AGAM) and defensin C of *A. aegypti* (Def-AAED) were clustered in an outgroup branch.

Lysozyme homologs, containing Lysozyme-like (IPR023346) domains, were identified in all hemipterans (Fig. 1, Additional file 1: Table S4). The maximum likelihood tree, based on lysozyme homolog sequences, showed the presence of seven clades (A-G) (Fig. 4). In clade A, all lysozyme homolog sequences of CLEC (CLEC009914, CLEC003818, CLEC013272 and CLEC003819) were grouped with low bootstrap support (286). Clade B, grouped a gene encoding lysozyme A of RPRO (LysA-RPRO) with one sequence each of RPRO (RPRC015441) and TDIM (TDIM_H9TUR5Q01CQN5O_5) with high bootstrap support (1000). A sequence of TDIM (TDIM_isotig05675_5) was located at the basal position of this clade. In clade C, a gene encoding lysozyme 1 of TINF (Lys1-TINF) and a sequence of TINF (TINF_isotig04514_1) were grouped with medium bootstrap support (451). A gene encoding lysozyme 1 of TBRA (Lys1-TBRA) was located at the basal position of the B and C clades, with medium bootstrap support value

(617). In clade D, a gene encoding lysozyme 2 of TINF (Lys2-TINF) was grouped with two sequences of TINF (TINF_isotig04526_5 and TINF_isotig04520_6), with high bootstrap support values (965–992). Clade E clustered sequences of TPAL (TPAL_isotig05649_2 and TPAL_isotig05202_3) and TDIM (TDIM_isotig05522_6, TDIM_isotig04659_6) with high bootstrap support (960). A RPRO (RPRC015442) and TINF (TINF_isotig01507_3) sequences were located at the basal position of this clade, with medium bootstrap support (481–673). All lysozyme homolog sequences of ACPI (ACYPI008509, ACYPI009125 and ACYPI002175) were grouped with variable bootstrap values (376–999) in clade F. Clade G clustered a gene encoding lysozyme B of RPRO (LysB-RPRO) and RPRO (RPRC015440) and TPAL (TPAL_isotig05093_2) sequences, with variable bootstrap support (494–997). Dipteran lysozyme 2 of *A. gambiae* (Lys-AGAM) and lysozyme A of *A. aegypti* (Lys-AAED) were clustered in an outgroup branch.

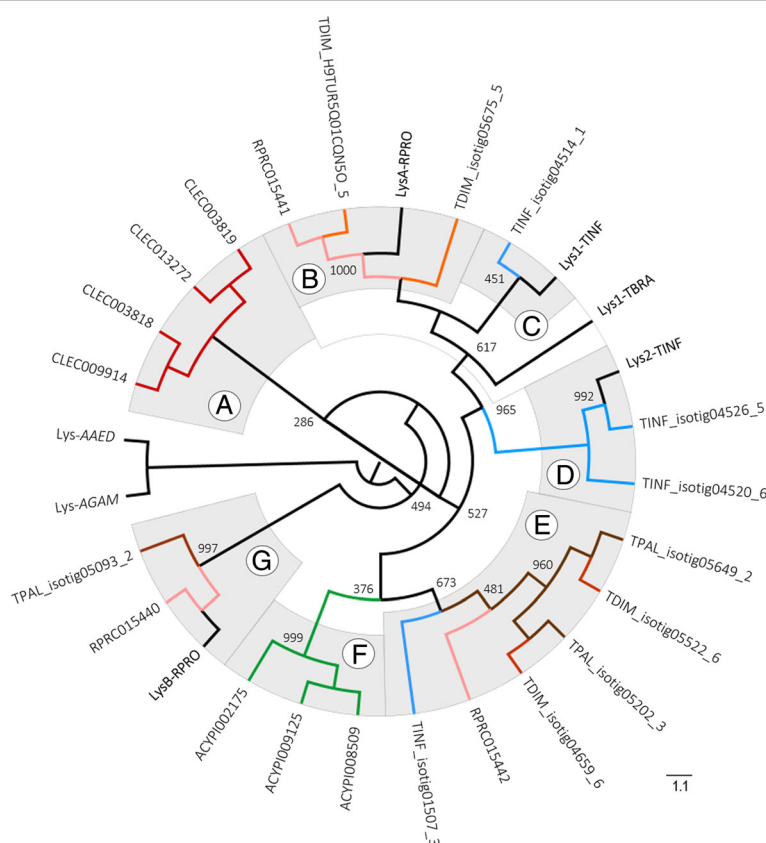


Fig. 4 Lysozyme phylogenetic analysis. Maximum Likelihood cladogram of hemipteran lysozymes. Branch colors indicate the different hemipteran species. Brown branches correspond to *T. pallidipennis* (TPAL), orange to *T. dimidiata* (TDIM), blue to *T. infestans* (TINF), pink to *R. prolixus* (RPRO), red to *C. lecticularis* (CLEC) and green to *A. pisum* (ACPI). The black branches correspond to genes encoding lysozyme A and B of RPRO (LysA, B-RPRO), lysozyme 1 of *T. brasiliensis* (TBRA) (Lys1-TBRA) and lysozyme 1 and 2 of TINF (Lys1-TINF) as well as lysozyme sequences of *A. gambiae* (Lys-AGAM) and *A. aegypti* (Lys-AAED), which were used as outgroups. The gray shading enclosing the branches represents the clusters formed. Branch labels show the bootstrap samples supporting that branch

We also identified prolixicin homologs in all triatomine bugs and CLEC, but we failed to detect prolixicin homologs in ACPI (Fig. 1, Additional file 1: Table S4).

No homologs of cecropin, attacin, andropin, drosomyacin, drosocin, listericin and metchnikowin were identified in any hemipteran (Fig. 1, Additional file 1: Table S4).

Melanization

We found PPO homolog sequences, containing hemocyanin C-terminal (IPR005203) and immunoglobulin E-set (IPR014756) domains, in TTTs (Fig. 1, Additional file 1: Table S4). We corroborate the presence of previously identified PPO homologs, containing the aforementioned domains and a tyrosinase copper-binding (IPR002227) domain, in RPRO, CLEC and ACPI [30, 31, 33] (Fig. 1, Additional file 1: Table S4).

Nitric oxide

We identified a nitric oxide synthase (Nos) homolog sequence, containing a nitric oxide synthase N-terminal (IPR004030) domain, in TINE, but not in TDIM and TPAL. Also, Nos homologs with a eukaryote nitric-oxide synthase (IPR012144) domain, were identified in RPRO, CLEC and ACPI (Fig. 1, Additional file 1: Table S4).

Reactive oxygen species

Duox homologs were not detected in TTTs. But, we identified homolog sequences, containing a Dual oxidase peroxidase (IPR034821) and Haem peroxidase (IPR010255; IPR019791) domains, in RPRO and ACPI; as well as a number of previously reported homologs of Duox enzymes in CLEC [31]; although, only one (CLEC009522) of these sequences contains a Haem peroxidase and a Dual oxidase peroxidase domains (Fig. 1, Additional file 1: Table S4).

