Journal of Insect Physiology 114 (2019) 1–9

Contents lists available at ScienceDirect

Journal of Insect Physiology

journal homepage: www.elsevier.com/locate/jinsphys

Domain structure and expression along the midgut and carcass of peritrophins and cuticle proteins analogous to peritrophins in insects with and without peritrophic membrane

Renata O. Dias[a], Christiane Cardoso[a], Camila S. Leal[b], Alberto F. Ribeiro[b], Clélia Ferreira[a], Walter R. Terra[a,*]

[a] Departamento de Bioquimica, Instituto de Quimica, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-000 São Paulo, Brazil
[b] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, C.P. 11461, 05422-970 São Paulo, Brazil

ARTICLE INFO

Keywords:
PM proteins
PM-less insects
Chitin synthase
CPAP
PMP
Ubiquitous-CBD-containing protein

ABSTRACT

Most insects have a peritrophic membrane (matrix) (PM) surrounding the food bolus. This structure, similarly to the cuticle, is mainly composed of chitin and proteins. The main proteins forming PM are known as peritrophins (PMP), whereas some of the cuticle proteins are the cuticle proteins analogous to peritrophins (CPAP). Both proteins are composed of one or more chitin binding peritrophin-A domain (CBD) and no other recognized domain. Furthermore, insects containing PM usually have two chitin synthase (CS) genes, one mainly expressed in carcass and the other in midgut. In this work we identified PMP, CPAP and CS genes in the genome of insects from the Polyneoptera, Paraneoptera and Holometabola cohorts and analyzed their expression profile in different species from each group. In agreement with the absence of PM, we observed less CBD-containing proteins and only one CS gene in the genome of Paraneoptera species, except for the Phthiraptera *Pediculus humanus*. The lack of PM in Paraneoptera species was also confirmed by the micrographs of the midgut of two Hemiptera species, *Dysdercus peruvianus* and *Mahanarva fimbriolata* which agreed with the RNA-seq data of both species. Our analyses also highlighted a higher number of CBD-containing proteins in Holometabola in relation to the earlier divergent Polyneoptera group, especially regarding the genes composed of more than three CBDs, which are usually associated to PM formation. Finally, we observed a high number of CBD-containing proteins being expressed in both midgut and carcass tissues of several species, which we named as ubiquitous-CBD-containing proteins (UCBP), as their function is unclear. We hypothesized that these proteins can be involved in both cuticle and PM formation or that they can be involved in immune response and/or tracheolae formation.

1. Introduction

Most insects have a peritrophic membrane (PM, also named peritrophic matrix), which is a film (not a mass as suggested by the name matrix) composed of chitin and proteins, mainly peritrophins, surrounding the food bolus. The primary role of PM is to enhance the digestive efficiency by compartmentalizing the digestive processes. This allows midgut countercurrent fluid fluxes that remove digestive enzymes and partially digested food molecules from inside the PM, preventing enzyme excretion, and removing oligomers that may inhibit depolymerases (Terra, 2001; Bolognesi et al., 2008). Moreover, PM also prevents non-specific binding of undigested food to the midgut surface, avoiding absorption inhibition (Terra, 2001; Bolognesi et al., 2008). Finally, PM may complement the action of the mucus layer (Dias et al., 2018) against mechanical damage (Peters, 1992), plant allelochemicals (Barbehenn, 2001), and bacterial infection (Kuraishi et al., 2011).

PMs are present in almost all insect lineages, except for some groups like Paraneoptera (Psocoptera, book lice; Phthyraptera, lice; Thysanoptera, thrips; Hemiptera, bugs) (Peters, 1992). Thysanoptera and Hemiptera have as extracellular layers microvilli-associated membranes. These lipoproteic membranes may surround the microvilli as glove fingers (Burgos and Gutiérrez, 1976; Lane and Harrison, 1979), known as perimicrovillar membranes (PMM, Terra, 1988, 1990); they can be associated with the tips of the microvilli (modified perimicrovillar membranes, MPMM, Cristofoletti et al., 2003) or even form bundles of microvilli (microvilli bundle-forming membranes, MBFM, Del Bene et al., 1991; Utiyama et al., 2016). The absence of PM in such groups is thought to result from its loss in their ancestral group, as an

* Corresponding author.
E-mail address: warterra@iq.usp.br (W.R. Terra).

https://doi.org/10.1016/j.jinsphys.2019.02.002
Received 14 November 2018; Received in revised form 22 January 2019; Accepted 4 February 2019
Available online 05 February 2019

adaption to a plant sap diet poor in proteins and starch, which does not require luminal digestion (Terra, 2001). Finally, some insect midguts have instead of PM a peritrophic gel, which is a gel-like and chitin-free structure that could entirely or partially cover regions lacking PM (Terra, 2001).

PM proteins are mainly composed of peritrophins (or peritrophic matrix/membrane proteins – PMPs), proteins having one or more chitin binding peritrophin-A domains (CBDs). These proteins are responsible for organizing the chitin fibrils, allowing PM formation. Similar to the PM, insect cuticles are also composed of chitin and proteins. In this structure, one of the main proteins are known as cuticular proteins analogous to peritrophins (CPAPs). CPAPs are typically composed of one or three chitin binding peritrophin-A domains (CBDs) and are essentials for, among other processes, insect development, molting, and cuticle integrity (Jasrapuria et al., 2012; Tetreau et al., 2015). PMPs and CPAPs are very similar proteins that could be only distinguished by expression, phylogenetical and functional studies (Jasrapuria et al., 2010; Tetreau et al., 2015).

PM evolutionary appearance seems to be also accompanied by the emergence of a second chitin synthase (CS) gene in the Arthropoda lineage (CSM or CHS2). This second CS gene is thought to come from a paraphyletic group that was lost in several Arthropoda lineages (Zakrzewski et al., 2014). The CS gene duplication allowed a specialization of these genes to code for proteins able to act in cuticle or in PM formation. In fact, these genes were mainly expressed in carcass (CSC) or in midgut (CSM) tissues (Dias et al., 2018).

This work was undertaken to test the hypothesis that associated with the loss of PMs, Paraneopterans lost PMPs and midgut chitin synthase and conserved CPAPs, in contrast to both Polyneopterans and Holometabolans that conserved all of them. The results showed that in Paraneoptera species there are fewer genes encoding CBD-containing proteins with more than 3 CBDs than in Polyneoptera or Holometabola species. The data also revealed the expression of CBD-containing proteins both in midgut and carcass (Ubiquitous CBD-containing Protein, UCBP), with unclear function. Micrography studies also confirmed the absence of PM in the entire midgut of the Paraneoptera *Dysdercus peruvianus* (Hemiptera, Prosorrhyncha) and *Mahanarva fimbriolata* (Hemiptera, Auchenorryncha) species. In agreement with the absence of PM, Paraneoptera species, except for Phthiraptera, have lost both a second CS gene and a CS gene expression in their midgut.

2. Material and methods

2.1. Genome databases

Protein sequences from three Polyneoptera, 18 Paraneoptera and three Holometabola genomes were retrieved from the Refseq Genome (O'Leary et al., 2016), i5K (Evans et al., 2013), VectorBase (Giraldo-Calderon et al., 2015), BIPAA (https://bipaa.genouest.org/), and Ensemble Genome (Kersey et al., 2018) databases. *Periplaneta americana* genome sequences were retrieved from Li et al. (2018). For genes with more than one alternative transcript, only the longest isoform was kept for further analysis. All genome, transcriptome, and microscopy data used in this work is described in Table 1.

2.2. Transcriptome assembling and RNA-seq analysis

Insects used to perform RNA-seq analysis were reared as described in Table 1. The dissection of insect midgut was done with great care to avoid contamination by other tissues. In spite of that, it was not possible to remove tracheolae from the midgut samples.

