Supplemental Text

Genomic and Transcriptomic Annotation Quality is Highly Similar. To identify possible differences in annotation quality between our transcriptomic and genomic OR gene annotation pipelines, we compared the *E. dilemma* OR set annotations based on the antennal transcriptome (35) and the whole-genome assembly (this study), respectively. While the transcriptome-based annotation resulted in a total of 86 annotated ORs (35), we were able to annotate 183 ORs based on the whole-genome assembly (Fig. 1). All antennal OR annotations were present in the genomic annotations of which nine ORs could only be incompletely annotated due to gaps in the genome assembly. With the exception of the incomplete genes, all ORs found in both annotation sets were completely identical in length and sequence, with a single exception. OR12 contained a 30 amino acid longer N terminus and was missing 14 amino acids on the C terminal end based on the transcriptome in comparison to the homologous genome annotation. Comparisons to homologous ORs in the other nine bee species in this study revealed that the transcriptome-based annotation for OR12 was incorrect. Despite this misannotation, our results show that both annotation pipelines used in this study are of equal quality but lead to a lower number of OR gene predictions in transcriptome based annotations.

Lineage-Specific Expansions in OBPs and GRs Lead to Unusually Large Gene Repertoires in Honey Bees and Bumblebees. Our analyses of five chemosensory gene families including the ORs, GRs, IRs, OBPs, and CSPs revealed that total sizes of the chemosensory gene repertoires vary widely between gene families but, with the exception of the OR family, are of similar size between species (see Main Text, Fig. 1). An exception is the honeybee, which has an increased number of 21 OBP genes compared to the 14 (*M. quadrifasciata*) to 16 (*B. terrestris* and *Ef. mexicana*) genes found in all other species. Similarly, the bumble bee has an expanded set of GRs (25 genes), which is more than 1.5 times as large than that of the species with the second largest repertoire (16 genes, *M. quadrifasciata* and *Ef. mexicana*). Based on our phylogenetic reconstructions of the chemosensory gene families, we identified multiple duplications and larger expansions in the honey bee OBP gene family and the bumble bee GR gene family. The honey bee genome carries orthologous OBPs in all but two of the 15 orthologous OBP groups (Fig. S2). In addition, we identified two lineage-specific duplications, as well as one large expansion including 9 paralogous OBPs in honey bees, without orthologs in any other bee species analyzed. Similarly, we found bumblebee GRs in each of the 13 orthologous GR groups. Three of these orthologous groups include bumble bee-specific expansions of five genes, each (Fig. S2). Interestingly, two of these expansions are accompanied by lineage-specific expansions in the stingless bee and the orchid bee *Eufriesea mexicana*, suggesting that these orthologous groups are more dynamic in terms of gene duplication than others. Most of the genes in the honey bee and bumble bee expansions do not contain pre-mature stop codons and are likely functional.

OR Gene Translocations are Ancestral.

Although the genomic location of most OR genes was conserved and generally corresponded to OR subfamilies, we identified several notable exceptions (Fig. 2a; Tab. S2). The sistergroups to all other ORs in the subfamilies G13A and G09C had a different but conserved genomic location relative to the rest of the genes within the respective subfamily. This pattern held true across all corbiculate bee groups (Fig. 2b). Additionally, we identified a translocation of an OR in subfamily G11B, where the two orthologous copies in the stingless bee and bumble bee genomes shared a different genomic location than the orthologous copies found in orchid bees and honey bees. Similarly, we found an inversion within the OR subfamily O that is shared between honey bees, bumble bees, and stingless bees but not the orchid bees. In fact, these ancestral translocations provide the first structural genomic support for a phylogenetic position of orchid bees as sister to all other corbiculate bee lineages (Node C, Fig. 1), with stingless bees + bumble bees forming a monophyletic group (Node A, Fig. 1). This topology is congruent with an earlier hypothesis based on morphological characters as well as recent phylogenomic analyses (83-85).

In addition to these genomic rearrangements shared by the common ancestor of multiple corbiculate bee lineages, we detected few translocations of single ORs in individual bee lineages. The most dynamic group in terms of genomic architecture was the OR subfamily GUnC. The genes belonging to the subfamily D were not located in a single tandem array but in species-specific genomic locations (Tab. S2). However, all duplicate OR genes observed in subfamily GUnC were located in the same taxon-specific tandem-arrays. Overall, our results indicate that during the approximately 80 million years of corbiculate bee evolution, ORs were rarely impacted by chromosomal rearrangements, resulting in a highly conserved genomic organization.

