Published in final edited form as:

Comp Biochem Physiol Part D Genomics Proteomics. 2017 December ; 24: 41–50. doi:10.1016/j.cbpd.2017.07.005.

PROFILES OF SOLUBLE PROTEINS IN CHEMOSENSORY ORGANS OF THREE MEMBERS OF THE AFRO-TROPICAL *ANOPHELES GAMBIAE* COMPLEX.

Immacolata Iovinella^{a,b}, Beniamino Caputo^b, Maria Calzetta^b, Laurence J. Zwiebel^d,
Francesca Romana Dani^{a,c,*}, Alessandra della Torre^b

^aBiology Department, Università di Firenze, Italy

^bDepartment of Public Health & Infectious Diseases, Laboratory affiliated to Istituto Pasteur Italia-Fondazione Cenci Bolognietti, University of Rome “Sapienza”, Rome, Italy

^cCISM, Mass Spectrometry Centre, Università di Firenze, Italy

^dDepartments of Biological Sciences and Pharmacology, Vanderbilt University, Nashville, USA

Abstract

In mosquitoes females, host-seeking, host-choice and other reproductive behaviours are driven by odorants. Species of the Afrotropical *Anopheles gambiae* complex display divergent host-preferences associated with significant differences in their efficiency as vectors of human malaria. Odorant perception starts in chemosensilla which are primarily housed in the antennae and maxillary palps. Taking advantage of recently completed genome sequencing efforts for species of the *Anopheles gambiae* complex, we applied shotgun proteomics to characterize the profile of soluble proteins of antennae and maxillary palps of three different species: *An. coluzzii*, *An. arabiensis* and *An. quadriannulatus*, showing remarkable differences in anthropophilic behaviour. Protein profiles with particular focus on soluble proteins involved in the olfactory-based detection of odours were compared. Several differences between species were found in the abundance of Insect Cuticle proteins. Moreover, proteins belonging to Glutathione S-transferase and Odorant binding proteins, the latter of which known to be directly involved in odour recognition, are also different among the three species, with *An. coluzzii* and *An. arabiensis* being less different with respect to *An. quadriannulatus*.

Keywords

antennae; maxillary palps; *An. coluzzii*; *An. arabiensis*; *An. quadriannulatus*; shotgun-proteomic

1. INTRODUCTION

The *Anopheles gambiae* complex includes eight morphologically indistinguishable mosquito species. Differences in olfaction driven behaviors of adult females, the most relevant being

*Corresponding author. Dipartimento di Biologia, Università degli Studi di Firenze, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy. francescaromana.dani@unifi.it (F.R. Dani).

host seeking and host choice, but also including search for oviposition sites, is well documented among some species (Takken and Knols, 1999; Takken and Verhulst, 2013; Lutz et al. 2013). Among the three most widely spread species in sub Saharan Africa, *An. gambiae* and *An. coluzzii* are highly anthropophilic, *An. arabiensis* has a more generalistic feeding behavior, while the southern *An. quadriannulatus* is strongly zoophilic and rarely if ever attacks humans (Dekker et al., 2001; Pates et al., 2006; Scott and Takken, 2012). All species of the complex are vectors of the human malaria parasite *Plasmodium falciparum*, however, due to considerable variation in species-specific host preference, their vectorial capacity varies from being the highest among all Afrotropical vectors in the case of *An. gambiae* and *An. coluzzii*, to intermediate levels of vectorial capacity for *An. arabiensis*, to null in the case of *An. quadriannulatus* (Takken et al., 1999; Habtewold et al., 2008)..

In anophelines and other mosquitoes, host-seeking behaviors and host preference are modulated by a range of sensory modalities that are largely comprised of chemosensory cues detected by the olfactory system of adult females (Montell and Zwiebel, 2016). Taken together, these elements are responsible for the ability of the mosquitoes to identify and respond to a range of discrete semiochemicals emitted by potential human and animal hosts probably in the form of species-specific odor blends (Costantini et al., 1998; Tirodos et al., 2006; Zwiebel and Takken, 2004; Takken and Verhulst, 2013). In mosquitoes, as generally in insects (Leal, 2013), semiochemicals are initially detected by parallel signal transduction pathways expressed in odorant receptor neurons (ORNs) found within sensory hairs known as sensilla that decorate the antennae and the maxillary palps as well as other chemosensory appendages (Takken and Knols, 1999; Lu et al., 2007; Know et al., 2006; Guidobaldi et al., 2014). Odorants diffuse through numerous surface pores on each sensillum to enter an aqueous lymph that must be traversed in order to reach the spectrum of molecular receptors present on the ORN dendrites (Steinbrecht, 1997).

At the molecular level, insect chemosensory signal transduction pathways initiate with the interplay of odorants with soluble proteins that are secreted into the extracellular sensillar lymph. These include several odorant binding proteins (OBPs) and chemosensory proteins (CSPs) that have been postulated to act as carriers of odorants, facilitating their access to ORNs and contributing to odour discrimination (Pelosi et al., 2014). Recently OBPs have been also proposed to buffer changes in odor environments (Larter et al., 2016). In addition, numerous classes of enzymes and other proteins collectively act as odorant degradation enzymes (ODEs) involved in clearance of odorants from the sensillar lymph (Suh et al., 2014). Once odorants reach the ORN dendrite, specific recognition and binding occurs via the action of specific sets of odorant receptors (ORs) (Su et al., 2009), ionotropic receptors (IRs) or gustatory receptors (GRs). These proteins are expressed in discrete sets of receptor neurons where, once sufficiently activated, they act to trigger action potentials that travel along axons projecting to discrete glomeruli in the antennal lobe which is the initial site of synaptic integration that eventually results in perception and behavioral responses (Montell and Zwiebel, 2016). Because of the direct interaction between the receptors and the odorants, differences in host preference between species are likely to be reflected in differences in the abundance or molecular structure of either the receptors or other olfactory proteins secreted into the extracellular sensillar lymph where they presumably directly interact with odorants (Rinker et al., 2013a).

The comprehensive annotation of several, in some instances, very large gene families that encode for chemosensory components of *An. gambiae* were first described as part of the *An. gambiae* genome project (Holt et al., 2002). Similar bioinformatics-based approaches have subsequently identified candidate chemosensory genes in other mosquito genomes, including the yellow fever and dengue vector, *Aedes aegypti* (Bohbot et al., 2007; Manoharan et al., 2013). ORs, GRs and to a lesser degree CSPs and OBPs display generally high levels of sequence divergence, while, in contrast, the IRs, and ODEs are significantly more highly-conserved. Furthermore, since their initial identification, numerous studies have examined the abundance and functionality of anopheline ORs, most notably in *An. coluzzii* where nearly all the odorant recognition specificities (sometimes denoted as “odor coding”) of the majority of ORs have been characterized (Lu et al, 2007, Wang et.al, 2010; Carey et al, 2010).

