Transcriptome profiling of chemosensory appendages in the malaria vector Anopheles gambiae reveals tissue- and sex-specific signatures of odor coding

R Jason Pitts1†, David C Rinker2†, Patrick L Jones1†, Antonis Rokas1,2 and Laurence J Zwiebel1,2*

Abstract

Background: Chemosensory signal transduction guides the behavior of many insects, including Anopheles gambiae, the major vector for human malaria in sub-Saharan Africa. To better understand the molecular basis of mosquito chemosensation we have used whole transcriptome RNA sequencing (RNA-seq) to compare transcript expression profiles between the two major chemosensory tissues, the antennae and maxillary palps, of adult female and male An. gambiae.

Results: We compared chemosensory tissue transcriptomes to whole body transcriptomes of each sex to identify chemosensory enhanced genes. In the six data sets analyzed, we detected expression of nearly all known chemosensory genes and found them to be highly enriched in both olfactory tissues of males and females. While the maxillary palps of both sexes demonstrated strict chemosensory gene expression overlap, we observed acute differences in sensory specialization between male and female antennae. The relatively high expression levels of chemosensory genes in the female antennae reveal its role as an organ predominately assigned to chemosensation. Remarkably, the expression of these genes was highly conserved in the male antennae, but at much lower relative levels. Alternatively, consistent with a role in mating, the male antennae displayed significant enhancement of genes involved in audition, while the female enhancement of these genes was observed, but to a lesser degree.

Conclusions: These findings suggest that the chemoreceptive spectrum, as defined by gene expression profiles, is largely similar in female and male An. gambiae. However, assuming sensory receptor expression levels are correlated with sensitivity in each case, we posit that male and female antennae are perceptive to the same stimuli, but possess inverse receptive prioritizations and sensitivities. Here we have demonstrated the use of RNA-seq to characterize the sensory specializations of an important disease vector and grounded future studies investigating chemosensory processes.
cellular basis of insect chemosensation has been studied for many years, its molecular underpinnings have only recently begun to be elucidated.

In mosquitoes, host-seeking behavior is driven largely by olfaction [8,9]. An. gambiae females display a strong preference for human hosts (anthrophily), which contributes substantially to their ability to transmit human diseases, including malaria [8-10]. Numerous studies have shown that the antennae of An. gambiae are the principle chemosensory organs that respond to volatile odors [8,9]. The maxillary palps of An. gambiae respond to carbon dioxide, one of the major activators of mosquito upwind flight and a synergistic attractant when combined with other volatile odors [8,9,11]. The identification of chemoreceptor gene families in the An. gambiae genome [12,13] has facilitated the correlation of receptor expression with behavioral observations and physiological sensitivities [14-16]. Specific chemoreceptors expressed in antennal and palpal neurons of An. gambiae are sensitive to host odors, including volatile components produced from bacteria associated with human skin [17-19]. As a consequence, the function of select chemoreceptor genes in An. gambiae has been linked to semiochemicals that are integral to specific host seeking behaviors. Despite this progress, very little of the downstream signaling events and regulation of chemoreceptor function is known. Moreover, the potential chemosensory bases of sexually distinct behaviors in An. gambiae are poorly understood [8,20,21] and studies of male An. gambiae chemosensory biology are particularly lacking [20].

RNA-seq offers great potential to efficiently and comprehensively study gene expression in the chemosensory head appendages of An. gambiae and to provide insight into the molecular foundations of chemoreception. While several microarray-based studies have examined global transcript abundance in An. gambiae [22-29], none has focused exclusively on chemoreceptive tissues. Moreover, unlike microarrays and older methods, RNA-seq provides transcriptome-wide sequence coverage with unbiased, highly quantitative results [30] and greatly improved sensitivity [31,32]. To date, RNA-seq has been used to address several functional and evolutionary questions pertaining to mosquito biology [33-37].

Here we have utilized RNA-seq to quantify global abundance levels of poly-adenylated transcripts of An. gambiae whole adults, antennae and maxillary palps between sexes, at a life stage when females are known to host seek [8,9]. By mapping the generated short read sequences against the full set of annotated An. gambiae transcripts we have generated six tissue- and sex-specific transcriptome profiles (Table 1). As expected, gene families with well-established chemosensory function display antenna- or palp-enhanced expression, with antennae showing enhancement of a larger number of these genes. We also have identified numerous members of other gene families that are enhanced in either antennae or maxillary palps, such as biotransformation enzymes, transcription factors, transmembrane receptors, ion channels, transporters and proteases which are likely to function in chemosensory pathways. Our data also revealed an unanticipated level of sexual monomorphism with respect to the expression and distribution of known chemoreceptive functional classes in the antenna and the maxillary palp. Taken as a whole, this study greatly broadens our understanding of the molecular processes in peripheral sensory appendages, and establishes an agnostic, quantitative data set that can be built upon by future research.

**Methods**

**Mosquito rearing**

An. gambiae sensu stricto, which originated from Sua-koko, Liberia [38], was reared as described [39]. Briefly, mosquitoes were reared in an isolated chamber at 27°C and 75% relative humidity with a photoperiod of 12:12 (L:D). Larvae were reared at low densities to ensure large adult size. Pupae were hand collected and allowed to eclose in small cages. Almost all pupae eclosed on the day after collection. The few pupae that failed to eclose were removed, such that adults in any single cage were the same age. Adult females and males were kept together in the same cages for 4-6 days and were sugar-fed with 10% sucrose *ad libitum* until the time of tissue collection. Females were not bloodfed prior to tissue collection, nor were they selected based upon any specific response to external stimuli. As a consequence of the rearing protocol, mosquitoes were intermittently exposed to the odor of their human caretakers. The vast majority of females were assumed to be mated based on numerous previous studies of cage-reared An. gambiae [40,41] and our own experience. Moreover, nearly all females reared as described above in our laboratory will bloodfeed when presented with an anesthetized mouse, indicating that they are physiologically competent to host seek.

