Gustatory receptor expression in the labella and tarsi of *Aedes aegypti*

Jackson T. Sparks\(^2\), Bryan T. Vinyard\(^b\), Joseph C. Dickens\(^a\,\,*\)

\(^a\) United States Department of Agriculture, Agricultural Research Service, Henry A. Wallace Beltsville Agricultural Research Center, Plant Sciences Institute, Invasive Insect Biocontrol and Behavior Laboratory, Beltsville, MD 20705, USA

\(^b\) United States Department of Agriculture, Agricultural Research Service, Henry A. Wallace Beltsville Agricultural Research Center, Biometrical Consulting Service, Beltsville, MD, USA

**ABSTRACT**

The yellow-fever mosquito, *Aedes aegypti*, infects a growing number of people every year with dengue, yellow fever and chikungunya viruses. Contact chemoreception in mosquitoes influences a number of behaviors including host-selection, oviposition and feeding. While these behaviors are in many instances well documented, the molecular mechanisms mediating them are not well understood. Here we report the results of sequencing total messenger RNA in the labella and tarsi of both male and female *Ae. aegypti* to reveal Gustatory Receptor (GR) gene expression profiles in these major gustatory appendages. Gene expression levels in each tissue were verified by RT-qPCR. We discuss potential functions for the GRs revealed here by considering homologous GRs in other insects. Specific GRs provide molecular targets for modification of gustatory-mediated behaviors in this important disease vector.

**1. Introduction**

*Aedes aegypti* (L.) (Diptera: Culicidae) is an important disease vector, contributing to the spread of dengue, yellow fever, chikungunya and West Nile viruses. Perhaps most alarming is the deadly dengue virus, as female *Ae. aegypti* infect between 50 and 390 million people yearly through blood-meal mediated viral transmission (Guzman et al., 2010; Beatty et al., 2011; Bhatt et al., 2013). The production of a dengue vaccine is challenging, and unpredictable viral mutation threatens long-term efficacy (Barban et al., 2012; Sabchareon et al., 2012; Wallace et al., 2013). Thus, prevention of mosquito bites remains one of the chief strategies to break this viral transmission cycle. Deterrents such as DEET can reduce the chances of mosquito bites and prevent disease transmission (Debboun and Strickman, 2013). Central to the discovery of other useful deterrents is the identification of the receptor molecules mediating responses to such compounds. While *Ae. aegypti* olfactory receptors have recently been implicated in the detection of volatile DEET (DeGennaro et al., 2013), receptors involved in gustatory responses to DEET and other deterrent compounds (Sanford et al., 2013) have not been determined for mosquitoes.

The repertoire of genes involved in insect gustatory reception is diverse (Liu et al., 2003; Lin et al., 2005; Al-Anzi et al., 2006; Fischler et al., 2007; Mitri et al., 2009; Cameron et al., 2010; Chen et al., 2010; Croset et al., 2010; Kim et al., 2010; Zhang et al., 2013) and includes the large, divergent gene family of Gustatory Receptors (GRs) (Clyne et al., 2000; Dunipace et al., 2001; Scott et al., 2001; Robertson et al., 2003; Isono and Morita, 2010). Membrane-bound insect GRs are generally expressed in gustatory receptor neurons (GRNs) whose dendritic processes innervate sensory hairs or sensilla distributed on the paired labella, labrum, inner mouthparts, wing margins, genitalia, and tarsal segments of the legs (Dunipace et al., 2001; Scott et al., 2001; Thorne et al., 2004; Wang et al., 2004; Dahanukar et al., 2007). The porous tip of these sensilla allows chemicals to come into contact with GRNs within the sensillar lumen (Stocker, 1994). At this interface, individual or complexes of GRs are selectively activated by chemicals, leading to neuronal responses resulting in action potentials that are carried by axonal processes to the primary gustatory centers of the brain or other central ganglia (Miyazaki and Ito, 2010). Here information from various sensory modalities is integrated, thus enabling insects, including mosquitoes, to assess and respond to their chemical environment.

Mosquitoes locate potential food sources and mates at a distance using olfactory, visual, and other cues like temperature and humidity (Peterson and Brown, 1952; Laarman, 1958; Bidlingmayer and Hem, 1980; Bowen, 1991; Clements, 1992; Klun et al., 2013). Upon landing, sensory hairs of the tarsal segments make the first
The yellow-fever mosquito, Aedes aegypti, infects a growing number of people every year with dengue yellow fever and chikungunya viruses. Contact chemoreception in mosquitoes influences a number of behaviors including host-selection, oviposition and feeding. While these behaviors are in many instances well documented, the molecular mechanisms mediating them are not well understood. Here we report the results of sequencing total messenger RNA in the labella and tarsi of both male and female Aedes aegypti to reveal Gustatory Receptor (GR) gene expression profiles in these major gustatory appendages. Gene expression levels in each tissue were verified by RT-qPCR. We discuss potential functions for the GRs revealed here by considering homologous GRs in other insects. Specific GRs provide molecular targets for modification of gustatory-mediated behaviors in this important disease vector.
contact assessment of the sugar source or blood-host. This initial gustatory assessment is followed by a brief exploratory phase wherein the labellum repeatedly contacts the surface of the plant or animal host to further evaluate nutritive resources (Clements, 1992). Contact gustatory reception may also be used by mosquitoes to assess the social status of conspecifics, as is the case for other insects (Sturtevant, 1915; Greenspan and Ferveur, 2000). In Ae. aegypti, gustatory sensilla are located in stereotyped positions on the labella and tarsi of the legs (Chaika and Elizarov, 1971; McIver and Siemicki, 1978; Hill and Smith, 1999; Lee and Craig, 2009), although the specific expression of GRs in these tissues has not yet been reported.

Putative Ae. aegypti GRs have been annotated from genomic sequences and compared to previously identified insect GRs (Kent et al., 2008). In general, GRs of the two main mosquito sub-families, Anopheline and Culicine, share very low amino acid sequence similarity. However, a few GRs, including those related to sugar receptors or those activated by CO2, are conserved across many insect orders including mosquitoes (Kent and Robertson, 2009; Robertson and Kent, 2009). Conservation of these GR sequences between long diverged insect species likely reflects indispensable gustatory sensitivities to a particular chemical or set of chemicals, a property that allows us to speculate as to their potential function.