More recently, the *Anopheles* Genomics Consortium comprehensively identified and annotated all the chemosensory genes (eg OBPs, CSPs, ODEs, ORs and other chemoreceptors) from 16 anopheline species (Neafsey et al., 2015). The availability of comprehensive genomic information has driven a range of studies that have used microarrays and Next Generation Sequencing (NGS) approaches, most notably massively parallel sequencing of RNA molecules (RNAseq) to comprehensively map the chemosensory transcriptomes of several mosquito vectors (Rinker et al., 2016). For example, in order to examine whether host preference is associated with the peripheral abundance profiles of chemosensory genes, Rinker et al (2013a) compared the transcriptome profiles of the antennae of non-blood fed, female *An. gambiae* and its zoophilic sibling species *An. quadriannulatus*. While because of their extraordinarily close evolutionary relationship, these species retain a nearly identical repertoire of OR, this study uncovered significant inter-specific enrichment in the abundance of a subset of ORs in *An. gambiae* that are specifically tuned to human-associated odors. Taken together, these data facilitate an understanding of how shifts in peripheral sensitivity may drive the development of anthropophilic host preference that underlies the profound vectorial capacity of *An. coluzzii*. While technical constraints make proteomic studies generally not informative for hydrophobic membrane proteins that notably are expected to include most, if not all, sensory receptors, they nevertheless complement transcriptome profiling studies and, by measuring directly the proteins rather than the transcripts, provide a more reliable profile of soluble proteins. However, thus far only three studies focused on the characterization of the proteome profile of *An. gambiae* antennae (Mastrobuoni et al., 2013; Rund et al., 2013; Zhou et al., 2016) and none of them has extended interspecific comparisons to the maxillary palps which represents an important, albeit secondary olfactory appendage. In the present work, we have applied a shotgun proteomic approach to characterize and compare the relative abundance of proteins in antennae and maxillary palps of host-seeking adult females of *An. coluzzii*, *An. arabiensis* and *An. quadriannulatus*, with a focus on the abundance patterns of soluble olfactory proteins.

2. MATERIALS AND METHODS

2.1 'Mosquito rearing and tissue preparation

An. coluzzii (GA-CAM), *An. arabiensis* (AR-KGB) and *An. quadriannulatus* (SANQUA) were reared in the insectaries of the Department of Public Health & Infectious Diseases of Sapienza University, at $26\pm1^{\circ}\text{C}$, $>70\%$ RH and 12-hour photoperiod and fed with 0.5% sugar solution. Three-day old adult females were deprived of sugar solution for 3–4 hours. After that time, at about dusk time, individuals landing on an operator hand, presumably searching for hosts, were collected and immediately anesthetized at -20°C . Antennae and maxillary palps were dissected under a microscope on a cold plate and for each appendage two pools, one from 50 and one from 25 *An. coluzzii* females, and two, both from 25 *An. arabiensis* and *An. quadriannulatus* females, were obtained and kept at -20°C until protein extraction.

2.2 Reagents

Ammonium bicarbonate, DTT, iodoacetamide, sodium chloride, formic acid, acetonitrile, trifluoroacetic acid, acetic acid, thiourea and bovine serum albumin were from Sigma-Aldrich (Milano, Italy), Tris and urea from Euroclone, Trypsin from Promega (Sequencing Grade Modified Trypsin), Lys-C from Thermo Scientific (MS grade) and the Protein Assay kit was from Bio-Rad. The hand-made desalting/purification STAGE column were prepared using three C18 Empore Extraction Disks (3M).

2.3 Protein Sample Preparation and Digestion

Antennae and maxillary palps samples were crushed in a mortar under liquid nitrogen and the proteins extracted with 6M Urea/2M Thiourea in Tris-Cl 50mM pH 7.4 with the addition of phenylmethanesulfonyl fluoride. The protein extracts were centrifuged at 14,000 rpm for 40 minutes at 4°C and the supernatants were collected for the analysis. The total amount of protein in each sample was assessed by the Bradford colorimetric assay (Bradford, 1976), with the "Bio-Rad Protein Assay" kit using serial dilutions of bovine serum albumin to generate a standard curve. Protein sample concentration was measured by Infinite PRO 200 reader (TECAN).

Protein digestion was carried out on 15 μg protein extracts. Reduction of disulfide bridges was performed by treating samples with DTT (1 μg of DTT/50 μg of proteins for 30 min at RT), followed by alkylation (5 μg of iodoacetamide/ 50 μg of proteins for 20 min at RT in the dark), as described by Foster and co-workers (Foster et al., 2003). Protein samples were diluted 3 times with 500 mM ammonium bicarbonate, to increase pH and reduce the concentration of urea/thiourea. An initial enzymatic digestion was performed by incubating the samples with Lys-C in a ratio 1:50 (w/w) for 3 h at 37°C . The digestion products were then incubated with trypsin in a ratio 1:50 (w/w) overnight at 37°C . The digested samples were then acidified by adding trifluoroacetic acid and desalted on STAGE tips (Rappsilber et al., 2007). The eluates were concentrated and reconstituted to 20 μL in 0.5% acetic acid, prior to HPLC-MS analyses.

2.4 Mass Spectrometric Analysis

For each sample, a volume containing the peptide mixture of corresponding to 2.25 µg of digested proteins, was submitted to a nanoLC-nanoESI-MS/MS analysis on an Ultimate 3000 HPLC (Dionex, San Donato Milanese, Milano, Italy) coupled to a LTQ-Orbitrap mass spectrometer (Thermo Fisher, Bremen, Germany), as described in detail in Iovinella et al. 2015.

