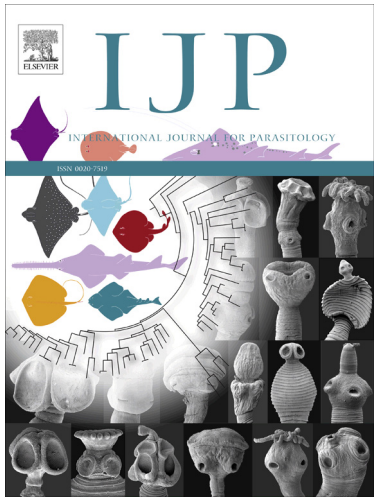


Transcriptome and toxin family analysis of the paralysis tick, *Ixodes holocyclus*

Manuel Rodriguez-Valle, Paula Moolhuijzen, Roberto A. Barrero, Chian Teng Ong, Greta Busch, Thomas Karbanowicz, Mitchell Booth, Richard Clark, Johannes Koehbach, Hina Ijaz, Kevin Broady, Kim Agnew, Aleta G. Knowles, Matthew I. Bellgard, Ala E. Tabor

PII: S0020-7519(17)30284-9
DOI: https://doi.org/10.1016/j.ijpara.2017.07.007
Reference: PARA 4000

To appear in: *International Journal for Parasitology*

Received Date: 26 May 2017
Revised Date: 26 July 2017
Accepted Date: 31 July 2017

Please cite this article as: Rodriguez-Valle, M., Moolhuijzen, P., Barrero, R.A., Ong, C.T., Busch, G., Karbanowicz, T., Booth, M., Clark, R., Koehbach, J., Ijaz, H., Broady, K., Agnew, K., Knowles, A.G., Bellgard, M.I., Tabor, A.E., Transcriptome and toxin family analysis of the paralysis tick, *Ixodes holocyclus*, *International Journal for Parasitology* (2017), doi: https://doi.org/10.1016/j.ijpara.2017.07.007

Transcriptome and toxin family analysis of the paralysis tick, *Ixodes holocyclus*

Manuel Rodriguez-Valle[a,f,*], Paula Moolhuijzen[b,g], Roberto A. Barrero[b], Chian Teng Ong[a], Greta Busch[a], Thomas Karbanowicz[a], Mitchell Booth[a], Richard Clark[c], Johannes Koehbach[c], Hina Ijaz[c], Kevin Broady[d], Kim Agnew[e,h], Aleta G. Knowles[e,i], Matthew I. Bellgard[b], Ala E. Tabor[a,b,*].

[a]*Queensland Alliance for Agriculture & Food Innovation, The University of Queensland, St. Lucia, Queensland 4072, Australia*

[b]*Centre for Comparative Genomics, Murdoch University, Perth, WA 6150, Australia*

[c]*School of Biomedical Sciences, The University of Queensland, St Lucia, Queensland 4072, Australia*

[d]*University of Technology Sydney, Ultimo, NSW 2007, Australia*

[e]*Elanco Animal Health, West Ryde, NSW 2114, Australia*

[f]*Present address: Department of Veterinary Biosciences, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria 3010, Australia. Affiliation with The University of Queensland*

[g]*Present address: Faculty of Science and Engineering, Curtin University, WA 6102, Australia*

[h]*Present address: Boehringer Ingelheim Animal Health, Sydney, Australia*

[i]*Present address: Virbac Aust. Pty. Ltd., Sydney, Australia*

*Corresponding authors. Manuel Rodriguez-Valle and Ala Tabor, Queensland Alliance for Agriculture and Food Innovation (QAAFI), Level 3, Queensland Biosciences Precinct, 306 Carmody Rd, UQ, St Lucia, Qld 4072, Australia. Tel.: +61 7 3346 2176; fax: +61 7 3365 0555.

E-mail addresses: m.rodriguezvalle@uq.edu.au; a.lewtabor@uq.edu.au

Abstract

The Australian paralysis tick (*Ixodes holocyclus*) secretes neuropathic toxins into saliva that induce host paralysis. Salivary glands and viscera were dissected from fully engorged female *I. holocyclus* ticks collected from dogs and cats with paralysis symptoms. cDNA from both tissue samples were sequenced using Illumina HiSeq 100 bp pair end read technologies. Unique and non-redundant holocyclotoxin (HT) sequences were designated as HT2 to HT19, as none were identical to the previously described HT1. Specific binding to rat synaptosomes was determined for synthetic HTs, and their neurotoxic capacity was determined by neonatal mouse assay. They induced a powerful paralysis in neonatal mice, particularly HT4 which produced rapid and strong respiratory distress in all animals tested. This is the first known genomic database developed for the Australian paralysis tick. The database contributed to the identification and subsequent characterization of the holocyclotoxin family that will inform the development of novel anti-paralysis control methods.

Keywords: Paralysis tick; *Ixodes holocyclus*; Transcriptome; Neurotoxin; Mouse model; Holocyclotoxin; Salivary gland; Viscera

1. Introduction

Ticks are hematophagous obligate ectoparasites that belong to the Class Arachnida which includes spiders and scorpions. They impair the health of livestock, pets and humans directly via blood loss, or indirectly as vectors of tick-borne diseases (Jongejan and Uilenberg, 2004). The parasitic stage of the tick life cycle is initiated when feeding larvae overcome the host's hemostatic and immunological responses (Rodriguez-Valle et al., 2010, 2012). Tick saliva mediates the disruption of host defenses through a complex mixture of molecules such as serine protease inhibitors, proteases and lipocalins (Francischetti et al., 2009; Rodriguez-Valle et al., 2010; Ribeiro et al., 2011). Additionally, the saliva of some tick species have venomous characteristics due to the presence of toxins that induce host paralysis (Masina and Broady, 1999; Mans et al., 2004). Stone et al. (1989) suggested that paralyzing tick toxins' 'most likely primary function is as a feeding stimulant, with toxicity as a secondary, probably accidental, additional function'. It has been reported that 69 tick species from a total of 892 described species are associated with tick paralysis, and most are members of the Ixodidae (hard ticks) while only nine are members of the Argasidae (soft ticks). The most economically significant of these tick species for animal and human health are *Dermacentor andersoni* and *Dermacentor variabilis* in North America; *Ixodes rubicundus*, *Rhipicephalus evertsi evertsi* and *Argas* (*Persicargas*) *walkerae* in Africa; and *Ixodes holocyclus* (the Australian paralysis tick) in Australia (Masina and Broady, 1999; Mans et al., 2004).

Ixodes holocyclus is present along the eastern coast of Australia where high humidity and moderate temperatures underpin this geographic distribution. The number of domestic animals affected annually by tick paralysis ranges from 10,000 to 100,000 with a death rate of approximately 10% (Atwell et al., 2001; Hall-Mendelin et al., 2011; Vink et al., 2014). Additionally, there is increasing concern associated with the allergic reactions caused by this tick species in humans (Gauci et al., 1990). In 1966, a purified protein fraction from *I. holocyclus* was found to contain toxins which produced paralysis in dogs (Kaire, 1966). The toxicological potency of salivary *I. holocyclus* holocyclotoxins (HTs) was subsequently demonstrated in a biological assay using neonatal mice (Stone et al., 1982, 1983). In 1992, *I. holocyclus* neurotoxin activity was

shown using a rat synaptosome binding assay. Three neurotoxins (HTs) with molecular weights of approximately 5 kDa, named HT1, HT2 and HT3, were identified by isolating the proteins bound to rat synaptosomes, and a partial protein and full-length DNA sequence for HT1 was reported (Thurn, 1992). Other studies attempted to produce a recombinant variant of HT1, which exhibited immunogenic properties against toxin present in salivary gland (SG) extracts of *I. holocyclus* but no protection from paralysis symptoms in immunized animals (Masina and Broady, 1999; Hall-Mendelin et al., 2011). Twenty-one nucleotide sequences related to HTs were recently submitted to GenBank (https://www.ncbi.nlm.nih.gov/pubmed/23193287) and were linked to a patent application by Zoetis, Australia (WO 2014018724 A1). However, the patent did not report the biological activity of these toxins in animal models, which is presented here.

Tick genomes are typically large, highly repetitive and difficult to assemble, and few whole genome projects have been undertaken, with the *Ixodes scapularis* genome being the first tick genome available in 2016 (Gulia-Nuss, 2016). With the difficulty in obtaining genomic sequences, many researchers have relied on transcriptomic analyses to provide the needed genomic datasets to allow further research. Transcriptomic and proteomic analyses of SG extracts from different ticks species have been reported including *Amblyomma variegatum*; *Amblyomma americanum*, *D. variabilis*, *Ixodes ricinus* and *D. andersoni*. Numerous full-length transcripts of tick proteins and unique tick protein families have been elucidated based on these analyses. The *I. scapularis* genome is often used as a reference for further gene discovery associated with tick – host interactions. Limited *I. holocyclus* sequence data is available despite the medical and veterinary significance of this tick species. Recently, a preliminary study of the *I. holocyclus* sialotranscriptome was conducted using fully engorged *I. holocyclus* female ticks collected from dogs and cats (incidental hosts) with paralysis symptoms (Illumina HiSeq), and from bandicoots (primary hosts) (454 FLX Roche). The data from this study showed that the host had an effect on the expression of *I. holocyclus* transcripts (Ong, 2016).

We report for the first known time a transcriptome dataset containing 13,872 transcripts unique to viscera (VISC) and 6,095 transcripts unique to SGs of adult female *I. holocyclus* ticks.

Toxin-related sequence descriptions were identified for 72 transcripts within the dataset. Data showed that HTs are members of a multigene family that encodes proteins with variable amino acid sequences, but with a highly conserved inhibitor cysteine knot (ICK) motif (Vink et al., 2014). HTs have a neurotoxic effect induced by their direct interaction with cells of the host nervous system. Synthetic forms of these toxins were capable of binding to rat synaptosomes in vitro, and induced symptoms of paralysis in neonatal mice in vivo. This will lead to the development of new targeted paralysis treatments and preventative vaccines to protect animals from paralysis. This research significantly expands the genomic information available for *Ixodes* spp., and contributes to the characterization of the Australian paralysis tick HT family. This is the first known comprehensive report showing tick toxin discovery and biological activity of synthetic HTs.

2. Materials and methods

2.1. Tick collection and mRNA isolation

SGs and VISC were dissected from two pooled collections (two biological replicates) of ~25-30 fully engorged female ticks collected from different locations within known geographic areas of the *I. holocyclus* habitat (eastern Australian coast). Veterinary clinics across these areas participated in the collection of fully engorged ticks from cats and dogs (the majority of ticks were collected from dogs) with confirmed tick paralysis symptoms. RNA was prepared from dissected SGs and VISC disrupted by repetitive passage through a syringe in the presence of TRIzol® reagent. Total RNA was extracted according to the manufacturer's protocol (GibcoBRL, USA). mRNA was isolated utilising the Poly (A) Purist™ MAG Kit (AMBION, USA).

