Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Appendage identity dominates chemosensory transcriptomic organization in the black soldier fly Hermetia illucens

Dor Perets
The Hebrew University of Jerusalem
Esther Yakir
The Hebrew University of Jerusalem
Philippos A. Papathanos
The Hebrew University of Jerusalem
Jonathan D. Bohbot
jonathan.bohbot@mail.huji.ac.il
The Hebrew University of Jerusalem

Research Article

Keywords: Hermetia illucens, black soldier fly, RNA-seq, transcriptomics, chemosensory appendages, sex-biased gene expression, mating, odorant receptors, Diptera, gene expression atlas

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11693501/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Hermetia illucens is a commercially important dipteran whose adult reproductive behaviors rely on chemosensory cues for mate finding and oviposition-site selection. How peripheral chemosensory gene expression is organized across adult sensory appendages, and how it is modulated by sex and reproductive state, remains unknown.

Results

We generated RNA-seq profiles of antennae, maxillary palps and tarsi from virgin males, virgin females and mated females. Appendage identity was the dominant axis of both global and chemosensory transcriptomic organization, whereas sex and mating status had modest effects involving genes associated with sex determination, circadian regulation and neuropeptide signaling. Sex-biased chemosensory gene expression was concentrated in the antennae, where female-biased odorant receptors predominated, and no chemoreceptor responded to mating in any appendage. Phylogenetic analysis of 567 transcript models (393 loci) in seven families (*Or*, *Gr*, *Ir*, *Obp*, *Csp*, *Ppk* and *Trp*) provided a curated catalog of the *H. illucens* chemosensory repertoire. The five gustatory receptors in the conserved dipteran CO_2 receptor clade were divided between appendages, three enriched in the maxillary palp as in mosquitoes and two in the antenna as in *Drosophila melanogaster*, so *H. illucens* presents both dipteran arrangements.

Conclusions

Adult chemosensory transcription in *H. illucens* is organized primarily by appendage, with sex and reproductive state acting through subtle regulatory changes. We propose a hierarchical model in which appendage-specific sensory identities were established early in dipteran evolution and receptor families diversified within these modules. The chemosensory catalog, transcriptomic atlas and data browser provide a foundation for comparative and functional studies across Diptera.

Background

Hermetia illucens (Diptera: Stratiomyidae) is a cosmopolitan saprophagous species whose ecological success is rooted in its ability to exploit microbe-rich decaying substrates [1, 2]. Beyond its commercial importance in waste bioconversion and sustainable protein production [2, 3], *H. illucens* is an early-diverging brachyceran [4] and therefore a useful comparative point for distinguishing ancestral features of dipteran peripheral chemosensation from innovations that emerged later during brachyceran evolution. Combined with data from other dipteran lineages, it allows us to ask whether appendage-specific sensory organization, reproductive plasticity, and chemosensory gene repertoires are deeply

conserved features of Diptera or innovations associated with brachyceran diversification. Recent genome assemblies, including a chromosome-level assembly [5, 6], syntheses of the genomic data [7], and evidence for lineage-specific expansions of gene families involved in digestion, detoxification, immunity, and chemosensation [5, 8], now provide the genomic resources needed to address these questions.

The adult life stage of *H. illucens* is brief [9] and behaviorally focused on reproduction, although adults retain a functional digestive system and can feed [10]. In this species mate location is light-dependent and lek-based [11, 12], and females respond behaviorally to substrate- and oviposition-associated cues [13]. In other Diptera, chemosensation additionally mediates close-range courtship and mate recognition through contact pheromones [14, 15] and guides gravid females to oviposition sites via microbial and substrate-derived volatiles [16], indicating that chemosensory cues shape multiple aspects of adult reproductive behavior. Electrophysiological recordings show that the antennae of both sexes respond to volatiles emitted by decaying organic matter of the kind used as larval rearing substrate [17], consistent with chemosensory tuning toward microbe-associated habitats. In cage assays, mated females perch preferentially on plant-based oviposition substrates [13], indicating a behavioral transition that accompanies the change in reproductive state. The antennae and maxillary palps, two of the principal chemosensory appendages alongside the tarsi and the labellum, are sexually dimorphic: the antennae differ morphometrically between the sexes [18], and female maxillary palps are larger and carry more sensory pits than those of males [19], suggesting differences in peripheral sensory sampling between the sexes.

In dipteran insects, the molecular basis of sex- and reproductive-state-dependent chemosensory function has been characterized in several species [20–23]. Sex-biased expression of odorant-binding proteins, chemosensory proteins, odorant receptors and ionotropic receptors has been documented in the antennae of several species, including *H. illucens*, reflecting the distinct sensory demands of courtship, mate recognition, host seeking, and oviposition [24–26]. Superimposed on these stable sex differences, reproductive state also remodels peripheral sensory function. In female mosquitoes, blood feeding induces widespread antennal transcriptional reprogramming, including changes in odorant receptor expression associated with host-seeking behavior [20]. Likewise, mating suppresses pheromone attraction in female *Drosophila melanogaster* through modulation of olfactory sensory neuron activity [21]. Together, these studies indicate that peripheral chemosensory systems are dynamically regulated to match changing reproductive and ecological demands.

Across insects, peripheral chemosensation is mediated by seven conserved gene families: odorant receptors (*Ors*), gustatory receptors (*Grs*), ionotropic receptors (*Irs*), odorant-binding proteins (*Obps*), chemosensory proteins (*Csps*), pickpocket channels (*Ppks*), and transient receptor potential channels (*Trps*). These families fulfill complementary roles in odorant transport, ligand detection, signal transduction, and thermo-, mechano-, and contact chemosensation, and exhibit characteristic distributions across sensory appendages [27–32]. Although genomic resources for *H. illucens* have expanded substantially in recent years [6–8], including a chromosome-level genome assembly and the

identification of several chemosensory gene families, the transcriptional organization of the adult chemosensory system remains largely unexplored. Previous studies have focused on antennal transcriptomes or larval physiology [26, 33–35], leaving unresolved how the complete repertoire of chemosensory genes is partitioned among adult appendages and how its expression is modulated by sex and reproductive state.

In other insects, mating modulates chemosensory gene expression in peripheral sensory organs. Mating increases sex peptide receptor expression in *Drosophila* olfactory and gustatory neurons, and experimentally elevating this receptor in virgin females is sufficient to shift their polyamine preference to that of mated females [36]. Mating-associated changes in chemosensory gene expression have been reported across insects [37]. We therefore hypothesized that the transition from mate searching to post-mating oviposition would be accompanied by substantial remodeling of peripheral chemosensory gene expression, producing pronounced transcriptional differences between sexes and reproductive states within each appendage. To test this hypothesis, we generated RNA-seq profiles from the antennae, maxillary palps, and tarsi of virgin males, virgin females, and mated females. This experimental design enabled a comprehensive comparison of appendage identity, sex, and reproductive state while establishing a curated catalog of 567 chemosensory transcript models (393 gene loci) spanning all seven major gene families together with phylogenetic annotation and appendage-specific expression profiles.

Appendage identity was the dominant axis of transcriptomic organization, whereas sex and mating status produced comparatively modest but biologically coherent changes involving conserved pathways associated with sex determination, circadian regulation, neuropeptide signaling, and reproductive physiology. Together with the *H. illucens* Olfactome Browser, a data browser that readers can run on their own computer (see Availability of data and materials), these resources provide an appendage-resolved atlas of the *H. illucens* chemosensory system, extending earlier de novo transcriptome surveys of this species [26] to defined sensory appendages and reproductive states, and provide a foundation for comparative and functional studies of how peripheral chemosensory systems diversified during the early evolution of the Brachycera.

Results

The H. illucens chemosensory repertoire comprises 393 gene loci across seven families

The three adult chemosensory appendages analyzed in this study, the antennae, maxillary palps, and tarsi (Fig. 1A), provide the principal sites for volatile, short-range, and contact chemosensation in adult *H. illucens*. Because *H. illucens* occupies an early-diverging position within the Brachycera [4] (Fig. 1B), it is well placed for examining the evolution of dipteran chemosensory systems.

To establish a comprehensive chemosensory gene repertoire, maximum-likelihood phylogenetic analyses were performed on 567 annotated *H. illucens* proteins spanning seven conserved gene families (OR, GR, IR, OBP, CSP, PPK and TRP), reconstructed together with homologous sequences from *Aedes*

aegypti, *Drosophila melanogaster*, *Musca domestica*, *Calliphora stygia*, *Bactrocera dorsalis* and *Glossina morsitans*. A circular summary tree illustrates the distribution of all seven families (Fig. 1C), "while the complete tree and the detailed tree of each family are provided in Additional file 1: Figs. S1–S8.

The 567 annotated *H. illucens* chemosensory transcript models correspond to 393 distinct gene loci: 148 *Or*, 83 *Ir*, 65 *Obp*, 35 *Ppk*, 33 *Trp*, 23 *Gr* and 6 *Csp* (Table 1; **Additional file 2**). Compared on a locus basis with other dipterans, the *Trp* family shows the most consistent expansion, being 2.4-fold, 3.0-fold and 2.8-fold larger than in *D. melanogaster*, *Ae. aegypti* and *M. domestica*, respectively, whereas the *Or* repertoire is 2.5-fold and 2.2-fold larger than in *D. melanogaster* (60 genes) and *M. domestica* and only 1.3-fold larger than in *Ae. aegypti* (148 versus 117). The *Gr* repertoire is reduced against all three comparators (23 versus 60, 72 and 46), and the *Ir* repertoire (83 loci) is larger than in *D. melanogaster* and *M. domestica* (66 and 77) but smaller than in *Ae. aegypti* (135). The *Trp* comparison should be read with caution, because our *Trp* set includes subfamilies that are not always counted as TRPs elsewhere (see Limitations). Family sizes across species are given in **Additional file 3**, whose legend states the counting basis of each published value [29, 38–42]. Orthology analyses identified conserved co-receptors and receptor orthologs of known function across all seven families (**Additional files 4–10**). Within the OR family, the two RefSeq transcript models annotated as HillORco and HillORco2 belong to a single gene locus (LOC119658571) and group with the conserved insect *Orco* lineage [43] (bootstrap 98). HillOR179 and HillOR180 fall within the indolergic receptor (indolOR) clade, an 18-sequence group recovered with full support that contains the mosquito receptors AaegOR2, AaegOR9 and AaegOR10 [44, 45] together with DmelOR30a, DmelOR43a, DmelOR49b and their *Calliphora* and *Musca* counterparts [46, 47]; the two *H. illucens* genes form their own two-gene lineage (bootstrap 100) sister to the higher-brachyceran receptors (**Additional file 1: Fig. S6**; **Additional file 11**). HillOR179 was expressed predominantly in the antennae, with lower expression detected in the tarsi. Within the IR family, HillIR7 was recovered as the ortholog of the obligate co-receptor *Ir8a* (nearest outgroup sequences DmelIR8a and CstyIR8a; bootstrap 86) and HillIR3 as a 1:1 ortholog of the co-receptor *Ir25a* [48] (nearest outgroup AaegIR25a; bootstrap 95), while the remaining IRs fall within the broader ionotropic receptor radiation. The tree does not resolve, with adequate bootstrap support, how these remaining IRs divide into the conserved antennal IRs and the divergent, largely gustatory IRs, so we do not assign counts to the two groups; per-gene nearest-outgroup assignments are given in **Additional file 6**. Within the GR family, HillGR1–HillGR5 fall inside a 10-sequence clade recovered with full support (bootstrap 100) that also contains the characterized dipteran CO_2 receptors DmelGR21a and DmelGR63a [49, 50] and the *Ae. aegypti* receptors AaegGR1–AaegGR3 [51, 52], with DmelGR59e and DmelGR59f as its sister group (**Additional file 11**); the *H. illucens* genes are therefore interdigitated with the CO_2 receptor lineage rather than outgroup to it. HillGR40 was identified as the ortholog of the DmelGR43a fructose receptor [53]. Expression of the five CO_2-clade receptors was divided between appendages: *Gr4* was maxillary palp-specific (mean normalized counts 2,493 in the palp against 2.6 and 2.8 in the antenna and tarsi), *Gr2* and *Gr5* were palp-biased (1,251 and 2,227 against 12 and 785 in the antenna), whereas *Gr1* and *Gr3* were antenna-biased (1,479 and 956 against 318 and 204 in the palp; adjusted $p \le 4.9 \times 10^{-10}$ in every contrast). All five were near-absent from the tarsi (≤ 41 counts). Additional orthology assignments

identified HillPPK3 as the closest *H. illucens* relative of the pheromone-sensing channel *ppk23* [14, 15] (bootstrap 100) and HillPPK18 as the ortholog of the *Nach* sodium channel of the degenerin/epithelial sodium channel (DEG/ENaC) family [54] (bootstrap 100).

Appendage identity dominates transcriptomic organization of the adult sensory appendages

Whole-transcriptome variation is primarily explained by appendage identity

Principal component analysis (PCA) of the 27 transcriptomic libraries revealed clear organization by appendage, with PC1 and PC2 explaining 39.60% and 16.61% of the variance, respectively (Fig. 2A). Between-appendage distances exceeded within-appendage distances (270.11 versus 137.59), and samples showed moderate clustering (mean silhouette = 0.427). Antennae were transcriptionally more similar to maxillary palps than to tarsi, whereas the maxillary palp–tarsi and antenna–tarsi pairs were similarly divergent (mean distances 295.4 and 294.4). Permutational multivariate analysis of variance (PERMANOVA) confirmed appendage identity as the principal determinant of transcriptomic variation ($R^2 = 0.664$, $F_{2,24} = 23.69$, $p \leq 0.001$), although dispersion differed among appendages (betadisper F = 3.94, $p = 0.028$), indicating that both centroid separation and within-group variability contributed to the observed structure.

The Pearson correlation matrix supported these relationships (Fig. 2B). Antennae and maxillary palps exhibited the highest similarity (mean $r \approx 0.954$), whereas maxillary palps and tarsi were the least similar (≈ 0.900). Within each appendage, correlations remained high across sex and mating status (0.977–0.991 in antennae, 0.986–0.993 in maxillary palps, and 0.980–0.990 in tarsi), indicating highly conserved transcriptional profiles despite differences in reproductive condition. Consistent with these observations, variation partitioning showed that appendage identity accounted for 69.14% of the total variance, whereas sex and mating status explained only 1.45% and 1.03%, respectively, leaving 28.38% unexplained (Fig. 2C).

Chemosensory genes exhibit enhanced appendage-specific organization

Restricting the analysis to the 567 annotated chemosensory transcript models strengthened the appendage-derived transcriptional specificity. PCA again separated the three appendages (Fig. 2E), with PC1 and PC2 explaining 36.06% and 30.25% of the variance, respectively, and clustering increasing relative to the complete transcriptome (mean silhouette = 0.637 versus 0.427). Between-appendage distances remained substantially greater than within-appendage distances (87.57 versus 28.12), although the similarity relationships changed: maxillary palps and tarsi became the closest pair, while antennae were equally distant from both.