Homologs of NADPH oxidase family (Nox), containing (NADPH) oxidase 5 (IPR029648) domains, were identified in all hemipterans except TPAL and TINF (Additional file 1: Table S4). In addition, we detected homologs of the phospholipase C- β (PLC β), containing phosphoinositide phospholipase C (IPR001192) domains, in all hemipterans (Fig. 1, Additional file 1: Table S4).

Antioxidant system

Antioxidant system

Homologs of a *Drosophila* catalase, with multiple catalase-like and catalase immune-responsive (IPR010582) domains were observed in all hemipterans, except TINF (Fig. 1, Additional file 1: Table S6). CysGPx homologs (that participate in the elimination of hydrogen and organic peroxides), containing both glutathione peroxidase (IPR000889) and thioredoxin-like fold (IPR012336) domains, were detected only in TPAL and TDIM. Several TPx homologs, containing thioredoxin-like fold and peroxiredoxin C-terminal (IPR019479) domains, were identified in all hemipterans.

Additionally, numerous Haem peroxidase homologs, with peroxidase (IPR010255; IPR019791) domains and *Drosophila* peroxiredoxin-2540 homolog sequences, containing both alkyl hydroperoxide reductase subunit C/thiol specific antioxidant (IPR000866) and peroxiredoxin C-terminal (IPR019479) domains, were detected in all hemipterans (Fig. 1, Additional file 1: Table S6).

RNA interference

siRNA pathway

We recognized key components of the siRNA pathway in all hemipterans. Of these, Dcr-2 homologs, with Dicer dimerization (IPR005034), helicase ATP-binding (IPR014001), P-loop NTPase (IPR027417), PAZ (IPR003100) and ribonuclease III (IPR000999) domains, were identified in RPRO, CLEC and ACPI, but not in TTTs. R2D2 homologs, with a double-stranded RNA-binding (IPR014720) domains, were observed in CLEC and all triatomine bugs except TDIM. In triatomine bugs, we identified Ago2 homologs with Piwi (IPR003165), ribonuclease H-like (IPR012337) and PAZ domains. While homologs of these molecules containing other Ago-related (IPR032474; IPR014811; IPR032472) domains were detected in CLEC and ACPI (Fig. 1, Additional file 1: Table S7).

We also identified double-stranded RNA-binding domain Loqs homologs in TPAL, TDIM, and ACPI, while sequences with the characteristic staufen C-terminal (IPR032478) domain were observed in TINE, RPRO and CLEC. Additionally, other key enzymes from several RNAi-related pathways were detected in all hemipterans (Fig. 1, Additional file 1: Table S7).

Coagulation

We detected *Hemolectin* homologs, containing both coagulation factor 5/8 C-terminal (IPR000421) and von Willebrand factor, type D (IPR001846) domains, in RPRO, CLEC and ACPI, but not in TTTs. Homolog sequences of Fondue were not identified in any hemipteran (Fig. 1, Additional file 1: Table S8).

Discussion

Here, using as reference immune molecules described in established invertebrate immunology models [1, 36], we present a compressive description of the humoral and cellular innate immune components of four important Chagas disease vectors (TPAL, TDIM, TINF and RPRO), along with two other hemipterans phylogenetically related (CLEC and ACPI). Key homologs of constitutive and induced immune responses were identified in all the studied hemipterans (Fig. 1). However, compared to other insects; important differences were observed (Fig. 2). Our results in the TTTs extend previous observations in other hemipterans lacking several components of the Imd signaling

pathway. Further comparison with other hexapods, using published data, revealed that lacking Imd canonical components is common in several hemimetabolous species.

Transcriptome analysis of organisms with no sequenced genomes could provide useful preliminary gene catalogues. Although we cannot exclude that genes expressed at low levels or restricted to few cells could not be detected, this possibility was reduced by the normalization of the libraries and their high coverage values assessed using two approaches (CEGMA and BUSCO). Almost two-thirds of the *de novo* TTT transcriptomes mapped to the RPRO genome and predicted proteome [33], and three quarters of the RPRO proteome had homologous matches in TTT datasets. Nevertheless, although the constructed TTT-immune landscape appears patchy, the presence of canonical immune homologs in at least one of the triatomines could be considered as extant in TTTs as a group; this was further supported by their presence in the RPRO genome. While their presence only in RPRO (different genera) could not be interpreted as present in all triatomine bugs. The use of the available genomes of CLEC [31] and ACPI [71] ensured the most complete data available for meaningful comparative analysis with other hemipterans.

Overall, we identified hemipteran homologs belonging to the major immune categories (microbial recognition and activation, signaling, effectors, regulation, antioxidant system, RNA interference and coagulation), but with particular compositions that could be attributed to lifestyle and environmental exposures of these insects.

Conserved microbial recognition GNBPs in all hemipterans is consistent with the long ancestry of these receptors, while the presence of PGRPs in all hematophagous hemipterans, but not in ACPI could be indicative of their involvement in surveillance and activation of immune signaling against pathogens encountered in their different environs. Likewise, hemipterans possess cellular receptors as Dscam and different classes of scavenger receptors involved in cellular internalization of foreign agents [20, 21]. This hemocyte response also involves CTLs that are conserved in hemipterans, and participate in the defense against flagellated parasites [72, 73]. In triatomine bugs, these pathogen-binding molecules facilitate the recruitment of hemocytes for the encapsulation and melanization of pathogens [74].

TEPs share similarities with vertebrate complement factors C3/C4/C5 and have thioester sites for microbial recognition in a common [23]. TEP1 of *A. gambiae* participates in the defense against bacterial and *Plasmodium* infections [75, 76]. At least one member of the TEP family was detected in RPRO, CLEC and ACPI, but none in TTTs, which reflects an apparent low proportion of these recognition-molecules in hemipterans [30, 77, 78]. In contrast, TEPs are subject to rapid

lineage-expansions in other insects such as *Drosophila* [34] and mosquitoes [37, 79]. Remarkably, *Musca domestica* possesses the largest TEP repertoire of the sequenced dipterans. This genic expansion is related to the coexistence of these insects with a wide diversity of microorganisms [80]; the low representation of TEPs in hemipterans could be related to a limited exposure to pathogens mainly due to restricted diets.

Similarly, although we detected several serine proteases in all hemipterans, serine proteases with CLIP domain were detected in low proportions. In other insects, CLIPs participate in the regulation of extracellular pathways involved in the proteolytic activation of PPO and the Toll signaling pathway [81, 82], and represent large families of genes [36–38, 83]. The relative paucity of CLIP-domain contrast with an enrichment of protease inhibitor such as pacifastins (IPR008037), previously detected in *Triatoma* species (Martinez-Barnette et al., unpublished observations). In arthropods, pacifastins are upregulated after immune challenge and act as downregulators of the melanization by preventing PO activation [84]. Nevertheless, the significance of this finding remains uncertain.