Here we employed two methods to determine expression profiles of GRs in the major gustatory appendages of Ae. aegypti. First, we identified and quantified putative AeagGR transcripts by RNA-seq analysis; total poly-adénylated (poly[A]) RNA was isolated from the labella, as well as the pro-, meso- and metathoracic tarsi of both males and females, converted to cDNA and sequenced (Illumina). Second, we employed RT-qPCR to independently validate the RNA-seq data. We discuss the significance of the expression profiles of AeagGRs with an emphasis on those GRs with sequence orthology to functional GRs in other insects. The GRs uncovered here provide targets for high-throughput screening of chemicals for use as feeding deterrents or stimulants aimed at disruption of mosquito behavior for the protection of humans and other animals.

2. Materials and methods

2.1. Animal rearing

Ae. aegypti eggs (Orlando strain) were obtained from the Center for Medical and Veterinary Entomology, USDA, ARS in Gainesville, FL, USA. Larvae were reared at 25 °C (12-hL:12-hD) and fed with ground TetraMin® fish food. Unsexed pupae were hand-collected daily and transferred to plastic dishes (9 cm × 5.5 cm) inside small containment buckets, thus establishing 24-h age groups. Greater than 95% of adults emerged 2 days post-pupation, after which all remaining pupae were removed from containment buckets. Adult mosquitoes were fed with a 10% sucrose solution which all remaining pupae were removed from containment buckets, thus establishing 24-h age groups. Non-quantitative amplification of gDNA using Platinum® Taq DNA Polymerase (Life Technologies) to confirm the correct identity and singularity of amplicons. PCR products were visualized using ethidium bromide on 2% agarose gels (Figure S1), purified with DNA Clean & Concentrator™-5 (Zymo Research, Irvine, CA, USA) and then directly sequenced to confirm

2.2. RNA isolation and sequencing

For RNA sequencing, paired labella from 500 males or 500 females were carefully dissected to limit inclusion of other adjacent proboscis tissues. Samples from legs were comprised of pro-, meso- or metathoracic tarsal segments of 400 males or 400 females. Dissected tissues were immediately stored on dry ice and mechanically ground in TRIzol® (Life Technologies, Carlsbad, CA, USA). Total RNA was isolated by RNeasy® Plus Mini Kit (Qiagen, Valencia, CA, USA), quantified on a Nanodrop ND-1000 spectrophotometer (Nanobag Products, Wilmington, DE, USA), and sent to the Genomics Services Lab at the Hudson Alpha Institute for Biotechnology (Huntsville, AL), where samples were first assessed by Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). mRNA isolation and cDNA synthesis were completed using NEBNext® agents (NEB, Ipswich, MA, USA) and standard protocols with custom GSL adaptors. cDNA libraries corresponding to distinct tissues were sequenced on an Illumina HiSeq2000 to generate 25 million 50 basepair, paired-end reads per sample.

2.3. Analysis of annotated and unannotated chemoreception genes

Reference genome (AaegL1, Scaffolds) and annotations (AaegL1.3) for Ae. aegypti were downloaded from VectorBase (http://aegypti.vectorbase.org/GetData/Downloads/). Output Fastq Illumina files were mapped to the genome and annotated gene build with TopHat 2.0.7 (Trapnell et al., 2009). The unambiguous sequence alignment files were uploaded into the Avadis NGS software (Strand Scientific Intelligence, CA, USA), where quantification and normalization were performed. Prior to quantification using the ‘Deseq’ normalization method, the read list was filtered to remove duplicate, single-end, mate-mismatched, mate-mate flip, single mate flip, and unaligned reads. Read quality metric values were sufficient to rule out low quality reads; Quality threshold ≥30, Ns allowed in read ≤0, Alignment score threshold ≥95, Mapping quality threshold ≥40. Picard software (picard.sourceforge.net) was used to determine the genomic location of all reads mapping to genome assembly. Transcript expression levels for all genes are reported in values of Reads Per Kilobase per Million reads mapped (RPKM). RPKM values represent a quantitative measure of the number of corresponding 50 bp sequence reads (sequenced in both directions) for a given gene. The 4 putative protein-coding loci associated with AaegGR40 (Kent et al., 2008) are unannotated and therefore not represented in our dataset. Attempts to resolve gene models manually for read mapping were unsuccessful due to limited genomic coverage and high similarity to the related AaegGR39 locus. In addition, specific read mapping assignments for AaegGR67a, AaegGR67b and AaegGR67c were grouped as one value, as alignments for each of these transcripts were difficult to distinguish from one another by the computational approach used with the other transcripts. We assigned no specific threshold for functional expression vs. background “noise.”

2.4. Quantitative RT-PCR validation of RNA sequencing

Thirteen chemoreception genes and one housekeeping gene were selected for qPCR analysis to evaluate gene expression over a dynamic range, both in predicted sequence abundance and presumed chemosensory gene function. Primer pairs were designed for each target gene to amplify a specific 100–180 basepair PCR product (Primer-BLAST Primer Designing tool, NCBI). At least one primer per set spans an intron boundary to exclude non-specific gDNA amplification. We dissected 150 paired labella and 200 tarsi of pro-, meso- and metathoracic legs from either male or female mosquitoes for each RNA extraction. Labella of both sexes were collected in biological triplicate (450 total). Total RNA was isolated as previously described, cDNA was synthesized using Superscript® III First-Strand Synthesis Supermix for qRT-PCR (Life Technologies, Carlsbad, CA, USA). Non-quantitative amplification of target genes was performed using Platinum® Taq DNA Polymerase (Life Technologies) to first confirm the correct identity and singularity of amplicons. PCR products were visualized with ethidium bromide on 2% agarose gels (Figure S1), purified with DNA Clean & Concentrator™-5 (Zymo Research, Irvine, CA, USA) and then directly sequenced to confirm
amplicon identity (data not shown) (Macrogen, Rockville, MD, USA). RT-qPCR was subsequently performed on each target gene using KicqStart® SYBR® Green qPCR ReadyMix™ iQ (Sigma–Aldrich, St. Louis, MO, USA) and an iCycler iQ™ Real-Time PCR Detection System (Bio-Rad, Hercules, California, USA). All Ct values (Table S2) were calculated by Bio-Rad IQ5 Optical System Software (Bio-Rad, Hercules, California, USA). Reactions were performed in technical triplicate 20 µL volumes. Three-step cycles plus melt curves were used for each reaction, using an annealing temperature of 56 °C for all primers. Efficiencies for each primer set were calculated from the slope of the standard curve using the formula

\[ E = 10^{(-1/\text{slope})} \]

(Pfaffl, 2001; Rasmussen, 2001). Primer efficiencies are based on three 1:10 serial dilutions of cDNA template used in side-by-side technical triplicate reactions. Efficiencies are listed in Table S1.

