Deciphering the combinatorial expression pattern and genetic regulatory mechanisms of Beats and Sides in the olfactory circuits of *Drosophila*

Qichen Duan[1], Sumie Okuwa[1], Rachel Estrella[1], Chun Yeung[1], Yu-Chieh David Chen[2], Laura Quintana Rio[3], Khanh M. Vien[1], Pelin C. Volkan[1, #]

1. Department of Biology, Duke University, Durham, NC 27708, USA
2. Department of Biology, New York University, New York, NY 10003, USA
3. Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10027, USA

Correspondence: pelin.volkan@duke.edu

Abstract

Over the past decades, many critical molecular players have been uncovered to control distinct steps in olfactory circuit assembly in Drosophila. Among these, multi-member gene families of cell surface proteins are of interest because they can act as neuron-specific identification/recognition tags in combinations and contribute to circuit assembly in complex brains through their heterophilic or homophilic interactions. Recently, a multi-protein interactome has been described between the Beat and Side families of IgSF proteins. Here, we use the publicly available single-cell RNA-seq datasets and newly generated gene trap transgenic driver lines to probe the *in vivo* spatial expression pattern of the *beat/side* gene families in odorant receptor neurons (ORNs) and their synaptic target projection neurons (PNs). Our results revealed that each ORN and its synaptic target PN class expresses a class-specific combination of *beat/side* genes, hierarchically regulated by lineage-specific genetic programs. Though ORNs or PNs from closer lineages tend to possess more similar *beat/side* profiles, we also found many examples of divergence from this pattern among closely related ORNs and closely related PNs. To explore whether the class-specific combination of *beats/sides* defines ORN-PN matching specificity, we perturbed presynaptic *beat-IIa* and postsynaptic *side-IV* in two ORN-PN partners. However, disruption of Beat-IIa-Side-IV interaction did not produce any significant mistargeting in these two examined glomeruli. Though without affecting general glomerular targeting, knockdown of *side* in ORNs leads to the reduction of synaptic development. Interestingly, we found conserved expression patterns of *beat/side* orthologs across ORNs in ants and mosquitoes, indicating the shared regulatory strategies specifying the expression of these duplicated paralogs in insect evolution. Overall, this comprehensive analysis of expression patterns lays a foundation for in-depth functional investigations into how Beat/Side combinatorial expression contributes to olfactory circuit assembly.

Introduction

The *Drosophila* olfactory system provides an excellent model for understanding the genetic basis of the neuronal class-specific circuit organization and assembly. The first-order sensory neurons in the circuit are olfactory receptor neurons (ORNs), which are housed in the peripheral sensilla covering the antennae (Barish and Volkan, 2015; Brochtrup and Hummel, 2011; Hong and Luo, 2014; Jefferis and Hummel, 2006; Rodrigues and Hummel, 2008). Each ORN typically expresses only a single identity-defining olfactory receptor (OR) gene or a unique combination of up to three OR genes (Barish and Volkan, 2015). The cell bodies of ORNs expressing the same OR genes, thus of the same class, are dispersed across the antenna. Yet, their axons converge onto a single, uniquely positioned class-specific glomerulus in the antennal lobe, where they synapse with the second-order projection neurons (PNs) (Barish and Volkan, 2015; Hong and Luo, 2014). ORNs are thus defined by their OR expression and glomerular target identity, comprising ~60 classes (Task et al., 2022). PNs, born from distinct neuroblast lineages surrounding antennal lobes, are also defined by their unique dendritic targeting of the dedicated glomeruli. At the same time, their axons extend to other higher brain regions, mushroom bodies, and lateral horns, which integrate the odor signal with other sensory cues and internal states to guide the animals' behaviors (Hong and Luo, 2014; Jefferis et al., 2001; Li et al., 2018; Yu et al., 2010). Importantly, within the antennal lobe, ORNs and PNs form a one-to-one match in the class-specific glomerulus. This precisely controlled organization brings many interesting questions. How are these neurons born? What molecules mediate the communication between the neurons of the same class, different classes, or pre- and post-synaptic partners? How do these neurons acquire the necessary cell surface signaling to mediate these interactions that eventually ensure the formation of this stereotyped glomerular map? As this one-to-one structure maintains over evolution from insects to mammals (Imai et al., 2010), the easy and genetically tractable *Drosophila* is a great platform to determine the lineage-specific and combinatorial expression of cell surface proteins, their establishment, function, and evolution.

Many molecular players have been identified to participate in organizing the *Drosophila* olfactory circuits, ranging from the critical regulatory hubs at the top of the developmental hierarchy, like transcription factors instructing the lineage and/or the expression of a broad range of cell surface molecules (CSMs) (Komiyama et al., 2004; Komiyama et al., 2003; Li et al., 2020a; Tichy et al., 2008; Xie et al., 2022), to the executors at the bottom of the hierarchy, which are critical CSMs themselves. These include many members of large protein families, like leucine-rich repeat superfamily members Toll-6/Toll-7 (Ward et al., 2015), Fili (Xie et al., 2019), and Capricious (Hong et al., 2009); Teneurins Ten-a and Ten-m (Hong et al., 2012; Mosca and Luo, 2014); Cadherin superfamily members Ncad (Hummel and Zipursky, 2004; Zhu and Luo, 2004), Flamingo (Arguello et al., 2021), and Fat2 (Kug) (Vien et al., 2023); Immunoglobulin Super Family (IgSF) members Dscam with extremely diverse isoform repertoire (Goyal et al., 2019; Hattori et al., 2007; Hummel et al., 2003; Zhu et al., 2006) and DIPs/Dprs, which consist of multiple paralogs (Barish et al., 2018). Our previous work showed that class-specific combinatorial expression of DIPs/Dprs organizes ORN axons within glomeruli (Barish et al., 2018). DIPs/Dprs are of particular interest because they form two multi-member subfamilies (11 DIPs and 21 Dprs), bind one another (Dprs to DIPs) primarily through heterophilic interactions, and exhibit striking cell-type-specific expression (Brovero et al., 2021; Carrillo et al., 2015; Cheng et al., 2019a; Cheng et al., 2019b;

Cosmanescu et al., 2018; Özkan et al., 2013; Wang et al., 2022). Besides the olfactory system, several trans-synaptically interacting DIPs and Dprs have been shown to control synapse selectivity and formation in the visual circuits and neuromuscular junction (Ashley et al., 2019; Carrillo et al., 2015; Courgeon and Desplan, 2019; Dombrovski et al., 2025; Menon et al., 2019; Tan et al., 2015; Venkatasubramanian et al., 2019; Xu et al., 2019; Xu et al., 2022; Xu et al., 2018).

There are around 130 IgSF-encoding genes in the Drosophila genome (Sanes and Zipursky, 2020). In addition to well-characterized DIPs/Dprs comprised of 32 members, another two families, the Beaten path (Beat) family (14 paralogs) and Sidestep (Side) family (8 paralogs) proteins (Figure 1A), are less studied. They share many features with DIPs/Dprs. They also belong to IgSF, with two or five extracellular Ig domains mediating adhesion (Figure 1B), and form a heterophilic interaction network (Figure 1A) (Li et al., 2017b; Özkan et al., 2013). Beats/Sides have been shown to control the neuromuscular junction formation in both larval and adult motor systems (de Jong et al., 2005; Fambrough and Goodman, 1996; Heymann et al., 2022; Kinold et al., 2021; Kinold et al., 2018; Pipes et al., 2001; Siebert et al., 2009; Sink et al., 2001). Very recently, studies have begun to reveal their roles in synaptic specificity and induction to assemble the adult visual system (Carrier et al., 2025; Dombrovski et al., 2025; Osaka et al., 2024; Yoo et al., 2023). However, little is known about whether and how they contribute to the olfactory circuit organization. We observed that the expression levels of *beats/sides* in antennal tissues increase over development, and most of them tend to have higher transcriptional levels in the latter half of the pupal stage, from 40h after puparium formation (APF) throughout adulthood (Figure S2A), during which the stereotyped glomerular map is being formed. Given their protein properties and known roles in neural development, we sought to illustrate the *beats/sides*' expression patterns and test their functions in building the *Drosophila* olfactory circuit.

In this study, we systematically characterized the expression pattern of *beat/side* family genes in both ORNs and PNs at single-cell and single-class levels. By analyzing the previously published single-cell RNA-seq datasets and genetically probing the native expression with *MiMIC*-based *beat/side* gene trap *GAL4* lines, we revealed that each ORN or PN class expresses a unique combination of *beats/sides*, and this *beat/side* profile is likely regulated by the lineage-specific genetic programs in a hierarchical manner. We also tested the functional relevance of one interacting pair, Beat-IIa and Side-IV, based on their matching expression pattern in partner ORNs and PNs. However, perturbation of either one pre- or post-synaptically did not result in apparent glomerular mistargeting. Nonetheless, we found pervasive ORN synapse defects when the *side* was knocked down in ORNs, suggesting its possible role in synaptic development. Interestingly, we found analogous expression principles for *beat/side* orthologs across ORNs in mosquitoes and ants, indicating the shared strategies across evolution for circuit assembly through lineage-specific regulation of cell surface protein combinations, particularly by coordinating the expression of multiple duplicated paralogs. Overall, our study reveals the combinatorial expression pattern and functional role that Beat/Side proteins play in olfactory circuit assembly and synaptic development.

Results

Genetically probing the *beat/side* expression *in vivo* by a collection of *MiMIC*-based gene trap driver lines

Recently, efforts to profile the transcriptional landscape of each cell across the whole fruit fly body have provided a valuable resource to examine the gene expression patterns in the cell types of interest. However, in the olfactory system, single-cell RNA-seq has only captured a limited portion of ORN or PN classes, leaving the transcriptome of many other ORN and PN classes unknown. To fully reveal the *beat/side* combinatorial profile, we used a *MiMIC*-based gene trap approach to generate a collection of transgenic *beat/side*-specific *GAL4* driver lines (Diao et al., 2015; Venken et al., 2011). By swapping the *GAL4* construct into 5' UTR-located *MiMIC* sites or the in-frame *T2A-GAL4* construct into introns between two coding exons, we could make GAL4 hijack the expression of the host gene (Figure 1C and Materials and Methods). We then obtained *GAL4* driver lines for 13 of the 14 *beat* members and 7 of the 8 *side* members, except *beat-Vb* and *side-VII* (Figure 1D, E). Notably, *beat/side* genes generally have one to four annotated isoforms, and our *beat/side-MiMIC GAL4* collection is expected to trap all isoforms of each gene except *beat-IIb*, of which the *GAL4* only captures the expression of one of two isoforms (Figure 1E). This near-complete driver line collection reveals the remarkable enrichment of *beat/side* expression in neurons, from larval, pupal, to adult stages, in both peripheral and central nervous systems (Figure 1D; Figure S1). At the gross brain level, *beats/sides* are differentially expressed across different brain regions, including the antennal lobe, central complex, mushroom body, etc. (Figure 1D). This driver collection thus provides a valuable toolkit to study *beat/side* functions in diverse neuronal contexts in the olfactory circuits and beyond. Next, we used intersectional genetic strategies to restrict the reporter expression to ORNs or PNs, which allowed us to map the *beat/side* expression across glomeruli innervated by ORNs or PNs, respectively.

A glomerular map of *beat/side* expression in ORNs

Our antennal bulk RNA-seq through pupal development shows that *beats/sides* are generally expressed at higher levels at later stages of glomerular formation, reaching the adult levels by mid-pupal stages (Figure S2A). To delineate the *beat/side* expression in each ORN class, we examined the publicly available single-ORN RNA-seq datasets from three distinct developmental stages: 24h APF (early pupal stage), 42-48h APF (mid-pupal stage), and adulthood (Li et al., 2020a; McLaughlin et al., 2021). At 24h APF, ORN axons have arrived at the antennal lobes and chosen a medial versus lateral antennal lobe trajectory, which positions them en route to their future glomerular regions. By the mid-pupal stage, the glomerular targeting is almost completed, and ORNs start forming synapses with their matching postsynaptic PNs. By the beginning of the adult stage, a stereotypical and discrete glomerular map has formed, and all glomeruli are now discernible. We visualized these datasets in bubble plots showing the fraction of positive cells and mean expression of genes of interest of each ORN class (Figure S2B). This reveals that each ORN class possesses a unique signature of *beat/side* combinations throughout the developmental stages (Figure S2B). Combining bulk antennal RNA-seq and ORN single-cell RNA-seq results, we found that *sides* are generally expressed at higher levels in ORNs than *beats* (Figures S2A, B). Some *beats*, including *beat-Va*, *beat-Vb*, and *beat-Vc*, are barely expressed in any cell types within

the antenna, whereas *beat-Ic* and *side-VIII* appear to be predominantly expressed in non-neuronal cells of the antenna but not in ORNs (Figure S2A, B).

As many ORN classes are not captured in the single-cell RNA-seq datasets, we used the *MiMIC-GAL4* lines to probe the ORN expression of different *beats/sides in vivo* in 3-5-day-old adult brains as the proxy of their developmental expression. We leveraged the *eyeless*-driven FLP recombinase to excise the *STOP* cassette from *UAS-FRT-STOP-FRT-mCD8.GFP*, such that membrane-localized GFP expression by *beat/side-GAL4* can be restricted to ORNs. This way, we can label any glomeruli innervated by ORNs expressing each *beat/side* while excluding the signal from post-synaptic PNs, where *eyeless* has no expression. We generated a near-complete glomerular *beat/side* expression map across all ORN classes (Figure 2A). Consistent with bulk and single-cell RNA-seq, *beat-Ic*, *beat-Va*, and *side-VIII* are not expressed in any ORN classes. Some family members, like *beat-Ib*, *beat-VI*, and *side-IV,* are only sparsely expressed in very few ORN classes (Figure 2A). On the other hand, other *beats/sides* are expressed in a much broader pattern, though at varying levels among different ORN classes (Figure 2A).

Based on the GAL4 labeling patterns, we binarized the expression of all examined *beat* and *side* genes, as positive (value = 1) or negative (value = 0) for each gene in each ORN target glomerulus. We summarized the results and hierarchically clustered the genes by their expression pattern and ORN classes by their *beat/side*-expression profile, as shown in Figure 2B. We observed three types of gene expression patterns for *beat* and *side* genes across ORN classes: (1) broadly expressed, (2) expressed at a restricted pattern, or (3) not expressed (Figure 2B). More interestingly, ORN classes from the same lineage tend to be clustered based on the combinatorial *beat/side* expression (Figure 2B). This suggests a lineage-specific mechanism in patterning the ORN class-specific combinatorial expression of *beats/sides*. We decided to investigate this in further detail.