As expected, genes responsible for constitutive primary effector mechanisms, widely conserved in insects, such as PPO enzymes responsible for the *de novo* synthesis of melanin, were present in all hemipterans. Melanization plays an important role in the elimination of a variety of pathogens, as well as in facilitating wound healing [85, 86]. In addition, this reaction is related to hemocyte-mediated processes such as phagocytosis of bacteria and parasite encapsulation [87]. We also detected PLC β homologs and numerous NADPH enzymes, responsible for the production of ROS. In triatomines, oxygen intermediates constitute a primary defense line against trypanosomatid parasites [88, 89]. ROS also play a key role in the regulation of intestinal bacteria, which undergo dramatic increases after blood meals [90]. In this context, antioxidants are particularly important for hematophagous insects continuously exposed to ROS, due to the release of heme after blood-feeding [88]. On the other hand, prolonged exposure to ROS leads to oxidative stress and cell damage [91], and the presence of enzymes responsible for the removal of hydrogen and organic peroxides, such as catalases, GPx and TPx, indicates the important role of these redox mechanisms for cellular homeostasis in hemipterans.

As in other insects, in hemipterans NO could act as a signaling and cytotoxic molecule after the damage produced by bacteria and parasites [8, 92–94]. NO is active in the hemolymph and the digestive tract of triatomines, where it contributes to resist trypanosomatid parasites [9, 94]. In addition, NO triggers the production of other effector molecules such as AMPs [8, 92, 95].

The synthesis of AMPs is consequence of the activation of NF- κ B and it is the hallmark of the induced humoral immune response in insects [4, 27]. As in other insects, the role of AMPs in triatomines is the defense against microbial agents, including *T. cruzi* [64]. Except for ACPI, we found defensin homologs in most hemipterans. The phylogenetic analysis revealed divergence between defensins homologs expressed in different tissues among species (Fig. 3). A group of RPRO defensins (clade A) appears related to DefA and DefB of RPRO that are upregulated in the fat body and midgut after immune challenge [61]. Similarity, a set of defensin sequences of TDIM and RPRO (clade F) seems related to DefC of RPRO that is also involved to immune functions [61]. While, a cluster of sequences of TINF, TPAL and TDIM (clade B and C) appears related to TBRA defensins (Def1–3) which are induced in the triatomine salivary glands and digestive tract after feeding [62, 63]. Both CLEC defensins appear to be related with TPAL sequences (clade C). Other not previously described TINF and TPAL sequences (clade D) grouped with a defensin expressed in the gut of TINF. Interestingly, a RPRO basal taxon (RPRC012182) and a set of defensins of this species (clade G) were grouped apart from other hemipterans, including the rest of RPRO sequences. A similar separation pattern was observed for a group of TPAL and TDIM sequences (clade E) that seem closely related, although the role of these groups of defensins is unclear.

Prolixicin homologs were only identified in triatomine bugs and CLEC. This AMP, related to the dipterin-attacin family, is expressed by the fat body and midgut of triatomines in response to bacterial infections, although it is not toxic for *T. cruzi* [96]. Lysozymes were conserved in all hemipterans. The function of lysozyme is still not clear in triatomines. These enzymes exhibit organ-dependent expression and are reportedly involved in both digestive and immune functions [64–66]. The phylogenetic analysis exposed separate clusters of CLEC (clade A), ACPI (clade F) and TTTs lysozyme sequences that seem to be related to their digestive and immune defense functions (Fig. 4). A group of RPRO and TDIM sequences appears related to LysA of RPRO associated to immune-related functions (clade B). This molecule is predominantly expressed in the intestinal tract after ingestion of *T. cruzi* in a blood meal, and after injection of bacteria into the haemocoel [63, 64]. LysB of RPRO that is expressed primarily in hemocytes and fat body after bacterial challenge [64] grouped with other not previously described RPRO and TPAL sequences (clade G) indicating their possible participation in immunity. A TINF lysozyme sequence, appears related to Lys1 of TINF and Lys1 of TBRA (clade C), which are upregulated in the stomach after feeding [62, 65]. This upregulation may reflect their digestive functions or their induction in response to the

drastic increase of bacterial populations in this organ after a blood meal [90]. Other two TINF lysozyme sequences clustered with Lys2 (clade D) expressed in the midgut of TINF, with not yet elucidated its physiological function [66]. While, a set of TPAL, TDIM, TINF and RPRO lysozymes grouped separately (clade E) from the other sequences, but their roles is still unknown.

In contrast, no homologs of other *Drosophila* AMPs such as cecropin, attacin, andropin, drosomyacin, drosocin, listericin and metchnikowin were identified in any hemipteran, corroborating previous observations in RPRO [97]. Although, the majority of AMPs, such as defensins, cecropins, proline-rich peptides and attacins have been found in several insect orders [98], some AMPs have been identified only in certain orders [99, 100]. While some AMPs, such as cecropins of *M. domestica*, exhibit significant duplication rates [101], it is possible that differences in insect AMPs repertoires could be consequence of different selection pressures exerted by exposure to pathogens and habitat conditions.

The Toll and Jak-STAT signaling cascades with most of their canonical components were documented in all hemipterans, corroborating previous observations in the RPRO, CLEC and ACPI genomes [30–33]. These ancestral pathways are widely conserved and participate in the development of bilaterally symmetric animals (such as worms, mollusks, arthropods and vertebrates) [102, 103]. In *Drosophila*, both Toll and Jak-STAT pathways serve a dual function in development and immunity [1].

Conversely, signaling canonical components Imd, FADD and DREED of the Imd pathway were not detectable in all studied the triatomines. The Imd signaling pathway is responsible for intestinal immune responses in dipterans. In *Drosophila*, intestinal dipterin, cecropin, drosocin and attacin regulated by this pathway, are constitutively expressed [104, 105], and the gut microbiota maintains basal their expression levels [104, 106, 107]. In mosquitoes, the growth of gut-dwelling bacteria induced by blood meals increases their expression through activation of Imd [105]. Thus, although we did not include microbial challenged insects in the preparations of our transcriptomes, blood-fed individual were included, and along with a basal expression, we expect similar inductions in these insects. As transcriptome datasets were normalized, the possibility that our transcriptomes were unable to detect at least one transcript of Imd pathway canonical members was minimal. The absence of these molecules in TTT is consistent previous observations in the RPRO, CLEC and ACPI genomes [30, 31, 33], which we corroborated by examining these genomes datasets.

Orthologs of Imd, FADD and DREED are highly conserved in a number of holometabolous insects from the orders Diptera, Hymenoptera, Lepidoptera and Coleoptera [34–42]; although these insects exhibit considerable

variations in the size and diversity of immune gene families [35–37]. In contrast, the absence of these Imd components appears to be a common feature in insects with incomplete metamorphosis (hemimetabolous), such as *Anasa tristis* (Hemiptera: Coreidae) [108], *Diaphorina citri* (Hemiptera: Liviidae) [77, 109], *Bemisia tabaci* (Hemiptera: Aleyrodidae) [110, 111] and *Pediculus humanus* (Phthiraptera: Pediculidae) [112]. Even in chelicerates (*Tetranychus*, *Metaseiulus* and *Ixodes*) [113, 114] and ametabolous hexapods of the subclass Collembola (*Folsomia candida* and *Orchesella cincta*) (without morphological transformations during their development) [115] appear to lack these key immune signaling molecules.

