



HHS Public Access

Author manuscript

Insect Biochem Mol Biol. Author manuscript; available in PMC 2022 September 22.

Published in final edited form as:

Insect Biochem Mol Biol. 2022 August ; 147: 103800. doi:10.1016/j.ibmb.2022.103800.

A deeper insight into the sialome of male and female *Ochlerotatus triseriatus* mosquitoes

Stephen Lu,

Ines Martin-Martin¹,

Jose M. Ribeiro,

Eric Calvo*

Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA

Abstract

Over the last 20 years, advancements in sequencing technologies have highlighted the unique composition of the salivary glands of blood-feeding arthropods. Further biochemical and structural data demonstrated that salivary proteins can disrupt host hemostasis, inflammation and immunity, which favors pathogen transmission. Previously, a Sanger-based sialome of adult *Ochlerotatus triseriatus* female salivary glands was published based on 731 expressed sequence tag (ESTs). Here, we revisited *O. triseriatus* salivary gland contents using an Illumina-based sequencing approach of both male and female tissues. In the current data set, we report 10,317 DNA coding sequences classified into several functional classes. The translated transcripts also served as a reference database for proteomic analysis of *O. triseriatus* female saliva, in which unique peptides from 101 proteins were found. Finally, comparison of male and female libraries allowed for the identification of female-enriched transcripts that are potentially related to blood acquisition and virus transmission.

1. Introduction

The mosquito *Ochlerotatus triseriatus*, also known as *Aedes triseriatus* (Reinert, 2000), is the main vector of La Crosse virus (LACV) in North America (Borucki et al., 2002). LACV belongs to the *Orthobunyavirus* genus and is the etiological agent of La Crosse encephalitis (LACE) (Beaty and Calisher, 1991). A predominately pediatric disease, LACE has a low mortality rate (0.5%) and generally causes mild symptoms, including fever, lethargy and headaches. However, a small subset of patients display long-term complications such as

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author. ecalvo@niaid.nih.gov (E. Calvo).

¹Current address: Laboratory of Medical Entomology, National Center for Microbiology, Instituto de Salud Carlos III, Madrid, Spain.

Author's contributions

Stephen Lu: Formal analysis, Investigation, Writing – Original draft & Editing.

Ines Martin-Martin: Methodology, Investigation, Writing – Reviewer & Editing.

Jose M. Ribeiro: Formal analysis, Investigation, Writing – Reviewer & Editing.

Eric Calvo: Conceptualization, Methodology, Writing – Reviewer & Editing, Funding acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ibmb.2022.103800>.

seizures and cognitive impairment (Rust et al., 1999; McJunkin et al., 2001). LACV is maintained in nature through a cycle involving *O. triseriatus* and small mammals such as chipmunks and tree squirrels. Additionally, infected female mosquitoes can vertically transmit LACV to their progeny (Beatty and Calisher, 1991).

It is well established that mosquito saliva contains a complex mixture of proteins with several potent pharmacological activities that interfere with vertebrate host hemostasis, inflammation and immunity (Ribeiro and Arca, 2009). Moreover, it has been demonstrated that mosquito saliva facilitates arbovirus transmission to vertebrate hosts (Le Coupanec et al., 2013; Agarwal et al., 2016; Sun et al., 2020), prompting the idea of using salivary proteins as potential targets for vaccine development (Manning et al., 2018). Mosquito saliva also plays a role in sugar feeding (James and Rossignol, 1991; Lehane, 2010).

In the last two decades, advancements in sequencing technologies and improvements in assembly and annotation tools (van Dijk et al., 2014) established a cost-efficient way to explore the salivary gland contents (sialome) from a plethora of hematophagous vectors (Ribeiro and Arca, 2009). Detailed sialome studies of several mosquito species paired with biochemical and structural data highlighted their unique composition and allowed for the identification of specific *Aedes* (Arca et al., 2007; Ribeiro et al., 2016), *Anopheles* (Francischetti et al., 2002; Valenzuela et al., 2003; Calvo et al., 2006a) and *Culex* (Ribeiro et al., 2004, 2018) salivary proteins. However, the composition of *O. triseriatus* salivary glands has been less studied, with only a single Sanger-based transcriptome study published (Calvo et al., 2010) and a handful of salivary activities demonstrated (Reno and Novak, 2005; Ribeiro et al., 1994). Here we present a deeper insight into the salivary gland contents of *O. triseriatus* females using a next generation sequencing method paired with mass spectrometry analysis. We also carried out a sialome analysis of *O. triseriatus* males, and sex comparisons allowed for the identification of female-enriched transcripts that can potentially contribute to blood feeding and virus transmission. This work not only represents the first transcriptome of male *O. triseriatus* mosquitoes but also expands the current repertoire of mosquito salivary proteins, serving as reference for further studies in the absence of an *O. triseriatus* genome.

2. Materials and methods

2.1. Mosquito rearing

The *O. triseriatus* colony originated from larvae collected from used tires in Waterford, CT, US that were generously donated by Dr. Doug E. Brackney (The Connecticut Agricultural Experiment Station, New Haven, CT). Briefly, all mosquito stages were maintained at 28 °C, 80% relative humidity, with a 12-h light/dark cycle. Eggs were collected from blood-fed females on wet fluted filter paper and were stored moist for at least two weeks to allow complete embryo development. Stored eggs were immersed in cups containing distilled water, and hatching was synchronized in a vacuum chamber. Stage 1 larvae were transferred to large trays containing 1 L of distilled water. Two days post-hatch, stage 2 larvae were distributed in trays at low density (~200 larvae per liter of distilled water per tray). Larvae were fed on fish food (TetraMin) until the pupa stage was reached. In our conditions, pupation usually occurred on day 7 post-hatch. Pupae were transferred to cups containing

distilled water without fish food and were placed inside a cage. Adults were fed with sugar (cotton balls immersed in 10% Karo syrup placed on top of the cages). Adults of at least 4 days old were used for blood feeding. The day before blood feeding, the sugar cotton balls were replaced with cotton balls soaked in distilled water. Mosquitoes were fed using artificial membrane feeders (NDS Technologies, Inc, Vineland, NJ, US) covered with Parafilm® M (Amcor, Ann Arbor, MI, US) filled with bovine whole blood in ACD obtained from Lampire Biological Laboratories (Pipersville, PA, US) supplemented with 2.4 mM ATP.

2.2. Salivary glands and saliva collection

Salivary glands from sugar-fed adult mosquitoes (1- to 3-day old) were dissected under a stereomicroscope in sterile PBS pH 7.4 and immediately transferred to an Eppendorf tube containing 100 µl of TRIzol (Invitrogen). Pools of salivary glands (50 pairs from females and 100 from males) were stored at -80 °C until RNA extraction. More male than female salivary glands were used to account for their lower number of transcripts (Ribeiro et al., 2016). Three biological replicates per male and per female were used for transcriptomic analysis. Saliva from starved female *O. triseriatus* mosquitoes was collected as described previously (Kern et al., 2021). Briefly, 5- to 8-day old female *O. triseriatus* mosquitoes were starved for 12 h, transferred to Petri dishes, and sedated on ice for 10 min. Mosquitoes were secured to double-sided tape by sticking their dorsa and wings to the tape. To promote salivation, mosquitoes were intrathoracically injected with 200 nL of 10.8 mg/mL pilocarpine hydrochloride (Sigma P6503, St. Louis, MO, US) and were incubated at 27 °C for 1 h prior to saliva collection. Mosquito salivation was initiated by inserting mouthparts in P10 pipette tips containing 1 µL of mineral oil. The hypopharynx was carefully separated from the outer sheath of the proboscis using fine forceps to encourage a higher yield of saliva. Mouthparts were left partially inserted to avoid submersion of the maxillary palps. Mosquitoes were left to salivate for 1 h, and the resulting emulsion of saliva and mineral oil was expelled into a Protein Lo-Bind tube (Eppendorf, Enfield, CT, US) containing 10 µL of PBS. The sample was centrifuged at 20,000×g for 5 min. The aqueous phase was recovered, and the concentration of the resulting protein solution was measured by spectrophotometry (DS-11, DeNovix). Two saliva samples were recovered from 50 mosquitoes each, containing 1.94 mg/mL and 1.03 mg/mL of protein, respectively.

2.3. Library preparation, sequencing, and analysis

Total RNA from mosquito salivary glands was isolated using TRIzol (Invitrogen) according to the manufacturer's instructions. RNA purity was checked using the NanoPhotometer® spectrophotometer (IMPLEN, CA, US), and RNA integrity and quantitation were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, US). A total amount of 1 µg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNALibrary Prep Kit for Illumina® (NEB, US) following the manufacturer's instructions. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X) or by sonication with Diagenode Bioruptor Pico for breaking RNA strands. First strand cDNA was synthesized

using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of DNA fragment 3' ends, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. To select cDNA fragments of ~150–200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, US). Three microliters of USER Enzyme (NEB, US) were used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The fastq files were trimmed of the Illumina adapters and low-quality sequences ($Q < 20$) using TrimGalore (<https://github.com/FelixKrueger/TrimGalore>), merged into a single file and assembled using Abyss (Simpson et al., 2009) with k values from 25 to 95 (with increments of 10) in single-stranded mode and Trinity (Grabherr et al., 2011) in single-stranded F mode. The assemblies from Abyss and Trinity were combined and filtered using CD-HIT (Fu et al., 2012). Coding DNA sequences (CDS) with an open reading frame (ORF) larger than 150 nucleotides (nt) were extracted based on BLASTp results to a subset of the non-redundant protein database and selected if they began with a methionine and if fragments shared 70% similarity with a matching protein. Additionally, all ORFs starting with a methionine and having at least 40 aa in length were submitted to the signalP program (v. 3.0) (Bendtsen et al., 2004). The fragments that presented a putative signal peptide were mapped to the ORFs, and the most 5' methionine was selected as the starting methionine of the transcript. To assess assembly quality, the BUSCO benchmark for universal single-copy orthologs using the Diptera database was used (Seppey et al., 2019). For annotation, we used an in-house program that scans a vocabulary of ~400 words and their order of appearance in the protein matches from BLASTp results, including their e-values and coverage. Relative quantification of CDS was performed by mapping the library reads to the extracted CDS using the RSEM tool (Li and Dewey, 2011). The annotated CDSs were exported to a hyperlinked Excel spreadsheet.