Lineage-regulated genetic programs specify the combinatorial expression of *beats/sides* in ORNs

ORNs differentiate from precursor cells via a multi-step hierarchical genetic program (Barish and Volkan, 2015; Li et al., 2018). Early larval and pupal patterning factors first prepattern antennal imaginal discs (Royet and Finkelstein, 1997), followed by the sensillar type assignment by critical transcription factors, including Lozenge (Lz), Atonal, and Amos (Figure S2C) (Goulding et al., 2000; Gupta et al., 1998; Gupta and Rodrigues, 1997; zur Lage et al., 2003). We first examined the bulk antennal RNA-seq datasets profiling the transcriptional changes in mutants of *amos* (lacking trichoid and basiconic ORNs) or *atonal* (lacking coeloconic ORNs) (Menuz et al., 2014; Mohapatra and Menuz, 2019). We indeed found a couple of *beats/sides* that are downregulated in *atonal* and *amos* mutant antennae (Figure S2D), suggesting these *beats/sides* are enriched in the cells housed in Amos/Atonal-specified sensillar types. A few genes are also upregulated in *amos* and *atonal* mutants (Figure S2D). These *beat* and *side* genes are either transcriptionally upregulated or are now enriched in the mutants due to changes in ORN ratios.

Upon sensillar type selection, sensillar subtype fates are further specified by an additional set of transcription factors like Rotund (Rn), Dachshund (Dac), and Engrailed (En) (Figure S2C) (Barish and Volkan, 2015; Blagburn, 2008; Li et al., 2016; Li et al., 2013; Song et al., 2012). Finally, within

203 each sensillar subtype (or sensillum), one multipotent precursor cell undergoes several
204 sequential asymmetric divisions that eventually produce one to four different ORNs expressing
205 distinct ORs and projecting to distinct glomeruli and four non-neuronal supporting cells (Barish
206 and Volkan, 2015; Li et al., 2018). This step is achieved by iterative recruitment of Notch signaling
207 bifurcating cell fates, coupled with other mechanisms like epigenetic modifiers and late
208 transcription factors (Barish and Volkan, 2015; Chai et al., 2019; Endo et al., 2007) (Figure S2C).
209 We here present a simplified decision tree to illustrate the kinships between ORNs and to
210 compare *beat/side* combinations based on lineage relationships (Figure S2C): ORNs mapped to
211 the same class, defined as "within ORN class"; ORNs of different classes but housed in the same
212 sensillar subtype, described as "within sensillar subtype"; ORNs from different sensillar subtypes
213 but belonging to the same sensillar type, namely, "within sensillar type"; and the furthest kinship,
214 ORNs housed in distinct sensillar types, i.e., "between sensillar types". We compared the
215 similarity of *beat/side* combinatorial expression between cells falling into these ORN kinship
216 categories to obtain insights into the developmental regulation of *beat/side* expression.

217 Firstly, based on our gene-ORN class expression matrix derived from the genetic labeling data
218 (Figure 2B), we calculated the similarity (Pearson's correlation) of the *beat/side* expression vector
219 between each ORN class pair, and found significantly higher similarity between ORN classes
220 belonging to the same sensillar types than those from different sensillar types (Figure 2C). In
221 contrast, this pattern is not observed from the randomized control, where we shuffled the ORN
222 class identity with the expression vector (Figure 2D). Additionally, we found this difference also
223 holds for the single-cell RNA-seq datasets. By calculating the similarity (measured by Spearman's
224 correlation) according to gene expression between each single ORN, we found that, as expected,
225 ORNs mapped to the same class display the highest pairwise similarity of *beat/side* profile in all
226 stages (Figure 2E; Figure S2E). Moreover, ORNs within the same sensillar type appear to possess
227 more similar *beat/side* profiles than ORNs from different sensillar types, in contrast to the shuffled
228 control (Figure 2E, F), supporting that the lineage-intrinsic mechanisms set the *beat/side*
229 expression.

230 However, we found that *beat/side* combinatorial expression tends to diverge between ORNs
231 housed in the same sensillar subtype (Figure S2E). We observed even lower pairwise similarity
232 according to the *beat/side* expression between these cells than between cells from different
233 sensillar subtypes but sharing the identical sensillar type (Figure S2E). This observation stands
234 true for all three stages (Figure S2E). In contrast, at 24h APF and mid-pupal stage, the similarity of
235 the pan-CSM profile and the whole transcriptome linearly increases with closer kinship (Figure
236 S2E). It suggests that the overall transcriptional profile and the cell surface gene expression
237 patterns are primarily set by lineage-specific factors, such that cells with closer kinship tend to
238 have more similar transcriptomic profiles and overall cell surface codes. Furthermore, *beats/sides*
239 expression adds new levels of complexity to cell surface signals, potentially diversifying ORN-
240 specific glomerular decisions, particularly among the most developmentally related ORNs in the
241 same sensillar subtype. Interestingly, at the adult stage, this divergence also exists for overall CSM
242 and cell-specific transcriptional profiles (Figure S2E). These results suggest that at the end of
243 development, ORNs with the closest kinship may have acquired broader transcriptional variance,
244 which may support the establishment of discrete glomerular maps necessary for olfactory
245 transduction, odor discrimination, and processing to drive odor-guided behaviors.

While the single-cell RNA-seq-based analysis only reflects the trend for the limited cell types captured in that dataset, our genetic labeling-based comprehensive analysis of *beat/side* across all ORN classes reveals the existence of both convergence and divergence: ORNs from the same sensillar subtype exhibit binary *beat/side* similarity: they could be clustered together, suggesting their shared *beat/side* combinatorial expression; they could also be segregated, indicating the divergence of *beat/side* profile from genetically very close ORNs (Figure 2B). We therefore propose a model: in some sensillar subtypes, at the step of ORN terminal selection, *beat/side* combinatorial expression is diversified from a putative initial lineage-specific "template" to ensure the closely related ORNs gain more variation in these cell surface molecular repertoire, whereas in some sensillar subtypes, *beat/side* profile is pre-specified and these closely related ORNs show similar *beat/side* combinations.

Collectively, our thorough characterization of *beat/side* expression from both single-cell RNA-seq and gene trap *GAL4* intersectional labeling suggests that Beats/Sides could be the cell surface "executors" underlying the hierarchical genetic programs that specify the ORN fate, likely participating in fine-tuning the ORN wiring.

A glomerular map of *beat/side* expression in PNs

ORNs synapse with dedicated PNs within each glomerulus to assemble the olfactory circuit, and Beats/Sides are involved in synaptogenesis (Osaka et al., 2024; Yoo et al., 2023). We also need to survey the *beat/side* expression on the PN side, which may inform whether Beats/Sides could function trans-synaptically to mediate ORN-PN matching. We started with the bulk PN RNA-seq (Li et al., 2020b) from the FACS-sorted PNs at 36h APF and adult stage. In general, *beats/sides* appear to be expressed at comparable levels between these two time points (Figure S3A). Next, we mined the published single-cell RNA-seq data (Li et al., 2017b; Xie et al., 2021) for developing PNs to gain insights into the PN type-specific expression of *beats/sides*. This dataset captured PNs at 0h APF, 24h APF, 48h APF (mid-pupal), and adult stage. Some PNs are born embryonically and are part of the larval olfactory circuit (Jefferis et al., 2001; Lin et al., 2012; Marin et al., 2005; Yu et al., 2010). At the beginning of the pupal stage, these PNs first prune their terminal axons and dendrites and re-extend their neurites to be integrated into the adult olfactory circuit following other larvally born PNs (Marin et al., 2005). From 0h to 24h APF, PNs project their axons to the mushroom body and lateral horn and their dendrites to the antennal lobes. PNs' dendrites create a prototypical glomerular map before ORN axons arrive in the antennal lobes at around 24h APF. Thereafter, PNs begin to match their presynaptic ORN partners. And starting from the mid-pupal stage, they build synapses, refine the discrete glomerular organization, and finally mature the olfactory circuit by the end of the pupal stage (Jefferis et al., 2001). In terms of PN class-specific expression of *beats/side*s at different time windows, we found: (1) *beats/sides* are enriched at 24h APF and thereafter, while expressed at minor levels at the beginning of metamorphosis (Figure S3B); (2) compared with ORNs, *beats/sides* are generally expressed in a broader pattern, and all *beats/sides* are expressed in PNs while some *beats* and *side-VIII* are undetectable in ORNs (Figures S2B, S3B). Suppose Beats/Side mediate adhesion between ORNs and PNs; this observation raises a hypothetical model where broadly and less specifically located postsynaptic

287 surface “locks” are matched by sparsely and more specifically distributed presynaptic surface
288 “keys”.

289 Next, we used these *GAL4* lines to drive the GFP reporter expression in postsynaptic PNs by
290 removing the *STOP* cassette from *UAS-FRT-STOP-FRT-mCD8.GFP* with the PN-specific *GH146-*
291 *flippase*. We found generally broader expression of *beats/sides* across PN-labeled glomeruli, with
292 a few sparsely distributed ones, like *beat-Ic* and *side-IV* (Figure 3A). We then summarized the
293 binary expression of all *beats/sides* examined across all glomeruli identified (Figure 3B).
294 Hierarchical clustering of PN types according to their *beat/side* expression shows, in general, a
295 correlation between *beat/side* combination and their PN lineages (Figure 3B). We therefore further
296 investigated this.

297

298 **Lineage-regulated genetic programs specify the combinatorial expression of *beats/sides* in**
299 **PNs**

300 PN fates are specified separately from ORNs by a sequential genetic program from three distinct
301 neuroblast lineages (Figure S3C) (Jefferis et al., 2001; Li et al., 2018). A common neuroblast
302 sequentially produces one ganglion mother cell and a self-renewed neuroblast. The ganglion
303 mother cell further divides into one terminal PN, which innervates a dedicated glomerulus, and
304 another daughter cell, which undergoes cell death (Lin et al., 2012; Lin et al., 2010). The younger
305 neuroblast continues this process. Thus, different PN classes are born in a sequentially
306 stereotyped order (Jefferis et al., 2001; Lin et al., 2012; Yu et al., 2010). Based on this, we defined
307 the relationship between PNs into three categories: (1) PNs from the same lineage targeting the
308 same glomerulus, (2) PNs from the same lineage targeting different glomeruli, and (3) PNs from
309 different lineages targeting different glomeruli.

310 We again asked whether the *beat/side* profile similarity between PN classes is based on their
311 developmental kinship. Calculating the PN class-specific *beat/side* expression similarity based on
312 the *GAL4* labeling expression matrix showed higher similarity between PN classes of the same
313 lineage than of different lineages (Figure 3C, D). Similarity analysis for *beat/side* expression in the
314 single-PN RNA-seq datasets also confirmed these observations (Figure 3E, F; Figure S3D). In the
315 same lineage, as expected, PNs targeting the same glomeruli also have more similar *beat/side*
316 expression than PNs targeting different glomeruli (Figure S3D). This is also true for the expression
317 of CSMs and the cellular transcriptional profiles at all four stages (Figure S3D). This suggests that,
318 as part of their cell surface repertoire, the *beat/side* profiles in PNs are primarily set by the lineage
319 and PN class-specific genetic mechanisms.

320 As PNs in each lineage are born in a stereotyped order, we also examined whether PN birth order
321 determines its *beat/side* profile. We focused on the adPN lineage and calculated the correlation
322 between each PN pair. We plotted their *beat/side* profile and sorted the PN classes in this lineage
323 based on the reported birth order (Figure 3G). We also plotted the Pearson’s correlation efficiency
324 between each pair of PN classes (Figure 3H). We found that the early-born PNs have more
325 divergent *beat/side* profiles while later-born PNs gradually obtain a shared Beat/Side
326 combinatorial “template” which gets less complex as new PN fates are generated (Figure 3G, H).
327 This trend indicates that PN precursor-specific factors set the *beat/side* expression potential of

328 daughter PNs and undergo a temporal fate specification along with the birth sequence of PNs. It is
329 also interesting that, similar to ORNs (Figure 2B, C), closely related PNs can possess shared or
330 divergent *beat/side* combinatorial expression, suggesting that different intrinsic genetic programs
331 may be present in different precursor cells to control the opposing *beat/side* expression similarity
332 in their daughter cells.

333

334 **Beats/Sides form a robust molecular interaction network between ORN and PN partners**

335 The first identified Beat-Ia and Side pair act trans-synaptically to promote synapse formation at
336 the neuromuscular junction (Fambrough and Goodman, 1996; Kinold et al., 2021; Kinold et al.,
337 2018; Siebert et al., 2009; Sink et al., 2001). Recent work also identified additional Beat-Side
338 interactions, including Side-IV and Beat-IIa/b, mediating neuronal recognition, connection
339 specificity, and synaptogenesis in the fly visual system (Osaka et al., 2024; Yoo et al., 2023). We,
340 therefore, hypothesized that Beat and Side interactions between ORNs and PNs can also regulate
341 the synaptic matching. Guided by the known Beat-Side interactions based on the *in vitro*
342 biochemical assays, we examined the *beat/side* expression between partner ORN and PN
343 targeting the same glomerulus (Figures 1A, 4A). This analysis revealed a putative Beat/Side-
344 mediated interactome between ORNs and PNs, which is promiscuous, redundant, and complex.

345 While several putative Beat-Side interactions exist between partner ORN and PN classes, we
346 found one elegant and straightforward case: presynaptic Beat-IIa and postsynaptic Side-IV. *beat-*
347 *IIa* is expressed in just a few ORN classes, including Or42b ORNs targeting one posterior
348 glomerulus DM1 and Or92a ORNs targeting one anterior glomerulus VA2 (Figure 4B). In parallel,
349 *side-IV* expression is restricted in PNs, including PNs projecting to DM1 and VA2 glomeruli (Figure
350 4C). Next, we tested whether perturbing *beat-IIa* expression in ORNs or *side-IV* expression in PNs
351 results in glomerular targeting defects. To do this, we used a pan-ORN driver, *peb-GAL4,* to knock
352 down *beat-IIa* in all ORNs while labeling Or42b ORNs or Or92a ORNs as a readout of glomerular
353 targeting. However, we did not observe any gross phenotype in glomerular integrity, morphology,
354 position, or ectopic mistargeting (Figure 4D, E). We also used a PN-specific driver, *GH146-GAL4,*
355 to express *side-IV RNAi* and test if Or42b or Or92a ORNs target their glomeruli normally. Again, we
356 found no visible differences between knockdowns and controls (Figure 4F, G). These results
357 suggest that either Beat-IIa/Side-IV interaction in matching ORNs and PNs is unnecessary for
358 glomerular targeting or is compensated by other Beat and Side proteins with redundant functions
359 expressed in the same neurons.

