Transcriptome Sequencing and Developmental Regulation of Gene Expression in *Anopheles aquasalis*

André L. Costa-da-Silva[1,2][9], Osvaldo Marinotti[3][9], José M. C. Ribeiro[4][9], Maria C. P. Silva[1], Adriana R. Lopes[5], Michele S. Barros[6], Anderson Sá-Nunes[2,6], Bianca B. Kojin[3], Eneas Carvalho[7], Lincoln Suesdek[1,8], Mário Alberto C. Silva-Neto[2,9], Anthony A. James[3,10], Margareth L. Capurro[1,2]*

1 Laboratório de Mosquitos Geneticamente Modificados, Departamento de Parasitologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, São Paulo, Brazil, 2 Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular, INCT-EM, Rio de Janeiro, Rio de Janeiro, Brazil, 3 Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, California, United States of America, 4 Section of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, United States of America, 5 Laboratório de Bioquímica e Biofísica, Instituto Butantan, São Paulo, São Paulo, Brazil, 6 Laboratório de Imunologia Experimental, Departamento de Imunologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, São Paulo, Brazil, 7 Centro de Biotecnologia, Instituto Butantan, São Paulo, São Paulo, Brazil, 8 Laboratório de Parasitologia, Instituto Butantan, São Paulo, São Paulo, Brazil, 9 Laboratório de Sinalização Celular, Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brazil, 10 Department of Microbiology and Molecular Genetics, University of California Irvine, Irvine, California, United States of America

Abstract

Background: *Anopheles aquasalis* is a major malaria vector in coastal areas of South and Central America where it breeds preferentially in brackish water. This species is very susceptible to *Plasmodium vivax* and it has been already incriminated as responsible vector in malaria outbreaks. There has been no high-throughput investigation into the sequencing of *An. aquasalis* genes, transcripts and proteins despite its epidemiological relevance. Here we describe the sequencing, assembly and annotation of the *An. aquasalis* transcriptome.

Methodology/Principal Findings: A total of 419 thousand cDNA sequence reads, encompassing 164 million nucleotides, were assembled in 7544 contigs of ≥2 sequences, and 1999 singletons. The majority of the *An. aquasalis* transcripts encode proteins with their closest counterparts in another neotropical malaria vector, *An. darlingi*. Several analyses in different protein databases were used to annotate and predict the putative functions of the deduced *An. aquasalis* proteins. Larval and adult-specific transcripts were represented by 121 and 424 contig sequences, respectively. Fifty-one transcripts were only detected in blood-fed females. The data also reveal a list of transcripts up- or down-regulated in adult females after a blood meal. Transcripts associated with immunity, signaling networks and blood feeding and digestion are discussed.

Conclusions/Significance: This study represents the first large-scale effort to sequence the transcriptome of *An. aquasalis*. It provides valuable information that will facilitate studies on the biology of this species and may lead to novel strategies to reduce malaria transmission on the South American continent. The *An. aquasalis* transcriptome is accessible at http://exon.niaid.nih.gov/transcriptome/An_aquasalis/Anaquexcel.xlsx.

Introduction

Anopheles aquasalis is a neotropical malaria vector, found along the northern coast of South America. Its geographical distribution extends from Brazil to Panama on the Atlantic shore and from Panama to Ecuador along the Pacific coast [1–6]. Their larvae develop preferentially in brackish water such as in mangrove swamps and coastal ground pools, but they also are capable of living in fresh water and often occur several kilometers from the coast [7].

The species has been reported as malaria vector in Venezuela, Brazil, Trinidad and The Caribbean [8–12]. *Plasmodium falciparum* and *P. vivax*, the two main human malaria parasites, are transmitted by *An. aquasalis* [12,13]. Its epidemiological importance was confirmed during a *P. vivax* malarial outbreak in 1991–1992 occurred in Trinidad which was linked to *An. aquasalis* [14].

Despite its importance as a malaria vector in Central and South America, regions responsible for 22% of the global area at risk of *P. vivax* transmission [15], little is known about its genome,

Author Summary

The mosquito *Anopheles aquasalis* is responsible for transmitting malaria parasites to humans in South America coastal areas. *An. aquasalis* females transmit *Plasmodium vivax* and *Plasmodium falciparum*, the two major malaria etiological agents in these endemic sites. Although the vectorial importance of this mosquito has been demonstrated, molecular aspects of its biology have been poorly explored. In this study, we present the transcriptome of *An. aquasalis* using 454 sequencing followed by automated bioinformatic analyses. Our study identified and annotated more than 9,000 putative proteins based on homology, gene ontology, and/or biochemical pathways, including putative secretory proteins. The comparison of RNAs present in samples extracted from larvae, sugar fed adult females, or blood fed adult females, reveal gene expression regulation during mosquito development. The present dataset provides a useful resource and adds greatly to our understanding of a human malaria vector from developing countries.

transcriptome and proteome. Efforts to colonize this species [12,16] provided the basis for studies assessing development, gene expression and immune responses to *Plasmodium* infection [12,17–24]. The role of reactive oxygen species and JAK-STAT pathway in the control of *P. vivax* infection were characterized in *An. aquasalis*. However, it was also demonstrated that this anopheline promotes an apparent weak response against *P. vivax* infection [18] which is supposed to be related to its susceptibility to this parasite species. Corroborating with this hypothesis, a recent work showed that colonized *An. aquasalis* presents higher infection rates and oocyst numbers when compared to other Neotropical anophelines [25].

Here we describe the transcriptomes and deduced proteomes of *An. aquasalis* larvae and adults fed on sugar and on blood. This data set provides indispensable information for the systematic and comprehensive analysis of molecules that may play an active role in mosquito biology and malaria transmission.

Methods

Mosquitoes

The *An. aquasalis* colony (gift from Dr. Paulo Filemon Paolucci Pimenta/Fiocruz) was maintained in the insectary at the Departamento de Parasitologia, ICB-USP (São Paulo, Brazil) at $27\pm1°C$, 75–80% relative humidity and 12 h light:12 h dark photoperiod. Larvae were kept in 0.2% marine salt (w/v), and were fed with powdered fish food (Tetramin, Blacksburg, VA, USA). Adult males and females were kept together in a cage with access to a 10% sucrose solution (w/v) *ad libitum*. Female mosquitoes aged 5–7 days after emerging from the pupal stage were allowed to feed on anaesthetized mice for 30 minutes. Eggs were collected 2–3 days post blood meal and hatched in 0.2% marine salt (w/v).

RNA extraction and quantification

Larvae RNA was extracted from a pool of third and fourth instar larvae (20 each). RNA also was extracted from one pool of twenty 5–7 day old adults females fed with sucrose and two additional samples, each composed of twenty 5–7 day old adult females at 24 h after blood meal. Frozen animals were used for mRNA extraction using magnetic beads covalently bound to oligo(dT) tags (Dynabeads mRNA DIRECT, Invitrogen, Grand Island, NY, USA), accordingly to the manufacturer's instructions. Aliquots of the purified mRNA samples were quantified using Quant-iTRiboGreen RNA Reagent (Invitrogen) and their integrity was checked in a microfluidics-based platform (Agilent 2100 Bioanalyzer, Santa Clara, CA, USA).

Sequencing

Approximately 400 ng of polyA+ RNA from each sample were used as template for sequencing. mRNA was fragmented with zinc chloride, resulting in molecules with a size distribution range from 300 to 800 bases (assessed by the Bioanalyzer), and used as a template for cDNA synthesis. Adaptors were linked to the fragment ends. Beads coated with oligonucleotides complementary to the adaptor sequences were incubated with the cDNA fragments, and a water-in-oil emulsion was produced, followed by emulsion PCR. Washed beads were deposited in picotiter plate wells, and other sequencing reagents were loaded on the 454 GS-Junior sequencer (Roche, Branford, CT, USA). Two hundred sequencing cycles were performed. Base-calling was performed by the 454 GS-Junior data processing software GS Run Processor, version 2.7.

Assembly and annotation

The blastn tool (performed locally from executables obtained at the NCBI FTP site ftp://ftp.ncbi.nih.gov/blast/executables/) [26] and CAP3 assembler [27] were used for sequence clustering, by a decreasing word size inclusion strategy as described in detail previously [28]. Coding sequences (CDS) were extracted as described before [28] based on matches to public databases or longer open reading frames with a signal peptide indicative of secretion. Contigs are named Anoaqua-XXX or megaclu_asb-SigP-XXX reflecting the two modes of data extraction, where XXX represents the number of the full length assembled contig. Reference to specific contigs in this will use an abbreviated notation, contigXXX, instead of the full CDS name. The data was organized in a hyperlinked spreadsheet (Anaquexcel) as described in [29]. The blastx [30] tool was used to compare the translated nucleotide sequences to the NR protein database of the NCBI and to the Gene Ontology (GO) database [31]. The tool, reverse position-specific BLAST (rpsblast) [30], was used to search for conserved protein domains in the Pfam [32], SMART [33], KOG [34] and conserved domains databases (CDD) [35]. Predicted protein segments starting with a methionine were submitted to the SignalP server [36] to identify translation products that could be secreted. Glycosylation sites on the proteins were predicted with the program NetOGlyc [37]. Functional annotation of the transcripts was based on all of the comparisons above. Transcripts and their encoded proteins were classified based on function and/or protein families. To compare gene expression between libraries, paired comparisons of the number of reads hitting each contig were calculated by X^2 tests to detect significant differences between samples when the minimum expected value was larger than 5 and P<0.05. A 2-fold change (up or down) was considered of interest when statistically significant. Normalized fold-ratios of the library reads were computed by adjusting the numerator by a factor based on the ratio of the total number of reads in each library, and adding one to the denominator to avoid division by zero. The complete Anaquexcel dataset (including links) may be downloaded from http://exon.niaid.nih.gov/transcriptome/An_aquasalis/Anaquexcel.xlsx and searched as an Excel spreadsheet. The raw data were deposited to the Sequence Read Archives (SRA) of the National Center for Biotechnology Information (NCBI) under bioproject number PRJNA210899, biosamples SRS455914 (adults) and SRS455922 (larvae) and runs

SRR927455 (female sucrose), SRR927456 (female blood fed) and SRR927458 (L3+L4 larvae). CDS representing >90% of known proteins or larger than 250 amino acids were deposited to the Transcriptome Shotgun Annotation (TSA) portal of the NCBI and received the accession numbers from GAMD01000001 to GAMD01003464.