PERMANOVA indicated even stronger appendage-level organization than observed for the whole transcriptome ($R^2 = 0.856$, $F_{2,24} = 71.16$, $p \leq 0.001$), with no strong evidence for differences in dispersion among appendages (betadisper $p = 0.054$), supporting interpretation primarily in terms of centroid separation. Pearson correlations similarly shifted (Fig. 2F), with reduced similarity between antennae and the other appendages (r = 0.617–0.686) and increased similarity between maxillary palps and tarsi (r = 0.772–0.831). Variation partitioning further emphasized the dominance of appendage identity, which explained 91.5% of the total chemosensory variance, whereas sex and mating status accounted for only 0.85% and 0.18%, respectively (Fig. 2G). These results indicate that chemosensory genes are even more strongly partitioned among appendages than the transcriptome as a whole.

Sex and mating status produce limited appendage-specific effects

Within individual appendages, the effects of sex and mating status were modest and differed among appendages (Fig. 2D, H). The maxillary palp showed the clearest evidence of biologically meaningful differentiation, with significant centroid separation for both the whole transcriptome and the chemosensory subset and no accompanying difference in dispersion. In the antenna, PERMANOVA was significant for both gene sets, but the absence of clustering together with significant dispersion heterogeneity indicates that this reflects differences in multivariate spread rather than coherent separation among reproductive groups. The tarsi showed significant but weak effects with limited clustering. Complete test statistics, including R^2, F, *p*, silhouette and betadisper values for every appendage and gene set, are provided in **Additional file 12**.

Appendage-biased gene expression reflects transcriptional divergence

Pairwise differential expression analyses among antennae, maxillary palps, and tarsi revealed extensive transcriptional divergence across all three appendages (Fig. 3A). Chemosensory genes made up the largest share of the antenna-biased differentially expressed genes (DEGs), 9.1% against the maxillary palp and 8.2% against the tarsi, whereas they represented only 2.2–3.3% of the DEGs biased to the maxillary palp or tarsi, which were dominated by non-chemosensory functions (Fig. 3B). Chemosensory genes were over-represented among antenna-biased genes relative to all expressed genes (odds ratio 6.26, 95% CI 5.23–7.48 against the maxillary palp; 5.33, 4.44–6.39 against the tarsi; both $p < 2 \times 10^{-16}$), whereas the palp- and tarsus-biased sets of these two comparisons were not (odds ratio 1.15, 95% CI 0.79–1.63 and 0.98, 0.79–1.21). In the third comparison, between tarsi and maxillary palp, both sets showed a modest but significant enrichment (odds ratio 1.66, 95% CI 1.38–1.99 and 1.49, 1.16–1.90), so these two appendages are distinguished from one another by chemosensory family composition rather than by chemosensory content. All enrichment tests are given in **Additional file 13**.

Genes consistently biased to one appendage define its functional specialization

To identify genes consistently biased in their expression in individual appendages, appendage-biased DEGs from each pairwise comparison were intersected to generate consensus appendage-biased gene sets. This analysis identified 494 antenna-biased, 539 maxillary palp-biased, and 4,161 tarsus-biased genes (**Additional file 1: Fig. S9**). Because bias was defined relative to the other two appendages rather than to the whole body, these genes are appendage-enriched within this three-way comparison and may also be expressed in other tissues.

The antenna displayed the most specialized molecular-function profile in Gene Ontology (GO) over-representation tests. Among the 278 antenna-biased genes with predicted molecular functions, olfactory receptor activity accounted for 127 genes, 45.7% of the set against 1.5% of the annotated background (odds ratio 166, $q = 1.2 \times 10^{-171}$, where q is the Benjamini–Hochberg-adjusted p) (Additional file 1: Fig. S10). This concentration is unique … or tarsus-biased genes (odds ratio 0.35, p = 1.0)(**Additional file 1: Fig. S10**). In the maxillary palp the most strongly over-represented functions were solute:inorganic anion antiporter activity (12 genes, odds ratio 64.7, $q = 4.1 \times 10^{-13}$), odorant binding (14 genes, odds ratio 14.9, $q = 3.2 \times 10^{-9}$), structural constituent of the ribosome (16 genes, odds ratio 8.8) and of the cuticle (19 genes, odds ratio 7.1), and monooxygenase activity (23 genes, odds ratio 5.0, $q = 3.0 \times 10^{-7}$), with no single function predominating (**Additional file 1: Fig. S11**). In the tarsi, which contained the largest appendage-biased repertoire, over-representation was confined to signaling and regulatory functions: G protein-coupled receptor activity (129 genes, odds ratio 6.9, $q = 3.1 \times 10^{-39}$), 'DNA-binding transcription factor activity, RNA polymerase II-specific' (89 genes, odds ratio 4.8, $q = 3.7 \times 10^{-20}$), regulation of alternative mRNA splicing (30 genes, odds ratio 9.6, $q = 6.4 \times 10^{-11}$) and neuropeptide hormone activity (all 12 annotated genes carrying the term, $q = 2.1 \times 10^{-7}$) (**Additional file 1: Fig. S12**). The numerically largest annotation categories in the tarsal set, protein binding (396 genes) and nucleotide binding (242 genes), are not over-represented against the annotated background (odds ratio 1.10, q = 0.77 and odds ratio 0.69, q = 1.00, respectively) and reflect the size of the set rather than functional specialization. The GO domain composition of each pairwise set of appendage-biased genes is shown in **Additional file 1: Fig. S13**.

Distinct chemosensory gene families define appendage specialization

Classification of the 510 non-redundant chemosensory transcript models into appendage-specific, appendage-biased, expressed and not expressed categories revealed striking differences in sensory specialization among appendages (Fig. 3C; **Additional file 14**). The antenna contained the largest repertoire of enriched chemosensory transcript models (95 appendage-specific and 54 appendage-biased non-redundant models, from 81 and 39 gene loci). This specialization was dominated by the odorant receptor family, with 85 antenna-specific and 36 antenna-biased *Or*s, representing 66.5% of the

182 non-redundant OR sequences (**Additional file 1: Fig. S14**). Ionotropic receptors contributed 16 enriched models (6 specific and 10 biased, the latter including the co-receptors *Ir8a* and *Ir25a*), while *Obp*s contributed 10 (4 specific and 6 biased). Two gustatory receptors, *Gr1* and *Gr3*, were also antenna-biased (**Additional file 1: Fig. S15**), both members of the CO_2 receptor clade described above, whereas no *Csp*, *Ppk* or *Trp* genes were preferentially enriched in the antenna. Antennal specialization is therefore dominated by, but not restricted to, volatile odor detection through OR and IR pathways. The maxillary palp contained comparatively few enriched chemosensory models (13 specific and 15 biased) yet exhibited a distinctive molecular profile. Odorant-binding proteins were the predominant family, comprising six palp-specific (*Obp27*, *Obp30*, *Obp48*, *Obp63*, *Obp71* and *Obp75*) and 10 palp-biased models (*Obp1*, *Obp12–Obp18*, *Obp29* and *Obp50*). Five *Or*s (*Or82*, *Or118*, *Or147*, *Or157* and *Or201*) were palp-specific, *Gr*s contributed four enriched models (*Gr4* specific; *Gr2*, *Gr5* and *Gr24* biased), *Csp*s two (*Csp2* and *Csp9*) and *Ppk*s one (*Ppk35*, palp-specific). Notably, one palp-biased *Obp* (*Obp50*) was additionally mating-responsive, with lower expression in mated females (Fig. 4A), suggesting that at least some odorant-binding proteins in the maxillary palp respond to reproductive state. *Obp19* showed a similar mating response but was expressed at comparable levels in the antenna and the palp and was therefore not palp-biased, whereas the palp-biased *Obp12* did not respond to mating.

The tarsi exhibited the broadest representation of chemosensory gene families (23 appendage-

Figure 4. **Sex- and mating-dependent transcriptional remodeling across adult chemosensory appendages.** Genes are called differentially expressed at $|\log_2 FC| \geq 1$ and adjusted $p < 0.001$. Mating-responsive denotes significance in the same direction against both virgin females and virgin males. **A**) Chemosensory loci responding to sex (left) or mating (right), one row per gene locus, grouped by family. Stems run from no change to the $\log_2$ fold change; dot area is mean normalized expression in that appendage. Color denotes appendage: antenna, dark blue; maxillary palp, mid blue; tarsi, gray. Panels A and B share one dot-size scale. **B**) Non-chemosensory genes, the top 20 per contrast ranked by $|\log_2 FC| \times -\log_{10}(\text{adjusted } p)$, sex above and mating below. Colors as in A. Names are abbreviated; an asterisk marks a name inferred from the best *Drosophila melanogaster* BLASTp match rather than an *H. illucens* annotation. **C**) Venn diagrams of genes shared among appendages for four programs: male-biased and female-biased in virgins, and reduced or induced in mated females. Representative genes are named beside their regions. **D**) Reorganization of sex bias by mating. Sex bias measured in virgins (virgin female versus virgin male, x) against sex bias measured in mated females (mated female versus virgin male, y), for every gene significant in either contrast. Because both axes share virgin males as the reference, a change of sign between them is a genuine reversal of sex bias. Genes are classified by their behavior across the two reproductive states: sex-biased in both states (blue), sex-biased in one state only (gray), or sex-biased in both with the direction reversed (ringed). Red marks chemosensory genes. Quadrant labels give the direction of bias in each group, and the counts give the genes in each quadrant. Axes are clipped at ± 10 $\log_2 FC$, with any gene beyond drawn on the boundary.

The tarsi exhibited the broadest representation of chemosensory gene families (23 appendage-specific and 75 appendage-biased models). Pickpocket channels were particularly enriched, comprising five

tarsus-specific members (*Ppk11*, *Ppk19*, *Ppk23*, *Ppk29* and *Ppk36*) and 15 tarsus-biased models, consistent with established roles of pickpocket channels in contact chemosensation and pheromone detection [14, 15]. Ionotropic receptors contributed six tarsus-specific models (*Ir12*, *Ir34*, *Ir77*, *Ir90*, *Ir95* and *Ir100*) together with 17 tarsus-biased models, whereas the co-receptors *Ir8a* (HillIR7) and *Ir25a* (HillIR3), although expressed in all three appendages, were antenna-biased. Transient receptor potential channels contributed 20 tarsus-biased models from six loci, 10 of them transcript variants of a single *Trpm* locus; TRP channels mediate thermo- and mechanosensation [55], although members of the TRPM subfamily also have non-sensory roles. Gustatory receptors were also more numerous among the enriched models of the tarsi (2 specific and 8 biased) than in either of the other appendages, whereas only 15 *Or*s (7 specific and 8 biased) were enriched, emphasizing the relatively minor contribution of volatile odor detection in this appendage.

Together, these analyses demonstrate that appendage-specific transcriptional programs are accompanied by distinct distributions of chemosensory gene families. Whereas the antenna is specialized for volatile olfaction through extensive *Or* and *Ir* enrichment, the maxillary palp exhibits a compact repertoire dominated by odorant-binding proteins and a small set of appendage-specific receptors, while the tarsi integrate multiple receptor classes associated with contact, mechanical, thermal, and gustatory sensing. Expression heatmaps for each chemosensory family across appendages and reproductive states are provided in **Additional file 1: Figs. S14–S20**.

Sex-biased transcription differs among chemosensory appendages

Comparison of virgin females and virgin males ($|\log_2 FC| \geq 1$; adjusted $p < 0.001$) revealed pronounced but appendage-specific sex-biased expression (Fig. 4C). The antenna and maxillary palp were dominated by male-biased transcription, with 403 of 535 and 470 of 555 DEGs, respectively, showing higher expression in males. In contrast, the tarsi were the only appendage in which female-biased genes predominated, accounting for 168 of 263 DEGs (63.9%). Chemosensory genes comprised only a small fraction of sex-biased DEGs, accounting for 6.0% of antennal, 3.2% of maxillary palp, and 2.7% of tarsal DEGs. Within the antenna, these comprised 19 *Or* transcript models from 10 loci, both transcript models of the *Orco* locus and two *Ir* transcript models of a single locus, all female-biased, together with nine male-biased *Trp* and *Obp* transcript models from six loci (Fig. 4A, which shows one row per locus).

Among non-chemosensory genes, Fig. 4B shows for each contrast the 20 genes with the strongest combination of fold change and significance. In the sex contrast, *transformer* was the only female-biased gene and was female-biased in all three appendages. The other 19 were male-biased, most strongly in the maxillary palp; they included *catalase*, *mucin-12*, *spätzle*, the endocuticle structural glycoproteins *ABD-4* and *ABD-5*, *pro-resilin*, *cuticle protein 18.7* and *cytochrome P450 4d2*, and six of them (*schlaff*, *spätzle*, *pro-resilin*, *cuticle protein 18.7*, *sialin* and a *farnesol dehydrogenase*) were male-biased in the antenna as well. The only tarsal gene among them was a male-biased *trehalose-phosphate synthase*. In the mating contrast, *vitellogenin-A1* was the only induced gene, strongly elevated in mated tarsi. The remaining 19 were reduced in mated females in the maxillary palp, 16 of them also in the antenna and

none in the tarsi, and comprised cuticular and extracellular-matrix proteins, among them endocuticle structural glycoproteins, *chitin deacetylase 1*, *obstructor-E*, *mucin-5AC* and *proteoglycan 4*, together with the neuropeptide *SIFamide receptor*.

Mating elicits appendage-dependent transcriptional remodeling

Comparison of mated with virgin females alone identified 196, 386 and 230 DEGs in the antennae, maxillary palps and tarsi, respectively; requiring significance against both virgin groups reduced these to 146, 318 and 33 (Fig. 4C). Because mated females were older than virgins at collection, adult age and mating status are confounded in this design (see Methods), and these differences cannot be attributed to mating alone. The GO-domain composition of mating-responsive genes did not differ significantly from that of sex-biased genes in the antenna ($\chi^2 = 5.09$, df = 2, $p = 0.079$) or the tarsi ($\chi^2 = 2.27$, df = 2, $p = 0.32$) and differed modestly in the maxillary palp ($\chi^2 = 7.62$, df = 2, $p = 0.022$); because each gene carries a single GO term this comparison is descriptive, and some expected counts in the antenna and tarsi were below five. GO over-representation among sex-biased and mating-responsive genes is shown in **Additional file 1: Figs. S21 and S22**; the two gene sets share 41, 168 and 3 genes in the antenna, maxillary palp and tarsi, respectively. Volcano plots of both contrasts in all three appendages are provided in **Additional file 1: Fig. S23**.