OR Gene Synteny within Tandem Arrays is Highly Conserved. We analyzed the micro-genomic architecture within each OR subfamily to determine its potential role in the observed evolutionary dynamics. OR genes belonging to the same OR subfamily tend to be clustered in tandem-arrays across conserved genomic regions (see Main Text). We determined the within tandem-array OR gene order (synteny) based on these homologous genomic regions. We restricted our analysis to the two orchid bee species with whole genome information (*Eg. dilemma* and *Ef. mexicana*) since high gene turnover rates increase the number of paralogous OR genes between distantly related species pairs, impairing the synteny analysis. All gene families with more than two OR copies in either species were in perfect synteny. The only exception was found in OR subfamily G12A where a single OR deviated from the strict syntenic pattern (Fig. S3). Species-specific duplications or larger expansions were found in five subfamilies (G02A, G04A, G09A, G12A, and G12B) revealing that duplicates were usually located directly adjacent to each other (Fig. S3). However, a small number of duplicates were located several OR genes apart, which was a result of segmental duplications of several kilo-bases long regions including up to three ORs (Fig. S3). The three smaller tandem arrays with duplicates (G04A, G09A, G12B) were dominated by species-specific paralog OR gene copies,

indicating high gene turn over. Interestingly, species-specific expansions in the two largest subfamilies (G02A and G12A) were restricted to fractions of the tandem array only (Fig. S3). For both species, genetic divergence and physical distance on the genome were positively correlated (Fig. S3; Pearson correlation coefficient: *Eg. dilemma* r^2=0.12, p=0; *Ef. mexicana* r^2=0.25, p=0).

Orthologous ORs are Affected by Positive Selection Independent of Divergence Time. In a recent study, we described strong positive selection shaping the divergence between orthologous chemosensory receptors in two closely related orchid bee species (*Eg. dilemma* and *Eg. viridissima,* Brand et al. 2015). Using our updated dataset and two additional sympatric species pairs representing divergence times between 0.15 and 13 my (Fig. 1), we further ascertained the importance of selection between orthologs including all chemosensory genes with copies found in both species of a pair. In the analysis described above we detected two, four and four out of 110, 131, and 136 orthologous chemosensory genes under positive selection between the *Eg. dilemma* – *Eg. viridissima*, *Eg. flammea* – *Eg. imperialis* and *El. bombiformis* – *El. meriana* species pair, respectively (Tab. S4). Furthermore, we conducted complementary pairwise dN/dS estimates of orthologs (22) for all three orchid bee species pairs, revealing one and three additional genes under positive selection between *Eg. dilemma* and *Eg. viridissima* and *Eg. flammea* and *Eg. imperialis*, respectively. In total, we detected 14 orthologs exhibiting signatures of positive selection in one species of a pair, of which one belonged to the OBP gene family and the remaining 13 belonged to chemosensory receptors (11 ORs, 2 GRs and 1 IR). Although genetic divergence is considerably lower between the recently diverged *Eg. dilemma* and *Eg. viridissima* pair (Fig. 1), positive selection has a similar impact on chemosensory genes in all three species pairs (*Eg. dilemma* - *Eg. viridissima*: 2.7% (3 of 110), *Eg. flammea* - *Eg. imperialis*: 5.3% (7 of 131) , *El. bombiformis* - *El. meriana*: 2.9% (4 of 136); Fisher's exact test, p>0.05). This indicates, that ORs can be subject to strong divergent selective pressures between species in the early phase of divergence.

High positive selective pressures on the GR gene family suggest highly dynamic evolution of the insect gustatory system. Our selection analysis of the five major insect chemosensory gene families (ORs, GRs, IRs, OBPs, and CSPs) in corbiculate bees revealed that the GRs evolve under significantly higher positive selection pressures in contrast to the other four gene families. We detected a significantly higher proportion of branches under positive selection in the GR gene family (p<0.0001; Fisher's Exact Test). Likewise, we found that 69.2% of the GR orthologous groups had at least one branch under positive selection, in contrast to 35.7% of all OR, 33.3% of all IR, 26.7% of all OBP, and 14.3% of all CSP orthologous groups (Tab. 1). This observation supports the hypothesis that GRs exhibit an elevated signature of positive selection in insects, formulated as a result of comparisons between the OR and GR gene family in *Drosophila* vinegar flies (12). Unlike vinegar flies, corbiculate bees show a four-fold to five-fold reduction in the number of GRs encoded by the genome (60 – 73 GRs in *Drosophila* species vs. 13 – 25 in corbiculate bees; Fig. 1). Accordingly, our results suggest that an elevation of

positive selective pressures on the GR gene family is independent of GR gene family size and a common pattern among insects.