Validation of RNA-seq data by qRT-PCR

To confirm the expression profile generated by the transcriptome sequencing, we validated the expression levels of twelve contigs identified in RNA-seq analysis using a qRT-PCR method. Contigs classified as enhanced or specific for larva (Anoaqua-397, Anoaqua-1598, Anoaqua-4095, Anoaqua-17360 and Anoaqua-1222), specific for adult (Anoaqua-436 and Anoaqua-457) and blood meal regulated (megaclu_asbSigP-9948, Anoaqua-3237, megaclu_asbSigP-2537, Anoaqua-5059 and Anoaqua-24500) were analyzed.

To perform these quantifications, TRIZOL reagent (Invitrogen) was used to extract total RNA from 3 independent biological pools of third and fourth instar larvae (10 each), ten 5–7 day old adults females fed with sucrose and ten 5–7 day old adult females at 24 h after blood meal. For each extraction, total RNA was quantified and 4.0 µg were treated with DNAse I (Invitrogen) and was reverse transcribed using superscript II (Invitrogen) and oligoDT (Invitrogen) in a 40 µL final reaction volume.

qRT-PCR assay was performed in Mastercycler Realplex 2 thermocycler (Eppendorf) with Maxima SYBR green Master Mix (Thermo Scientific). Reactions were performed in a 20 µL final volume containing 2 µL of cDNA template and 0.5 µM of each primer (Table S3). Primers pair amplification efficiency was estimated using original cDNA in seven-fold serial dilutions to generate a standard curve. All primers pair showed efficiency greater than 90% (Table S3) and only one specific peak was observed in the melting curve for each analyzed transcript. Each sample was measured in triplicate and three biological replicates were quantified. Expression levels of *An. aquasalis* Rp49 constitutive gene [18,38] (Tables S3 and S4) was used to normalize variation in total cDNA concentration as an endogenous control. Fold-changes in gene expression were estimated by delta-delta C_T method [39] and sample with lower expression levels for each gene was defined as calibrators.

Statistical analysis

Statistical significance was evaluated using Graph Pad Prism5 software. Data was checked in relation to normality using D'Agostino and Pearson omnibus normality test. One way ANOVA followed by Tukey's Multiple Comparison posttest were applied when the data adequate to parametric model. Non-parametric data was analyzed by Kruskal-Wallis test followed by Dunn's Multiple Comparison posttest. Confidence intervals of 95% were defined.

Results/Discussion

General description of the *An. aquasalis* transcriptome

Sequencing returned $1.1–1.7 \times 10^5$ reads among the samples classes, with averages sizes ranging from 350–420 nucleotides in length and $48–60 \times 10^6$ total bases sequenced (Table 1). Approximately 7% of these were ribosomal RNAs and therefore were excluded in the subsequent analyses.

Assembled and annotated sequences are available in Anaquexcel database at http://exon.niaid.nih.gov/transcriptome/An_aquasalis/Anaquexcel.xlsx (a condensed table with basic information was also provided at http://exon.niaid.nih.gov/transcriptome/An_aquasalis/Anoaqua-Summarized.xlsx). A summary of the assembly compared to the raw reads is shown in Figure S1. The Anaquexcel database contains 7544 contigs assembled from ≥2 sequences and 1999 singletons. The number of sequences that compose each contig varies widely (from 2 to 5,207, average of 35 sequences per contig); 43% of the assembled contigs contained 10 or more sequences (Figure 1).

Blastp analyses of the deduced *An. aquasalis* protein sequences indicated that 70% of them have their closest counterpart in another insect (39% *An. darlingi*; 22% *An. gambiae*; 3% *Ae. aegypti*; 2% *Cu. quinquefasciatus*; and 4% other insects) (Figure 2). The percentage coverage was at least 50% for 4,671 proteins and 2,057 had coverage of at least 90%. Furthermore, 6,495 of the blastx searches resulted in best hits with at least 50% identity. These results support the proposed functional annotation of the majority of the deduced *An. aquasalis* proteins (Figure 3). The translated products of 4,821 contigs or singletons were not assigned a putative function (classified as unknown) either because efforts failed to identify similar products in all the searched databases or because similar proteins identified in other organisms have no assigned biological role or activity. Among those with unknown functions, 1,434 encode products similar to proteins found in other organisms (conserved hypothetical proteins). The remaining 2,387 transcripts classified as having unknown functions either encode novel proteins or alternatively correspond to fragments from mRNAs untranslated regions (UTRs) or non-protein encoding RNAs. The predicted presence of amino terminal secretory signal peptide-like sequences supports the conclusion that 15% of the translation products are putatively secreted proteins. These data complement and extend the analyses of EST databases derived from *An. gambiae* mosquitoes in similar physiological conditions [40–44].

To better detail the data generated by RNA-seq, we elaborated 7 sections below to discuss results related to larva and adult enhanced and specific transcripts, the adult female *An. aquasalis* sialome, blood meal regulated transcripts, immunity-related transcripts, signaling networks in *An. aquasalis* and conservation of gene regulation between *An. aquasalis* and *An. gambiae*. In order to support our data, we validated the expression levels of 12 contigs by qRT-PCR experiments, being 5 contigs chosen from

Table 1. *Anopheles aquasalis* sequencing results.

Sample	No. reads	average size (nt)	total sequenced bases	% rRNA
sucrose-fed adult female	131,678	418	55,152,581	7.4
blood-fed adult female[1]	172,244	350	60,423,895	7.0
larvae	115,771	420	48,674,928	7.2