**RNA isolation and sequencing**

Approximately 1500 antennae or maxillary palps were hand dissected from randomly selected, age-matched cohorts of 4-6 d.o. adults (ZT10-12). Additionally, approximately 20 whole bodies of 4-6 d.o. adults were collected of each sex (ZT10-12). All collected tissues were immediately placed in RNA Later Ice (Ambion Corp.; Austin, TX) on ice prior to RNA extraction. Total RNA was isolated from each sample using RNeasy columns (Qiagen Inc.; Carlsbad, CA) according to the manufacturer’s protocol. mRNA isolation and cDNA
library preparation were carried out using the Illumina mRNA sequencing kit (Illumina Inc.; San Diego, CA). Libraries were sequenced using an Illumina Genome Analyzer II or HiSeq2000. A single biological replicate, representing a large sample size was used in the subsequent analysis.

AgOr and AgObp reannotations
Novel AgOrs were identified by tBLASTn searches (http://blast.ncbi.nlm.nih.gov/Blast.cgi; default parameters) using previously identified AGOR peptides as queries. Two new candidate AgOrs were identified and have been named AgOrs 80 and 81. Furthermore, AgOrs 12, 67, 78 and 79 have been purged from the AgOr family as apparent duplication errors in the original assembly (Table 2). Three new candidate AgObps (69, 70 and 71) were identified using similar tBLASTn searches and were added to the family based on two criteria: the candidate genes possessed motifs that exemplify the Obp family [42-45], each gene model encoded a unique transcript. Other genes resembling Obps were identified, but have not been included in the named members of the AgObp family. However we recognize the possibility that these genes may ultimately prove to be unique, or function as odor-carriers. These will be discussed in more detail below. Similarly, AgObps 16, 17, 24, 58, 59, 60, 61, and 65 were purged from the AgObp family as apparent duplication errors in assembly. All modifications to the AgOr and AgObp families have been submitted to VectorBase.

Data processing and expression profiling
Individual Illumina read files were mapped to the recently updated (Dec. 2010) version of the assembled An. gambiae genome, to the annotated An. gambiae transcripts (http://www.VectorBase.org). For mapping purposes, all transcript isoforms for a given gene were condensed under that gene’s AGAP designation. Prior to mapping, individual reads were quality checked and uniformly trimmed by 4 and 12 nucleotides on their 5’ and 3’ ends, respectively, to account for spurious adapter incorporation (5’-end) and for sequencing reaction degeneration (3’-end). Mapping was carried out using seqmap software, configured to allow for a maximum of three mismatches per read. Processed mapping data was then consolidated based upon AGAP number and the results summarized by rseq software. Expression level output by rseq was reported in terms of unique reads, total weighted reads, and transcript length. Total weighted reads and AGAP transcript lengths were used to calculate a normalized transcript abundance level in units of Reads Per Kilobase per Million reads mapped (RPKMs), for every AGAP in every tissue type [32].

PfamA categorization
Peptide sequences from AgamP3.6 conceptual peptides (n = 12669) were compared to the PfamA dataset [46], using the default e-value threshold of 1.0.

Comparison of tissue expression profiles
Statistical significance was assigned to each pairwise tissue comparison (antenna:bodies, palps:bodies, bodies: bodies) by setting up a Fisher’s Exact test, comparing the number of weighted, mapped reads for each gene to the total number of mapped reads for that tissue sample. The Agam3.6 transcript annotation contains 13319 unique, annotated transcripts and the statistical significance of the Fisher’s Test was evaluated against a Bonferroni corrected p-value of 3.9 × 10^{-6}.
Results and Discussion

RNA Sequencing and Gene Mapping

As a means of inferring gene expression in chemosensory appendages we have employed single-end short read (43bp) RNA-seq technology to characterize the relative abundances of poly-adenylated RNAs in antennae, maxillary palps and whole bodies of female and male adult mosquitoes. We established tissue-specific gene expression profiles for each of our six samples by mapping the read sequence files against the annotated *An. gambiae* transcriptome, using an approach that quantitated transcript abundance per gene and which accounted for all annotated transcripts per gene (see Materials and Methods). As our reference transcriptome, we used the AgamP3.6 version of the *An. gambiae* gene annotation, which contains 12669 protein-coding genes and 650 non-coding RNAs (http://www.VectorBase.org). For each of the tissue types assayed, we obtained an average of 30.5 million sequence reads per tissue type and mapped them to the *An. gambiae* transcriptome, nuclear and mitochondrial genomes (Table 1). Additional file 1 contains the complete RNA-seq data set described above, including the number of reads from each tissue sample that mapped to all 13319