360 Indeed, the combinatorial *beat/side* codes can distinguish ORNs targeting DM1 or VA2 glomeruli
361 from ORNs innervating the neighboring glomeruli (Figure 4H), such that removing *beat-IIa* in ORNs
362 doesn't ambiguate the CSM repertoire and still preserves a level of diversity among the local axon
363 fibers in a glomerular neighborhood. Similarly, PNs targeting DM1 or VA2 also harness this robust
364 combinatorial "barcode", and deleting *side-IV* in PNs may not abolish the dendritic diversity in the
365 glomerular neighborhood (Figure 4I). To test for general defects in glomerular organization, we also
366 conducted a pan-ORN driver *peb-GAL4* driven *UAS-RNAi* screening against *beat* and *side* genes
367 that are expressed in ORNs (Figure S4A). This screen revealed mild defects in the VA1v glomerulus
368 targeted by Or47b ORNs by knockdown of *beat-IIb, beat-IIIa/b/c,* and *side-III*, and a low-
369 penetrance phenotype of split DM3 glomerulus targeted by Or47a ORNs (Figure S4A, B). However,

370 additional controls showed that the VA1v phenotypes of *beat-IIIa/b/c* perturbation appeared to
371 arise from the transgenic background that sensitizes the VA1v glomerular disorganization (Duan et
372 al., 2023)(Figure S4C). These negative results might be because the Beat/Side combinatorial
373 expression can theoretically act as a robust cell surface recognition code for each ORN or PN
374 class, given its error-tolerance feature.

375 Given the lack of glomerular targeting defects, we next asked whether *beat* and *side* genes have
376 roles in regulating synapse development and maintenance within glomeruli. For this, we used a
377 previously generated transgenic reporter, *UAS-SynLight*, which expresses the presynaptic active
378 zone marker, Bruchpilot-Short (sBrp), and the neurite membrane marker mCD8.GFP in the same
379 transcript (Aimino et al., 2023). We expressed this transgene in ORNs to label and quantify the
380 synaptic density in the antennal lobes while doing perturbation in ORNs. We calculated the ratio
381 of the mean fluorescence intensity of RFP and GFP in the antennal lobe as a proxy of synaptic
382 density. This way normalizes the potential variability in transcription levels, imaging conditions, or
383 neuron size across groups. As inhibiting neural activity has been reported to decrease synapse
384 numbers, we first silenced ORN neuronal activity by overexpressing the mutant *shaker* channel
385 EKO. We observed synaptic density reduction globally in the antennal lobe, compared with the
386 control group where an inactive variant of tetanus toxin light-chain (TNT-) was introduced instead
387 (Figure S4D). This confirmed the feasibility of using this transgenic reporter to measure synapses
388 and the effect of neural activity on synaptogenesis. We next used *peb-GAL4*-mediated *RNAi*
389 knockdown of Side*,* which, like many other *side* genes, is expressed in the majority of ORNs
390 (Figures 2A, B, 4A; Figure S2B), and interacts with four Beat partners (Figure 1A). We also knocked
391 down *side-VIII*, which is not expressed in ORNs (Figure 2A, B; Figure S2B). We found a pervasive
392 synaptic reduction in *side* knockdowns compared to the control *RNAi*, but not *side-VIII*
393 knockdowns (Figure S4E). Interestingly, analysis of previously reported antennal RNA-seq data
394 (Deanhardt et al., 2023) from *Or47b* and *Or67d* mutants revealed differentially expressed *beat* and
395 *side* genes (Figure S4F). These results suggest that *side* genes might contribute to ORN-specific
396 synapse development and maintenance. Notably, the transcriptional changes induced by OR-
397 dependent neural activity in addition to lineage factors, can contribute to the final combinatorial
398 expression and synaptic function of Beat and Side proteins.

399

400 **Evolutionarily conserved expression of *beat/side* orthologs across insect ORNs**

401 Next, we sought to interrogate the ORN class-specific expression of *beats/sides* in other species
402 to gain some evolutionary insights into the regulation and function of these families of IgSF
403 proteins. Recently reported transcriptome atlas of ORNs in two additional insect species, yellow
404 fever mosquitoes *Aedes aegypti*, and clonal raider ants *Ooceraea biroi*, allowed us to perform
405 comparative analyses of *beat/side* expression in the peripheral olfactory tissue at single-cell
406 resolution. Yellow fever mosquito and clonal raider ants belong to the order Diptera and a distant
407 order Hymenoptera, and diverged from the fruit fly *Drosophila melanogaster*, a Dipteran species,
408 260 million and 300 million years ago, respectively (Figure 5A). We first set out to identify the bona
409 fide orthologs of *beats/sides* in mosquitoes and ants. To do this, we queried the protein sequence
410 of *Drosophila* Beat-Ia and Side against the proteome database of two other species and ran
411 multiple sequence alignment after selecting the hits with comparable lengths and similar

412 AlphaFold-predicted structures (two Ig domains for Beats; five Ig domains and one Fibronectin
413 domain for Sides, also see Materials and Methods, Figure S5A). We then built the phylogenetic
414 tree of Beat/Side protein orthologs across the three species. This analysis yielded 13 *beat* genes
415 and 10 *side* genes in the yellow fever mosquito, while the clonal raider ant has 10 *beats* and 9
416 *sides* in its genome (Figure 5B, C). We observed several notable gene duplication/loss events in
417 these species. For example, there is a single ant gene in the Beat-I clade, whereas there are three
418 copies in mosquitoes and flies (Figure 5B). In contrast, three ant genes encode the Side-IV clade
419 and three mosquito genes encode the Side-II clade, while only one fruit fly gene is in each clade
420 (Figure 5C). Additionally, the *side-V* gene appears to be lost in clonal raider ants (Figure 5C). The
421 phylogenetic tree indicates that: (1) gene duplication events generated ancestral *beat/side*
422 paralogs earlier than insect species divergence; (2) additional gene duplication/loss events also
423 occurred post-divergence of fruit fly, mosquito, and ant species.

424 Based on the phylogenetic tree, we renamed the mosquito and ant genes according to their
425 sequence similarities with the fruit fly orthologs. We examined their expression in ORNs using
426 published single-cell datasets (Brahma et al., 2023; Herre et al., 2022). Strikingly, we found
427 combinatorial expression principles of *beat/side* genes in different ORN classes across the three
428 insect species are conserved (Figure 5D, E, and F; Figure S5B): (1) genes that are not expressed in
429 *Drosophila* ORNs, like *beat-Va/b/c*, *beat-VI*, and *side-VIII*, are also not expressed in ORNs of yellow
430 fever mosquitoes (*beat-Va/b*, *beat-VI*, *side-VIII*) or clonal raider ants (*beat-V*, *beat-VIa/b*, *side-VIII*);
431 (2) genes that are broadly expressed in *Drosophila* ORNs, like *beat-VII*, *side-II*, and *side-III*, are also
432 broadly expressed in ORNs of yellow fever mosquitoes (*beat-VII*, *side-IIa/b*, and *side-III*) and clonal
433 raider ants (*beat-VII*, *side-II*, and *side-III*); (3) genes that are expressed in a restricted pattern
434 across *Drosophila* ORNs, like *beat-IIa/b*, beat-*IV*, and *side-IV* (Figure 2A, B), are also expressed in
435 only a portion of ORN clusters in mosquitoes (*beat-IIa*, beat-*IVa/b*, and *side-IV*) and ants (*beat-II*,
436 beat-*IV*, and *side-IVb/c*). Even though there is significant conservation of *beat/side* gene
437 expression patterns across ORNs, there are also a few exceptions. For example, *beat-Ia*,
438 expressed in many ORN classes in fruit flies and at high levels in restricted ORN classes in
439 mosquitoes, is barely expressed in ant ORNs. In addition, *side* is expressed in many *Drosophila*
440 ORN classes but is undetectable in mosquito and ant ORNs (Figure 5D, E, and F). These
441 comparative analyses suggest that both the evolutionarily conserved and plastic *beat* and *side*
442 genes in peripheral sensory neurons likely emerged due to genome evolution events that deleted
443 or duplicated the members of the *beat* and *side* gene families and the regulatory sequences.

444 Interestingly, recent single-cell RNA-seq reports from clonal raider ant ORNs revealed regulatory
445 differences in sensory receptor genes, where multiple *OR* transcripts are expressed in each ORN.
446 In detail, ORNs that share a more similar transcriptome, and thus are clustered together in UMAP
447 plots, usually express *ORs* located in a tandem array in the genome. In each ORN, a single start
448 site is selected that transcribes *OR* genes downstream, while only the most upstream mRNA in
449 the transcript is translated into a functional OR protein. This means the ORNs in each cluster are
450 transcriptionally heterogeneous, containing multiple ORN classes defined by the functional
451 expression of a singleton OR protein. To zoom in on ORN classes within each single-cell RNA-seq
452 cluster, we plotted *beat/side* expression and the *ORs* of the tandem expressed in the cluster of
453 each single cell as heatmaps (Figure 5G; Figure S5D). We found that each ORN class, defined by
454 the expression of the most upstream *OR*, exhibits class-specific expression of *beat/sides*. This

455 highlights the similarities in the lineage mechanisms coupling *OR* transcription initiation and
456 *beat/side* expression specification between the clonal raider ant and fruit flies.

457

458 Discussion

459 In this study, we comprehensively delineated the expression pattern of *beat/side* family genes in
460 ORNs and PNs at single-cell and single-class levels. We found that each ORN or PN class exhibits
461 a unique combinatorial expression of *beats/sides*, and this *beat/side* profile appears to be
462 specified by the lineage-intrinsic genetic programs. In addition to analyzing the single-cell
463 transcriptome of ORNs and PNs, we also generated *MiMIC*-based gene trap driver lines to probe
464 the *beat/side* expression across ORNs and PNs *in vivo*. This approach enabled mapping the
465 *beat/side* expression to additional ORN/PN types not covered in the single-cell RNA-seq datasets.
466 Collectively, we found that *beat/side* profiles can diverge between some closely related ORNs or
467 some closely related PNs, suggesting that the Beat/Side combination may add variability to cell
468 surface codes, which biases their glomerular targeting. We also investigated one interacting pair,
469 Beat-IIa and Side-IV, based on their matching expression pattern between partner ORNs and PNs,
470 and found no apparent glomerular mistargeting in knockdowns of *Beat-IIa* in ORNs or *Side-IV* in
471 PNs. Yet, knockdown of the *side* gene in ORNs resulted in diminished synapses in the antennal
472 lobes, suggesting a role in synapse development. Moreover, we found evolutionarily conserved
473 expression patterns and ORN-specific combinatorial signature principles for *beat* and *side*
474 orthologs in the olfactory systems of mosquitoes and ants, suggesting the shared genetic
475 programs establishing the *beat/side* profile. In sum, our studies implicate ORN/PN class-specific
476 combinatorial Beat and Side protein expression and their roles in olfactory circuit assembly.

477

478 Developmental regulation of *beat/side* expression in ORNs and PNs

479 We found the lineage-correlated expression profile of *beats/sides* in ORNs and PNs to be a general
480 principle, suggesting the lineage-specific programs in setting the *beat/side* combinatorial
481 expression. Our genetic labeling results and single-cell RNA-seq analyses also reveal the
482 existence of both shared and divergent expression patterns of *beat/side* between closely related
483 ORNs. We found some sibling ORNs have different *beat/side* profiles and project to distant
484 antennal lobe regions, while some maintain the putative "template" combinations found in many
485 ORNs and project to neighboring or close glomeruli. Adjacently born PNs can also have shared or
486 divergent *beat/side* expression profiles. We therefore propose that: (1) at the higher level of the
487 developmental hierarchy, i.e., in ORN sensillar type decisions or PN lineage specification
488 decisions, the *beat/*side "template" is set by lineage-specific factors; (2) the latest common
489 precursor of ORNs (i.e. sensilla subtype precursors) or PNs may have diverse intrinsic factors that
490 bias the expression of additional *beat/side* combinatorial expression while some do not; (3) such
491 that at the final terminal selection decision, Beat/Side combinations may mediate the synapse
492 programs in closely related ORNs or PNs. Furthermore, expression of additional cell surface
493 proteins can increase the cell surface protein repertoire, further diversifying neuronal
494 identification tags displayed on neurons utilized during circuit assembly. This two-step model,
495 setting the lineage-specific cell surface molecular template, followed by within-lineage

diversification, seems to be an efficient way to establish the cell type-specific surface molecular repertoire by genetically deterministic programs. This might be a developmental strategy to coordinate the combinatorial expression of duplicated paralogs.

Molecular complexity and redundancy of Beats/Sides as cell surface "codes"

Side was first identified as the postsynaptic ligand, expressed in muscles, to attract motor neurons. Motor neurons expressing the receptor Beat-Ia follow the Side-labeled muscles and leave axonal fasciculations to innervate muscle targets (de Jong et al., 2005; Fambrough and Goodman, 1996; Pipes et al., 2001; Siebert et al., 2009; Sink et al., 2001). The embryonic expression pattern analysis expanding to other Beat/Side family paralogs also suggests that Beats are neuronal receptors for Sides expressed on peripheral tissues (Li et al., 2017b). For example, Side-VI is expressed in muscle fibers while its receptors, Beat-Vs, are expressed in motor neurons (Li et al., 2017b). This evidence jointly points to the fact that Sides are present primarily in postsynaptic sites, signaling to presynaptic Beats to promote synaptogenesis. Outside of the neuromuscular junction, recent reports in the fly visual circuits also suggest presynaptic Side-II and postsynaptic Beat-VI mediate synaptic recognition (Yoo et al., 2023). Additional studies revealed presynaptic (photoreceptor-expressed) Side-IV can induce synaptogenesis by acting as both ligands and receptors when interacting with postsynaptic Beat-IIa/b, forming synaptogenic complexes through their interactions with Beat-IIa/b and coreceptor Kirre (Osaka et al., 2024).

In the olfactory circuits consisting of ORNs and PNs, we found that *beat* and *side* genes are expressed in the same neurons. ORNs generally lack Beats but are abundant for Sides. This would point to a model where Sides primarily interact with PN-expressed Beats. However, ORNs do express a few Beats with their binding partner Sides as well as some orphan Beats. And PNs generally express both Sides and Beats. Similarly, DIPs and Dprs can also be expressed in the same cells, like adult ORNs (Barish et al., 2018) and larval motor/sensory neurons (Wang et al., 2022). There are several possible functional explanations based on Beat and Side interactions. First, Sides and Beats may have different subcellular localizations that position places of protein-protein interactions and cellular functions. For example, Beats might localize to PN dendrites to interact with Sides on ORN axon terminals, while Side proteins localized to PN axon terminals may interact with Beats expressed on the dendrites of third-order neurons. If this is the case, this binary subcellular preference may vary in different cell types to mediate interactions with other neurons. In motor neurons, Beats are preferentially sorted to axon terminals, whereas in the central nervous system or sensory neurons, Sides are preferentially sorted to axon terminals. Second, Beats/Sides can function in both dendrites and axons simultaneously and act as ligands and receptors bidirectionally. This points to a model in which the combinatorial code of multiple surface proteins establishes synaptic specificity recognition, likely by providing differential adhesive forces. Third, Beats and Sides in the same neuron might interact *in cis* on the same membrane in addition to their interactions *in trans* with Beat/Side proteins on other neurons. Interestingly, a recent study showed that DIP-Dpr interactions favor *cis* over *trans* when present in the same cell (Morano et al., 2025). It might also be the case that when Beats/Sides are localized together in the same membrane regions, it is the competition between *cis* and *trans* interactions with all present binding partners in the proximity that sets the "net" Beat/Side surface codes.