2.4. Proteomic analysis of mosquito saliva

Saliva samples from *O. triseriatus* females were reduced in solution containing 50 mM HEPES, pH 8.0, 10% acetonitrile and 5 mM DTT at 37 °C for 40 min. After cooling to room temperature, the samples were supplemented with iodoacetamide to give a final concentration of 15 mM. After 15 min of alkylation, 200 ng of trypsin were added, and the samples were incubated at 37 °C for 15 h in a final volume of 40 μ l. The solution was evaporated to near dryness under vacuum at 50 °C. Twenty-five microliters of 0.1% trifluoroacetic acid were added, and the pH was adjusted to 2.5 with the addition of 10% trifluoroacetic acid. Samples with an estimated protein content of less than 2 μ g were desalted and concentrated with C18 μ Zip Tips. Samples containing up to 10 μ g of protein were desalted with C18 OMIX 10 solid phase extraction tips. The digests were eluted with 0.1% TFA, 50% acetonitrile and dried under vacuum. The peptides were dissolved in 12 μ l 0.1% formic acid, 3% acetonitrile, which was used as the injection solvent. Digested peptides were subjected to the LC-MS analysis using Orbitrap Fusion mass spectrometer (ThermoFisher Scientific, West Palm Beach, FL, US) connected with an EASY nLC 1000

liquid chromatography system. Nano-LC was carried out with a 5 μ L injection onto a PepMap 100 C18 3- μ m trap column (2 cm, ID 75 μ m) and a 2 μ m PepMap RSLC C18 column (25 cm, ID 75 μ m), both from ThermoFisher Scientific. The LC was operated at a 300 μ L/min flow rate with a 100-min linear gradient from 100% solvent A (0.1% formic acid, and 99.9% water) to 40% solvent B (0.1% formic acid, 20% water, and 79.9% acetonitrile) followed by a column wash. A standard data-dependent acquisition was performed with a full MS spectrum and obtained by the Orbitrap for m/z 400–2000 at the resolution of 120000 with EASY-IC calibration. The precursor ions, with charges from 2 to 8, were selected, isolated (1.6 m/z window), fragmented by CID, then scanned by the Ion Trap. Survey scans were performed every 2 s, and the dynamic exclusion was enabled for 30 s.

Acquisitions were searched against the NCBI nr (06/2020) proteome using PEAKS v10 (Bioinformatics Solutions Inc, Ontario, CA) and a semitryptic search strategy with tolerances of 6 ppm for MS and 0.5 Da for MS/MS, and carbamidomethylation of cysteine as a fixed modification and oxidation of methionine as a dynamic modification allowing for two missed cleavages. Peptides were filtered with a 0.5% false discovery rate (FDR) using a decoy database approach and a 2 spectral matches/peptide requirement.

2.5. Statistical analysis

Differential expression analysis between females and males was carried out using the EdgeR package (Robinson et al., 2010) for R (Team, 2020). Statistical significance was considered when $p < 0.05$ and FDR < 0.05 . The heat map was made using the gplots package for R.

2.6. Data availability

The transcriptome data was deposited to the National Institute for Biotechnology Information (NCBI) under Bioproject PRJNA793857 and Biosample accessions SAMN24587577. The raw reads were deposited to the Short Reads Archive of the NCBI under accessions SRR17427549 – SRR17427554, and the assembled CDS was deposited to the Transcriptome Shotgun Assembly (TSA) under accession GJYJ00000000. The raw proteomic data was deposited to the ProteomeX change platform under the accession number PXD033588.

3. Results and discussion

3.1. Overall description of the *O. triseriatus* sialome

After removing adapter and low-quality sequences, we obtained a total of 112,826,168 reads. The *de novo* assembly using abyss and trinity generated over 180,000 sequences from which 20,772 CDS were extracted based on their homology with previously identified proteins or by the presence of a putative signal peptide. The relative quantification of each CDS was performed using the RSEM tool. The libraries from female *O. triseriatus* presented $41.3\% \pm 1.4\%$ of mapped reads, which is in range with other sialomes from blood feeders (Ribeiro et al., 2017; Tirloni et al., 2020). Similar values were observed in two male mosquito libraries (40.8% and 43.6%). However, one of our male samples (male sample 1) presented 28.7% of mapped reads and had the highest number of primer

contaminants (~40%) in comparison to the other samples ($24\% \pm 1.4\%$). Therefore, this sample was removed from downstream analyses, including differential expression analysis between males and females.

The remaining unmapped reads (~58.3%) were potentially from the 5' and 3' CDS UTRs, from non-coding RNA, or from any sequence that failed to be extracted because it did not generate an ORF with at least 150 nt, lacked a putative signal peptide, or failed to produce matches against previously deposited proteins. Another possible explanation for the unmapped reads could be the presence of a bias in our *de novo* strategy used to assemble the CDS. To address this issue, we performed a BUSCO analysis of completeness using the Diptera database as reference. The current dataset displayed a completeness of 70.7% (65% single and 5.7% duplicated), 3.1% of fragmentation and 26.2% of missing sequences, which is within range of previous genomic assemblies and other transcriptome assemblies (Tirloni et al., 2020; Waterhouse et al., 2019; Lu et al., 2022), indicating no major bias in our strategy.

For annotation and transcript differential expression analysis between females and males, we selected CDS with at least 150 nt and 3 transcripts per million (TPM), resulting in 10,317 CDS. The heat map plot using the normalized TPM from each CDS resulted in the clusterization of the female samples, while higher variability was found in male libraries (Fig. 1). The annotated CDS were exported to a hyperlinked Excel spreadsheet (Supplementary file 1). Functional classification of the filtered CDS into 27 groups revealed that proteins belonging to the 'secreted' group were the most abundant salivary proteins (Fig. 2) in both males, with an average of 862 CDS representing 52.3% of all mapped reads, and females, with an average of 997 CDS accounting for 57.6% of all mapped reads. Proteins involved in carbohydrate metabolism were the second most abundant functional group with 138 CDS (14.3%) in males and 166 CDS (4.5%) in females (Supplementary Table 1). These findings are in accordance with the previous sialome of adult *O. triseriatus* female mosquitoes (Calvo et al., 2010).

We identified 33 protein families in the 'secreted' class in the current dataset. Similar to other mosquito sialomes (Ribeiro et al., 2016; Valenzuela et al., 2003), the most abundant salivary protein family found in *O. triseriatus* females was the D7 family, with 13 CDS accounting for 44% of all secreted proteins (Table 1), followed by aegyptin-like proteins (4 CDS, 12%) and conserved proteins with unknown function (177 CDS, 12%). In males, conserved proteins with unknown function and unknown proteins (728 CDS) accounted for 75% of all secreted proteins, highlighting the major knowledge gap between male and female salivary proteins.

Comparison of the relative CDS quantification from females and males using the Edge R package identified 285 differentially expressed transcripts ($p < 0.05$ and $FDR < 0.05$), with 200 transcripts upregulated and 85 transcripts downregulated in females (Table 2, Supplementary file 2). As expected, transcripts from the 'secreted' class, which have been previously shown to be related to blood acquisition, presented the highest number of differentially expressed transcripts (Table 3), highlighting the 'magic potion' that blood feeding arthropods possess to counteract host homeostasis and immunity.

3.2. Overall description of the saliva from *O. triseriatus* females

In addition to male and female salivary gland transcriptomics, we performed mass spectrometry on female *O. triseriatus* saliva. After data filtering, peptides from 101 proteins were identified, from which 80 were identified in sample 1 and 91 identified in sample 2 (Supplementary file 3). Despite the presence of proteins that were only identified in one of the samples, the heat map plot of the Log_2 (AUC) revealed a similar pattern between the two samples (Fig. 3). Additionally, most proteins that were observed in a single sample were identified in low abundances, suggesting that these proteins are present in lower concentrations in *O. triseriatus* saliva, which could explain why they were not found in both samples.

Further inspection of the LC-MS data revealed, in accordance with the transcriptome, that the D7 family is the most abundant and most represented salivary protein from *O. triseriatus* saliva. In the current data set, we identified 11 proteins belonging to the D7 family, and together they accounted for 50.6% of all proteins identified by the LC-MS analysis (Table 4). Other highly abundant protein families we identified included members of the SGS family (7 proteins, 6.9%), aegyptin-like proteins (5 proteins, 10.6%), Antigen-5-like proteins (4 proteins, 6.5%), inhibitors from the serpin, TIL (trypsin inhibitor-like) and Kazal sub-families (7 proteins, 3.2%) and apyrases (4 proteins, 4.2%). These proteins are commonly reported in the saliva of other mosquitoes (Almeras et al., 2009; Fontaine et al., 2012). Together, these six families account for 82% of all proteins identified. Interestingly, the same protein families were found to be among the most abundant in our transcriptome analysis, suggesting a correlation between transcripts with high TPM values and their protein counterparts. To compare our proteomic and transcriptomic data, we determined the Spearman correlation coefficient between the relative quantification of proteins identified in the LC-MS analysis by the relative quantification of their transcripts (TPM), resulting in a $\text{Rho} = 0.5274$ with $p = 1.037 \times 10^{-8}$ (Supplementary Fig. 1). This result indicates a positive correlation between transcripts and proteins.