| gene class       | PfamA         | PfamA description                           | # *An. gambiae* | enhanced >2x |
|------------------|---------------|---------------------------------------------|----------------|--------------|
| 7tm Receptor     | PF00001       | 7tm receptor (rhodopsin family)             | 84             | 28 20 18 14  |
| 7tm Receptor     | PF02949       | 7tm Odorant receptor (Or)                    | 78             | 56 31 3 3   |
| 7tm Receptor     | PF08395       | 7tm Chemosensory receptor (Gr)              | 52             | 1 4 3 4    |
| 7tm Receptor     | PF00002       | 7tm receptor (Secretin family)              | 11             | 2 1 0 2    |
| 7tm Receptor     | PF00003       | 7tm sweet-taste receptor of 3 GCPR          | 7              | 4 5 2 1    |
| lipophilic carrier| PF01395       | PBP/GOBP family                               | 62             | 18 17 6 4  |
| lipophilic carrier| PF00650       | CRAL/TRIO domain                              | 43             | 17 9 17 16  |
| lipophilic carrier| PF06585       | Haemolymph juvenile hormone bind. (JHBP)     | 24             | 10 5 15 9  |
| lipophilic carrier| PF00188       | Cysteine-rich secretory protein family       | 20             | 7 2 9 7    |
| lipophilic carrier| PF03392       | Insect pheromone-bind. family, A10/OS-D      | 7              | 2 2 4 1    |
| channel/transporter| PF01130     | CD36 family                                   | 14             | 5 1 7 5    |
| channel/transporter| PF07690       | Major Facilitator Superfamily                | 65             | 21 16 16 13 |
| channel/transporter| PF00083       | Sugar (and other) transporter                | 49             | 7 4 7 8    |
| channel/transporter| PF00060       | Ligand-gated ion channel                      | 29             | 22 20 5 3  |
| channel/transporter| PF00520       | Ion transport protein                         | 27             | 15 10 9 3  |
| channel/transporter| PF02931       | Neurotrans.-gated ion-channel ligand bind.   | 24             | 10 6 4 0    |
| channel/transporter| PF00858       | Aminioide-sensitive sodium channel            | 23             | 5 2 1 1    |
| channel/transporter| PF01061       | ABC-2 type transporter                        | 19             | 10 4 12 11 |
| channel/transporter| PF00005       | ABC transporter                               | 18             | 4 3 5 2    |
| channel/transporter| PF00664       | ABC transporter transmemb.                   | 15             | 4 2 2 4    |
| channel/transporter| PF07885       | Ion channel                                   | 9              | 3 3 1 1    |
| biotransformation| PF00067       | Cytochrome P450                               | 113            | 30 19 34 24 |
| biotransformation| PF00135       | Carboxylesterase                              | 50             | 15 13 14 14 |
| biotransformation| PF00043       | Glutathione S-transferase, C-term.            | 18             | 6 1 4 1    |
| biotransformation| PF02798       | Glutathione S-transferase, N-term.            | 17             | 5 3 4 3    |
| transcription factor| PF00095      | Zinc finger, C2H2 type                        | 114            | 21 50 21 24 |
| transcription factor| PF00046      | Homeobox domain                               | 76             | 17 19 14 13 |
| transcription factor| PF00651      | BTB/POZ domain                                | 54             | 17 26 5 7  |
| transcription factor| PF00010      | Helix-loop-helix DNA-binding                  | 41             | 6 6 5 6    |
| transcription factor| PF00250      | Fork head domain                              | 19             | 6 8 3 4    |
| transcription factor| PF00716      | Basic region leucine zipper                   | 14             | 3 4 1 3    |
| transcription factor| PF00292      | vPaired box\ domain                           | 10             | 3 5 3 3    |
| transcription factor| PF00078      | T-box                                        | 11             | 8 6 8 5    |
| transcription factor| PF00170      | bZIP transcription factor                     | 8              | 3 3 2 2    |
| transcription factor| PF00157      | Pou domain - N-terminal to homebox           | 4              | 2 3 3 1    |

Cells in each row contain information corresponding to the gene class listed. PfamA: PfamA family number. PfamA description: PfamA family description. # *An. gambiae*: number of genes identified in PfamA searches of *An. gambiae* transcriptome. enhanced > 2x: number of genes in each PfamA family that were enhanced relative to bodies in the specified tissues, relative to bodies. FA - female antennae, MA - male antennae, FP - female palps, MP - male palps.
annotated *An. gambiae* genes. On average, 57.4% of the reads per sample mapped to annotated genes, 91.5% to the nuclear genome (Table 1), and 2% to the mitochondrial genome (Table 1). Of the reads that mapped only to the genome, many of them are likely to represent unannotated 5’ or 3’ untranslated regions (UTRs). Moreover, there likely remain regions of the genome, most notably the Y-chromosome, where novel exons and transcripts remain [47].

On a whole-genome level, comparison of the mapping density of reads sequenced from the female body along all chromosomes showed a high degree of correspondence between the number of reads mapped to the nuclear genome and the number of reads mapped to the transcriptome (Figure 1). There are a few areas of asymmetry where a higher degree of mapping to either the transcriptome or to the genome was observed, most noticeably in the gene-rich autosomal telomeres and in several regions of the X chromosome (Figure 1). Greater mapping frequency to the genome represents actively transcribed regions that are unannotated as distinct genes. Greater mapping frequency to the transcriptome can generally be explained as reads that map to exon-exon junctions, which by their nature would not map to the genome. For example, the observed asymmetry in the 2R telomeric region is due to the high number of exon junction reads that mapped to two rhodopsin-family genes (Figure 1).

To quantify relative differences in gene expression levels within each tissue, we calculated a Reads Per Kilobase per Million (RPKM) reads mapped value for each gene within a sample [32]. Mean and median RPKM values for each tissue type in this study were very similar across samples, as were the number of genes showing basal or greater levels of transcription (Table 1). RPKM values spanned more than 6 orders of magnitude for each of the tissue types examined (see Additional file 1).

We assessed fold-differences in transcript abundance by independently comparing ratios of RPKM values between pairs of tissues within each sex: antennae to bodies and maxillary palps to bodies. For each of these pairwise comparisons we performed a Fisher’s Exact Test on counts of mapped reads and assigned statistical significance using a Bonferroni-corrected p-value (p < 3.9 × 10^-6; see Materials and Methods). Furthermore, we have used the term “enhanced” to describe any gene that displayed at least a 2-fold, significant difference in transcript abundance between the samples being compared (Figure 2). These conservative criteria were applied to avoid false positives stemming from variations within the samples themselves, as well as to reduce the numbers of genes that were used for subsequent analyses [48,49].

**Gene Expression Profiling in Chemosensory Tissues**

To examine global gene expression patterns, we have compared RPKM values pairwise for whole bodies versus either antennae or maxillary palps in both sexes. One such comparison is shown in Figure 2 where 4587 genes displayed enhancement in the female antenna to body comparison (Figure 2, black dots). Of those, 2277 were enhanced in the antenna (Figure 2, right half). Similarly, we found that 1906 genes were enhanced in
female palps, 3037 genes were enhanced in male antennae, and 2284 genes were enhanced in male palps. These 4 gene sets formed the basis of our subsequent analyses where we compared enhanced gene profiles between chemosensory tissues and across sexes (Figure 3).