2.5 Data processing

Different sets of analyses were performed using MaxQuant software (version 1.5.2.6) (Cox and Mann, 2008). Firstly, in order to highlight differences between antennae and palps within the same species, raw files of these tissues were analyzed together. Secondly, raw files of antennae of the three species as well as raw files of palps were separately analyzed, in order to compare the protein abundance among species in the same appendages. In each analysis, peak lists were searched with Andromeda search engine (Cox et al., 2011). Given the small number of annotated sequences in Uniprot for *An. coluzzii*, *An. arabiensis* and *An. quadriannulatus*, we used a combined database of all the proteins of the *Anopheles* genus obtained from Uniprot (September 2015) and FASTA files containing the list of predicted peptide sequences for the three species downloaded from VectorBase (Bioinformatics Resource for Invertebrate Vectors of Human Pathogens). In parameter section, we set as enzyme Trypsin and Lys-C, allowing up to two missed cleavages. The minimum required peptide length was seven amino acids. Carbamidomethylation of cysteine and oxidation of methionine were set as variable modifications. As no labeling was performed, multiplicity was set to 1. During the main search, parent masses were allowed an initial mass deviation of 4.5 ppm and fragment ions were allowed a mass deviation of 0.5 Da. PSM (Peptide Spectrum Match) and protein identifications were filtered using a target-decoy approach at a false discovery rate (FDR) of 1%. The second peptide feature was enabled.

The match between runs option was enabled with a match time window of 2.5 min and an alignment time window of 20 min. Relative label-free quantification (LFQ) of proteins was performed using the MaxLFQ algorithm integrated into MaxQuant. Default parameters were used. For protein quantification, we used the following parameters: 1 as LFQ Minimum ratio count, “Unique+Razor” peptides (i.e. those exclusively shared by the proteins of the same group), peptides with variable modifications, and unchecked “discard unmodified counterpart peptide”. All the informatics data are available as supplemental files in the “proteinGroups” output files, containing the full list of identified and quantified proteins (Tables S1, S2, S3, S4 and S5).

2.6 Data Analysis

Further analysis of the MaxQuant-processed data was performed using Perseus software (version 1.5.1.6) and the “proteingroups.txt” output files from these analyses were evaluated separately, as follows.

First, hits to the reverse database, contaminants and proteins only identified with modified peptides were eliminated. Samples were first grouped according to replicates. Annotations of *An. gambiae* regarding gene ontology (GO) categories, Pfam and Interpro were

downloaded from the link available in Perseus software (<http://141.61.102.106:8080/share.cgi?ssid=0q4b6sT>) and associated at each protein identifiers. Only the category of the leading protein was considered in the data analysis.

In the analysis of antennae and palps of the single species, LFQ intensity values were averaged between replicates and data were filtered for proteins identified with at least 2 “Unique + Razor” peptides. Proteins exclusive for each body part in each studied species were searched using the Perseus tool “Numeric Venn diagram”. Results are reported in supplementary file S6. To measure the difference in protein level between antennae and palps, a fold change was calculated. Proteins of each species having a fold change (LFQ antenna/LFQ palp) greater than 2 or lower than 0.5 were graphically reported in a bar chart according to their Pfam annotation.

To study differences in single protein levels between antennae and palps of the same species, LFQ intensity values were transformed by $\log_2$. Missing values were imputed according to the default settings of Perseus software; however, zero was manually assigned when a protein was not quantified in both replicates of the same appendage. Data were then filtered for proteins identified with at least 2 “Unique+ Razor” peptides. A t-test was applied to determine proteins significantly different in the two structures, with a permutation based False Discovery Rate (FDR) set to 0.05, number of randomization set to 1000 and S0 set to 0.1. This latter value is an artificial within-groups variance which controls both the relative importance of t-test p-value and difference between means (Tusher et al., 2001). Results of t-tests are reported in the supplementary file Table S7.

To compare the three species, we transformed by $\log_2$ the LFQ values of the antenna and palp “protein group” files where raw data of the three species were analyzed together. Data were filtered for proteins having all valid values in each sample (i.e. proteins whose LFQ intensity was calculated in each biological replicates of each species). The proteome profiles of antennae and palps of the three species were compared by applying a t-test where the Bonferroni correction was used and the FDR was set to 0.016. The number of randomization was set to 1000 and S0 to 0.1. Results of t-test are reported in the supplementary file Table S8 and visualized in the form of a volcano-plot (Figure S1). Unsupervised hierarchical clustering on the imputed data of proteins identified with 2 “Unique+ Razor” peptides, after assigning manually zero to proteins that have not been quantified in any replicates of one single species, was performed using the default settings of Perseus tool (Euclidean distance for the clustering process; Average as clustering method; 300 as number of clusters created by the k-means algorithm).

3. RESULTS & DISCUSSION

We carried out nanoLC-nanoESI-MS/MS coupled with a LTQ-Orbitrap mass spectrometer analysis to drive the interrogation of Vectorbase’s proteomic database and identified a total of 464 and 502 proteins in antennae and palps, respectively. Proteins were grouped according to Pfam annotation; families listing thresholds set at a minimum of 3 identified proteins in one chemosensory appendage of one of the three species examined are reported in Table 1.

Insect cuticle protein family groups are the higher number of identified proteins in all the three species; these proteins are characterized by a conserved 35-36 amino acid chitin-binding domain (Rebers and Willis, 2001). This is in agreement with transcriptome profiling data for several genes encoding members of this Pfam which indicates those transcripts are enriched in *An. coluzzii* (at the time solely classified as *An. gambiae*) chemosensory organs as compared to the rest of the body, especially in females (Pitts et al, 2011). In addition, several members of the single and paired EF-hand domain Pfams that play important regulatory and structural roles in cellular metabolism through this motifs role in binding intracellular calcium (Nelson et al, 2002) were found to be present in the proteome profile of the antennae and maxillary palps of the three species.

The remainder of Pfams above a representation threshold of 5 proteins/appendage/species are all associated with chemosensory processes. The GST_C (Glutathione S-transferase, C-terminal domain) is the most represented Pfam among the identified GST families, which includes important classes of enzymes that play an essential role in protection from oxidative stress and detoxification from xenobiotic compounds (Enayati et al., 2005). The transcripts of several GST genes have been found to be more abundant in chemosensory organs, compared to the rest of the body, in *An. coluzzii* (Pitts et al., 2011). In addition, three GST proteins are suggested to act as Odorant Degrading Enzymes (ODEs) in the olfactory system of in *Drosophila melanogaster* and other insects (Younus et al., 2014). High levels of these enzymes have been reported in insecticide resistant insects (Ding et al., 2005; Enayati et al., 2005) and elevated GST-activity phenotypes have been shown to affect mosquito longevity and vectorial capacity (Tripathy and Kar, 2015).