For the Polyneoptera *P. americana* species, total RNA from salivary gland and midgut tissues was extracted using the Trizol® protocol and sequenced in an Illumina HiSeq 2500 equipment using a paired-end strategy (2 × 300 bp). In order to remove bad quality and adaptor sequences, the obtained reads were trimmed using the Trimmomatic software (Bolger et al., 2014). The resultant paired-end reads were then aligned to the reference genome scaffolds (Li et al., 2018) using the STAR v.2.6 software (Dobin et al., 2013).

For *Mahanarva fimbriolata*, *Dysdercus peruvianus*, *Abracris flavolineata*, *Musca domestica*, *Spodoptera frugiperda*, and *Tenebrio molitor* total RNA samples were extracted using the Trizol® protocol and sequenced in a HiSeq 2500 equipment, using a paired-end strategy (2 × 100pb). Transcriptome assembly and RNA-seq analyses for those species used the same strategy described in Dias et al. (2018).

For *Rhodnius prolixus*, total RNA samples from anterior and posterior midgut, and rectum were extracted and sequenced using a 454 Genome Sequencer FLX Titanium machine, as detailed in Ribeiro et al. (2014). The sequenced reads were trimmed using fastx-toolkit tools and aligned to the reference genome scaffolds (Mesquita et al., 2015) using the STAR software v.2.6 (Dobin et al., 2013).

For all RNA-seq analyses, the number of aligned reads per gene was normalized as Transcripts per Million (TPM), according to Wagner et al. (2012).

2.3. CPAP, PMP, ubiquitous CBD -containing protein (UCBP) and chitin synthase sequence identification

CPAP, PMP, and UCBP protein sequences from each analyzed genome were identified as those having at least one chitin-binding Peritrophin-A domain (CBD - PFAM family ID: PF01607) and no other predicted domain by InterproScan version 5.28–67.0 (Jones et al., 2014). A list with all accession identifiers corresponding to CPAP, PMP, UCBP and chitin synthase protein sequences identified in the analyzed genomes are described in Supplementary Table 1.

For species in which only a transcriptome assembly was available and transcripts are usually incomplete, the CPAP/PMP/UCBP sequences were identified in two steps. First, the translated sequences identified in the genome searches were used as query to search homologous proteins in these assemblies using the BLASTp algorithm (Altschul et al., 1990) (threshold of: protein sequence identity ≥ 70% and query coverage by the subject ≥ 50%). Second, the sequences identified in this step were used as query to search homologous proteins in the genomes sequences (threshold of: protein sequence identity ≥ 50%). This two-steps procedure (reciprocal blast) was performed to avoid the incorrect annotation of incomplete sequences.

CBD domains were also checked for their amino acid composition, as proposed by Jasrapuria et al. (2010), but allowing a variability of two amino acid less or more between two consecutive cysteine residues, except for the five residues between the second and the third cysteine: $CX_{9\text{-}26}CX_5CX_{7\text{-}16}CX_{10\text{-}18}CX_{4\text{-}10}C$.

Chitin synthase (CS) protein sequences were identified as those possessing the IPR004835 domain, according to the Interproscan version 5.28–67.0 (Jones et al., 2014) prediction. A CS phylogenetic proposition was performed using the RAxML software version 8.2.11 (Stamatakis, 2014) using the PROTGAMMAAUTO parameter to automatically define the best protein model.

All transcriptome sequences from de novo transcriptome assemblies used in this work were described in the Supplementary Table 2.

2.4. Light and electron microscopy

For histological studies with the light microscope, animals were dissected in the fixative (Bouin's solution) under the stereomicroscope and the midguts were pulled apart. The tissue was kept in the fixative for 6 h at room temperature, upgraded in ethanol and embedded in historesin (Leica, Heidelberg). Serial sections (3–4 μm thickness) were cut using a Leica RM 2145 microtome, stained with haematoxylin and eosin, and mounted in glass slides with Entellan (Merck, Darmstadt).

For fluorescent visualization of chitin (or structures very rich in N-acetyl-glucosamine or sialic acid), the samples were fixed and incubated with wheat germ agglutinin coupled to fluorescein

Table 1
Description of the genome, transcriptome and microscopy data used in this work.

Cohort	Order	Species	Type of data	Source	Observation
Pol.	Orthoptera	*Locusta migratoria*	Genome	i5K	
		Abracris flavolineata	Transcriptome / RNA-seq	Our data	Adults, collected in the field, fed with collard leaves and with ample access to water
	Dictyoptera	*Periplaneta americana*	Genome	Li et al., 2018	
			RNA-seq	Our data	Adults, maintained in laboratory, and fed on oats and chayote
		Zootermopsis nevadensis	Genome	RefSeq Genome	
Par.	Thysanoptera	*Frankliniella occidentalis*	Genome	i5K	
	Hemiptera/Auchenorrhyncha	*Homalodisca vitripennis*	Genome	i5K	
		Nilaparvata lugens	Genome	RefSeq Genome	
		Mahanarva fimbriolata	Transcriptome/RNA-seq/ Microscopy	Our data	Adults, collected in the field
	Hemiptera/Sternorrhyncha	*Bemisia tabaci*	Genome	RefSeq Genome	
		Aphis glycines	Genome	BIPAA	
		Rhopalosiphum padi	Genome	BIPAA	
		Acyrthosiphon pisum	Genome	Ensemble Genome	
		Diuraphis noxia	Genome	RefSeq Genome	
		Myzus persicae	Genome	RefSeq Genome	
		Myzus cerasi	Genome	BIPAA	
		Pachypsylla venusta	Genome	i5K	
		Diaphorina citri	Genome	i5K	
	Hemiptera/Prosorrhyncha	*Gerris buenoi*	Genome	i5K	
		Cimex lectularius	Genome	RefSeq Genome	
		Rhodnius prolixus	Genome	VectorBase	
			RNA-seq	Ribeiro et al., 2014	Adults, maintained in laboratory, and fed on rabbit blood according Ribeiro et al., 2014
		Oncopeltus fasciatus	Genome	i5K	
		Halyomorpha halys	Genome	RefSeq Genome	
		Dysdercus peruvianus	Transcriptome/RNA-seq/ Microscopy	Our data	Adults, maintained in laboratory and fed on cotton seeds with ample access to water
	Phthiraptera	*Pediculus humanus*	Genome	VectorBase	
Hol.	Coleoptera	*Tribolium castaneum*	Genome	RefSeq Genome	
		Tenebrio molitor	Transcriptome/RNA-seq	Our data	Larvae, maintained in laboratory, and fed on wheat bran
	Lepidoptera	*Spodoptera litura*	Genome	RefSeq Genome	
		Spodoptera frugiperda	Transcriptome/RNA-seq	Our data	Larvae, maintained in laboratory, and fed on artificial diet according Parra (1986)
	Diptera	*Musca domestica*	Genome	VectorBase	
			RNA-seq	Our data	Larvae, maintained in laboratory, and reared in a mixture of fermented commercial pig food and rice hull (1:2) (Targa and Peres, 1979)

*Pol: Polyneoptera, Par: Paraneoptera, and Hol: Holometabola.

isothiocyanatein (WGA-FITC) in the presence of excess N-acetylglucosamine, as detailed in Caldeira et al. (2007).

For transmission electron microscopy, midgut pieces were fixed in 3% glutaraldehyde in cacodylate buffer, followed by 1% osmium tetroxide in the same buffer and processed as described by Caldeira et al. (2007).