Relative gene quantification was calculated as

\[ E_{\text{target}} = \frac{C_{\text{target}}}{C_{\text{-reference}}} \]

for each target gene (14 total) and averaged for each replicate, both biological and technical. Ae. aegypti housekeeping gene lysosomal Aspartic Protease (vectorbase ID: AAEI006169) was used to normalize Ct values between biological replicates and to root the Y-axis scale in comparisons of relative expression.

The relationship between the RNA-seq and RT-qPCR relative gene expression data was examined using the least-squares linear regression model, RNA-seq = \( b_0 + b_1 \text{qPCR} \), to data observed on the target genes, for each gender and tissue; and by testing the hypothesis of statistical equivalence of these lines with the identity line of unity slope and zero intercept. As most of the 14 target genes exhibited relatively low abundance, close to reference value 1, the observed abundance (\( \sim 10 \)) of AaegOBP11 was excluded from the statistical analyses. Including AaegOBP11 overly influenced the line estimating the RNA-seq and RT-qPCR relationship, to primarily connect the AaegOBP11 abundance to the group of closely aggregated genes near reference value 1. Excluding AaegOBP11 allowed the RNA-seq and RT-qPCR relationship to be estimated using genes with similarly low abundance values, each having similar influence on the regression line fit. Coefficients of determination (\( r^2 \)) and 95% confidence intervals for intercept (\( b_0 \)) and slope (\( b_1 \)) (Table 1) are fit statistics. \( r^2 \) indicates the proportion of the total variability in the observed data that is explained by the regression model and \( 1 - r^2 \), indicating how widely the gene abundance values are scattered around the model. Confidence intervals for intercept and slope were used to determine that the RNA-seq vs. RT-qPCR relationship was “not statistically different” from the line of zero intercept and unit slope. To test the hypothesis that RNA-seq and RT-qPCR abundance values were “statistically equivalent,” a more statistically rigorous test was needed. To assist in assessing how similar the observed RNA-seq vs. RT-qPCR was to the identity line; the regression-based, bootstrap equivalence procedure (Robinson et al., 2005) was iteratively applied, to the data from each gender and tissue, to identify the smallest “region of similarity” that can be constructed around the observed RNA-seq and RT-qPCR data, and allow the relative gene expression as measured by RT-qPCR to be considered statistically equivalent to the measurement of relative gene expression by RNA-seq. To assist with visual assessment of the RNA-seq vs. RT-qPCR relationship, graphs were produced containing the observed data points, fitted regression line, identity line and 95% region of slope similarity. Linear regression analyses and hypothesis tests for statistical equivalence were conducted using the lm function of the stats package in the R computing environment (www.r-project.org).

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Least-squares regression line RNA-seq = ( b_0 + b_1 \text{qPCR} )</th>
<th>( r^2 )</th>
<th>( b_0 )</th>
<th>( b_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male labellum</td>
<td>0.00335 (−0.037, 0.066)</td>
<td>1.040 (0.99, 1.44)</td>
<td>0.983</td>
<td></td>
</tr>
<tr>
<td>Male prothoracic tarsus</td>
<td>−0.00132 (−0.129, 0.107)</td>
<td>1.189 (0.58, 1.90)</td>
<td>0.857</td>
<td></td>
</tr>
<tr>
<td>Male mesothoracic tarsus</td>
<td>0.02234 (−0.102, 0.147)</td>
<td>1.167 (0.73, 2.35)</td>
<td>0.806</td>
<td></td>
</tr>
<tr>
<td>Male metathoracic tarsus</td>
<td>0.00687 (−0.118, 0.800)</td>
<td>1.157 (0.42, 1.74)</td>
<td>0.913</td>
<td></td>
</tr>
<tr>
<td>Female labellum</td>
<td>−0.03121 (−0.059, 0.030)</td>
<td>0.953 (0.58, 1.01)</td>
<td>0.961</td>
<td></td>
</tr>
<tr>
<td>Female prothoracic tarsus</td>
<td>−0.00486 (−0.106, 0.073)</td>
<td>1.081 (0.51, 1.54)</td>
<td>0.923</td>
<td></td>
</tr>
<tr>
<td>Female mesothoracic tarsus</td>
<td>−0.04290 (−0.131, 0.047)</td>
<td>1.031 (0.30, 1.22)</td>
<td>0.951</td>
<td></td>
</tr>
<tr>
<td>Female metathoracic tarsus</td>
<td>−0.01106 (−0.059, 0.030)</td>
<td>1.003 (0.20, 1.47)</td>
<td>0.953</td>
<td></td>
</tr>
</tbody>
</table>

Pair values in parentheses besides Intercept (\( b_0 \)) and Slope (\( b_1 \)) estimates are 95% confidence intervals. \( r^2 \) (coefficient of determination) is the square of the correlation of RT-qPCR estimation of relative gene abundance and that of RNA-seq, with 1 showing perfect correlation and 0 showing no correlation.