[1]Includes all blood-fed samples.
doi:10.1371/journal.pntd.0003005.t001
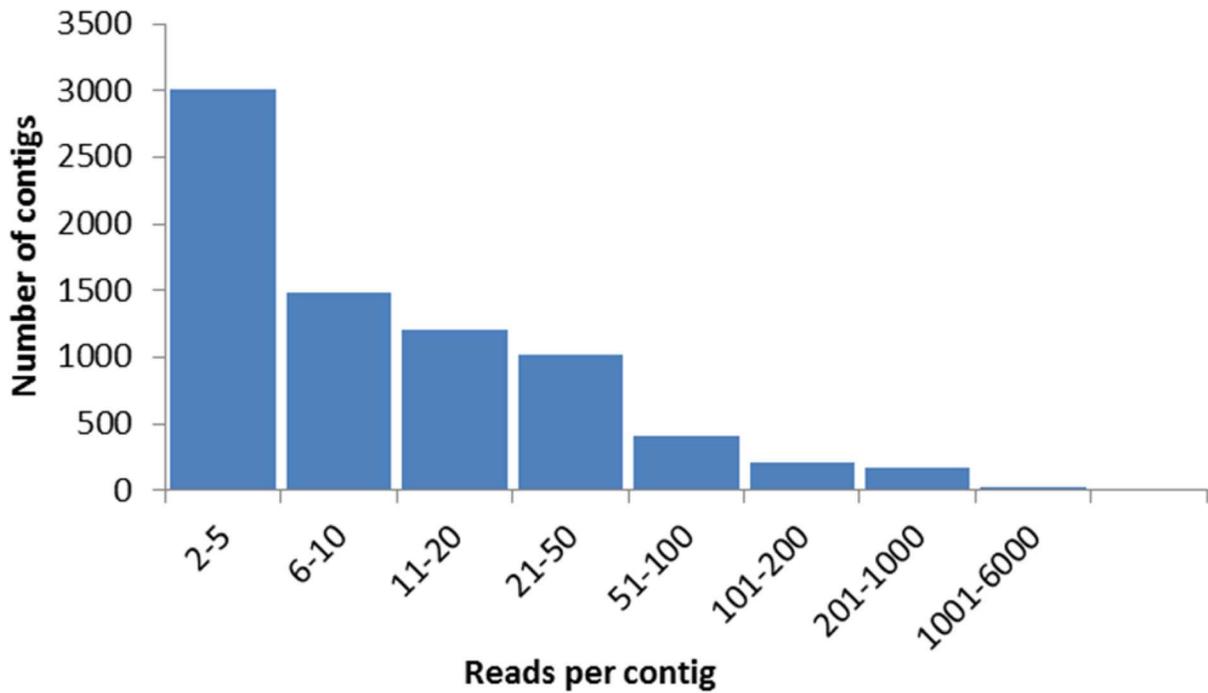
Figure 1. Number of sequences composing the assembled *An. aquasalis* contigs. A total of 7544 contigs were assembled from ≥2 sequences. The number of sequences that compose each contig varies from 2 to 5,207, with an average of 35 sequences per contig. Forty-three percent of the assembled contigs contained 10 or more sequences.
doi:10.1371/journal.pntd.0003005.g001
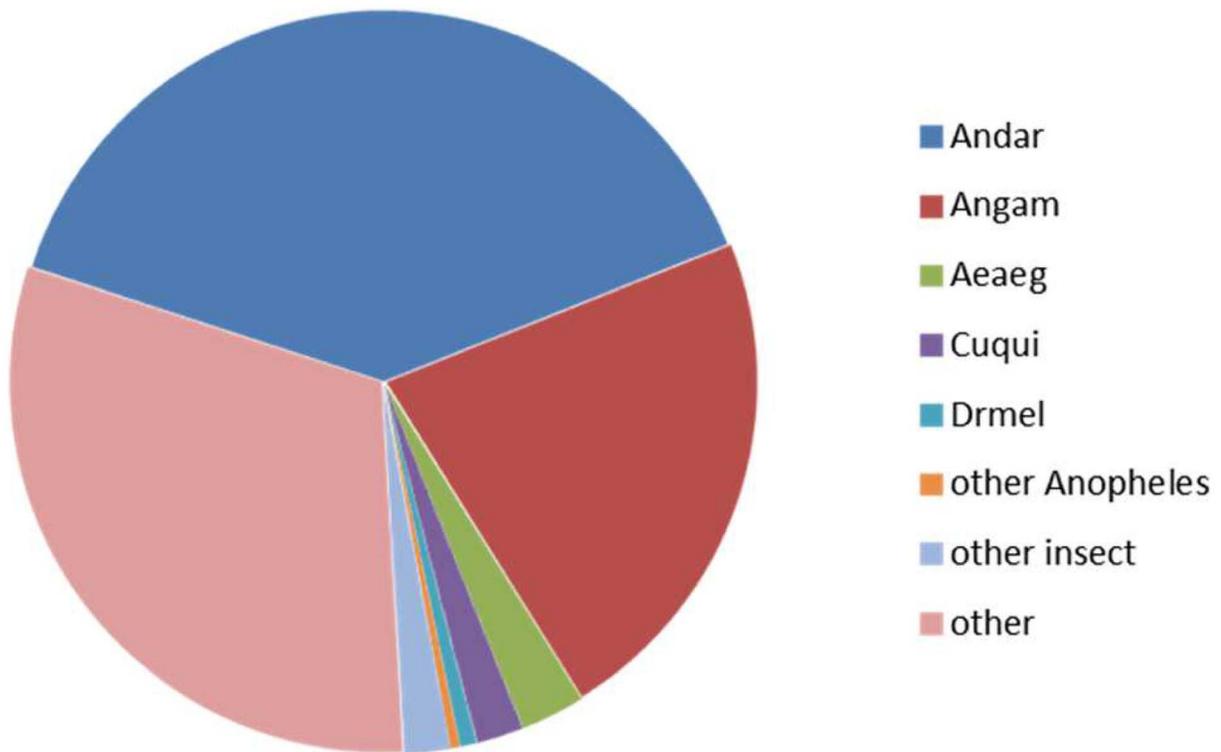
Figure 2. Distribution of the best matches of all *An. aquasalis* predicted proteins by organisms: Andar- *An. darlingi*; **Angam-** *An. gambiae*; **Aeaeg,** *Ae. aegypti*; **Cuqui,** *Cu. quinquefasciatus*; **Drmel,** *D. melanogaster*; **Other** *Anopheles* **species; Other insects- not of the** *Anopheles* **genus; Other- non-insect organisms.**
doi:10.1371/journal.pntd.0003005.g002

Figure 3. Classification of the annotated *An. aquasalis* proteins in functional categories. The numbers in parenthesis indicate the number of proteins in each category. apoptosis (64), cs- Cytoskeletal (278), detox- Detoxification (123), extmat- Extracellular matrix an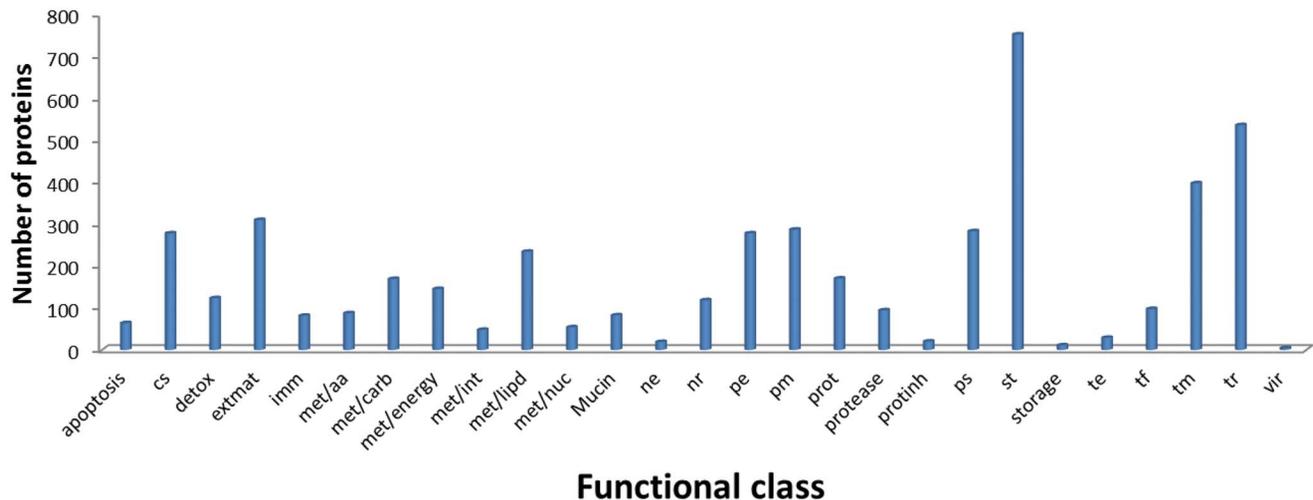d adhesion (310), imm-Immunity (81), met/aa- Amino acids metabolism (87), met/carb- Carbohydrate metabolism (169), met/energy- Energy metabolism (145), met/int-Intermediary metabolism (48), met/lipd- Lipid metabolism (234), met/nuc- Nucleotide metabolism (54), mucin (82), ne- Nuclear export (19), nr-Nuclear regulation (118), pe- Protein export machinery (278), pm- Protein modification machinery (287), prot- Proteasome machinery (170), protease (94), protinh- Protease inhibitor (20), ps- Protein synthesis (283), st- Signal transduction (752), storage (11), te- Transposable element (29), tf-Transcription factor (97), tm- Transcription machinery (397), tr- Transporters and storage (536), vir- Viral product (4).
doi:10.1371/journal.pntd.0003005.g003

blood meal regulated transcripts, 5 chosen from larva enhanced and specific transcripts and 2 chosen from adult enhanced and specific transcripts (Supplementary data, Figure S2, S3 and S4).

Larva enhanced and specific transcripts

One hundred twenty one transcripts were represented by at least 10 reads in the larval sample and were not detected in adults. Transcripts encoding hexamerins were among those accumulated most in larvae (Table S1 and S2). This finding is consistent with previous descriptions of hexamerins, also referred to as larval serum proteins (LSP) or insect storage proteins, as abundant proteins at the late larval stages of holometabolous insects [45,46]. Hexamerins are synthesized in the larval fat body, secreted into the hemolymph and taken up by fat body shortly before pupation. These proteins degrade during metamorphosis providing a source of amino acids for energy production and adult protein synthesis. Both LSP-1 like (contigs 1020, 1221, 1222, 1226, 3211 and 23555) and LSP-2 like (contigs 637, 2009 and 7437) hexamerins were found. Hexamerin-encoding transcripts corresponding to contig 1221 were found highly accumulated in both larvae and adults. Hexamerins expressed in adult mosquitoes have been reported [47] but their function during this developmental stage is unknown. Abundant larval mRNAs encoding ribosomal proteins, translation initiation factors and elongation factors represent the active protein synthesis machinery, consistent with the rapid growth rate during the third and fourth instars of mosquito larvae. Several transcripts encoding cuticle proteins have >50 reads in larvae and were not detected in adults (contigs 397, 983, 1598, 4095 and 17360). Developmental stage specific-cuticular protein transcripts were reported in *An. gambiae* [48], and those and our data support the hypothesis that the stage-specific expressed proteins are components of the different cuticular structures that characterize each metamorphic stage.