Few sex-biased genes are shared across appendages

To identify transcriptional programs conserved across sensory appendages, sex-biased and mating-responsive gene sets were compared among appendages using Venn diagrams (Fig. 4C; **Additional files 15–20**). Shared sex-biased expression was limited, indicating that most sex-biased expression is appendage-specific. Among male-biased genes, 82 were shared between the antenna and maxillary palp, including 76 genes exclusive to these two appendages and six shared across all three appendages. The universally male-biased genes comprised five transcript variants of a single *cytochrome P450 4d2-like* locus together with *daywake*, a clock-linked gene [56]. The antenna–maxillary palp subset included many structural and cytoskeletal genes, among them paralogs of *serine/arginine repetitive matrix protein 1*, *obstructor-E*, *major royal jelly protein 1*, adult cuticle proteins, and additional genes associated with muscle function and extracellular architecture (Fig. 4C; **Additional files 15 and 17**).

In contrast, only five genes were female-biased across all three appendages. These included two transcript variants of the sex-determination gene *transformer* [57] together with three *adult cuticle protein 1-like* paralogs.

The mating response is shared between the olfactory appendages but not with the tarsi

No gene responded to mating in all three appendages once significance against both virgin groups was required (Fig. 4C). The 18 genes that appear shared when mated females are compared with virgin

females alone, among them the neuropeptide *SIFamide receptor* and mucin paralogs, fail that criterion in the tarsi, so the apparent universality of the post-mating response depends on using virgin females as the sole reference. At the level of individual genes the mating response is appendage-specific.

The antenna and maxillary palp shared 95 mating-responsive genes that were not mating-responsive in the tarsi, and every one of them was reduced in mated females rather than induced. This program of the two olfactory appendages was dominated by cuticular and extracellular-matrix components, including *chitin deacetylase 1* and chitin synthase *chs-2*, *mucin-5AC* paralogs, *obstructor-E*, *serine/arginine repetitive matrix protein 1* paralogs and *ice-structuring glycoprotein*, together with the neuropeptide *SIFamide receptor*. Many of these genes are associated with maturation of the adult cuticle after eclosion, so this program may reflect the age difference between mated and virgin females as much as mating itself. Forty-three of the 95 carry no functional annotation. In the tarsi, by contrast, 29 of the 33 mating-responsive genes were specific to that appendage.

No chemoreceptor responds to mating, whereas sex-biased transcription is reorganized in an appendage-specific manner

Comparing mated females with virgin females and with virgin males allowed mating-associated changes to be distinguished from expression differences peculiar to virgin females (**Additional files 21–23**). Across the 497 appendage-level calls that met both criteria (146, 318 and 33 in the antenna, maxillary palp and tarsi, respectively), corresponding to 398 distinct transcript models (**Additional file 1: Fig. S24**), no odorant, gustatory or ionotropic receptor was differentially expressed, and none met the criterion even when the threshold was relaxed to adjusted $p < 0.05$; the entire chemosensory contribution comprised three odorant-binding proteins (*Obp19*, *Obp36* and *Obp50*) and two pickpocket transcript models of a single locus (*Ppk12* and *Ppk21*). Sex, by contrast, was associated with receptor-level differences: 21 *Or* transcript models (11 loci) were sex-biased in virgin antennae and 29 (16 loci) in mated females, together with ionotropic receptor, odorant-binding protein and transient receptor potential genes, while in the maxillary palp the sex-biased chemosensory genes of virgins were mostly odorant-binding proteins and pickpocket channels, with gustatory and ionotropic receptors appearing only in the comparison of mated females with virgin males.

The direction of the mating response differed among appendages (Fig. 4C). In the antenna and maxillary palp it consisted predominantly of reduced expression relative to both virgin groups (140 of 146 and 280 of 318 genes, respectively), whereas in the tarsi it was predominantly an induction (27 of 33 genes, corresponding to 20 distinct loci). Genes induced in mated tarsi were dominated by reproductive and metabolic functions, including *vitellogenin-A1*, two transcript variants of a *lipase member H-A-like* gene, *tyrosinase* and two monooxygenases. The single chemosensory-family gene among them, *Obp36*, was elevated in mated tarsi relative to both virgin females and virgin males ($\log_2$FC 3.02 and 2.10, adjusted $p = 5.8 \times 10^{-8}$ and 2.8×10^{-5}).

Sex-biased expression itself was reorganized differently in each appendage (Fig. 4D). Of 535 genes sex-biased between virgin females and virgin males in the antenna, 398 remained significantly sex-biased in mated females, with the median magnitude little changed ($|\log_2 FC|$ 1.27 versus 1.38), and none reversed direction. The maxillary palp behaved similarly: 487 of 555 genes retained significant bias in the same direction and eight reversed, and the median difference increased from 1.43 to 2.10. The tarsi differed sharply: only 15 of 263 sex-biased genes remained significant in mated females and the median sex difference fell from 1.45 to 0.31, so tarsal sex-biased expression was largely lost rather than inverted. For 46 of the 248 genes that lost significant sex bias, virgin females differed significantly from both mated females and virgin males, so part of the tarsal sex bias of virgins reflects expression peculiar to virgin females. Three tarsal genes reversed direction, representing two loci: the *lipase member H-A-like* transcript variants and *vitellogenin-A1*, the latter shifting from male-biased in virgins ($\log_2 FC$ − 2.49, adjusted $p = 2.0 \times 10^{-4}$) to strongly female-biased in mated females ($\log_2 FC$ + 5.10, adjusted $p = 6.6 \times 10^{-37}$). None of the 11 reversals involved a chemosensory gene; in the maxillary palp they comprised three *adult cuticle protein 1-like* paralogs, a *glycine-rich cell wall structural protein* and four loci with no functional annotation (Fig. 4D). *Obp31*, among the most strongly male-biased tarsal genes in virgins ($\log_2 FC$ − 8.15, adjusted $p = 1.9 \times 10^{-7}$; the virgin-female tarsal mean of 0.36 normalized counts makes this magnitude imprecise), was no longer significantly sex-biased in mated females ($\log_2 FC$ − 2.02, adjusted $p = 0.057$). Because mated females were older than virgin males at collection, this second sex contrast carries the same age confound as the mating comparison, so the retention, loss and reversal of sex bias cannot be attributed to mating independently of adult age.

Discussion

One of the central questions in sensory evolution is how much behavioral innovation arises from changes in peripheral sensory repertoires as against modulation of downstream neural circuits [58, 59]. The two are not alternatives, and what is rarely measured is their relative contribution. We hypothesized that the transition from mate searching to post-mating oviposition would involve extensive remodeling of peripheral chemosensory gene expression [20, 22]. Instead, appendage identity was the dominant determinant of transcriptomic organization, whereas sex and reproductive state exerted comparatively modest effects. Within the peripheral compartment sampled here, the sensory architecture of *H. illucens* is therefore stable, and the differences between sexes and reproductive states involve a small set of regulatory genes rather than restructuring of receptor repertoires: sex determination (*transformer*) and circadian regulation (*daywake*) among sex-biased genes, and neuropeptide signaling (the *SIFamide receptor*) together with cuticular and extracellular-matrix genes among the genes responding to mating in the two olfactory appendages. Whether this pattern holds more widely cannot be settled with data from a single species. In the *Ae. aegypti* neurotranscriptome, samples cluster by tissue rather than by sex or blood-feeding state [24], and in *Anopheles gambiae* the antennae and maxillary palps differ sharply in chemosensory gene expression [25], so appendage-level organization is unlikely to be a brachyceran innovation. The proposition that appendage-specific sensory architectures became stabilized early in dipteran evolution while behavioral flexibility was achieved through regulation of conserved circuits is

our own and is best treated as a hypothesis, which appendage-resolved expression data from further early-diverging lineages would test. Host specialization in *Drosophila sechellia*, by comparison, involved the evolution of both receptors and circuits [60].

Appendage identity is stable; receptor expression is not

Two processes need to be distinguished here. Appendage identity in the developmental sense, the programs that make an antenna rather than a leg, is specified in the imaginal discs in *Drosophila* [61] and is not addressed by our data. What our data describe is the second process: which members of the chemosensory gene families are expressed in each appendage of the adult, and in what amounts. On that measure the three appendages have sharply distinct identities that are stable across sex and reproductive state. Because a single species was sampled, we cannot date when those identities arose. The observation is consistent with, but does not demonstrate, early stabilization of appendage-specific expression followed by diversification within receptor families, and the mosquito datasets cited above indicate that appendage-specific expression is at least as old as the divergence of nematocerans and brachycerans.

The three appendages therefore represent complementary sensory modules rather than redundant sensory organs: the antenna specializes in volatile olfaction, the maxillary palp in close-range chemical evaluation, and the tarsi in contact, mechanical, and thermal sensing [32, 55, 62–65]. Such functional partitioning likely maximizes sensory coverage while minimizing redundancy. Interestingly, despite the reduction in the number of maxillary palp segments across dipteran evolution, from the multi-segmented condition of nematocerans, through the two-segmented palps of early brachycerans such as *H. illucens*, to the single-segmented palps of cyclorrhaphan flies [19, 66], the palp has retained an olfactory function in both nematocerans and cyclorrhaphans [65, 67]. This decoupling of appendage morphology from sensory function points to a distinction that runs through our results: the identity of an appendage, which is stable here across sex and reproductive state, is a different thing from the particular receptors it expresses, which can change between lineages. The candidate CO_2 receptors show how far the two can come apart. Of the five *H. illucens* genes in this clade, three are enriched in the maxillary palp, as in mosquitoes, while two are antenna-biased, as in *D. melanogaster*, the best-characterized higher brachyceran in this respect [49, 51, 67, 68]. A single receptor clade is therefore distributed across two appendages within one species, so which appendage expresses a given receptor is evidently labile even where the appendages themselves are stable and the receptor lineage is conserved. On this reading the mosquito and *Drosophila* arrangements are unlikely to represent a binary switch: palp and antennal expression are not mutually exclusive, and the ancestral dipteran condition may have been dual, with each crown lineage subsequently biasing one appendage over the other. A lineage branching between mosquitoes and the higher flies is therefore useful less for sorting characters into ancestral and derived states than for showing which apparent dichotomies are not dichotomies at all. The indole-sensing OR lineage illustrates the opposite case, a receptor lineage conserved in both its sequence and its appendage. IndolORs form an ancient, functionally conserved dipteran clade that unites the mosquito indole and skatole receptors AaegOR2, AaegOR9 and AaegOR10 [44, 45] with the *Drosophila* indolORs

DmelOR30a, DmelOR43a and DmelOR49b [46] and is retained across Diptera [47]. In *H. illucens* this clade is represented by the antenna-biased HillOR179 together with HillOR180, whose mean normalized expression remained below 10 counts in all three appendages sampled here. HillOR179 is therefore a candidate receptor for indolic volatiles in this species, consistent with indole detection being an ancestral feature of the dipteran peripheral olfactory system [44, 46] rather than a lineage-specific innovation; functional characterization will be needed to test this.

Reproductive plasticity fine-tunes conserved sensory modules

Despite clear behavioral differences between virgin and mated females [13], reproductive state explained only a small fraction of transcriptomic variation. The clearest chemosensory difference between the sexes was in the antenna, where 19 *Or* transcript models from 10 loci and the co-receptor *Orco* were female-biased, consistent with the sexual dimorphism of the antenna [18]; these receptors are candidates for detecting odors used by females. Rather than broadly altering receptor repertoires, sex and mating were associated with a small number of regulatory genes: sex determination and circadian regulation among sex-biased genes, and neuropeptide signaling among mating-responsive genes, together with cuticular and extracellular-matrix genes whose response may partly reflect adult age. These observations suggest that reproductive plasticity operates by adjusting the physiological state of pre-existing sensory circuits while preserving appendage-specific sensory identities, so that behavioral transitions associated with mating would be accomplished primarily by changing the operating state of existing circuits rather than by rebuilding peripheral sensory repertoires [21, 23]. This separates adult physiological plasticity from the developmental specialization of the appendages themselves, which our design does not address. Comparable evidence from other species is limited and rarely isolates mating itself: in *D. melanogaster* mating raises sex peptide receptor expression in olfactory and gustatory neurons, and this increase is sufficient to shift female preference for polyamines [36], and mating-associated changes in chemosensation are reported across insects [37], whereas the mosquito datasets that resolve individual appendages track blood feeding and the gonotrophic cycle rather than mating [24]. Direct comparisons of mated and virgin females at the level of single chemosensory appendages therefore remain scarce.

The female bias of *transformer* in all three appendages is consistent with continued activity of the sex-determination pathway in adult sensory appendages [57, 69]; in non-drosophilid flies, *transformer* autoregulation maintains female fate [70]. The male bias of *daywake* and the reduced expression of the *SIFamide receptor* after mating point to possible circadian and neuromodulatory inputs, whose role in peripheral sensory function remains to be tested [56, 71–74]. Each of these pathways is represented here by a single gene, so they are best regarded as candidate regulatory inputs to peripheral sensory physiology rather than as an established regulatory core. Overall, sex and mating status remodeled appendage transcriptomes primarily through genes associated with general cellular, structural and physiological processes, whereas chemosensory genes represented only a small proportion of the

regulated transcripts; the contrasting directional biases among appendages indicate that reproductive state engages distinct transcriptional programs in each appendage.

The maxillary palp showed the largest number of mating-responsive genes, but the tarsi were the only appendage in which the mating response consisted mainly of increased expression (27 of 33 genes, against 38 of 318 in the palp and 6 of 146 in the antenna). This is consistent with the multifunctional role of insect tarsi as the final sensory interface between the animal and its substrate [75]; in *Ae. aegypti*, a leg-expressed pickpocket channel mediates the choice of egg-laying site [76]. The change is not a reversal of sex bias: in mated females tarsal sex-biased expression was instead largely lost, and the genes induced in mated tarsi are dominated by reproductive and metabolic functions. Only one gene from a chemosensory family, *Obp36*, was induced relative to both virgin groups. Because *Obp36* is a soluble binding protein within a lineage-specific expansion rather than a receptor, and given the documented non-chemosensory roles of this family [28], its functional significance remains open.

Homology-based annotation extended this picture only partially. When the differentially expressed genes that carry no GO term were compared with the *D. melanogaster* proteome by BLASTp, fewer than half had a homolog covering at least half of the query sequence (42% in the antenna, 38% in the maxillary palp and 28% in the tarsi; **Additional file 24**), so a substantial part of each appendage transcriptome remains functionally uncharacterized and cannot yet be assigned to a pathway.