2.2. Transcript quality control and assembly

The cDNA libraries obtained from two biological replicates of SG and VISC samples were sequenced through the Australian Genome Research Facility Ltd (AGRF) using Illumina HiSeq 100 bp paired-end read technologies. Sequencing produced 65,035,631 paired reads for the SG and 76,180,419 for the VISC libraries. Quality was assessed using FastQC

(http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Content Dependent Trimmer (ConDeTri) to standardise the paired-end reads by trimming the low quality reads before assembly (Smeds and Kunstner, 2011). Samples were filtered for potential contamination using Bowtie2 and SAMtools (Li et al., 2009; Langmead and Salzberg, 2012), using bacterial genome (taxid:2) sequences downloaded from National Center for Biotechnology Information, USA (NCBI). The Illumina paired read data sets were assembled using Velvet's Oases (Schulz et al., 2012) across odd kmers ranging between 59 and 79.

2.3. Transcript annotation

The longest transcript/open reading frame (ORF) (start to stop codon) for each Oases assigned locus was selected as the representative locus sequence, annotated using the Automatic Functional Annotation and Classification Tool (AutoFACT V3.4) (Koski et al., 2005) decision matrix based on BLASTX 2.2.23 searches (Altschul et al., 1990) of Clusters of Orthologous Groups COG) (Tatusov et al., 2003), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2012), NCBI non-redundant protein (Wheeler et al., 2008), UniRef100 (Boutet et al., 2007), and UniRef90 (Boutet et al., 2007; Marchler-Bauer et al., 2009) databases. AutoFACT Reverse Position Specific BLAST (RPS-BLAST 2.2.23) (Marchler-Bauer et al., 2009; Finn et al., 2010) comparison of sequences was undertaken with conserved protein domains using the Protein Family Database (Pfam v.21.0) (Finn et al., 2010) and Simple Modular Architecture Research Tool (SMART v4.0) (Letunic et al., 2009) database; followed by the assignment of Gene Ontologies (GO) (Harris et al., 2004).

2.4. Creating a non-redundant dataset for each sample

The Cluster Database with High Identity with Tolerance (CD-HIT) was used to collapse the sequences with 100% identity to create non-redundant datasets for the salivary gland and viscera samples.

2.5. Protein clustering of orthologous groups

The longest translated ORFs were selected from EMBOSS version 6.3.1 (Rice et al., 2000) 'getorf' ORF translations of regions between the start and stop codons. The sample protein data sets were then clustered with OrthoMCL (Fischer et al., 2011) at 50% identity and 50% alignment coverage.

2.6. Secreted protein searches

Sample transcripts were analysed for the presence of signal peptides using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) showing D cut-off scores using default settings of 0.45 – no TM networks and 0.5 – TM networks. The transmembrane region numbers were determined using the TMHMM 2.0 server (http://www.cbs.dtu.dk/services/TMHMM/).

2.7. Gene expression analysis

Assembled transcripts for SG and VISC samples were merged using CD-HIT (Fu et al., 2012), using a minimal sequence similarity threshold of 95%. This yielded 68,065 representative transcripts that were used for expression analysis. High quality Illumina reads sequenced for SG and VISC samples were aligned onto 68,065 *I. holocyclus* representative transcripts using Bowtie 2 (Langmead and Salzberg, 2012). Aligned raw counts for each transcript were calculated using SAMtools (Li et al., 2009). Transcripts with • 1 read per million (RPM) mapped tags for at least three samples were selected for statistical analysis. Differentially expressed transcripts between SG and VISC samples were calculated using edgeR (Robinson et al., 2010). The tool eXpress (Roberts and Pachter, 2013) was used to calculate fragments per kilobase of exon per million mapped reads (FPKM).

2.8. HT sequence analysis

HT sequences identified within the transcript libraries and confirmed by protein BLAST were aligned with the previously reported HT1 protein sequence (GenBank Accession No.

AAV34602). HT sequences were assembled using ClustalW default parameters within Geneious Version 9.1.5 (Biomatters Ltd., New Zealand). The signal peptide cleavage sites of each HT were determined using the SignalP 4.1 Server. Potential glycosylation sites were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). Using the ClustalW alignment, a Jukes-Cantor distance model with 1000 bootstrap replicates was undertaken to construct a Neighbor-Joining tree using a branching support threshold of 70% (Geneious Version 9.1.5). Duplicate and incomplete HT protein sequences were removed from the alignment for tree building using the Versutoxin from the venom of the funnel-web spider *Atrax versutus* (GenBank Accession No. **P13494.1**) as an out-group (Brown et al., 1988). Twenty-one nucleotide sequences (Sequence 1 – Sequence 21) identified by the patent WO 2014018724 were translated into predicted protein sequences and those with full ORFs were included into the Neighbor-Joining tree analysis.

2.9. Synthesis of the HTs

HT1, HT2, HT3 and HT4 sequences were synthesized as the following sequences: SCTNPGKKRCNAKCSTHCDCKDGPTHNFGAGPVQCKKCTYQFKGEAYCKQ, SCSNPGKRNCNDPCRTHCDCIGGKKYDNGAGLVLCQKCTYQLGSKVGYCKF, SCSTPGRRNCNQDCYTHCDCVGGKEYNNGAGMVLCKTCTYPLGKKVGFCKF, SCTKPGKRACNAKCKYHCDCKDGPASKGPQGRYHCKTCEVALRDLQGYCVQ, respectively. Peptides were automatically assembled using Fmoc chemistry on 2-chlorotrityl resin at a 0.25 mmol scale using a CS-Bio peptide synthesiser. Couplings were performed for 30 min with four equivalents (eq) of amino acids and 0.5 M N,N,N′,N′-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate (4 eq) and *N,N*-diisopropyletylamine (4 eq) in *N,N*-dimethylformamide (DMF). Fmoc deprotection was carried out for 2 x 5 min using 20% piperidine in DMF. DMF was used for washing steps in between couplings and deprotection. After completion of the chain assembly and final deprotection of the terminal, Fmoc group peptides were cleaved off the resin using a mixture of 92.5% trifluoroacetic acid (TFA), 2.5% 2,2′-(ethylenedioxy) diethanethiol, 2.5% triisopropylsilane and 2.5% water (all v/v). TFA was removed under reduced

pressure and peptides were precipitated in ice-cold ether, resolved in a 50:50 mix of buffers A (100% water, 0.05% TFA) and B (90% acetonitrile, 10% water, 0.045% TFA) (all v/v), and lyophilized. Peptides were purified by Reversed- Phase HPLC on a Shimadzu Prominence HPLC unit using preparative and semi-preparative Vydac C18 columns with linear gradients of 1%/min of buffer B at flow rates of 8 and 3 mL/min, respectively. Oxidation was performed using either 0.1 M NH_4HCO_3 with 13 μM reduced glutathione (HTs 1, 2 and 4) or a mixture of 35% DMSO, 5% dodecyl-β-maltoside and 2/2 mM reduced/oxidized glutathione (HT 3) at pH 8.2 for 48 - 96 h. Purity (>95%) of the correct products was assessed by mass spectrometry and analytical HPLC. Purity (>95%) of the major product for each peptide was assessed by mass spectrometry and analytical HPLC before subsequent analysis by Nuclear Magnetic Resonance (NMR) spectroscopy. [^{1}H] NMR spectra for each peptide dissolved in 90% H_2O/10% D_2O (pH 3.5 – 4.0) were recorded at 298 K on a Bruker Avance 600 NMR spectrometer (Bruker Corporation, MA, USA). Spectra were analyzed using Topspin 1.3 (Bruker). All peptides gave well-dispersed peaks in the amide region of the ^{1}H NMR spectra (Supplementary Fig. S1) which is characteristic of a folded, defined three dimensional structure. In addition, Total Correlated Spectroscopy (TOCSY) and Nuclear Overhauser Effect Spectroscopy (NOESY) data were acquired for HT1 and were found to be consistent with the previously published NMR structural data (Vink et al., 2014).

2.10. Synaptosome purification

Six male Wistar rats at 12 weeks of age were euthanized using CO_2 and their brains dissected (Animal Ethics approval QAAFI/502/12, The University of Queensland, Australia). The brains were transferred into a 15 mL Corning tube containing homogenizing buffer (HB: 0.32 M sucrose, 1 mM EDTA, 0.25 mM DTT, 5 mM Tris, pH 7.4). The excess blood was removed by rinsing the brains in HB several times. The brains obtained were homogenised in 9 mL of ice cold HB per gram of brain tissue in a pre-chilled tissue glass homogenizer mounted to a 1.5 A overhead stirrer (Laboratory Supply Pty Ltd, Australia) at speed 6 for 10 up/down cycles. The homogenate was centrifuged at 1000 *g* for 10 min at 4°C and the supernatant collected (S1 fraction) in ~ 9 mL

and diluted by the addition of 10 mL of HB. The diluted S1 preparation was kept on ice until the protein concentration was determined by the Bradford method (Brandford, 1976). Synaptosomes were purified by discontinuous Percoll gradient (Dunkley et al., 2008). Briefly, 2 mL of the S1 homogenate were diluted to a protein concentration of ~4 mg/mL and carefully pipetted onto the 3% Percoll layer of the Percoll gradient (3%, 10%, 15% and 23% vol/vol). The gradient was centrifuged at 31,000 *g* for 5 min at 4 °C. The F3 and F4 fractions were collected from the gradient by inserting a wide bore pipette tip and placing these fractions into a pre-chilled tube on ice. Ice-cold SET buffer (SET: 0. 32 mM sucrose + 1 mM EDTA) was added to obtain a total volume of 80 mL that was allocated into two 50 mL polycarbonate tubes. These SET diluted fraction were centrifuged at 20,000 x *g* for 30 min at 4°C. The supernatant was discarded, and the pellet containing the synaptosome was kept on ice. Finally, the synaptosome preparation was frozen as described previously (Daniel et al., 2012).

2.11. Binding of labelled HT to synaptic membranes

Synthetic HTs were labelled using the Alexa 488 kit (ThermoFisher Scientific Inc., Australia). BSA and the Taipan venom (*Oxyuranus scuttelatus scuttelatus*) (B.0189.13, Venom Supplies Pty Ltd, Australia) were used as negative and positive controls, respectively. The standard binding procedure was conducted in triplicate by incubating 75 µg of synaptosomes with 5 µg of toxins in the dark at 37°C for 90 min in Eppendorf tubes. The binding reaction was centrifuged for 7 min at 3,000 *g* and the supernatant was discarded. The pellet was washed three times with 180 µL of Krebs buffer (143 mM NaCl, 0.1 mM $CaCl_2$, 1.2 mM $MgCl_2$, 24.9 mM $NaCO_3$) followed by gentle agitation and centrifuged at 3,000 *g* to remove the unbound toxins. The pellet was resuspended in 100 µL of Krebs buffer and transferred to a Nunc F96 Microwell Plate. The fluorescence was read in a SpectraMax 340PC384 microplate reader with absorbance of 495 nm and emission at 519 nm, and processed by SoftMAX Pro (Molecular Devices LLC, Australia).

The quantification of the labelled HT bound in the synaptosome-binding assay was expressed as uptake units of labelled HT. A unit was defined as the total amount of labelled toxin

bound minus the background. The amount of labelled toxin and BSA non-specifically bound to the empty Eppendorf tubes and synaptosomes were considered the background components of the assay.