3.3. Transcripts differentially expressed in the salivary glands of *O. triseriatus* females and males

Below we present a discussion of the secreted salivary protein families found differently expressed between male and female mosquitoes, and their possible functions in *O. triseriatus* physiology.

3.3.1. 13 kDa family—Members of the 13 kDa protein family have been previously reported in sialome studies of *Ae. albopictus* (Arca et al., 2007), *Ae. aegypti* (Ribeiro et al., 2016), *Cx. quiquefasciatus* (Ribeiro et al., 2004) and the previous *O. triseriatus* sialome (Calvo et al., 2010). Although named the 13 kDa family, mature proteins of this family range from 11 to 14 kDa and share 10 conserved cysteine residues. In this study, we identified two 13 kDa CDS in females (TPM = 3441 and 13.31) and one in males (TPM = 0.11) (Supplementary file 1). Differential expression analysis revealed that both transcripts are upregulated in females (Table 3). Additionally, unique peptides from members of this family were identified by LC-MS (Supplemental Table 3), confirming their presence in female saliva and suggesting a potential role in blood acquisition. Interestingly, members of this

family were also reported in the salivary glands of male *Anopheles* mosquitoes, and it was suggested that they may act as an antimicrobial agent or be related to sugar feeding (Arca et al., 2007). The biological function of this protein remains elusive thus far.

3.3.2. Aegyptin/30 kDa salivary allergen—Members of the 30 kDa protein allergen family are among the most abundant proteins reported in sialome studies of mosquitoes, including *Anopheles* sp. (Francischetti et al., 2002; Calvo et al., 2007a) and *Aedes* sp. (Ribeiro et al., 2007). In the current data set, we identified four CDS belonging to this protein family, and all were upregulated in female salivary glands (LogFC ranged from 8.2 to 16.7) compared to male salivary glands (Supplementary file 1). Mass spectrometry analysis of *O. triseriatus* female saliva also identified unique peptides belonging to members of this family, indicating that these proteins are present in the saliva and are injected into the host during blood feeding. Together, they were the third most abundant protein family in mosquito saliva (Table 4). The function of aegyptin, originally identified in *Ae. aegypti* salivary glands, is the most well understood. Aegyptin inhibits platelet aggregation by specifically binding to collagen, disrupting its interaction with the physiological ligands GPVI, integrin $\alpha 2\beta 2$ and the Von Willebrand factor. However, aegyptin has no effect on platelet aggregation induced by thrombin and other agonists (Calvo et al., 2007b). Functional studies using transgenesis showed that mosquitoes with reduced aegyptin expression in their salivary glands had longer probing times and significantly reduced feeding success (Chagas et al., 2014). Interestingly, a second member of the 30 kDa protein, the *Ae. albopictus* aegyptin-like protein (alALP), was recently reported, and its function was characterized *in vitro* and *in vivo*. Recombinant alALP prolonged the activate partial thromboplastin time (APPT), prothrombin time (PT) and thrombin time (TT) *in vitro*, in addition to prolonging tail bleed time in a mouse model. Although these findings suggest alALP is an anticoagulant, its mechanism of action still remains to be elucidated (Li et al., 2020). Despite the difference between aegyptin and alALP, it appears that the 30 kDa protein family plays an important role during blood acquisition by modulating host homeostasis.

3.3.3. 34 kDa family—Members of the 34 kDa protein family have been reported in *Aedes* sp. (Arca et al., 2007; Ribeiro et al., 2016) and *Culex* sp. (Ribeiro et al., 2004, 2018), but not in *Anopheles* (Calvo et al., 2006a, 2007a; Arca et al., 2005). In the previous *O. triseriatus* sialome, a single CDS with 61% identity to an *Ae. albopictus* sequence was reported (Calvo et al., 2010). Here, we report three CDS classified as members of the 34 kDa protein family. All three CDS were found exclusively in female salivary glands with high TPM values (617–2479) (Supplementary file 1). Moreover, unique peptides from two proteins were also identified in the LC-MS analysis of *O. triseriatus* saliva (Table 4), confirming that proteins from this family are secreted. Together they represent the fourth most abundant protein family in the mosquito saliva. Similar findings were reported in *Ae. aegypti*, in which members of this family were overexpressed 50-fold in female salivary glands compared to males (Ribeiro et al., 2016). As expected, our differential expression analysis revealed that all three CDS were upregulated in females (Table 3). Although the function of this family remains unknown, their high abundance in female salivary glands suggests they are related to hematophagy. It is worth mentioning that a member of this family of salivary protein was found to enhance dengue virus replication in cultured

keratinocytes (Surasombatpattana et al., 2014). This work suggests that the 34 kDa protein may suppress antiviral immune response in the earliest stages of infection.

3.3.4. 41 kDa family—In the current data set, we classified four CDS as members of the 41 kDa protein family in both males and females (Supplemental file 1), while the previous *O. triseriatus* sialome reported one CDS belonging to this family (Calvo et al., 2010). Unique peptides from three were also identified in low concentrations from *O. triseriatus* female saliva by LC-MS (Table 4). Members of this family were also reported in the sialome of *Ae. aegypti* males and females (Ribeiro et al., 2016) and in the non-blood feeding mosquito *Toxorhynchites amboinensis* (Calvo et al., 2008). Additionally, DEG (Differentially Expressed Genes) analysis revealed the presence of one CDS (Ocseq_188668) downregulated in females (Table 3). To date, no members of this family have been functionally characterized; however, the abundance of this protein family in both male and female mosquitoes suggest a potential role related to sugar feeding rather than to blood acquisition (Calvo et al., 2010).

3.3.5. Antigen 5—Antigen 5-like proteins belong to the CAP (cysteine-rich secretory proteins, antigen 5 and pathogenesis-related protein 1) superfamily that is widely distributed in different organisms (Tadokoro et al., 2020). Although this family is commonly reported in the sialome of blood feeding arthropods (Tirloni et al., 2020; Lu et al., 2022; Charlab et al., 1999), their functions are better understood in hornets and wasps because they are the most abundant allergens in vespid venom (Lu et al., 1993; Monsalve et al., 2020). In mosquitoes, antigen 5-like proteins have been reported in the salivary glands of both males and females (Ribeiro et al., 2016; Arca et al., 2005), including in the non-blood feeding mosquito *T. amboinensis* (Calvo et al., 2008). In the first *O. triseriatus* sialome (Calvo et al., 2010), only a single truncated antigen 5-like sequence was identified. In the present study, we identified 10 (4 complete and 6 truncated CDS) putative antigen 5-like transcripts in both male and female salivary glands (Supplemental file 1). Differential expression analysis revealed the presence of one upregulated and one downregulated CDS in females (Table 3). From the 10 CDS identified in female salivary glands, unique peptides from four were also found in female saliva (Table 4). The role of antigen 5-like proteins in mosquito physiology is still uncharacterized. Despite the presence of antigen 5-like proteins in both males and females, a recent transcriptome comparing the salivary glands of *Ae. aegypti* males and females reported the presence of antigen 5-like sequences 10- and 100-fold overexpressed in females (Ribeiro et al., 2016), suggesting a possible role in blood acquisition. Additionally, the role of antigen 5-like molecules in blood acquisition has been demonstrated in other blood feeders. In the horsefly *Tabanus yao*, antigen 5-like proteins were isolated from salivary glands and inhibited thrombosis, angiogenesis, and platelet aggregation (Ma et al., 2009, 2010; Xu et al., 2008). Moreover, in *Dipetalogaster maxima* and *Triatoma infestans*, antigen 5-like proteins had antioxidant potential and disrupted platelet aggregation induced by low concentrations of collagen (Assumpcao et al., 2013). Together, these studies suggest the potential role of antigen 5-like proteins in blood acquisition. In a recent study, Valenzuela-Leon et al. (2022) found that a member of this family bound to the Zika virus envelope protein with nanomolar affinities. However, this interaction had no effect on viral replication in cultured endothelial cells and keratinocytes.

3.3.6. Apyrases—Apyrases are enzymes that hydrolyse ADP and ATP to AMP (Plesner, 1995) and are further classified into three subfamilies: the 5'-nucleotidase family, the *Cimex*-type apyrases and the homologs of the human B cell antigen CD-39 (Hughes, 2013). Apyrases are commonly reported in the saliva of hematophagous vectors, including in ticks (Mans et al., 1998), kissing bugs (Sarkis et al., 1986) and fleas (Cheeseman, 1998; Ribeiro et al., 1990). In mosquitoes, apyrases belonging to the 5'-nucleotidase subfamily (Champagne et al., 1995) have been found in several species (Reno and Novak, 2005; Ribeiro et al., 1984; Cupp et al., 1994). Early studies comparing the composition of male and female *Ae. aegypti* salivary glands revealed that the apyrase protein family is expressed specifically in the adult female (Smartt et al., 1995). In agreement with previous studies, in the current data set, we identified five CDS belonging to the 5'-nucleotidase subfamilies exclusively in female samples with a high TPM range (3.1–2155) (Supplementary file 1). Mass spectrometry analysis of *O. triseriatus* female saliva revealed the presence of unique peptides from four different apyrases (Table 4), confirming that these proteins are secreted.

Apyrases from different blood feeders facilitate blood acquisition by inhibiting ADP-induced platelet aggregation at the bite site (Ribeiro et al., 1985; Valenzuela et al., 1998; Sun et al., 2006). Moreover, phylogenetic analysis of apyrases from different blood-feeding arthropods supports the hypothesis that members of this family were selected independently several times during evolution (Hughes, 2013), reinforcing their key role in blood acquisition.