Adult enhanced and specific transcripts

A total of 424 transcripts represented by at least 10 reads in the adult dataset were not found in larvae. Transcripts enhanced in non-blood fed adults include components of visual sensory organs (Table S1 and S2). Opsins and arrestins (contigs 618, 1471 and 1843) are abundant components of the adult insect compound eyes and expressed highly in adult insects. Arrestins are important components for desensitization of G protein-coupled receptor cascades that mediate neurotransmission as well as olfactory and visual sensory reception [49,50]. Other components of the olfactory system with enhanced or specific expression in adults include16 identified contigs encoding odorant binding proteins, 10 of which were not represented in the sequenced larval RNA sample. These latter are indicative of new functions in the olfactory system that are specific to adult mosquitoes, such as host finding and breeding site selection.

Structural and enzymatic components of the digestive system also were among the adult specific transcripts. An adult-specific peritrophin (contig 1702) component of the peritrophic matrix (PM) was identified. A number of functions have been attributed to the PM, including protection against pathogens and abrasion, and compartmentalization of digestion [51,52]. The PM may delay digestion in adult mosquitoes [53] and modulate malaria parasites development [54]. Digestive trypsins with an expression pattern similar to those of *An. gambiae* trypsins 3, 4 and 7 were detected. These trypsins (contigs 457,1977 and 8935) are expressed exclusively in adults and are down-regulated following a blood meal [55]. These trypsin-like enzymes are probably necessary at the initial steps of digestion, but are dispensable later. Alternatively, their functions may be unrelated to digestion and they could associate with other processes regulated by limited proteolysis of precursors. Precursor proteins often require processing at specific sites in order to release their bioactive products [56,57]. Contig 383 corresponds to a transcript encoding an adult specific cuticle protein further supporting the hypothesis that the stage-specific

cuticular proteins make up the different cuticular structures of larvae and adult insects [48].

The adult female *An. aquasalis* sialome

Although the *An. aquasalis* transcriptome presented here was performed with whole body-extracted RNA and the salivary glands are only a small percentage of the total tissue, possibly less than 0.1%, (generally containing 1–3 µg protein), several putative salivary proteins of the adult female *An. aquasalis* were identified, based on their similarity to a database of salivary proteins from blood feeding Nematocera [58]. We considered only transcripts that are significantly up-regulated in the adult libraries as compared to the larval library as indicated in the methods section. Among putative salivary enzymes, contigs 5562, 17508 and 8234 encode members of the 5′nucleotidase/apyrase families, which are inhibitors of platelet aggregation. These contigs were assembled from 22–62 reads from adult libraries and 0–1 from larvae; similarly, contig 9665 codes for an alkaline phosphatase 66% identical to the *Aedes aegypti* salivary enzyme and was assembled from 60 reads from adults but zero from larvae. Although many serine peptidases were found in the *An. aquasalis* transcriptome, those encoded by contigs 2013 and 20556 are most similar to previously-described mosquito salivary enzymes that may play a role in blood feeding, such as destroying fibrin clots. Three peroxiredoxins similar to previously-described salivary proteins are also enriched in the adult libraries (contigs 4115, 19752 and 2971). Maltases have been described in mosquito salivary glands and are found in both male and female glands and assist sugar feeding. Contig 10431 with 45 reads has no larval reads and is a candidate for encoding a salivary enzyme.

Mosquito sialomes also contain antimicrobial and immune components found in both male and females including pathogen-recognition proteins and classical antimicrobial peptides. The antimicrobial peptide gambicin (contig 120, 70 reads from adults, 2 from larvae, 8.4 relative fold enrichment in adult library), three chitinase-like proteins (contigs 568, 3554 and 3923), a protein with an ML domain involved in pathogen lipid recognition (contig 4200) and a GGY family peptide (contig 10227) are possible salivary enriched gene products. The antigen-5 family of proteins is found ubiquitously in animal and plants, and specific family members are expressed in virtually all sialomes studied so far. Contig 436 (76 reads from adults and zero from larvae) is most similar to previously-annotated salivary members of this family.

A more specific set of proteins found only in mosquito or Nematoceran sialomes also were discovered. This group includes four members of the Aegyptin family of inhibitors of collagen-induced platelet aggregation (contigs 2621, 5929, 5148 and 1583) assembled from 843 adult reads but only three from the larval library. Members of the D7 family, involved in binding of host biogenic amines and inflammatory prostanoids also were assembled solely from adult-derived reads (contigs 2060, 950, 4474 and 19124). The *Anopheles*-specific antithrombin, anophelin, was matched in contig 801, assembled from 139 and 5 adult and larval reads, respectively. The uniquely anopheline salivary protein family, SG1/Trio, with unknown function, is represented by contigs 148, 3403, 21751, 2539 and 1630, assembled from 410 and 21 adult and larval reads, respectively. Further studies demonstrating the salivary specificity of these transcripts are needed.

Blood meal regulated transcripts

Fifty one transcripts not detected in larvae and sugar-fed mosquito samples were represented by ≥10 reads in the RNA samples of blood-fed females (Table S1 and S2).

Blood Digestion - The up-regulation of enzymes and proteins involved in digestion following a blood meal is well-documented in mosquitoes [55,59–61], including *An. aquasalis* [22,23,62]. Accordingly, the transcription products of several serine peptidases (contigs 299 and 305, AnaqTryp-1 and -2, respectively; contigs 2013 and 10625, Anachy1 and Anachy2 respectively), aminopeptidases and carboxypeptidases (contigs 2537 and 31670) were identified in our study as increasing in abundance following blood feeding.

Mucins, peritrophins (contigs 5059 and 9948) and G12 microvillar proteins (contig 6859) are up-regulated following blood meal in *An. aquasalis* and other insects. The accumulation of these proteins in the mosquito midgut lumen at a time when malaria parasites are traversing the midgut epithelium has been linked to a possible role in modulation of malaria parasites development [54,63–65].

Oogenesis - Genes encoding vitellogenins, other yolk components and eggshell-related proteins (vitellogenins, contigs 407 and 408; vitellogenic cathepsin, contig 1783) were identified as having large increases in transcript abundance after a blood meal. Our data confirm previously-described increase in vitellogenic protein expression in blood-fed mosquitoes [66]. Contig 24500 encodes lipophorin, a lipid transporter crucial for oogenesis, up-regulated following blood meal in *An. aquasalis* as well as other mosquitoes [43]. In addition to their indispensable function in mosquito reproduction, lipophorin and vitellogenin reduce parasite-killing by the antiparasitic factor TEP1 [67]. In the absence of either one, TEP1 binding to the ookinete surface becomes more efficient.

Eggshell proteins are expressed abundantly in the ovaries of vitellogenic insects. These proteins and related gene products have been studied extensively in model organisms such as *Drosophila melanogaster* and *Bombyx mori* [68], and more recently in *An. gambiae* [69]. We identified up-regulated transcripts encoding structural (vitelline membrane proteins, contigs 296, 2546 and 2709) as well enzymatic components of the mosquito eggshell (chorion peroxidase, contig 3237). Several other transcripts associated with egg formation and embryo development also were included in the group of blood meal induced transcripts (for example, Maternal protein exuperantia, contig 1219; Tyrosine-protein kinase receptor torso, contig 19837).

Transgenic mosquito strains are being developed to contribute to the control of malaria transmission [70]. The genetics and resulting phenotypes of a female-specific RIDL strategy, previously developed for dengue vector mosquitoes [71] has been adapted to a vector of human malaria, *An. stephensi* [72], indicating that the approach is applicable to other anopheline species. Moreover, production of malaria parasites-resistant transgenic mosquitoes has also been achieved in several laboratories [73]. The translation of these new, genetics-based technologies to new world anopheline species is likely feasible and the successful colonization of *An. aquasalis* places these mosquitoes in the list of target species for transgenesis. Therefore, regulatory sequences that drive transgenes expression in the appropriate developmental stage and organ, and produce an optimum amount of product are required [74,75] and the blood meal-activated genes identified in our work provide candidates for isolation and characterization of *An. aquasalis* functional promoters.

One hundred and ninety transcripts decrease significantly in abundance after the blood meal, most of which encode proteins with unknown functions. These likely represent genes expressed in tissues that are not involved in digestive and oogenesis functions. Alternatively, these genes could encode products such as early trypsins (contig 989) necessary at the initial steps of the digestion, but dispensable later [76,77].

Immunity-related transcripts

The mosquito immune system plays a critical role in limiting the spread of malaria and other vector-borne diseases. We identified a series of components of the innate immune system of *An. aquasalis*, including the antimicrobial peptides defensin (contigs 30436, 21901 and 1968), attacin (contig 20438), cecropin (contig 120), gambicin (contig 3403) lysozyme-c (contigs 4061, 1957, 14486, 16360 and 13712) and lysozyme-i (contigs 28407 and 27701). Additionally, members of the Toll pathway (Toll, contig 15063, cactus, contig 1543, dorsal, contig 17521, Kenny, contig 11522), members of the IMD pathway (IMD) (DIAP2, contig 6046, IKKbeta, contig 10934) and thioester proteins (TEPs) (contigs 21929, 29052 and 32589), the three major immune response systems in dipterans insects were identified.

Recent research supported differences between the responses of *An. aquasalis* to *P. vivax* infection when compared to immune response of *An. gambiae* to *P. falciparum* [18,19,78]. The immunity-related transcripts identified in this study will allow a more detailed study of the immune response of this neotropical vector to both *P. vivax* and *P. falciparum* infection.