2.12. Neonatal mouse assay

The experiment was conducted under approval of the Animal Ethics Committee, The University of Queensland, Australia (QAAFI/502/12). A total of 40 Quackenbush White mice at 5 – 6 days of age and approximately 4–5 g in weight were randomly allocated into test groups of six neonate mice per group. In the experiment each neonate mouse was injected s.c with 30 µg of synthetic HT1 (Group A), HT2 (Group B), HT3 (Group C), HT4 (Group D, this group had four mice); 30 µg of each HT1, HT2 and HT3 were mixed for Group E. Group F was the positive control and consisted of neonatal mice inoculated with 30 µg of salivary gland extract diluted in 100 µL of 1X PBS. Group G was injected with 1X PBS as the negative control of the experiment. Mice were observed every 30 min for 10 h and the degrees of paralysis at these time points were scored using the system previously described (Stone et al., 1982):

0 - no paralysis, 1 - dropping of hips, 2 - partial paralysis in one hind limb, 3 - partial paralysis in both hind limbs, 4 - complete paralysis in one hind limb, 5 - progressing paralysis in other hind limb, 6 - complete paralysis in both hind limbs, 7 - complete paralysis in both hind limbs and partial paralysis in forelimbs, 8 - complete paralysis in all limbs, 9 - respiratory distress, 10 - euthanasia. The mice were euthanized at score 9 or at the end of the experiment.

3. Results

3.1. Illumina sequencing and assembly of the I. holocyclus transcriptome

A total of 200,208 transcripts • 200 bp long were obtained from SG (*n*=134,039) and VISC samples (*n*=66,169). The average contig lengths were 613 bp and 848 bp in SG and VISC samples, respectively. High genomic coverage was obtained with a total of 138.1 Mb of sequences with N50 = 865 bp and N50 = 1,543 bp in both tissue samples (Table 1). Even though a large number of

genes (*n*=19,967) were identified, it is likely that a significant fraction of non/low-expressed transcripts are missing.

A large number of transcripts were obtained for the SG samples (*n*=134,276) compared with the VISC samples (*n*=66,282). The majority of the transcripts were assigned to the category of 'novel proteins' for the SG (*n*=96,278) and VISC (*n*=36,717) samples representing 71.19% and 55.39% of all transcripts, respectively, (see Figs. 1A, 2A). In contrast, the proportion of transcripts in the 'putative protein' category was highly represented in VISC (*n*=16,601) compared with SG (*n*=19,523), (Figs. 1A, 2A). The transcripts of most interest in this study are identified in the 'known secreted' and 'putative secreted' protein categories, contributing to approximately 2% and 3% of the total transcripts of both experimental samples. The analysis of the 'known secreted' proteins category of both samples showed that the highly represented protein families were 'enzyme', 'protease inhibitor' and 'Salp15' (Figs. 1B, 2B). The 'putative secreted' protein transcripts in Figs. 1C and 2C showed similar representations for both protein categories. The 'enzyme' (Enz) and 'novel protein' (NP) related transcripts were highly represented in SG compared with the VISC sample. In contrast, transcripts related to 'immunity' (Imm) and 'glycine rich' families (GR) were more highly represented in VISC than in SG. The proportion of transcripts related to 'histamine binding proteins' (HBP) was relatively unchanged in both Illumina sequenced samples. Other categories of interest as mentioned above remained relatively unchanged between these two experimental samples. Specific analysis of the enzyme-related transcripts within the known secreted proteins (Figs. 1B, 2B) revealed a higher proportion of 'metalloprotease' and 'endonuclease' transcripts in the SG sample. In VISC the 'serine proteases', 'carboxyl esterases', and 'carboxypeptidases' were the most significant transcripts. The VISC and SG share 24 groups of enzymes, however multiple 'inositol polyphosphate phosphatases', 'alkaline phosphatases', 'esterases' and 'phospholipases A2' were unique to SG within the known 'secreted protein' category. The 'metalloprotease' expressed sequence tags (ESTs) within the 'putative secreted' category (Figs. 1C, 2C) were higher in the SG, although 'endonucleases' were under-represented.

By combining the enzymes of both 'known secreted' and putative secreted categories, there were no enzymes unique to either SG or VISC samples.

3.2. Functional annotation of the I. holocyclus transcriptome

The longest assembled transcript for each Oases assigned locus was selected as the representative sequence. The possible function of assembled *I. holocyclus* coding sequence (CDS) was conducted by GO assignments (Harris et al., 2004) to classify the distinct sequences. The sequence homology of 13,872 and 6,095 transcripts unique to VISC and SG, respectively, were categorized into functional groups (Supplementary Fig. S2). GO analysis of the transcripts illustrated an overall higher statistically significant ($P<0.05$) proportion of VISC GO terms for all categories (Supplementary Fig. S2). The majority of genes were limited to the 'Cellular Component' Classification, with 'Biological Process' and 'Molecular Function' following, respectively. A higher proportion of 'Cell Killing', 'Reproduction', 'Reproductive Process', 'Viral Reproduction' and 'serine/threonine phosphatase complex' was observed in the VISC sample in comparison to the SG sample. 'Auxiliary Transport Proteins' and 'Nutrient Reservoir' were represented by ~ 0.01% of SG genes, compared with 0.001% in VISC.

The transcript clustering analysis showed small numbers of clusters, which were unique to each library with 433 and 447 clusters for SG and VISC (Supplementary Tables S1, S2), respectively. A total of 8,167 clusters were common to both tissue samples. The putative and housekeeping protein clusters were the majority of the common clusters with 4,269 and 1,480 transcripts, respectively. Additionally, a high proportion of known (n=18) and putative (n= 61) secreted proteins clusters were unique to SG (Supplementary Tables S1, S2). However, these were under-represented in the VISC (Fig. 3). The same trend was reflected in the number of 'hypothetical protein' clusters with more occurring in the SG (n= 61) than VISC (n=48), but a higher number of 'putative protein' clusters were identified in VISC (n=153) compared with SG (n=87). All other categories were similar between the two samples (Fig. 3).

3.3. Highly expressed transcripts in the transcriptome of I. holocyclus

Transcripts assembled for SG and VISC samples were clustered using CD-HIT (Fu et al., 2012) to merge redundant transcripts. This analysis yielded 68,065 *I. holocyclus* representative transcripts that were used to measure expression differences between the SG and VISC samples. A total of 11,735 transcripts were identified to have • 1 RPM mapped tags in all samples and these were selected for further statistical analysis. We identified 5,250 differentially expressed genes between SG and VISC samples (False Discovery Rate (FDR) <0.001) of which 2,766 and 2,484 were up regulated in SG and VISC samples, respectively (Supplementary Table S3). Table 2 shows the classification of the top 1,000 differentially expressed transcripts between the SG and VISC samples. VISC samples show a larger fraction of gene categories that are highly expressed compared with SG (Table 2, Supplementary Table S3) including cytochrome P450, sulfotransferases, peritrophic membrane chitin binding protein, and serine proteinase inhibitor serpin-3. In contrast, SG samples show an increase in metalloproteases that are differentially expressed compared with VISC samples (Table 2, Supplementary Table S3).

3.4. Identification and characterization of the holocyclotoxin family

A total of 18 different and non-redundant holocyclotoxin sequences were found in this *I. holocyclus* transcriptome analysis which were designated as HT2, HT3, HT4 up to HT19, (Fig. 4A) including the original HT1 accession. The protein sequence of each HT comprised 68-79 amino acids that include a signal peptide (18-26 amino acids). The mature toxins have an average size of 22 amino acids that represents the predicted molecular weight of ~ 5 kDa. The conserved cysteines that form the cysteine knot motif (Cys 2- Cys 6, Cys 3-Cys 7 and Cys 5-Cys 8 disulfide bonds) were found in all identified HTs as previously reported (Fig. 4A). Additionally, these toxin sequences are characterized by high content of G, K, and L-amino acids. A predicted signal secretion peptide was found in all the toxins, and only HT16 has potential N-glycosylation sites at positions 49 and 54 (Fig. 4A).

The DNA sequences (n=21) of the holocyclotoxins listed in GenBank from Patent WO

2014018724 A1 were translated into predicted HT proteins, and 14 of these 21 HT sequences had full ORFs. These ORFs showed similar sequence characteristics to the HTs described above. Of the 14 full ORFs from this patent, five had 100% identity to five of the 19 HTs outlined in this transcriptome including the previously published HT1 sequence:

Sequence 1 (GenBank: **JC962178**) = HT1 (**AAV34602**)

Sequence 4 (GenBank: **JC962181**) = HT8 (**KP096307**)

Sequence 11 (GenBank: **JC962188**) = HT13 (**KP963968**)

Sequence 12 (GenBank: **JC962189**) = HT14 (**KP963969**)

Sequence 13 (GenBank: **JC962190**) = HT12 (**KP963967**)

Fig. 4B shows a Neighbor-Joining tree built from the mature protein alignments of all HTs including *Atrax versutus* spider venom 'Versutoxin' (GenBank Accession no. **P13494.1**) as 'outgroup' for the tree and excluding the five HT duplicates described above and the seven incomplete HT ORFs reported in the Patent WO 2014018724 (Sequence 15 (**JC962192**); Sequence 16 (**JC962193**); Sequence 17 (**JC962194**); Sequence 18 (**JC962195**); Sequence 19 (**JC962196**); Sequence 20 (**JC962197**); and Sequence 21 (**JC962198**)). It is of note that Sequence 20 has an ORF with ~5 N-terminal amino acids missing and otherwise appears to be identical to HT11 (**KP096310**). Fig. 4B shows branches with over 70% bootstrap support and although there are small clusters of HTs with high similarity, there were many branches not supported in the tree construction. HT16 is the most distant sequence (25-34% identity to all other HTs) followed by HT17 (33-43 % identity to all other HTs) (Supplementary Table S4). Also, HT4 and Sequence 14 (**JC962191**) are 96% similar to each other but collectively only 30-48% similar to the other HTs.

A subset of these holocyclotoxins - HT1, HT2, HT3 and HT4 - were synthesized. HT1, HT2 and HT4 were labelled with Alexa 488 to determine their capability to bind rat synaptosome preparations in vitro. Fig. 5 shows specific binding and high relative fluorescent units (RFU) for HT1, HT2 and HT3 compared with the negative control (BSA). The level of binding of HT3 was similar to the positive sample of the assay Taipan venom (*Oxyuranus scuttelatus scuttelatus*) (Fig.

5).

3.5. Paralysis symptoms in the neonate mouse assay

This sensitive biological test of paralysis tick toxins in neonatal mice was used to measure the quantitative paralysis index (PI) for each of the synthetic HTs used in the experiment. HT1, HT2 and HT3 scored PIs of 3-4 in all animals under experimentation except one animal in the HT1 group with a score of PI = 9-10 after 8 h (Fig. 6). HT4 was the only holocyclotoxin capable of inducing severe paralysis and respiratory distress in all neonate mice after 2 h of administration (Fig. 6D). Mice treated with an HT1, HT2 and HT3 mixture reached PI values equal to 7-8 after 3 h p.i.. In this experimental group, the paralysis symptoms appeared rapidly compared with the positive control inoculated with *I. holocyclus* SG extract. However, the SG extract induced strong paralysis in the range of PI = 9 after 5-8 h p.i. in the neonatal mice. The negative control did not show any symptoms of paralysis during the experiment (Fig. 6).