3. Results

3.1. Chitin synthase, CPAP, PMP and UCBP in selected genomes from insects of three cohorts

Insects with PM usually have two genes coding for chitin synthases (CSs). One is preferentially expressed outside the midgut, that is in carcass, and codes for cuticle chitin synthase (CSC or CHS1). The other is preferentially expressed in the midgut tissues, coding for PM chitin synthase, implied in the production of PM (CSM or CHS2). In agreement with the presence of PM, two chitin synthase genes were found in all Holometabola and in two (see below) of the three Polyneoptera analyzed species, whereas only one gene was found in Paraneoptera genomes, except for *P. humanus* that has two (Table 2). All CS protein sequences were further clustered by a Maximum Likelihood phylogenetic tree (Fig. 1). The predicted tree corroborates the classification of these proteins as CSC or CSM. A protein sequence alignment of all CSs confirmed they must be functional (Supplementary Fig. 1)

Cuticular Protein Analogous to Peritrophin (CPAP), Peritrophic Membrane/matrix Protein (PMP) and Ubiquitous Chitin-Binding domain containing Protein (UCBP) sequences from the three cohorts were selected by InterproScan analyses, as those proteins with at least one CBD and no other identified domain. It is important to notice that we decided to include in our work a new protein class, the UCBP, proteins that differ from the other two classes in having significant expression values in both midgut and carcass. A classification of PMP, CPAP, and UCBP based only on their protein sequence. Table 2 shows a sum of the number of those proteins found in each species.

The highest number of genes coding for PMP, CPAP, and UCBP proteins were described in the Holometabola and Polyneoptera, with an average of 52 and 40 genes identified by the InterproScan analysis, respectively (Table 2). Paraneoptera species have in average 25 predicted genes, with the *Sternorrhyncha* species (except for *D. citri*) possessing the smaller number (less than 24 genes).

For all species, CPAP/PMP/UCBP predicted proteins were mainly composed of sequences with a single predicted CBD. These proteins represent in average 64, 65 and 39% of all predicted sequences of these classes for Polyneoptera, Paraneoptera and Holometabola species, respectively (Table 2). Proteins with more than three CBDs were found in all cohorts but mainly in Holometabola species, in agreement with their putative role in PM formation. It is noticeable, however, that Polyneoptera species that also possess PM have peritrophins with fewer CBD per sequence than Holometabola species. This suggests that Polyneoptera PMs may have different properties from those of Holometabola.

Table 2
Number of chitin synthase and CPAP, PMP, and UCBP sequences found in selected genomes from insects pertaining to the three insect cohorts.

			Chitin synthase	CPAP, PMP, and UCBP				
				Sequences per number of CBDs				Total number of sequences
Cohort	Order	Species		1 CBD	2 CBDs	3 CBDs	greater than 3 CBDs	
Pol	Orthoptera	*L. migratoria*	2	26	6	10	0	42
	Dictyoptera	*P. americana*	1*	25	8	4	4	41
		Z. nevadensis	2	23	5	7	2	37
Par	Thysanoptera	*F. occidentalis*	1	19	5	3	1	28
	Hemiptera/Auchenor-ryncha	*H. vitripennis*	0	13	4	4	1	22
		N. lugens	1	24	5	10	2	41
	Hemiptera/Sternor-rhyncha	*B. tabaci*	1	15	2	5	1	23
		A. glycines	1	11	2	3	2	18
		R. padi	1	13	1	6	1	21
		A. pisum	1	12	1	6	1	20
		D. noxia	1	11	2	5	0	18
		M. persicae	1	13	1	5	1	20
		M. cerasi	1	11	2	5	1	19
		P. venusta	1	7	3	2	0	12
		D. citri	1	19	8	1	0	28
	Hemiptera/Prosor-rhyncha	*G. buenoi*	1	24	5	5	0	34
		C. lectularius	1	21	3	6	2	32
		R. prolixus	1	23	3	3	2	31
		O. fasciatus	1	23	6	2	1	32
		H. halys	1	23	2	6	1	32
	Phthiraptera	*P. humanus*	2	15	2	3	3	23
Hol	Coleoptera	*T. castaneum*	2	22	9	8	6	45
	Lepidoptera	*S. litura*	2	14	6	10	12	42
	Diptera	*M. domestica*	2	24	15	16	14	69

Pol: Polyneoptera, Par: Paraneoptera, and Hol: Holometabola. Chitin synthase sequences without all critical signatures were discarded.
*The unique chitin synthase for *P. americana* is possibly the result of a mispredicted gene fusion during the genome annotation (see details in text Section 3.2). The number of the combined CPAP, PMP and UCBP sequences were shown in columns according to their amount of CBDs.

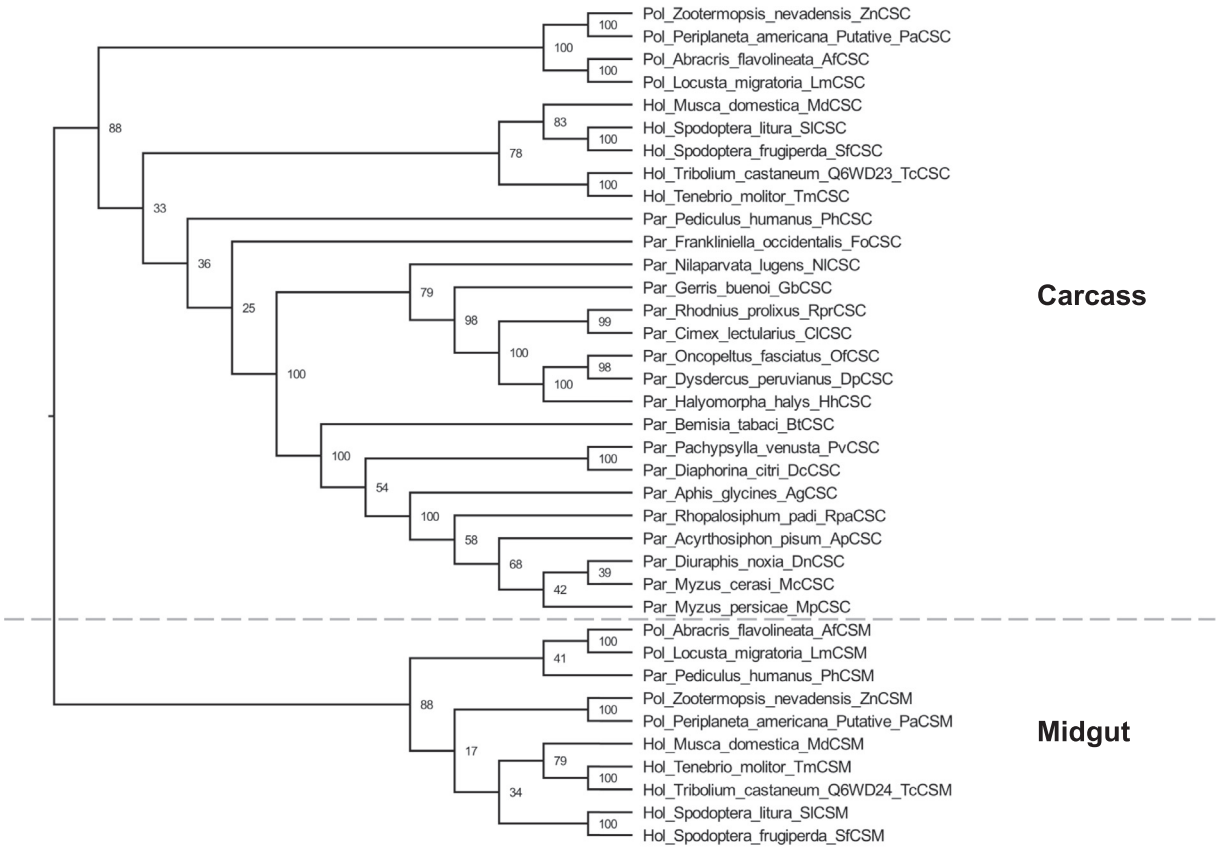

Fig. 1. Phylogenetic tree inference for Insecta chitin synthase genes. Pol: Polyneoptera, Par: Paraneoptera, and Hol: Holometabola. The taxonomic information for the selected species can be found in Table 1. Numbers on the nodes represent bootstrap values for 100 replicates.