### 3. Results

#### 3.1. RNA-seq analyses

Illumina based RNA-seq analysis generates sequences or “reads” of 50 bp length. Each of our analyses (labella and pro-, meso- and metathoracic tarsal segments of male or female mosquitoes) consistently yielded 18–23 million of these reads (Fig. 1). Between 8 and 12% of these reads mapped to the genome scaffold assembly for Ae. aegypti (AaegL1), of which ~50% were subsequently mapped to annotated genes (AaegL13) including many putative GRs (Fig. 1; Kent et al., 2008). Thus, we focused on 85 uniquely transcribed loci, representing 87 putative AaegGRs (Fig. 2). Each AaegGR gene was assigned an RPKM value, which represents the level of expression of a given gene (see Methods). RPKM values of 1–3 may also signify functional transcripts (see Section 4.4). RPKM values for AaegGR loci in these tissues ranged from 0 (undetectable) to 80; 68 (80%) of all loci displayed quantifiable read data (Fig. 2). Labella expressed 28% of the 87 GRs at RPKM values at or above 3. Tarsal tissues expressed fewer GRs with this expression level; only 8% of the 87 GRs had RPKM values at or above 3 in any tarsal sample. Expression of putative pseudogenic GRs (Kent et al., 2008) is not reported, as RPKM values higher than 1 for pseudogenic transcripts were not observed.

Relatively high AaegGR sequence conservation was not predictive of higher expression in labella and tarsi in general; however, highly expressing AaegGRs with likely orthologs in Anopheles gambiae almost always demonstrated conservation in Drosophila melanogaster as well. Twelve out the total 40 AaegGRs with likely orthologs in An. gambiae had RPKM values greater than 3 in labella or tarsal tissues (Fig. 2): AaegGR3, AaegGR4, AaegGR5, AaegGR6, AaegGR7, AaegGR9, AaegGR10, AaegGR11, AaegGR14, AaegGR19c, AaegGR20d and AaegGR34. Eleven of these 12 GRs showed strong to moderate conservation in D. melanogaster (Kent et al., 2008), the exception being AaegGR20d. Similarly, 13 of the remaining 47 AaegGRs that lacked clear homologous relationships in An. gambiae showed RPKM values greater than 3 (Fig. 2): AaegGR15, AaegGR16, AaegGR17, AaegGR18, AaegGR36, AaegGR49, AaegGR55, AaegGR60, AaegGR65, AaegGR66, AaegGR72, AaegGR76 and AaegGR79.

Among the AaegGRs with the highest RPKM values were AaegGR4, AaegGR5, AaegGR6, AaegGR7, AaegGR9, AaegGR10 and AaegGR11. These 7 genes are closely related to D. melanogaster sugar receptors (DmelGR5a, Dmel61a and DmelGR64a-f; Kent and Robertson, 2009) and represented more than 50% of all GR loci reads for all tissues (Fig. 1). RPKM values of each of these GRs
ranged from less than 2 to over 80 in labella, and relatively highly expressing GRs in labella also expressed relatively highly in tarsal tissues (Fig. 2). Other AaeGGRs expressing at greater than 3 RPKM with functionally characterized orthologs in D. melanogaster included AaegGR3 (DmelGR63a, required for volatile CO2 detection), AaegGR14 (DmelGR66a, required for normal caffeine and DEET detection), AaegGR19c (DmelGR28b, associated with light detection and thermosensation), and AaegGR34 (DmelGR43a, a fructose receptor in GRNs and neurons of the CNS).

3.2. RT-qPCR validation of RNA-seq

We next used RT-qPCR amplification of labellar and tarsal cDNA samples to validate expression trends of a small set of genes expressing GRs in labella also expressed relatively highly in tarsal tissues (Fig. 2). Other AaeGGRs expressing at greater than 3 RPKM with functionally characterized orthologs in D. melanogaster included AaeegGR3 (DmelGR63a, required for volatile CO2 detection), AaeegGR14 (DmelGR66a, required for normal caffeine and DEET detection), AaeegGR19c (DmelGR28b, associated with light detection and thermosensation), and AaeegGR34 (DmelGR43a, a fructose receptor in GRNs and neurons of the CNS).

Table 1; black lines, Fig. 4) indicate how closely qPCR gene expression estimates agree with actual RNA-seq read counts in each tissue, with perfect equivalence represented by lines having slope = 1 and intercept = 0 (red lines, Fig. 4). Fig. 4 illustrates the variability (r²; Table 1) of the observed abundance values measured by RNA-seq and RT-qPCR, relative to both the least-squares (black) line and the identity (red) line. The 95% confidence intervals for intercept and slope (Table 1), contain 0 (intercept) and 1 (slope) for all tissues observed for both genders; indicating there was not evidence that any of the observed RNA-seq vs. RT-qPCR relationships were statistically different from being identical.

Increasing the statistical rigor, and testing the hypothesis that the observed RNA-seq vs. RT-qPCR relationships were statistically equivalent to the identity relationship; the area between the dashed lines (Fig. 4), biessed by the (red) identity line, illustrates the smallest "region of similarity" within which the slope of the RNA-seq vs. RT-qPCR regression line can vary and still be considered statistically equivalent to slope 1. Instead of specifying the region of similarity (i.e., slopes of the dashed lines) and then testing the hypothesis that the slope of the least-squares regression line was statistically equivalent to the identity line, we used the equivalence algorithm (iteratively) to calculate the smallest region of similarity of slope for which it could be confidently concluded that the observed RNA-seq and RT-qPCR values exhibit a relationship that was statistically equivalent to the identity. The slopes of the regions of similarity boundaries (dashed lines, Fig. 4) deviate much too widely from unity slope to be practically considered equivalent to unity; thus, our equivalence analyses detected the fundamental differences between the RNA-seq and RT-qPCR methods. However, the fit statistics for the least-squares regression lines (Table 1) indicate the relationships between the RNA-seq and RT-qPCR values observed in this study were statistically strong.
Fig. 2. RPKM values for putative Gustatory Receptors in Ae. aegypti. Top: Dissected tissues from males (left) and females (right) are indicated by blue shading. Table: Cells report RPKM values for each Gustatory Receptor annotation or gene cluster. Numerical values are rounded to the nearest tenth. Heat-map color intensity is capped at RPKM value 10 in order to better distinguish values between 1 and 10. All values greater than 10 display maximum color intensity.
enough to provide supporting and complementary information beneficial to making conclusions about our study objectives. Quantitative estimation of transcript abundance by RT-qPCR of 12 chemoreception genes and the housekeeping gene *AaegLAsP* support the RNA-seq read quantification data most strongly for all female tissue and male labella, and less strongly for male tarsi.