Comparing the enhanced gene sets between the female antennae and palps revealed significant overlap, with 1158 genes (61% of palp set) enhanced in both tissues (see Additional file 2). Similarly, male antennae and palps showed significant overlap with 1208 genes enhanced in both tissues (53% of palp set; see Additional file 2). Interestingly, the most well-represented gene families in both of these overlapping sets were 7-transmembrane receptors (PF00001), protein kinases (PF00069), cytochrome P450s (PF00067), trypsins (PF00089), carboxylesterases (PF00135), and potential transcription factors (PFs 00046 and 00096; see Additional files 2 and 3, bottom tables). However, we also observed several differentially enhanced gene sets between the antennae and palps (see Additional files 1 and 2). The An. gambiae Ors (AgOrs; [12]; PfamA family PF02949) were the most prevalent class in female antennae (see Additional file 2, left table) and second-most in male antennae (see Additional file 3, left table). Other chemosensory gene families, such as ligand-gated ion channels, which include the recently identified ionotropic receptors (AgIrs; [13]; PfamA family PF02949) were the most prevalent class in female antennae (see Additional file 2, left table) and second-most in the male antennae (see Additional file 3, left table). Other chemosensory gene families, such as ligand-gated ion channels, which include the recently identified ionotropic receptors (AgIrs; [13]; PF00060), and odorant binding proteins (AgObps; [42]; PF01395), were highly
represented in the antennae (see Additional files 2 and 3). It is clear from these antennae-to-palp analyses that both extensive overlap and significant distinctions in gene expression profiles exist. The consistent identification of the same PfamA families in all enhanced gene sets implicates functional groups that can be studied in greater detail to elucidate their potential roles in mosquito chemosensation.

To evaluate gene expression patterns between sexes, we have compared female and male antennal-enhanced gene sets and palp-enhanced gene sets. In antennae, 2277 female, and 3037 male antennal-enhanced genes shared a common set of 1381 genes (Figure 3 and Additional file 4). Once again, this set included a common set of 1381 genes (Figure 3 and Additional file 4). In antennae, we have compared female and male antennal-enhanced gene sets and palp-enhanced gene sets. In antennae, we have compared female and male antennal-enhanced gene sets as a subset of genes that are significantly enhanced in both tissues.

Figure 3 An. gambiae Enhanced Gene Pairwise Tissue Comparisons. Proportional Venn diagrams showing the various pairwise comparisons made in this study. Overlap represents the subset of genes that are significantly enhanced in both tissues.

Their gene expression profiles is not necessarily straightforward. Chemosensory sensilla, and AgOr-containing neurons in particular, are distributed over the entire length of the female antenna, whereas male antennae house ~3-fold fewer chemosensilla that are restricted to distal segments 12 and 13 [4,50-52]. Furthermore, while female antennae are predominantly chemosensory, male antennae are also highly specialized for hearing [53,54]. Accordingly, the An. gambiae orthologs of the D. melanogaster trpV channels Nanchung and inactive, which are required for hearing in the fruit fly, were enhanced in antennae of both An. gambiae sexes (AGAPs 012241 and 000413, respectively; Table 2), but their expression levels were much higher in male antennae (RPKM of 183.92 and 104.49 in males and 20.54 and 7.66 respectively, in females) [55,56]. This elevated expression of auditory-associated genes in the male antenna is consistent with male An. gambiae mating behavior where an acute sense of hearing facilitates the recognition of female wing beats within the context of a male swarm [40,53,54]. Given that wild females are likely to mate just once, while males swarm daily in search of a mate [20,40], the specialization shift away from olfaction and toward audition in the principle male sensory organ is reasonable, presumably as a mechanism to increase male mating success.

These comparisons also revealed multiple classes of genes beyond the known chemosensory gene families that displayed enhanced tissue expression. A detailed examination of the expression patterns of a subset of other gene families is provided in Table 2, many of which are represented in figures found in Additional files 2, 3, 4, and 5. Nearly half of the members of the large superfamily of 7-transmembrane (7tm) receptors (114 of the 241 recognized by PfamA) were enhanced in at least one of the chemosensory tissues examined (Table 2). This may indicate unrecognized roles in sensory reception or regulation of chemoreceptor neuron or accessory cell function. Importantly, efferent projections from serotonergic, or tachykinin neuroendocrine cells have been identified in mosquito chemosensory appendages [57-59]. Thus the expression of serotonin (AGAPs 002232, 002679, 004222, 004223, 007136, and 011481), and tachykinin (AGAPs 001592 and 012824) receptor homologs in An. gambiae antennae and maxillary palps (see Additional file 1) is consistent with a neuromodulatory role for these compounds.

Other gene families with multiple members that displayed chemosensory enhancement include the CD36 family, some members of which function in insect olfaction [60,61] and ion channels and transporters, including the recently identified chemosensory ionotropic receptors [13,62,63]. Enhanced levels of such biotransformation enzymes as carboxylesterases and cytochrome
P450s could hint at a potential role in odor degradation [64-66]. Carbonic anhydrases involved in carbon dioxide detection in mammals [67] and transcription factors, including the *An. gambiae* homologs of *acj6* and *pdm3*, *D. melanogaster pou*-type transcription factors involved in *DmOr* gene regulation and ORN axon targeting [68-74] were also enhanced in chemosensory tissues (Table 2).

We also identified a number of genes encoding small, soluble proteins with enhanced expression in chemosensory tissues of both sexes (Table 2). Transcripts encoding the CRAL-TRIO (PF00650), cysteine-rich secretory protein (PF00188), and haemolymph juvenile hormone binding proteins (JHBP, PF06585) were highly enhanced. To our knowledge, the first two gene families have not been linked to chemosensation, but the members of the JHBP family have been identified in screens of highly expressed genes in mosquito antennae [75,76]. Moreover the JHBP gene, *takeout*, links the circadian clock and feeding behavior in *D. melanogaster* [77], and modulates aggregation behavior in *Locusta migratoria* [78]. The extremely high expression levels of some members of these 3 gene families suggest potential chemosensory functions analogous to other soluble lipophilic carriers such as the *Obps*.