It has been suggested that the absence of key immune signaling components, particularly in the Imd pathway, may be the result of largely free-of-microbes diets (phloem sap or blood) that do not require specific defenses within the digestive tract, and do not exert selective pressures to maintain the high cost of immune defense [30].

In addition, these insects need to harbor populations of obligate symbionts that synthesize essential amino acids and vitamins that are poorly represented in their restricted diets [30, 116–119]. In triatomines, extracellular symbionts are mainly acquired through the consumption of feces of conspecifics during their first life stages and inhabit the midgut lumen, where they play digestive roles (hemolysis) [90, 120]. As depletion of symbionts results in drastic physiological and pathological alterations [121], it has been suggested that lacking a complete Imd (mainly responsible for the intestinal immune response) is an adaptation to ensure functional symbiosis. Nevertheless, other insects with obligate symbiotic relationships employ AMPs as coleoptericin to control symbiont populations [122, 123]. The induction of the coleoptericin family members is mainly regulated by the Imd pathway [124].

Hemimetabolous insects hatch as nymphs, morphologically similar to adults and grow progressively through molts until the adult stage. Adults differ from nymphs for the presence of functional wings and genitalia. In contrast, the holometabolous insects hatch as larvae and undergo drastic anatomical changes to pupa and adult. Among these, they suffer the complete remodeling of the larval midgut, which is then replaced by a new pupal epithelium that matures to the adult epithelium [125]. The destruction of larval intestinal epithelium is directed by the hormone 20-hydroxyecdysone (ecdysone) and mediated by processes of programmed cell death, involving the activation of caspases and regulation of IAP2, a key component of the Imd pathway [126]. In *Drosophila*, Imd is expressed at high levels during the pupariation stage, when massive apoptotic events occur [127].

The overexpression of Imd results in the activation of *reaper* in adult flies [127]. *Reaper* is a key pro-apoptotic gene in *Drosophila* [128]. Induction of *reaper* occurs in a stage-specific manner during larval midgut histolysis [126]. At this time, intestinal cells are exposed to the microorganisms present in the gut during the larva-pupa transition, coinciding with the release of different antimicrobial components into the intestine [129–132]. This may also contribute to the protection of the pupa and the adult from the bacterial threats that could originate during midgut remodeling [130]. This intestinal immune process is thought to be regulated, but is still not described. Although other proteins linked to the Imd cascade were identified in hemipterans, these may represent homologs involved in more general cellular processes such as ubiquitination and apoptosis [126, 133], but are not true Imd pathway orthologs. Thus, we speculate that these proteins, along new components (Imd, FADD, DREDD) that constitute the Imd signaling pathway were recruited by holometabolous insects in response to pathogenic bacterial threats during the intestinal remodeling in the course of metamorphosis. This concept is phylogenetically sound, as ametabolous and hemimetabolous insects (lacking Imd) are more ancient than holometabolous insects [134, 135].

Although it is difficult to establish the absence of genes from transcriptomes, the completeness of our assemblies (BUSCO and CEGMA assessments), and the datasets normalization from insects expected to respond to midgut microbiota, support the absence of Imd components in TTTs. Nevertheless, although Imd, FADD and DREDD are absent in the genomes of another Reduviid (RPRO) and the hemipterans CLEC and ACPI, further evidence is needed to corroborate our assumption. The high-resolution genomic-scale data derived from 1 K (Insect Transcriptome Evolution) with more than a thousand of insect-transcriptomes from all recognized taxonomic orders [135] and the 5000 arthropod genomes initiative (i5K) may represent useful resources to prove our hypothesis [136].

Conclusions

We provide evidence for the presence of major constitutive and inducible immune components in four important Chagas disease vectors (*T. pallidipennis*, *T. dimidiata*, *T. infestans* and *R. prolixus*) and two related hemipterans (*C. lecticularis* and *A. pisum*). Homologs involved in microbial recognition and immune activation (GNBPs, PGRPs, CTLs, TEPs, SRs and CLIPs) were documented in most species. But differences, like low proportions of TEPs and CLIPs, attributable to lifestyle and limited pathogen exposure were observed.

Conserved constitutive immune components responsible for *de novo* synthesis of melanin (PO), nitric oxide (Nos), and ROS production (PLC β and NADPH enzymes)

in all hemipterans reflect the relevance of these effector molecules in insect defense. A number of catalases, GPx and TPx reveal the importance of antioxidant mechanisms in hemipterans.

Several AMPs were found in most species; although differences in AMP repertoires were detected. Lysozymes related to digestive and immune defense functions were identified in all hemipterans. Defensins were detected only in the hematophagous hemipterans, exhibiting divergence according their differential expression in insect tissues. In contrast, no cecropins and attacins were detected in hemipterans, corroborating previous observations. Similarly, homologs of other *Drosophila* AMPs were not detected in any hemipteran species.

Most of the canonical components of the Toll and Jak-STAT signaling cascades are conserved in the studied insects. In contrast, key components of the Imd pathway (Imd, FADD and DREED) were absent from all hemipterans. Orthologs of Imd, FADD and DREED were documented in a number of holometabolous insects that undergo complete larval midgut remodeling of during metamorphosis. Conversely, the lack of Imd, FADD and DREED appears to be a common feature in more ancient insects with incomplete metamorphosis (hemimetabolous insects), including hemipterans. We speculate that these Imd signaling members were recruited by holometabolous insects in response to pathogenic bacterial threats during midgut remodeling.

Additional file

Additional file 1: Table S1. Immunity-genes reference dataset. Dataset of immunity-genes sequences retrieved from GenBank, VectorBase, UniProt, FlyBase and Ensembl databases and published information used to query for immune-related homologs in TTTs transcriptomes and RPRO, CLEC and ACPI predicted-peptide sets. **Table S2.** Immune-related homologs involved in immune recognition and activation (PGRPs; GNBNPs; Lectins; TEPs; Scavenger receptors; Clip-domain serine-proteases) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S3.** Immune-related homologs involved in immune signaling (Toll, Jak-STAT and Imd pathways) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S4.** Immune-related homologs involved in immune effectors (antimicrobial peptides; PPOs; nitric oxide synthase; NADPH enzymes) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S5.** Immune-related homologs involved in immune regulation (immune negative regulators) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S6.** Immune-related homologs involved in antioxidant mechanisms (Catalases; Peroxidases) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S7.** Immune-related homologs involved in RNA interference (RNAi machinery components) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. **Table S8.** Immune-related homologs involved in coagulation (Fondue and Hemolectin proteins) and their respective transcriptome/genome sequence-IDs in the six hemipteran species. (XLSX 53 kb)

Abbreviations

ACPI: *Acyrtosiphon pisum*; Ago: argonaut protein; AMP: antimicrobial peptide; CLIP: CLIP domain serine proteases; CTL: C-type lectin;