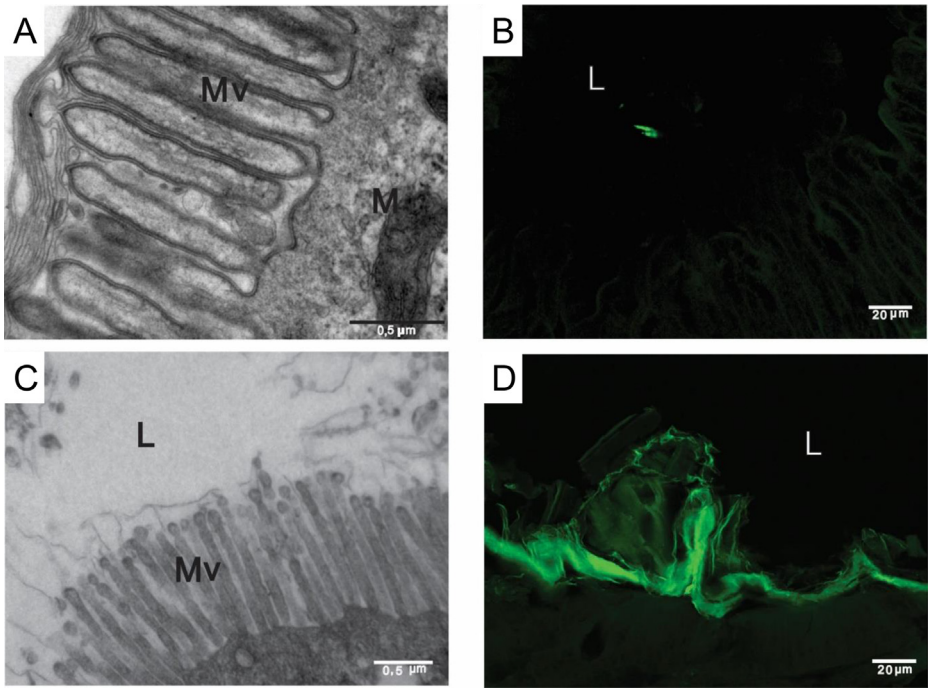

Fig. 2. Luminal and apical structures in the midguts of several insects. Electron micrograph (A) and light microscopy (B) of the midgut from *Dysdercus peruvianus,* showing microvilli (Mv) enclosed by perimicrovilar membranes in A and absence of WGA-FITC fluorescence in B, indicating the lack of chitin at the luminal surface. Electron microscopy (C) of the apex of midgut cells in *Mahanarva fimbriolata,* with microvilli (Mv) surrounded by microvilli-bundle-forming membranes, and light microscopy (D) of the same region with faint WGA-FITC fluorescence after addition of excess N-acetylglucosamine, suggesting the absence of chitin at the luminal surface.

3.2. Chitin synthase, CPAP, PMP, and UCBP genes in Polyneoptera insects

In Polyneoptera, two chitin synthase genes were found in two of the three analyzed genomes. It is important to notice that in *Z. nevadensis* genome CS genes occur in tandem. This gene distribution may also be true for *P. americana*, leading to the prediction by mistake of only one large CS gene in the current version of *P. americana* genome (Li et al., 2018). This sequence would code for a protein of 2652 amino acids of length, which has as best blast hits proteins from *Z. nevadensis* cuticle CS (ZnCSC), from residue 1 to 1530 and midgut CS (ZnCSM), from residue 1577 to 2642. In agreement with this, our reads from *P. americana* midgut tissue were almost exclusively aligned to the second portion of the predicted gene (Supplementary Fig. 2). Furthermore, the anterior half of the gene sequence (putative PaCSC) branches with other CSC sequences, whereas the posterior half (putative PaCSM), with CSM sequences (Fig. 1). This means that in the annotation of the genome of *P. americana* chitin synthase genes were predicted as being fused. Based on these observations, we concluded that all analyzed Polyneoptera species have both CS genes.

In *P. americana* (Dictyoptera), 41 genes coding for putative CPAP/PMP/UCBP proteins were identified in its genome. The expression profile of these genes in midgut (MG) and salivary gland (SG) tissue samples were analyzed by RNA-seq strategy and are displayed in Table 3. Seven PMP and eight CPAP sequences were identified in this analysis. Except for one CPAP gene (PaCPAP6), all proteins were composed by one to three CBD domains.

In *A. flavolineata* (Orthoptera), 23 PMP sequences with up to four CBDs and two chitin synthase genes (AfCSC and AfCSM) were previously described by our group in the insect *de novo* transcriptome assembly (Dias et al., 2018). In the present work, we searched and found two CPAP proteins in this dataset, with two and three predicted CBD domains (Supplementary Table 3).

3.3. Chitin synthase, CPAP, PMP, and UCBP genes in Holometabola insects

In *M. domestica* (Diptera), 22 PMP sequences and two chitin synthase genes (MdCSC and MdCSM) were previously identified by our group (Dias et al., 2018). In the present work, six proteins with TPM values at least 20 times higher in carcass than in midgut CPAP) were found in the species genome (Supplementary Table 4), all of them possessing one to three predicted CBD domains. We also identified five UCBP proteins in this analysis with one to four predicted CBD domains.

In *S. frugiperda* (Lepidoptera), 38 PMP and two chitin synthase (SfCSC and SfCSM) sequences were previously described by our group in the insect *de novo* transcriptome assembly (Dias et al., 2018). In the present work, four putative CPAP and one UCBP sequence were found in this dataset, all of them with one or three predicted CBD domains (Supplementary Table 5).

In *T. molitor* (Coleoptera), 16 PMP and two chitin synthase (TmCSC and TmCSM) sequences were previously described by our group (Dias et al., 2018). In the present work, ten putative CPAP and two UCBP protein sequences were identified in the insect *de novo* transcriptome assembly, all of them having one to three predicted CBD domains (Supplementary Table 6).

3.4. Chitin synthase, CPAP, PMP, and UCBP genes in Paraneoptera insects

In Paraneoptera, only *P. humanus* possesses two complete CS genes annotated in its genome. The other species have a single gene (Table 2) with the CSC genes from *Z. nevadensis, M. domestica, S. frugiperda* or *T. castaneum* as best blast hits.

In *D. peruvianus* (Hemiptera, Prosorrhyncha), seven putative CPAP sequences were identified in its *de novo* transcriptome assembly (Table 4). All these sequences have up to three CBD domains. Furthermore, only one chitin synthase gene (DpCSC) was identified in the species transcriptome and it was expressed only in the carcass (Table 4). These results agree with the absence of PM in this species, which have instead perimicrovillar membranes (Silva et al., 1995), seen in Fig. 2A. Those membranes are lipoproteic membranes that contain glycoproteins, but not enough to maintain their WGA-FITC fluorescence on the addition of excess N-acetylglucosamine (Fig. 2B).

In *R. prolixus* (Hemiptera, Prosorrhyncha), 31 putative CPAP/PMP/UCBP sequences were found in its genome (Table 2). The expression profile of these proteins in the anterior (AM) and posterior (PM) midgut and in the cuticle-containing rectum (REC) tissues are presented in the Table 5. Two of these proteins were classified as CPAPs as they were detected in rectum but were absent from midgut. Nineteen sequences were classified as UCBP, as they were detected in both midgut and

Table 3
Periplaneta americana CPAP, PMP, and chitin synthase expression.