Several dehydrogenases (Pfam: short chain dehydrogenase) have also been reported as ODEs (Vogt, 2003). In addition, members of the Pheromone Binding Protein and General-Odorant Binding Protein families which are grouped together in the PBP_GOBP Pfam are also highly represented in both chemosensory organs, as also expected based on their recognized role in chemical perception (Pelosi et al., 2006) and transcript abundance (Pitts et al., 2011).

3.1 Differential protein abundance between antennae and palps

The large majority (95.5% in *An. coluzzii*, 94.1% in *An. arabiensis* and 89.2% in *An. quadriannulatus*) of the identified proteins are found both in the antennae and maxillary palps which comprise the primary and secondary chemosensory appendages in Anophelines, respectively. That said, our analysis indicates that in each species there are several notable proteins exclusively found in either the antennae or palps (Table S6). Of these, it is noteworthy that 3 OBPs (OBP2, OBP3 and OBP47) are antennal-specific in *An. coluzzii*, and one of these, OBP2, is also found exclusively in the antennae of *An. quadriannulatus*. Furthermore, the maxillary palps of all 3 anophelines exclusively contain OBP57 (Table S6).

Beyond those proteins found exclusively on either the antennae and maxillary palps, we looked for those that display significantly differential abundance. To measure the changes of relative protein levels between chemosensory appendages, we calculated the fold change, i.e. the ratio between the average LFQ values in antennae and palps (and vice versa). These results are summarized (Figures 1a, 2a, 3a) for Pfams that contain at least 2 proteins

identified only in palps or antennae or showing greater than 2-fold differential abundance. A list of the individual proteins found to be significantly different between the antennae and maxillary palps is reported in Table S7. Of these, 7 insect cuticle proteins are more abundant in palps than in antennae in *An. arabiensis*, while members of the same family show variable patterns of abundance between the two chemosensory organs in *An. coluzzii* and *An. quadriannulatus*. In addition, significant differences in the levels of several soluble olfactory proteins were detected.

Consistent with transcriptomic data available for *An. gambiae* (Pitts et al., 2011), the levels of several GST_C Pfam members are consistently more abundant in the antennae than in the palps, and in *An. coluzzii* one GST_C is exclusively found in the antennae. Furthermore, a number of PBP_GOBP proteins were differentially present in the two tissues in each species as well as chemosensory proteins (CSPs) belonging to the OS-D (insect pheromone-binding family A10/OS-D) family which are differentially abundant in *An. coluzzii* and *An. quadriannulatus*.

A more detailed comparison of the LFQ values calculated for olfactory soluble proteins belonging to PBP_GOBP and OS-D Pfams, as well as other antennal proteins that have been postulated to act as odorant carriers, are reported in Figures 1b, 2b, 3b. OBP1 and OBP9 are the most abundant soluble proteins found in the antennae of the 3 species, an observation that is consistent with results previously obtained from *An. coluzzii* (Mastrobuoni et al., 2013). OBP57 is exclusively observed in the maxillary palps, although traces of this protein were previously reported in *An. coluzzii* antennae (Mastrobuoni et al., 2013), probably due both to the higher number of specimens available for sample preparation as well as the higher instrument sensitivity. Similarly, other proteins previously found in trace amounts by Mastrobuoni et al. (e.g. OBP12), were not detected in the present analysis.

With the single exception of GR47, which was detected in the proteome profiles of the antennae and palps of *An. coluzzii* and *An. arabiensis* (Figures 1b, 2b, 3b), our analyses did not detect any chemosensory receptors. This is inconsistent with functional (Lu et al., 2007) and transcriptomic studies which identified a distinctive set of *Gr/Ir/Or* receptors (that did not include GR 47) in the maxillary palps as well as a much larger number of Or and Ir receptors in the antennae *An. coluzzii*, *Aedes aegypti* and other mosquitoes (Pitts et al., 2011; Rinker et al., 2013a; Zhou et al., 2014; Bohbot et al., 2014; Matthews et al., 2016). While the paucity of receptors primarily reflects the relatively mild extraction methods used here that marginalizes the representation of membrane components, it is consistent with other proteomic analyses of insect chemosensory organs, including species with large antennae that typically display broad and complex patterns of olfactory sensitivity (Zhao et al., 2015). From an odor coding perspective - and apart from the notable exceptions of the Orco co-receptor which is uniformly expressed across all antennal and maxillary palp ORNs in *An. gambiae* and the triad of Anopheline palpal GRs (*AgGrs* 22, 23, 24) that act as volatile CO₂ receptors expressed in a distinct population of sensory neurons that populate every capitata peg sensilla on the maxillary palps of these mosquitoes - this likely reflects the presence of large numbers of differently “tuned” chemoreceptors expressed across even larger numbers of receptor neurons resulting in low overall abundance levels of these transmembrane proteins.

3.2 Differential protein abundance between species

Data search for this analysis was based on a merged database of the three *An. gambiae* s.l. taxa. As this analysis uses a merged database of the three taxa, proteins exclusive to one or two species could include orthologous proteins assigned to different protein groups based on identified peptides containing a few divergent amino acids (razor+unique peptide). Since this could lead to an overestimation of exclusive proteins, Venn diagrams were not computed in the comparison between species. Furthermore, our analyses were based only on proteins identified with a threshold of at least 2 peptides in each replicate, which reduced the dataset from 502 to 168 proteins in the palps, and to 464 to 159 in the antennae. Proteins showing statistically significant differences are listed in Table 2 and graphically reported in Volcano plots (Figure S1). Results of t-tests are reported in Table S8.

Major differences in protein abundance are found between *An. coluzzii* and *An. quadriannulatus* antennae and between *An. arabiensis* and *An. quadriannulatus* palps. The comparison between *An. arabiensis* and *An. quadriannulatus* reveals that most inter-specific differences were found for cuticular proteins both in antennae and in palps, and for proteins involved in muscle contraction solely in the maxillary palps. The gene encoding for the cuticular protein RR-2 family (CPR125) -ACOM036100-PA-, which is more abundant in antennae of *An. quadriannulatus* than in *An. arabiensis*, has been reported to be an important candidate target of selection in *An. coluzzii*, as a result of its position on the region of the X-chromosome which has undergone a recent insecticide-resistance gene introgression from sister species *An. gambiae* (Main et al., 2015).