A hierarchical model of dipteran chemosensory evolution

Taken together, our findings support a hierarchical model for the evolution of dipteran chemosensation, operating across three biological levels (Fig. 5), that builds on the view that olfactory evolution acts on both receptors and circuits [58–60]. At the deepest level we propose, as a hypothesis rather than as a result of this study, that appendage-specific sensory identities became evolutionarily stabilized; what the present data establish is that these identities are unchanged across sex and reproductive state within this species, and testing the evolutionary claim requires comparable appendage-resolved data from additional dipteran lineages. At the intermediate level, if that is so, selection would have acted primarily on the expression and diversification of receptor families within conserved appendage architectures rather than on the evolution of entirely new peripheral sensory organs [29, 77, 78]. At the adult level, physiological states fine-tune these conserved sensory modules through comparatively subtle regulation of conserved signaling pathways rather than wholesale remodeling of receptor repertoires. Because *H. illucens* is an early-diverging lineage of the lower Brachycera (Stratiomyomorpha) [4], the chemosensory catalog and appendage-resolved transcriptomic atlas presented here provide a comparative reference for reconstructing how peripheral sensory systems diversified during one of the earliest radiations within Diptera. Future studies integrating receptor deorphanization, electrophysiology, single-cell transcriptomics, and comparative analyses across early-diverging dipterans will determine how these conserved peripheral sensory modules generate the remarkable ecological diversity of modern flies.

Limitations and future directions

Several limitations should temper interpretation. First, only the three chemosensory appendages were sequenced, with no whole-body or non-sensory reference tissue, so "appendage-specific" and "appendage-biased" are defined relative to the three appendages sampled and do not exclude expression elsewhere in the body. Second, bulk RNA-seq cannot distinguish neuronal from non-neuronal sources of expression; single-cell or single-nucleus transcriptomics will be needed to map chemosensory genes to specific sensory neuron types and accessory cells. Third, whole-appendage dissections may include fat-body cells and hemocytes, which could contribute transcripts of reproductive metabolism such as *vitellogenin*, normally produced by the fat body [79]; in situ hybridization or cell-type-specific approaches will be needed to confirm where such transcripts are expressed. Fourth, tarsi were analyzed as pooled leg segments, although individual tarsomeres may differ in chemosensory function, and the labellum, the principal gustatory organ of the mouthparts, was not sampled. Fifth, the design included virgin males, virgin females and mated females but not mated males, so female-specific and general post-mating effects cannot be separated. Sixth, mated females were necessarily older than virgins at collection (approximately 54–59 h versus 28 h post-emergence), so adult age and mating status are confounded and some mating-associated differences, in particular those involving cuticle maturation, may reflect age rather than mating itself. Seventh, because each library pooled 100–400 appendages, the three replicates per group estimate variance at the pool level rather than between individuals. Finally, cross-species comparisons of family size should be read with caution. The *H. illucens* counts derive from automated RefSeq gene models recovered under a stringent similarity filter, whereas the published comparator counts come from manually curated repertoires assembled under different conventions, and the two biases are not symmetric: divergent gustatory receptors are the most likely to be missed on our side, while our transient receptor potential set includes subfamilies that are not always counted as TRPs elsewhere.

Conclusions

Overall, our findings show that the adult *H. illucens* sensory transcriptome is organized primarily by appendage identity rather than by sex or mating status. Chemosensory genes are strongly partitioned among antennae, maxillary palps, and tarsi, reflecting their specialized roles in volatile olfaction, short-range chemical detection, and contact chemosensation. Sex and mating status do not broadly restructure peripheral sensory repertoires, but instead modulate a focused set of regulatory and physiological pathways, including sex determination (*transformer*), circadian regulation (*daywake*) and *SIFamide receptor* signaling. Together, the chemosensory catalog and appendage-resolved expression atlas, which can be explored with the accompanying *H. illucens* Olfactome Browser, provide a foundation for functional genetics, neurophysiology, and comparative studies of sensory evolution in this early-diverging brachyceran fly.

Methods

Fly rearing, appendage collection and RNA extraction

Hermetia illucens were reared until adulthood in the FreezeM rearing facility under controlled conditions (28–31°C, 60% relative humidity, 14:10 h light:dark photoperiod). Antennae, tarsi, and maxillary palps were collected 28 h post-emergence (virgin males and virgin females) or 30–35 h after mating (mated females). Mating pairs were identified by direct visual observation, and each mated female was isolated individually immediately after the observed copulation event. Mating occurred approximately 24 h after adult emergence, so mated females were approximately 54–59 h post-emergence at collection, compared with 28 h for virgin males and virgin females. Total RNA libraries consisted of 100, 200 and 400 antennae, tarsi and maxillary palps, respectively. Appendages were submerged in TRIzol reagent (Invitrogen 15596026) and RNA was extracted using the Direct-zol RNA MiniPrep (Zymo Research R2050) kit. Libraries were sequenced at the Novogene facility in Singapore using Illumina NovaSeq X, and each library was sequenced to a depth of 30 million 150-bp paired-end reads. Read quality was assessed with FastQC [80] and summarized across libraries with MultiQC [81]. Reads were aligned to the reference genome assembly GCF_905115235.1 [6] with HISAT2 v2.1.0 [82], and reads were assigned with featureCounts (Subread v2.0.1) [83] to the 24,828 RefSeq protein-coding transcript models of the same assembly, with multi-mapping reads counted (-M); alternative transcripts of one locus can therefore carry identical counts, and gene counts reported in this study refer to transcript models unless loci are specified.

Annotation of chemosensory genes and phylogenetic analysis

Representative dipteran species were selected to represent nematoceran, early-diverging brachyceran and higher brachyceran lineages [4], while maximizing the availability of well-annotated chemosensory repertoires. The dataset includes the nematoceran mosquito *Aedes aegypti*, the early-diverging brachyceran *Hermetia illucens*, and five higher brachyceran species (*Bactrocera dorsalis*, *Drosophila melanogaster*, *Glossina morsitans*, *Musca domestica*, and *Calliphora stygia*), thereby providing a phylogenetic context for distinguishing conserved, lineage-specific, and derived features of dipteran chemosensory gene families.

For identifying *H. illucens* orthologs in the GCF_905115235.1 assembly [6], tBLASTn (BLAST+ v2.9.0 [84]) was used for each gene family with *Ae. aegypti*, *D. melanogaster*, *M. domestica* and *C. stygia* protein sequences as queries. *Ae. aegypti* sequences were obtained from VectorBase [41] and *D. melanogaster* sequences from FlyBase [39]; the *M. domestica* and *C. stygia* queries, and the outgroup sequences of the two remaining species, were taken from the dataset of Xu et al. [26]. BLAST results were filtered by e-value < 1×10^{-20} and query coverage ≥ 70%. Of 579 candidate chemosensory transcript models initially identified by tBLASTn, 12 corresponded to NCBI RefSeq gene models annotated as non-coding RNA (XR_ accessions; 9 *Or*, 2 *Gr*, and 1 *Obp*) that lack a protein product and were therefore excluded from the protein-based phylogenetic and amino-acid classification analyses, yielding 567 protein-coding (XM_) annotated sequences; these were obtained from the GCF_905115235.1 genome assembly and aligned using MAFFT v7 (G-INS-1) [85, 86]. All 567 annotated sequences were carried forward into the

expression analysis. Pairwise amino acid identity, calculated from the MAFFT alignment of each family, identified 28 clusters of sequences sharing 100% amino acid identity, together comprising 85 gene models and therefore 57 redundant models (**Additional files 25–31**; Table 1). All cluster members were additionally compared at the nucleotide level (**Additional file 32**). In all 151 within-cluster comparisons the two models belong to the same RefSeq gene and their genomic intervals overlap, and no pair shares an identical full-length mRNA; in 23 of the 28 clusters the models are alternative transcript variants of a single gene that differ in their untranslated regions. The coding sequence of the shorter transcript is contained within that of the longer one over its entire length in 123 of the 151 comparisons, of which 82 are identical base for base. In the other five clusters identical proteins are encoded by tandem-duplicated coding exons within a single annotated locus: in three the copies differ only by synonymous substitutions (*Gr16* and *Gr17*, 99.23% coding nucleotide identity; *Gr19* and *Gr20*, 99.13%; *Obp23* and *Obp24*, 99.24%), and in two (*Obp29* and *Obp31*; *Trp14*, *Trp17*, *Trp18* and *Trp19*) they are identical at the nucleotide level. Because clusters were formed by transitive grouping, not every pairwise comparison within a cluster reaches 100% identity: 129 of the 151 comparisons are identical at the amino acid level and the remainder range from 85.8% to 99.4%. The identity matrices report identity over aligned positions only. To avoid inflating appendage-specificity and sex- and mating-bias counts with potentially redundant gene models, 100% amino-acid-identity clusters were collapsed to a single representative, yielding 510 non-redundant sequences (182 OR, 34 GR, 99 IR, 71 OBP, 37 PPK, 8 CSP, 79 TRP) for appendage-bias classification. Collapsing identical proteins does not remove transcript variants that encode different proteins, so the non-redundant set still contains several models for some loci (for example, 79 TRP models from 33 loci). Per-family pairwise amino acid identity matrices are provided as **Additional files 25–31**; phylogenetic orthology assignments for each *H. illucens* gene (based on smallest mixed-clade topology with outgroup sequences) are provided in **Additional files 4–10** (OR, GR, IR, OBP, PPK, TRP, CSP respectively); amino acid identity between each 1:1 ortholog and its nearest outgroup sequence is given in **Additional file 33**; although members of 100%-identity clusters are collapsed for appendage-bias classification, every transcript model is reported individually in the Additional files. Maximum-likelihood phylogenetic trees were constructed using IQ-TREE v2.0.3 [87], with the best-fit substitution model for each alignment selected automatically by ModelFinder [88] under the Bayesian information criterion, and node support estimated by standard non-parametric bootstrapping [89] with 5,000 replicates for the combined tree of all seven families and 1,000 replicates for each family tree. Trees were visualized and annotated in iTOL v6 [90]. The seven family trees are provided in Newick format, with branch lengths and support values, as **Additional file 11**. Numbered gene designations within each family (e.g., *Ppk11*, *Or3*) reflect numerical assignment and do not imply one-to-one orthology with *Drosophila* or *Ae. aegypti* paralogs of the same number, nor does a consecutive number imply adjacency on the same chromosome; for example, *H. illucens Ppk23* is not the ortholog of *D. melanogaster ppk23*, which is HillPPK3.

Differential expression analysis

Differential expression was analyzed using DESeq2 [91] with size-factor normalization to correct for library-depth differences. The master count table (**Additional file 34**) reports these size-factor normalized counts. Fragments per kilobase of transcript per million mapped fragments (FPKM), computed as (normalized counts / kilobase of transcript length, the summed exon length of each RefSeq transcript model) / (sum of normalized counts in the library / 10^6), are provided for every transcript model and library in **Additional file 35** as a reference unit for user convenience only; all differential expression analysis was performed using DESeq2 on raw count matrices with size-factor normalization, not on FPKM values. Three sets of pairwise comparisons were performed: (i) appendage-level comparisons among antenna, maxillary palp, and tarsi (n = 9 libraries per appendage; design formula: ~ appendage); (ii) sex-effect comparisons between virgin females and virgin males within each appendage (n = 3 per group; design formula: ~ sex); and (iii) mating-effect comparisons between mated females and virgin females within each appendage (n = 3 per group; design formula: ~ mating_status). Because replicates represent pools of 100–400 appendages, variance estimates reflect between-pool rather than between-individual variability. A gene was classified as differentially expressed if it displayed $|\log_2 FC| \geq 1$ and a Benjamini–Hochberg-adjusted [92] $p < 0.001$. Relaxing the adjusted-p threshold to 0.05 increased DEG counts without changing qualitative directional or appendage-level patterns.

All differential expression results supporting this study are provided in **Additional files 15–23 and 36–38**. **Additional files 36–38** contain pairwise appendage-vs-appendage DESeq2 results (antenna versus tarsi, antenna versus maxillary palp, and tarsi versus maxillary palp, respectively), each reporting all 24,828 transcript models with baseMean, $\log_2 FC$, lfcSE, Wald statistic, *p*-value, and Benjamini–Hochberg-adjusted *p*-value. **Additional files 15–20** contain within-appendage DE results for the sex contrast (virgin female versus virgin male) and the mating-effect contrast (mated female versus virgin female): **Additional files 15 and 16** for the antenna, **Additional files 17 and 18** for the maxillary palp, and **Additional files 19 and 20** for the tarsi; these files additionally include per-replicate normalized count columns to allow direct inspection of group-level expression. **Additional files 21–23** contain the mated female versus virgin male contrast of each appendage (see Sex- and mating-effect directional analysis). The full normalized count matrix across all 27 libraries (**Additional file 34**) and the GO annotation (**Additional file 39**) are provided as additional reference files. Pairwise amino acid identity matrices for each of the seven chemosensory gene families (CSP, GR, IR, OBP, OR, PPK, TRP), used for 100%-identity deduplication, are provided as **Additional files 25–31**. Note: In **Additional files 15–20**, $\log_2$FoldChange is reported in the DESeq2 output convention (sex files: Vm/VF; mating files: VF/MF), such that positive values indicate higher expression in virgin males (sex comparisons) or virgin females (mating comparisons). In the manuscript text and figures, this is negated throughout (positive $\log_2 FC$ = female-biased or higher in mated females).

Global transcriptomic and multivariate analysis

Transcriptomic analyses were performed in R (v4.5.2) [93] using normalized count matrices from 27 libraries representing antenna, maxillary palp, and tarsi across three reproductive states (virgin male,

virgin female, and mated female; n = 3 biological replicates per group; full matrix provided as **Additional file 34**). The core packages were DESeq2 [91] (differential expression), vegan [94] (PERMANOVA, varpart, betadisper), cluster [95] (silhouette analysis) and ggplot2 [96], with dplyr, tidyr, tibble, stringr, scales, ggrepel, ggforce, cowplot and gridExtra (Figs. 2 and 3). Package versions were DESeq2 1.50.2, vegan 2.7-2, permute 0.9-8, cluster 2.1.8.1, ggplot2 4.0.1, dplyr 1.1.4, tidyr 1.3.1, tibble 3.3.0, stringr 1.6.0 and scales 1.4.0; the full sessionInfo() output is appended to each analysis script (**Additional file 12** and the analysis code archived at Zenodo [97]). Genes with zero variance across samples were removed, and expression values were $\log_2$-transformed after addition of a pseudocount of 1. For the chemosensory subset, genes annotated as *Or*, *Ir*, *Gr*, *Obp*, *Csp*, *Ppk*, and *Trp* were retained based on curated annotations.

Principal component analysis (PCA) was conducted on scaled, $\log_2$-transformed expression matrices using prcomp, with samples as observations. Cluster geometry was characterized by mean within- and between-group Euclidean distances in PC space, convex hull areas, and 95% confidence ellipse areas in the PC1–PC2 plane; geometric metrics including convex hull areas and confidence ellipses are provided in **Additional file 12**. All analyses were implemented in custom R scripts (Figure_2.R, Figure_3.R and Figure_4.R, with the helper scripts Figure_4A_lollipop.R, Figure_4B_lollipop.R and Figure_4_names.R).