CysGPx: cysteine-containing glutathione peroxidase; Dcr: dicer protein; dsRNA: double-stranded RNA; Duox: dual oxidase; GNBP: Gram-negative binding protein; GPx: glutathione peroxidase; Imd: immunodeficiency; Jak-STAT: Janus kinase/Signal transducers and activators of transcription; Loqs: loquacious protein; NO: nitric oxide; Nox: NADPH oxidase; PAMP: pathogen-associated molecular pattern; PGRP: peptidoglycan receptor protein; PLC β : phospholipase C- β ; PNG: peptidoglycan; PO: phenoloxidase; PPO: prophenoloxidase; PRR: pathogen recognition receptor; Px: peroxidase; RLC: RISC-loading complex; RNAi: RNA interference; ROS: reactive oxygen species; RPRO: *Rhodnius prolixus*; siRNA: small interfering RNA; SR: scavenger receptor; SR-B: scavenger receptor class B; SR-C: scavenger receptor class C; TBRA: *Triatoma brasiliensis*; TDIM: *Triatoma dimidiata*; TEP: thioester-containing protein; TINF: *Triatoma infestans*; TPAL: *Triatoma pallidipennis*; TPx: thioredoxin peroxidase; TTTs: *Triatoma pallidipennis*, *Triatoma dimidiata* and *Triatoma infestans*

Funding

The present work was supported by the grant 240838 from Consejo Nacional de Ciencia y Tecnología (CONACyT, México). FAZE received a Ph.D. scholarship from CONACyT, México (CVU/Becario: 376792/244049).

Availability of data and materials

The conclusions of this work are supported by the information included within the article and its additional files. The datasets supporting the results of this article are available at <http://201.131.57.23:8080/data/triatoma>.

Authors' contributions

FAZE participated in the study design, interpretation data, writing the manuscript and carried out the transcriptomic analysis. JMB and MHR participated in the study design, the interpretation data and helped draft the manuscript. AL and RRP participated in the study design, and interpretation of data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study was approved by the Instituto Nacional de Salud Pública, México Institutional Bioethical Committee (approval No. 1681).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Centro de Investigación Sobre Enfermedades Infecciosas (CISEI), Instituto Nacional de Salud Pública, Cuernavaca, México. ²Centro de Bioinvestigaciones (CeBio) and Centro de Investigación y Transferencia del Noroeste de Buenos Aires (CITNOBA-CONICET), Universidad Nacional del Noroeste de la Provincia de Buenos Aires, Pergamino, Argentina. ³Laboratorio de Genética y Genómica Funcional. Centro Regional de Estudios Genómicos. Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina.

Received: 18 August 2017 Accepted: 28 November 2017

Published online: 22 January 2018

References

- Lemaitre B, Hoffmann J. The host defense of *Drosophila melanogaster*. *Annu Rev Immunol*. 2007;25:697–743.
- Gottar M, Gobert V, Matskevich AA, Reichhart JM, Wang C, Butt TM, et al. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. *Cell*. 2006;127:1425–37.
- Hultmark D. *Drosophila* immunity: paths and patterns. *Curr Opin Immunol*. 2003;15:12–9.
- Bulet P, Stocklin R. Insect antimicrobial peptides: structures, properties and gene regulation. *Protein Pept Lett*. 2005;12:3–11.