4. Discussion

Gustatory receptors provide mosquitoes with the capability to detect chemical signals important for feeding, oviposition and conspecific recognition. Moreover, the feeding deterrent DEET both stimulates *Ae. aegypti* labellar GSNs (Sanford et al., 2013) and
Fig. 4. Statistical equivalence plots of RNA-seq and RT-qPCR datasets. The vertical axes represent relative gene abundance as demonstrated by RNA-seq data, with a value of 1 being attributed to housekeeping gene AaegLAsP and all other gene expression being reported relative to this value. The horizontal axes represent relative gene abundance as predicted by RT-qPCR, with a value of 1 being attributed to housekeeping gene AaegLAsP and all other gene expression being reported relative to this value. Red lines represent exact equivalence. Dashed lines represent boundaries of the "region of similarity" for slope, calculated by the statistical equivalence algorithm (Robinson et al., 2005). Solid black lines represent the least-squares regression of RT-qPCR equivalence to RNA-seq quantification for the 12 target genes and AaegLAsP housekeeping gene, each represented by circle data point.
evokes avoidance behavior in anosmic Ae. aegypti (DeGennaro et al., 2013). Here we discuss specific AaegGRs likely involved in these and other behaviors and their expression profiles in the primary gustatory appendages.

Using RNA-seq, we detected 41 AaegGRs expressing in labella and 19 AaegGRs expressing in at least one tarsal sample, 16 of which were common to both tissue types (Venn diagram, Fig. 1; Fig. 2). We then validated this sequencing using RT-qPCR to estimate relative expression levels of a small set of chemoreception genes (Fig. 3). A statistical comparison of RNA-seq and RT-qPCR data for Ae. aegypti chemoreception genes demonstrated a similar trend (Table 1) between the data collected by each method, but not exact equivalence (Fig. 4). As previously stated, the two techniques are limited by different chemistries, and thus did not produce identical data. However, relative expression trends of all genes tested are broadly reflected in side-by-side comparisons (Fig. 3).

4.1. Expression of conserved AaegGRs

AaegGR1, AaegGR2 and AaegGR3 represent the most highly conserved GRs in insects (Jones et al., 2007; Kent et al., 2008; Robertson and Kent, 2009). AaegGR1 and AaegGR3 are each required for normal CO2 activation of sensory neurons on the maxillary palps (Erdelyan et al., 2012), a function conserved in D. melanogaster (orthologs DmelGR21a and DmelGR63a, Jones et al., 2007). A function has not been attributed to the AaegGR1 paralog AaegGR2, although AaegGR2 is also expressed in the maxillary palp and well conserved (Kent et al., 2008; Kent and Robertson, 2009). Interestingly, while significant expression of AaegGR3 (RPKM = 13.2 female, 9.7 male) occurs in the labella, AaegGR1 (RPKM = 0.2 female, 0.0 male) and AaegGR2 (RPKM = 0.2 female, 0.4 male) expression are barely detectable in this tissue (Figs. 2 and 3). AaegGR1 and AaegGR3 are thought to function as hetero-dimeric sensors of volatile CO2, however no function has been attributed to AaegGR3 singularly or in combination with GRs other than AaegGR1. Proboscis ablation does not appear to significantly affect CO2 initiated host-seeking behavior in An. stephensi (Maekawa et al., 2011); therefore, it is unlikely that AaegGR3 expression in the labellum is related to CO2 detection. It is possible AaegGR3 is either non-functional though transcribed in this context or has a yet to be described role.

All seven of the putative sugar receptors identified by sequence homology (Kent et al., 2008) are expressed in the labella of both sexes (Figs. 2 and 3). This expression is consistent with sugar GRs from D. melanogaster (Clyne et al., 2000; Scott et al., 2001; Dunipace et al., 2001 Clyb et al., 2003; Dahanukar et al., 2007). The most highly expressing GRs within this clade in labella (AaegGR4, AaegGR7 and AaegGR9) also express highly in tarsi relative to all other putative AaegGRs. Since insect GRs may function as hetero-dimers or multimers (Jiao et al., 2008; Moon et al., 2009; Lee et al., 2009; Weiss et al., 2011) with particular GRs serving as ubiquitous co-receptors to subsets of other GRs with more narrowly tuned chemical sensitivities, these three AaegGRs are the most salient candidate sugar co-receptors. Sugar perception in mosquitoes is crucial for the identification of suitable nutritive sources for feeding (Galun and Fraenkel, 1957; Salama, 1967; Briegel and Kaiser, 1973; Clements, 1992). Sugar stimulation of tarsi induces labella probing behaviors in both males and females leading to active sugar feeding (Frings and Hamrum, 1950; Feir et al., 1961; Pappas and Larsen, 1978; Clements, 1992). Both physiological responses of sugar sensitive hairs on the tarsi and labella, and feeding behavior are well documented (Salama, 1966; Pappas and Larsen, 1976; Angioy et al., 1982). Our data confirm the expression of all conserved sugar receptors in the labella and demonstrate expression of at least some of them in legs. This strongly suggests that the AaegGRs of this group are required for normal physiological and behavioral responses of mosquitoes to sugar stimuli.

An even more highly conserved sugar sensitive GR, DmelGR43a, is required for labellar and tarsal GRN responses to fructose as well as normal feeding behavior involving sensing of hemolymph fructose levels in the CNS in D. melanogaster (Miyamoto et al., 2012). Interestingly, AaegGR34, the likely ortholog of DmelGR43a, shows about 4-fold enrichment in female labella with respect to male labella (Fig. 2). If this GR serves as a fructose receptor in mosquitoes, this differential expression between the sexes may reflect a sensory adaptation to fructose in females.