Figure S2a and S2b reports unsupervised hierarchical clustering respectively for antennae and palps, of the three species. In both analyses, each sample replicate cluster together underscoring their integrity. For antennal samples, this analysis revealed that the chemosensory proteome profile of *An. arabiensis* was distinct from that of *An. coluzzii* and *An. quadriannulatus*, which can be discriminated from each other based on their resolution into two separate clusters on the same branch. In contrast, in the maxillary palp proteome profiles - where, in light of the uniform population of chemosensory sensilla, the complexity of appendage-specific proteins with each species is expected to be lower than in antennae (Lu et al., 2007) - *An. quadriannulatus* clusters separately from the branch of *An. coluzzii* and *An. arabiensis*. This is likely due a consequence of the *An. quadriannulatus* maxillary palp proteome having the fewest proteins in common with *An. coluzzii* and *An. arabiensis*.

The heat map representation of the antennae clustered matrix (Figure 4, panel A) - previously filtered for proteins belonging to Pfam reported to be involved in olfaction (i.e. PBP_GOBP, OS-D, GST and dehydrogenase) either as odorant carrier or as Odorant Degrading Enzymes (Pelosi et al., 2006; Younus et al., 2014) - showed that *An. coluzzii* and *An. arabiensis* are resolved into two separate clusters on the same branch distinct from that of *An. quadriannulatus*. The same result was obtained for maxillary palps (Figure 4, panel B), where *An. quadriannulatus* proteome profile was separated from those of the other two species, that in this case are not separated in two distinct clades.

In conclusion, this study provides the first wide scale quantitative proteomic characterization of the main chemosensory appendages (i.e. antennae and maxillary palps) of host-seeking

adult mosquito females from three members of the *An. gambiae* complex characterized by different host-preferences.

The data reported here significantly complements results from transcriptome profiling studies insofar as soluble proteins in those tissues. Significant differences are observed in the presence and abundance of several groups of proteins between these sensory appendages. Moreover, our observation that the protein abundance pattern differs between the three species, supports the hypothesis that molecular components underlying peripheral mechanisms of olfactory signal transduction may directly contribute to the distinctive odor mediated behaviors known for *Anopheles*, including search for different host species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This study was funded by “Futuro in Ricerca 2010” grant to BC (Grant No. RBFR106NTE) and by AWARD 2013 projects (SAPIENZA University) to AdT. We thank Elena Michelucci for technical assistance during LC-MS/MS analyses and Gemma Paciotti for great help in mosquitoes breeding and dissections. We are grateful to Willem Takken, Jeroen Spitzen and Leon Westerd from Wageningen University for providing *An. arabiensis* and *An. quadriannulatus* eggs from their colonies.

REFERENCES

- Bohbot J, Pitts RJ, Kwon HW, Rutzler M, Robertson HM, Zwiebel LJ, 2007 Molecular characterization of the *Aedes aegypti* odorant receptor gene family. *Insect Mol Biol* 16, 525–537. [PubMed: 17635615]
- Bohbot JD, Sparks JT, Dickens JC, 2014 The maxillary palp of *Aedes aegypti*, a model of multisensory integration. *Insect Biochem Mol Biol* 48, 29–39. [PubMed: 24613607]
- Bradford MM, 1976 Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. [PubMed: 942051]
- Carey AF, Wang G, Su CY, Zwiebel LJ, Carlson JR, 2010 Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* 464, 66–71. [PubMed: 20130575]
- Costantini C, Sagnon NF, della Torre A, Diallo M, Brady J, Gibson G, Coluzzi M 1998 Odor-mediated host preferences of West African mosquitoes, with particular reference to malaria vectors. *Am J Trop Med Hyg.* 58, 56–63. [PubMed: 9452293]
- Cox J, Mann M, 2008 MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. [PubMed: 19029910]
- Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M, 2011 Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res* 10, 1794–1805. [PubMed: 21254760]
- Dekker T, Takken W, Braks MAH, 2001 Innate preference for host-odour blends modulates degree of anthropophagy of *Anopheles gambiae* sensu lato (Diptera: Culicidae). *J Med Entomol* 38, 868–871. [PubMed: 11761386]
- Ding Y, Hawkes N, Meredith J, Eggleston P, Hemingway J, Ranson H, 2005 Characterization of the promoters of Epsilon glutathione transferases in the mosquito *Anopheles gambiae* and their response to oxidative stress. *Biochem J* 387, 879–88. [PubMed: 15631620]
- Enayati AA, Ranson H, Hemingway J, 2005 Insect glutathione transferases and insecticide resistance. *Insect Mol Biol* 14, 3–8. [PubMed: 15663770]