Protein	Accession ID	Length (AA)	Expression (TPM)		Number of CBDs	Domain structure
			MG	SG		
PaPMP1	PaOGS06982	498	9978	–	1	CM
PaPMP2	PaOGS05622	424	1723	–	2	CMCM
PaPMP3	PaOGS29773	987	1400	–	1	CM
PaPMP4	PaOGS12742	268	121	–	2	CMC
PaPMP5	PaOGS12221	421	71	–	1	CM
PaPMP6	PaOGS09645	1317	58	–	2	CMCM
PaPMP7	PaOGS26537	267	50	–	1	MC
PaCPAP1	PaOGS07029	149	–	22,778	2	CMC
PaCPAP2	PaOGS07027	159	–	2466	1	MCM
PaCPAP3	PaOGS07026	175	–	1832	1	MC
PaCPAP4	PaOGS07030	141	–	1654	1	C
PaCPAP5	PaOGS26524	432	–	228	3	CMCMC
PaCPAP6	PaOGS35822	1852	–	67	4	MCMCMCMC
PaCPAP7	PaOGS03017	132	–	32	1	C
PaCPAP8	PaOGS05547	240	–	5	3	CCC
Chitin synthase						
PaCS	PaOGS16634		80.9	0.1		

Only sequences with at least five TPMs of expression are presented on the table. "-", TPM values less than five; C, CBD; M, mucin domain; MG, midgut; SG, salivary gland. Sequences were classified as: PMP, when presenting TPM values at least 20 times higher in MG than in SG, and as CPAP, when presenting TPM values at least 20 times higher in SG than in midgut.

rectum tissues.

In *M. fimbriolata* (Hemiptera, Auchenorryncha), eight putative CPAP/UCBP genes were described in its *de novo* transcriptome assembly (Table 6). All these proteins have one or three predicted CBDs. Five of these proteins have TPM values at least twenty times higher in carcass than in the filter chamber (FC) and in the midgut tissues (A, M, and P). However, besides the absence of PM and chitin synthase expression in the midgut of this insect, three UCBP proteins with three CBD domains showed high TPM values in at least one midgut section. *M. fimbriolata* UCBPs are closer to CPAPs than to PMPs, since they branch together with *T. castaneum* CPAPs in a cladogram predicted using sequences of the first CBD of each of these proteins (Supplementary Fig. 3). *M. fimbriolata* microvilli form bundles enclosed by a membrane (Fig. 2C). The microvilli bundle-forming membranes have glycoconjugates which fluorescence with WGA-FITC is reduced but not abolished in the presence of excess N-acetyl glucosamine (Fig. 2D)

4. Discussion

4.1. Chitin synthase distribution among Polyneoptera, Paraneoptera and Holometabola species and the adaptation of insects to suck plant sap

A peritrophic membrane (PM) was described in the majority of the insect lineages, except for some insects from Paraneoptera (Peters, 1992). As a PM occurs in the earlier divergent Polyneoptera group, it is believed that its absence in Paraneoptera species is due to its loss in their lineage ancestor (Terra, 2001).

In insects with PM, two chitin synthase (CS) genes are usually present, one of them involved in the cuticle (CSC) and the other in the PM formation (CSM). For this, the CSC gene is most expressed in carcass, whilst the CSM gene is most expressed in midgut tissues. In agreement with this, in the present work, two CS genes were found in Holometabola and Polyneoptera species, whereas only one CS gene was found in the analyzed Paraneoptera species, except for *P. humanus* (Phthiraptera) (Table 2). This last observation deserves further discussion. Inside the Paraneoptera group, besides the PM absence, Hemiptera and Thysanoptera species are characterized by having microvilli-associated membranes (Burgos and Gutiérrez, 1976; Lane and Harrison, 1979; Del Bene et al., 1991; Silva et al., 1995; Cristofoletti et al., 2003; Utiyama et al., 2016). In contrast, no microvilli-associated membranes were described in *P. humanus* midgut. Instead, this species is supposed to have a peritrophic gel covering the midgut surface (Silva et al., 2004). This gel-like layer is observed in several insect groups and is not associated to chitin (Terra, 2001). So, the role of these two CS genes in Phthiraptera is yet unclear. Actually, the evolutionary history of Phthiraptera is different from Condylognata (Hemiptera, true bugs plus Thysanoptera, thrips). Until recently, the Paraneoptera orders, which includes Psocodea (Psocoptera, bark lice plus Phthiraptera, true lice)

Table 4
D. peruvianus CPAP and chitin synthase expression values.

Protein	Length (AA)	Expression (TPM)				Num. of CBD	Comple-teness	Best hit		
		A	M	P	CAR			*Species*	Iden-tity	Domain structure
DpCPAP1	194	–	–	–	17	1	NtF	*H. halys*	72.3	MC
DpCPAP3	791	–	–	–	10	1	NtF	*H. halys*	57.3	CMC
DpCPAP4	231	–	–	–	234	3	CtF	*H. halys*	88.6	CCCM
DpCPAP5	330	–	–	–	57	2	NtF	*H. halys*	86.8	CC
DpCPAP6	260	–	–	–	34	3	NtF	*H. halys*	90.4	CCC
DpCPAP7	274	–	–	–	13	3	Full	*H. halys*	80.6	CCC
DpCPAP8	308	–	–	–	10	2	Full	*H. halys*	76.3	CCC
Chitin synthase										
DpCSC	1579	0.0	0.0	0.0	8.8		Full			

Only sequences with at least five TPMs of expression are shown in the table. "-", TPM values less than five; C, CBD; M, mucin domain; A, M and P, anterior, middle and posterior portions of the midgut, respectively; CAR, carcass (insect less midgut); Full, complete sequences; NtF, N-terminal fragments; CtF, C-terminal fragments. Sequences were classified as CPAP, when presenting TPM values at least 20 times higher in carcass than in any midgut tissue.

Table 5
R. prolixus CPAP, UCBP, and chitin synthase expression values.

Protein	Accession ID	Length (AA)	Expression (TPM)			Number of CBDs	Domain structure
			A	P	REC		
RpUCBP1	RPRC011432-PA	243	4878	6243	5033	1	CM
RpUCBP2	RPRC002963-PA	231	3221	2270	3048	1	C
RpUCBP3	RPRC008497-PA	130	1777	1032	1800	1	MC
RpUCBP4	RPRC013809-PA	143	1715	1278	1338	1	MC
RpUCBP5	RPRC003885-PA	263	1102	1552	738	1	C
RpUCBP6	RPRC014043-PA	123	921	556	749	1	C
RpUCBP7	RPRC005415-PA	414	708	522	404	1	MC
RpUCBP8	RPRC015128-PA	270	692	801	462	1	C
RpUCBP9	RPRC002964-PA	99	517	139	54	1	C
RpUCBP10	RPRC013481-PA	229	185	4064	1609	1	C
RpUCBP11	RPRC003944-PA	252	170	212	246	1	C
RpUCBP12	RPRC000663-PA	142	–	214	207	1	C
RpUCBP13	RPRC006446-PA	251	–	14	11	1	C
RpCPAP1	RPRC001569-PA	206	–	–	29	1	MC
RpUCBP14*	RPRC013478-PA	267	–	175	–	1	MC
RpUCBP15	RPRC005413-PA	712	375	761	634	2	MCMCM
RpUCBP16	RPRC010253-PA	231	–	14	148	2	CC
RpUCBP17	RPRC013059-PA	239	15	16	157	3	CCC
RpUCBP18	RPRC000042-PA	258	–	46	53	3	CCCM
RpCPAP2	RPRC013845-PA	223	–	–	48	3	CMCC
RpUCBP19	RPRC010255-PA	474	13	40	31	6	MCCCCCMC
Chitin synthase							
RpCSC**	RPRC008031-PA	1495	0	0	0		

Only sequences with at least five TPMs of expression are shown in the table. "-", TPM values less than five; C, CBD; M, mucin domain; A, anterior midgut; P, posterior midgut; REC, rectum (part of the hindgut). Sequences were classified as CPAP, when expressed in rectum but not in any midgut tissue; and as UCBP, when expressed in both midgut (A or P) and rectum tissues.