Differences in multivariate expression structure were evaluated using PERMANOVA [98] (adonis2, vegan, 999 permutations). Homogeneity of dispersion was assessed using betadisper [99]; PERMANOVA *p*-values were interpreted jointly with silhouette coefficients [100] to distinguish centroid separation from variance-driven effects (full metrics in **Additional file 12**). Variation partitioning used a three-factor model (appendage identity, sex, mating status) via partial redundancy analysis (varpart, vegan) with Ezekiel-corrected adjusted R^2 values [101]. Note that PERMANOVA and DESeq2 address complementary analytical questions: PERMANOVA tests whether multivariate expression centroids or dispersion structures differ between groups, whereas DESeq2 tests for statistically significant changes in individual gene expression levels. The absence of PERMANOVA-detected centroid separation in a given comparison does not preclude identification of individual DEGs by DESeq2, as these two approaches operate on fundamentally different dimensions of the data.

Pearson correlation coefficients were calculated on group-averaged expression profiles to characterize global similarity among appendages and reproductive groups. For variation partitioning, the response matrix comprised $\log_2$-transformed, zero-variance-filtered expression values across all 27 libraries (24,608 transcript models after removing 220 with zero variance). Explanatory variables were treatment-coded: appendage identity (two dummy variables), sex (binary: 0 = male, 1 = female), and mating status (binary: 0 = virgin, 1 = mated). The small fractions attributed to sex (1.45%) and mating status (1.03%) carry substantial uncertainty at n = 27 and should be interpreted as order-of-magnitude estimates. Full statistical outputs are provided as **Additional file 12**.

Chemosensory gene expression classification

To classify chemosensory gene expression patterns, a curated set of 510 non-redundant chemosensory transcript models across seven gene families (*Or*, *Gr*, *Ir*, *Obp*, *Csp*, *Ppk*, *Trp*) was assembled from the *H. illucens* genome annotation (the 567 protein-coding models retained from 579 tBLASTn candidates, after collapsing 100%-identity clusters; per-family statistics in Table 1). Sequences with 100% amino acid identity were collapsed to a single representative prior to analysis (per-family pairwise amino acid identity matrices provided as **Additional files 25–31**). All 567 sequences, including members of 100%-identical clusters, were retained in the expression analysis; Table 1 provides the number of non-redundant representatives per family. For each appendage, each chemosensory gene was assigned to one of four expression categories based on DESeq2 results from pairwise appendage comparisons: appendage-specific (mean normalized count ≥ 10 in that appendage and below that threshold in both other appendages; no statistical test is applied to this category), appendage-biased (expressed in that appendage and significantly higher than each of the other two appendages in the pairwise DESeq2 comparisons, adjusted $p < 0.001$ and $\log_2 FC \geq 1$ against both), expressed (mean normalized count ≥ 10 but neither specific nor biased), or not expressed (mean normalized count < 10). Sex-biased and mating-responsive classifications were applied independently using the same thresholds from the respective pairwise comparisons. The classification of every transcript model in each appendage is provided in **Additional file 14**. The GO annotation used for all enrichment analyses is provided as **Additional file 39**.

Table 1

Per-family chemosensory gene counts. Annotated transcript models, non-redundant sequences remaining after collapsing 100% amino-acid-identity clusters, and distinct gene loci obtained by mapping every model to its RefSeq GeneID.

Family	Annotated sequences	Non-redundant (after collapsing 100% amino acid identity clusters)	Distinct gene loci
Or (incl. *Orco*)	192	182	148
Gr	39	34	23
Ir	101	99	83
Obp	75	71	65
Ppk	39	37	35
Csp	10	8	6
Trp	111	79	33
Total	**567**	**510**	**393**

Excluded non-coding models. The 12 candidate models annotated by NCBI RefSeq as non-coding RNA (XR_ accessions; 9 *Or*, 2 *Gr*, 1 *Obp*) and excluded from the protein-based analyses are *Or22* (XR_005249864.1), *Or46* (XR_005250427.1), *Or47* (XR_005250426.1), *Or107* (XR_005249937.1), *Or121* (XR_005250358.1), *Or122* (XR_005250357.1), *Or126* (XR_005249998.1), *Or131* (XR_005249944.1), *Or149* (XR_005249890.1); *Gr23* (XR_005249778.1), *Gr29* (XR_005250024.1); *Obp4* (XR_005249469.1).

Sex- and mating-effect directional analysis

Mated females were additionally compared with virgin males within each appendage, using the same DESeq2 design (n = 3 per group), so that mating-associated changes could be distinguished from expression peculiar to virgin females (**Additional files 21–23**). A gene was classified as mating-responsive only if it was significant ($|\log_2 FC| \geq 1$; adjusted $p < 0.001$) in the same direction in both the mated-female versus virgin-female and the mated-female versus virgin-male comparisons. Sex bias was assessed in virgins (virgin female versus virgin male) and again after mating (mated female versus virgin male), and each gene sex-biased in virgins was classified as retaining, losing or reversing that bias according to its significance and sign in the mated contrast; reported medians of $|\log_2 FC|$ are taken over all genes sex-biased in virgins, re-measured in the mated contrast. Correlation coefficients between fold changes are not reported. If sex and mating effects are both expressed relative to virgin females, a negative correlation between them is expected under the null, and a change of sign between those two contrasts is algebraically equivalent to the virgin-female mean being the highest or the lowest of the three group means; sex bias in mated females was therefore measured against virgin males. In 18 of the 4,697 significant calls across all contrasts and appendages (0.4%), one of the two groups had a mean of zero normalized counts; for these the reported fold change reflects the model's treatment of an absent group rather than a measured ratio and should be read as presence/absence rather than as a magnitude. The most extreme is an antennal *laccase-2*-like gene, absent from mated females, at $\log_2 FC - 21.8$. Two carry chemosensory annotations: *Or17* (antenna, absent from virgin males) and *Obp34* (tarsi, absent from virgin females).

Gene Ontology [102] annotations for all 24,828 transcript models were obtained by mapping the RefSeq transcript accessions of the *H. illucens* genome assembly GCF_905115235.1 to GO terms with g:Profiler g:Convert [103]; term names and domains were then added with a custom Python script, and the resulting annotation is provided as **Additional file 39**. Genes carrying no GO term in this annotation, labeled "Unknown" domain although most of them do carry an NCBI RefSeq product name, were additionally compared with the *D. melanogaster* proteome (UniProt/FlyBase reference proteome [39, 104]) by BLASTp using a threshold of e-value $< 1 \times 10^{-5}$; percentage identity and query coverage are reported for every hit in **Additional file 24** and were not applied as filters, so hits of low query coverage are present in that file and should be interpreted accordingly. Three-set Venn diagrams were used to identify genes shared among appendages within each program (male-biased and female-biased in virgins, and reduced or induced in mated females; Fig. 4C); the overlap between the sex-biased and mating-responsive sets within each appendage is reported in the Results. Scatter plots display sex bias in virgins against sex bias in mated females for genes sex-biased in either contrast; no correlation coefficient is reported, for the reason given above. Over-representation of GO terms within each gene set was tested with a one-sided Fisher's exact test against all annotated genes tested in the same contrast, and *p* values were corrected by the Benjamini–Hochberg procedure within each gene set, appendage and GO domain; odds ratios are conditional maximum-likelihood estimates. Results for the consensus appendage-biased, sex-biased and mating-responsive gene sets are given in **Additional files 40 and 41**.

Data deposition

Raw RNA-seq reads from all 27 libraries have been deposited in the NCBI Sequence Read Archive under BioProject accession PRJNA1490337 [105].

Use of AI-assisted tools

Claude (Anthropic) was used to assist with language editing, reference checking and the writing of analysis scripts. All outputs were reviewed and verified by the authors, who take full responsibility for the content.

Declarations

Ethics approval and consent to participate

Not applicable. This study did not involve human participants, human data, or vertebrate animals.

Consent for publication

Not applicable.

Availability of data and materials

All supplementary data are provided as **Additional files 1–41** accompanying this publication. In **Additional file 34**, sample columns are prefixed Ant_ (antenna), Leg_ (tarsi), and P_ (maxillary palp). **Additional file 24** contains BLASTp homology results for differentially expressed genes carrying no GO term, searched against the UniProt/FlyBase *D. melanogaster* reference proteome; percentage identity and query coverage are reported for every hit. Full statistical reports for **Fig. 2**, including PCA geometry metrics, PERMANOVA summaries, silhouette analyses, Pearson correlation tables, and variation partitioning results, are provided as **Additional file 12** (ten CSV tables and one full-text report). Raw RNA-seq reads are available in the NCBI Sequence Read Archive under BioProject accession PRJNA1490337 [105]. The expression data can also be explored with the *H. illucens* Olfactome Browser, an interactive application for per-gene queries of appendage-, sex- and mating-state-resolved expression, browsing of the phylogenetic trees and download of the underlying tables. The browser is intended to be run locally, from its source code with installation instructions (https://github.com/dorperets-lab/Hillucens_olfactome_app) or from the snapshot archived at Zenodo; a hosted instance (https://hillucensolfactomeapp-rk3chuhczmmvbpfq2ctvan.streamlit.app/) is provided for convenience but may not remain available. The R scripts that generate **Figs. 2–4** and **Additional file 1: Figs. S9–S24**, the protein sequences and phylogenies of the seven chemosensory families, and a snapshot of the browser and its underlying tables are archived at Zenodo (https://zenodo.org/records/23452179) [97], so that they remain available independently of any web host. **Additional files 11 and 12** are provided as archives containing multiple files.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by the Israeli Ministry of Innovation, Science and Technology (grant No. 85152).

Author Contributions Statement

D.P. and J.D.B. designed the study. D.P. performed all experimental work, including the dissections and RNA extraction, and all analyses. D.P. and J.D.B. wrote the manuscript. P.A.P. provided scientific guidance and critically reviewed the manuscript. E.Y. provided technical support. All authors approved the final version.

Acknowledgements

We thank FreezeM for providing adult black soldier flies used in this study.

References

1. Kaya C, Generalovic TN, Ståhls G, Hauser M, Samayoa AC, Nunes-Silva CG, et al. Global population genetic structure and demographic trajectories of the black soldier fly, *Hermetia illucens*. BMC Biol. 2021;19:94. https://doi.org/10.1186/s12915-021-01029-w.
2. Rehman KU, Hollah C, Wiesotzki K, Rehman RU, Rehman AU, Zhang J, et al. Black soldier fly, *Hermetia illucens* as a potential innovative and environmentally friendly tool for organic waste management: A mini-review. Waste Manag Res. 2023;41:81–97. https://doi.org/10.1177/0734242x221105441.
3. Sheppard DC, Tomberlin JK, Joyce JA, Kiser BC, Sumner SM. Rearing Methods for the Black Soldier Fly (Diptera: Stratiomyidae). J Med Entomol. 2002;39:695–8. https://doi.org/10.1603/0022-2585-39.4.695.
4. Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, et al. Episodic radiations in the fly tree of life. Proc Natl Acad Sci U S A. 2011;108:5690–5. https://doi.org/10.1073/pnas.1012675108.
5. Zhan S, Fang G, Cai M, Kou Z, Xu J, Cao Y, et al. Genomic landscape and genetic manipulation of the black soldier fly *Hermetia illucens*, a natural waste recycler. Cell Res. 2020;30:50–60. https://doi.org/10.1038/s41422-019-0252-6.
6. Generalovic TN, McCarthy SA, Warren IA, Wood JMD, Torrance J, Sims Y et al. A high-quality, chromosome-level genome assembly of the Black Soldier Fly (*Hermetia illucens* L.). G3 (Bethesda). 2021;11:jkab085. https://doi.org/10.1093/g3journal/jkab085

7. Muchina P, Kinyua J, Khamis F, Tanga CM, Gebreyesus G, Cai Z, et al. The genomic landscape of *Hermetia illucens*: exploring the potential of the black soldier fly through molecular insights. J Insects Food Feed. 2025;11:2049–74. https://doi.org/10.1163/23524588-bja10200.
8. Zhou W, Kunz D, Jiggins CD. Comparative analysis of gene family evolution demonstrates expansion of digestive, immunity and olfactory functions in the black soldier fly (*Hermetia illucens*) lineage. Heredity (Edinb). 2026;135:187–98. https://doi.org/10.1038/s41437-025-00805-6.
9. Tomberlin JK, Sheppard DC, Joyce JA. Selected Life-History Traits of Black Soldier Flies (Diptera: Stratiomyidae) Reared on Three Artificial Diets. Ann Entomol Soc Am. 2002. https://doi.org/10.1603/0013-8746(2002)095[0379:slhtob]2.0.co;2. 95:379 – 86.
10. Bruno D, Bonelli M, Cadamuro AG, Reguzzoni M, Grimaldi A, Casartelli M, et al. The digestive system of the adult *Hermetia illucens* (Diptera: Stratiomyidae): morphological features and functional properties. Cell Tissue Res. 2019;378:221–38. https://doi.org/10.1007/s00441-019-03025-7.
11. Tomberlin JK, Sheppard DC. Factors Influencing Mating and Oviposition of Black Soldier Flies (Diptera: Stratiomyidae) in a Colony. J Entomol Sci. 2002;37:345–52. https://doi.org/10.18474/0749-8004-37.4.345.
12. Tomberlin JK, Sheppard DC. Lekking Behavior of the Black Soldier Fly (Diptera: Stratiomyidae). Fla Entomol. 2001;84:729. https://doi.org/10.2307/3496413.
13. Laksanawimol P, Singsa S, Thancharoen A. Behavioral responses of different reproductive statuses and sexes in *Hermetia illucens* (L) adults to different attractants. PeerJ. 2023;11:e15701. https://doi.org/10.7717/peerj.15701.
14. Thistle R, Cameron P, Ghorayshi A, Dennison L, Scott K. Contact Chemoreceptors Mediate Male-Male Repulsion and Male-Female Attraction during *Drosophila* Courtship. Cell. 2012;149:1140–51. https://doi.org/10.1016/j.cell.2012.03.045.
15. Toda H, Zhao X, Dickson BJ. The *Drosophila* Female Aphrodisiac Pheromone Activates ppk23 + Sensory Neurons to Elicit Male Courtship Behavior. Cell Rep. 2012;1:599–607. https://doi.org/10.1016/j.celrep.2012.05.007.
16. Ponnusamy L, Xu N, Nojima S, Wesson DM, Schal C, Apperson CS. Identification of bacteria and bacteria-associated chemical cues that mediate oviposition site preferences by *Aedes aegypti*. Proc Natl Acad Sci U S A. 2008;105:9262–7. https://doi.org/10.1073/pnas.0802505105.
17. Piersanti S, Rebora M, Marri GC, Salerno G. Antennal olfactory responses in the black soldier fly *Hermetia illucens*. J Insect Physiol. 2024;159:104722. https://doi.org/10.1016/j.jinsphys.2024.104722.
18. Pezzi M, Leis M, Chicca M, Falabella P, Salvia R, Scala A, et al. Morphology of the Antenna of *Hermetia illucens* (Diptera: Stratiomyidae): An Ultrastructural Investigation. J Med Entomol. 2017;54:925–33. https://doi.org/10.1093/jme/tjx055.
19. Pezzi M, Scapoli C, Bharti M, Faucheux MJ, Chicca M, Leis M, et al. Fine Structure of Maxillary Palps in Adults of *Hermetia illucens* (Diptera: Stratiomyidae). J Med Entomol. 2021;58:658–65. https://doi.org/10.1093/jme/tjaa251.