- Foster LJ, De Hoog CL, Mann M, 2003 Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc Natl Acad Sci U S A.* 100, 5813–5818. [PubMed: 12724530]
- Guidobaldi F, May-Concha IJ, Guerenstein PG, 2014 Morphology and physiology of the olfactory system of blood-feeding insects. *J Physiol Paris* 108, 96–111. [PubMed: 24836537]
- Habtewold T, Povelones M, Blagborough AM, Christophides GK, 2008 Transmission blocking immunity in the malaria non-vector mosquito *Anopheles quadriannulatus* species A. *PLoS Pathog* 4(5):e1000070. [PubMed: 18497855]
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, 2002 The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298, 129–149. [PubMed: 12364791]
- Iovinella I, Caputo B, Michelucci E, Dani FR, Della Torre A, 2015 Candidate biomarkers for mosquito age-grading identified by label-free quantitative analysis of protein expression in *Aedes albopictus* females. *J Proteomics* 128, 272–279. [PubMed: 26271156]
- Kwon HW, Lu T, Rutzler M, Zwiebel LJ, 2006 Olfactory responses in a gustatory organ of the malaria vector mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* 103, 13526–13531. [PubMed: 16938890]
- Larter NK, Sun JS, Carlson JR, 2016 Organization and function of *Drosophila* odorant binding proteins. *eLife* 5: e20242. [PubMed: 27845621]
- Leal WS, 2013 Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annual Review of Entomology* 58, 373–391.
- Lu T, Qiu YT, Wang G, Kwon JY, Rutzler M, Kwon HW, Pitts RJ, van Loon JJ, Takken W, Carlson JR, et al. 2007 Odor coding in the maxillary palp of the malaria vector mosquito *Anopheles gambiae*. *Curr Biol* 17, 1533–1544. [PubMed: 17764944]
- Lutz EK, Lahondere C, Vinauger C, Riffell JA 2017 Olfactory learning and chemical ecology of olfaction in disease vector mosquitoes: a life history perspective. *Curr Opin Insect Sci* 20, 75–83 [PubMed: 28602240]
- Main BJ, Lee Y, Collier TC, Norris LC, Brisco K, Fofana A, Cornel AJ, Lanzaro GC, 2015 Complex genome evolution in *Anopheles coluzzii* associated with increased insecticide usage in Mali. *Mol Ecol* 24, 5145–57. [PubMed: 26359110]
- Manoharan M, Ng Fuk Chong M, Vařtinadapoulé A, Frumence E, Sowdhamini R, Offmann B, 2013 Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. *Genome Biol Evol* 5, 163–80. [PubMed: 23292137]
- Mastrobuoni G, Qiao H, Iovinella I, Sagona S, Niccolini A, Boscaro F et al., 2013 A proteomic investigation of soluble olfactory proteins in *Anopheles gambiae*. *PLoS ONE* 8:e75162. [PubMed: 24282496]
- Matthews BJ, McBride CS, DeGennaro M, Despo O, Vossall LB, 2016 The neurotranscriptome of the *Aedes aegypti* mosquito. *BMC Genomics* 17, 32. [PubMed: 26738925]
- Montell and Zwiebel, 2016 Mosquito Sensory Systems, in: Elsevier, *Advances in Insect Physiology*. Vol 51, pp 294–328. doi: 10.1016/bs.aip.2016.04.007
- Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, et al., 2015 Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes. *Science* 347, 1258522–1258529. [PubMed: 25554792]
- Nelson MR, Thulin E, Fagan PA, Forsén S, Chazin WJ, 2002 The EF-hand domain: a globally cooperative structural unit. *Protein Sci.* 2002 11, 198–205.
- Pates HV, Takken W, Curtis CF, Jamet H, 2006 Zoophilic *Anopheles quadriannulatus* species B found in a human habitation in Ethiopia. *Ann Trop Med Parasitol* 100, 177–9. [PubMed: 16492366]
- Pelosi P, Zhou JJ, Ban LP, Calvellido M (2006). Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences* 63, 1658–1676. [PubMed: 16786224]
- Pelosi P, Iovinella I, Felicioli A, Dani FR, 2014 Soluble proteins of chemical communication: an overview across arthropods. *Front Physiol* 5, 320. [PubMed: 25221516]
- Pitts RJ, Rinker DC, Jones PL, Rokas A, Zwiebel LJ, 2011 Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. *BMC Genomics* 12, 271. [PubMed: 21619637]

- Pitts RJ, Fox AN, Zwiebel LJ, 2004 A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector *Anopheles gambiae*. *Proc Natl Acad Sci U S A* 101, 5058–5063. [PubMed: 15037749]
- Rappsilber J, Mann M, Ishihama Y, 2007 Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc.* 2, 1896–1906. [PubMed: 17703201]
- Rebers JE, Willis JH, 2001 A conserved domain in arthropod cuticular proteins binds chitin. *Insect Biochem Mol Biol* 31, 1083–1093. [PubMed: 11520687]
- Rinker DC, Zhou X, Pitts RJ, AGC Consortium, Rokas A, Zwiebel LJ, 2013a Antennal transcriptome profiles of anopheline mosquitoes reveal human host Olfactory specialization in *Anopheles gambiae*. *BMC Genomics* 14, 749. [PubMed: 24182346]
- Rinker DC, Pitts RJ, Zhou X, Suh E, Rokas A, Zwiebel LJ, 2013b Blood meal-induced changes to antennal transcriptome profiles reveal shifts in odor sensitivities in *Anopheles gambiae*. *Proc Natl Acad Sci U S A* 110, 8260–8265. [PubMed: 23630291]
- Rinker DC, Pitts RJ, Zwiebel LJ, 2016 Disease vectors in the era of next generation sequencing. *Genome Biol* 17, 95. [PubMed: 27154554]
- Rund SS, Bonar NA, Champion MM, Ghazi JP, Houk CM, Leming MT, Syed Z, Duffield GE, 2013 Daily rhythms in antennal protein and olfactory sensitivity in the malaria mosquito *Anopheles gambiae*. *Sci Rep* 3, 2494. [PubMed: 23986098]
- Scott TW, Takken W, 2012 Feeding strategies of anthropophilic mosquitoes result in increased risk of pathogen transmission. *Trends Parasitol* 28, 114–121. [PubMed: 22300806]
- Steinbrecht R, 1997 Pore structures in insect olfactory sensilla: A review of data and concepts. *Int J Insect Morphol* 26, 229–245.
- Su CY, Menuz K, Carlson JR, 2009 Olfactory perception: receptors, cells, and circuits. *Cell* 139, 45–59. [PubMed: 19804753]
- Suh E, Bohbot J, Zwiebel LJ, 2014 Peripheral olfactory signaling in insects. *Curr Opin Insect Sci* 6, 86–92. [PubMed: 25584200]
- Takken W, Verhulst NO, 2013 Host Preferences of Blood-Feeding Mosquitoes. *Annual Review of Entomology* 58, 433–453.
- Takken W, Knols BG, 1999 Odor-mediated behavior of Afrotropical malaria mosquitoes. *Annu. Rev. Entomol.* 44, 131–157. [PubMed: 9990718]
- Takken W, Eling W, Hooghof J, Dekker T, Hunt R, Coetzee M, 1999 Susceptibility of *Anopheles quadriannulatus theobaldi* (Diptera: Culicidae) to *Plasmodium falciparum*. *Trans R Soc Trop Med Hyg* 93, 578–580. [PubMed: 10717736]
- Tirados I, Costantini C, Gibson G, Torr SJ 2006 Blood-feeding behaviour of the malarial mosquito *Anopheles arabiensis*: implications for vector control. *Med Vet Entomol* 20, 425–37. [PubMed: 17199754]
- Tripathy A, Kar SK, 2015 Feeding stage, species, body part and sex-specific activity of Glutathione S-transferase in mosquito. *Trop Biomed* 32, 65–75. [PubMed: 25801255]
- Tusher VG, Tibshirani R, Chu G, 2001 Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116–21. [PubMed: 11309499]
- Vogt RG (2003) in *Insect Pheromone Biochemistry and Molecular Biology: The Biosynthesis and Detection of Pheromones and Plant Volatiles*, eds. Bomquist GJ & Vogt RG (Elsevier Academic, London), pp. 391–445.
- Wang G, Carey AF, Carlson JR, Zwiebel LJ, 2010 Molecular basis of odor coding in the malaria vector mosquito *Anopheles gambiae*. *Proc Natl Acad Sci U S A* 107, 4418–23. [PubMed: 20160092]
- Younus F, Chertemps T, Pearce SL, Pandey G, Bozzolan F, Coppin CW, Russell RJ, Maibèche-Coisne M, Oakeshott JG, 2014 Identification of candidate odorant degrading gene/enzyme systems in the antennal transcriptome of *Drosophila melanogaster*. *Insect Biochem Mol Biol* 53, 30–43. [PubMed: 25038463]
- Zhao Y, Li H, Miao X, 2015 Proteomic Analysis of Silkworm Antennae. *J Chem Ecol.* 41, 1037–42. [PubMed: 26515890]