538 Additionally, combinations of cell surface adhesion molecules can form multi-molecular
539 complexes, so the absence of any component may not abolish but just reduce the recognition
540 efficacy or shift these complexes to new stoichiometric configurations (Honig and Shapiro, 2020;
541 Sanes and Zipursky, 2020). Given this sophisticated context, it is not surprising that removing a
542 single gene is insufficient to change glomerular organization. In terms of Beat/Side combinatorial
543 complexity, deleting *side-IV* in PNs or *beat-IIa* in ORNs doesn't ambiguate DM1 from adjacent
544 glomeruli (Figure 4), and ORN-PN synaptic matching can still occur. Thus, the promiscuous and
545 redundant functions of the Beat/Side family have the capacity to form an error-tolerant adhesion
546 code, enabling recognition robustness. These are supported by our efforts to perturb a single
547 *beat/side* gene that caused minor to no defects (Figure S4A). Previously reported screens for
548 CSMs mediating ORN-PN synaptic matching also didn't report any defects in glomerular targeting
549 with any *beat/side* knockdowns (Xie et al., 2019). Our expression dataset provides a roadmap to
550 further crack the combinatorial coding nature of Beats and Sides by multiplexing perturbation.

551 Intriguingly, we observed weakened synapses in the antennal lobes when we knocked down the
552 *side* gene, which suggests that Side and Beat protein interactions might mediate synapse
553 formation in the olfactory circuits. This is supported by other recent studies in the visual system
554 reporting Side-IV as a synapse-inducing protein (Osaka et al., 2024), and that Beat-VI expression
555 gradients regulate synaptic density gradients through their interactions with Side-II (Dombrovski et
556 al., 2025). These results suggest that Beat/Side interactions may inherently have synaptogenic
557 functions and participate in the subglomerular synaptic organization during olfactory circuit
558 development. Furthermore, since many *beats/sides* show persistent high expression levels from
559 the pupal stage to adulthood, they might be recruited for synapse induction, formation, and
560 stabilization, and modulated by OR signaling and/or neuronal activity.

561

562 **Evolutionarily conserved expression of *beats/sides* in ORNs of insects**

563 We observed similar combinatorial expression patterns across ORN clusters among recently
564 duplicated paralogs within the same species, and orthologs in ants and mosquitoes, suggesting
565 evolutionarily conserved mechanisms that regulate their expression in ORNs. We speculate that
566 ancestral paralogs of *beats/sides* with their regulatory elements acquired cell-type-specific
567 expression patterns, which were generally retained during evolution and shared between species.
568 More recent gene duplication events after speciation generated additional paralogs within each
569 *beat/side* clade, where we observe both shared (*Drosophila beat-IIIa/b* and *beat-Va/b/c*, *O. biroi*
570 *beat-VIa/b*, *A. aegypti beat-IVa/b* and *side-IIa/b*) and divergent expression patterns (*A. aegypti*
571 *beat-Ia1/a2* and *side-IIb/c*, O. *biroi side-IVa/b/c*). Though some *beat* and *side* genes exhibit
572 evolutionary plasticity in their expression patterns, conservation of the expression of many
573 *beat/side* genes in ORNs appears to be a general rule. This implies that there are selection
574 pressures on the operative roles that these genes play in the ORN development and function.

575

576 **Limitations of the study**

577 Many *beat/side* genes transcribe multiple splice isoforms, which adds an additional layer of
578 molecular diversity. This study didn't examine the expression of *beat* and *side* splice isoforms in

579 the olfactory system. Our transgenic *GAL4* driver lines generally trap the common regions across
580 isoforms and cannot address isoform-specific expression. Single-cell RNA-seq approaches that
581 resolve *beat* and *side* splice isoforms or expression analysis of isoform-specific *GAL4*s would be
582 needed to this end.

583 This study only reveals the *beat* and *side* expression at the transcriptional level and does not
584 address the Beat and Side protein-level distributions. Some studies reported large discrepancies
585 between transcripts encoding surface molecules and their protein abundance in developing PNs
586 (Li et al., 2020b), suggesting regulation at the level of translation, protein stability, modification,
587 sorting, and delivery to the cell surface.

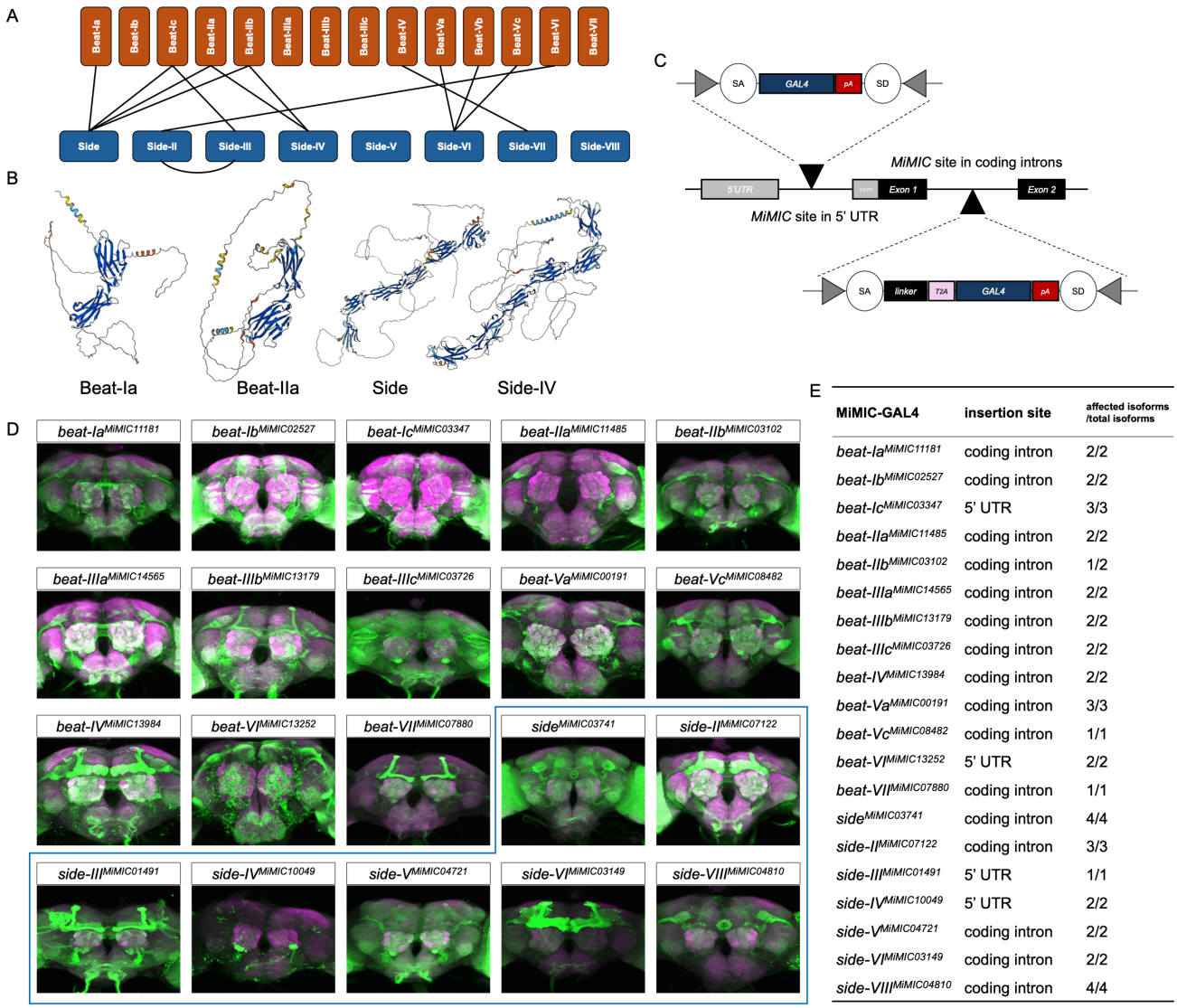
588

589 **Future directions**

590 In the future, it will be interesting to investigate how regulatory sequence evolution contributes to
591 the conserved and divergent expression patterns of *beat* and *side* paralogs across species.
592 Additional biochemical characterization of Beats/Sides protein-protein interactions in different
593 insect species will provide novel insights into the co-evolution of coding sequences with
594 regulatory sequences to reconcile their functional evolution in circuit assembly.

595 Understanding the biochemical features of Beat and Side proteins will greatly inform the
596 functional analysis in the circuit assembly context. This includes developing a more sensitive and
597 specific test to deorphanize Beat-Side interactions, characterizing whether Beats/Sides are
598 transmembrane proteins or membrane-anchored like DIPs/Dprs (Lobb-Rabe et al., 2024), and
599 testing whether they interact *in trans* and/or *in cis*. If they are transmembrane proteins, then what
600 are the functions in synapse formation? How are these functions mediated by intracellular and
601 extracellular protein domains? Proximity labeling-based proteome profiling like BIO-ID can
602 identify these protein interactors, which will inform further functional investigation. A recent study
603 reported that the N-terminal domain of Beat-Ia is localized to the cell surface while the C-
604 terminus accumulates in the nucleus; this implies that Beat-Ia might undergo proteolytic cleavage
605 for proper function (Heymann et al., 2022). Future biochemical, genetic, and neurophysiological
606 studies will help reveal the function of Beat and Side proteins in recognition specificity and
607 synaptic transmission.

608



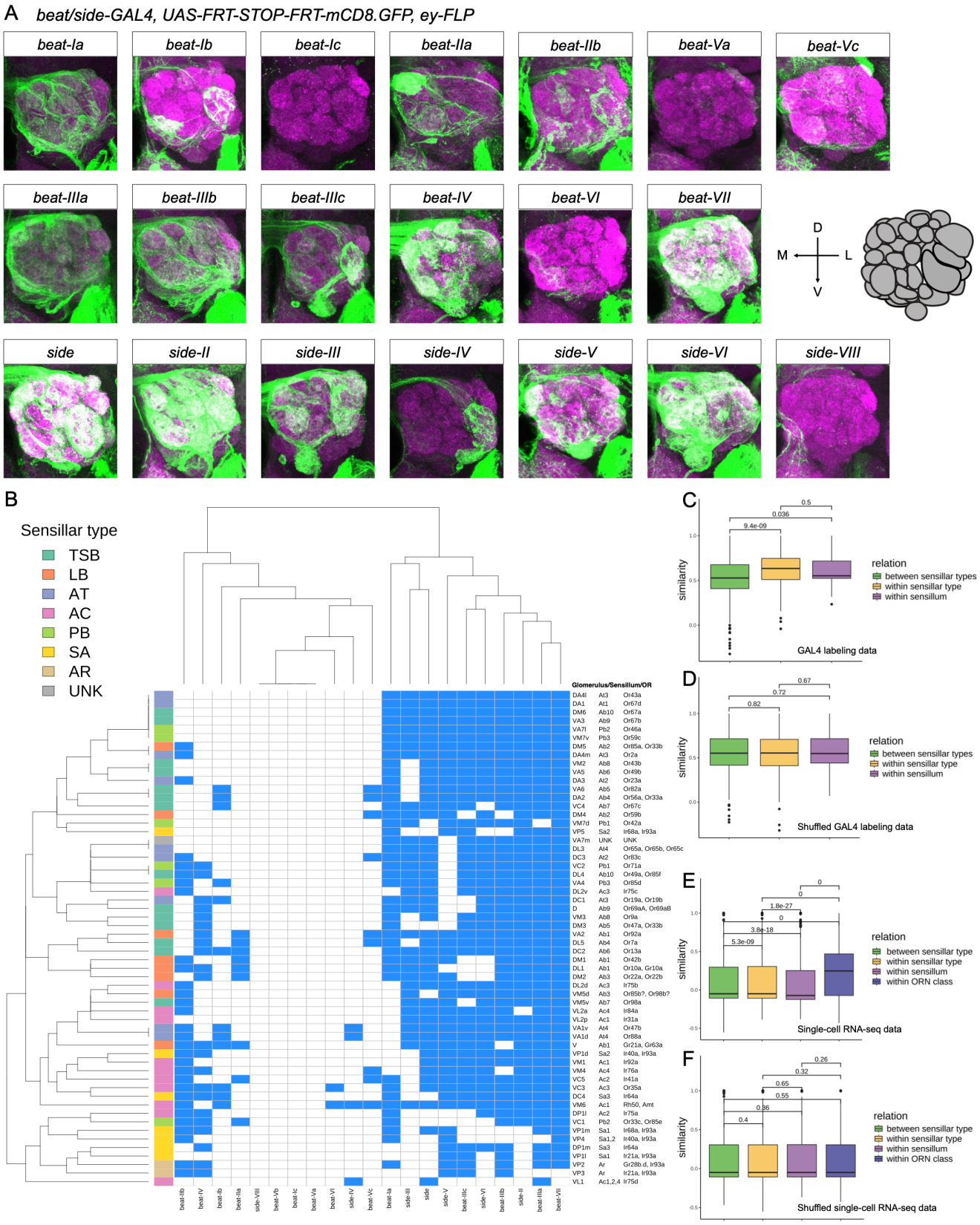
MiMIC-GAL4	insertion site	affected isoforms /total isoforms
beat-Ia^{MiMIC11181}	coding intron	2/2
beat-Ib^{MiMIC02527}	coding intron	2/2
beat-Ic^{MiMIC03347}	5' UTR	3/3
beat-IIa^{MiMIC11485}	coding intron	2/2
beat-IIb^{MiMIC03102}	coding intron	1/2
beat-IIIa^{MiMIC14565}	coding intron	2/2
beat-IIIb^{MiMIC13179}	coding intron	2/2
beat-IIIc^{MiMIC03726}	coding intron	2/2
beat-IV^{MiMIC13984}	coding intron	2/2
beat-Va^{MiMIC00191}	coding intron	3/3
beat-Vc^{MiMIC08482}	coding intron	1/1
beat-VI^{MiMIC13252}	5' UTR	2/2
beat-VII^{MiMIC07880}	coding intron	1/1
side^{MiMIC03741}	coding intron	4/4
side-II^{MiMIC07122}	coding intron	3/3
side-III^{MiMIC01491}	5' UTR	1/1
side-IV^{MiMIC10049}	5' UTR	2/2
side-V^{MiMIC04721}	coding intron	2/2
side-VI^{MiMIC03149}	coding intron	2/2
side-VIII^{MiMIC04810}	coding intron	4/4

Figure 1. Beats/Sides are heterophilically interacting IgSF proteins.
(A) Schematics showing the known interactions between Beat and Side proteins.
(B) AlphaFold-predicted structures of example Beat and Side proteins. Beats have two extracellular Ig domains, while Sides' extracellular parts consist of five Ig domains and one Fibronectin type III domain. Transmembrane and intracellular domains are poorly predicted. Structures were downloaded from AlphaFold.ebi.ac.uk (Jumper et al., 2021).
(C) Schematic showing *MiMIC*-based insertion of the GAL4-coding construct into either the 5' UTR or the "coding" intron, which will hijack the expression of the host gene. If *GAL4* is inserted into introns within the 5' UTR, the *GAL4* artificial exon will first be kept from splicing and then will be translated as it has its own start codon. If the in-frame *T2A-GAL4* construct is inserted into an intron between two coding exons. In that case, this whole artificial exon will also be prevented from splicing and then translated along with the upstream coding sequence. During translation, the *T2A* sequence causes the ribosome to fail at synthesizing the peptide bond, and thereby, a truncated peptide of the host gene and a GAL4 protein will be produced separately. In both cases, GAL4 is expected to be expressed in the same cell where the native gene is expressed. SA, splice

acceptor; SD, splice donor; pA, Hsp70 polyadenylation signal to terminate transcription; linker sequence ensures *T2A-GAL4* will be translated in frame.