3.3.7. D7 family—The D7 protein family was first described as a novel protein family in *Ae. aegypti* salivary glands (James et al., 1991). Later, it was shown that D7 proteins are, in fact, distantly related to the odorant-binding protein (OBP) superfamily (Hekmat-Scafe et al., 2000). Currently, the D7 family is subclassified into short forms that contain a single OBP domain (15–20 kDa), or long forms that contain two OBP domains (24–30 kDa). Both forms are commonly reported in mosquito sialome studies as the most abundant secreted protein family. Interestingly, in *Anopheles*, short forms are the most highly expressed transcripts, while in *Aedes*, the long forms are the most abundant (Arca et al., 2007; Ribeiro et al., 2016; Francischetti et al., 2002; Calvo et al., 2007a). In addition to mosquito D7s, sand fly D7s have been described (Valenzuela et al., 2002; Jablonka et al., 2019). The previous *O. triseriatus* sialome reported four D7 family members (two short and two long forms) (Calvo et al., 2010), and in the current study, we identified a total of four long and nine short form D7s (Supplemental file 1). As expected, DEG analysis revealed that 11 of the 13 D7 transcripts were upregulated in females (Table 3). LC-MS analysis also identified unique peptides from nine of the 13 D7 proteins in the saliva of *O. triseriatus* females, and the long forms were the most abundant, corresponding to 30% of all quantified salivary proteins. The short D7 was the second most abundant protein (Table 4).

The role of D7 proteins in blood-feeding has been under intense study over the last few years, resulting in the characterization of several members of this family across different mosquito species. Early studies with short D7 proteins from the malaria vector *An. gambiae* shed light on their ability to bind small agonists relevant to host hemostasis (serotonin, epinephrine, norepinephrine and histamine). Structural data revealed that the D7 domain has two additional helices when compared to the classic OBP domain (Martin-Martin et

al., 2020a). Moreover, D7-complexed crystals allowed for the identification of key residues involved in ligand stabilization (Ribeiro and Arca, 2009; Calvo et al., 2006b; Mans et al., 2007). The functions of *An. gambiae* D7 long forms (AngaD7L1, L2, and L3) were recently discovered (Smith et al., 2022). AngaD7L1 binds leukotriene C4 and thromboxane A2 analog U-46619 while AngaD7L2 weakly binds leukotrienes B4 and D4. Interestingly, AngaD7L3 is the first D7 long form from anophelines to bind serotonin. This protein does not bind lipids. AngaD7L2 had a dose-dependent anticoagulant effect via the intrinsic coagulation pathway by interacting with factors XII, XIIa, and XI. Although anopheline D7 short forms also display anti-coagulant activity, AngD7L2 is the only D7 long form to act as a direct anticoagulant. AnSt-D7L1, a D7 long form from *An. stephensi*, bound cysteinyl leukotrienes and inhibited platelet activation by binding to thromboxane A₂ (Alvarenga et al., 2010). In *Aedes*, only long-form D7s have been characterized, and they bind cysteinyl leukotrienes at their N-terminal domain, as well as biogenic amines at their C-terminal domain (Calvo et al., 2009; Martin-Martin et al., 2020b, 2021). In *Cx. quinquefasciatus*, the long D7 form CxD7L1 bound the platelet aggregator ADP while CxD7L2 retained biogenic amine and biolipid binding activity also seen in *Aedes* D7 long forms (Martin-Martin et al., 2020a). The function of short D7s in *Aedes* remains unknown, and it has been speculated that these proteins do not have a biogenic binding function because they lack the conserved residues relevant for this interaction (Mans et al., 2007). Yet the presence of several upregulated transcripts in female salivary glands suggests their potential role in blood acquisition.

3.3.8. Protease inhibitors—Serine peptidase inhibitors are frequently reported in sialomes of blood-feeding vectors and are usually associated with the inhibition of host complement and blood clotting cascades (Ribeiro and Arca, 2009). Here we report several transcripts belonging to different types of serine peptidase inhibitors, including serpins, Kazal, Kunitz, pacifastins and the TIL (trypsin inhibitor-like) subfamilies in both male and female salivary glands (Supplemental file 1). The LC-MS analysis of female saliva identified seven proteins classified as protease inhibitors, from which five were of the serpin subfamily (Supplemental file 3), confirming that these inhibitors are secreted into the host.

Early studies with salivary gland extracts from different anophelines and culicines highlighted the ability of mosquito saliva to interfere with human blood coagulation. Additional experiments suggested that the saliva of anophelines targeted thrombin, while the saliva of culicines targeted factor Xa (Stark and James, 1996). This was further confirmed by the isolation of *Ae. aegypti* and *Ae. albopictus* factor Xa that inhibited serpin (Stark and James, 1998; Calvo et al., 2011) and the anti-thrombin anopheline from *An. albimanus* (Valenzuela et al., 1999). Our DEG analysis revealed the presence of two upregulated transcripts in females belonging to the serpin family, Ocseq_196173 (LogFC = 5.6) and Ocseq_184646 (LogFC = 7.6), suggesting a similar role in blood acquisition as observed with *Ae. aegypti* serpin. Interestingly, two transcripts containing TIL domains were downregulated in females, Ocseq_203815 (LogFC = -2.3) and Ocseq_192061 (LogFC = -2.1). In a previous study, similar transcripts were identified in *An. gambiae* adult males, and it was proposed that these inhibitors act as antimicrobial peptides (Calvo et al., 2006a).

3.3.9. Odorant-binding proteins (OBP)—OBPs are small proteins with a hydrophobic binding pocket commonly found in the sensory organs and salivary glands of arthropods. In the sensory organs, OBPs play a key role in insect chemoreception, contributing to the insect's ability to locate food, hosts and mating partners (Zwiebel and Takken, 2004). In our sialome analysis, we identified 13 transcripts with similarities to *Ae. aegypti* OBPs. However, all transcripts presented low TPM values, with a maximum value of 145 (Supplementary file 1). Moreover, only one OBP was identified in our LC-MS analysis (Table 4), indicating that such proteins are in small concentrations in mosquito saliva. With the exception of the D7 family, little is known about salivary OBPs. A microarray study comparing salivary transcripts from *Ae. aegypti* 1, 3, 24 and 48 h post-feeding (hpf) identified an OBP-like transcript to be downregulated at 1, 3 and 48 hpf and upregulated at 24 hpf in comparison to naïve-mosquitoes (Thangamani and Wikel, 2009). Additionally, our differential expression analysis revealed the presence of two transcripts (Ocseq_221634 and Ocseq_221635) upregulated in females (Table 4), suggesting a possible role in mosquito blood-feeding physiology.

3.3.10. Salivary gland surface protein (SGS)—Members of the SGS family are multidomain proteins similar to bacterial proteins, and they appear to be exclusive to mosquitoes (Arca et al., 2005; Korochkina et al., 2006). Interestingly, *Ae. aegypti* aaSGS1 was only found in female salivary glands and was not modulated during blood-feeding. Moreover, the authors demonstrated that aaSGS1 is a potential receptor for *Plasmodium gallinaceum* invasion (Korochkina et al., 2006; Kojin et al., 2021). The same study also identified four SGS proteins in *An. gambiae*, with agSGSG4 and agSGS5 transcripts expressed only in female salivary glands and agSGS2 and agSGS3 found in the thorax (where the salivary glands are located in both males and females). In contrast to aaSGS1, a second study exploring the role of salivary SGS from *An. gambiae* revealed that agSGS4 and agSGS5 transcripts increase when a blood meal is provided (King et al., 2011).

In our current data set, we identified seven SGS-like transcripts upregulated in females. LC-MS identified unique peptides from all seven proteins in mosquito saliva (Supplementary files 1 and 3), confirming that, like *Ae. aegypti* and *An. gambiae*, the *O. triseriatus* SGS proteins are also secreted into the host. In addition to the first SGS protein reported in *Ae. aegypti*, a recent sialome study comparing *Ae. aegypti* males and females identified 10 SGS-like transcripts overexpressed (over 100-fold) in female salivary glands (Ribeiro et al., 2016). The sex differences, along with the distinct expression profiles in mosquito tissues, suggest that SGS proteins play multiple roles in mosquito physiology.

3.3.11. Unknown proteins—In addition to the previously described families, we identified differentially expressed transcripts that we failed to classify into other protein families. These CDS either presented similarities to previously deposited sequences with no known function or are novel sequences that have low or no homology to other sequences. Therefore, they are potential new targets for future studies. In the current data set, we identified 16 upregulated secreted transcripts and 8 downregulated transcripts in female salivary glands (Table 4).

Two upregulated transcripts in female salivary glands, Ocseq_181581 (LogFC = 11.4) and Ocseq_182568 (LogFC = 8.4), were similar to a putative secreted protein in *Ae. aegypti* (Supplementary Fig. 2). Unique peptides from both proteins were also found in *O. triseriatus* female saliva (Supplementary file 3). Similarly, the CDS seqSigP-183016 was upregulated in female salivary glands (LogFC = 17.5) and found in saliva. This CDS displayed a high degree of homology with previously deposited putative secreted proteins from *Ae. albopictus* and *Ae. aegypti* (Supplementary Fig. 3). Together, these data support the hypothesis that these proteins are secreted into saliva and injected into the host, potentially contributing to blood acquisition and pathogen transmission.