- Zhou Y, Badgett MJ, Bowen JH, Vannini L, Orlando R, Willis JH, 2016 Distribution of cuticular proteins in different structures of adult *Anopheles gambiae*. *Insect Biochem Mol Biol* 75, 45–57. [PubMed: 27179905]
- Zhou X, Rinker DC, Pitts RJ, Rokas A, Zwiebel LJ, 2014 Divergent and conserved elements comprise the chemoreceptive repertoire of the nonblood-feeding mosquito *Toxorhynchites amboinensis*. *Genome Biol Evol* 6, 2883–2896. [PubMed: 25326137]
- Zwiebel LJ, Takken W, 2004 Olfactory regulation of mosquito-host interactions. *Insect Biochem Mol Biol* 34, 645–652. [PubMed: 15242705]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1. Bar chart of Pfam and olfactory proteins in *An. coluzzii*.

A) Bar charts of protein families (Pfam) containing at least 2 proteins among those exclusively identified in *Anopheles coluzzii* palps or antennae or showing a fold change higher than 2 (2FC); B) Bar charts showing $\log_{10}$ LFQ (Label-free quantification) intensities of the identified olfactory proteins. Protein names are assigned according to Uniprot gene name. Accession number is reported for D7-related proteins (D7r), putative antennal proteins and salivary proteins. Proteins significantly different at t-test ($p < 0.05$) are indicated with an asterisk.

Figure 2. Bar chart of Pfam and olfactory proteins in *An. arabiensis*.

A) Bar charts of protein families (Pfam) containing at least 2 proteins among those exclusively identified in *Anopheles arabiensis* palps or antennae or showing a fold change higher than 2 (2FC); B) Bar charts showing $\log_{10}$ LFQ (Label-free quantification) intensities of the identified olfactory proteins. Accession number is reported for D7-related proteins (D7r), putative antennal proteins and salivary proteins. Proteins significantly different at t-test ($p < 0.05$) are indicated with an asterisk.

Figure 3. Bar chart of Pfam and olfactory proteins in *An. quadriannulatus*.

A) Bar charts of protein families (Pfam) containing at least 2 proteins among those exclusively identified in *Anopheles quadriannulatus* palps or antennae or showing a fold change higher than 2 (2FC); B) Bar charts showing $\log_{10}$ LFQ (Label-free quantification) intensities of the identified olfactory proteins. Accession number is reported for D7-related proteins (D7r), putative antennal proteins and salivary proteins. Proteins significantly different at t-test ($p < 0.05$) are indicated with an asterisk.

Panel B

Figure 4. Cluster of olfactory proteins.

Unsupervised hierarchical clustering for antennae (panel A) and palps (panel B) of 3 species of the *Anopheles gambiae* complex based on LFQ (Label-free quantification) values of proteins, identified with at least 2 “Unique + Razor” peptides, belonging to Pfams reported to be involved in olfaction (PBP_GOBP, OS-D, GST and adh_short). For each proteins we reported ID and Pfam. Colour scale reports log2 transformed LFQ intensity values.

Table 1.
Pfam of proteins identified in antennae and palps of the three species.

List of protein families (Pfams) containing at least 3 members in palps or antennae of one of the 3 species of the *Anopheles gambiae* complex. For each tissue we report the number of proteins belonging to the relative Pfam and their percentage with respect to the total number of identified proteins.

	<i>An. coluzzi</i>				<i>An. arabiensis</i>				<i>An. quadriannulatus</i>			
	Antennae		Palps		Antennae		Palps		Antennae		Palps	
	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%
Insect cuticle protein	30	6,42	30	6,26	29	8,19	31	8,24	33	10,38	30	11,19
Glutathione S-transferase, C-terminal domain	10	2,14	9	1,88	6	1,69	7	1,86	4	1,26	3	1,12
PBP/GOBP family	10	2,14	7	1,46	8	2,26	8	2,13	5	1,57	4	1,49
EF-hand domain pair	9	1,93	10	2,09	7	1,98	9	2,39	8	2,52	4	1,49
Short chain dehydrogenase	6	1,28	5	1,04	1	0,28	2	0,53	1	0,31	0	0
Aldehyde dehydrogenase family	6	1,28	6	1,25	2	0,56	3	0,8	5	1,57	2	0,75
Glutathione S-transferase, N-terminal domain	6	1,28	5	1,04	3	0,85	4	1,06	3	0,94	2	0,75
Trypsin	6	1,28	9	1,88	5	1,41	7	1,86	2	0,63	2	0,75
Glutathione S-transferase, N-terminal domain	5	1,07	5	1,04	4	1,13	4	1,06	2	0,63	2	0,75
C-terminal domain of 1-Cys peroxiredoxin	4	0,86	4	0,84	3	0,85	3	0,8	2	0,63	2	0,75
AhpC/TSA family	4	0,86	4	0,84	3	0,85	3	0,8	2	0,63	2	0,75
ATP synthase alpha/beta family, nucleotide-binding domain	4	0,86	4	0,84	4	1,13	4	1,06	4	1,26	4	1,49
ATP synthase alpha/beta family, beta-barrel domain	4	0,86	4	0,84	4	1,13	4	1,06	4	1,26	4	1,49
Biotin-requiring enzyme	4	0,86	4	0,84	2	0,56	2	0,53	0	0	0	0
Calponin homology (CH) domain	4	0,86	5	1,04	4	1,13	4	1,06	1	0,31	1	0,37
Regulatory CLIP domain of proteinases	4	0,86	6	1,25	2	0,56	4	1,06	1	0,31	1	0,37
EF-hand domain	4	0,86	3	0,63	3	0,85	3	0,8	1	0,31	1	0,37
Hsp20/alpha crystallin family	4	0,86	4	0,84	3	0,85	4	1,06	4	1,26	3	1,12
Hsp70 protein	4	0,86	5	1,04	3	0,85	3	0,8	2	0,63	3	1,12