* RpUCBP14 gene had a similar expression value in P and in a whole insect sample (175 and 170 TPMs, respectively), which made us believe that this gene may also be expressed on carcass (insect less midgut) tissue.

** RpCSC gene had an estimated TPM value of 19.8 in a whole insect sample, which may represent carcass.

and Condylognatha were thought to form a monophyletic clade (Grimaldi and Engel, 2005). At the present, phylogenomics data revealed that Condylognatha diverged from a grouping formed by Psocodea + Holometabola, which later on separated into Psocodea and Holometabola (Misof et al., 2014). Condylognatha divergence may have resulted from the adaption to suck plant sap described above, whereas Psocodea is the sister grouping of Holometabola, with an evolutionary history to be described that could explain their differences in relation to Codylognatha.

4.2. Peritrophins and chitin in the midgut of Hemiptera

Hemipterans lack a PM (Peters, 1992) but have as extracellular layers microvilli-associated membranes, which are the lipoproteic membranes described before.

The existence of some kind of PM in hemipteran midguts was suggested by Alvarenga et al. (2016), based on the finding of peritrophins in *Rhodnius prolixus* (Ribeiro et al., 2014) and of chitin in their midguts, despite the fact that chitin is mainly at the hemal side of the midgut, probably corresponding to tracheolae. The midgut morphological alterations reported after chitin synthase suppression (Alvarenga et al., 2016) may have resulted from hypoxia as a consequence of a decrease in tracheolae.

The hypothesis of the existence of PM-like structures in Hemiptera is not supported by the discovery that there is only a cuticle-type chitin synthase in Hemiptera genomes (including *R. prolixus*) and that

Table 6
M. fimbriolata CPAP, UCBP and chitin synthase expression values.

Protein	Length (AA)	Expression (TPM)					Number of CBDs	Comple-teness	Best hit		
		FC	A	M	P	CAR			Species	Iden-tity	Domain structure
MfUCBP1	253	–	–	–	22	34	3	NtF	*N. lugens*	70.1	CCCM
MfUCBP2	266	–	–	–	17	70	3	NtF	*H. hallys*	82.5	CCC
MfUCBP3	289	–	–	5	–	51	3	NtF	*N. lugens*	77.8	CCC
MfCPAP1	270	–	–	–	–	78	3	Full	*H. vitripennis*	88.5	CC
MfCPAP2	289	–	–	–	–	74	3	Full	*H. vitripennis*	78.1	CCM
MfCPAP3	303	–	–	–	–	36	3	Full	*N. lugens*	71.1	MCCC
MfCPAP4	114	–	–	–	–	5	1	Inc	*D. citri*	85.5	C
MfCPAP5	255	–	–	–	–	16	3	NtF	*H. vitripennis*	79.0	CCMC
Chitin synthase											
MfCSC	77	0.0	0.0	0.2	0.0	2.1		Inc			
MfCSC	73	0.0	0.3	0.0	0.3	5.1		Inc			
MfCSC	52	0.0	0.0	0.3	0.0	3.1		Inc			

Only sequences with at least five TPMs of expression are presented in the table. "-", TPM values less than five; C, CBD; M, mucin domain; FC, filter chamber; A, M and P, anterior, middle and posterior regions of the midgut, respectively; CAR, carcass (insect less midgut); Full, complete sequences; NtF, N-terminal fragments; and Inc, fragment without N- and C-terminals. Sequences were classified as: CPAP, when presenting TPM values at least 20 times higher in carcass than in any midgut tissue, and as UCBP for the remaining cases.

peritrophins expressed in Hemiptera midguts (including *R. prolixus*) are not PMP. Furthermore, WGA-FITC fluorescence at the apex of *R. prolixus* midgut cells is reduced by sialic acid (Albuquerque-Cunha et al., 2009), whereas for *D. peruvianus* and *M. fimbriolata* (this paper) fluorescence is reduced by N-acetylglucosamine. Fluorescence reduction with these additions indicates the existence of glycoconjugates, not chitin. Glycoconjugates were actually biochemically and cytochemically characterized at the midgut luminal surface *R. prolixus* (Alves et al., 2007; Albuquerque-Cunha et al., 2009).

4.3. Differences between CPAP, PMP and UCBP composition among insects

CPAP, PMP and UCBP proteins are composed of at least one chitin binding Peritrophin-A domain (CBD). CPAP proteins have been described as proteins composed by one or three of these domains (Jasrapuria et al., 2012). In our work, genes were considered as coding for putative CPAPs when having TPM values at least 20 times higher in carcass than in any midgut tissue. According to this criterion, only *S. frugiperda, R. prolixus,* and *M. fimbriololata* have all putative CPAP sequences with one or three identified CBDs.

Sequences with more than three predicted CBD are thought to be involved in PM formation. Our group previously reported these sequences as putative PMP in *M. domestica, S. frugiperda, T. molitor* and *A. flavolineata* (Dias et al., 2018). It is curious, however, that we found genes coding for proteins with more than three CBDs in almost all analyzed Paraneoptera genomes (14 of 18), in spite of the absence of PM in these insects (Table 1). Moreover, one of these sequences was expressed in all *R. prolixus* tissues (Table 5). *R. prolixus* expression data is yet more unusual, as genes coding for proteins with one to three CBDs were also expressed in its midgut and rectum, the last one displaying a cuticle as all hindgut structures.

Genes with similar expression values in both midgut and cuticle-containing tissues were commonly found in our work for different species and motivated us to create a new protein class, the ubiquitous-CBD-containing proteins (UCBP). UCBP proteins were identified in the Holometabola: *M. domestica* (5), *S. frugiperda* (1), and *T. molitor* (2), and in the Paraneoptera: *R. prolixus* (19), and *M. fimbriolata* (3) species. As UCBP resembles more CPAPs than PMPs (at least in *M. fimbriolata*) we may speculate they are involved in tracheolae formation in midgut and other tissues (included in carcass samples). However, the *M. domestica* gene (MdCPAP5) homologous to the tracheolae peritrophin-like gene of *D. melanogaster* (Gasp) (Barry et al., 1999) is expressed only in *M. domestica* carcass (Supplementary Table 4). Although this does not necessarily discard a role of UCBP in tracheolae, it does not favor our hypothesis. CPAP-coding genes with one and three predicted CBDs were expressed in the fat body of *Bactrocera dorsalis*, suggesting that they may be involved in insect immunity (Chen et al., 2018) and/or in tracheolae formation in this tissue.

Genes coding for proteins with only one CBD were the most abundant putative CPAP/PMP/UCBP in all the analyzed genomes (Table 2). These proteins were also expressed in *P. americana* midgut and salivary glands, *D. peruvianus* carcass, *M. fimbriolata* carcass; and in all analyzed tissues from *R. prolixus*. Some proposed roles for CPAP/PMP proteins with only one CBD includes: a) capping the ends of chitin fibrils, helping the PMP and cuticle formation (Devenport et al., 2005); and b) sequestering free chitin molecules from the midgut lumen (Jochim et al., 2008). However, the exact role of these proteins is still unclear, as only three of the ten analyzed CPAP1 genes in *T. castaneum* exhibited lethal phenotypes in RNAi experiments and were required in the pupal-to-adult molt (Jasrapuria et al., 2012). Furthermore, besides the high similarity to the CBD peritrophin-A domain, some of the identified domains may not bind chitin or may act as in the *Drosophila melanogaster* mind-the-gap (MTG) protein (Uniprot ID: A0A0C4DHF5). This extracellular protein is composed of only a signal peptide and one CBD and acts in the synapse assembly (Rushton et al., 2009).