(D) Expression pattern of each *beat/side* in the brain revealed by *beat/side-MiMIC-GAL4*. Green is the GFP signal driven by the gene trap GAL4 and magenta is anti-Ncad staining showing the brain architecture.

(E) Table summary of the insertion site and the isoforms labeled by each *beat/side-MiMIC-GAL4*.


A
beat/side-GAL4, UAS-FRT-STOP-FRT-mCD8.GFP, ey-FLP
beat-Ia
beat-Ib
beat-Ic
beat-IIa
beat-IIb
beat-Va
beat-Vc
beat-IIIa
beat-IIIb
beat-IIIc
beat-IV
beat-VI
beat-VII
D
M
L
V
side
side-II
side-III
side-IV
side-V
side-VI
side-VIII
B
Sensillar type
TSB
LB
AT
AC
PB
SA
AR
UNK
Glomerulus/Sensillum/OR
C
similarity
relation
between sensillar types
within sensillar type
within sensillum
GAL4 labeling data
D
Shuffled GAL4 labeling data
E
between sensillar type
within ORN class
Single-cell RNA-seq data
F
Shuffled single-cell RNA-seq data

632

633 **Figure 2. Lineage-dependent combinatorial expression of *beats/sides* across ORN classes.**
634 **(A)** A glomerular map of *beat/side* expression in ORNs revealed by transgenic *GAL4* driver lines.
635 Glomerular structures are stained by anti-Ncad (magenta). Anti-GFP staining (green) highlights the
636 glomeruli innervated by the indicated *beat/side*-positive ORNs. All antennal lobes shown are right
637 antennal lobes. M, medial; L, lateral; D, dorsal; V, ventral. Each image is shown as the intensity
638 projection over the Z-axis.
639 **(B)** Hierarchical clustering of *beat/side* expression across ORNs based on the GAL4 labeling
640 analysis. Row and column dendrograms are based on hierarchical clustering results. ORN class
641 (named by its target glomerulus), the sensillum (sensillar subtype) where the ORN is housed, and
642 the corresponding ORs expressed are shown on the right. The left row label is color-coded based
643 on the ORN lineage. Blue means positive for the given gene expression, whereas white denotes
644 negative. TSB, thin and small basiconics; LB, large basiconics; AT, antennal trichoids; AC, antennal
645 coeloconics; PB, palp basiconics; SA, sacculus; AR, arista; UNK, unknown. VM6 glomerulus has
646 been recently shown to be comprised of three subglomeruli, VM6v, VM6m, and VM6l, innervated
647 by ORNs from Ac1 sensillum and Sacculus chamber III, respectively. As it is difficult to distinguish
648 Ac1-originated VM6v from the other two without additional colabeling, we counted VM6 as a
649 single intact glomerulus for simplicity. We did not have *beat-Vb-GAL4* driver line, but inferred they
650 are not expressed in any ORNs based on the bulk and single-cell RNA-seq.
651 **(C)** Similarity (Pearson's correlation) of *beat/side* expression between ORN class pairs based on
652 the expression matrix of (B) in three relation categories.
653 **(D)** Similarity (Pearson's correlation) of *beat/side* expression between ORN class pairs based on
654 the expression matrix of (B), with row labels shuffled in three relation categories.
655 **(E)** Similarity (Spearman's correlation) of *beat/side* expression between ORN cell pairs based on
656 the single-cell RNA-seq in four relation categories.
657 **(F)** Similarity (Spearman's correlation) of *beat/side* expression between ORN cell pairs based on
658 the single-cell RNA-seq with ORN class annotation shuffled in four relation categories. In (C) to
659 (F), P values from the Mann-Whitney U test of each comparison are shown.
660

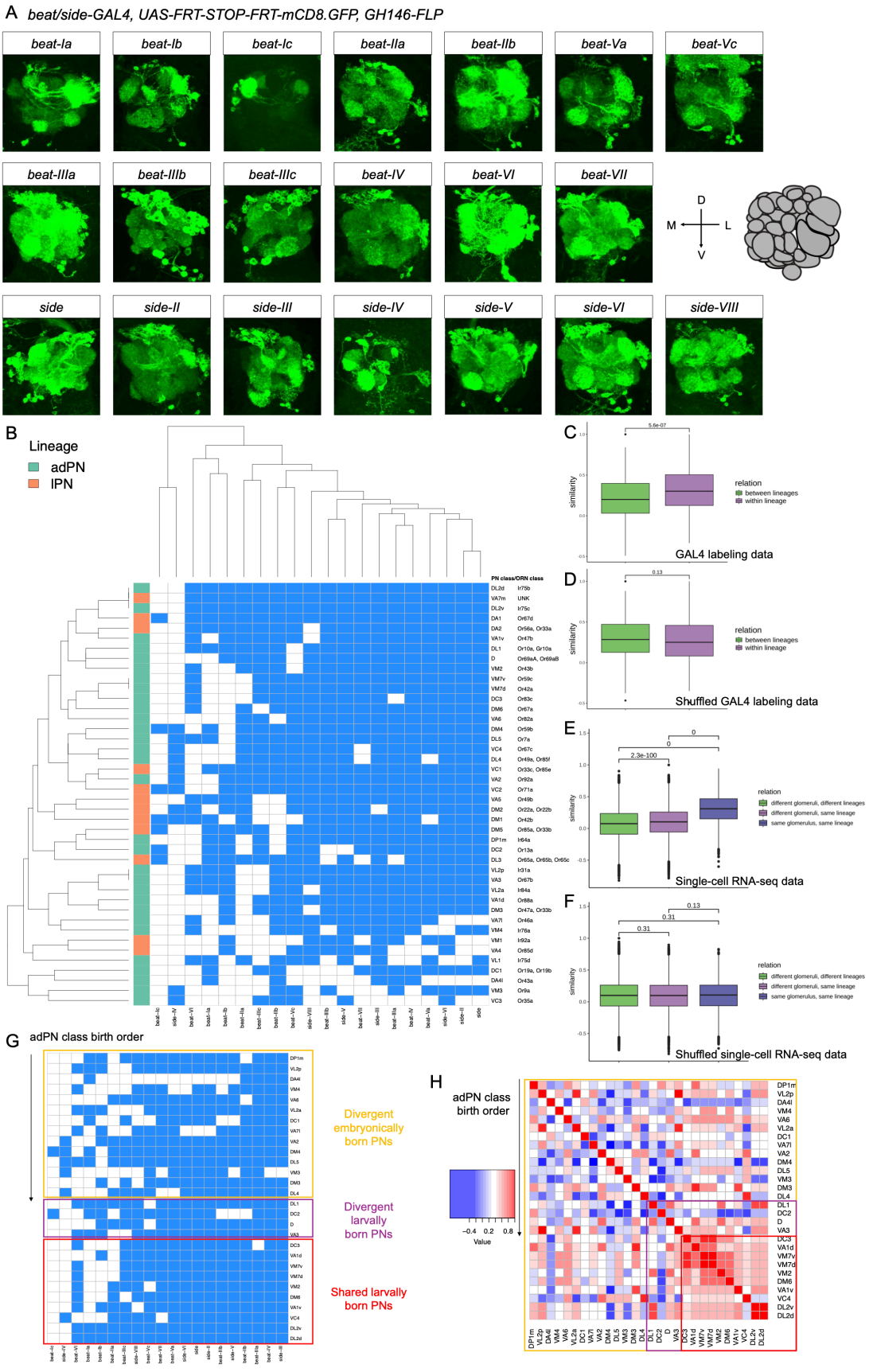
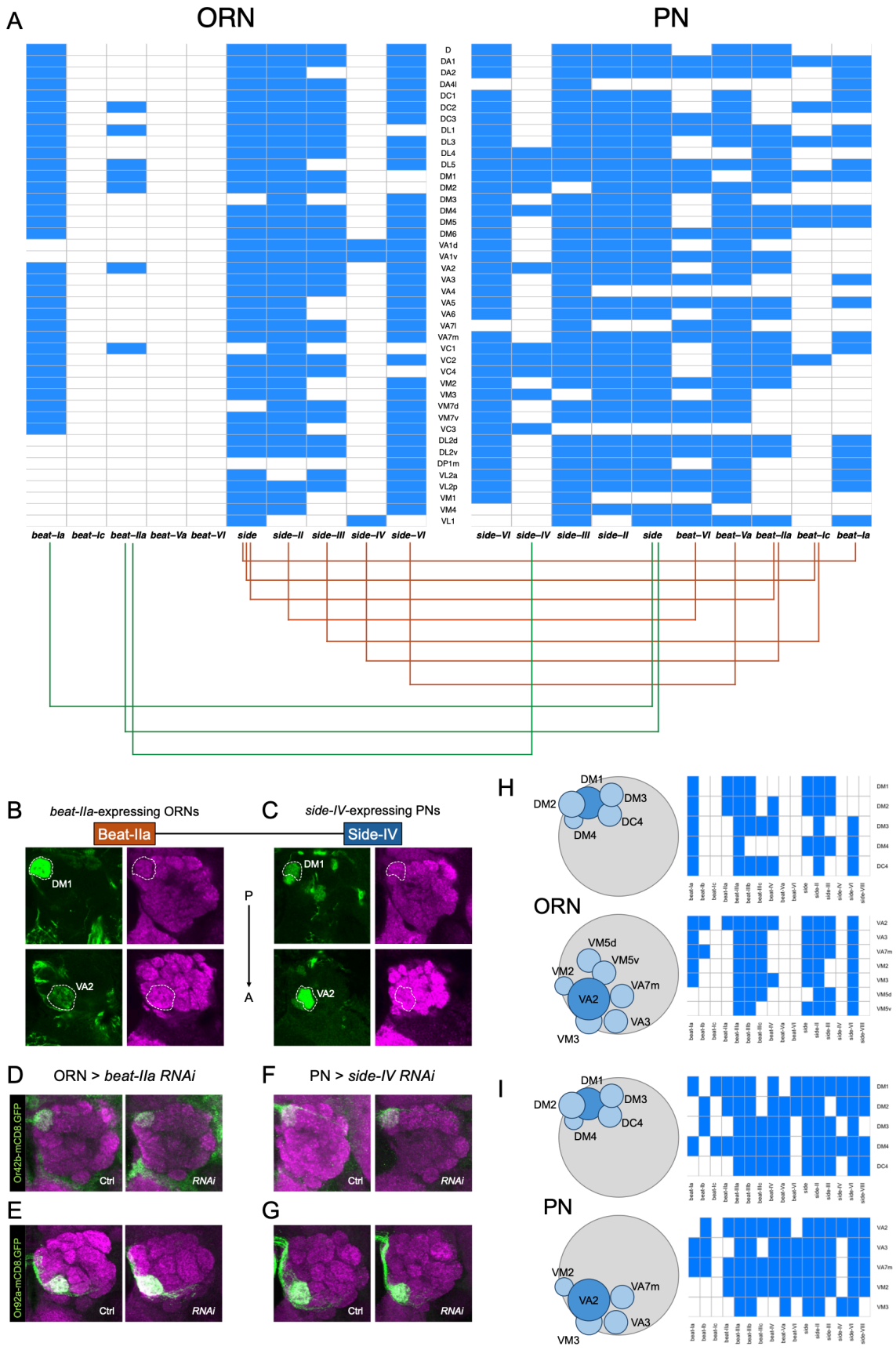

A beat/side-GAL4, UAS-FRT-STOP-FRT-mCD8.GFP, GH146-FLP
beat-Ia
beat-Ib
beat-Ic
beat-IIa
beat-IIb
beat-Va
beat-Vc
beat-IIIa
beat-IIIb
beat-IIIc
beat-IV
beat-VI
beat-VII
D
M
L
V
side
side-II
side-III
side-IV
side-V
side-VI
side-VIII
B
Lineage
adPN
lPN
PN class/ORN class
C
similarity
relation
between lineages
within lineage
GAL4 labeling data
D
Shuffled GAL4 labeling data
E
different glomeruli, different lineages
different glomeruli, same lineage
same glomerulus, same lineage
Single-cell RNA-seq data
F
Shuffled single-cell RNA-seq data
G
adPN class birth order
Divergent embryonically born PNs
Divergent larvally born PNs
Shared larvally born PNs
H
adPN class birth order
Value

Figure 3. Lineage-correlated combinatorial expression of *beats/sides* across PN classes.
(A) A glomerular map of *beat/side* expression in PNs revealed by transgenic *GAL4* driver lines. Only the GFP signal (green) highlights the glomeruli innervated by the indicated *beat/side*-positive PNs. Neuropils were stained with anti-Ncad antibody to help determine the glomerular identity but are not shown for visualization contrast. All antennal lobes shown are right antennal lobes. M, medial; L, lateral; D, dorsal; V, ventral. Each image is shown as the intensity projection over the Z-axis.
(B) Hierarchical clustering of *beat/side* expression across PNs based on the GAL4 labeling analysis. Row and column dendrograms are based on hierarchical clustering results. PN class and its matching ORN class (named by corresponding ORs expressed) are shown on the right. The left row label is color-coded based on the uniglomerular PN lineage. Blue means positive for the given gene expression, whereas white denotes negative.
(C) Similarity (Pearson's correlation) of *beat/side* expression between PN class pairs based on the expression matrix of (B) in two relation categories.
(D) Similarity (Pearson's correlation) of *beat/side* expression between PN class pairs based on the expression matrix of (B) with row labels shuffled in two relation categories.
(E) Similarity (Spearman's correlation) of *beat/side* expression between PN cell pairs based on the single-cell RNA-seq in three relation categories.
(F) Similarity (Spearman's correlation) of *beat/side* expression between PN cell pairs based on the single-cell RNA-seq with ORN class annotation shuffled in three relation categories. In (C) to (F), P values from the Mann-Whitney U test of each comparison are shown.
(G) The summary matrix is replotted from (B), but only PN classes from the adPN lineage with decoded birth order are shown. Rows representing PN classes are ordered based on the known birth sequence, while *beat/side* genes are hierarchically clustered.
(H) Heatmap showing the correlation matrix between PN classes based on the *beat/side* expression. PN classes are the same as (G) and are ordered based on the birth sequence.