4. Discussion

The ecological interaction of the mammalian host with their indigenous ectoparasites is characterized by the development of different host defense mechanisms for rejecting parasite infestations. For example, grooming is an important behavior in monkeys, cats and cattle to reduce the number of external parasites (Nekaris et al., 2013). The primary line of host defense also relies on complex hemostatic and immunological responses of the host against ectoparasites. For example, ticks counteract the host response and successfully survive attached to the skin of their host via the secretion of a complex mixture of molecules present in tick saliva (Rodriguez-Valle et al., 2010; Chmelar, 2016; Xu et al., 2016). Host - parasite interactions have been studied using proteomic and transcriptomic analyses of economically important ticks such as *A. americanus* (Aljamali et al., 2009), *D. andersoni* (Alarcon-Chaidez, 2007), *Ixodes pacificus* (Francischetti et al., 2005), *Rhipicephalus microplus* (Rodriguez-Valle et al., 2010), and others. These reports concluded

that tick proteins secreted into the tick's attachment site inhibit host hemostatic and immunological responses by diverse mechanisms (Francischetti et al., 2009). Blood coagulation, platelet aggregation and vasoconstriction are the hemostatic reactions fundamentally affected by proteins secreted in the tick saliva. A detailed description of these anti-hemostatic molecules, such as IRS-2, apyrase, TSGP2 and 3, savignygrin, ixoderin, rhipilin, Ir-CPI, serine protease inhibitors, among others, involved in the inhibition of different enzymes of the coagulation pathway, has been reported (Chmelar, 2016). The data obtained in the present transcriptomic study showed that proteins secreted by *I. holocyclus* SG and VISC tissues did not differ from the protein expression patterns observed in other tick transcriptomic studies except for the holocyclotoxins (Alarcon-Chaidez, 2007; Aljamali et al., 2009; Rodriguez-Valle et al., 2010; Ong, 2016).

The common proteins identified in the transcriptomes include protease inhibitors, metalloproteases, glycine rich proteins, lipocalins, and *Ixodes*-specific SG transcripts (SALP15) associated with pathogen transmission. Protease inhibitors (Alarcon-Chaidez, 2007; Aljamali et al., 2009; Rodriguez-Valle et al., 2010) are characterised by a high variability in their amino acid sequences which leads to important structural modifications that affect their specific functions (Schwarz et al., 2014; Chmelar, 2016). Also, peptide members of the Kunitz protein family have diverse activities in venomous scorpions and spiders eg LmKKT-1a, huwentoxin-1, huwentoxin-2, and magi-1 families that inhibit proteases and ion channels (Chen et al., 2012). RasKLP from *Rhipicephalus appendiculatu,* is a member of Kunitz family with a highly modified structure that activates maxi K channels but loses its protease inhibition activity, suggesting its potential role in the regulation of host blood influx during tick feeding (Paesen et al., 2009; Schwarz et al., 2014). In the *I. holocyclus* SG, the serine protease inhibitors (serpins) would also be important factors for overcoming the hemostatic response of the host by affecting the coagulation, platelet aggregation and complement activation (Rodriguez Valle et al., 2015; Chmelar, 2016; Xu et al., 2016). In this study, serpins were also identified in the viscera transcriptome which would also be associated with the digestion of host blood proteins. Serpins have been reported in *I. ricinus* (Iris) (Prevot et al.,

2009), *A. americanum* (Mulenga et al., 2013), *Haemaphysalis longicornis* (Imamura et al., 2005), *Rhipicephalus haemaphysaloides* (Yu et al., 2013) and *R. microplus* (Rodriguez Valle et al., 2015). Metalloproteases (found in the *I. holocyclus* VISC and SGs) are also commonly found in other Ixodid ticks (including *I. scapularis*) and are known to affect platelet disaggregation and blood coagulation, thus facilitating blood feeding (Francischetti et al., 2003; Chmelar et al., 2012). These enzymes are regulated by the uptake of the blood meal, pathogen acquisition and transmission (Francischetti et al., 2003).

GR proteins (Anatriello et al., 2010) and lipocalins (as HBPs) are usually reported in tick sialome studies. Cutaneous inflammation is a usual skin reaction that affects tick attachment by reducing feeding and reproductive fecundity (Wikel, 1982). Histamine is released from mast cells and induces vasodilation, leukocyte recruitment and edema formation. Ticks secrete HBPs at the tick attachment site to inhibit physiological reactions of the host against tick infestation, mediated by histamine. SG Salp15 (or *Iho*-Salp15) transcripts are associated with tick pathogen transmission and have been previously identified in *I. scapularis* (Das et al., 2001), *I. ricinus* (Hovius et al., 2007), *Ixodes sinensis* (GenBank Accession Nos. **ACV32166**, **AFP59046**, **AFP59045**) and *Ixodes persulcatus* (Hojgaard et al., 2009). Salp15 adheres to the outer surface protein C (OspC) of *Borrelia burgdorferi* to form a complex (Salp15 – OspC) that protects the spirochete from antibody-mediated cytotoxicity, enabling host infection. Additionally, it has been reported that *I. scapularis* Salp15 inhibits CD4+ cell proliferation by the suppression of the calcium flux stimulated by T Cell receptors (TCR), resulting in reduction of the IL-2 level (Anguita et al., 2002).

The presence of HTs in ticks has been explained as derivative of a symbiotic organism in a tick (Mans et al., 2004) or a vestigial function preserved in ticks throughout their evolution toward a parasitic lifestyle (Mans and Neitz, 2004). This latter statement is supported by phylogenetic evidence that shows the parasitiformes (*Acari*) and pseudoscorpions share a common predecessor within the phylum Arthropoda (Cabezas-Cruz, 2014). Additionally, there is extensive evolution of tick proteins as reported for the Kunitz/bovine pancreatic trypsin inhibitor (BPTI) domains in

R. appendiculatus. The Ra-KLP secreted in the saliva of *R. appendiculatus* has an extensive modification of the basic Kunitz fold resulting in a loss of its protease inhibitory capability against a broad range of proteases. However, this protein has a stimulatory effect on Ca^{2+}-activated K^{+} channels (Paesen et al., 2009). There are diverse tick species directly related with paralysis of the host as *D. andersoni*, *D. variabilis* in North America, *I. rubicundus* and others, but the protein sequences for these tick toxins are yet to be determined (Mans et al., 2004). Tick toxins have also been identified in tick egg extracts from other hard tick species such as *Amblyomma hebraeum, R. e. evertsi, R. microplus, Rhipicephalus decoloratus* and *Hyalomma truncatum*. The presence of these toxins in eggs was related to egg protection against predation in natural environments (Mans et al., 2004). The larvae of *Argas (Persicargas) walkerae* contain other tick toxins associated with neuropathogenic properties at a molecular weight of 11 kDa (Maritz et al., 2000).

In this study, a specific family of proteins was identified containing the ICK motif, eight conserved cysteines and homology to holocyclotoxin HT1 (**AAV34602**) at sequence identities ranging from 30-90% (Vink et al., 2014). Peptides with the ICK motif usually contain three disulfide bonds (Smith et al., 2011) and are highly represented in the venoms of spiders and other unrelated organisms. HTs have an average molecular weight of approximately 5.9 kDa, a predicted basic pi, and a total of eight cysteines in highly conserved positions. Consequently, HTs have an additional disulfide bond similar to the spider *Hadronyche versuta* versutoxin, and the ergtoxin from the scorpion, *Centruroïdes noxius*, but they have different three-dimensional structures as was reported for HT1 (Vink et al., 2014). A proteomics study confirmed that these toxins are secreted in the saliva of *I. holocyclus* (Busch, G., 2016. Identification and characterisation of *Ixodes holocyclus* toxins to develop novel treatment methods, Queensland Alliance for Agriculture and Food Innovation. Masters Thesis, The University of Queensland, Australia http://espace.library.uq.edu.au/uqgbusch). Electrophysiological studies recently conducted using extensor digitorum longus (EDL) muscles showed that these toxins are functionally similar. The study also showed that HT1, HT3 and HT12 induced muscle paralysis by inhibiting

neurotransmitter release through a calcium-dependent mechanism; however, the molecular target of these toxins is currently unknown (Chand et al., 2016).

Some holocyclotoxins such as HT1, HT2 and HT3 did not induce severe paralysis in the neonatal mouse model; however, when a mixture of these toxins was used, high paralysis scores were achieved. This could be the result of different forms of toxins interacting with each other in saliva to form stable molecular aggregates for a strong toxic effect or by individual toxins acting on different molecular targets. This statement is in agreement with results obtained with toxins from *A. walkerae* and *I. holocyclus* (Thurn, 1992; Maritz et al., 2000). Differences in the paralytic action of the holocyclotoxins in the neonatal mouse model may not reflect their activity in other host animals severely affected by the tick paralyses such as dogs and cats. For instance HT4 as a single treatment induced quick and strong paralysis of the host hind and fore limbs and was characterized by significant respiratory distress in the neonate mouse model. However, this does not exclude the possible molecular aggregation of HT4 molecules under normal physiological conditions during host-tick interactions. Finally, this research significantly expands the genomic information for *Ixodes* spp. and, in particular, for *I. holocyclus,* and contributes to the characterization of the Australian paralysis tick toxin family. We believe this is the first comprehensive analysis and biological study of this family of toxins derived from a tick species.

Acknowledgments

The Australian Research Council Linkage Grant LP120200836 and Elanco Animal Health, Australia, supported this research. The rat brain synaptosome extractions and neonatal mouse assays were conducted in accordance with requirements of the Animal Ethics Committee, The University of Queensland, Australia (QAAFI/502/12). The authors wish to thank Bronwyn Venus for project resource management and technical support. Drs Paula Moolhuijzen, Kim Agnew and Aleta Knowles, formerly with Murdoch University (Australia) and Elanco are currently affiliated with Curtin University (Australia), Boehringer Ingleheim Animal Health (Australia) and Virbac Australia Pty. Ltd., respectively. We thank Barbara Arnts, Kevin Wathen-Dunn and the staff at the

Australian Institute for Bioengineering and Nanotechnology Animal Facility, The University of Queensland Biological Resources, for their expert assistance and training. We thank Professor Kelly Brayton of Washington State University, USA for critical review of the manuscript. The authors declare that they have no competing interests.

References

Alarcon-Chaidez, F.J., Sun, J. X., Wikel, S. K., 2007. Transcriptome analysis of the salivary glands of *Dermacentor andersoni* Stiles (Acari : Ixodidae). Insect Biochem. Mol. Biol. 37, 48-71.

Aljamali, M.N., Hern, L., Kupfer, D., Downard, S., So, S., Roe, B.A., Sauer, J.R., Essenberg, R.C., 2009. Transcriptome analysis of the salivary glands of the female tick *Amblyomma americanum* (Acari: Ixodidae). Insect Mol. Biol. 18, 129-154.

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.

Anatriello, E., Ribeiro, J.M.C., de Miranda-Santos, I.K.F., Brandão, L.G., Anderson, J.M., Valenzuela, J.G., Maruyama, S.R., Silva, J.S., and Ferreira, B.R., 2010. An insight into the sialotranscriptome of the brown dog tick, *Rhipicephalus sanguineus*. BMC Genomics 11, 450.

Anguita, J., Ramamoorthi, N., Hovius, J.W., Das, S., Thomas, V., Persinski, R., Conze, D., Askenase, P.W., Rincon, M., Kantor, F.S., Fikrig, E., 2002. Salp15, an *Ixodes scapularis* salivary protein, inhibits CD4(+) T cell activation. Immunity 16, 849-859.

Atwell, R.B., Campbell, F.E., Evans, E.A., 2001. Prospective survey of tick paralysis in dogs. Aust. Vet. J. 79, 412-418.