Signaling networks in *An. aquasalis*

The dynamics of tyrosine phosphorylation-dephosphorylation constitutes a master biochemical regulator of cell biology [79]. It is mediated by a set of three major components: protein tyrosine kinases (PTKs), protein tyrosine phosphatases (PTPs), and Src Homology 2 (SH2) domains. It has been demonstrated previously the role of such a mechanism during tick embryogenesis [80], mosquito early adult development [81] and parasite infection [82]. The analysis of the *An. aquasalis* tyrosine phosphorylation-dephosphorylation regulatory enzymes revealed two contigs encoding PTPs (contigs 10198 and 21444). The first, contig 10198, encodes the classical non-receptor PTP (PTPn9) and the number of its reads increases in the larval/adult transition but decreases after blood feeding. This is a soluble tyrosine phosphatase that down-regulates prolactin- and EGF-mediated STAT5 activation [83]. STAT5 regulates expression of genes that promote cell survival and proliferation in breast cancer cells. The levels of phosphorylated EGF also are increased upon the suppression of PTPn9 mediated by MicroRNA miR-24 [84].

The *Anopheles* family of STAT transcription factors (Ag-STAT) was reported to be activated by bacterial challenge which then results in their nuclear translocation. This pathway is activated by inhibitors of PTPs [85]. We speculate that when bacterial loads increase following a blood meal, the suppression of PTPn9 expression allows the establishment of a STAT-mediated immunity and the induction of cell growth by incoming nutrients derived from blood digestion.

The second transcript, contig 21444, encodes for a dual-specificity phosphatase (DUSP), a PTP that also increases during the larval to adult transition and is strongly suppressed after blood feeding. These enzymes dephosphorylate both phosphotyrosine and phosphothreonine residues in target proteins and act as deactivators of mitogen-activated PK (MAPK) cascades. The complete set of genes in *An. gambiae* encoding for MAPKs and their activation profiles were described [86]. The level of phosphorylation of MAPKs in *Anopheles* was demonstrated to be responsive through treatment with insulin, TGF-B1, and LPS [86]. Curiously, p38 phosphorylation also is affected by hydrogen peroxide treatment, a common inhibitor of PTPs. So it is likely that contig 21444 is the enzyme that ultimately down-regulates the level of MAPK activation after cell treatment with the above-mentioned agents. The induction of MAPK activation following both

metabolic and immune challenges after blood feeding coincides with the down-regulation of contig 21444 transcripts.

An overall analysis of signaling molecules involved in direct phosphorylation-dephosphorylation circuits whose transcripts are up-regulated during different stages of mosquito development revealed one contig (1237) encoding a CBL-interacting serine/threonine-protein kinase 10 present in larvae (Table S2). Calcineurin B-like proteins (CBLs) are calcium binding proteins that interact in the presence of calcium with a group of serine/threonine kinases designated as CBL-interacting protein kinases (CIPKs). This signaling network in plants allows the coupling of several different types of stress to a specific response. The most common is the regulation of salt stress [87]. Since *An. aquasalis* larvae are highly tolerant to salt stress, the CIPKs could be part of a similar response in which overexpression of these genes promotes salt tolerance [88,89].

The DUSP (contig 21444) mentioned above and a serine/threonine-protein phosphatase 4 regulatory subunit 1 (PP4) (contig 20076) are among the most down-regulated transcripts following a blood meal (Table S2). Serine/threonine phosphatases are divided in two main families, PPP and PPM. PPP are divided in five subtypes (PP1, PP2A, PP3, PP5 and PP7). PP4 protein belongs to the PP2A subfamily and like this enzyme is modulated by R regulatory subunits [90,91]. PP4 is likely involved in the repair of DNA double-strand breaks but it was recently demonstrated that it also acts as a negative regulator of negative regulator of NF-κB activity in T lymphocytes [92,93]. PP4R1 provides the interaction between the IκB kinase (IKK) complex and the phosphatase PP4c, thereby dephosphorylating and inactivating the IKK complex. The inactivation of IKK complex blocks NF-kB activation once its inhibitors, called IκBs (Inhibitor of κB), remain bound to the NF-kB complex. Deficiency of PP4R1 caused sustained and increased IKK activity and thus the permanent inhibition of immune responses [93]. Mosquito Rel1 and Rel2 members of the NF-kB transcription factors are activated after a blood meal and their silencing blocks the establishment of an immune response against malaria parasite. This occurs due to the inhibition of the basal expression of the anti-plasmodium genes TEP1 and LRIM,1 which are involved in the mosquito resistance to malaria parasite [94]. Thus PP4 might represent a long term down regulator of NF-kB activation and its 9-fold down regulation after a blood meal supports the hypothesis that it is required to enhance mosquito refractoriness to eventual pathogen infection. Future molecular analysis of such pathways together with the precise identification of the phosphorylation sites affected may reveal novel targets to overcome disease transmission by anophelines.

Conservation of gene regulation between *An. aquasalis* and *An. gambiae*

The changes in transcript abundance between larvae and adults and between sugar fed and blood fed females observed in our study were compared with those previously described for *An. gambiae* in similar developmental stages. A total of 8,355 contigs presented here had a homolog *An. gambiae* transcript (best Blast match) represented in the GeneChip *Plasmodium/Anopheles* Genome Array [42,95]. The pairwise comparisons including all *An. aquasalis/An. gambiae* homologous pairs demonstrated a lack of conservation of developmental changes in gene expression between the two mosquito species. Approximately half of the genes showed consistent up or down regulation in both species while the remaining showed up regulation in one mosquito and down in the other (Figure 4).

A more stringent analysis was performed by restricting the transcript list to putative 1:1 orthologous pairs defined by
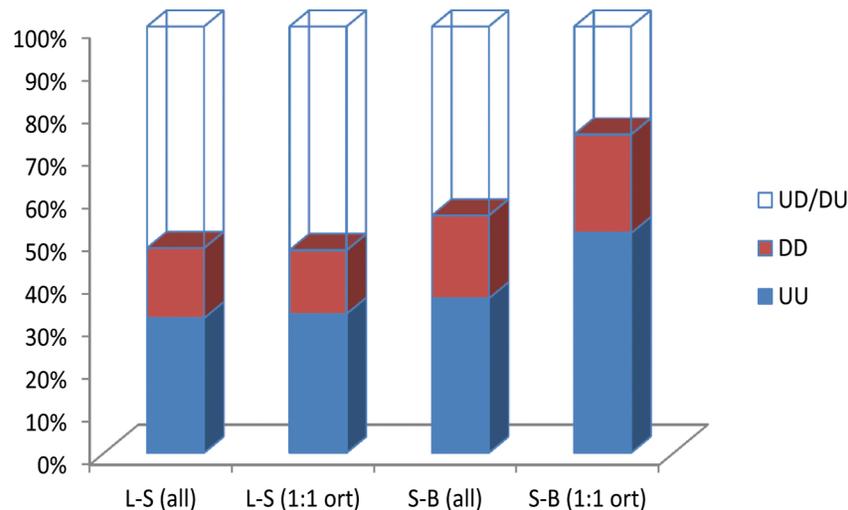
Figure 4. Comparisons of developmental changes in gene expression between *An. aquasalis* and *An. gambiae*. Developmental gene regulations [up(U) or down(D)-regulation] between larvae and sugar fed females (L-S all) or between sugar fed females and blood fed females (S-B all) of *An. aquasalis* transcripts that have a homolog *An. gambiae* (best Blast match) represented in the GeneChip *Plasmodium/Anopheles* Genome Array [42,95] were compared. The pairwise comparisons including all *An. aquasalis/An. gambiae* homologous pairs of genes demonstrated a lack of conservation of developmental changes in transcript abundance between the two mosquito species. Similar analyses restricting the transcript list to putative 1:1 ortholog pairs, defined by reciprocal blast and only those significantly regulated in *An. aquasalis*, with at least 3 fold change between two compared samples (L-S 1:1 ort or S-B 1:1 ort) showed that 75% the transcripts regulated by blood feeding were consistently up or down regulated in both species. Using the same restricted list of transcripts, only 49% of the transcripts were consistently up- or down-regulated between L-S in both species. Genes up-regulated or down-regulated in both species are indicated by (UU) or (DD), respectively. Transcripts differentially regulated between the two species are indicated by (UD/DU).
doi:10.1371/journal.pntd.0003005.g004

reciprocal blast. Additionally, the list included only those regulated significantly in *An. aquasalis*, and with ≥3-fold change between two compared samples (larvae/sugar fed adult females or sugar fed females/blood fed females). Applying these more stringent parameters, 75% of the transcripts regulated by blood feeding were consistently up- or down-regulated in both species. The same was not observed for the transcripts accumulated differentially between larvae and adults where only 49% were consistently up- or down-regulated in both species (Figure 4).

This study represents the first effort to sequence the transcriptome of the New World malaria vector, *An. aquasalis*. We have explored the transcriptomes of larva and adult *An. aquasalis*, providing valuable information about protein-coding transcripts involved in biological processes relevant to mosquito development, blood feeding, blood digestion, reproduction, and the *Plasmodium* life cycle. This study, together with other recently published and ongoing efforts to sequence the genomes and transcriptomes of malaria vectors (vectorbase.org) [96–98], will provide a needed and more complete understanding of malaria vector biology.