**Chemosensory Gene Families**

In light of the existing literature on the molecular mechanisms underlying the processes of peripheral chemosensation in vector mosquitoes, we examined in detail the expression patterns of *AgOrs*, *AgIrs*, *AgObps* and gustatory receptors (*AgGrs*). As expected, the vast majority of *AgOrs* were highly enhanced in antennae. Of the 76 *AgOrs*, 58 showed enhanced expression in female antennae as compared to only 36 in male antennae (Figure 5). The entire set of male-enhanced *AgOrs* was contained within the female enhanced set. None of the larval-specific *AgOrs*: 37, 40, 52, or 58, was enhanced in adult antennae or palps, supporting previous observations [79]. In the palps, only *AgOrs*8 and 28 and *AgOrco* (formerly *AgOr7*, recently renamed to reflect its capacity as an obligate Or co-receptor) were enhanced in female maxillary palps (Figure 5), a result consistent with our previous study on odor coding in the *An. gambiae* maxillary palps [11]. The same 3 *AgOrs* were enhanced in male palps (Figure 5).

Several members of the recently described *AgIr* gene family [13,62], displayed significant enhancement in antennae of both sexes (Figure 6), further supporting their potential roles as chemosensory receptors in *An. gambiae*. A high degree of overlap was observed...
between the sexes, where 21 AgIrs were enhanced in both. Similar to the AgOrs, there were many fewer AgIrs enhanced in the palps compared to the antennae, with 7 and 6 enhanced in female and male palps, respectively. Furthermore, the degree of overlap (3 genes) between the sexes was much less pronounced in the palp (Figure 6).

The enhanced AgGrs were the only class that did not overlap in the antennae between the sexes, with very few showing enhancement in either females or males (Figure 7). Only AgGr1 was enhanced in female antennae, while AgGrs, 33, 48, 49, and 50 were enhanced in male antennae. Notably, one member of this large gene family, AgGr33 was strongly enhanced in male antennae (Figure 7), perhaps indicating a specialized function in males. Interestingly, AgGr33 shares significant homology with D. melanogaster Gr28 [12], some splice forms of which are expressed in non-chemosensory tissues, including the Johnston’s organ [80]. In contrast to the acute sexual dimorphism displayed in the antennae, both sexes showed enhanced expression of AgGrs 22, 23, and 24, in their maxillary palps (Figure 6). These three AgGrs are homologs of the D. melanogaster carbon dioxide receptors [81-83], and are expressed in capitate peg sensilla on the maxillary palps where they have been directly implicated in An. gambiae CO2 sensing [11].

Enhanced chemosensory expression of members of the large AgObp family was evident across all tissues and
Sixteen classical and 3 Plus-C AgObps were significantly enhanced in the female antennae (Figure 8). Of these, 17 were also significantly enhanced in the male antennae (Figure 8) including the D. melanogaster LUSH homolog, AgObp4 [84]. AgObp19 was the only one that demonstrated significantly enhanced expression in the female antennae and in no other tissue. In the maxillary palp, enhancement of AgObp transcripts also displayed substantial overlap between sexes, where the 4 male enhanced AgObps were all similarly elevated in females. Overall, the AgObp expression pattern was nearly identical between male and female chemosensory tissues (Figure 8). Several AgObps were also enhanced in bodies, with a greater number of them being enhanced in male bodies (Figure 8). A direct comparison of AgObp expression in female and male bodies revealed 17 AgObps that were specifically enhanced in males, while only 5 were enhanced in females. The high representation of AgObps in the male body without the coordinate expression of known chemoreceptors (Figures 5, 6, and 7) suggests uncharacterized roles for these AgObps, perhaps as general lipophilic carriers in male-specific physiology or in chemosensory processes in tarsal and wing sensilla.

In contrast, atypical AgObps were not enhanced in any of the tissues examined, which is consistent with previous results suggesting that expression of this subfamily is limited to pre-adult stages [42]. With the exception of AgObps 47, 48, 57, which had RPKMs of >1000, expression levels of the members of the Plus-C
AgObp subfamily was very low. Of these, it is noteworthy that AgObp48 was one of the most highly expressed genes (RPKM = 32311) in any tissue, with greatly elevated expression levels in both the male and female olfactory tissues (see Additional file 1). While AgObps, and insect Obps in general are among the most highly expressed gene families in chemosensory tissues their role in non-pheromone chemosensation remains largely undefined.

With regard to olfaction, it has been hypothesized that Obps act principally as molecular shuttles/chaperones, which deliver odorants to receptors and/or transiently protect specific odorants from biotransformation enzymes [43]. If individual Obps bind a subset of odorants, it is reasonable to hypothesize, that in tissues with high Or and therefore odor-coding complexity such as the antennae, the Obp landscape would need to be similarly complex in order to bind the required range of odorants. The converse would also be expected for tissues with reduced odor coding complexity such as the maxillary palp.

Female antennae showed enhanced expression of 58 conventional AgOrs, while only 3 AgOrs were enhanced in the female palp (Figure 5). Furthermore, the odorant response profiles of the palp-expressed AgOrs 8 and 28 were also vastly different from the deorphianized antennal AgOrs [11,17,18]. These differences in AgOr coding capacity and their expression profiles strongly suggest that the ability of the female antennae to sense odors is much greater than the maxillary palp.
In *An. gambiae* females both the antennae and maxillary palps expressed 21 AgObp family members with an RPKM >10, of which 19 were found in both (see Additional file 1). While not all of these AgObps' transcript levels met our criteria for enhancement, they were nevertheless expressed in these tissues. Although the AgObp complexity was almost identical in these two appendages, they displayed vastly different AgOr complexity and thus odor coding capacities. This analysis confounds standing theories about Obp function; if a large number of Obps are required in the antennae for signaling, then their presence in the palp, with its more limited odor coding capacity, would appear superfluous. Given the broad expression of AgObps and a demonstrated lack of functional overlap between the antennae and palps, our analysis suggests that in at least some instances, Obps act as low-pass filters for environmental odorants rather than as specific odorant-carrier agents. Therefore, Obps may act to solubilize odors in some cases, but as molecular sinks in others, adding yet another dimension to peripheral odor coding. In addition, the near ubiquitous expression in both sexes of some AgObps suggests that they are playing completely uncharacterized roles outside of chemosensory processes.