20. Rinker DC, Pitts RJ, Zhou X, Suh E, Rokas A, Zwiebel LJ. Blood meal-induced changes to antennal transcriptome profiles reveal shifts in odor sensitivities in *Anopheles gambiae*. Proc Natl Acad Sci U S A. 2013;110:8260–5. https://doi.org/10.1073/pnas.1302562110.
21. Lebreton S, Grabe V, Omondi AB, Ignell R, Becher PG, Hansson BS, et al. Love makes smell blind: mating suppresses pheromone attraction in *Drosophila* females via Or65a olfactory neurons. Sci Rep. 2014;4:7119. https://doi.org/10.1038/srep07119.
22. Crava CM, Sassù F, Tait G, Becher PG, Anfora G. Functional transcriptome analyses of *Drosophila suzukii* antennae reveal mating-dependent olfaction plasticity in females. Insect Biochem Mol Biol. 2019;105:51–9. https://doi.org/10.1016/j.ibmb.2018.12.012.
23. Gadenne C, Barrozo RB, Anton S. Plasticity in Insect Olfaction: To Smell or Not to Smell? Annu Rev Entomol. 2016;61:317–33. https://doi.org/10.1146/annurev-ento-010715-023523.
24. Matthews BJ, McBride CS, DeGennaro M, Despo O, Vosshall LB. The neurotranscriptome of the *Aedes aegypti* mosquito. BMC Genomics. 2016;17:32. https://doi.org/10.1186/s12864-015-2239-0.
25. Pitts RJ, Rinker DC, Jones PL, Rokas A, Zwiebel LJ. Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. BMC Genomics. 2011;12:271. https://doi.org/10.1186/1471-2164-12-271.
26. Xu Q, Wu Z, Zeng X, An X. Identification and Expression Profiling of Chemosensory Genes in *Hermetia illucens* via a Transcriptomic Analysis. Front Physiol. 2020;11:720. https://doi.org/10.3389/fphys.2020.00720.
27. Leal WS. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. Annu Rev Entomol. 2013;58:373–91. https://doi.org/10.1146/annurev-ento-120811-153635.
28. Pelosi P, Iovinella I, Zhu J, Wang G, Dani FR. Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. Biol Rev Camb Philos Soc. 2018;93:184–200. https://doi.org/10.1111/brv.12339.
29. Robertson HM. Molecular Evolution of the Major Arthropod Chemoreceptor Gene Families. Annu Rev Entomol. 2019;64:227–42. https://doi.org/10.1146/annurev-ento-020117-043322.
30. Joseph RM, Carlson JR. *Drosophila* Chemoreceptors: A Molecular Interface Between the Chemical World and the Brain. Trends Genet. 2015;31:683–95. https://doi.org/10.1016/j.tig.2015.09.005.
31. Hallem EA, Carlson JR. Coding of Odors by a Receptor Repertoire. Cell. 2006;125:143–60. https://doi.org/10.1016/j.cell.2006.01.050.
32. Dahanukar A, Hallem EA, Carlson JR. Insect chemoreception. Curr Opin Neurobiol. 2005;15:423–30. https://doi.org/10.1016/j.conb.2005.06.001.
33. Auger L, Bouslama S, Deschamps MH, Vandenberg G, Derome N. Absence of microbiome triggers extensive changes in the transcriptional profile of *Hermetia illucens* during larval ontogeny. Sci Rep. 2023;13:2396. https://doi.org/10.1038/s41598-023-29658-x.
34. Feng X, Memon FU, Ma C, Zhu J, Huang Z, Tettamanti G, et al. Proteomic and transcriptomic analysis of cold- and heat-tolerant black soldier fly (*Hermetia illucens*) larvae. J Insects Food Feed.

2025;11:543–61. https://doi.org/10.1163/23524588-00001215.

35. Shah PN, Maistrou S, Willemsen I, van Loon JJA, Dicke M. Transcriptomic response of *Hermetia illucens* L. (Diptera: Stratiomyidae) to wounding and Gram-negative bacterial infection. J Insects Food Feed. 2025;11:751–71. https://doi.org/10.1163/23524588-00001211.
36. Hussain A, Üçpunar HK, Zhang M, Loschek LF, Grunwald Kadow IC. Neuropeptides Modulate Female Chemosensory Processing upon Mating in *Drosophila*. PLoS Biol. 2016;14:e1002455. https://doi.org/10.1371/journal.pbio.1002455.
37. Nanfack-Minkeu F, Sirot LK. Effects of Mating on Gene Expression in Female Insects: Unifying the Field. Insects. 2022;13:69. https://doi.org/10.3390/insects13010069.
38. Scott JG, Warren WC, Beukeboom LW, Bopp D, Clark AG, Giers SD, et al. Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. Genome Biol. 2014;15:466. https://doi.org/10.1186/s13059-014-0466-3.
39. Öztürk-Çolak A, Marygold SJ, Antonazzo G, Attrill H, Goutte-Gattat D, Jenkins VK, et al. FlyBase: updates to the *Drosophila* genes and genomes database. Genetics. 2024;227:iyad211. https://doi.org/10.1093/genetics/iyad211.
40. Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, et al. Improved reference genome of *Aedes aegypti* informs arbovirus vector control. Nature. 2018;563:501–7. https://doi.org/10.1038/s41586-018-0692-z.
41. Giraldo-Calderón GI, Harb OS, Kelly SA, Rund SS, Roos DS, McDowell MA. VectorBase.org updates: bioinformatic resources for invertebrate vectors of human pathogens and related organisms. Curr Opin Insect Sci. 2022;50:100860. https://doi.org/10.1016/j.cois.2021.11.008.
42. Robertson HM, Warr CG, Carlson JR. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 2003;100(Suppl 2):14537–42. https://doi.org/10.1073/pnas.2335847100.
43. Larsson MC, Domingos AI, Jones WD, Chiappe ME, Amrein H, Vosshall LB. Or83b Encodes a Broadly Expressed Odorant Receptor Essential for *Drosophila* Olfaction. Neuron. 2004;43:703–14. https://doi.org/10.1016/j.neuron.2004.08.019.
44. Bohbot JD, Jones PL, Wang G, Pitts RJ, Pask GM, Zwiebel LJ. Conservation of Indole Responsive Odorant Receptors in Mosquitoes Reveals an Ancient Olfactory Trait. Chem Senses. 2011;36:149–60. https://doi.org/10.1093/chemse/bjq105.
45. Ruel DM, Yakir E, Bohbot JD. Supersensitive odorant receptor underscores pleiotropic roles of indoles in mosquito ecology. Front Cell Neurosci. 2019;12:533. https://doi.org/10.3389/fncel.2018.00533.
46. Ruel DM, Vainer Y, Yakir E, Bohbot JD. Identification and functional characterization of olfactory indolergic receptors in *Drosophila melanogaster*. Insect Biochem Mol Biol. 2021;139:103651. https://doi.org/10.1016/j.ibmb.2021.103651.
47. Arbel M, Yakir E, Bohbot JD. Functional evolution of olfactory indolergic receptors in Diptera. bioRxiv [Preprint]. 2026. https://doi.org/10.64898/2026.01.06.697762

48. Abuin L, Bargeton B, Ulbrich MH, Isacoff EY, Kellenberger S, Benton R. Functional Architecture of Olfactory Ionotropic Glutamate Receptors. Neuron. 2011;69:44–60. https://doi.org/10.1016/j.neuron.2010.11.042.
49. Jones WD, Cayirlioglu P, Grunwald Kadow I, Vosshall LB. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. Nature. 2007;445:86–90. https://doi.org/10.1038/nature05466.
50. Kwon JY, Dahanukar A, Weiss LA, Carlson JR. The molecular basis of CO_2 reception in *Drosophila*. Proc Natl Acad Sci U S A. 2007;104:3574–8. https://doi.org/10.1073/pnas.0700079104.
51. McMeniman CJ, Corfas RA, Matthews BJ, Ritchie SA, Vosshall LB. Multimodal Integration of Carbon Dioxide and Other Sensory Cues Drives Mosquito Attraction to Humans. Cell. 2014;156:1060–71. https://doi.org/10.1016/j.cell.2013.12.044.
52. Erdelyan CNG, Mahood TH, Bader TSY, Whyard S. Functional validation of the carbon dioxide receptor genes in *Aedes aegypti* mosquitoes using RNA interference. Insect Mol Biol. 2012;21:119–27. https://doi.org/10.1111/j.1365-2583.2011.01120.x.
53. Miyamoto T, Slone J, Song X, Amrein H. A Fructose Receptor Functions as a Nutrient Sensor in the *Drosophila* Brain. Cell. 2012;151:1113–25. https://doi.org/10.1016/j.cell.2012.10.024.
54. Zelle KM, Lu B, Pyfrom SC, Ben-Shahar Y. The genetic architecture of degenerin/epithelial sodium channels in *Drosophila*. G3 (Bethesda). 2013;3:441 – 50. https://doi.org/10.1534/g3.112.005272
55. Fowler MA, Montell C. *Drosophila* TRP channels and animal behavior. Life Sci. 2013;92:394–403. https://doi.org/10.1016/j.lfs.2012.07.029.
56. Yang Y, Edery I. Daywake, an Anti-siesta Gene Linked to a Splicing-Based Thermostat from an Adjoining Clock Gene. Curr Biol. 2019;29:1728–34. https://doi.org/10.1016/j.cub.2019.04.039. .e4.
57. Boggs RT, Gregor P, Idriss S, Belote JM, McKeown M. Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the *transformer* gene. Cell. 1987;50:739–47. https://doi.org/10.1016/0092-8674(87)90332-1.
58. Ramdya P, Benton R. Evolving olfactory systems on the fly. Trends Genet. 2010;26:307–16. https://doi.org/10.1016/j.tig.2010.04.004.
59. Hansson BS, Stensmyr MC. Evolution of Insect Olfaction. Neuron. 2011;72:698–711. https://doi.org/10.1016/j.neuron.2011.11.003.
60. Auer TO, Khallaf MA, Silbering AF, Zappia G, Ellis K, Álvarez-Ocaña R, et al. Olfactory receptor and circuit evolution promote host specialization. Nature. 2020;579:402–8. https://doi.org/10.1038/s41586-020-2073-7.
61. Casares F, Mann RS. Control of antennal versus leg development in *Drosophila*. Nature. 1998;392:723–6. https://doi.org/10.1038/33706.
62. Stocker RF. The organization of the chemosensory system in *Drosophila melanogaster*: a rewiew. Cell Tissue Res. 1994;275:3–26. https://doi.org/10.1007/bf00305372.

63. Vosshall LB, Stocker RF. Molecular Architecture of Smell and Taste in *Drosophila*. Annu Rev Neurosci. 2007;30:505–33. https://doi.org/10.1146/annurev.neuro.30.051606.094306.
64. Couto A, Alenius M, Dickson BJ. Molecular, Anatomical, and Functional Organization of the *Drosophila* Olfactory System. Curr Biol. 2005;15:1535–47. https://doi.org/10.1016/j.cub.2005.07.034.
65. de Bruyne M, Clyne PJ, Carlson JR. Odor Coding in a Model Olfactory Organ: The *Drosophila* Maxillary Palp. J Neurosci. 1999;19:4520–32. https://doi.org/10.1523/jneurosci.19-11-04520.1999.
66. Schneeberg K, Beutel RG. The evolution of head structures in lower Diptera. ScienceOpen Res. 2015. https://doi.org/10.14293/s2199-1006.1.sor-life.altce1.v2.
67. Lu T, Qiu YT, Wang G, Kwon JY, Rutzler M, Kwon HW, et al. Odor Coding in the Maxillary Palp of the Malaria Vector Mosquito *Anopheles gambiae*. Curr Biol. 2007;17:1533–44. https://doi.org/10.1016/j.cub.2007.07.062.
68. Suh GSB, Wong AM, Hergarden AC, Wang JW, Simon AF, Benzer S, et al. A single population of olfactory sensory neurons mediates an innate avoidance behaviour in *Drosophila*. Nature. 2004;431:854–9. https://doi.org/10.1038/nature02980.
69. Robinett CC, Vaughan AG, Knapp JM, Baker BS. Sex and the Single Cell. II. There Is a Time and Place for Sex. PLoS Biol. 2010;8:e1000365. https://doi.org/10.1371/journal.pbio.1000365.
70. Pane A, Salvemini M, Bovi PD, Polito C, Saccone G. The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. Development. 2002;129:3715–25. https://doi.org/10.1242/dev.129.15.3715.
71. Terhzaz S, Rosay P, Goodwin SF, Veenstra JA. The neuropeptide SIFamide modulates sexual behavior in *Drosophila*. Biochem Biophys Res Commun. 2007;352:305–10. https://doi.org/10.1016/j.bbrc.2006.11.030.
72. Martelli C, Pech U, Kobbenbring S, Pauls D, Bahl B, Sommer MV, et al. SIFamide Translates Hunger Signals into Appetitive and Feeding Behavior in *Drosophila*. Cell Rep. 2017;20:464–78. https://doi.org/10.1016/j.celrep.2017.06.043.
73. Krishnan B, Dryer SE, Hardin PE. Circadian rhythms in olfactory responses of *Drosophila melanogaster*. Nature. 1999;400:375–8. https://doi.org/10.1038/22566.
74. Sakai T, Ishida N. Circadian rhythms of female mating activity governed by clock genes in *Drosophila*. Proc Natl Acad Sci U S A. 2001;98:9221–5. https://doi.org/10.1073/pnas.151443298.
75. Sparks JT, Vinyard BT, Dickens JC. Gustatory receptor expression in the labella and tarsi of *Aedes aegypti*. Insect Biochem Mol Biol. 2013;43:1161–71. https://doi.org/10.1016/j.ibmb.2013.10.005.
76. Matthews BJ, Younger MA, Vosshall LB. The ion channel ppk301 controls freshwater egg-laying in the mosquito *Aedes aegypti*. eLife. 2019;8:e43963. https://doi.org/10.7554/elife.43963.
77. Sánchez-Gracia A, Vieira FG, Rozas J. Molecular evolution of the major chemosensory gene families in insects. Heredity (Edinb). 2009;103:208–16. https://doi.org/10.1038/hdy.2009.55.

78. McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, Sang R, et al. Evolution of mosquito preference for humans linked to an odorant receptor. Nature. 2014;515:222–7. https://doi.org/10.1038/nature13964.
79. Wu Z, Yang L, He Q, Zhou S. Regulatory Mechanisms of Vitellogenesis in Insects. Front Cell Dev Biol. 2021;8:593613. https://doi.org/10.3389/fcell.2020.593613.
80. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics; 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 14 Oct 2026.
81. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8. https://doi.org/10.1093/bioinformatics/btw354.
82. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37:907–15. https://doi.org/10.1038/s41587-019-0201-4.
83. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30. https://doi.org/10.1093/bioinformatics/btt656.
84. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421. https://doi.org/10.1186/1471-2105-10-421.
85. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20:1160–6. https://doi.org/10.1093/bib/bbx108.
86. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80. https://doi.org/10.1093/molbev/mst010.
87. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37:1530–4. https://doi.org/10.1093/molbev/msaa015.
88. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9. https://doi.org/10.1038/nmeth.4285.
89. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39:783–91. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.
90. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78–82. https://doi.org/10.1093/nar/gkae268.
91. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8.
92. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Stat Methodol. 1995;57:289–300.

https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

93. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. 2025. https://www.R-project.org/. Accessed 14 Oct 2026.
94. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR et al. vegan: community ecology package. R package version 2.7-2. 2025. https://CRAN.R-project.org/package=vegan. Accessed 14 Oct 2026.
95. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. cluster: cluster analysis basics and extensions. R package version 2.1.8.1. 2025. https://CRAN.R-project.org/package=cluster. Accessed 14 Oct 2026.
96. Wickham H. ggplot2: elegant graphics for data analysis. 2nd ed. Cham: Springer; 2016. https://doi.org/10.1007/978-3-319-24277-4.
97. Perets D, Yakir E, Papathanos PA, Bohbot JD. Data and code for: Appendage identity dominates chemosensory transcriptomic organization in the black soldier fly *Hermetia illucens*. Version 1.1. Zenodo. 2026. https://zenodo.org/records/23452179
98. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001;26:32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x.
99. Anderson MJ. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics. 2006;62:245–53. https://doi.org/10.1111/j.1541-0420.2005.00440.x.
100. Rousseeuw PJ, Silhouettes. A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7.
101. Peres-Neto PR, Legendre P, Dray S, Borcard D. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology. 2006;87:2614–25. https://doi.org/10.1890/0012-9658(2006)87[2614:vposdm]2.0.co;2.
102. The Gene Ontology Consortium, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224:iyad031. https://doi.org/10.1093/genetics/iyad031.
103. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler: interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. 2023;51:W207–12. https://doi.org/10.1093/nar/gkad347.
104. The UniProt Consortium, Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023;51:D523–31. https://doi.org/10.1093/nar/gkac1052.
105. Perets D, Yakir E, Papathanos PA, Bohbot JD. Appendage-resolved RNA-seq atlas of adult chemosensory tissues in the black soldier fly, *Hermetia illucens*. NCBI BioProject PRJNA1490337. 2026. https://identifiers.org/bioproject:PRJNA1490337
106. Zhou JJ, He XL, Pickett JA, Field LM. Identification of odorant-binding proteins of the yellow fever mosquito *Aedes aegypti*: genome annotation and comparative analyses. Insect Mol Biol. 2008;17:147–63. https://doi.org/10.1111/j.1365-2583.2007.00789.x.

Figures

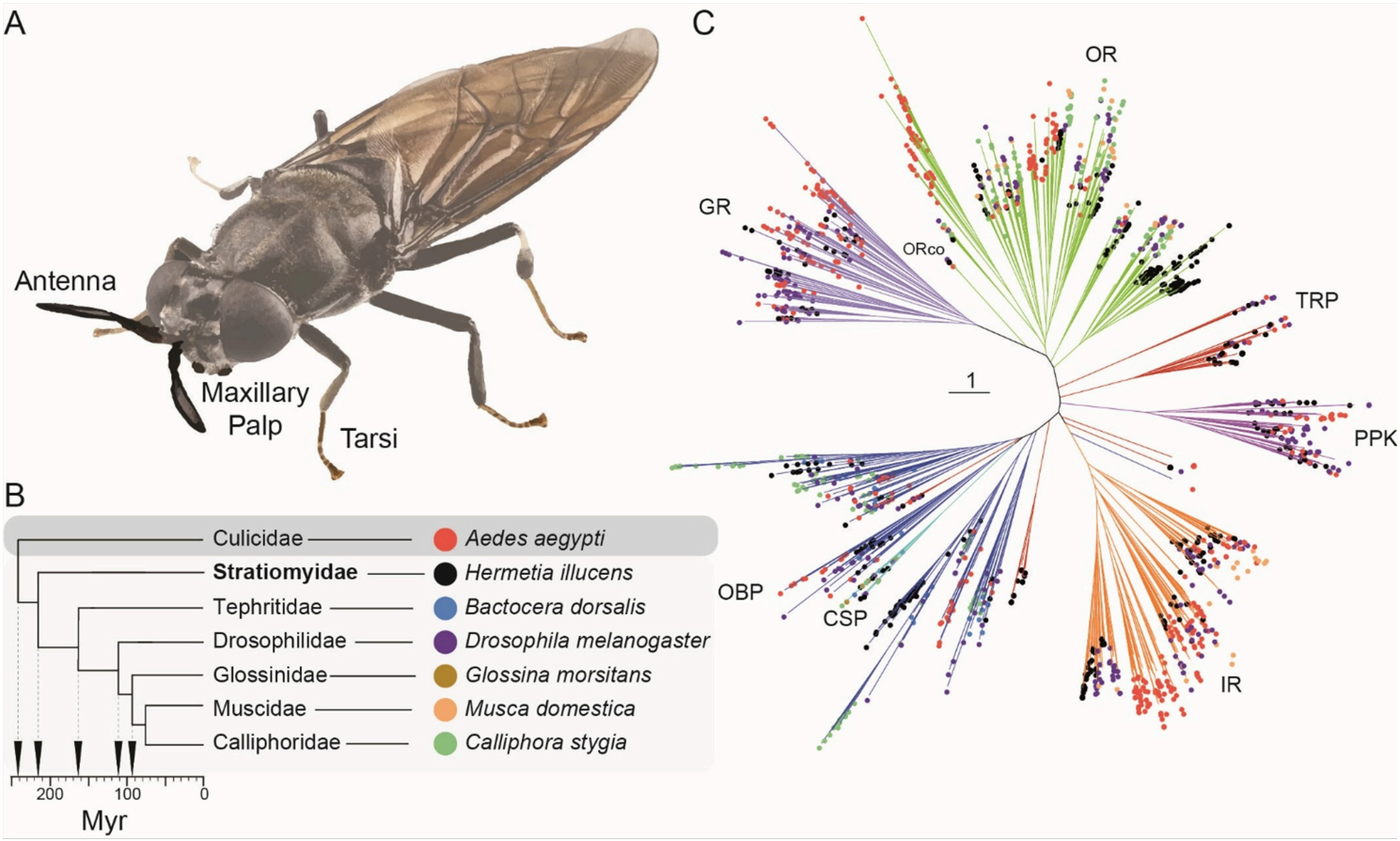

Figure 1

Phylogenetic analysis of *H. illucens* chemosensory gene families. A) Adult *H. illucens* highlighting the three chemosensory appendages analyzed in this study: antennae, maxillary palps, and tarsi (image adapted from Beta Bugs Ltd., UK). **B**) Simplified phylogeny of representative dipteran families included in the comparative analyses, illustrating the early-diverging position of *H. illucens* (Stratiomyidae) within the Brachycera. The time scale (Myr) indicates the approximate divergence of major dipteran lineages based on Wiegmann et al. [4]. **C**) Unrooted maximum-likelihood phylogeny of 1,632 chemosensory proteins, comprising 567 *H. illucens* sequences and 1,065 homologs from *Aedes aegypti*, *Drosophila melanogaster*, *Musca domestica*, *Calliphora stygia*, *Bactrocera dorsalis*, and *Glossina morsitans*. Proteins are grouped into the seven conserved chemosensory gene families: odorant receptors (OR), gustatory receptors (GR), ionotropic receptors (IR), odorant-binding proteins (OBP), chemosensory proteins (CSP), pickpocket channels (PPK), and transient receptor potential channels (TRP). Because the seven families are not homologous to one another, the combined tree is an overview of repertoire composition, and relationships between families are not interpreted. Branches are color-coded by gene family. Expanded phylogenies with protein names, orthology assignments, and bootstrap support values are provided in **Additional file 1: Figs. S1–S8**.

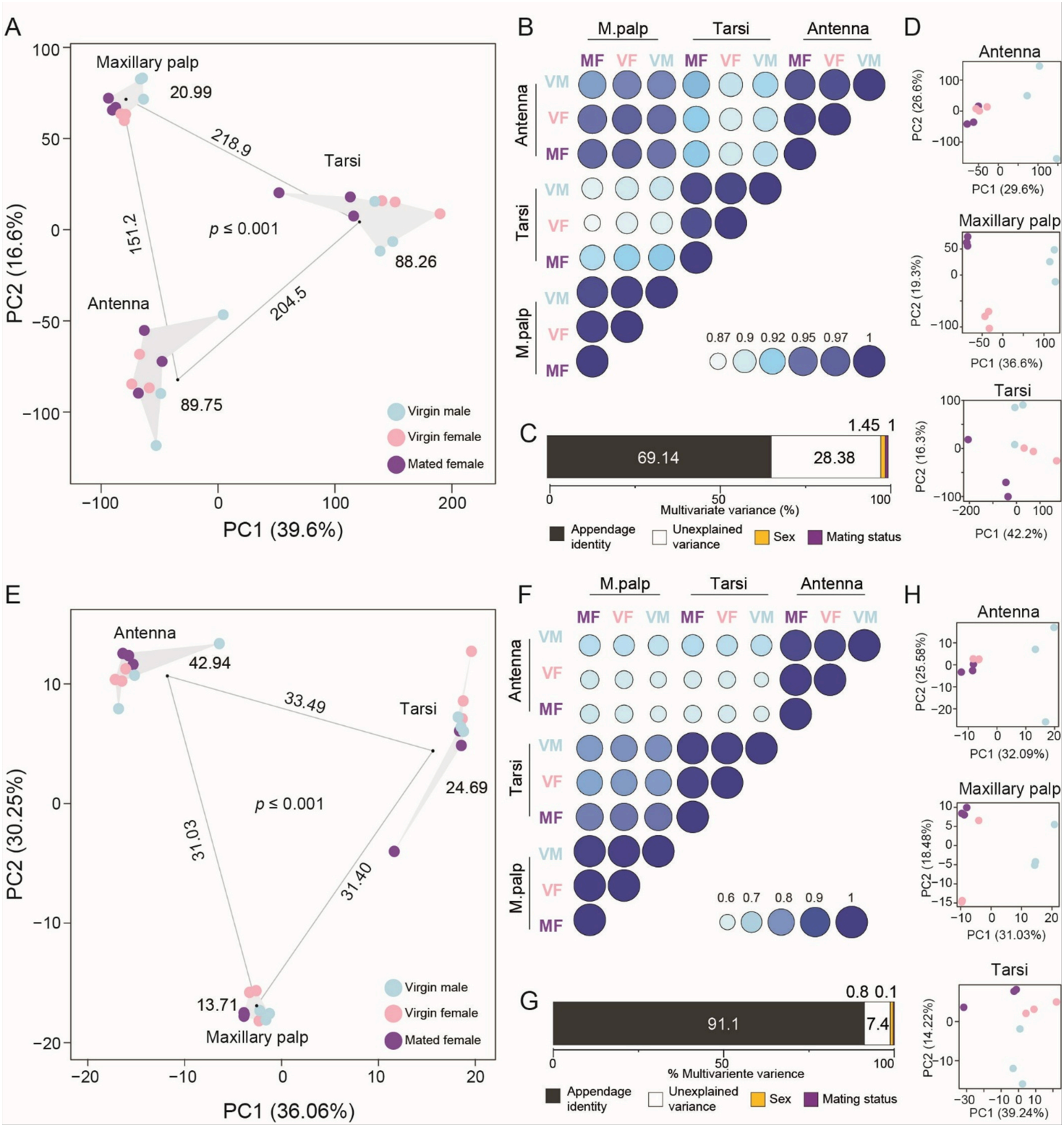

Figure 2

Appendage identity dominates transcriptomic organization in *H. illucens*. A) Principal component analysis (PCA) of global transcriptomic profiles from antennae, maxillary palps, and tarsi of virgin males (VM), virgin females (VF), and mated females (MF) (n = 3 biological replicates per group). Samples cluster primarily by appendage. Gray polygons delineate appendage clusters, numbers within polygons indicate cluster area in PC1–PC2 space, and gray lines connect appendage centroids. Numbers along

the connecting lines indicate Euclidean distances between appendage centroids in PCA space. **B**) Pearson correlation matrix of mean global gene expression profiles across appendages and reproductive groups. Circle size and color intensity are proportional to the Pearson correlation coefficient. **C**) Variation partitioning of the global transcriptome showing the proportion of multivariate variance explained by appendage identity, sex, mating status, and unexplained variance. **D**) PCA of the global transcriptome within each appendage, illustrating variation among virgin males, virgin females, and mated females. **E**) PCA of the 567 annotated chemosensory transcript models showing stronger appendage-specific separation than observed for the global transcriptome. Gray polygons, centroid connections, cluster areas, and centroid distances are displayed as in panel A. **F**) Pearson correlation matrix of chemosensory gene expression profiles. **G**) Variation partitioning of chemosensory gene expression. **H**) Within-appendage PCA of chemosensory gene expression across reproductive groups. Together, these analyses show that appendage identity is the primary determinant of both global and chemosensory transcriptomic variation, whereas sex and mating status contribute comparatively modest effects.