Our findings on gene functionalities shed light on the essential physiology of *An. aquasalis* and thus may help one to develop new control strategies. Moreover, present data may act as shortcuts to investigate genes of other congeneric pathogen-vectors. Data may also be used as taxonomic molecular markers or in future phylogenetic inferences (of genes or species) based on exons which are not under differential between-taxa natural selection.

Finally, one limitation of the sequencing project reported here is that transcripts present only at developmental stages not included in this study (embryos, pupae, adult males) could not be detected. It also is important to be aware that accumulation levels and variations in transcript abundance may not correlate with a similar variation in the amount of the encoded protein. Furthermore,

enzyme activity may be subject to regulation by feedback inhibition by the corresponding pathway product, allosteric interactions, reversible covalent modifications or programmed proteolytic cleavage.

Differently from other Anophelines, the complete genome sequence of the *An. aquasalis* was not obtained until now. This fact imposes a limitation to estimate how complete this transcriptome is, and the size of coding genome as well as orthology comparison between related species needs to be adopted with this assumption.

Further studies to generate a comprehensive picture of gene expression, protein synthesis and function throughout the mosquito development are needed to uncover biological processes in mosquitoes and to help in the efforts to control malaria transmission.

Supporting Information

Figure S1 Summary of assembly compared to the raw reads. A. Number of sequences in raw file in relation to sequence size. B. Number of reads in assembled results in relation to sequence size. (TIF)

Figure S2 Validation of blood meal regulated transcripts by qRT-PCR. Expression levels of 5 genes (megaclu_asbSigP-9948, Anoaqua-3237, megaclu_asbSigP-2537, Anoaqua-5059 and Anoaqua-24500) were determined in samples from *An. aquasalis* L3/L4 larvae pools (Larvae), adult females fed on sucrose (Sucrose-fed) and adult females 24 hs post blood meal (Blood-fed). Relative mRNA levels are displayed as the mean of fold differences in relation to the calibrator sample (Larvae) and error bars represent the standard error of the mean (± S.E.M.). Expression levels of megaclu_asbSigP-9948 in blood-fed samples are significant different in relation to other samples (Kruskal-Wallis test. *, $p <$ 0.0001). Expression levels of Anoaqua-3237, megaclu_asbSigP-

2537, Anoaqua-5059 and Anoaqua-24500 in blood-fed samples are significant different in relation to other measured samples (One way ANOVA test. *, $p<0.0001$).
(TIF)

Figure S3 Validation of larva enhanced and specific transcripts by qRT-PCR. Expression levels of 5 genes (Anoaqua-397, Anoaqua-4095, Anoaqua-1222, Anoaqua-1598 and Anoaqua-17360) were determined in samples from *An. aquasalis* L3/L4 larvae pools (Larvae), adult females fed on sucrose (Sucrose-fed) and adult females 24 hs post blood meal (Blood-fed). Relative mRNA levels are displayed as the mean of fold differences in relation to calibrator sample (Sucrose-fed for Anoaqua-397 and Anoaqua-4095; Blood-fed for Anoaqua-1222, Anoaqua-1598 and Anoaqua-17360) and error bars represent the standard error of the mean (± S.E.M.). Expression levels of Anoaqua-397, Anoaqua-4095, Anoaqua-1222 and Anoaqua-1598 in Larvae samples are significant different in relation to other measured samples (One way ANOVA test. *, $p<0.0001$). Expression levels of Anoaqua-17360 in Larvae samples are significant different in relation to other samples (Kruskal-Wallis test. *, $p<0.0001$).
(TIF)

Figure S4 Validation of adult enhanced and specific transcripts by qRT-PCR. Expression levels of 2 genes (Anoaqua-436 and Anoaqua-457) were determined in samples from *An. aquasalis* L3/L4 larvae pools (Larvae), adult females fed on sucrose (Sucrose-fed) and adult females 24 hs post blood meal (Blood-fed). Relative mRNA levels are displayed as the mean of fold differences in relation to calibrator sample (Larvae) and error bars represent the standard error of the mean (± S.E.M.). Expression levels of Anoaqua-436 and Anoaqua-457 in Larvae samples are significant

different in relation to other measured samples (One way ANOVA test. *, $p<0.0001$).
(TIF)

Table S1 Most abundant transcripts in larvae, sugar fed and blood fed adult female *An. aquasalis*.
(XLSX)

Table S2 Most up or down-regulated transcripts during *An.aquasalis* development and following a blood meal.
(XLSX)

Table S3 Primer pairs used for qRT-PCR validation and PCR efficiency of the targets
(XLSX)

Table S4 Validation of Rp49 as housekeeping gene to normalize expression values of selected candidates.
(XLSX)

Author Contributions

Conceived and designed the experiments: EC JMCR OM ALCdS ASN MLC. Performed the experiments: ALCdS EC. Analyzed the data: ALCdS OM JMCR BBK. Contributed reagents/materials/analysis tools: MLC JMCR. Wrote the paper: OM JMCR ALCdS MCPS ARL MSB ASN BBK EC LS MACSN AAJ MLC.

References

1. Deane LM (1986) Malaria Vectors in Brazil. Mem Inst Oswaldo Cruz 81: 5–14.
2. Zimmerman RH (1992) Ecology of malaria vectors in the Americas and future direction. Mem Inst Oswaldo Cruz 87 Suppl 3: 371–383. doi:10.1590/S0074-02761992000700064.
3. Rozendaal JA (1987) Observations on the biology and behaviour of Anophelines in the Suriname rainforest with special reference to Anopheles darlingi Root. Cah ORSTOM sér Entomol méd Parasitol 25: 33–43.
4. Laubach HE, Validum L, Bonilla JA, Agar A, Cummings R, et al. (2001) Identification of Anopheles aquasalis as a possible vector of malaria in Guyana, South America. West Indian Med J 50: 319–321.
5. Sinka ME, Rubio-Palis Y, Manguin S, Patil AP, Temperley WH, et al. (2010) The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. Parasit Vectors 3: 72. doi:10.1186/1756-3305-3-72.
6. Hay SI, Snow RW (2006) The malaria Atlas Project: developing global maps of malaria risk. PLoS Med 3: e473. doi:10.1371/journal.pmed.0030473.
7. Faran M, Linthicum K (1981) A handbook of the Amazonian species of Anopheles (Nyssorhynchus) (Diptera: Culicidae). Mosq Syst 13: 85.
8. Molina D, Saume F, Bisset J, Hidalgo O, Castillo M, et al. (1997) Establecimiento de la línea de susceptibilidad de la fase adulta de Anopheles spp. a insecticidas químicos. Bol Dir Malar San Amb 37: 55–69. Bol Dir Malar San Amb 37: 55–69.
9. Davis NC (1931) A note on the malaria-carrying Anophelines in Belém, Pará, and in Natal, Rio Grande do Norte, Brazil. Riv Malariol 10: 43–51.
10. Galvão ALA, Damasceno RG, Marques AP (1942) Algumas observações sobre a biologia dos anofelinos de importância epidemiológica. Arch Hig 12: 51–110.
11. Póvoa MM, Conn JE, Schlichting CD, Amaral JCOF, Segura MNO, et al. (2003) Malaria vectors, epidemiology, and the re-emergence of Anopheles darlingi in Belém, Pará, Brazil. J Med Entomol 40: 379–386.
12. Da Silva AR, Tauil PL, Bastos Júnior JL, de Matos WB, Costa EAP, et al. (2006) Aspects of the focal transmission of malaria in the Island of São Luis, Maranhão. Rev Soc Bras Med Trop 39: 250–254. doi:10.1590/S0037-86822006000300004.
13. Downs WG, Gillette HPS, Shannon RC (1943) A malaria survey of Trinidad and Tobago, British West Indies. J Natl Malar Soc 2: 5–44.
14. Chadee DD, Kitron U (1999) Spatial and temporal patterns of imported malaria cases and local transmission in Trinidad. Am J Trop Med Hyg 61: 513–517.
15. Gething PW, Elyazar IRF, Moyes CL, Smith DL, Battle KE, et al. (2012) A long neglected world malaria map: Plasmodium vivax endemicity in 2010. PLoS Negl Trop Dis 6: e1814. doi:10.1371/journal.pntd.0001814.
16. Arruda ME, Rios RI, Arantes PC, Oliveira AC, P NL (1982) Manutenção em laboratório de Anopheles albitarsis e Anopheles aquasalis por copulacao induzida. Mem Inst Oswaldo Cruz 77: 89–91.
17. Carvalho SCG de, Martins Junior A de J, Lima JBP, Valle D (2002) Temperature influence on embryonic development of Anopheles albitarsis and Anopheles aquasalis. Mem Inst Oswaldo Cruz 97: 1117–1120. doi:10.1590/S0074-02762002000800009.
18. Bahia AC, Kubota MS, Tempone AJ, Pinheiro WD, Tadei WP, et al. (2010) Anopheles aquasalis Infected by Plasmodium vivax Displays Unique Gene Expression Profiles when Compared to Other Malaria Vectors and Plasmodia. PLoS One 5: e9795. doi:10.1371/journal.pone.0009795.
19. Bahia AC, Oliveira JHM, Kubota MS, Araújo HRC, Lima JBP, et al. (2013) The role of reactive oxygen species in Anopheles aquasalis response to Plasmodium vivax infection. PLoS One 8: e57014. doi:10.1371/journal.pone.0057014.
20. Oliveira SB, Ibraim IC, Tadei WP, Ruiz JC, Nahum LA, et al. (2011) Identification of a fibrinogen-related protein (FBN9) gene in neotropical anopheline mosquitoes. Malar J 10: 21. doi:10.1186/1475-2875-10-21.
21. Souza-Neto JA, Machado FP, Lima JB, Valle D, Ribolla PEM (2007) Sugar digestion in mosquitoes: Identification and characterization of three midgut α-glucosidases of the neo-tropical malaria vector Anopheles aquasalis (Diptera: Culicidae). Comp Biochem Physiol A Mol Integr Physiol 147: 993–1000. doi:10.1016/j.cbpa.2007.03.008.
22. De Almeida RW, Tovar FJ, Ferreira II, Leoncini O (2003) Chymotrypsin genes in the malaria mosquitoes Anopheles aquasalis and Anopheles darlingi. Insect Biochem Mol Biol 33: 307–315. doi:10.1016/S0965-1748(02)00245-X.
23. Caroci AS, Calvo E, Ribolla PE, De Bianchi AG, Marinotti O (2003) Two digestive trypsins occur in three species of neotropical anophelines. J Med Entomol 40: 991–995.
24. Valle D, Monnerat AT, Soares MJ, Rosa-Freitas MG, Pelajo-Machado M, et al. (1999) Mosquito embryos and eggs: polarity and terminology of chorionic layers. J Insect Physiol 45: 701–708.
25. Rios-Velásquez CM, Martins-Campos KM, Simões RC, Izzo T, Dos Santos E V, et al. (2013) Experimental Plasmodium vivax infection of key Anopheles species from the Brazilian Amazon. Malar J 12: 460. doi:10.1186/1475-2875-12-460.
26. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. J Mol Biol 215: 403–410.
27. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868–877.