**Diverse Roles for Chemosensory Tissues**

To explore the effect of morphology on observed AgOr expression, we have attempted to normalize sex-specific differences in transcript abundance by scaling up male...
AgOrs in proportion to the number of female chemosensilla. AgOrs are expressed in the trichoid sensilla, the predominant sensillar type, and not in grooved peg (GP) sensilla [14]. Sensilla counts have indicated that female antennae house an average of 630 trichoid sensilla while male antennae house an average of 225 trichoid sensilla [4,51,85]. Multiplying the male AgOr RPKM levels by a factor of 2.8 (630/225), resulted in a sex-normalized, AgOr expression profile that was qualitatively very similar in both sexes (Figure 9, top panel), yet male AgOr RPKM values remained lower than those in females. Alternatively, normalization of male AgOr expression by a factor equal to the fold-difference in AgOrco RPKM levels between female and male antennae (916/186 or 4.9) was also performed. In the case of AgOrco normalization, AgOr expression profiles appeared to be very similar between sexes (Figure 9, middle panel). Based on the same logic, we also normalized AgIr expression in male antennae (Figure 9, bottom panel). We postulated that AgIrs are localized in GP sensilla neurons, as they are in D. melanogaster [13,61], we used a GP normalization factor of 4.2, which is the fold difference in GP numbers between female and male An. gambiae antennae [4]. As with AgOrs, the AgIr gene expression patterns were qualitatively similar in both sexes after normalization (Figure 9). Taken together these results suggest that male antennae express AgOr and AgIr chemoreceptor repertoires that are highly similar to those expressed in female antennae.

The AgOr and AgGr expression profiles in the maxillary palps support a similar conclusion. Although AgOrco, AgOrs 8 and 28, and AgGrs 22, 23 and 24 were enhanced in both sexes, their expression levels were lower in males than in females (Figures 5 and 7). As is the case for An. gambiae antennae, the maxillary palps are sexually dimorphic with males having about 4-fold fewer chemosensilla [4,11]. This could account for the apparent lower chemosensory gene transcript abundances in males. Normalizing male palp AgOrs and AgGrs by this factor elevated their RPKM values closer

**Figure 9 Normalized AgOr and AgIr Expression Profiles.** Top Panel - Sensilla-normalized (factor 2.8) AgOr expression profile map. Middle Panel - AgOrco-normalized (factor 4.9) AgOr expression profile map. Bottom Panel - Sensilla-normalized (factor 4.2) AgIr expression profile map. Green color intensity scale (below maps) indicates increasing RPKM values from left to right. MA - male antennae RPKMs. FA - female antennae RPKMs. MA norm - normalized male antennae RPKMs.
palmal chemosensilla in Mansonia least 2 species of mammalian hosts at greater distances than conspecific processing could contribute significantly to sexually chemosensory abilities. Additionally, higher order distinguishing qualitatively and quantitatively female and male functionally relevant and serve as the basis for distinc-
differences in gene expression profiles could also be specific females [88]. In either case, the aforementioned differences in gene expression profiles could also be functionally relevant and serve as the basis for distin-
guishing qualitatively and quantitatively female and male chemosensory abilities. Additionally, higher order processing could contribute significantly to sexually dimorphic behaviors, thus the perception of the same odors may elicit very different responses in the sexes. These competing hypotheses are directly testable using a combination of electrophysiological recording and behavioral response assays. Moreover, the requirement in chemoreception for any of the differentially expressed genes could potentially be explored by gene silencing.

Conclusion
We are interested in understanding the molecular components of the chemosensory pathways that modulate the physiology and behaviors that distinguish female mosquitoes which blood-feed from males that do not. Inasmuch as differential gene expression between the sexes may serve as a potential mechanism for modulating peripheral sensitivity, we have carried out a compre-
hensive comparative analysis of the chemosensory transcriptomes of adult male and non-bloodfed female An. gambiae. Broadly, we identified several novel classes of protein coding genes whose expression is strongly biased toward chemosensory tissues, which have to date, not been associated with chemosensory pathways. Principal among these are several cytochrome P450s and a wide range of cysteine-rich secretory proteins. These genes exhibit the localization, expression and physical properties consistent with a role in semiochemical binding.

With regard to AgObps, a known class of semiochemical binding proteins, our data suggest that in chemosensory tissues the number of enhanced AgObps is discordant with the number of enhanced chemoreceptors. This would belie the accepted and singular role for AgObps in odorant binding and clearance. It is therefore reasonable to hypothesize that AgObps play multi-
faceted and as yet, not fully characterized, roles in the physiology of An. gambiae.

For the principal chemosensory gene classes, we observed an unexpected pattern of conserved AgOr expression between male and female antennae, lending strong empirical reinforcement to prior speculation that male and female mosquitoes share a similar range of odor coding capacity. However, the relative levels of AgOr transcripts were much higher in the female antennae, a finding consistent with the females’ greater number of chemosensory sensilla and indicative of enhanced odor sensitivity. A similar situation was also observed in the antennal expression levels of genes associated with hearing which are expressed in both male and female antennae albeit at much higher in males. Taken together, these findings reveal the antenna of An. gambiae to be a bi-modal sensory appendage, one that shares a surprisingly similar suite of sensory genes between the sexes. The difference between male and female antenna seems to be less one of transcript type and more one of transcript quantity, reflective of sexually dimorphic sensory prioritizations. Male mosquito antennae are more specialized for audition and female antenna for olfaction.

We have used RNA-sequencing to conduct a high resolution and quantitative assessment of whole-trans-
scriptome gene-expression profiles in chemosensory tissues and bodies of an organism of great medical importance. This study has begun to explore the potential of this dataset insofar as its implications for odor coding mechanisms in An. gambiae thereby establishing a precedent for the use of these approaches for the study of insect chemosensory processes.