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., Bairoch, A., 2007. UniProtKB/Swiss-Prot. Methods Mol. Biol. 406, 89-112.

Brandford M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254.

Brown, M.R., Sheumack, D.D., Tyler, M.I., Howden, M.E., 1988. Amino acid sequence of versutoxin, a lethal neurotoxin from the venom of the funnel-web spider *Atrax versutus*. Biochem. J. 250, 401-405.

Cabezas-Cruz, A., Valdés, J.J., 2014. Are ticks venomous animals? Front. Zool. 11, 47.

Chand, K.K., Lee, K.M., Lavidis, N.A., Rodriguez-Valle, M., Ijaz, H., Koehbach, J., Clark, R.J., Lew-Tabor, A., Noakes, P.G., 2016. Tick holocyclotoxins trigger host paralysis by presynaptic inhibition. Sci. Rep. 6, 29446.

Chen, Z.Y., Hu, Y.T., Yang, W.S., He, Y.W., Feng, J., Wang, B., Zhao, R.M., Ding, J.P., Cao, Z.J., Li, W.X., Wu, Y.L., 2012. Hg1, novel peptide inhibitor specific for Kv1.3 channels from first scorpion Kunitz-type potassium channel toxin family. J. Biol. Chem. 287, 13813-13821.

Chmelar, J., Calvo, E., Pedra, J.H., Francischetti, I.M., Kotsyfakis, M., 2012. Tick salivary secretion as a source of antihemostatics. J. Proteomics 75, 3842-3854.

Chmelar, J., Kotál, J., Kopecký, J., Pedra, J.H., Kotsyfakis, M., 2016. All for one and one for all on the tick–host battlefield. Trends Parasitol. 32, 368-377.

Daniel, J.A., Malladi, C.S., Kettle, E., McCluskey, A., Robinson, P.J., 2012. Analysis of synaptic vesicle endocytosis in synaptosomes by high-content screening. Nat. Protoc. 7, 1439-1455.

Das, S., Banerjee, G., DePonte, K., Marcantonio, N., Kantor, F.S., Fikrig, E., 2001. Salp25D, an *Ixodes scapularis* antioxidant, is 1 of 14 immunodominant antigens in engorged tick salivary glands. J. Infect. Dis. 184, 1056-1064.

Dunkley, P.R., Jarvie, P.E., Robinson, P.J., 2008. A rapid Percoll gradient procedure for preparation of synaptosomes. Nat. Protocol. 3, 1718-1728.

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A., 2010. The Pfam protein families database. Nucleic Acids Res. 38, D211-D222.

Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., Shanmugam, D., Roos, D.S., Stoeckert Jr., C.J., 2011. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. In: Bateman, A., Draghici, S., Khurana, E., Orchard, S., Pearson, W.R. (Eds.). Current Protocols in Bioinformatics, John Wiley & Sons Inc., Supplement 35, Unit 6.12.11-19.

Francischetti, I.M., Sa-Nunes, A., Mans, B.J., Santos, I.M., Ribeiro, J.M., 2009. The role of saliva in tick feeding. Front. Biosci. 14, 2051-2088.

Francischetti, I.M.B., Pham, V.M., Mans, B.J., Andersen, J.F., Mather, T.N., Lane, R.S., Ribeiro, J.M.C., 2005. The transcriptome of the salivary glands of the female western blacklegged tick *Ixodes pacificus* (Acari : Ixodidae). Insect Biochem. Mol. Biol. 35, 1142-1161.

Francischetti, M.I., Mather, N.T., Ribeiro, M.J., 2003. Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrinogenlytic activities in the saliva of the Lyme disease tick vector *Ixodes scapularis*. Biochem. Biophys. Res. Commun. 305, 869-875.

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150-3152.

Gauci, M., Stone, B.F., Thong, Y.H., 1990. An immunoblot technique for identification of allergens of the Australian paralysis tick *Ixodes holocyclus*. J. Immunol. Methods 126, 51-55.

Gulia-Nuss, M., Nuss, A.B., Meyer, J.M., Sonenshine, D.E., Roe, R.M., Waterhouse, R.M., Sattelle, D.B., de la Fuente, J., Ribeiro, J.M., Megy, K., Thimmapuram, J., Miller, J.R., Walenz B.P., Koren, S., Hostetler, J.B., Thiagarajan, M., Joardar, V.S., Hannick, L.I., Bidwell, S., Hammond, M.P., Young, S., Zeng, Q., Abrudan, J.L., Almeida, F.C., Allón, N., Bhide, K., Bissinger, B.W., Bonzon-Kulichenko, E., Buckingham, S.D., Caffrey, D.R., Caimano, M.J., Croset, V., Driscoll, T., Gilbert, D., Gillespie, J.J., Giraldo-Calderón, G.I., Grabowski, J.M., Jiang, D., Khalil, S.M.S., Kim, D., Kocan, K.M., Koči, J., Kuhn, R.J., Kurtti, T.J., Lees, K., Lang, E.G., Kennedy, R.C., Kwon, H., Perera, R., Qi, Y., Radolf, J.D., Sakamoto, J.M., Sánchez-Gracia, A., Severo, M.S., Silverman, N., Šimo, L., Tojo, M., Tornador, C., Van Zee, J.P., Vázquez, J., Vieira, F.G., Villar, M., Wespiser, A.R., Yang, Y., Zhu, J., Arensburger, P., Pietrantonio, P.V., Barker, S.C., Shao, R., Zdobnov, E.M., Hauser, F., Grimmelikhuijzen, C.J.P., Park, Y., Rozas, J., Benton, R., Pedra, J.H.F., Nelson, D.R., Unger, M.F., Tubio, J.M.C., Tu, Z., Robertson, H.M., Shumway, M., Sutton, G., Wortman, J.R., Lawson, D., Wikel, S.K., Nene, V.M., Fraser, C.M., Collins, F.H., Birren, B., Nelson, K.E., Caler, E., Hill, C.A., 2016. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. Nat. Commun. 7, 10507.

Hall-Mendelin, S., Craig, S.B., Hall, R.A., O'Donoghue, P., Atwell, R.B., Tulsiani, S.M., Graham, G.C., 2011. Tick paralysis in Australia caused by *Ixodes holocyclus* Neumann. Ann. Trop. Med. Parasitol. 105, 95-106.

Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., Richter, J., Rubin, G.M., Blake, J.A., Bult, C., Dolan, M., Drabkin, H., Eppig, J.T., Hill, D.P., Ni, L., Ringwald, M., Balakrishnan, R., Cherry, J.M., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S., Fisk, D.G., Hirschman, J.E., Hong, E.L., Nash, R.S., Sethuraman, A., Theesfeld, C.L., Botstein, D., Dolinski, K., Feierbach, B., Berardini, T., Mundodi, S., Rhee, S.Y., Apweiler, R., Barrell, D., Camon, E., Dimmer, E., Lee, V., Chisholm, R., Gaudet, P., Kibbe, W., Kishore, R., Schwarz, E.M., Sternberg, P., Gwinn, M., Hannick, L., Wortman, J., Berriman, M., Wood, V., de la Cruz, N., Tonellato, P., Jaiswal, P., Seigfried, T., White, R., Gene Ontology, C., 2004. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. 32, D258-D261.

Hojgaard, A., Biketov, S.F., Shtannikov, A.V., Zeidner, N.S., Piesman, J., 2009. Molecular identification of Salp15, a key salivary gland protein in the transmission of lyme disease spirochetes, from *Ixodes persulcatus* and *Ixodes pacificus* (Acari: *Ixodidae*). J. Med. Entomol. 46, 1458-1463.

Hovius, J.W., Ramamoorthi, N., Van't Veer, C., de Groot, K.A., Nijhof, A.M., Jongejan, F., van Dam, A.P., Fikrig, E., 2007. Identification of Salp15 homologues in *Ixodes ricinus* ticks. Vector Borne Zoonotic Dis. 7, 296-303.

Imamura, S., da Silva Vaz Junior, I., Sugino, M., Ohashi, K., Onuma, M., 2005. A serine protease inhibitor (serpin) from *Haemaphysalis longicornis* as an anti-tick vaccine. Vaccine 23, 1301-1311.

Jongejan, F., Uilenberg, G., 2004. The global importance of ticks. Parasitology 129, S3-S14.

Kaire, G.H., 1966. Isolation of tick paralysis toxin from *Ixodes holocyclus*. Toxicon 4, 91–97.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 40, D109-114.

Koski, L.B., Gray, M.W., Lang, B.F., Burger, G., 2005. AutoFACT: an automatic functional annotation and classification tool. BMC Bioinformatics 6, 151.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357-359.

Letunic, I., Doerks, T., Bork, P., 2009. SMART 6: recent updates and new developments. Nucleic Acids Res. 37, D229-D232.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Mans, B.J., Gothe, R., Neitz, A.W., 2004. Biochemical perspectives on paralysis and other forms of toxicoses caused by ticks. Parasitology 129 Suppl, S95-111.

Mans, B.J., Neitz, A.W.H., 2004. The sand tampan, *Ornithodoros savignyi*, as a model for tick-host interactions. S. Afr. J. Sci. 100, 283-288.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Bryant, S.H.,

2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37, D205-210.

Maritz, C., Louw, A.I., Gothe, R., Neitz, A.W., 2000. Detection and micro-scale isolation of a low molecular mass paralysis toxin from the tick, *Argas (Persicargas) walkerae*. Exp. Appl. Acarol. 24, 615-630.

Masina, S., Broady, K.W., 1999. Tick paralysis: development of a vaccine. Int. J. Parasitol. 29, 535-541.

Mulenga, A., Kim, T., Ibelli, A.M., 2013. *Amblyomma americanum* tick saliva serine protease inhibitor 6 is a cross-class inhibitor of serine proteases and papain-like cysteine proteases that delays plasma clotting and inhibits platelet aggregation. Insect Mol. Biol. 22, 306-319.

Nekaris, K.A., Moore, R.S., Rode, E.J., Fry, B.G., 2013. Mad, bad and dangerous to know: the biochemistry, ecology and evolution of slow loris venom. J. Venom. Anim. Toxins Incl. Trop. Dis. 19, 21.

Ong, C.T., Rodriguez-Valle M., Moolhuijzen, P.M., Barrero, R.A., Hunter, A., Szabo, T., Bellgard, M.I., Lew-Tabor, A.E., 2016. Exploring the transcriptomic data of the Australian paralysis tick, *Ixodes holocyclus*. GSTF J. Vet. Sci. 3, 10.5176/2345-7880_5173.5171.5114.

Paesen, G.C., Siebold, C., Dallas, M.L., Peers, C., Harlos, K., Nuttall, P.A., Nunn, M.A., Stuart, D.I., Esnouf, R.M., 2009. An ion-channel modulator from the saliva of the brown ear tick has a highly modified Kunitz/BPTI structure. J. Mol. Biol. 389, 734-747.

Prevot, P.P., Beschin, A., Lins, L., Beaufays, J., Grosjean, A., Bruys, L., Adam, B., Brossard, M., Brasseur, R., Zouaoui Boudjeltia, K., Vanhamme, L., Godfroid, E., 2009. Exosites mediate the anti-inflammatory effects of a multifunctional serpin from the saliva of the tick *Ixodes ricinus*. FEBS 276, 3235-3246.