AaegGR14 is the most abundant AaegGR transcript in the labella of both sexes in our survey and is the only putative bitter receptor based on sequence homology (Kent et al., 2008). Unlike its D. melanogaster ortholog DmelGR66a, AaegGR14 is not expressed in tarsi (Fig. 2), noting that all leg expression data for DmelGR66a to date is promoter-based and not directly determined from localized RNA (Scott et al., 2001; Dunipace et al., 2001; Bray and Amrein, 2003; Moon et al., 2006). DmelGR66a is required (along with DmelGR32a and DmelGR33a) but not sufficient for activation of D. melanogaster GRNs and aversive feeding response elicited by the deafferent DEET (Lee et al., 2010). DmelGR66a is also involved in the detection of caffeine, a bitter stimulus, and is expressed in every bitter sensitive GRN in the labellum (Moon et al., 2006, 2009). Both the orthology of AaegGR14 with DmelGR66a and its relatively high expression suggest it may be involved in the detection of aversive substances in Ae. aegypti. Indeed, a recent electrophysiological study identified a GRN housed within sensilla on the labella of Ae. aegypti that responded to bitter quinine as well as DEET and other known insect repellents (Sanford et al., 2013). Subsequently, avoidance responses specifically associated with contact chemoreception of DEET were observed in behavioral studies of transgenic Ae. aegypti in which the olfactory co-receptor Orco was mutated (DeGennaro et al., 2013). These behaviors were separate from avoidance of volatile DEET cues, a caveat only indirectly tested in prior feeding assays involving insect repellents (Bar-Zeev and Schmidt, 1959; Kuhn et al., 2006).

Significant expression (RPKM range 6–59) of AaegGR19c, the third most abundant transcript in our survey, is evident in the labella and tarsi of both sexes (Fig. 2). Alternative splice forms AaegGR19a and AaegGR19b show much lower RPKM values in legs and labella (RPKM range 0–1.7). The alternatively spliced AaegGR19 locus is orthologous to a similarly spliced locus in D. melanogaster, DmelGR28b (Kent et al., 2008). Three of the five splice variants of the DmelGR28b(A-E) locus show atypical GR expression patterns in addition to commonly observed peripheral sensory neuron expression. Promoter-driven reporter and transcript expression is evident for these three genes in Johnston’s organ, the base of the aristae, campaniform sensilla of the wing, neurons within the CNS, oenocytes and neurons of the abdominal wall and gut (Thorne and Amrein, 2008; Park and Kwon, 2011). In addition, the DmelGR28b locus is required for light-induced responses of class IV dendritic arborization neurons in larvae (Xiang et al., 2010) and the DmelGR28b-D transcript confers thermosensitivity in the antennae and aristae of D. melanogaster (Ni et al., 2013). DmelGR28b-E is the most downstream of the genes in the DmelGR28 locus and is only expressed in labella and tarsi. Similarly, AaegGR19c is the most downstream of the genes in the AaegGR19 locus and is expressed in the labella and tarsi (Figs. 2 and 3). Whether these consistencies reflect conservation of expression pattern or function between the respective loci remains a topic for future investigations.
4.2. Expression of non-conserved AaegGRs

Thirteen Ae. aegypti-specific GRs showed expression values above 3 RPKM in any tissue (Fig. 2). Without strong homology to previously characterized GRs, we are unable to speculate as to their possible functions. A majority (33/38) of the GR promoters driving expression in D. melanogaster labella label bitter sensitive GSNs (Weiss et al., 2011); all exceptions being well-described sugar receptors labeling sugar sensitive GRNs. Furthermore, large expansions of bitter sensitive GRs have been proposed for the silkmoth Bombyx mori (Wanner and Robertson, 2008). Thus, it is possible that a majority of the Ae. aegypti-specific GRs expressing in labella are involved in the detection of bitter substances.

Unlike the other GRs in this group, AaegGR76 is only expressed in tarsi (Fig. 2); being the third most abundant GR in male prothoracic tarsi and the fourth most abundant GR in female prothoracic tarsi. AaegGR76 may be involved in a non-labellar gustatory modality. Furthermore, AaegGR76 RPKM values of 10.0 in male and 7.3 in female prothoracic tarsi contrast lower values in mesothoracic tarsi and zero values in metathoracic tarsi for both sexes (Fig. 2). These tarsal expression patterns may reflect sensory pathways unique to a single tarsal type, thus highlighting differences in chemoreception function between tarsal types.

4.3. Sex-specific AaegGR expression

Sex-specific expression biases greater than 2-fold among AaegGRs in the labella and tarsi tissues are numerous (Fig. 2); male-enriched: AaegGR20e, AaegGR27, AaegGR36, AaegGR39e, AaegGR43; female-enriched: AaegGR5, AaegGR22 AaegGR34, AaegGR44, AaegGR56. These differences in expression may indicate involvement of these GRs in sex-specific recognition of social cues or in the case of females, blood-host attractants. Identifying AaegGRs that elicit neuronal responses to non-volatile sex-specific cuticular hydrocarbons could provide a basis for uncovering attractive or repulsive neuronal circuitry in the tarsi and labella of Ae. aegypti. Likewise, AaegGRs that elicit neuronal responses to non-volatile blood-host cues could provide clues to neuronal circuitry involved in the mosquito’s choice to bite. Female, but not male, Ae. aegypti take host blood-meal to supplement egg production (Clements, 1992); therefore, it is possible that female-enriched AaegGR5, AaegGR22, AaegGR44 or AaegGR56 are involved in the reception of blood-host cues.

The relatively balanced expression between males and females of all other AaegGRs does not exclude them from mediating sex-specific behaviors, as differences in peripheral GRN circuitry rather than GRN receptor repertoire are sometimes more predictive of sex-linked behavior in insects (Villella and Hall, 2008; Lu et al., 2012; Thistle et al., 2012). The neuronal signals elicited by AaegGRs expressing equally in both sexes may be differentially interpreted by downstream circuitry that is sex-specific.

4.4. Challenges of sequencing AaegGRs

Surprisingly, only 8–12% of total RNA-seq reads mapped to the genome scaffold assembly of Ae. aegypti for labella and tarsi samples (Fig. 1). This proportion was much lower than those observed in prior RNA-sequencing of mosquito chemosensory tissue as between 90 and 94% reads mapped to the genome assembly for An. gambiae for antennal preparations (Pitts et al., 2011); however, it was closer to proportions reported for Ae. aegypti whole-body (less than 40%; Bonizzoni et al., 2011) and larva (47%; Paris et al., 2012). Even so, a more recent developmental study reports between 80 and 95% read mapping to the Aedes genome, depending on the tissue type (Akbari et al., 2013). Global BLASTs (blastn) were performed on the four most redundant, non-mapped 50 basepair reads to test for possible contamination in all eight tissue samples (data not shown) (pers. comm. Nripesh Prasad, Hudson-Alpha Institute for Biotechnology); for all tests, reads produced high similarity hits with insect genes, the most similar genes being from mosquito species. Therefore, it is unlikely that the unmapped reads correspond to contaminant nucleotides. Rather, it is likely the majority of raw reads for labella and tarsi correspond to actively transcribing transposable elements (~50% of Ae. aegypti genome; Nene et al., 2007) that are not well represented in the genome assembly.