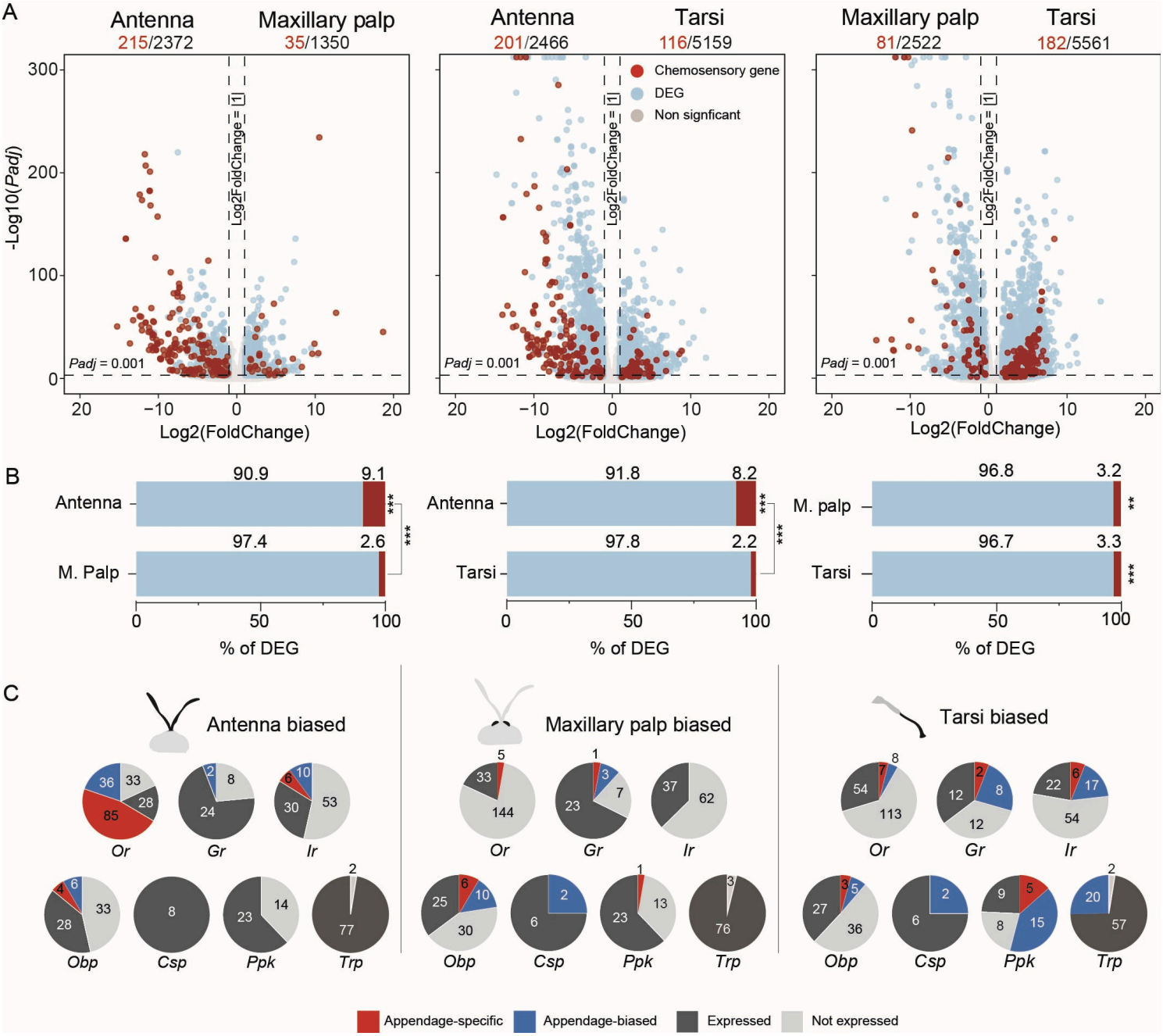

Figure 3

Appendage-specific transcriptional divergence and chemosensory specialization in *H. illucens*. A) Pairwise differential expression analyses comparing antennae, maxillary palps, and tarsi ($|\log_2 FC| \geq 1$; adjusted $p < 0.001$). Volcano plots show the genes biased to each appendage of a pair; chemosensory DEGs are shown in red, other DEGs in light blue and genes that are not significant in gray. Numbers above each plot give the chemosensory DEGs and the total DEGs biased to each appendage. **B**) Percentage of the DEGs biased to each appendage that belong to the seven chemosensory families (red) or not (blue); genes of the seven families are counted as chemosensory irrespective of their GO domain. Asterisks indicate the significance of Fisher's exact tests (**Additional file 13**). **C**) Classification of the 510 non-redundant chemosensory transcript models into appendage-specific, appendage-biased, expressed and not expressed categories across the seven chemosensory gene families (*Or*, *Gr*, *Ir*, *Obp*, *Csp*, *Ppk*,

and *Trp*). Classification criteria are described in the Methods. Together, these analyses demonstrate that each appendage possesses a distinct transcriptional identity supported by characteristic enrichment of chemosensory gene families, with *Or*s predominating in the antennae, *Obp*s in the maxillary palps, and *Ppk*, *Ir*, *Gr*, and *Trp* genes in the tarsi.

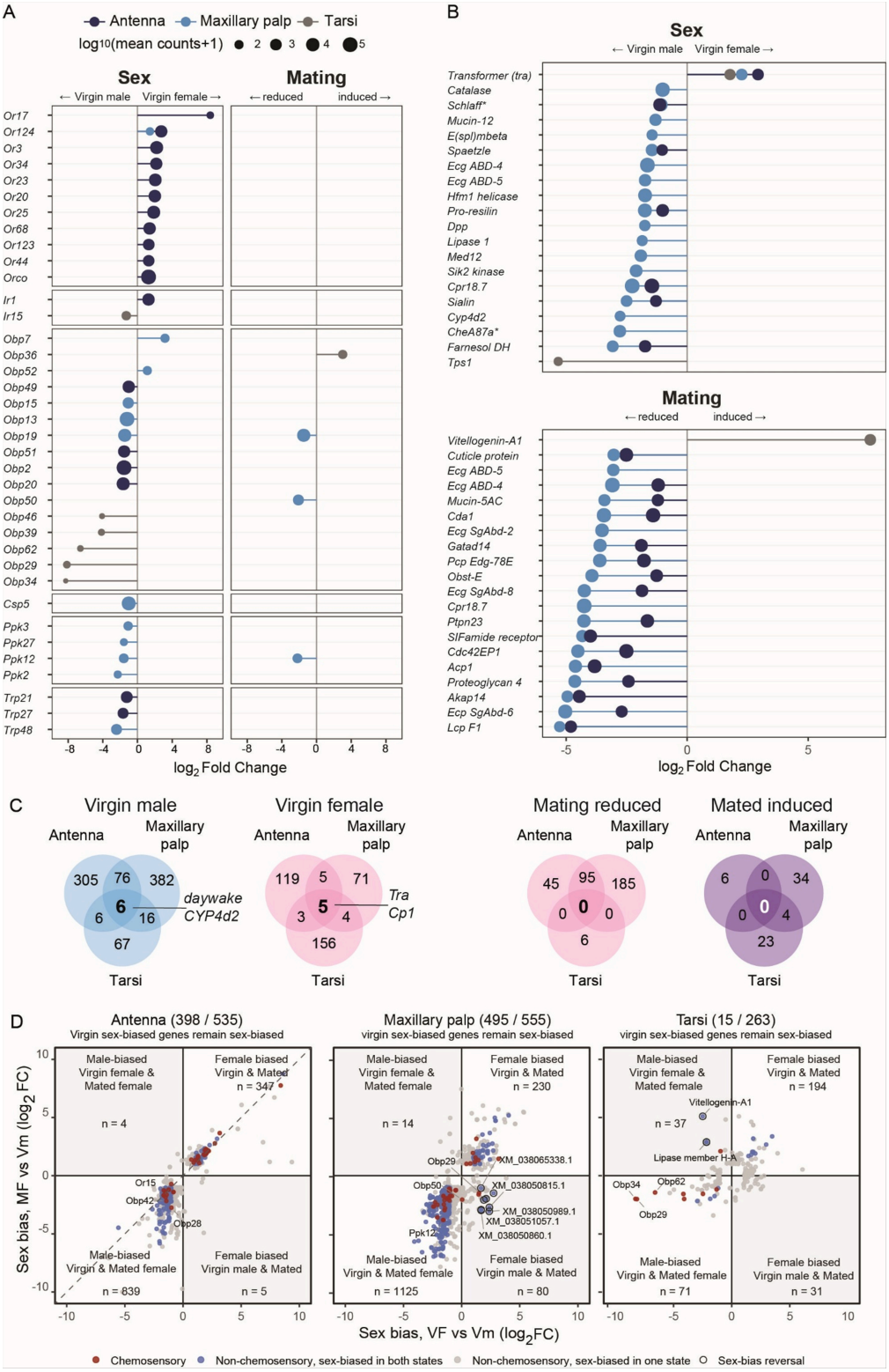

Figure 4

Sex- and mating-dependent transcriptional remodeling across adult chemosensory appendages. Genes are called differentially expressed at $|\log_2 FC| \geq 1$ and adjusted $p < 0.001$. Mating-responsive denotes significance in the same direction against both virgin females and virgin males. **A**) Chemosensory loci responding to sex (left) or mating (right), one row per gene locus, grouped by family. Stems run from no change to the $\log_2$ fold change; dot area is mean normalized expression in that appendage. Color denotes appendage: antenna, dark blue; maxillary palp, mid blue; tarsi, gray. Panels A and B share one dot-size scale. **B**) Non-chemosensory genes, the top 20 per contrast ranked by $|\log_2 FC| \times -\log_{10}$(adjusted *p*), sex above and mating below. Colors as in A. Names are abbreviated; an asterisk marks a name inferred from the best *Drosophila melanogaster* BLASTp match rather than an *H. illucens* annotation. **C**) Venn diagrams of genes shared among appendages for four programs: male-biased and female-biased in virgins, and reduced or induced in mated females. Representative genes are named beside their regions. **D**) Reorganization of sex bias by mating. Sex bias measured in virgins (virgin female versus virgin male, x) against sex bias measured in mated females (mated female versus virgin male, y), for every gene significant in either contrast. Because both axes share virgin males as the reference, a change of sign between them is a genuine reversal of sex bias. Genes are classified by their behavior across the two reproductive states: sex-biased in both states (blue), sex-biased in one state only (gray), or sex-biased in both with the direction reversed (ringed). Red marks chemosensory genes. Quadrant labels give the direction of bias in each group, and the counts give the genes in each quadrant. Axes are clipped at ±10 $\log_2 FC$, with any gene beyond drawn on the boundary.

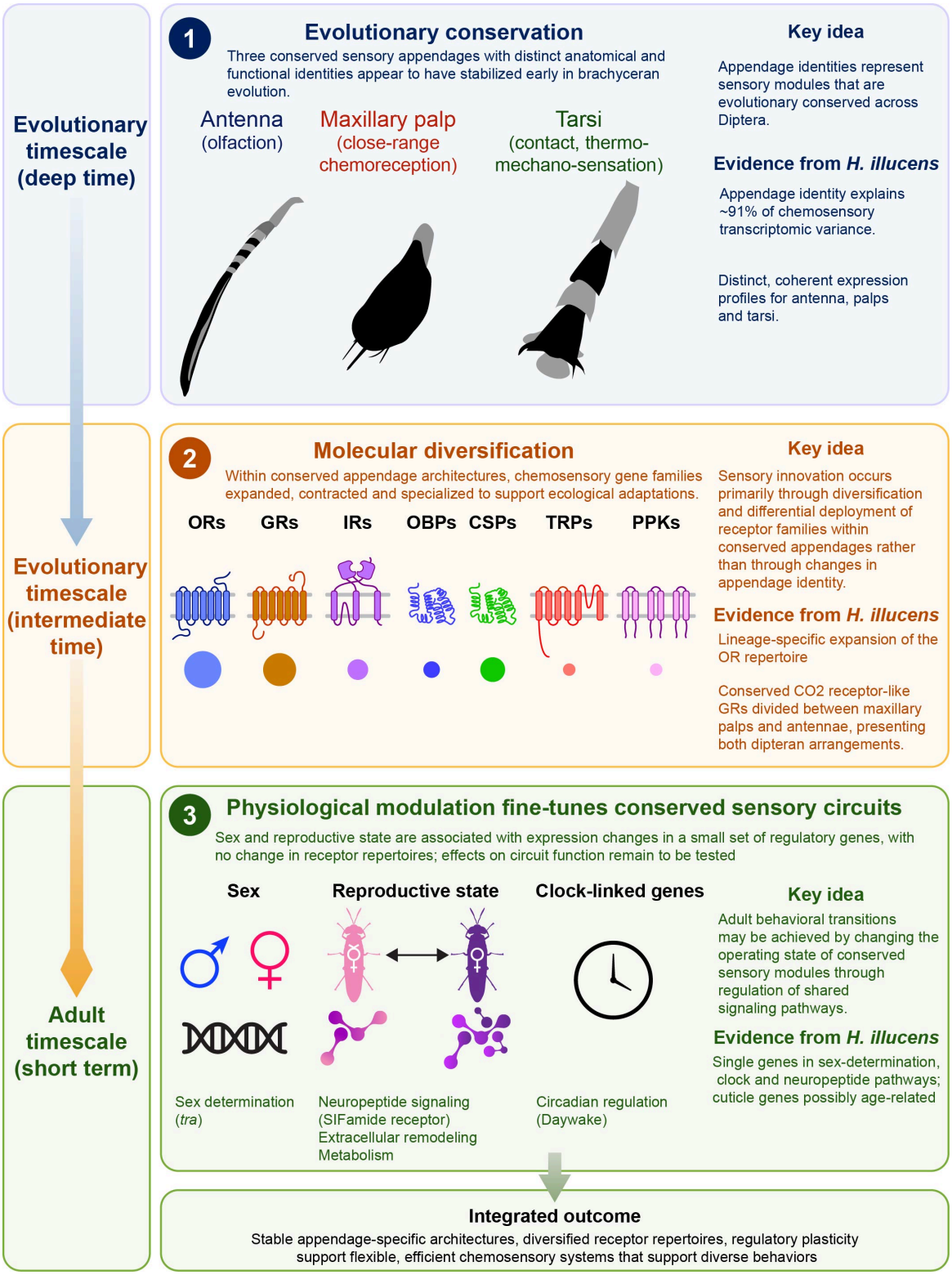

Figure 5

A hierarchical model for the evolution and modulation of peripheral chemosensory systems in Diptera. A conceptual model summarizing the evolutionary and physiological organization of the *Hermetia illucens* chemosensory system inferred from this study. The model proposes three hierarchical levels operating across distinct evolutionary timescales. (1) Evolutionary conservation (deep evolutionary timescale; proposed, not tested here): the antenna, maxillary palp and tarsi carry distinct anatomical and

functional identities, which we hypothesize became stabilized early in dipteran evolution. The present data establish that these identities are invariant across sex and reproductive state in *H. illucens*; dating their origin requires appendage-resolved data from further lineages. (2) Molecular diversification (intermediate evolutionary timescale): within these conserved appendage architectures, chemosensory gene families (*Or*, *Gr*, *Ir*, *Obp*, *Csp*, *Ppk*and *Trp*) diversified through lineage-specific expansion, contraction, and differential deployment, enabling ecological specialization without fundamentally altering appendage identity. (3) Physiological modulation (adult timescale): sex and reproductive state are associated with expression changes in a small set of regulatory genes, including the sex-determination gene *transformer*, the clock-linked gene *daywake* and the *SIFamide receptor*, together with cuticular and extracellular-matrix genes that may partly reflect adult age, with no change in receptor repertoires; effects on circuit function remain to be tested. In the tarsi, sex-biased expression is largely lost in mated females and the mating response consists mainly of induced genes, with no accompanying change in receptor-gene expression. Together, these findings support a hierarchical model in which stable appendage-specific sensory architectures provide the anatomical setting in which receptor-family diversification and physiological regulation generate behavioral flexibility.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfileslegends.docx
- Additionalfile1SupplementaryFiguresS1S24.pdf