	<i>An. coluzzi</i>				<i>An. arabiensis</i>				<i>An. quadriannulatus</i>			
	Antennae		Palps		Antennae		Palps		Antennae		Palps	
	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%
Insect pheromone-binding family, A10/OS-D	4	0,86	3	0,63	3	0,85	3	0,8	4	1,26	3	1,12
Papain family cysteine protease	4	0,86	4	0,84	3	0,85	3	0,8	2	0,63	2	0,75
Proteasome subunit	4	0,86	3	0,63	2	0,56	1	0,27	4	1,26	3	1,12
Thiolase, C-terminal domain	4	0,86	5	1,04	1	0,28	1	0,27	2	0,63	0	0
Thiolase, N-terminal domain	4	0,86	5	1,04	1	0,28	1	0,27	2	0,63	0	0
Thioredoxin	4	0,86	5	1,04	5	1,41	5	1,33	3	0,94	3	1,12
2-Oxoacid dehydrogenases acyltransferase (catalytic domain)	3	0,64	3	0,63	2	0,56	2	0,53	0	0	0	0
Actin	3	0,64	2	0,42	3	0,85	3	0,8	1	0,31	1	0,37
Aldo/keto reductase family	3	0,64	3	0,63	2	0,56	2	0,53	3	0,94	2	0,75
ATP synthase alpha/beta chain, C terminal domain	3	0,64	3	0,63	3	0,85	3	0,8	3	0,94	3	1,12
Pupal cuticle protein C1	3	0,64	3	0,63	3	0,85	3	0,8	3	0,94	2	0,75
Dehydrogenase E1 component	3	0,64	3	0,63	1	0,28	2	0,53	1	0,31	1	0,37
Immunoglobulin I-set domain	3	0,64	3	0,63	2	0,56	1	0,27	0	0	0	0
Major royal jelly protein	3	0,64	3	0,63	0	0	2	0,53	1	0,31	1	0,37
Insulinase (peptidase family M16)	3	0,64	3	0,63	2	0,56	2	0,53	2	0,63	2	0,75
Peptidase M16 inactive domain	3	0,64	3	0,63	2	0,56	2	0,53	2	0,63	2	0,75
Proteasome subunit A N-terminal signature	3	0,64	2	0,42	2	0,56	1	0,27	3	0,94	2	0,75
Drosophila retinilin like protein	3	0,64	3	0,63	1	0,28	2	0,53	2	0,63	2	0,75
Serpin	3	0,64	4	0,84	2	0,56	3	0,8	0	0	1	0,37
Copper/zinc superoxide dismutase (SODC)	3	0,64	3	0,63	3	0,85	3	0,8	3	0,94	3	1,12
Spectrin repeat	3	0,64	4	0,84	2	0,56	2	0,53	1	0,31	0	0
Transketolase, pyrimidine binding domain	3	0,64	3	0,63	1	0,28	1	0,27	1	0,31	0	0
Tropomyosin	3	0,64	3	0,63	3	0,85	2	0,53	2	0,63	2	0,75
Tubulin	3	0,64	3	0,63	2	0,56	2	0,53	3	0,94	3	1,12

	<i>An. coluzzi</i>				<i>An. arabiensis</i>				<i>An. quadriannulatus</i>			
	Antennae		Palps		Antennae		Palps		Antennae		Palps	
	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%	nr proteins	%
Tubulin C-terminal domain	3	0,64	3	0,63	2	0,56	2	0,53	3	0,94	3	1,12
Tim10/DDP family zinc finger	3	0,64	3	0,63	3	0,85	3	0,8	1	0,31	1	0,37
Calcineurin-like phosphoesterase	2	0,43	3	0,63	1	0,28	1	0,27	1	0,31	1	0,37
WD domain, G-beta repeat	2	0,43	2	0,42	2	0,56	3	0,8	4	1,26	1	0,37
Core histone H2A/H2B/H3/H4	2	0,43	2	0,42	2	0,56	1	0,27	3	0,94	3	1,12
RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)												

Table 2.
T-test between antennae and palps.

List of proteins significantly different (t-test) between antennae or palps of 3 species of the *Anopheles gambiae* complex.

	Protein ID	Description	Comparison
Antennae	ACOM031698-PA	ATP carrier protein	<i>An. arabiensis</i> > <i>An.coluzzii</i>
	Q7PNP1	Cuticular protein 3	<i>An. coluzzii</i> > <i>An.arabiensis</i>
	ACOM025349-PA	Alpha-crystallin chain A	<i>An. arabiensis</i> > <i>An.quadrinnulatus</i>
	Q7PXZ2	Cuticular protein RR-2 family (CPR6)	<i>An. quadrinnulatus</i> > <i>An.arabiensis</i>
	ACOM036100-PA	Cuticular protein RR-2 family (CPR125)	<i>An. quadrinnulatus</i> > <i>An. coluzzii</i>
	AQUA010281-PA	Cuticular protein RR-2 family (CPR123)	<i>An. quadrinnulatus</i> > <i>An. coluzzii</i>
	AARA010500-PA	Acyl-CoA-binding protein	
	ACOM025349-PA	Alpha-crystallin chain A	<i>An.coluzzii</i> > <i>An.quadrinnulatus</i>
	ACOM026599-PA	ATP synthase subunit beta	
	ACOM030749-PA	Sarcoplasmic calcium-binding protein	
Palps	ACOM026924-PA	Cuticular protein CPLCP12	<i>An. arabiensis</i> > <i>An. coluzzii</i>
	AARA002622-PA	Cuticular protein RR-1 family (CPR127)	
	ACOM035844-PA	Chitin binding Peritrophin-A domain	<i>An.quadriannulatus</i> > <i>An.arabiensis</i>
	Q7PIX0	Cuticular protein RR-2 family (CPR125)	
	ACOM032061-PA	Cathepsin L	
	ACOM039137-PA	Cuticular protein RR-1 family (CPR16)	
	T1DGR6	Putative tropomyosin-2	
	Q7Q270	Putative tropomyosin-2	<i>An.arabiensis</i> > <i>An.quadriannulatus</i>
	ACOM036920-PA	Secreted ferritin G subunit	
	F5HME4	Putative tropomyosin-1	
	Q7Q978	Paramyosin	
	F5HME4	Putative tropomyosin-1	<i>An. coluzzii</i> > <i>quadriannulatus</i>
	Q7Q978	Paramyosin	
	AARA002622-PA	Cuticular protein RR-1 family (CPR127)	<i>An. quadriannulatus</i> > <i>An. coluzzii</i>