Ribeiro, J.M., Anderson, J.M., Manoukis, N.C., Meng, Z., Francischetti, I.M., 2011. A further insight into the sialome of the tropical bont tick, *Amblyomma variegatum*. BMC Genomics 12, 136.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genetics 16, 267-277.

Roberts, A., Pachter, L., 2013. Streaming fragment assignment for real-time analysis of sequencing experiments. Nat. Methods 10, 71-73.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Rodriguez Valle, M., Xu, T., Kurscheid, S., Lew-Tabor, A.E., 2015. *Rhipicephalus microplus* serine protease inhibitor family: annotation, expression and functional characterisation assessment. Parasit. Vectors 8, 7.

Rodriguez-Valle, M., Lew-Tabor, A.E., Gondro, C., Moolhuijzen, P., Vance, M., Guerrero, F.D., Bellgard, M., Jorgensen, W., 2010. Comparative microarray analysis of *Rhipicephalus (Boophilus) microplus* expression profiles of larvae pre-attachment and feeding adult female stages on *Bos indicus* and *Bos taurus* cattle. BMC Genomics 11, 47.

Rodriguez-Valle, M., Vance, M., Moolhuijzen, P.M., Tao, X., Lew-Tabor, A.E., 2012. Differential recognition by tick-resistant cattle of the recombinantly expressed *Rhipicephalus microplus* serine protease inhibitor-3 (RMS-3). Ticks Tick Borne Dis. 3, 159-169.

Schulz, M.H., Zerbino, D.R., Vingron, M., Birney, E., 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28, 1086-1092.

Schwarz, A., Cabezas-Cruz, A., Kopecky, J., Valdes, J.J., 2014. Understanding the evolutionary structural variability and target specificity of tick salivary Kunitz peptides using next generation transcriptome data. BMC Evol. Biol. 14, 4.

Smeds, L., Kunstner, A., 2011. ConDeTri--a content dependent read trimmer for Illumina data. PLoS ONE 6, e26314.

Smith, J.J., Hill, J.M., Little, M.J., Nicholson, G.M., King, G.F., Alewood, P.F., 2011. Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif. Proc. Natl. Acad. Sci. U.S.A. 108, 10478-10483.

Stone, B.F., Commins, M.A., Kemp, D.H., 1983. Artificial feeding of the Australian paralysis tick, *Ixodes holocyclus* and collection of paralysing toxin. Int. J. Parasitol. 13, 447-454.

Stone, B.F., Cowie, M.R., Kerr, J.D., Binnington, K.C., 1982. Improved toxin/antitoxin assays for studies on the Australian paralysis tick *Ixodes holocyclus*. Aust. J. Exp. Biol. Med. Sci. 60, 309-318.

Stone, B.F., Binnington, K.C., Gauci, M., Aylward, A.H., 1989. Tick/host interactions for *Ixodes holocyclus*: Role, effects, biosynthesis and nature of its toxic and allergenic oral secretions. Exp. Appl. Acarol. 7, 59-69.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4, 41.

Thurn, M.J., Gooley, A., Broady, K.W., 1992. Identification of the neurotoxin from the Australian paralysis tick, *Ixodes holocyclus*, In: Gopalakrishnakone, P., Tan, C.K.Ž (Ed.), Recent Advances in Toxinology Research. Venom and Toxin Research Group, National University of Singapore, Singapore, pp. 243 - 256.

Vink, S., Daly, N.L., Steen, N., Craik, D.J., Alewood, P.F., 2014. Holocyclotoxin-1, a cystine knot toxin from *Ixodes holocyclus*. Toxicon 90, 308-317.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., Dicuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L.Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D.J., Madden, T.L., Maglott, D.R., Miller, V., Ostell, J., Pruitt, K.D., Schuler, G.D., Shumway, M., Sequeira, E., Sherry, S.T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R.L., Tatusova, T.A., Wagner, L., Yaschenko, E., 2008. Database resources of the National Center for Biotechnology Information, USA. Nucleic Acids Res. 36, D13-D21.

Wikel, S.K., 1982. Histamine content of tick attachment sites and the effects of H1 and H2 histamine antagonists on the expression of resistance. Ann. Trop. Med. Parasitol. 76, 179-185.

Xu, T., Lew-Tabor, A., Rodriguez-Valle, M., 2016. Effective inhibition of thrombin by *Rhipicephalus microplus* serpin-15 (RmS-15) obtained in the yeast *Pichia pastoris*. Ticks Tick Borne Dis. 7, 180-187.

Yu, Y., Cao, J., Zhou, Y., Zhang, H., Zhou, J., 2013. Isolation and characterization of two novel serpins from the tick *Rhipicephalus haemaphysaloides*. Ticks Tick Borne Dis. 4, 297-303.

Figure legends

Fig. 1. Summary of *Ixodes holocyclus* salivary gland transcriptome collected from fully engorged adult female ticks. (A) Salivary Gland (SG) distribution of transcript counts described by the functionally annotated categories. (B) 'Known Secreted Protein', and (C) 'Putative Secreted Protein' categories were expanded. All segments show the category name and percentage of total transcript counts for each section derived from the SG library of fully engorged *I. holocyclus* adult female ticks.

Fig. 2. Summary of *Ixodes holocyclus* viscera transcriptome collected from fully engorged adult female ticks. (A) Viscera (VISC) distribution of transcript counts as described by the functionally annotated categories. (B) 'Known Secreted Protein', and (C) 'Putative Secreted Protein' categories were expanded. All segments show the category name and percentage of total transcript counts for each section derived from the VISC library of fully engorged *I. holocyclus* female adult ticks.

Fig. 3. Quantitative distribution of the sequence clusters unique to (A) salivary gland (SG) and (B) viscera (VISC) libraries.

Fig. 4. Sequence alignments and phylogeny of holocyclotoxin sequences. (A) Alignment of the non-redundant holocyclotoxin sequences (HT2-HT19) with the original HT1 (GenBank Accession # **AAV34602**). The highly conserved cysteine residues 'C' are shaded in grey. N-terminal residues are highlighted in italics. The hypothetical signal peptide region (SP) is indicated. Characteristic spatial distribution of the eight cysteine residues present in the mature 19 holocyclotoxins is represented as: C-X3G-X3-CN-X2-C-X2-HCDC-X2-G-X7-G-X3-C-X2-C-X10-C with numbers indicating the typical inter cysteine spacing between these residues. N-glycosylation sites are highlighted in HT16 with Asn-Xaa-Ser/Thr in the sequence highlighted in blue font and the Asparagines predicted to be N-glycosylated are highlighted in red font. (B) Neighbour-Joining tree

prepared from a ClustalW alignment of the 19 HT sequences (without the SP regions) described in A aligned with nine of the WO2014018724 patent sequences using the funnel web spider's Versutoxin (GenBank Accession No. **P13494.1**) as the outgroup using the Jukes-Cantor distance model (1000 bootstrap replicates) with 70% bootstrap branching support cut-off for tree branches. The percentage number indicates the branch support; transcriptome sequences identified in this study are highlighted with an asterisk (*).

Fig. 5. Synaptosome-binding assay. Venom from Taipan (*Oxyuranus scuttelatus scuttelatus*) (OSS) (Venom Supplies Pty Ltd, Australia) was the positive control for synaptosome binding and BSA was used as the negative control. Synthetic *Ixodes holocyclus* holocyclotoxins HT1, HT2 and HT3 were tested in this assay. The relative unit of fluorescence (RFU) was obtained by SpectraMax 340PC384 microplate reader with absorbance of 495 nm and emission at 519 nm. Data are represented as media ± S.D.

Fig. 6. Quantitative index of paralysis (PI) induced in neonatal mice after inoculation with *Ixodes holocyclus* holocyclotoxins (HTs). (A - D) Neonatal mice were inoculated with 30 µg of HT1, HT2, HT3 and HT4, respectively. (E) Neonatal mice were inoculated with a mixture containing 30 µg of each synthetic holocyclotoxin (HT1, HT2 and HT3). (F) Salivary gland extract (SGE) from *Ixodes holocyclus* (positive control) and PBS (negative control). Data shows the PI scored per neonatal mouse at different time points.

Table 1. Summary of Illumina HiSeq reads obtained from two biological replicates of pooled salivary gland (SG) and viscera (VISC) samples from fully engorged adult female *Ixodes holocyclus* ticks collected from paralysed companion animals.

***I. holocyclus* origin**	**Tissue sample**	**Number of transcripts ≥ 200 bp**	**L50**[a]	**Mean (b)**	**N50**	**Longest transcript**	**Total number of bases (Mb)**
Dogs & Cats	SG	41,669	7101	616	892	13,897	25.7
Dogs & Cats	SG	92,370	13,581	610	838	19,364	56.4
SG Total		134,039	20,682				82.1
Dogs & Cats	VISC	38,700	5565	830	1577	18,292	32.2
Dogs & Cats	VISC	27,469	4457	866	1509	18,582	23.8
VISC Total		66,169	10,022				56.0
Total		200,208	30,704				138.1

[a]L50, number of contigs at N50

Table 2. A summary of highly expressed transcript categories identified in the salivary gland (SG) and viscera (VISC) transcriptomes of *Ixodes holocyclus* adult female fully engorged ticks. Supplementary Tables S1 and S2 provide detailed lists of transcripts.

Annotation category	VISC	SG
Conserved Hypothetical Proteins	158	231
Hypothetical proteins	50	25
Secreted protein	31	8
Sulfotransferase	30	2
Secreted salivary gland peptide	24	2
Cytochrome P450	15	1
Metalloproteases	1	11
Peritrophic membrane chitin binding protein	8	0
Serine proteinase inhibitor serpin-3	7	1
Glutathione S-transferase	7	0
Thrombin inhibitor	5	0
Serine carboxypeptidase	5	1
Cystatin-2 precursor	5	1
Chymotrypsin-C precursor	5	1
Others	291	74
Total	**642**	**358**

Highlights

- Translated *Ixodes holocyclus* transcriptome sequences are predicted to counteract host immune responses
- Transcriptomic analysis identified 18 new members of the holocyclotoxin (HT) family
- HTs are members of a multi-gene family with conserved Inhibitory Cysteine Knot motifs
- Synthetic HTs bound specifically to rat synaptosomes and induced paralysis in mice
- HT-4 caused rapid and strong respiratory distress in neonatal mice