A
ORN
PN
D
DA1
DA2
DA4l
DC1
DC2
DC3
DL1
DL3
DL4
DL5
DM1
DM2
DM3
DM4
DM5
DM6
VA1d
VA1v
VA2
VA3
VA4
VA5
VA6
VA7l
VA7m
VC1
VC2
VC4
VM2
VM3
VM7d
VM7v
VC3
DL2d
DL2v
DP1m
VL2a
VL2p
VM1
VM4
VL1
beat–Ia beat–Ic beat–IIa beat–Va beat–VI side side–II side–III side–IV side–VI
side–VI side–IV side–III side–II side beat–VI beat–Va beat–IIa beat–Ic beat–Ia
B
beat-IIa-expressing ORNs
Beat-IIa
C
side-IV-expressing PNs
Side-IV
DM1
VA2
P
A
D
ORN > beat-IIa RNAi
Or42b-mCD8.GFP
Ctrl
RNAi
F
PN > side-IV RNAi
E
Or92a-mCD8.GFP
G
H
DM1
DM2
DM3
DC4
DM4
ORN
VM5d
VM5v
VM2
VA7m
VA2
VA3
VM3
I
PN

688

Figure 4. Beats-Sides interactions between matching ORN-PN partners can form a redundant, error-tolerant surface "barcode", enabling recognition robustness.

(A) Potential trans-synaptic Beat/Side interactions between matching ORN-PN classes targeting each glomerulus. The ORN (left) and PN (right) profile of beat/side expression is based on Figures 2B, 3B, but replotted to highlight the ORN class and class innervating the same glomerulus. Based on *in vitro* biochemical characterization (Figure 1A), known binding interactions between Beats and Sides are indicated as lines. Green lines represent Beats -> Sides interactions, and red lines represent Sides -> Beats interactions (ORN -> PN).

(B) ORNs targeting DM1 and VA2 glomeruli expressing *beat-IIa*, labeled by the intersection of *beat-IIa-T2A-GAL4* and *ey-FLP* to drive *UAS>STOP>mCD8.GFP*.

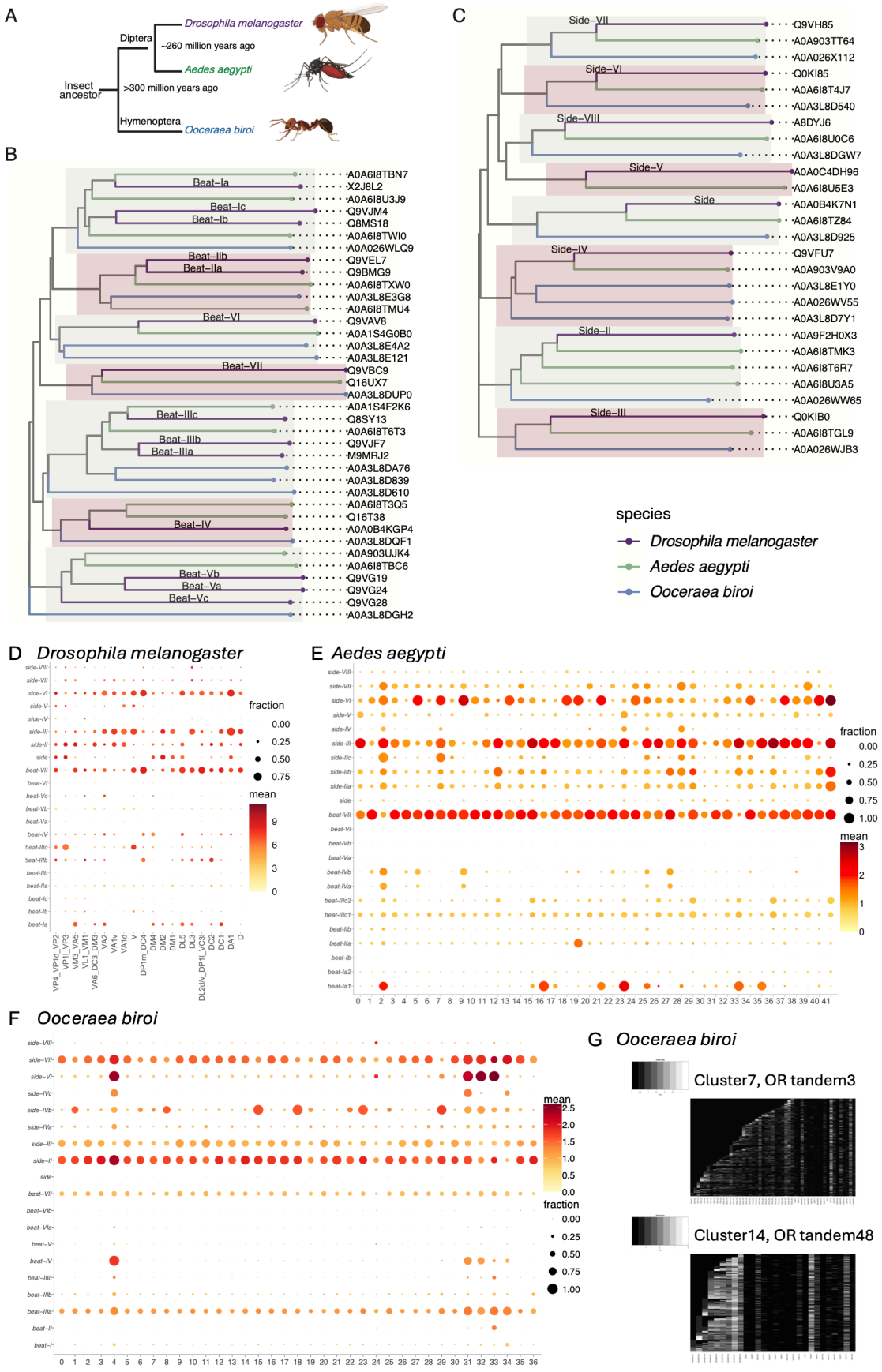
(C) PNs targeting DM1 and VA2 glomeruli expressing *side-IV*, labeled by the intersection of *side-IV-GAL4* and *GH146-FLP* to drive *UAS>STOP>mCD8.GFP*. In (B) and (C), the left columns show anti-GFP staining (green) and the right columns show anti-Ncad staining (magenta).

(D, E) Knockdown of *beat-IIa* in ORNs by a pan-ORN driver, *peb-GAL4,* while labeling Or42b ORNs targeting DM1 glomerulus (D) and Or92a ORNs targeting VA2 glomerulus (E).

(F, G) Knockdown of *side-IV* in PNs by a PN driver *GH146-GAL4* while labeling Or42b ORNs targeting DM1 glomerulus (F) and Or92a ORNs targeting VA2 glomerulus (G). Neuropils were stained with anti-Ncad antibody, shown in magenta. Representative images from 7 to 12 brains examined in each condition are shown (D-G).

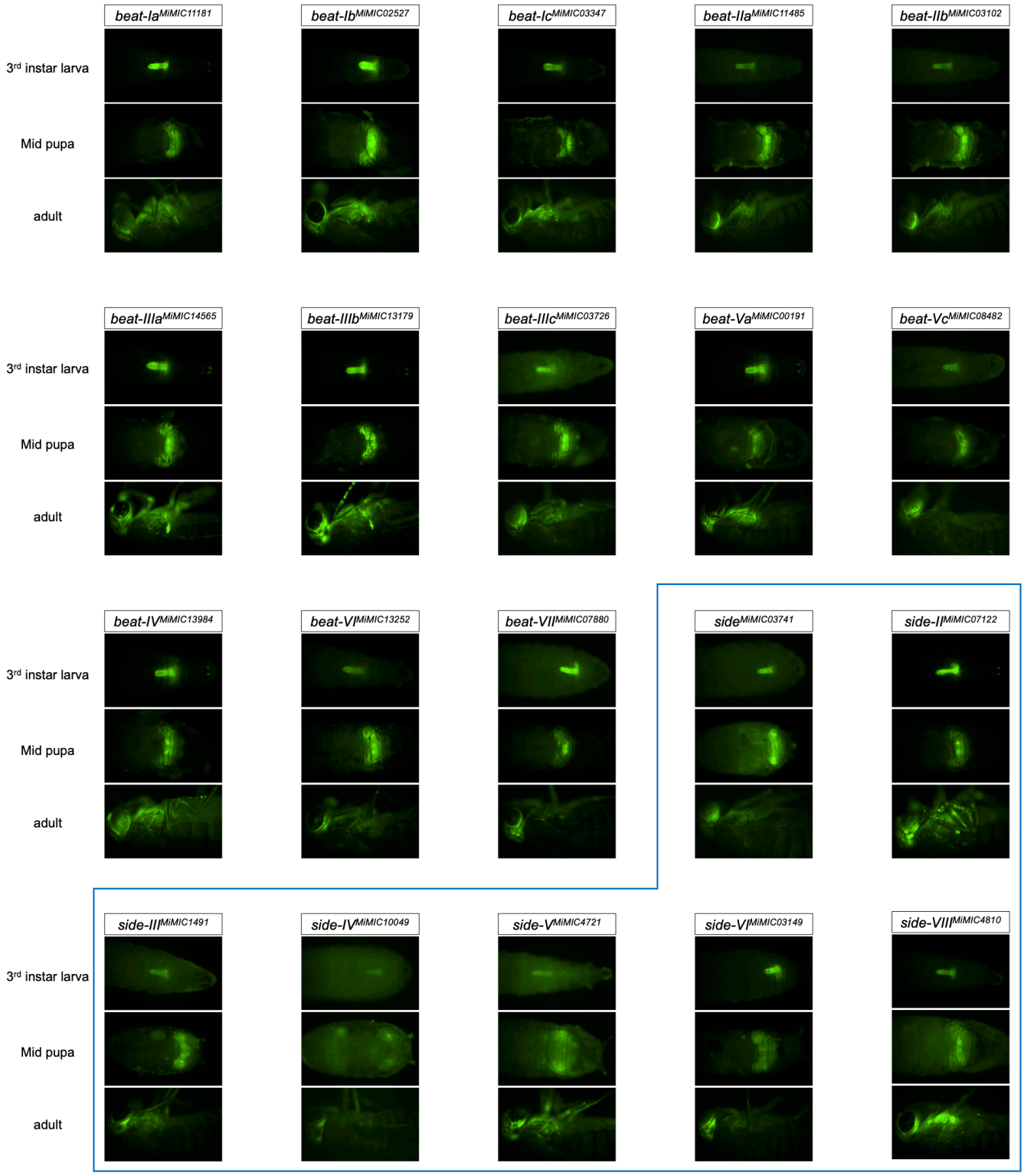
(H) The Beat/Side "barcode" in ORNs targeting DM1 or VA2 and their neighboring glomeruli, based on the *beat/side-GAL4* labeling results in Figure 2B.

(I) The Beat/Side "barcode" in PNs targeting DM1 or VA2 and their neighboring glomeruli based on the *beat/side-GAL4* labeling results in Figure 3B. VM5d/v PN expression is unknown because these glomeruli are not labeled by the PN-FLP we used. In (H) and (I), the schematics show the antennal lobe (gray circle) and relative positions of the indicated glomeruli (blue circles).


A
Drosophila melanogaster
Diptera
~260 million years ago
Aedes aegypti
Insect ancestor
>300 million years ago
Hymenoptera
Ooceraea biroi
B
Beat–Ia
Beat–Ic
Beat–Ib
Beat–IIb
Beat–IIa
Beat–VI
Beat–VII
Beat–IIIc
Beat–IIIb
Beat–IIIa
Beat–IV
Beat–Vb
Beat–Va
Beat–Vc
A0A6I8TBN7
X2J8L2
A0A6I8U3J9
Q9VJM4
Q8MS18
A0A6I8TWI0
A0A026WLQ9
Q9VEL7
Q9BMG9
A0A6I8TXW0
A0A3L8E3G8
A0A6I8TMU4
Q9VAV8
A0A1S4G0B0
A0A3L8E4A2
A0A3L8E121
Q9VBC9
Q16UX7
A0A3L8DUP0
A0A1S4F2K6
Q8SY13
A0A6I8T6T3
Q9VJF7
M9MRJ2
A0A3L8DA76
A0A3L8D839
A0A3L8D610
A0A6I8T3Q5
Q16T38
A0A0B4KGP4
A0A3L8DQF1
A0A903UJK4
A0A6I8TBC6
Q9VG19
Q9VG24
Q9VG28
A0A3L8DGH2
C
Side–VII
Side–VI
Side–VIII
Side–V
Side
Side–IV
Side–II
Side–III
Q9VH85
A0A903TT64
A0A026X112
Q0KI85
A0A6I8T4J7
A0A3L8D540
A8DYJ6
A0A6I8U0C6
A0A3L8DGW7
A0A0C4DH96
A0A6I8U5E3
A0A0B4K7N1
A0A6I8TZ84
A0A3L8D925
Q9VFU7
A0A903V9A0
A0A3L8E1Y0
A0A026WV55
A0A3L8D7Y1
A0A9F2H0X3
A0A6I8TMK3
A0A6I8T6R7
A0A6I8U3A5
A0A026WW65
Q0KIB0
A0A6I8TGL9
A0A026WJB3
species
Drosophila melanogaster
Aedes aegypti
Ooceraea biroi
D Drosophila melanogaster
fraction
0.00
0.25
0.50
0.75
mean
9
6
3
0
E Aedes aegypti
fraction
0.00
0.25
0.50
0.75
1.00
mean
3
2
1
0
F Ooceraea biroi
mean
2.5
2.0
1.5
1.0
0.5
0.0
fraction
0.00
0.25
0.50
0.75
1.00
G Ooceraea biroi
Cluster7, OR tandem3
Cluster14, OR tandem48