5. Conclusions

The results showed that on adapting to sap feeding, Paraneopterans (actually represented here by the Hemiptera) lost PM and associated with this also lost the CSM and PMP genes, whereas maintaining CPAP and UCBP genes. The last genes are arguably involved in tracheolae formation; however, this demands further research.

Holometabola species have a higher number of PMP proteins composed by more than three CBDs in relation to species from the earlier divergent Polyneoptera group. This observation can be related to different PM characteristics for these two cohorts and/or be a result of an adaptation to a peptidase-rich gut content (Wang et al., 2004).

Acknowledgements

This work was supported by the Brazilian research agencies FAPESP (Grant “Tematico” no. 2017/08103-4) and CNPq. We thank M. V. Cruz and W. Caldeira for technical assistance and an anonymous referee for the suggestion of naming ubiquitous chitin-binding-containing proteins (UCBP) those proteins containing a CBD expressed both at midgut and carcass. Camila S. Leal is a graduate fellow of CAPES. Renata O. Dias is a post-doctoral fellow of FAPESP. C. Cardoso, A. F. Ribeiro, C. Ferreira, and W.R. Terra are staff members of their respective departments. C. Ferreira and W.R. Terra are also research fellows of the CNPq and members of the Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular (INCT-EM).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jinsphys.2019.02.002.

References

Albuquerque-Cunha, J.M., Gonzales, M.S., Garcia, E.S., Mello, C.B., Azambuja, P., Almeida, J.C.A., de Souza, W., Nogueira, N.F.S., 2009. Cytochemical characterization of microvillar and perimicrovillar membranes in the posterior midgut epithelium of *Rhodnius prolixus*. Arthropod Struct. Dev. 38, 31–44.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Alvarenga, E.S., Mansur, J.F., Justi, S.A., Figueira-Mansur, J., Dos Santos, V.M., Masuda, H., Lara, F.A., Melo, A.C., Moreira, M.F., 2016. Chitin is a component of the *Rhodnius prolixus* midgut. Insect Biochem. Mol. Biol. 69, 61–70.

Alves, C.R., Albuquerque-Cunha, J.M., Mello, C.B., Garcia, E.S., Nogueira, N.F., Bourguingnon, S.C., de Souza, W., Azambuja, P., Gonzales, M.S., 2007. *Trypanosoma cruzi*: attachment to perimicrovillar membrane glycoproteins of *Rhodnius prolixus*. Exp. Parasitol. 116, 44–52.

Barbehenn, R.V., 2001. Roles of peritrophic membranes in protecting herbivorous insects from ingested plant allelochemicals. Arch. Insect Biochem. Physiol. 47, 86–99.

Barry, M.K., Triplett, A.A., Christensen, A.C., 1999. A peritrophin-like protein expressed in the embryonic tracheae of *Drosophila melanogaster*. Insect Biochem. Mol. Biol. 29, 319–327.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2114–2120.

Bolognesi, R., Terra, W.R., Ferreira, C., 2008. Peritrophic membrane role in enhancing digestive efficiency. Theoretical and experimental models. J. Insect Physiol. 54, 1413–1422.

Burgos, M.H., Gutiérrez, L.S., 1976. The intestine of *Triatoma infestans*. I. Cytology of the midgut. J. Ultrastruct. Res. 57, 1–9.

Caldeira, W., Dias, A.B., Terra, W.R., Ribeiro, A.F., 2007. Digestive enzyme compartmentalization and recycling and sites of absorption and secretion along the midgut of Dermestes maculates (Coleoptera) larvar. Arch. Insect Biochem. Physiol. 64, 1–18.

Chen, E.H., Hou, Q.L., Dou, W., Wei, D.D., Yue, Y., Yang, R.L., Yang, P.J., Yu, S.F., De Schutter, K., Smagghe, G., Wang, J.J., 2018. Genome-wide annotation of cuticular proteins in the oriental fruit fly (Bactrocera dorsalis), changes during pupariation and expression analysis of CPAP3 protein genes in response to environmental stresses. Insect Biochem. Mol. Biol. 97, 53–70.

Cristofoletti, P.T., Ribeiro, A.F., Deraison, C., Rahbé, Y., Terra, W.R., 2003. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrtosiphon pisum*. J. Insect Physiol. 49, 11–24.

Del Bene, G., Dallai, R., Marchini, D., 1991. Ultrastructure of the midgut and the adhering tubular salivary glands of Frankliniella occidentalis (Pergande) (Thysanoptera: Thripidae). Int. J. Insect Morphol. Embryol. 20, 15–24.

Devenport, M., Fujioka, H., Donnelly-Doman, M., Shen, Z., Jacobs-Lorena, M., 2005. Storage and secretion of Ag-Aper14, a novel peritrophic matrix protein, and Ag-Muc1