Figure 1

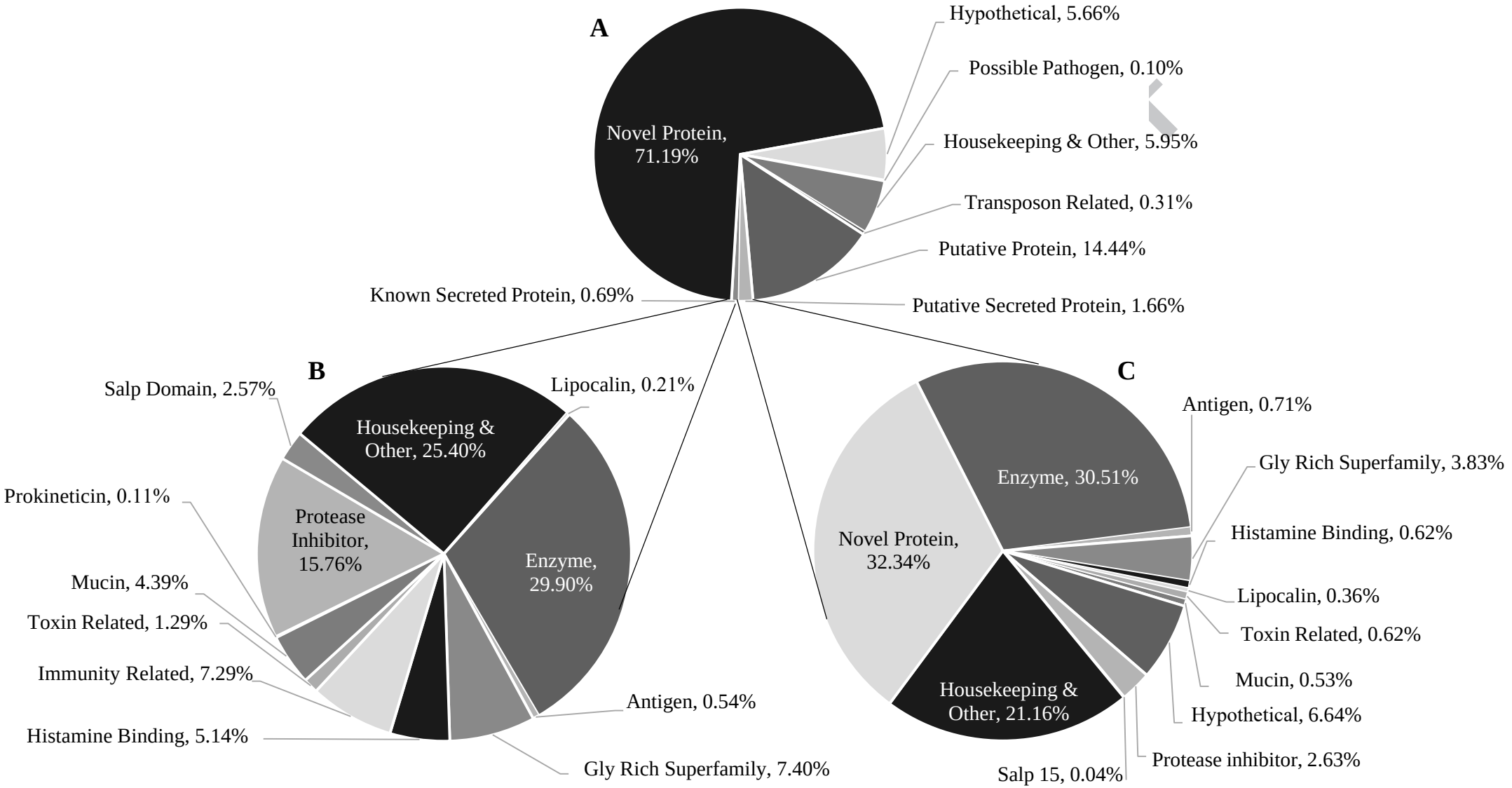


Figure 2

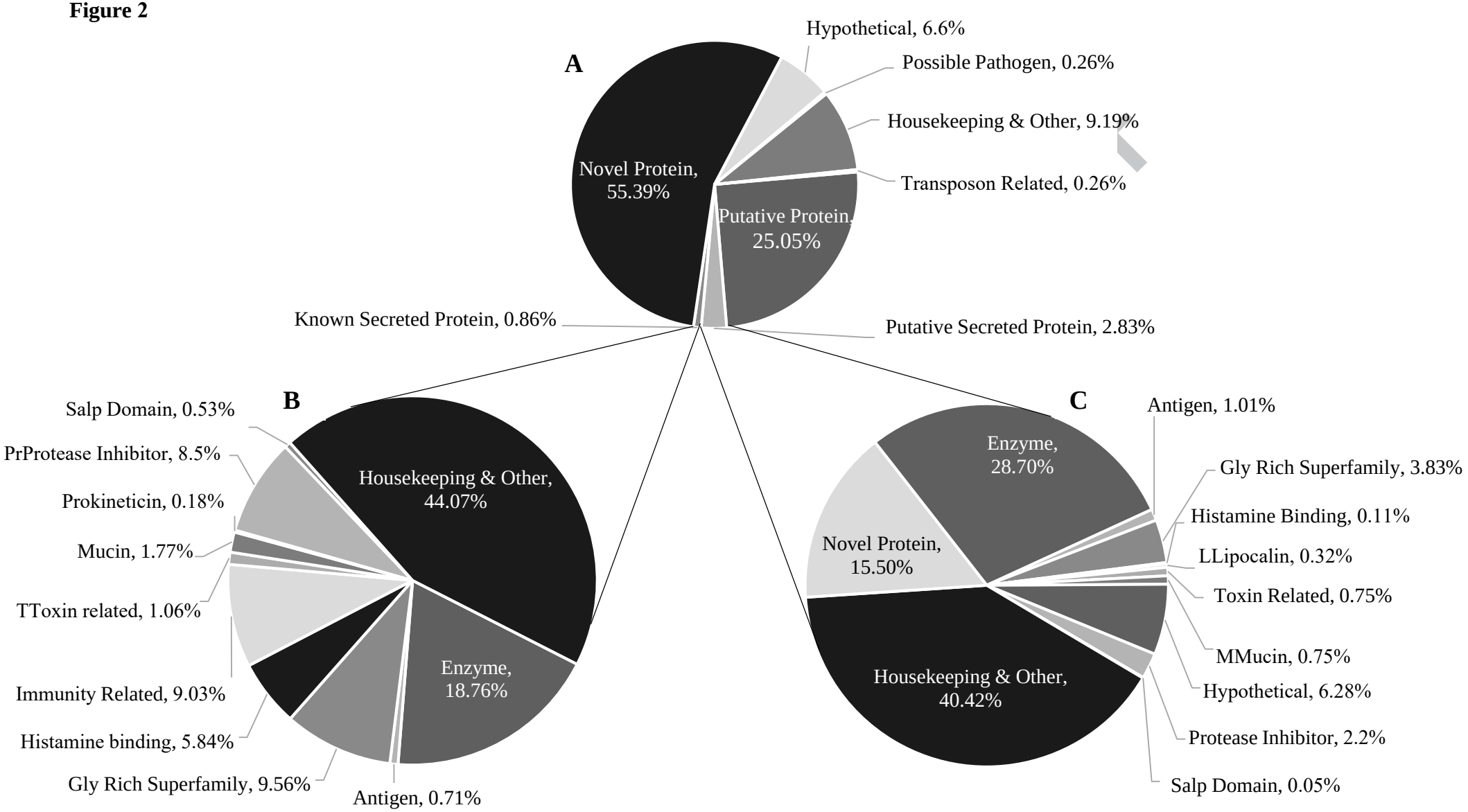
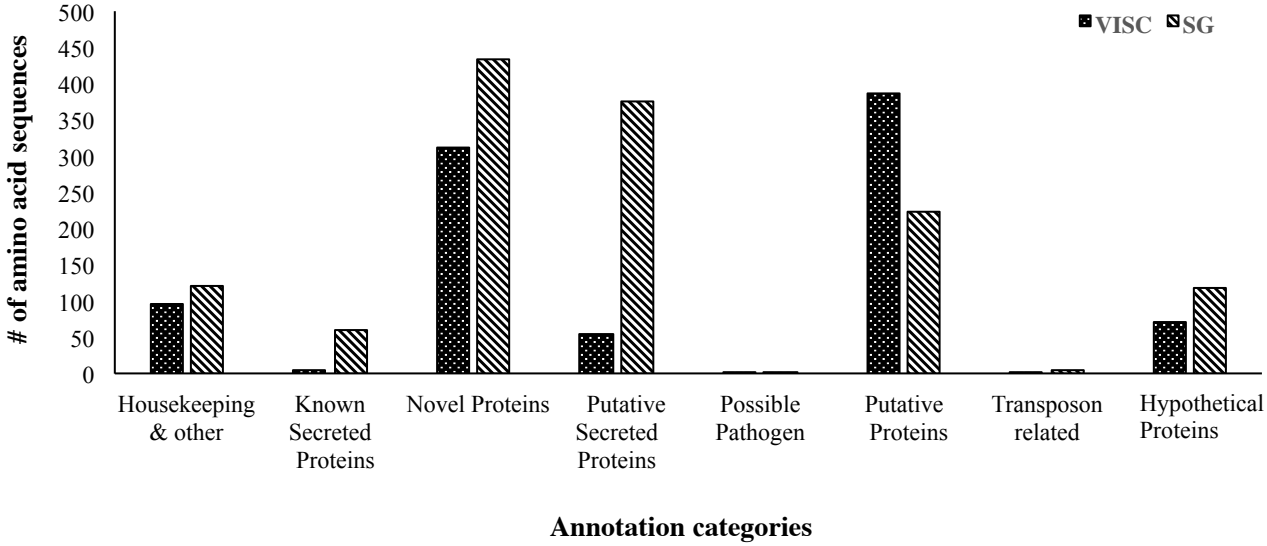
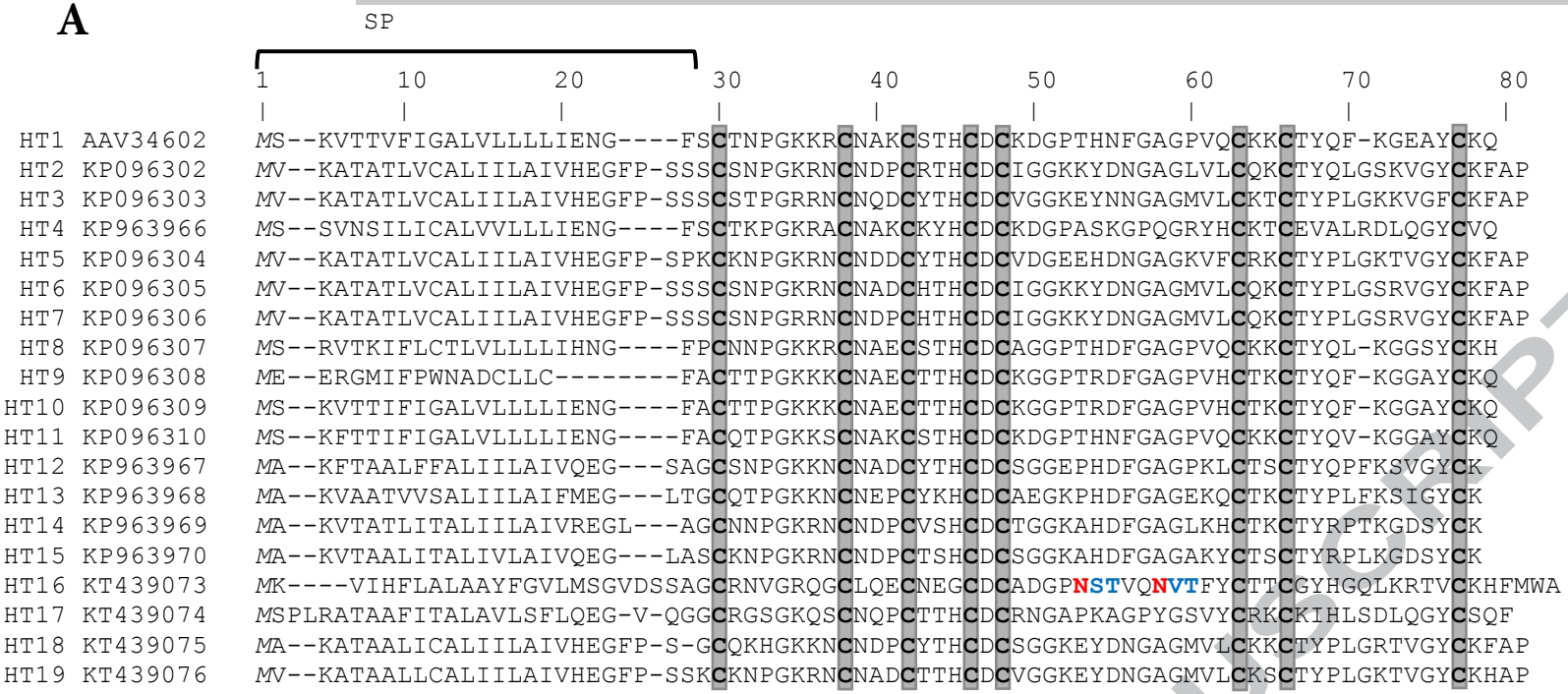