5. Ma C, Kanost MR. A beta1,3-glucan recognition protein from an insect, *Manduca sexta*, agglutinates microorganisms and activates the phenoloxidase cascade. *J Biol Chem*. 2000;275:7505–14.
6. Lee MH, Osaki T, Lee JY, Baek MJ, Zhang R, Park JW, et al. Peptidoglycan recognition proteins involved in 1,3-beta-D-glucan-dependent prophenoloxidase activation system of insect. *J Biol Chem*. 2004;279:3218–27.
7. Binggeli O, Neyen C, Poidevin M, Lemaitre B. Prophenoloxidase activation is required for survival to microbial infections in *Drosophila*. *PLoS Pathog*. 2014;10:e1004067.
8. Nappi AJ, Vass E, Frey F, Carton Y. Nitric oxide involvement in *Drosophila* immunity. *Nitric Oxide*. 2000;4:423–30.
9. Whitten MM, Mello CB, Gomes SA, Nigam Y, Azambuja P, Garcia ES, et al. Role of superoxide and reactive nitrogen intermediates in *Rhodnius prolixus* (Reduviidae)/*Trypanosoma rangeli* interactions. *Exp Parasitol*. 2001;98:44–57.
10. Ha EM, Oh CT, Bae YS, Lee WJ. A direct role for dual oxidase in *Drosophila* gut immunity. *Science*. 2005;310:847–50.
11. Bae YS, Choi MK, Lee WJ. Dual oxidase in mucosal immunity and host-microbe homeostasis. *Trends Immunol*. 2010;31:278–87.
12. Jones RM, Luo L, Ardita CS, Richardson AN, Kwon YM, Mercante JW, et al. Symbiotic lactobacilli stimulate gut epithelial proliferation via Nox-mediated generation of reactive oxygen species. *EMBO J*. 2013;32:3017–28.
13. Bae YA, Cai GB, Kim SH, Zo YG, Kong Y. Modular evolution of glutathione peroxidase genes in association with different biochemical properties of their encoded proteins in invertebrate animals. *BMC Evol Biol*. 2009;9:72.
14. Margis R, Dunand C, Teixeira FK, Margis-Pinheiro M. Glutathione peroxidase family - an evolutionary overview. *FEBS J*. 2008;275:3959–70.
15. Hillyer JF. Insect immunology and hematopoiesis. *Dev Comp Immunol*. 2016;58:102–18.
16. Hillyer JF, Strand MR. Mosquito hemocyte-mediated immune responses. *Curr Opin Insect Sci*. 2014;3:14–21.
17. Ding SW, Voinnet O. Antiviral immunity directed by small RNAs. *Cell*. 2007;130:413–26.
18. Karlikow M, Goic B, Saleh MC. RNAi and antiviral defense in *Drosophila*: setting up a systemic immune response. *Dev Comp Immunol*. 2014;42:85–92.
19. Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, et al. RNA interference directs innate immunity against viruses in adult *Drosophila*. *Science*. 2006;312:452–4.
20. Ramet M, Pearson A, Manfruell P, Li X, Koziel H, Gobel V, et al. *Drosophila* scavenger receptor CI is a pattern recognition receptor for bacteria. *Immunity*. 2001;15:1027–38.
21. Pearson A, Lux A, Krieger M. Expression cloning of dSR-CI, a class C macrophage-specific scavenger receptor from *Drosophila melanogaster*. *Proc Natl Acad Sci USA*. 1995;92:4056–60.
22. Watson FL, Puttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science*. 2005;309:1874–8.
23. Lagueux M, Perrodou E, Levashina EA, Capovilla M, Hoffmann JA. Constitutive expression of a complement-like protein in toll and JAK gain-of-function mutants of *Drosophila*. *Proc Natl Acad Sci USA*. 2000;97:11427–32.
24. Tanji T, Ohashi-Kobayashi A, Natori S. Participation of a galactose-specific C-type lectin in *Drosophila* immunity. *Biochem J*. 2006;396:127–38.
25. Cambi A, Koopman M, Fidgeor CG. How C-type lectins detect pathogens. *Cell Microbiol*. 2005;7:481–8.
26. Kanost MR, Jiang H. Clip-domain serine proteases as immune factors in insect hemolymph. *Curr Opin Insect Sci*. 2015;11:47–55.
27. Hoffmann JA, Reichhart JM. *Drosophila* innate immunity: an evolutionary perspective. *Nat Immunol*. 2002;3:121–6.
28. Hoffmann JA. The immune response of *Drosophila*. *Nature*. 2003;426:33–8.
29. Flores-Villegas AL, Salazar-Schettino PM, Cordoba-Aguilar A, Gutierrez-Cabrera AE, Rojas-Wastavino GE, Bucio-Torres MI, et al. Immune defence mechanisms of triatomines against bacteria, viruses, fungi and parasites. *Bull Entomol Res*. 2015;105:523–32.
30. Gerardo NM, Altincicek B, Anselme C, Atamian H, Barribeau SM, de Vos M, et al. Immunity and other defenses in pea aphids, *Acyrtosiphon pisum*. *Genome Biol*. 2010;11:R21.
31. Benoit JB, Adelman ZN, Reinhardt K, Dolan A, Poelchau M, Jennings EC, et al. Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nat Commun*. 2016;7:10165.
32. Ribeiro JM, Genta FA, Sorgine MH, Logullo R, Mesquita RD, Paiva-Silva GO, et al. An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *PLoS Negl Trop Dis*. 2014;8:e2594.
33. Mesquita RD, Vionette-Amaral RJ, Lowenberger C, Rivera-Pomar R, Monteiro FA, Minx P, et al. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc Natl Acad Sci USA*. 2015;112:14936–41.
34. Sackton TB, Lazzaro BP, Schlenke TA, Evans JD, Hultmark D, Clark AG. Dynamic evolution of the innate immune system in *Drosophila*. *Nat Genet*. 2007;39:1461–8.
35. Evans JD, Aronstein K, Chen YP, Hetru C, Imler JL, Jiang H, et al. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Mol Biol*. 2006;15:645–56.
36. Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC, et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science*. 2007;316:1738–43.
37. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science*. 2002;298:159–65.
38. Zou Z, Evans JD, Lu Z, Zhao P, Williams M, Sumathipala N, et al. Comparative genomic analysis of the *Tribolium* immune system. *Genome Biol*. 2007;8:R177.
39. Johnston PR, Makarova O, Rolff J. Inducible defenses stay up late: temporal patterns of immune gene expression in *Tenebrio molitor*. *G3 (Bethesda)*. 2013;4:947–55.
40. Tanaka H, Ishibashi J, Fujita K, Nakajima Y, Sagisaka A, Tomimoto K, et al. A genome-wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. *Insect Biochem Mol Biol*. 2008;38:1087–110.
41. Gunaratna RT, Jiang H. A comprehensive analysis of the *Manduca sexta* immunotranscriptome. *Dev Comp Immunol*. 2013;39:388–98.
42. Chetia H, Kabiraj D, Singh D, Mosahari PV, Das S, Sharma P, et al. De novo transcriptome of the muga silkworm, *Antheraea assamensis* (Helfer). *Gene*. 2017;611:54–65.
43. Rassi A Jr, Rassi A, Marin-Neto JA. Chagas disease. *Lancet*. 2010;375:1388–402.
44. Enger KS, Ordóñez R, Wilson ML, Ramsey JM. Evaluation of risk factors for rural infestation by *Triatoma pallidipennis* (Hemiptera: Triatominae), a Mexican vector of Chagas disease. *J Med Entomol*. 2004;41:760–7.
45. World Health Organization. Control of Chagas disease. In: World Health Organization technical report series; 2002. p. 1–109.
46. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27:863–4.
47. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
48. Parra G, Bradnam K, Korl I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*. 2007;23:1061–7.
49. Martínez-Barnette J, Gomez-Barreto RE, Ovilla-Munoz M, Tellez-Sosa J, Garcia Lopez DE, Dinglasan RR, et al. Transcriptome of the adult female malaria mosquito vector *Anopheles albimanus*. *BMC Genomics*. 2012;13:207.
50. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–2.
51. Giraldo-Calderon GI, Emrich SJ, MacCallum RM, Maslen G, Dialynas E, Topalis P, et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res*. 2015;43(Database issue):D707–13.
52. Legeai F, Shigenobu S, Gauthier JP, Colbourne J, Rispe C, Collin O, et al. AphidBase: a centralized bioinformatic resource for annotation of the pea aphid genome. *Insect Mol Biol*. 2010;19:5–12.
53. Brucker RM, Funkhouser LJ, Setia S, Pauly R, Bordenstein SR. Insect innate immunity database (IIID): an annotation tool for identifying immune genes in insect genomes. *PLoS One*. 2012;7:e45125.
54. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank. *Nucleic Acids Res*. 2005;33(Database issue):D34–8.
55. Pundir S, Martin MJ, O'Donovan C. UniProt protein knowledgebase. *Methods Mol Biol*. 2017;1558:41–55.
56. Gramates LS, Marygold SJ, Santos GD, Urbano JM, Antonazzo G, Matthews BB, et al. FlyBase at 25: looking to the future. *Nucleic Acids Res*. 2017;45:D663–71.
57. Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, et al. The Ensembl gene annotation system. *Database*. 2016;2016:baw093.
58. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7.
59. Eddy SR. Profile hidden Markov models. *Bioinformatics*. 1998;14:755–63.

60. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40.
61. Lopez L, Morales G, Ursic R, Wolff M, Lowenberger C. Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of Chagas disease. *Insect Biochem Mol Biol*. 2003;33:439–47.
62. Araujo CA, Waniek PJ, Stock P, Mayer C, Jansen AM, Schaub GA. Sequence characterization and expression patterns of defensin and lysozyme encoding genes from the gut of the reduviid bug *Triatoma brasiliensis*. *Insect Biochem Mol Biol*. 2006;36:547–60.
63. Waniek PJ, Castro HC, Sathler PC, Miceli L, Jansen AM, Araujo CA. Two novel defensin-encoding genes of the Chagas disease vector *Triatoma brasiliensis* (Reduviidae, Triatominae): gene expression and peptide-structure modeling. *J Insect Physiol*. 2009;55:840–8.
64. Ursic-Bedoya RJ, Nazzari H, Cooper D, Triana O, Wolff M, Lowenberger C. Identification and characterization of two novel lysozymes from *Rhodnius prolixus*, a vector of Chagas disease. *J Insect Physiol*. 2008;54:593–603.
65. Kollien AH, Fechner S, Waniek PJ, Schaub GA. Isolation and characterization of a cDNA encoding for a lysozyme from the gut of the reduviid bug *Triatoma infestans*. *Arch Insect Biochem Physiol*. 2003;53:134–45.
66. Balczun C, Knorr E, Topal H, Meiser CK, Kollien AH, Schaub GA. Sequence characterization of an unusual lysozyme gene expressed in the intestinal tract of the reduviid bug *Triatoma infestans* (Insecta). *Parasitol Res*. 2008;102:229–32.
67. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010;59:307–21.
68. FigTree version 1.4.3 2016. Available at <http://tree.bio.ed.ac.uk/software/figtree/>. Accessed 14 June 2017.
69. Fernando MD, Kounatidis I, Ligoxygakis P. Loss of Trabid, a new negative regulator of the *Drosophila* immune-deficiency pathway at the level of TAK1, reduces life span. *PLoS Genet*. 2014;10:e1004117.
70. Morris O, Liu X, Domingues C, Runchel C, Chai A, Basith S, et al. Signal integration by the IkkappaB protein pickle shapes *Drosophila* innate host defense. *Cell Host Microbe*. 2016;20:283–95.
71. The International Aphid Genomics Consortium. Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol*. 2010;8:e1000313.
72. Garcia ES, Castro DP, Figueiredo MB, Azambuja P. Parasite-mediated interactions within the insect vector: *Trypanosoma rangeli* strategies. *Parasit Vectors*. 2012;5:105.
73. Mello CB, Nigam Y, Garcia ES, Azambuja P, Newton RP, Ratcliffe NA. Studies on a haemolymph lectin isolated from *Rhodnius prolixus* and its interaction with *Trypanosoma rangeli*. *Exp Parasitol*. 1999;91:289–96.
74. Azambuja P, MCFD. Influence of the triatominae cellular and humoral defense system on the development of trypanosomatids. In: Fiocruz E, editor. Atlas of Chagas disease vectors in the Americas. Rio de Janeiro: Editora Fiocruz; 1997. p. 709–33.
75. Levashina EA, Moita LF, Blandin S, Vriend G, Lagueux M, Kafatos FC. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell*. 2001;104:709–18.
76. Blandin S, Shiao SH, Moita LF, Janse CJ, Waters AP, Kafatos FC, et al. Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell*. 2004;116:661–70.
77. Saha S, Hosmani PS, Villalobos-Ayala K, Miller S, Shippey T, Flores M, et al. Improved annotation of the insect vector of citrus greening disease: biocuration by a diverse genomics community. *Database*. 2017;2017:bax032.
78. Wang L, Tang N, Gao X, Guo D, Chang Z, Fu Y, et al. Understanding the immune system architecture and transcriptome responses to southern rice black-streaked dwarf virus in *Sogatella furcifera*. *Sci Rep*. 2016;6:36254.
79. Obbard DJ, Callister DM, Jiggins FM, Soares DC, Yan G, Little TJ. The evolution of TEP1, an exceptionally polymorphic immunity gene in *Anopheles gambiae*. *BMC Evol Biol*. 2008;8:274.
80. Scott JG, Warren WC, Beukeboom LW, Bopp D, Clark AG, Giers SD, et al. Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. *Genome Biol*. 2014;15:466.
81. An C, Zhang M, Chu Y, Zhao Z. Serine protease MP2 activates prophenoloxidase in the melanization immune response of *Drosophila melanogaster*. *PLoS One*. 2013;8:e79533.
82. Kambris Z, Brun S, Jang IH, Nam HJ, Romeo Y, Takahashi K, et al. *Drosophila* immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for toll activation. *Curr Biol*. 2006;16:808–13.
83. Cao X, He Y, Hu Y, Zhang X, Wang Y, Zou Z, et al. Sequence conservation, phylogenetic relationships, and expression profiles of nondigestive serine proteases and serine protease homologs in *Manduca sexta*. *Insect Biochem Mol Biol*. 2015;62:51–63.
84. Liu H, Jiravanichpaisal P, Cerenius L, Lee BL, Soderhall I, Soderhall K. Phenoloxidase is an important component of the defense against *Aeromonas hydrophila* infection in a crustacean, *Pacifastacus leniusculus*. *J Biol Chem*. 2007;282:33593–8.
85. Cerenius L, Lee BL, Soderhall K. The proPO-system: pros and cons for its role in invertebrate immunity. *Trends Immunol*. 2008;29:263–71.
86. Tang H. Regulation and function of the melanization reaction in *Drosophila*. *Fly (Austin)*. 2009;3:105–11.
87. Nappi AJ, Christensen BM. Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. *Insect Biochem Mol Biol*. 2005;35:443–59.
88. Finzi JK, Chiavegatto CW, Corat KF, Lopez JA, Cabrera OG, Mielniczki-Pereira AA, et al. *Trypanosoma cruzi* response to the oxidative stress generated by hydrogen peroxide. *Mol Biochem Parasitol*. 2004;133:37–43.
89. Cosentino-Gomes D, Rocco-Machado N, Meyer-Fernandes JR. *Rhodnius prolixus*: modulation of antioxidant defenses by *Trypanosoma rangeli*. *Exp Parasitol*. 2014;145:118–24.
90. Azambuja P, Feder D, Garcia ES. Isolation of *Serratia marcescens* in the midgut of *Rhodnius prolixus*: impact on the establishment of the parasite *Trypanosoma cruzi* in the vector. *Exp Parasitol*. 2004;107:89–96.
91. Sies H. Oxidative stress. London: Academic Press; 1985.
92. Foley E, O'Farrell PH. Nitric oxide contributes to induction of innate immune responses to gram-negative bacteria in *Drosophila*. *Genes Dev*. 2003;17:115–25.
93. Rivero A. Nitric oxide: an antiparasitic molecule of invertebrates. *Trends Parasitol*. 2006;22:219–25.
94. Whitten M, Sun F, Tew I, Schaub G, Soukou C, Nappi A, et al. Differential modulation of *Rhodnius prolixus* nitric oxide activities following challenge with *Trypanosoma rangeli*, *T. cruzi* and bacterial cell wall components. *Insect Biochem Mol Biol*. 2007;37:440–52.
95. Carton Y, Frey F, Nappi AJ. Parasite-induced changes in nitric oxide levels in *Drosophila paramelanica*. *J Parasitol*. 2009;95:1134–41.
96. Ursic-Bedoya R, Buchhop J, Joy JB, Durvasula R, Lowenberger C. Prolixicin: a novel antimicrobial peptide isolated from *Rhodnius prolixus* with differential activity against bacteria and *Trypanosoma cruzi*. *Insect Mol Biol*. 2011;20:775–86.
97. Ursic-Bedoya RJ, Lowenberger CA. *Rhodnius prolixus*: identification of immune-related genes up-regulated in response to pathogens and parasites using suppressive subtractive hybridization. *Dev Comp Immunol*. 2007;31:109–20.
98. Yi HY, Chowdhury M, Huang YD, Yu XQ. Insect antimicrobial peptides and their applications. *Appl Microbiol Biotechnol*. 2014;98:5807–22.
99. Hara S, Yamakawa M. Moricin, a novel type of antibacterial peptide isolated from the silkworm, *Bombyx mori*. *J Biol Chem*. 1995;270:29923–7.
100. Axen A, Carlsson A, Engstrom A, Bennich H. Gloverin, an antibacterial protein from the immune hemolymph of *Hyalophora pupae*. *Eur J Biochem*. 1997;247:614–9.
101. Sackton TB, Lazzaro BP, Clark AG. Rapid expansion of immune-related gene families in the house fly, *Musca domestica*. *Mol Biol Evol*. 2017;34:857–72.
102. Erwin DH, Davidson EH. The last common bilaterian ancestor. *Development*. 2002;129(13):3021–32.
103. Erwin DH, Davidson EH. The evolution of hierarchical gene regulatory networks. *Nat Rev Genet*. 2009;10:141–8.
104. Buchon N, Broderick NA, Chakrabarti S, Lemaitre B. Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. *Genes Dev*. 2009;23:2333–44.
105. Barletta AB, Nascimento-Silva MC, Talyuli OA, Oliveira JH, Pereira LO, Oliveira PL, et al. Microbiota activates IMD pathway and limits Sindbis infection in *Aedes aegypti*. *Parasit Vectors*. 2017;10:103.
106. Ryu JH, Kim SH, Lee HY, Bai JY, Nam YD, Bae JW, et al. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in *Drosophila*. *Science*. 2008;319:777–82.
107. Lhocine N, Ribeiro PS, Buchon N, Wepf A, Wilson R, Tenev T, et al. PIMS modulates immune tolerance by negatively regulating *Drosophila* innate immune signaling. *Cell Host Microbe*. 2008;4:147–58.
108. Shelby KS. Functional Immunomics of the squash bug, *Anasa tristis* (de Geer) (Heteroptera: Coreidae). *Insects*. 2013;4:712–30.
109. Arp AP, Hunter WB, Pelz-Stelinski KS. Annotation of the asian citrus psyllid genome reveals a reduced innate immune system. *Front Physiol*. 2016;7:570.