from the mosquito Anopheles gambiae. Cell Tissue Res. 320, 175–185.
Dias, R.O., Cardoso, C., Pimentel, A.C., Damasceno, T.F., Ferreira, C., Terra, W.R., 2018. The roles of mucus-forming mucins, peritrophins and peritrophins with mucin domains in the insect midgut. Insect Mol. Biol. 27, 46–60.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
Evans, J.D., Brown, S.J., Hackett, K.J., Robinson, G., Richards, S., Lawson, D., Elsik, C., Coddington, J., Edwards, O., Emrich, S., Gabaldon, T., Goldsmith, M., Hanes, G., Misof, B., Munoz-Torres, M., Niehuis, O., Papanicolaou, A., Pfrender, M., Poelchau, M., Purcell-Miramontes, M., Robertson, H.M., Ryder, O., Tagu, D., Torres, T., Zdobnov, E., Zhang, G.J., Zhou, X.I.K.C., 2013. The i5K Initiative: Advancing Arthropod Genomics for Knowledge, Human Health, Agriculture, and the Environment i5K CONSORTIUM. J. Hered. 104, 595–600.
Giraldo-Calderon, G.I., Emrich, S.J., Maccallum, R.M., Maslen, G., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Madey, G., Collins, F.H., Lawson, D., Vectorbase, C., 2015. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Res. 43, D707–D713.
Grimaldi, D., Engel, M.S., 2005. Evolution of Insects. Cambridge University Press, Cambridge.
Jasrapuria, S., Arakane, Y., Osman, G., Kramer, K.J., Beeman, R.W., Muthukrishnan, S., 2010. Genes encoding proteins with peritrophin A-type chitin-binding domains in Tribolium castaneum are grouped into three distinct families based on phylogeny, expression and function. Insect Biochem. Mol. Biol. 40, 214–227.
Jasrapuria, S., Specht, C.A., Kramer, K.J., Beeman, R.W., Muthukrishnan, S., 2012. Gene Families of Cuticular Proteins Analogous to Peritrophins (CPAPs) in Tribolium castaneum Have Diverse Functions. Plos One, 7.
Jochim, R.C., Teixeira, C.R., Laughinghouse, A., Mu, J., Oliveira, F., Gomes, R.B., Elnaiem, D.E., Valenzuela, J.G., 2008. The midgut transcriptome of Lutzomyia longipalpis: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and Leishmania infantum chagasi-infected sand flies. BMC Genomics 9, 15.
Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W.Z., Mcanulla, C., Mcwilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.
Kersey, P.J., Allen, J.E., Allot, A., Barba, M., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z.C., Maurel, T., Moore, B., Mcdowall, M.D., Maheswari, U., Naamati, G., Newman, V., Ong, C.K., Paulini, M., Pedro, H., Perry, E., Russell, M., Sparrow, H., Tapanari, E., Taylor, K., Vullo, A., Williams, G., Zadissia, A., Olson, A., Stein, J., Wei, S.R., Tello-Ruiz, M., Ware, D., Luciani, A., Potter, S., Finn, R.D., Urban, M., Hammond-Kosack, K.E., Bolser, D.M., De Silva, N., Howe, K.L., Langridge, N., Maslen, G., Staines, D.M., Yates, A., 2018. Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. Nucleic Acids Res. 46, D802–D808.
Kuraishi, T., Binggeli, O., Opota, O., Buchon, N., Lemaitre, B., 2011. Genetic evidence for a protective role of the peritrophic matrix against intestinal bacterial infection in Drosophila melanogaster. PNAS 108, 15966–15971.
Lane, N.J., Harrison, J.B., 1979. An unusual cell surface modification: a double plasma membrane. J. Cell Sci. 39, 355–372.
Li, S., Zhu, S.M., Jia, Q.Q., Yuan, D.W., Ren, C.H., Li, K., Liu, S.N., Cui, Y.Y., Zhao, H.G., Cao, Y.H., Fang, G.Q., Li, D.Q., Zhao, X.M., Zhang, J.Z., Yue, Q.Y., Fan, Y.L., Yu, X.Q., Feng, Q.L., Zhan, S., 2018. The genomic and functional landscapes of developmental plasticity in the American cockroach. Nat. Commun. 9.
Mesquita, R.D., Vionette-Amaral, R.J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F.A., Minx, P., Spieth, J., Carvalho, A.B., Panzera, F., Lawson, D., Torres, A.Q., Ribeiro, J.M.C., Sorgine, M.H.F., Waterhouse, R.M., Montague, M.J., Abad-Franch, F., Alves-Bezerra, M., Amaral, L.R., Araujo, H.M., Araujo, R.N., Aravind, L., Atella, G.C., Azambuja, P., Berni, M., Bittencourt-Cunha, P.R., Braz, G.R.C., Calderon-Fernandez, G., Carareto, C.M.A., Christensen, M.B., Costa, I.R., Costa, S.G., Dansa, M., Daumas, C.R.O., De-Paula, I.F., Dias, F.A., Dimopoulos, G., Emrich, S.J., Esponda-Behrens, N., Fampa, P., Fernandez-Medina, R.D., Da Fonseca, R.N., Fontenele, M., Fronick, C., Fulton, L.A., Gandara, A.C., Garcia, E.S., Genta, F.A., Giraldo-Calderon, G.I., Gomes, B., Gondim, K.C., Granzotto, A., Guarneri, A.A., Guigo, R., Harry, M., Hughes, D.S.T., Jablonka, W., Jacquin-Joly, E., Juarez, M.P., Koerich, L.B., Latorre-Estivalis, J.M., Lavore, A., Lawrence, G.G., Lazoski, C., Lazzari, C.R., Lopes, R.R., Lorenzo, M.G., Lugon, M.D., Majerowicz, D., Marcet, P.L., Mariotti, M., Masuda, H., Megy, K., Melo, A.C.A., Missirlis, F., Mota, T., Noriega, F.G., Nouzova, M., Nunes, R.D., Oliveira, R.L.L., Oliveira-Silveira, G., Ons, S., Pagola, L., Paiva-Silva, G.O., Pascual, A., Pavan, M.G., Pedrini, N., Peixoto, A.A., Pereira, M.H., Pike, A., Polycarpo, C., Prosdocimi, F., Ribeiro-Rodrigues, R., Robertson, H.M., Salerno, A.P., Salmon, D., Santesmasses, D., Schama, R., Seabra, E.S., Silva-Cardoso, L., Silva-Neto, M.A.C., et al., 2015. Genome of Rhodnius prolixus, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. PNAS 112, 14936–14941.
Misof, B., Liu, S., Meusemann, K., Peters, R.S., et al., 2014. Phylogenomics resolves the timing and pattern ofinsect evolution. Science 346, 763–767.
O'leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., Mcveigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., Mcgarvey, K.M., Murphy, M.R., O'neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S. S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., Dicuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D., Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44 (2016) D733–45.
Parra, J.R.P., 1986. Criação de insetos para estudos com patógenos. In: Alves, S.B. (Ed.), Controle Microbiano de Insetos. São Paulo, Editora Manole, pp. 348–373.
Peters, W., 1992. Peritrophic membranes. Berlin: Springer-Verlag. 23.
Ribeiro, J.M., Genta, F.A., Sorgine, M.H., logullo, R., et al., 2014. An insight into the transcriptome of the digestive tract of the blood sucking bug, Rhodnius prolixus. PLOS Neglected Trop. Dis. 8: e2594.
Rushton, E., Rohrbough, J., Broadie, K., 2009. Presynaptic Secretion of Mind-the-Gap Organizes the Synaptic Extracellular Matrix-Integrin Interface and Postsynaptic Environments. Dev. Dyn. 238, 554–571.
Silva, C.P., Ribeiro, A.F., Gulbenkian, S., Terra, W.R., 1995. Organization, origin and function of the outer microvillar (perimicrovillar) membranes of Dysdercus peruvianus (Hemiptera) midgut cells. J. Insect Physiol. 41, 1093–1103.
Silva, C.P., Silva, J.R., Vasconcelos, F.F., Petretski, M.D.A., Damatta, R.A., Ribeiro, A.F., Terra, W.R., 2004. Occurrence of midgut perimicrovillar membranes in para-neopteran insect orders with comments on their function and evolutionary significance. Arthropod Struct. Dev. 33, 139–148.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Targa, H.J., Peres, C.A., 1979. Radiation-induced dominant lethal mutations in oocytes of *Musca domestica*. Mutat. Res. 63, 153–160.
Tetreau, G., Dittmer, N.T., Cao, X., Agrawal, S., Chen, Y.-R., Muthukrishnan, S., Haobo, J., Blissard, G.W., Kanost, M.R., Wang, P., 2015. Analysis of chitin-binding proteins from Manduca sexta provides new insights into evolution of peritrophin A-type chitin-binding domains in insects. Insect Biochem. Mol. Biol. 62, 127–141.
Terra, W.R., 1990. Evolution of insect digestive systems. Annu. Rev. Entomol. 35, 181–200.
Terra, W.R., 1988. Physiology and biochemistry of insect digestion: an evolutionary perspective. Braz. J. Med. Biol. Res. 21, 675–734.
Terra, W.R., 2001. The origin and functions of the insect peritrophic membrane and peritrophic gel. Arch. Insect Biochem. Physiol. 47, 47–61.
Utiyama, A.H., Terra, W.R., Ribeiro, A.F., 2016. The digestive system of the leafhopper Bucephalogonia xanthophis (Hemiptera, Cicadellidae): the organization of the luminal membrane complex. J. Entomol. Res. 40, 339–346.
Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285.
Wang, P., Li, G., Granados, R.R., 2004. Identification of two new peritrophic membrane proteins from larval *Trichoplusia ni*: structural characteristics and their functions in the protease rich insect gut. Insect Biochem. Mol. Biol. 34, 215–227.
Zakrzewski, A.C., Weigert, A., Helm, C., Adamski, M., Adamska, M., Bleidorn, C., Raible, F., Hausen, H., 2014. Early Divergence, Broad Distribution, and High Diversity of Animal Chitin Synthases. Genome Biol. Evol. 6, 316–325.