Figure 3



A



B

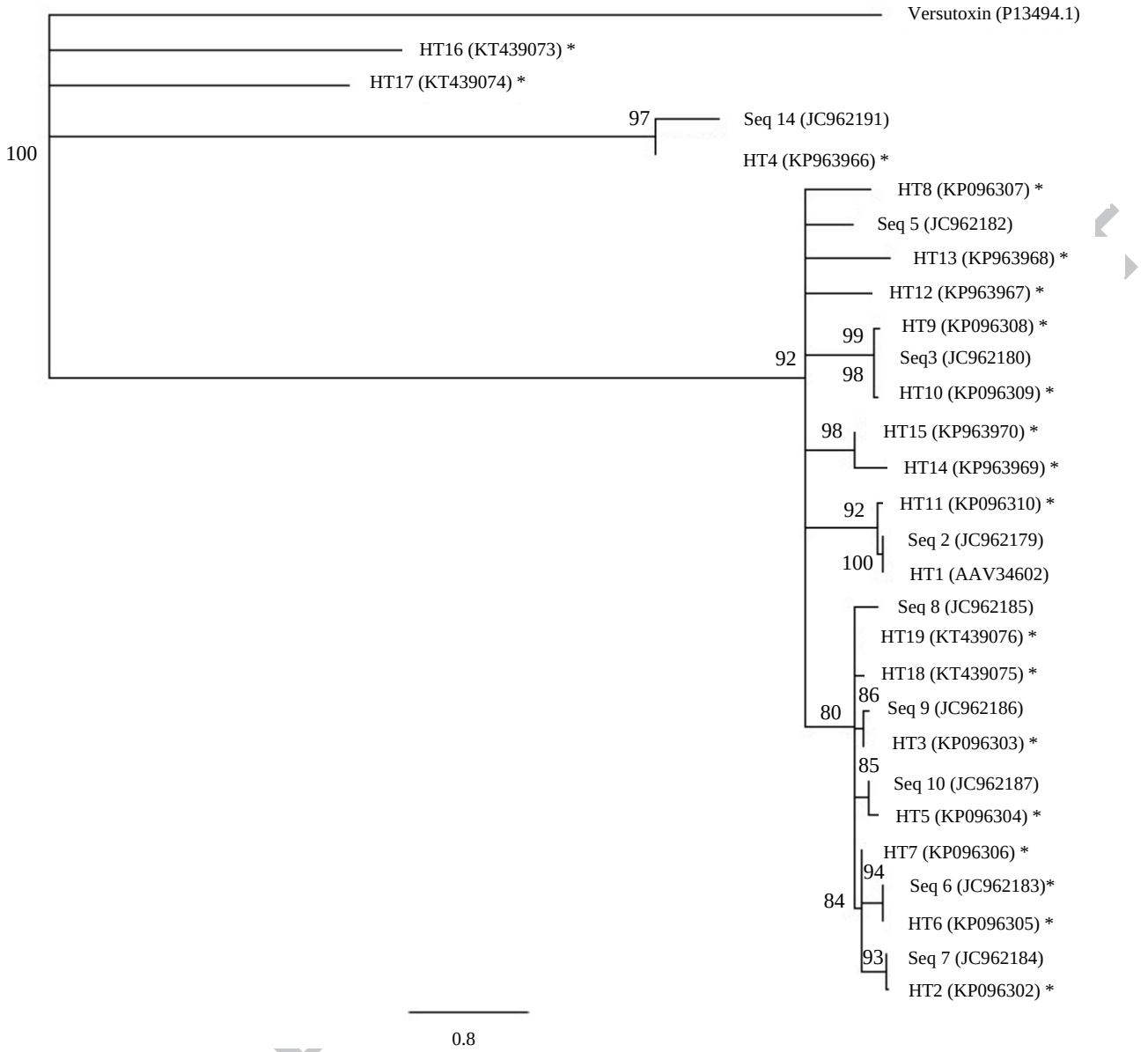
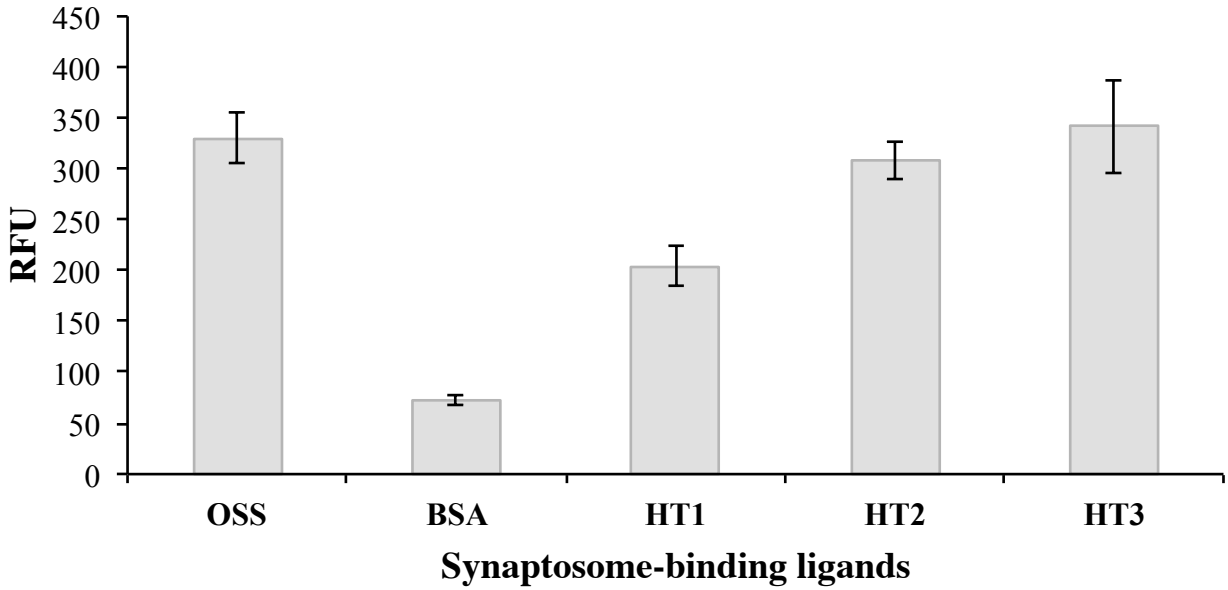
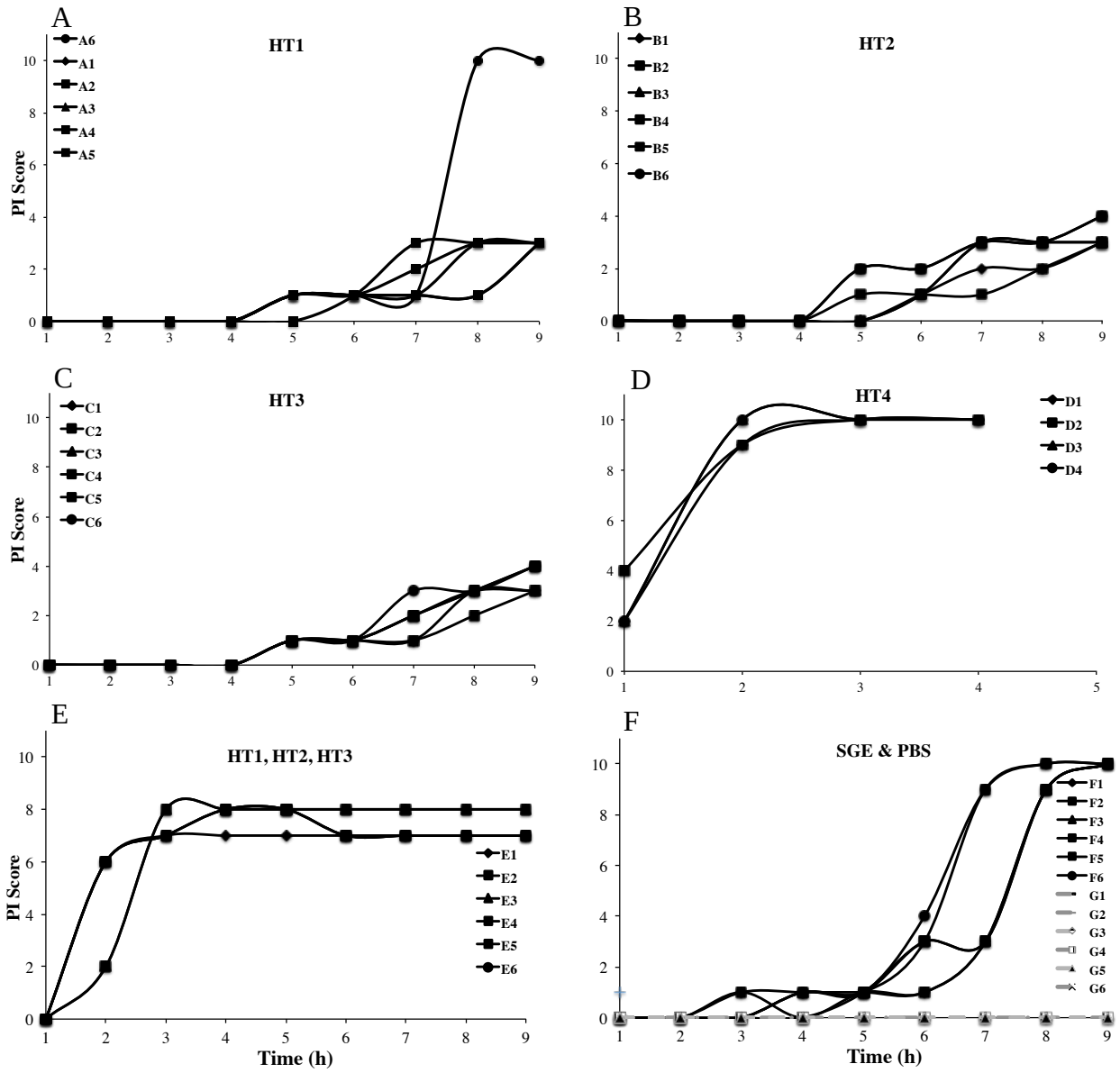
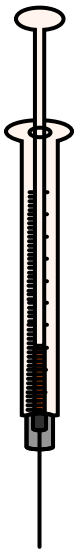


Figure 5







Viscera and salivary gland transcriptome: 19 putative holocyclotoxins (HTs)