28. Karim S, Singh P, Ribeiro JMC (2011) A deep insight into the sialotranscriptome of the gulf coast tick, Amblyomma maculatum. PLoS One 6: e28525. doi:10.1371/journal.pone.0028525.

29. Ribeiro JMC, Topalis P, Louis C (2004) AnoXcel: an Anopheles gambiae protein database. Insect Mol Biol 13: 449–457. doi:10.1111/j.0962-1075.2004.00503.x.

30. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.

31. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29. doi:10.1038/75556.

32. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, et al. (2000) The Pfam protein families database. Nucleic Acids Res 28: 263–266.

33. Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000) SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res 28: 231–234.

34. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, et al. (2003) The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4: 41. doi:10.1186/1471-2105-4-41.

35. Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, et al. (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 30: 281–283.

36. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein eng 10: 1–6.

37. Julenius K, Mølgaard A, Gupta R, Brunak S (2005) Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology 15: 153–164. doi:10.1093/glycob/cwh151.

38. Gentile C, Lima JB, Peixoto AA (2005) Isolation of a fragment homologous to the rp49 constitutive gene of Drosophila in the Neotropical malaria vector Anopheles aquasalis (Diptera: Culicidae). Mem Inst Oswaldo Cruz 100: 545–547. doi:/S0074-02762005000600008.

39. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402–408. doi:10.1006/meth.2001.1262.

40. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. (2002) The genome sequence of the malaria mosquito Anopheles gambiae. Science 298: 129–149. doi:10.1126/science.1076181.

41. Ribeiro JMC (2003) A catalogue of Anopheles gambiae transcripts significantly more or less expressed following a blood meal. Insect Biochem Mol Biol 33: 865–882.

42. Marinotti O, Nguyen QK, Calvo E, James AA, Ribeiro JMC (2005) Microarray analysis of genes showing variable expression following a blood meal in Anopheles gambiae. Insect Mol Biol 14: 365–373. doi:10.1111/j.1365-2583.2005.00567.x.

43. Marinotti O, Calvo E, Nguyen QK, Dissanayake S, Ribeiro JMC, et al. (2006) Genome-wide analysis of gene expression in adult Anopheles gambiae. Insect Mol Biol 15: 1–12. doi:10.1111/j.1365-2583.2006.00610.

44. Dissanayake SN, Ribeiro JM, Wang M-H, Dunn WA, Yan G, et al. (2010) aeGEPUCI: a database of gene expression in the dengue vector mosquito, Aedes aegypti. BMC Res Notes 3: 248. doi:10.1186/1756-0500-3-248.

45. Haunerland NH (1996) Insect storage proteins: gene families and receptors. Insect Biochem Mol Biol 26: 755–765.

46. Burmester T, Massey HC, Zakharkin SO, Benes H (1998) The evolution of hexamerins and the phylogeny of insects. J Mol Evol 47: 93–108.

47. Totten DC, Vuong M, Litvinova O V, Jinwal UK, Gulia-Nuss M, et al. (2013) Targeting gene expression to the female larval fat body of transgenic Aedes aegypti mosquitoes. Insect Mol Biol 22: 18–30. doi:10.1111/imb.12005.

48. Togawa T, Dunn WA, Emmons AC, Nagao J, Willis JH (2008) Developmental expression patterns of cuticular protein genes with the R&R Consensus from Anopheles gambiae. Insect Biochem Mol Biol 38: 508–519. doi:10.1016/j.ibmb.2007.12.008.

49. Merrill CE, Riesgo-Escovar J, Pitts RJ, Kafatos FC, Carlson JR, et al. (2002) Visual arrestins in olfactory pathways of Drosophila and the malaria vector mosquito Anopheles gambiae. Proc Natl Acad Sci U S A 99: 1633–1638. doi:10.1073/pnas.022505499.

50. Biessmann H, Nguyen QK, Le D, Walter MF (2005) Microarray-based survey of a subset of putative olfactory genes in the mosquito Anopheles gambiae. Insect Mol Biol 14: 575–589. doi:10.1111/j.1365-2583.2005.00590.x.

51. Terra WR (2001) The origin and functions of the insect peritrophic membrane and peritrophic gel. Arch Insect Biochem Physiol 47: 47–61. doi:10.1002/arch.1036.

52. Devenport M, Fujioka H, Jacobs-Lorena M (2004) Storage and secretion of the peritrophic matrix protein Ag-Aper1 and trypsin in the midgut of Anopheles gambiae. Insect Mol Biol 13: 349–358. doi:10.1111/j.0962-1075.2004.00488.x.

53. Villalon JM, Ghosh A, Jacobs-Lorena M (2003) The peritrophic matrix limits the rate of digestion in adult Anopheles stephensi and Aedes aegypti mosquitoes. J Insect Physiol 49: 891–895.

54. Abraham EG, Jacobs-Lorena M (2004) Mosquito midgut barriers to malaria parasite development. Insect Biochem Mol Biol 34: 667–671. doi:10.1016/j.ibmb.2004.03.019.

55. Müller HM, Catteruccia F, Vizioli J, della Torre A, Crisanti A (1995) Constitutive and blood meal-induced trypsin genes in Anopheles gambiae. Exp Parasitol 81: 371–385.

56. Seidah NG, Prat A (2002) Precursor convertases in the secretory pathway, cytosol and extracellular milieu. Essays Biochem 38: 79–94.

57. Neurath H (1989) Proteolytic processing and physiological regulation. Trends Biochem Sci 14: 268–271.

58. Ribeiro JMC, Mans BJ, Arcà B (2010) An insight into the sialome of blood-feeding Nematocera. Insect Biochem Mol Biol 40: 767–784. doi:10.1016/j.ibmb.2010.08.002.

59. Shen Z, Jacobs-Lorena M (1998) Nuclear factor recognition sites in the gut-specific enhancer region of an Anopheles gambiae trypsin gene. Insect Biochem Mol Biol 28: 1007–1012.

60. Shen Z, Dimopoulos G, Kafatos FC, Jacobs-Lorena M (1999) A cell surface mucin specifically expressed in the midgut of the malaria mosquito Anopheles gambiae. Proc Natl Acad Sci U S A 96: 5610–5615.

61. Vizioli J, Catteruccia F, della Torre A, Reckmann I, Müller HM (2001) Blood digestion in the malaria mosquito Anopheles gambiae: molecular cloning and biochemical characterization of two inducible chymotrypsins. Eur J Biochem 268: 4027–4035.

62. Chadee DD, Beier JC (1995) Blood-digestion kinetics of four Anopheles species from Trinidad, West Indies. Ann Trop Med Parasitol 89: 531–540.

63. Nolan T, Petris E, Müller H-M, Cronin A, Catteruccia F, et al. (2011) Analysis of two novel midgut-specific promoters driving transgene expression in Anopheles stephensi mosquitoes. PLoS One 6: e16471. doi:10.1371/journal.pone.0016471.

64. Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Dimon MT, et al. (2011) RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, Aedes aegypti. BMC Genomics 12: 82. doi:10.1186/1471-2164-12-82.