4. Conclusion

In 2010, a Sanger-based sialome study of adult female *O. triseriatus* reported 159 protein sequences of which 66 encoded putative secreted proteins (Calvo et al., 2010). Here, using an Illumina-based approach, we reported 10,317 total CDS of which 1312 were considered secreted proteins, providing a more complete overview of *O. triseriatus* salivary gland contents. Mass spectrometry analysis of mosquito saliva identified unique peptides from several proteins classified as secreted, confirming that these abundant salivary proteins are secreted and injected into the host. Finally, sequencing male salivary glands allowed for the identification of female enriched secreted transcripts that are potentially associated with blood acquisition and virus transmission. In the absence of a *O. triseriatus* genome, the data presented here serves as a comprehensive reference for *O. triseriatus* male and female sequences for future studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Brian Bonilla for excellent technical support and mosquito rearing. We also thank Glenn A. Nardone, Renee Olano and Motoshi Suzuki for conducting the mass spectrometry experiments. We also thank Dr. Adeline E. Williams for editorial assistance. The graphical abstract presented in this article was created with BioRender.com. This work was supported by the Intramural Research Program of the Division of Intramural Research (AI001246 and AI000810), National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH). This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

References

- Agarwal A, Joshi G, Nagar DP, Sharma AK, Sukumaran D, Pant SC, Parida MM, Dash PK, 2016. Mosquito saliva induced cutaneous events augment Chikungunya virus replication and disease progression. *Infect. Genet. Evol* 40, 126–135. [PubMed: 26925703]
- Almeras L, Orlandi-Pradines E, Fontaine A, Villard C, Boucomont E, de Senneville LD, Baragatti M, Pascual A, Pradines B, Corre-Catelin N, Pages F, Reiter P, Rogier C, Fusai T, 2009. Sialome individuality between *Aedes aegypti* colonies. *Vector Borne Zoonotic Dis* 9, 531–541. [PubMed: 18847318]
- Alvarenga PH, Francischetti IM, Calvo E, Sa-Nunes A, Ribeiro JM, Andersen JF, 2010. The function and three-dimensional structure of a thromboxane A2/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. *PLoS Biol* 8, e1000547. [PubMed: 21152418]

- Arca B, Lombardo F, Valenzuela JG, Francischetti IM, Marinotti O, Coluzzi M, Ribeiro JM, 2005. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. *J. Exp. Biol* 208, 3971–3986. [PubMed: 16215223]
- Arca B, Lombardo F, Francischetti IM, Pham VM, Mestres-Simon M, Andersen JF, Ribeiro JM, 2007. An insight into the sialome of the adult female mosquito *Aedes albopictus*. *Insect Biochem. Mol. Biol* 37, 107–127. [PubMed: 17244540]
- Assumpcao TCF, Ma D, Schwarz A, Reiter K, Santana JM, Andersen JF, Ribeiro JMC, Nardone G, Yu LL, Francischetti IMB, 2013. Salivary antigen-5/CAP family members are Cu²⁺-dependent antioxidant enzymes that scavenge O₂(-). and inhibit collagen-induced platelet aggregation and neutrophil oxidative burst. *J. Biol. Chem* 288, 14341–14361. [PubMed: 23564450]
- Beatty BJ, Calisher CH, 1991. *Bunyaviridae—Natural History* Springer.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S, 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol* 340, 783–795. [PubMed: 15223320]
- Borucki MK, Kempf BJ, Blitvich BJ, Blair CD, Beatty BJ, 2002. La Crosse virus: replication in vertebrate and invertebrate hosts. *Microb. Infect* 4, 341–350.
- Calvo E, Pham VM, Lombardo F, Arca B, Ribeiro JM, 2006a. The sialotranscriptome of adult male *Anopheles gambiae* mosquitoes. *Insect Biochem. Mol. Biol* 36, 570–575. [PubMed: 16835022]
- Calvo E, Mans BJ, Andersen JF, Ribeiro JM, 2006b. Function and evolution of a mosquito salivary protein family. *J. Biol. Chem* 281, 1935–1942. [PubMed: 16301315]
- Calvo E, Dao A, Pham VM, Ribeiro JM, 2007a. An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. *Insect Biochem. Mol. Biol* 37, 164–175. [PubMed: 17244545]
- Calvo E, Tokumasu F, Marinotti O, Villeval JL, Ribeiro JMC, Francischetti IMB, 2007b. Aegyptin, a novel mosquito salivary gland protein, specifically binds to collagen and prevents its interaction with platelet glycoprotein VI, integrin alpha2beta1, and von Willebrand factor. *J. Biol. Chem* 282, 26928–26938. [PubMed: 17650501]
- Calvo E, Pham VM, Ribeiro JM, 2008. An insight into the sialotranscriptome of the non-blood feeding *Toxorhynchites amboinensis* mosquito. *Insect Biochem. Mol. Biol* 38, 499–507. [PubMed: 18405828]
- Calvo E, Mans BJ, Ribeiro JM, Andersen JF, 2009. Multifunctionality and mechanism of ligand binding in a mosquito antiinflammatory protein. *Proc. Natl. Acad. Sci. U. S. A* 106, 3728–3733. [PubMed: 19234127]
- Calvo E, Sanchez-Vargas I, Kotsyfakis M, Favreau AJ, Barbian KD, Pham VM, Olson KE, Ribeiro JM, 2010. The salivary gland transcriptome of the eastern tree hole mosquito, *Ochlerotatus triseriatus*. *J. Med. Entomol* 47, 376–386. [PubMed: 20496585]
- Calvo E, Mizurini DM, Sa-Nunes A, Ribeiro JM, Andersen JF, Mans BJ, Monteiro RQ, Kotsyfakis M, Francischetti IM, 2011. Alboserpin, a factor Xa inhibitor from the mosquito vector of yellow fever, binds heparin and membrane phospholipids and exhibits antithrombotic activity. *J. Biol. Chem* 286, 27998–28010. [PubMed: 21673107]
- Chagas AC, Ramirez JL, Jasinskiene N, James AA, Ribeiro JM, Marinotti O, Calvo E, 2014. Collagen-binding protein, Aegyptin, regulates probing time and blood feeding success in the dengue vector mosquito, *Aedes aegypti*. *Proc. Natl. Acad. Sci. U. S. A* 111, 6946–6951. [PubMed: 24778255]
- Champagne DE, Smartt CT, Ribeiro JM, James AA, 1995. The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a member of the 5'-nucleotidase family. *Proc. Natl. Acad. Sci. U. S. A* 92, 694–698. [PubMed: 7846038]
- Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM, 1999. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc. Natl. Acad. Sci. U. S. A* 96, 15155–15160. [PubMed: 10611354]
- Cheeseman MT, 1998. Characterization of apyrase activity from the salivary glands of the cat flea *Ctenocephalides felis*. *Insect Biochem. Mol. Biol* 28, 1025–1030. [PubMed: 9887518]
- Cupp EW, Cupp MS, Ramberg FB, 1994. Salivary apyrase in African and New World vectors of *Plasmodium* species and its relationship to malaria transmission. *Am. J. Trop. Med. Hyg* 50, 235–240. [PubMed: 8116818]

- Fontaine A, Fusai T, Briolant S, Buffet S, Villard C, Baudelet E, Pophillat M, Granjeaud S, Rogier C, Almeras L, 2012. Anopheles salivary gland proteomes from major malaria vectors. *BMC Genom* 13, 614.
- Francischetti IM, Valenzuela JG, Pham VM, Garfield MK, Ribeiro JM, 2002. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. *J. Exp. Biol* 205, 2429–2451. [PubMed: 12124367]
- 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. [PubMed: 23060610]
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol* 29, 644–652. [PubMed: 21572440]
- Hekmat-Scafe DS, Dorit RL, Carlson JR, 2000. Molecular evolution of odorant-binding protein genes OS-E and OS-F in *Drosophila*. *Genetics* 155, 117–127. [PubMed: 10790388]
- Hughes AL, 2013. Evolution of the salivary apyrases of blood-feeding arthropods. *Gene* 527, 123–130. [PubMed: 23791653]
- Jablonka W, Kim IH, Alvarenga PH, Valenzuela JG, Ribeiro JMC, Andersen JF, 2019. Functional and structural similarities of D7 proteins in the independently-evolved salivary secretions of sand flies and mosquitoes. *Sci. Rep* 9, 5340. [PubMed: 30926880]
- James AA, Rossignol PA, 1991. Mosquito salivary glands: parasitological and molecular aspects. *Parasitol. Today* 7, 267–271. [PubMed: 15463385]
- James AA, Blackmer K, Marinotti O, Ghosn CR, Racioppi JV, 1991. Isolation and characterization of the gene expressing the major salivary gland protein of the female mosquito, *Aedes aegypti*. *Mol. Biochem. Parasitol* 44, 245–253. [PubMed: 2052024]
- Kern O, Valenzuela Leon PC, Gittis AG, Bonilla B, Cruz P, Chagas AC, Ganesan S, Ribeiro JMC, Garboczi DN, Martin-Martin I, Calvo E, 2021. The structures of two salivary proteins from the West Nile vector *Culex quinquefasciatus* reveal a beta-trefoil fold with putative sugar binding properties. *Curr. Res. Struct. Biol* 3, 95–105. [PubMed: 34235489]
- King JG, Vernick KD, Hillyer JF, 2011. Members of the salivary gland surface protein (SGS) family are major immunogenic components of mosquito saliva. *J. Biol. Chem* 286, 40824–40834. [PubMed: 21965675]
- Kojin BB, Martin-Martin I, Araujo HRC, Bonilla B, Molina-Cruz A, Calvo E, Capurro ML, Adelman ZN, 2021. *Aedes aegypti* SGS1 is critical for *Plasmodium gallinaceum* infection of both the mosquito midgut and salivary glands. *Malar. J* 20, 11. [PubMed: 33407511]
- Korochkina S, Barreau C, Pradel G, Jeffery E, Li J, Natarajan R, Shabanowitz J, Hunt D, Frevert U, Vernick KD, 2006. A mosquito-specific protein family includes candidate receptors for malaria sporozoite invasion of salivary glands. *Cell Microbiol* 8, 163–175. [PubMed: 16367875]
- Le Coupanec A, Babin D, Fiette L, Jouvion G, Ave P, Misse D, Bouloy M, Choumet V, 2013. *Aedes* mosquito saliva modulates Rift Valley fever virus pathogenicity. *PLoS Neglected Trop. Dis* 7, e2237.
- Lehane MJ, 2010. *The Biology of Blood-Sucking in Insects*, 2 ed. Cambridge University Press.
- Li B, Dewey CN, 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf* 12, 323.
- Li XP, Lin D, Zhang Y, Chen SQ, Bai HQ, Zhang SN, Liu WQ, Liang SH, 2020. Expression and characterization of anticoagulant activity of salivary protein aALP from Asian tiger mosquito *Aedes albopictus*. *Trop. Biomed* 37, 116–126. [PubMed: 33612723]
- Lu G, Villalba M, Coscia MR, Hoffman DR, King TP, 1993. Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. *J. Immunol* 150, 2823–2830. [PubMed: 8454859]
- Lu S, Andersen JF, Bosio CF, Hinnebusch BJ, Ribeiro JMC, 2022. Integrated analysis of the sialotranscriptome and sialoproteome of the rat flea *Xenopsylla cheopis*. *J. Proteomics*, 104476.
- Ma D, Wang Y, Yang H, Wu J, An S, Gao L, Xu X, Lai R, 2009. Anti-thrombosis repertoire of blood-feeding horsefly salivary glands. *Mol. Cell. Proteomics* 8, 2071–2079. [PubMed: 19531497]