Additional material

Additional file 1: An. gambiae Transcriptome Expression Data. Table of mapped reads to AnagP3.6 transcripts for all 6 data sets. VectorBase ID: Unique VectorBase (http://www.vectorbase.org) identification number for each An. gambiae gene. transcript length: length in base pairs of the longest annotated transcript for each gene. chromosomal location: chromosome arm, location of the first base pair of the initiator codon, location of the last base pair of the stop codon, reading frame (1 for plus strand or -1 for minus strand), gene name (if any). best match to NR database: best match to non-redundant protein database (http://www.ncbi.nlm.nih.gov/BLAST/blastcgihelp) with An. gambiae proteins removed. %ID: percent amino acid identity between An. gambiae and best match peptides. PfamA best hit: best match to protein family identified in PfamA searches (http://pfam.janelia.org/). e-value: relevance value as
returned in PfamA searches. PfamA description: protein family description. gene: AgQ1, AgT, AgG, and AgGp gene families identified for easy reference. RPKM: normalized transcript abundance values for each gene in the indicated tissues. unique hits: number of RNA-seq reads that map uniquely to each gene. total hits: weighted number of RNA-seq reads (unique plus fraction of non-unique) that map to a given gene. PfamA values in bold type indicate significantly enhanced transcript abundance (≥2-fold) in the antenna or palp relative to body for a given gene.

Additional file 2: Female Antennae vs. Palps Enhanced Gene Sets

Venn diagram showing the numbers of genes that are significantly enhanced in female antennae and maxillary palps. Overlap represents the subset of genes that are significantly enhanced in both sexes. Boxes contain ranked lists of the most prevalent PfamA families in each data set.

Additional file 3: Male Antennae vs. Palps Enhanced Gene Sets

Venn diagram showing the numbers of genes that are significantly enhanced in male antennae and maxillary palps. Overlap represents the subset of genes that are significantly enhanced in both sexes. Boxes contain ranked lists of the most prevalent PfamA families in each data set.

Additional file 4: Female vs. Male Antennae Enhanced Gene Sets

Venn diagram showing the numbers of genes that are significantly enhanced in female and male antennae. Overlap represents the subset of genes that are significantly enhanced in both sexes. Boxes contain ranked lists of the most prevalent PfamA families in each data set.

Additional file 5: Female vs. Male Palps Enhanced Gene Sets

Venn diagram showing the numbers of genes that are significantly enhanced in female and male maxillary palps. Overlap represents the subset of genes that are significantly enhanced in both sexes. Boxes contain ranked lists of the most prevalent PfamA families in each data set.

Acknowledgements

We thank the following individuals: Zhen Li for mosquito rearing, Travis Clark, Chelsea Baker and the Vanderbilt Genomics Shared Resource for sample preparation and Illumina sequencing; Shawn Levy and the Genomic Services Lab at HudsonAlpha Institute for Biotechnology for Illumina sequencing; John Gibbons and Kris McGary for helpful discussions of RNA-seq mapping and statistical methods. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. It was supported by the Searle Scholars Program and the National Science Foundation (DEB-0844968) to AR as well as grants from the Innovation University. It was supported by the Searle Scholars Program and the National Science Foundation (DEB-0844968) to AR as well as grants from the Innovation

Authors' contributions

RJP, DCR and PLJ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work. RJP, DCR, PLJ, AR and LJZ contributed equally to this work.

Received: 24 March 2011 Accepted: 27 May 2011

Page 15 of 17

References

1. Gilot C. Entomology. Springer, 3 2005.

2. Sato K, Oukawa K. Insect olfaction: receptors, signal transduction, and behavior. Results Probl Cell Differ 2009, 47:121-138.

3. de Bruyne M, Baker TC. Odor detection in insects: volatile codes. J Chem Ecol 2008, 34(7):882-897.

4. Mcker SB. Sensilla mosquitoes (Diptera: Culicidae). J Med Entomol 1982, 19(5):489-533.

5. Kaisling KE. Olfactory perireceptor and receptor events in moths: a kinetic model revised. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 2009, 195(10):895-922.

6. Touhara K, Vosshall LB. Sensing odors and pheromones with chemosensory receptors. Annu Rev Physiol 2009, 71:307-332.

7. Hallem EA, Carlson JR. Coding of odors by a receptor repertoire. Cell 2006, 125(1):143-160.

8. Zwiebel Lj, Takken W. Odor-mediated regulation of mosquito-host interactions. Insect Biochem Mol Biol 2004, 34(7):645-652.

9. Takken W, Knols BG. Odor-mediated behavior of Afrotropical malaria mosquitoes. Annu Rev Entomol 1999, 44:131-157.

10. Costantini C, Sagnon N, della Torre A, Coluzzi M. Mosquito Behavioral Aspects of Vector-Human Interactions in the Anopheles gambiae complex. Parasitology 1999, 41:209-217.

11. Lu T, Qiu YT, Wang G, Kwun JY, Rutzler M, Kwon HW, Pitts RJ, van Loon JJ, Takken W, Carlson JR. Zwiebel LJ. Odor coding in the maxillary palp of the malaria vector mosquito Anopheles gambiae. Curr Biol 2007, 17(18):1533-1544.

12. Hill CA, Fox AN, Pitts RJ, Kent LB, Tan PL, Chrysal MA, Cravchik A, Collins FH, Robertson HM, Zwiebel LJ. G protein-coupled receptors in Anopheles gambiae. Science 2002, 298(5591):176-178.

13. Liu C, Pitts RJ, Bobbot JD, Jones PL, Wang G, Zwiebel LJ. Distinct olfactory signaling mechanisms in the malaria vector mosquito Anopheles gambiae. PLoS Biol 2010, 8(8).

14. Fox AN, Pitts RJ, Robertson HM, Carlson JR, Zwiebel LJ. Candidate odorant receptors from the malaria vector mosquito Anopheles gambiae and evidence of down-regulation in response to blood feeding. Proc Natl Acad Sci USA 2001, 98(25):14693-14697.

15. Pitts RJ, Fox AN, Zwiebel LJ. A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector Anopheles gambiae. Proc Natl Acad Sci USA 2004, 101(14):5058-5063.