65. Berois M, Romero-Severson J, Severson DW (2012) RNAi knock-downs support roles for the mucin-like (AeIMUC1) gene and short-chain dehydrogenase/reductase (SDR) gene in Aedes aegypti susceptibility to Plasmodium gallinaceum. Med Vet Entomol 26: 112–115. doi:10.1111/j.1365-2915.2011.00965.x.

66. Ahmed AM, Maingon R, Romans P, Hurd H (2001) Effects of malaria infection on vitellogenesis in Anopheles gambiae during two gonotrophic cycles. Insect Mol Biol 10: 347–356.

67. Rono MK, Whitten MMA, Oulad-Abdelghani M, Levashina EA, Marois E (2010) The major yolk protein vitellogenin interferes with the anti-plasmodium response in the malaria mosquito Anopheles gambiae. PLoS Biol 8: e1000434. doi:10.1371/journal.pbio.1000434.

68. Margaritis LH (1985) The egg-shell of Drosophila melanogaster III. Covalent crosslinking of the chorion proteins involves endogenous hydrogen peroxide. Tissue Cell 17: 553–559.

69. Amenya D a, Chou W, Li J, Yan G, Gershon PD, et al. (2010) Proteomics reveals novel components of the Anopheles gambiae eggshell. J Insect Physiol 56: 1414–1419. doi:10.1016/j.jinsphys.2010.04.013.

70. Terenius O, Marinotti O, Sieglaff D, James AA (2008) Molecular genetic manipulation of vector mosquitoes. Cell Host Microbe 4: 417–423. doi:10.1016/j.chom.2008.09.002.

71. Fu G, Lees RS, Nimmo D, Aw D, Jin L, et al. (2010) Female-specific flightless phenotype for mosquito control. Proc Natl Acad Sci U S A 107: 4550–4554. doi:10.1073/pnas.1000251107.

72. Marinotti O, Jasinskiene N, Fazekas A, Scaife S, Fu G, et al. (2013) Development of a population suppression strain of the human malaria vector mosquito, Anopheles stephensi. Malar J 12: 142. doi:10.1186/1475-2875-12-142.

73. Isaacs AT, Li F, Jasinskiene N, Chen X, Nirmala X, et al. (2011) Engineered resistance to Plasmodium falciparum development in transgenic Anopheles stephensi. PLoS Pathog 7: e1002017. doi:10.1371/journal.ppat.1002017.

74. Wang S, Jacobs-Lorena M (2013) Genetic approaches to interfere with malaria transmission by vector mosquitoes. Trends Biotechnol 31: 185–193. doi:10.1016/j.tibtech.2013.01.001.

75. Sperança MA, Capurro ML (2007) Perspectives in the control of infectious diseases by transgenic mosquitoes in the post-genomic era: a review. Mem Inst Oswaldo Cruz 102: 425–433. doi:10.1590/S0074-02762007000500054.

76. Shen Z, Edwards MJ, Jacobs-Lorena M (2000) A gut-specific serine protease from the malaria vector Anopheles gambiae is downregulated after blood ingestion. Insect Mol Biol 9: 223–229.

77. Giannoni F, Müller HM, Vizioli J, Catteruccia F, Kafatos FC, et al. (2001) Nuclear factors bind to a conserved DNA element that modulates transcription of Anopheles gambiae trypsin genes. J Biol Chem 276: 700–707. doi:10.1074/jbc.M005540200.

78. Bahia AC, Kubota MS, Tempone AJ, Araújo HRC, Guedes BAM, et al. (2011) The JAK-STAT pathway controls Plasmodium vivax load in early stages of Anopheles aquasalis infection. PLoS Negl Trop Dis 5: e1317. doi:10.1371/journal.pntd.0001317.

79. Lim WA, Pawson T (2010) Phosphotyrosine signaling: evolving a new cellular communication system. Cell 142: 661–667. doi:10.1016/j.cell.2010.08.023.

80. Silveira AB, Castro-Santos J, Senna R, Logullo C, Fialho E, et al. (2006) Tick vitellin is dephosphorylated by a protein tyrosine phosphatase during egg development: effect of dephosphorylation on VT proteolysis. Insect Biochem Mol Biol 36: 200–209. doi:10.1016/j.ibmb.2006.01.015.

81. Jablonka W, Senna R, Nahu T, Ventura G, Menezes L, et al. (2011) A transient increase in total head phosphotyrosine levels is observed upon the emergence of Aedes aegypti from the pupal stage. Mem Inst Oswaldo Cruz 106: 546–552. doi:10.1590/S0074-02762011000500005.

82. Gazos-Lopes F, Mesquita RD, Silva-Cardoso L, Senna R, Silveira AB, et al. (2012) Glycoinositolphospholipids from Trypanosomatids subvert nitric oxide production in Rhodnius prolixus salivary glands. PLoS One 7: e47285. doi:10.1371/journal.pone.0047285.

83. Furth PA, Nakles RE, Millman S, Diaz-Cruz ES, Cabrera MC (2011) Signal transducer and activator of transcription 5 as a key signaling pathway in normal mammary gland developmental biology and breast cancer. Breast Cancer Res 13: 220. doi:10.1186/bcr2921.

84. Du WW, Fang L, Li M, Yang X, Liang Y, et al. (2013) MicroRNA miR-24 enhances tumor invasion and metastasis by targeting PTPN9 and PTPRF to promote EGF signaling. J Cell Sci 126: 1440–1453. doi:10.1242/jcs.118299.

85. Barillas-Mury C, Han YS, Seeley D, Kafatos FC (1999) Anopheles gambiae Ag-STAT, a new insect member of the STAT family, is activated in response to bacterial infection. EMBO J 18: 959–967. doi:10.1093/emboj/18.4.959.

86. Horton AA, Wang B, Camp L, Price MS, Arshi A, et al. (2011) The mitogen-activated protein kinome from Anopheles gambiae: identification, phylogeny and functional characterization of the ERK, JNK and p38 MAP kinases. BMC Genomics 12: 574. doi:10.1186/1471-2164-12-574.

87. Hashimoto K, Eckert C, Anschütz U, Scholz M, Held K, et al. (2012) Phosphorylation of calcineurin B-like (CBL) calcium sensor proteins by their CBL-interacting protein kinases (CIPKs) is required for full activity of CBL-CIPK complexes toward their target proteins. J Biol Chem 287: 7956–7968. doi:10.1074/jbc.M111.279331.

88. Osborn FR, Díaz S, Gómez CJ, Moreno M, Hernández G (2006) Oviposition preference and egg eclosion in different salt concentrations in the coastal malaria vector Anopheles aquasalis Curry. J Am Mosq Control Assoc 22: 42–46.

89. Tripathi V, Parasuraman B, Laxmi A, Chattopadhyay D (2009) CIPK6, a CBL-interacting protein kinase is required for development and salt tolerance in plants. Plant J 58: 778–790. doi:10.1111/j.1365-313X.2009.03812.x.

90. Brewis ND, Street AJ, Prescott AR, Cohen PT (1993) PPX, a novel protein serine/threonine phosphatase localized to centrosomes. EMBO J 12: 987–996.

91. Hu MC, Shui JW, Mihindukulasuriya KA, Tan TH (2001) Genomic structure of the mouse PP4 gene: a developmentally regulated protein phosphatase. Gene 278: 89–99.

92. Pfeifer GP (2012) Protein phosphatase PP4: role in dephosphorylation of KAP1 and DNA strand break repair. Cell cycle 11: 2590–2591. doi:10.4161/cc.21234.

93. Brechmann M, Mock T, Nickles D, Kiessling M, Weit N, et al. (2012) A PP4 holoenzyme balances physiological and oncogenic nuclear factor-kappa B signaling in T lymphocytes. Immunity 37: 697–708. doi:10.1016/j.immuni.2012.07.014.

94. Frolet C, Thoma M, Blandin S, Hoffmann JA, Levashina EA (2006) Boosting NF-kappaB-dependent basal immunity of Anopheles gambiae aborts development of Plasmodium berghei. Immunity 25: 677–685. doi:10.1016/j.immuni.2006.08.019.

95. Marinotti O, Calvo E, Nguyen QK, Dissanayake S, Ribeiro JMC, et al. (2006) Genome-wide analysis of gene expression in adult Anopheles gambiae. Insect Mol Biol 15: 1–12. doi:10.1111/j.1365-2583.2006.00610.x.

96. Martínez-Barnetche J, Gómez-Barreto RE, Ovilla-Muñoz M, Téllez-Sosa J, López DEG, et al. (2012) Transcriptome of the adult female malaria mosquito vector Anopheles albimanus. BMC Genomics 13: 207. doi:10.1186/1471-2164-13-207.

97. Marinotti O, Cerqueira GC, de Almeida LGP, Ferro MIT, Loreto EL d. S, et al. (2013) The Genome of Anopheles darlingi, the main neotropical malaria vector. Nucleic Acids Res 41: 7387–7400 doi:10.1093/nar/gkt484.

98. Neafsey DE, Christophides GK, Collins FH, Emrich SJ, Fontaine MC, et al. (2013) The Evolution of the Anopheles 16 Genomes Project. G3 (Bethesda) 3: 1191–1194. doi:10.1534/g3.113.006247.