A reasonable distribution of genomic locations was observed for reads that did map to the Ae. aegypti genome (Fig. 1). At least half of these reads mapped to coding regions for each tissue sample, suggesting we have obtained an accurate cross-section of expressed genes in all tissues. Non-mapped reads displayed a similar GC content to that of mapped reads (45%). We also mapped all unmapped reads to the 16,665 basepair mitochondrial genome (Behura et al., 2011) to rule out an abundance of mitochondria in our dissected tissue types. Only 2% of our unmapped reads mapped to this genome for the female labella sample (~900,000 reads; data not shown), therefore this is not a consideration. To date there are no other RNA-seq datasets for insect tarsi or labella, thus direct comparisons of read mapping percentages for these tissue types are not possible. A few instances of low genome mapping percentages have been reported for insect RNA-seq datasets, but these do not represent a majority of published data (14.4% for late stage D. melanogaster embryos, Daines et al., 2011).

In another study by our group, RNA-seq analysis of the maxillary palps of female Ae. aegypti of the same age and strain yielded raw reads of which ~80% mapped to the genome assembly (data not shown). These reads were mapped to the same genome assembly file using the same software parameters presented here. The starting RNA quality was identical as determined by Bioanalyzer (data not shown) and average read quality was actually lower than that of reads used in this study (Table S3). Therefore, our relatively low mapping percentage may reflect a unique transcriptional feature of labella and tarsi tissues in Ae. aegypti.

Low GR transcript levels observed in insects have made identification of specific GR genes challenging without the use of promoter-driven reporters (Isono and Morita, 2010). RNA-seq is sensitive enough to identify these rare transcripts, even as their expression levels approach that of background “noise” (Wang et al., 2009). In the labella and tarsi, the primary gustatory organs of adult Ae. aegypti, 28% (24/85) of the putative AaegGR loci displayed RPKM values higher than 3, a level proposed to represent expression well above background levels associated with non-functional or otherwise leaky genomic regions (Ramskold et al., 2009). Lower RPKM values of 1–3 account for another 22% (19/85) of AaegGRs in these tissues. The remaining 50% of putative AaegGRs are below the threshold of 1 RPKM, which approaches levels indistinguishable from background or non-significant expression. Nevertheless, we cannot rule out the possibility that these transcripts, if truly expressed, may produce functional GRs in the labella or legs of the mosquito, albeit in as few as one cell or in a very low copy number. This is less of a concern in labella tissue, as there are about 30 large gustatory sensilla tightly distributed over a majority of the outer surface, each innervated with 3–5 GRNs (Chaika and Elizarov, 1971; Lee and Craig, 2009). GRNs within sensilla on the tarsi of mosquito species. Therefore, it is unlikely that the unmapped reads correspond to contaminant nucleotides. Rather, it is likely the majority of raw reads for labella and tarsi correspond to actively transcribing transposable elements (~50% of Ae. aegypti genome; Nene et al., 2007) that are not well represented in the genome assembly.

A reasonable distribution of genomic locations was observed for reads that did map to the Ae. aegypti genome (Fig. 1). At least half of these reads mapped to coding regions for each tissue sample, suggesting we have obtained an accurate cross-section of expressed genes in all tissues. Non-mapped reads displayed a similar GC content to that of mapped reads (45%). We also mapped all unmapped reads to the 16,665 basepair mitochondrial genome (Behura et al., 2011) to rule out an abundance of mitochondria in our dissected tissue types. Only 2% of our unmapped reads mapped to this genome for the female labella sample (~900,000 reads; data not shown), therefore this is not a consideration. To date there are no other RNA-seq datasets for insect tarsi or labella, thus direct comparisons of read mapping percentages for these tissue types are not possible. A few instances of low genome mapping percentages have been reported for insect RNA-seq datasets, but these do not represent a majority of published data (14.4% for late stage D. melanogaster embryos, Daines et al., 2011).

In another study by our group, RNA-seq analysis of the maxillary palps of female Ae. aegypti of the same age and strain yielded raw reads of which ~80% mapped to the genome assembly (data not shown). These reads were mapped to the same genome assembly file using the same software parameters presented here. The starting RNA quality was identical as determined by Bioanalyzer (data not shown) and average read quality was actually lower than that of reads used in this study (Table S3). Therefore, our relatively low mapping percentage may reflect a unique transcriptional feature of labella and tarsi tissues in Ae. aegypti.

Low GR transcript levels observed in insects have made identification of specific GR genes challenging without the use of promoter-driven reporters (Isono and Morita, 2010). RNA-seq is sensitive enough to identify these rare transcripts, even as their expression levels approach that of background “noise” (Wang et al., 2009). In the labella and tarsi, the primary gustatory organs of adult Ae. aegypti, 28% (24/85) of the putative AaegGR loci displayed RPKM values higher than 3, a level proposed to represent expression well above background levels associated with non-functional or otherwise leaky genomic regions (Ramskold et al., 2009). Lower RPKM values of 1–3 account for another 22% (19/85) of AaegGRs in these tissues. The remaining 50% of putative AaegGRs are below the threshold of 1 RPKM, which approaches levels indistinguishable from background or non-significant expression. Nevertheless, we cannot rule out the possibility that these transcripts, if truly expressed, may produce functional GRs in the labella or legs of the mosquito, albeit in as few as one cell or in a very low copy number. This is less of a concern in labella tissue, as there are about 30 large gustatory sensilla tightly distributed over a majority of the outer surface, each innervated with 3–5 GRNs (Chaika and Elizarov, 1971; Lee and Craig, 2009). GRNs within sensilla on the tarsi of mosquito species. Therefore, it is unlikely that the unmapped reads correspond to contaminant nucleotides. Rather, it is likely the majority of raw reads for labella and tarsi correspond to actively transcribing transposable elements (~50% of Ae. aegypti genome; Nene et al., 2007) that are not well represented in the genome assembly.
expressed. Moreover, taking D. melanogaster tarsal GRNs as a model (Dunipace et al., 2001; Scott et al., 2001), only a small subset of these neurons may be expressing a given GR. Thus, it may be necessary to target more specifically the gustatory organs of tarsi for RNA extraction in order to better resolve the expression of AaeGGRs.