- Ma D, Gao L, An S, Song Y, Wu J, Xu X, Lai R, 2010. A horsefly saliva antigen 5-like protein containing RTS motif is an angiogenesis inhibitor. *Toxicon* 55, 45–51. [PubMed: 19635491]
- Manning JE, Morens DM, Kamhawi S, Valenzuela JG, Memoli M, 2018. Mosquito saliva: the hope for a universal arbovirus vaccine? *J. Infect. Dis* 218, 7–15. [PubMed: 29617849]
- Mans BJ, Gasper AR, Louw AI, Neitz AW, 1998. Purification and characterization of apyrase from the tick, *Ornithodoros savignyi*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol* 120, 617–624. [PubMed: 14598857]
- Mans BJ, Calvo E, Ribeiro JM, Andersen JF, 2007. The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. *J. Biol. Chem* 282, 36626–36633. [PubMed: 17928288]
- Martin-Martin I, Paige A, Valenzuela Leon PC, Gittis AG, Kern O, Bonilla B, Chagas AC, Ganesan S, Smith LB, Garboczi DN, 2020a. ADP binding by the *Culex quinquefasciatus* mosquito D7 salivary protein enhances blood feeding on mammals. *Nat. Commun* 11, 1–15. [PubMed: 31911652]
- Martin-Martin I, Smith LB, Chagas AC, Sa-Nunes A, Shrivastava G, Valenzuela-Leon PC, Calvo E, 2020b. *Aedes albopictus* D7 salivary protein prevents host hemostasis and inflammation. *Biomolecules* 10.
- Martin-Martin I, Kern O, Brooks S, Smith LB, Valenzuela-Leon PC, Bonilla B, Ackerman H, Calvo E, 2021. Biochemical characterization of AeD7L2 and its physiological relevance in blood feeding in the dengue mosquito vector, *Aedes aegypti*. *FEBS J* 288, 2014–2029. [PubMed: 32799410]
- McJunkin JE, de los Reyes EC, Irazuzta JE, Caceres MJ, Khan RR, Minnich LL, Fu KD, Lovett GD, Tsai T, Thompson A, 2001. La Crosse encephalitis in children. *N. Engl. J. Med* 344, 801–807. [PubMed: 11248155]
- Monsalve RI, Gutierrez R, Hoof I, Lombardero M, 2020. Purification and molecular characterization of phospholipase, antigen 5 and hyaluronidases from the venom of the Asian hornet (*Vespa velutina*). *PLoS One* 15, e0225672. [PubMed: 31923175]
- Plesner L, 1995. Ecto-ATPases: identities and functions. *Int. Rev. Cytol* 158, 141–214. [PubMed: 7721538]
- Reinert JF, 2000. New classification for the composite genus *Aedes* (Diptera: Culicidae: Aedini), elevation of subgenus *Ochlerotatus* to generic rank, reclassification of the other subgenera, and notes on certain subgenera and species. *J. Am. Mosq. Control Assoc* 16, 175–188. [PubMed: 11081644]
- Reno HE, Novak RJ, 2005. Characterization of apyrase-like activity in *Ochlerotatus triseriatus*, *Ochlerotatus hendersoni*, and *Aedes aegypti*. *Am. J. Trop. Med. Hyg* 73, 541–545. [PubMed: 16172478]
- Ribeiro JMC, Arca B, 2009. From sialomes to the Sialoverse: an insight into salivary potion of blood-feeding insects. *Adv. Insect Physiol* 37, 59.
- Ribeiro JM, Sarkis JJ, Rossignol PA, Spielman A, 1984. Salivary apyrase of *Aedes aegypti*: characterization and secretory fate. *Comp. Biochem. Physiol B* 79, 81–86.
- Ribeiro JM, Makoul GT, Levine J, Robinson DR, Spielman A, 1985. Antihemostatic, antiinflammatory, and immunosuppressive properties of the saliva of a tick, *Ixodes dammini*. *J. Exp. Med* 161, 332–344. [PubMed: 2982989]
- Ribeiro JM, Vaughan JA, Azad AF, 1990. Characterization of the salivary apyrase activity of three rodent flea species. *Comp. Biochem. Physiol B* 95, 215–219.
- Ribeiro JM, Nussenzweig RH, Tortorella G, 1994. Salivary vasodilators of *Aedes triseriatus* and *Anopheles gambiae* (Diptera: Culicidae). *J. Med. Entomol* 31, 747–753. [PubMed: 7966179]
- Ribeiro JM, Charlab R, Pham VM, Garfield M, Valenzuela JG, 2004. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. *Insect Biochem. Mol. Biol* 34, 543–563. [PubMed: 15147756]
- Ribeiro JM, Arca B, Lombardo F, Calvo E, Phan VM, Chandra PK, Wikel SK, 2007. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. *BMC Genom* 8, 6.
- Ribeiro JM, Martin-Martin I, Arca B, Calvo E, 2016. A deep insight into the sialome of male and female *Aedes aegypti* mosquitoes. *PLoS One* 11, e0151400. [PubMed: 26999592]

- Ribeiro JM, Slovak M, Francischetti IM, 2017. An insight into the sialome of *Hyalomma excavatum*. *Ticks Tick Borne Dis* 8, 201–207. [PubMed: 28049606]
- Ribeiro JMC, Martin-Martin I, Moreira FR, Bernard KA, Calvo E, 2018. A deep insight into the male and female sialotranscriptome of adult *Culex tarsalis* mosquitoes. *Insect Biochem. Mol. Biol* 95, 1–9. [PubMed: 29526772]
- Robinson MD, McCarthy DJ, Smyth GK, 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. [PubMed: 19910308]
- Rust RS, Thompson WH, Matthews CG, Beaty BJ, Chun RW, 1999. La Crosse and other forms of California encephalitis. *J. Child Neurol* 14, 1–14. [PubMed: 10025535]
- Sarkis JJ, Guimaraes JA, Ribeiro JM, 1986. Salivary apyrase of *Rhodnius prolixus*. Kinetics and purification. *Biochem. J* 233, 885–891. [PubMed: 3010945]
- Sepey M, Manni M, Zdobnov EM, 2019. BUSCO: Assessing genome assembly and annotation completeness. *Methods Mol. Biol* 1962, 227–245. [PubMed: 31020564]
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I, 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res* 19, 1117–1123. [PubMed: 19251739]
- Smartt CT, Kim AP, Grossman GL, James AA, 1995. The Apyrase gene of the vector mosquito, *Aedes aegypti*, is expressed specifically in the adult female salivary glands. *Exp. Parasitol* 81, 239–248. [PubMed: 7498420]
- Smith LB, Duge E, Valenzuela-Leon PC, Brooks S, Martin-Martin I, Ackerman H, Calvo E, 2022. Novel salivary anti-hemostatic activities of long-form D7 proteins from the malaria vector *Anopheles gambiae* facilitate hematophagy. *J. Biol. Chem.* 101971 [PubMed: 35460690]
- Stark KR, James AA, 1996. Salivary gland anticoagulants in culicine and anopheline mosquitoes (Diptera: Culicidae). *J. Med. Entomol* 33, 645–650. [PubMed: 8699461]
- Stark KR, James AA, 1998. Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, *Aedes aegypti*. *J. Biol. Chem* 273, 20802–20809. [PubMed: 9694825]
- Sun D, McNicol A, James AA, Peng Z, 2006. Expression of functional recombinant mosquito salivary apyrase: a potential therapeutic platelet aggregation inhibitor. *Platelets* 17, 178–184. [PubMed: 16702045]
- Sun P, Nie K, Zhu Y, Liu Y, Wu P, Liu Z, Du S, Fan H, Chen CH, Zhang R, Wang P, Cheng G, 2020. A mosquito salivary protein promotes flavivirus transmission by activation of autophagy. *Nat. Commun* 11, 260. [PubMed: 31937766]
- Surasombatpattana P, Ekchariyawat P, Hamel R, Patramool S, Thongrunkiat S, Denizot M, Delaunay P, Thomas F, Luplertlop N, Yssel H, Misse D, 2014. *Aedes aegypti* saliva contains a prominent 34-kDa protein that strongly enhances dengue virus replication in human keratinocytes. *J. Invest. Dermatol* 134, 281–284. [PubMed: 23752041]
- Tadokoro T, Modahl CM, Maenaka K, Aoki-Shioi N, 2020. Cysteine-rich secretory proteins (CRISPs) from venomous snakes: an overview of the functional diversity in A large and underappreciated superfamily. *Toxins* 12. [PubMed: 33374361]
- Team RC, 2020. R: A Language and Environment for Statistical Computing
- Thangamani S, Wikel SK, 2009. Differential expression of *Aedes aegypti* salivary transcriptome upon blood feeding. *Parasites Vectors* 2.
- Tirioni L, Lu S, Calvo E, Sabadin G, Di Maggio LS, Suzuki M, Nardone G, da Silva Vaz I Jr., Ribeiro JMC, 2020. Integrated analysis of sialotranscriptome and sialoproteome of the brown dog tick *Rhipicephalus sanguineus* (s.l.): insights into gene expression during blood feeding. *J. Proteomics* 229, 103899.
- Valenzuela JG, Charlab R, Galperin MY, Ribeiro JM, 1998. Purification, cloning, and expression of an apyrase from the bed bug *Cimex lectularius*. A new type of nucleotide-binding enzyme. *J. Biol. Chem* 273, 30583–30590. [PubMed: 9804829]
- Valenzuela JG, Francischetti IM, Ribeiro JM, 1999. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*. *Biochemistry* 38, 11209–11215. [PubMed: 10460178]