110. Zhang CR, Zhang S, Xia J, Li FF, Xia WQ, Liu SS, et al. The immune strategy and stress response of the Mediterranean species of the *Bemisia tabaci* Complex to an orally delivered bacterial pathogen. *PLoS One*. 2014;9:e94477.
111. Chen W, Hasegawa DK, Kaur N, Klot A, Pinheiro PV, Luan J, et al. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biol*. 2016;14:110.
112. Kim JH, Min JS, Kang JS, Kwon DH, Yoon KS, Strycharz J, et al. Comparison of the humoral and cellular immune responses between body and head lice following bacterial challenge. *Insect Biochem Mol Biol*. 2011;41:332–9.
113. Palmer WJ, Jiggins FM. Comparative genomics reveals the origins and diversity of arthropod immune systems. *Mol Biol Evol*. 2015;32:2111–29.
114. Smith AA, Pal U. Immunity-related genes in *Ixodes scapularis*—perspectives from genome information. *Front Cell Infect Microbiol*. 2014;4:116.
115. Faddeeva A, Studer RA, Kraaijeveld K, Sie D, Ylstra B, Marien J, et al. Collembolan transcriptomes highlight molecular evolution of hexapods and provide clues on the adaptation to terrestrial life. *PLoS One*. 2015; 10:e0130600.
116. Douglas AE, Prosser WA. Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrtosiphon pisum*) symbiosis. *J Insect Physiol*. 1992;38:565–8.
117. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc Natl Acad Sci USA*. 2010; 107:769–74.
118. Beard CB, Cordon-Rosales C, Durvasula RV. Bacterial symbionts of the triatominae and their potential use in control of Chagas disease transmission. *Annu Rev Entomol*. 2002;47:123–41.
119. Sasser D, Epis S, Pajoro M, Bandi C. Microbial symbiosis and the control of vector-borne pathogens in tsetse flies, human lice, and triatomine bugs. *Pathog Glob Health*. 2013;107:285–92.
120. Otálora-Luna F, Pérez-Sánchez AJ, Sandoval C, Aldana E. Evolution of hematophagous habit in Triatominae (Heteroptera: Reduviidae). *Rev Chil Hist Nat*. 2015;88:1–13.
121. Eichler S, Schaub GA. Development of symbionts in triatomine bugs and the effects of infections with trypanosomatids. *Exp Parasitol*. 2002; 100:17–27.
122. Login FH, Heddi A. Insect immune system maintains long-term resident bacteria through a local response. *J Insect Physiol*. 2013;59:232–9.
123. Anselme C, Perez-Brocal V, Vallier A, Vincent-Monegat C, Charif D, Latorre A, et al. Identification of the weevil immune genes and their expression in the bacteriome tissue. *BMC Biol*. 2008;6:43.
124. Yokoi K, Koyama H, Minakuchi C, Tanaka T, Miura K. Antimicrobial peptide gene induction, involvement of toll and IMD pathways and defense against bacteria in the red flour beetle, *Tribolium castaneum*. *Results Immunol*. 2012;2:72–82.
125. Hakim RS, Baldwin K, Smaghe G. Regulation of midgut growth, development, and metamorphosis. *Annu Rev Entomol*. 2010;55:593–608.
126. Jiang C, Baehrecke EH, Thummel CS. Steroid regulated programmed cell death during *Drosophila* metamorphosis. *Development*. 1997;124:4673–83.
127. Georgel P, Naitza S, Kappler C, Ferrandon D, Zachary D, Swimmer C, et al. *Drosophila* immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis. *Dev Cell*. 2001;1:503–14.
128. White K, Tahaoglu E, Steller H. Cell killing by the *Drosophila* gene reaper. *Science*. 1996;271:805–7.
129. Tsuji Y, Nakajima Y, Homma K, Natori S. Antibacterial activity of a novel 26-kDa serine protease in the yellow body of *Sarcophaga peregrina* (flesh fly) pupae. *FEBS Lett*. 1998;425:131–3.
130. Franzetti E, Romanelli D, Caccia S, Cappelozza S, Congiu T, Rajagopalan M, et al. The midgut of the silkworm *Bombyx mori* is able to recycle molecules derived from degeneration of the larval midgut epithelium. *Cell Tissue Res*. 2015;361:509–28.
131. Tettamanti G, Grimaldi A, Casartelli M, Ambrosetti E, Ponti B, Congiu T, et al. Programmed cell death and stem cell differentiation are responsible for midgut replacement in *Heliothis virescens* during prepupal instar. *Cell Tissue Res*. 2007;330:345–59.
132. Russell VW, Dunn PE. Lysozyme in the midgut of *Manduca sexta* during metamorphosis. *Arch Insect Biochem Physiol*. 1991;17:67–80.
133. Severo MS, Sakhon OS, Choy A, Stephens KD, Pedra JH. The 'ubiquitous' reality of vector immunology. *Cell Microbiol*. 2013;15:1070–8.
134. Truman JW, Riddiford LM. The origins of insect metamorphosis. *Nature*. 1999;401:447–52.
135. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. *Science*. 2014;346:763–7.
136. i5K Consortium. The i5K initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *J Hered*. 2013; 104:595–600.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