While our survey reveals expression of at least half of the putative AaeGGRs in labella and tarsi, a subset of the remaining AaeGGRs likely express in other tissues including the labrum, cibarial organ and perhaps elsewhere (Lee, 1974; Clements, 1992). Indeed, AaeGGR expression has been reported in the maxillary palp (Erdelyan et al., 2012; Bobbot et al., 2013), and in An. gambiae, GR expression is evident in both the antennae and maxillary palps (Pitts et al., 2011). In D. melanogaster, GRs express in neuronal tissues of the antennae, Johnston’s organ, wings, genitalia, pharynx and CNS (Clane et al., 2000; Dunipace et al., 2001; Scott et al., 2001; Ikeda et al., 2002; Thorne and Amrein, 2008; Ejima and Griffith, 2008). Promoter-based expression studies also indicate DmelGR expression in digestive and somatosensory tissues (Thorne and Amrein, 2008; Park and Kwon, 2011). Several of these DmelGRs are required for physiology and behavior associated with these tissues. Additionally, some mosquito GRs may be expressed in larval or pupal stages and have specific functions associated with an aquatic environment. Future GR expression profiling in A. aegypti should extend to these tissues, in addition to those surveyed here.

4.5. Future considerations

AaeGGRs expressed in labella and tarsal tissues mediate discrimination of many contact chemical stimuli, and those well-conserved GRs may exhibit predictable sensitivities, some of which are directly linked to host-selection or assessment of sugar sources. Most AaeGGRs are divergent however, and greater resolution as to their expression pattern is needed to better understand their roles. The AaeGGRs identified here are potential guideposts to resolve AaeGGR co-expression in single sensory neurons by double labeling with RNA probes or promoter-driven reporters. The AaeGGRs showing highest expression levels in our survey, like those related to sugar GRs, may be tested first to assess the feasibility of single cell labeling in labella. Such an expression map would help determine the combinatorial code for AaeGGR function in GRNs with sugar or bitter sensitivities. AaeGGRs also provide suitable subjects for future in vivo gene disruption or ex vivo heterologous expression studies, and are targets for the disruption or modification of gustatory-mediated behavior in mosquitoes. Several techniques have recently been used to successfully alter target genes in Ae. aegypti (Aryan et al., 2013a, 2013b; DeGennaro et al., 2013), thus opening the door to comprehensively studying the function of individual AaeGGRs.

Acknowledgments

The authors thank Dr. Shawn E. Levy and Nripesh Prasad of the Genomics Services Lab at the Hudson-Alpha Institute for Biotechnology for RNA sequencing and data analyses; Dr. Hugh Robertson, Department of Entomology, University of Illinois at Urbana-Champaign for Ae. aegypti putative GR sequences and Dr. Jonathan D. Bobbot of our group for the mosquito illustration in Fig. 2 and graphical abstract. We are grateful to Dr. Rick Jones USDA, ARS, GIPFL, Beltsville, MD for advice and Drs. Richard Vogt, Department of Biological Sciences, University of South Carolina and Daniel Strickman for critical review of manuscript. We also thank Kevin Nyberg, Department of Biology, University of Maryland for assistance with bioinformatic screens. This work was supported in part by a grant to J.C.D. from the Deployed War Fighter Protection (DWFP) Research Program funded by the Department of Defense through the Armed Forces Pest Management Board (AFPMB).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ibmb.2013.10.005.

References


Bonizzoni, M., Dunn, W.A., Campbell, C.L., Olson, K.E., Dimon, M.T., Marinotti, O., James, A.A., 2011. RNA-Seq analyses of blood-induced changes in gene expression in the mosquito vector species, Aedes aegypti, BMC Genomics 12, 82.


strong preference for humans and are not repelled by volatile DEET. J. Nat. Lond. 498, 487–491.
Dunlap, J.W., Steiner, S., McNealy, C., Amrein, H., 2001. Spatially restricted expres-
Ejima, A., Griffith, L.C., 2008. Courtship initiation is stimulated by acoustic signals in
the carbon dioxide receptor genes in Aedes aegypti mosquitos using RNA interference.
Insect Mol. Biol. 21, 119–127.
Feir, D., Lengy, J.I., Owen, W.B., 1961. Contact chemoreception in the mosquito
Culissa inornata: sensitivity of the tarsi and labelia to sucrose and glucose.
J. Insect Physiol. 6, 13–20.
Fischer, W., Gong, P., Marella, S., Scott, K., 2007. The detection of carbonation by the
Frings, H., Hamrump, C.L., 1950. The contact chemoreceptors of adult yellow fever
Galun, R., Frakenel, G., 1957. Physiological effects of carbohydrates in the nutrition of a
Hall, S.R., Smith, J.J.B., 1999. Consistent pattern in the placement of taste sensilla on
Jiao, Y., Moon, S.J., Wang, X., Ren, Q., Montell, C., 2008. Gr64f is required in com-
Lee, Y., Kim, S.H., Montell, C., 2010. Avoiding DEET through insect gustatory re-
Lee, Y., Moon, S.J., Montell, C., 2009. Multiple gustatory receptors required for the
Lin, H., Mann, K.J., Starostina, E., Kinser, R.D., Pikielny, C.W., 2005. A
Lin, H., Mann, K.J., Starostina, E., Kinser, R.D., Pikielny, C.W., 2005. A
A fine structure of the mouthparts of the malaria vector
Luo, B., LaMora, A., Sun, Y., Welsh, M.J., 2005. A gustatory receptor paralogue controls rapid
Luo, B., LaMora, A., Sun, Y., Welsh, M.J., 2005. A gustatory receptor paralogue controls rapid
Luo, B., LaMora, A., Sun, Y., Welsh, M.J., 2005. A gustatory receptor paralogue controls rapid
Luo, B., LaMora, A., Sun, Y., Welsh, M.J., 2005. A gustatory receptor paralogue controls rapid