- Valenzuela JG, Charlab R, Gonzalez EC, de Miranda-Santos IK, Marinotti O, Francischetti IM, Ribeiro JM, 2002. The D7 family of salivary proteins in blood sucking diptera. *Insect Mol. Biol* 11, 149–155. [PubMed: 11966880]
- Valenzuela JG, Francischetti IM, Pham VM, Garfield MK, Ribeiro JM, 2003. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochem. Mol. Biol* 33, 717–732. [PubMed: 12826099]
- Valenzuela-Leon PC, Shrivastava G, Martin-Martin I, Cardenas JC, Londono-Renteria B, Calvo E, 2022. Multiple salivary proteins from *Aedes aegypti* mosquito bind to the Zika virus envelope protein. *Viruses* 14.
- van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C, 2014. Ten years of next-generation sequencing technology. *Trends Genet* 30, 418–426. [PubMed: 25108476]
- Waterhouse RM, S M, Simão FA, Zdobnov EM, 2019. Using BUSCO to Assess Insect Genomic Resources Humana Press, New York, NY.
- Xu X, Yang H, Ma D, Wu J, Wang Y, Song Y, Wang X, Lu Y, Yang J, Lai R, 2008. Toward an understanding of the molecular mechanism for successful blood feeding by coupling proteomics analysis with pharmacological testing of horsefly salivary glands. *Mol. Cell. Proteomics* 7, 582–590. [PubMed: 18087067]
- Zwiebel LJ, Takken W, 2004. Olfactory regulation of mosquito-host interactions. *Insect Biochem. Mol. Biol* 34, 645–652. [PubMed: 15242705]

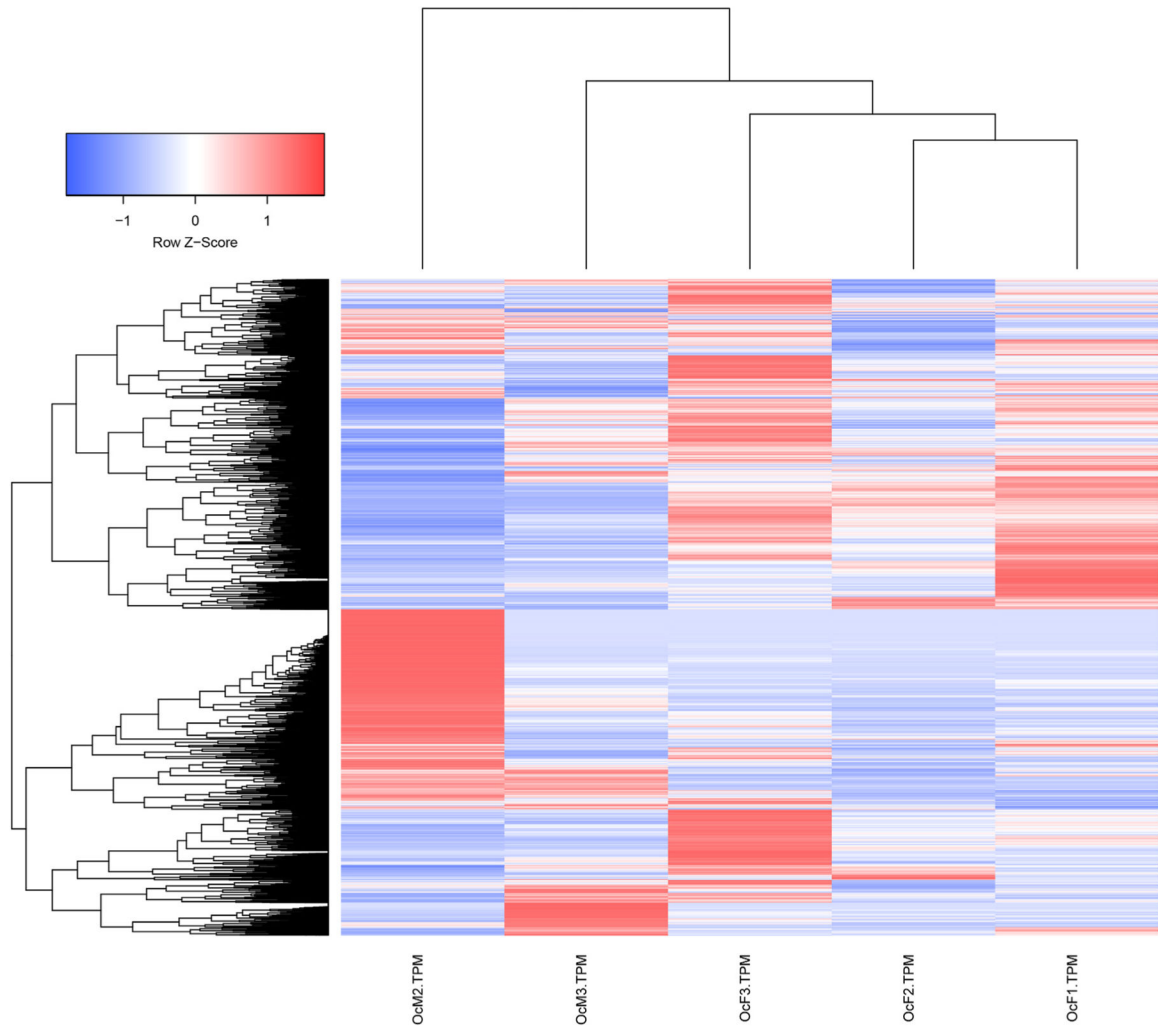


Fig. 1. Heat map plot of normalized TPM values of the 10,317 CDS identified in the sialome of *O. triseriatus* females and males. The number of biological samples is shown at the bottom.

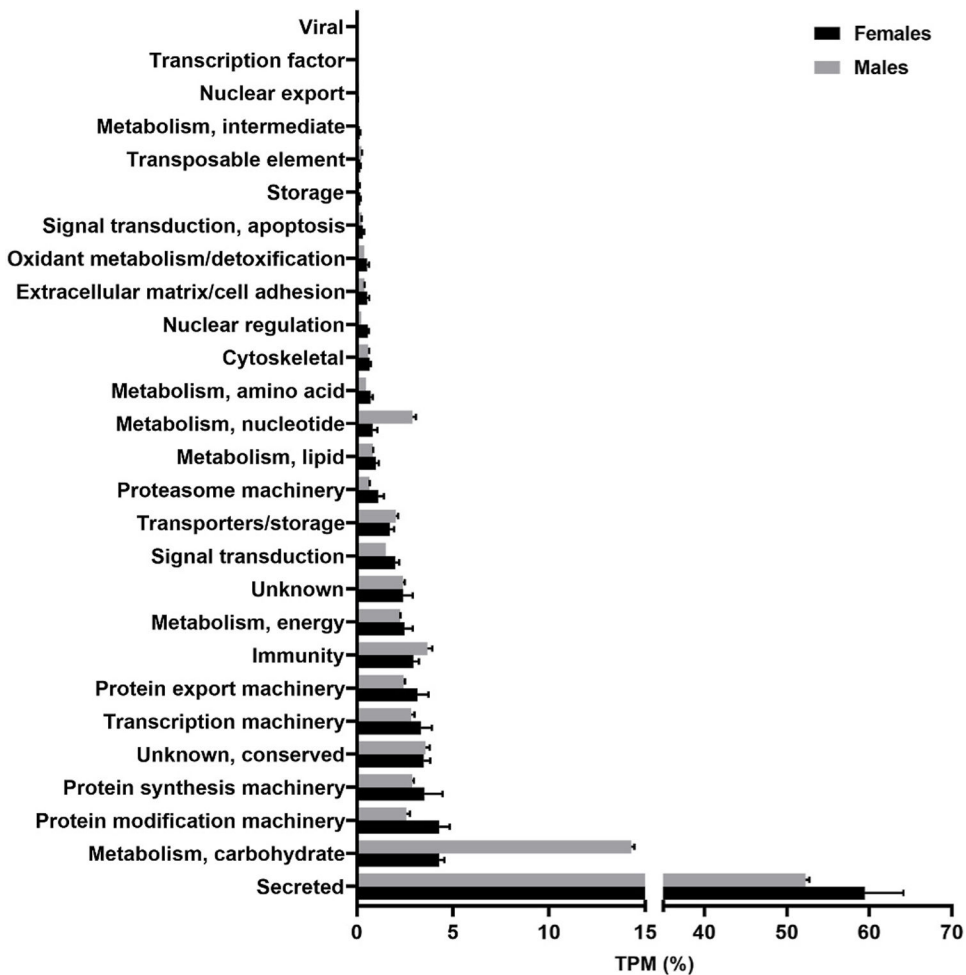


Fig. 2. Relative quantification of each functional class of the 10,317 CDS identified in the salivary gland homogenates of male and female *O. triseriatus* adults. Error bars represent the standard deviation of the mean from three (females) or two (males) biological replicates.