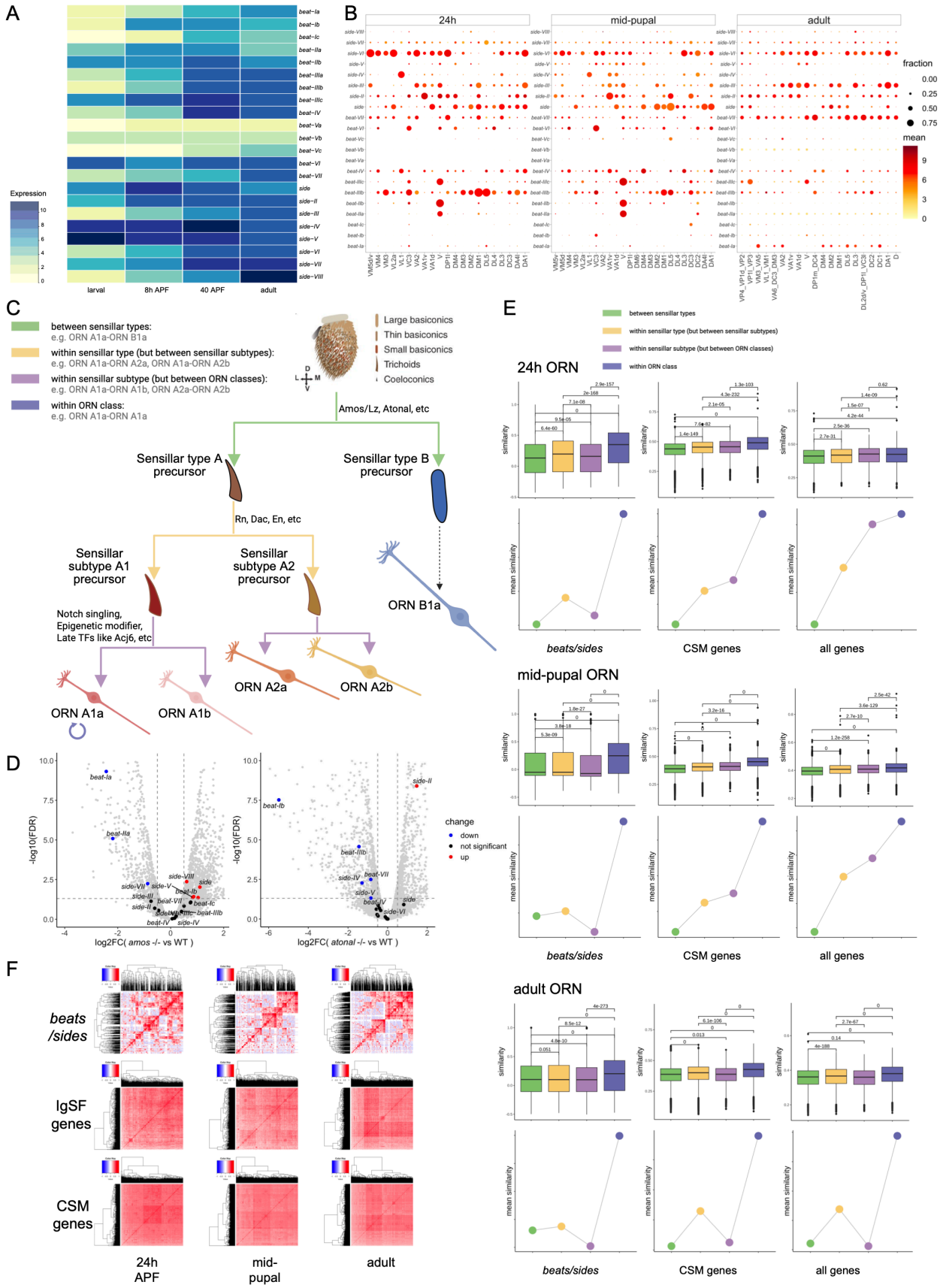
715

716 **Figure 5. Evolutionarily conserved expression pattern of *beats/sides* across ORNs in insects.**
717 **(A)** Schematic phylogeny showing the evolutionary distance among three insect species.
718 **(B)** Phylogeny tree of Beat orthologs in fruit flies *Drosophila melanogaster*, yellow fever
719 mosquitoes *Aedes aegypti*, and clonal raider ants *Ooceraea biroi*. Each leaf is color-coded by the
720 species, shown in the right bottom legend. Each orthologs is displayed as the UniProt protein ID,
721 while the known *Drosophila* Beat names are labeled on the corresponding branch.
722 **(C)** Phylogenetic tree of Side orthologs across three insect species.
723 **(D)** Expression of *beats/sides* in the adult ORN classes of fruit flies based on the single-cell RNA-
724 seq data from (McLaughlin et al., 2021). Each row is a gene, and each column represents an
725 annotated ORN class. The size of each dot represents the percentage of positive cells in the given
726 class (log2(CPM+1) > 0.5), and the color denotes the mean expression levels. This panel was
727 reproduced from Figure S2B for comparison.
728 **(E)** Expression of *beats/sides* in the adult ORN clusters of yellow mosquitoes based on the single-
729 cell RNA-seq data from (Herre et al., 2022). The mosquito *beats/sides* were renamed based on
730 their phylogenetic relations with the fruit fly orthologs.
731 **(F)** Expression of *beats/sides* in the adult ORN clusters of clonal raider ants based on the single-
732 cell RNA-seq data from (Brahma et al., 2023). The ant *beats/sides* were renamed based on their
733 phylogenetic relations with the fruit fly orthologs.
734 **(G)** Heatmap showing the expression of *beats/sides* and *OR* genes in two example ORN clusters of
735 clonal raider ants, based on the single-cell RNA-seq data from (Brahma et al., 2023). *OR* genes are
736 ordered from 5' end to 3' end in the tandem.
737


beat-Ia MiMIC11181
beat-Ib MiMIC02527
beat-Ic MiMIC03347
beat-IIa MiMIC11485
beat-IIb MiMIC03102
beat-IIIa MiMIC14565
beat-IIIb MiMIC13179
beat-IIIc MiMIC03726
beat-Va MiMIC00191
beat-Vc MiMIC08482
beat-IV MiMIC13984
beat-VI MiMIC13252
beat-VII MiMIC07880
side MiMIC03741
side-II MiMIC07122
side-III MiMIC1491
side-IV MiMIC10049
side-V MiMIC4721
side-VI MiMIC03149
side-VIII MiMIC4810
3rd instar larva
Mid pupa
adult

738

739 **Figure S1. *beat* and *side* expression are enriched in the nervous system throughout**
740 **development revealed by *MiMIC*-based gene trap drivers.**
741 Whole-animal fluorescence images of each *beat/side* gene trap GAL4 driving myr.GFP in the 3rd
742 instar larval stage, mid-pupal stage (~48h APF), and adult.


A
Expression
larval
8h APF
40 APF
adult
B
24h
mid-pupal
adult
fraction
mean
C
between sensillar types:
e.g. ORN A1a-ORN B1a
within sensillar type (but between sensillar subtypes):
e.g. ORN A1a-ORN A2a, ORN A1a-ORN A2b
within sensillar subtype (but between ORN classes):
e.g. ORN A1a-ORN A1b, ORN A2a-ORN A2b
within ORN class:
e.g. ORN A1a-ORN A1a
Large basiconics
Thin basiconics
Small basiconics
Trichoids
Coeloconics
Amos/Lz, Atonal, etc
Sensillar type A precursor
Sensillar type B precursor
Rn, Dac, En, etc
Sensillar subtype A1 precursor
Sensillar subtype A2 precursor
Notch singling, Epigenetic modifier, Late TFs like Acj6, etc
ORN B1a
ORN A1a
ORN A1b
ORN A2a
ORN A2b
D
-log10(FDR)
log2FC(amos -/- vs WT)
log2FC(atonal -/- vs WT)
change
down
not significant
up
E
between sensillar types
within sensillar type (but between sensillar subtypes)
within sensillar subtype (but between ORN classes)
within ORN class
24h ORN
mid-pupal ORN
adult ORN
similarity
mean similarity
beats/sides
CSM genes
all genes
F
beats /sides
IgSF genes
CSM genes
24h APF
mid-pupal
adult

743

Figure S2. Characterization of *beat/side* expression in the published datasets of bulk antenna and single-cell RNA-seq of ORNs.

(A) Dynamic *beats/side* expression in antennal tissues over development. Heatmap showing the temporal dynamic expression of *beats/sides* from 3rd instar larval antennal discs (larval), 8h APF pupal antennal discs, 40h APF antennae, and adult antennae. The color indicates the log-scaled relative expression values (base mean) of each gene at each stage from DESeq results.

(B) *beat/side* genes are differentially expressed across ORN classes at single-cell resolution. Each column is an annotated ORN class, and each dot represents the expression level of the row gene in the given ORN class, summarizing from the single-cell (24h APF, mid-pupal stage) or single-nucleus RNA-seq datasets (adult). The size of the dot denotes the fraction of "positive" cells in the given ORN class (log2(CPM+1) > 0.5), and the color denotes the mean expression of the positive cells.

(C) Hierarchical genetical programs control ORN lineage and glomerular targeting. A simplified model illustrating the ORN-ORN kinship. ORNs are housed in sensillar subtypes within each sensillar type, representing different genetic lineages. ORNs mapped to the same class are defined as "within ORN class", like ORN A1a and ORN A1a; ORNs of different classes but housed in the same sensillar subtype are defined as "within sensillar subtype", like ORN A1a and ORN A1b; ORNs from different sensillar subtypes but belonging to the same sensillar type, are defined as "within sensillar type", like ORN A1a and ORN A2a; and the furthest kinship is ORNs housed in distinct sensillar types, i.e., "between sensillar types", like ORN A1a and ORN B1a. Schematic was generated by BioRender. The antenna structure is reproduced from (Li et al., 2013).

(D) *beat/side* expression in the antenna is regulated by sensillar type-specification factors *amos* and *atonal*. Volcano plot showing the differentially expressed genes in *amos* mutant antennae compared with wild-type antennae (left panel) or *atonal* mutant antennae compared with wild-type antennae (right panel). Significantly downregulated *beat/side* genes are colored in blue, and significantly upregulated *beat/side* genes are colored in red. Significance is determined by FDR < 0.05 from EdgeR results. The horizontal dashed line in each plot is FDR = 0.05. Two vertical dashed lines in each plot are log2FC = -0.5 and 0.5, respectively. Gray dots are all other genes detected.

(E) Pairwise correlation between ORNs at three stages reveals the *beat/side* combinatorial expression is correlated with the ORN kinship. Top row in each stage: boxplot showing the similarity measured by Spearman's correlation between two cells from the indicated stage. The pairwise relation is categorized into four groups depending on the annotated ORN class and the corresponding sensillar type/subtype lineage. Bottom row in each stage: each dot is the mean similarity of the corresponding category in the top boxplot. The Spearman's correlation was calculated based on the expression of all beat/side genes, all CSM-encoding genes, or the whole transcriptome (all genes). P values are from Mann–Whitney U tests without multiple comparison adjustments. Figure 2E is reproduced here for comparison.

(F) *beat/side* combinatorial expression is more cell-population specific than pan-IgSF or pan-CSM genes across ORNs. Heatmap showing the Spearman's correlation of combinatorial gene expression between the row cell and the column cell across three developmental stages (24h APF, mid-pupal, and adult). The correlation was computed based on the combinatorial expression of *beat-side* genes, IgSF-encoding genes, or CSM-encoding genes. Cells are hierarchically clustered, as shown in the phylogeny tree in each heatmap. Correlation is shown as a spectrum from blue (-

1) to red (1). The diagonal is the similarity of each cell to itself, and thus the correlation always equals 1.

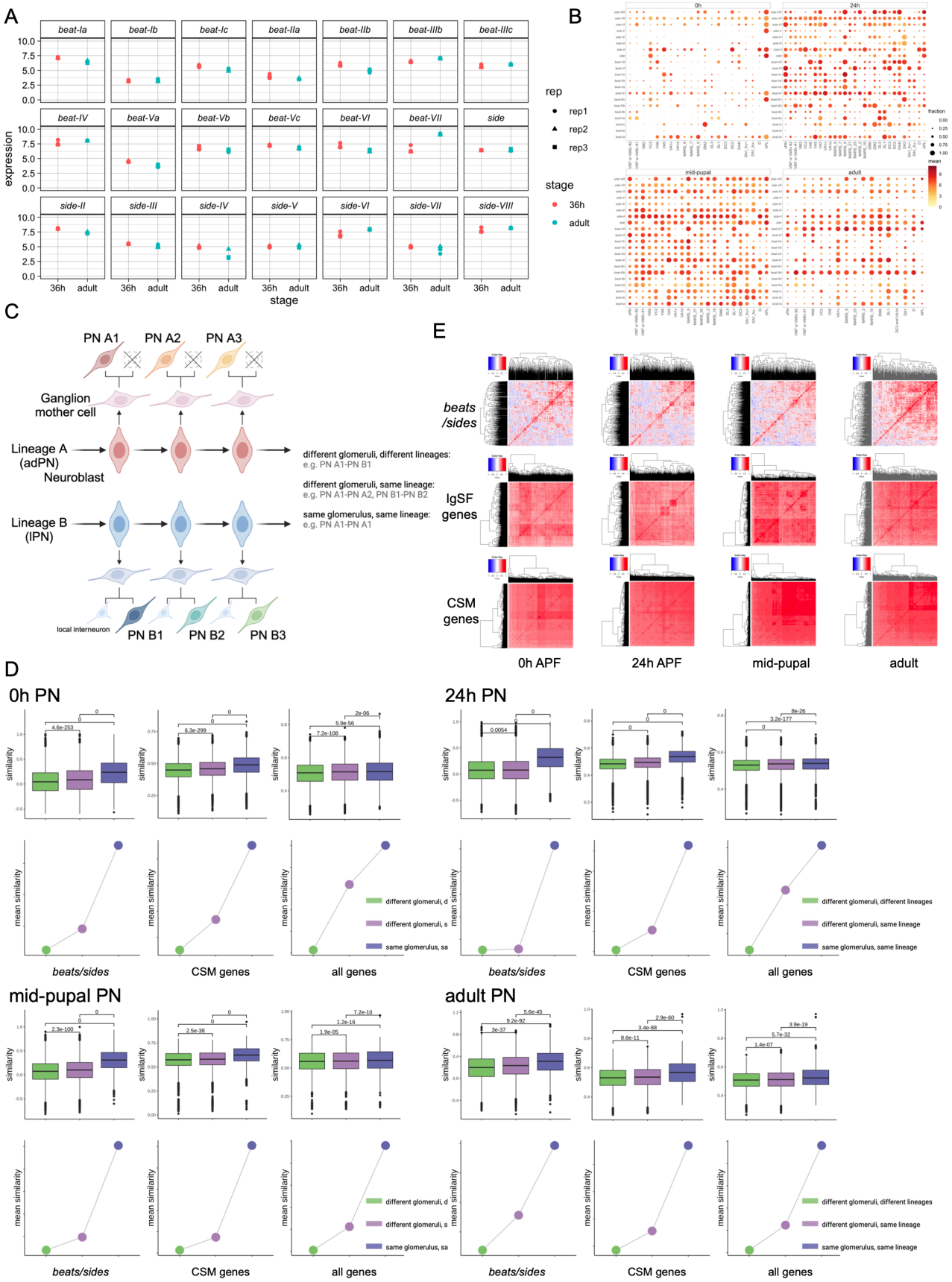

A
beat-Ia
beat-Ib
beat-Ic
beat-IIa
beat-IIb
beat-IIIb
beat-IIIc
beat-IV
beat-Va
beat-Vb
beat-Vc
beat-VI
beat-VII
side
side-II
side-III
side-IV
side-V
side-VI
side-VII
side-VIII
expression
stage
36h
adult
rep
rep1
rep2
rep3
B
0h
24h
mid-pupal
adult
fraction
mean
C
PN A1
PN A2
PN A3
Ganglion mother cell
Lineage A (adPN)
Neuroblast
Lineage B (lPN)
local interneuron
PN B1
PN B2
PN B3
different glomeruli, different lineages:
e.g. PN A1-PN B1
different glomeruli, same lineage:
e.g. PN A1-PN A2, PN B1-PN B2
same glomerulus, same lineage:
e.g. PN A1-PN A1
E
beats /sides
IgSF genes
CSM genes
0h APF
24h APF
mid-pupal
adult
D
0h PN
24h PN
mid-pupal PN
adult PN
similarity
mean similarity
beats/sides
CSM genes
all genes
different glomeruli, different lineages
different glomeruli, same lineage
same glomerulus, same lineage

Figure S3. Characterization of *beat/side* expression in the published datasets of bulk and single-cell RNA-seq of PNs.