16. Kwon HW, Lu T, Rutzler M, Zwiebel LJ. Odorant responses in a gustatory organ of the malaria vector mosquito Anopheles gambiae. Proc Natl Acad Sci USA 2006, 103(36):13526-13531.

17. Wang G, Carey AF, Carlson JR, Zwiebel LJ. Molecular basis of odor coding in the malaria vector mosquito Anopheles gambiae. Proc Natl Acad Sci USA 2010, 107(9):4418-4423.

18. Carey AF, Wang G, Su CY, Zwiebel LJ, Carlson JR. Odorant receptor in the malaria mosquito Anopheles gambiae. Nature 2010, 464(7285):166-171.

19. Verhults ND, Takken W, Dicke M, Schraa G, Smagghe GC. Chemical ecology of interactions between human skin microbiota and mosquitoes. FEMS Microbiol Ecol 2010, 74(1):1-9.

20. Howell PI, Knols BG. Male mating biology. Malar J 2009, 8 Suppl 2:S8.

21. Clements AN. The Biology of Mosquitoes. London, Glasgow, New York, Tokyo, Melbourne, Madrid: Chapman & Hall, 1992.

22. Aguilar R, Simard F, Kamdem C, Shields T, Glass GE, Garver LS. Dimopoulos G. Genome-wide analysis of transcriptomic divergence between laboratory colony and field Anopheles gambiae mosquitoes of the M and S molecular forms. Insect Mol Biol 2010, 19(3):695-705.

23. Das S, Radtke A, Choi YJ, Mendes AM, Valenzuela JG, Dimopoulos G. Transcriptomic and functional analysis of the Anopheles gambiae salivary gland in relation to blood feeding. BMC Genomics 2010, 11:566.

24. Cook PE, Sinkins SP. Transcriptional profiling of Anopheles gambiae mosquitoes for adult age estimation. Insect Mol Biol 2010, 19(6):745-751.

25. Baton LA, Robertson A, Warr E, Strand MR, Dimopoulos G. Genome-wide transcriptomic profiling of Anopheles gambiae hemocytes reveals pathogen-specific signatures upon bacterial challenge and Plasmodium berghei infection. BMC Genomics 2009, 10:257.

26. Warr E, Aguilar R, Dong Y, Maharai V, Dimopoulos G. Spatial and specific dissection of the Anopheles gambiae midgut transcriptome. BMC Genomics 2007, 8:37.

27. Marinotti O, Calvo E, Nguyen QK, Dissanyake S, Ribeiro JM, James AA. Genome-wide analysis of gene expression in adult Anopheles gambiae. Insect Mol Biol 2008, 15(1):1-12.

28. Marinotti O, Nguyen QK, Calvo E, James AA, Ribeiro JM. Microarray analysis of genes showing variable expression following a blood meal in Anopheles gambiae. Insect Mol Biol 2005, 14(4):365-373.

29. Aguilar R, Jedlicka AE, Mintz M, Maharai V, Scott AL, Dimopoulos G. Global gene expression analysis of Anopheles gambiae responses to microbial challenge. Insect Biochem Mol Biol 2005, 35(7):709-719.
75. Justice RW, Dimitratos S, Walter MF, Biessmann H: Sexual Dimorphic Expression of Putative Antennal Carrier Protein Genes in the Malaria Vector Anopheles gambiae. *Insect Molecular Biology* 2003, 12(6):581-594.

76. Bohbot J, Vogt RG: Antennal expressed genes of the yellow fever mosquito (Aedes aegypti L): characterization of odorant-binding protein 10 and takeout. *Insect Biochem Mol Biol* 2005, 35(9):961-979.

77. Sarov-Blat L, So VW, Liu L, Rosbash M: The Drosophila takeout gene is a novel molecular link between circadian rhythms and feeding behavior. *Cell* 2000, 101(6):647-656.

78. Guo W, Wang X, Ma Z, Xue L, Han J, Yu D, Kang L: CSP and Takeout Genes Modulate the Switch between Attraction and Repulsion during Behavioral Phase Change in the Migratory Locust. *PLoS Genet* 2011, 7(2): e1001291.

79. Xia Y, Wang G, Buscariollo D, Pitts RJ, Wenger H, Zwiebel LJ: The molecular and cellular basis of olfactory-driven behavior in Anopheles gambiae larvae. *Proc Natl Acad Sci USA* 2008, 105(17):6433-6438.

80. Thorne N, Amrein H: Atypical expression of Drosophila gustatory receptor genes in sensory and central neurons. *Journal of Comparative Neurology* 506:548-568.

81. Kwon JY, Dahanukar A, Weiss LA, Carlson JR: The molecular basis of CO2 reception in Drosophila. *Proc Natl Acad Sci USA* 2007, 104(9):3574-3578.

82. Cayirlioglu P, Kadow IG, Zhan X, Okamura K, Suh GS, Gunning D, Lai EC, Zipursky SL: Hybrid neurons in a microRNA mutant are putative evolutionary intermediates in insect CO2 sensory systems. *Science* 2008, 319(5867):1256-1260.

83. Robertson HM, Kent LB: Evolution of the gene lineage encoding the carbon dioxide receptor in insects. *J Insect Sci* 2009, 9:19.

84. Kim MS, Repp A, Smith DP: LUSH odorant-binding protein mediates chemosensory responses to alcohols in Drosophila melanogaster. *Genetics* 1998, 150(2):711-721.

85. Zwiebel LJ, Pitts RJ: Antennal sensilla of two female anopheline sibling species with differing host ranges. *Malar J* 2006, 5:26.

86. Blackwell A, Johnson SN: Electrophysiological investigation of larval water and potential oviposition chemo-attractants for Anopheles gambiae s.s. *Ann Trop Med Parasitol* 2000, 94(4):389-398.

87. McVey SB, Wilkes TJ, Gilles MT: Attraction to mammals of male Mansonia (Mansonioides) (Diptera: Culicidae). *Bulletin of Entomological Research* 1980, 70(01):11-16.

88. McVey SB: Sensory aspects of mate-finding behavior in male mosquitoes (Diptera: Culicidae). *Journal of Medical Entomology* 1980, 1754-57.