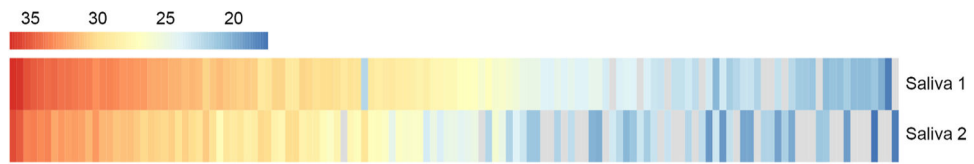


Fig. 3. Heat map plot of proteins identified in *O. triseriatus* female saliva by mass spectrometry. The area under the curve (AUC) of the features identified by each protein were determined and normalized by applying the Log₂ function. Proteins that were not identified in one of the two biological replicates are shown in gray color.

Table 1

Families of secreted proteins within the sialotranscriptome of *O. triseriatus*.

| Protein family | Females | | | Males | | | Female TPM/(Male TPM +1) |
|--------------------|-------------|------------|---------|-------------|------------|---------|--------------------------|
| | No. Contigs | TPM | TPM (%) | No. Contigs | TPM | TPM (%) | |
| Immunity | 1 | 8854.90 | 1.60 | - | - | - | 8854.90 |
| Apyrase | 5 | 6191.70 | 1.10 | - | - | - | 6191.70 |
| 62 kDa | 2 | 5135.00 | 0.94 | - | - | - | 5135.00 |
| Basic-tail | 1 | 4891.40 | 0.90 | - | - | - | 4891.40 |
| 34 kDa | 3 | 4606.30 | 0.84 | - | - | - | 4606.30 |
| Pro-rich | 2 | 3965.50 | 0.73 | - | - | - | 3965.50 |
| Aegyptin/30 kDa | 4 | 67,640.40 | 12.00 | 4 | 17.90 | 0.0037 | 3578.86 |
| 13 kDa | 2 | 3454.41 | 0.63 | - | - | - | 3454.41 |
| D7 | 13 | 242,786.60 | 44.00 | 5 | 95.70 | 0.0200 | 2510.72 |
| SGS | 5 | 1167.60 | 0.21 | 1 | 0.25 | 0.0001 | 934.08 |
| Leu-rich | 2 | 14.40 | 0.00 | - | - | - | 14.40 |
| Gly-rich | 5 | 1401.50 | 0.26 | 5 | 134.80 | 0.0280 | 10.32 |
| Glu-rich | 2 | 105.80 | 0.02 | 2 | 15.90 | 0.0033 | 6.26 |
| OBP | 13 | 308.10 | 0.06 | 10 | 50.40 | 0.0100 | 5.99 |
| Antigen 5 | 10 | 16,661.40 | 3.00 | 9 | 4163.90 | 0.8600 | 4.00 |
| Inhibitors | 42 | 11,594.40 | 2.10 | 38 | 5757.90 | 1.2000 | 2.01 |
| Met, nucleotides | 13 | 539.50 | 0.10 | 12 | 318.30 | 0.0650 | 1.69 |
| Lipocalin | 4 | 106.20 | 0.02 | 3 | 64.00 | 0.0130 | 1.63 |
| Proteases | 151 | 10,666.70 | 2.00 | 96 | 7716.50 | 1.6000 | 1.38 |
| Met, lipids | 17 | 3629.30 | 0.66 | 12 | 5522.10 | 1.1000 | 0.66 |
| 56 kDa | 1 | 4040.20 | 0.74 | 1 | 7748.00 | 1.6000 | 0.52 |
| Hormone binding | 12 | 131.70 | 0.02 | 10 | 305.50 | 0.0630 | 0.43 |
| Mucin | 15 | 14,114.80 | 2.60 | 15 | 34,953.50 | 7.2000 | 0.40 |
| Unknown | 608 | 55,379.00 | 10.00 | 574 | 158,051.10 | 33.0000 | 0.35 |
| Unknown, conserved | 177 | 66,739.50 | 12.00 | 154 | 206,561.50 | 42.0000 | 0.32 |
| 41 kDa | 4 | 7244.20 | 1.30 | 4 | 27,295.50 | 5.6000 | 0.27 |
| 8.5 kDa | 1 | 4930.20 | 0.90 | 1 | 27,430.90 | 5.6000 | 0.18 |

| Protein family | Females | | | Males | | | Female TPM/(Male TPM +1) |
|----------------|-------------|------|---------|-------------|------|---------|--------------------------|
| | No. Contigs | TPM | TPM (%) | No. Contigs | TPM | TPM (%) | |
| | | | | | | | |
| 15.3 kDa | 1 | 0.12 | 0.00 | 1 | 2.94 | 0.0006 | 0.03 |
| 16.8 kDa | 1 | 0.08 | 0.00 | 1 | 2.21 | 0.0005 | 0.02 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Functional classification of differentially expressed transcripts identified in the salivary glands of *O. triseriatus* when comparing females to males.

| Class | Up regulated | Down regulated |
|------------------------------------|--------------|----------------|
| Cytoskeletal | 2 | 2 |
| Extracellular matrix/cell adhesion | 4 | 4 |
| Immunity | 8 | 5 |
| Metabolism, amino acid | 9 | 1 |
| Metabolism, carbohydrate | 7 | 3 |
| Metabolism, energy | 2 | 1 |
| Metabolism, intermediate | 2 | – |
| Metabolism, lipid | 16 | 5 |
| Nuclear regulation | 1 | 1 |
| Oxidant metabolism/detoxification | 5 | 2 |
| Proteasome machinery | 3 | – |
| Protein export machinery | 4 | 1 |
| Protein modification machinery | 12 | 3 |
| Protein synthesis machinery | 4 | – |
| Secreted | 63 | 20 |
| Signal transduction | 9 | 5 |
| Transcription machinery | 6 | 5 |
| Transporters/storage | 9 | 3 |
| Transposable element | 3 | 3 |
| Unknown | 14 | 1 |
| Unknown, conserved | 14 | 6 |
| Metabolism, nucleotide | – | 2 |
| Signal transduction, apoptosis | – | 2 |
| Transcription factor | – | 1 |

Table 3

Secreted protein families found differentially expressed when comparing females and males *O. triseriatus*.

| Protein family | Up regulated | Down regulated |
|-----------------|--------------|----------------|
| 13 kDa | 2 | - |
| 34 kDa | 3 | - |
| 41 kDa | - | 1 |
| 62 kDa | 1 | - |
| Aegyptin/30 kDa | 3 | - |
| Antigen 5 | 1 | 1 |
| Apyrase | 4 | - |
| Basic tail | 1 | - |
| D7 | 11 | - |
| Hormone binding | - | 1 |
| Inhibitor | 2 | 2 |
| Leu-rich | 1 | - |
| Mucin | - | 1 |
| OBP | 2 | - |
| Pro-rich | 1 | - |
| Protease | 8 | 6 |
| SGS | 7 | - |
| Unknown | 16 | 8 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Functional classification of secreted proteins identified in the LC-MS/MS analysis of *O. triseriatus* female saliva.

| Family | No. Prot. | Avg. Area | Area (%) |
|-------------------------------|-----------|-----------|----------|
| 13kDa family | 1 | 8.46E+08 | 0.919 |
| 34 kDa family | 2 | 7.57E+09 | 8.220 |
| 41 kDa family | 3 | 9.48E+06 | 0.010 |
| 56 kDa family | 2 | 8.55E+07 | 0.093 |
| 62 kDa family | 2 | 2.10E+09 | 2.280 |
| Adenosine-deaminase | 4 | 8.46E+08 | 0.919 |
| Aegyptin/30 kDa | 5 | 9.79E+09 | 10.631 |
| Alpha-amylase | 2 | 1.20E+07 | 0.013 |
| Angiotensin-converting enzyme | 1 | 6.12E+08 | 0.665 |
| Antigen 5 | 4 | 6.02E+09 | 6.537 |
| Apyrase | 4 | 3.91E+09 | 4.246 |
| Basic-tail | 1 | 7.93E+07 | 0.086 |
| Cholinesterase | 1 | 1.64E+09 | 1.781 |
| D7/long | 3 | 2.84E+10 | 30.840 |
| D7/short | 6 | 1.82E+10 | 19.764 |
| Immunity | 7 | 1.02E+09 | 1.108 |
| Inhibitor | 7 | 2.92E+09 | 3.169 |
| Leu-rich | 1 | 4331500 | 0.005 |
| Maltase | 2 | 2.82E+08 | 0.306 |
| Odorant-binding protein | 1 | 15600000 | 0.017 |
| Proteases | 5 | 2.58E+07 | 0.028 |
| SGS | 7 | 6.32E+09 | 6.863 |
| Unknown | 29 | 1.37E+09 | 1.48 |