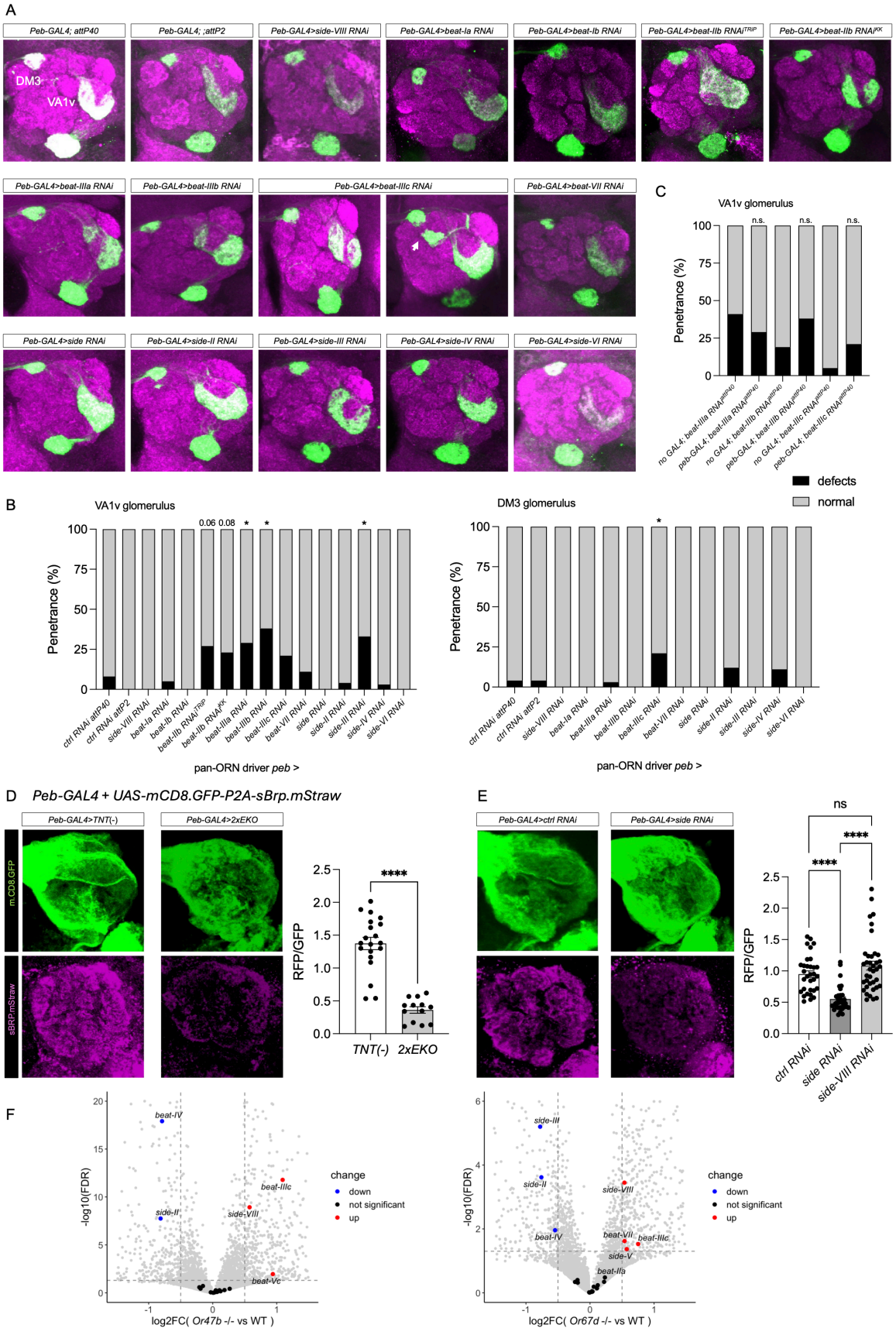
(A) *beat/side* expression in bulk PNs over development. Point plots showing the *beat/side* expression in the sorted bulk PNs from 36h APF and adult stage. The shape of each point denotes the biological replicate. The expression value is log2(CPM+1).

(B) *beat/side* genes are differentially expressed across PN classes at single-cell resolution. Each column is an annotated PN class, and each dot represents the expression level of the row gene in the given PN class, summarizing from the single-cell RNA-seq datasets. The size of the dot denotes the fraction of "positive" cells in the given PN class (log2(CPM+1) > 0.5), and the color denotes the mean expression of the positive cells. APL, anterior paired lateral neurons. APL and vPN clusters were not used for the correlation analysis in D.

(C) Sequential genetical programs control PN lineage and glomerular targeting. We only used PNs mapped to adPN and lPN lineage for correlation analysis. PNs from these two lineages are excitatory and project to a single glomerulus, whereas the third lineage, vPN neurons, which were removed for downstream analysis, are inhibitory GABAergic and can project to multiple glomeruli (Liang et al., 2013; Marin et al., 2002). In the adPN lineage, one of the two post-mitotic neurons survives and develops into a PN. In the lPN lineage, both post-mitotic neurons survive, one becoming a PN and the other becoming a local interneuron (Lin et al., 2012; Lin et al., 2010). Schematic was generated by BioRender.

(D) Pairwise correlation between PNs at four stages reveals the *beat/side* combinatorial expression is correlated with the PN kinship. Top row in each stage: boxplot showing the similarity measured by Spearman's correlation between two cells from the indicated stage. The pairwise relation is categorized into three groups depending on the annotated PN class identity and the corresponding lineage identity. Bottom row in each stage: each dot is the mean similarity of the corresponding category in the top boxplot. The Spearman's correlation was calculated based on the expression of all beat/side genes, all cell surface molecule (CSM)-encoding genes, or the whole transcriptome (all genes). P values are from Mann–Whitney U tests without multiple comparison adjustments. Figure 3E is reproduced here for comparison.

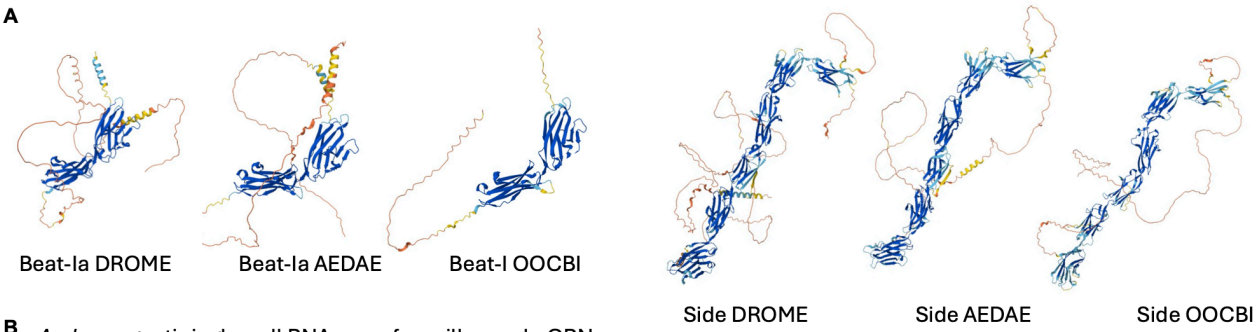
(E) *beat/side* combinatorial expression is more cell-population specific than pan-IgSF or pan-CSM genes across PNs. Heatmap showing the Spearman's correlation of combinatorial gene expression between the row cell and the column cell across four developmental stages (0h, 24h APF, mid-pupal, and adult). The correlation was computed based on the combinatorial expression of *beat-side* genes, IgSF-encoding genes, or cell surface molecule-encoding genes. Correlation is shown as a spectrum from blue (-1) to red (1). Cells are hierarchically clustered, as shown in the phylogeny tree in each heatmap. The diagonal is the similarity of each cell to itself, and thus the correlation always equals 1.


A
Peb-GAL4; attP40
Peb-GAL4; ;attP2
Peb-GAL4>side-VIII RNAi
Peb-GAL4>beat-Ia RNAi
Peb-GAL4>beat-Ib RNAi
Peb-GAL4>beat-IIb RNAi TRiP
Peb-GAL4>beat-IIb RNAi BDK
DM3
VA1v
Peb-GAL4>beat-IIIa RNAi
Peb-GAL4>beat-IIIb RNAi
Peb-GAL4>beat-IIIc RNAi
Peb-GAL4>beat-VII RNAi
Peb-GAL4>side RNAi
Peb-GAL4>side-II RNAi
Peb-GAL4>side-III RNAi
Peb-GAL4>side-IV RNAi
Peb-GAL4>side-VI RNAi
C
VA1v glomerulus
Penetrance (%)
n.s.
defects
normal
B
VA1v glomerulus
DM3 glomerulus
0.06 0.08
pan-ORN driver peb >
D Peb-GAL4 + UAS-mCD8.GFP-P2A-sBrp.mStraw
Peb-GAL4>TNT(-)
Peb-GAL4>2xEKO
mCD8.GFP
sBRP.mStraw
RFP/GFP
****
TNT(-)
2xEKO
E
Peb-GAL4>ctrl RNAi
Peb-GAL4>side RNAi
ns
ctrl RNAi
side RNAi
side-VIII RNAi
F
beat-IV
beat-IIIc
side-II
side-VIII
beat-Vc
-log10(FDR)
log2FC(Or47b -/- vs WT)
change
down
not significant
up
side-III
side-II
side-VIII
beat-IV
beat-VII
beat-IIIc
side-V
beat-IIa
log2FC(Or67d -/- vs WT)

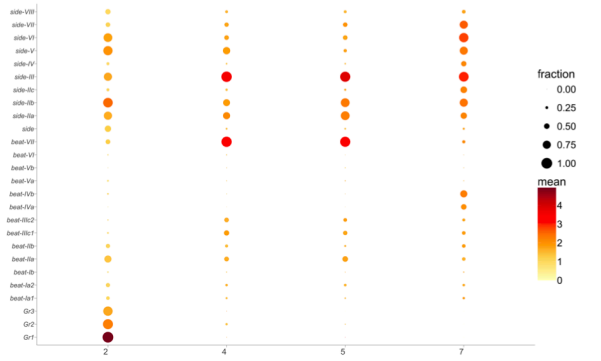
828

829 **Figure S4. Perturbation of *beats/sides* in ORNs led to no or minor defects in glomerular**
830 **organization while knockdown of *side* caused synaptic density reduction in ORNs.**
831 **(A)** Representative antennal lobes of controls and knockdowns of *beat/side* genes. Knockdown of
832 *beat-IV* or *side-V* with this *peb-GAL4* driver caused lethality, and thus the data were missing.
833 **(B)** Quantification of the percentage of lobes showing glomerular defects of VA1v and DM3
834 glomeruli in each condition. N=20-50 antennal lobes examined. P values were determined by
835 comparison to control groups. *, $p<0.05$.
836 **(C)** Quantification of the percentage of lobes showing glomerular defects of VA1v glomerulus of
837 *beat-IIIa/b/c* perturbation with additional *RNAi* transgenic background controls. N=22-34 antennal
838 lobes. n.s., not significant.
839 **(D, E)** Representative antennal lobes and quantification of the effects of blocking neuronal activity
840 of ORNs (D) or knockdown of *sides* (E) on gross glomerular synaptic density. N=14-38 antennal
841 lobes examined. ****, $p<0.0001$. n.s., not significant.
842 **(F)** Volcano plot showing the differentially expressed genes in *Or47b* mutant antennae (left) and
843 *Or67d* mutant antennae (right) compared with wild-type antennae. Significantly downregulated
844 *beat/side* genes are colored in blue, and significantly upregulated *beat/side* genes are colored in
845 red. Significance is determined by FDR<0.05 from DESeq2 results. The horizontal dashed line in
846 each plot is FDR = 0.05. Two vertical dashed lines in each plot are log2FC = -0.5 and 0.5,
847 respectively. Gray dots are all other genes detected. Data were reanalyzed from (Deanhardt et al.,
848 2023).
849

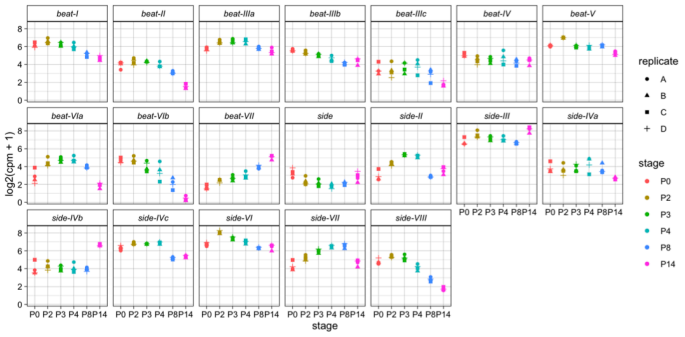
A



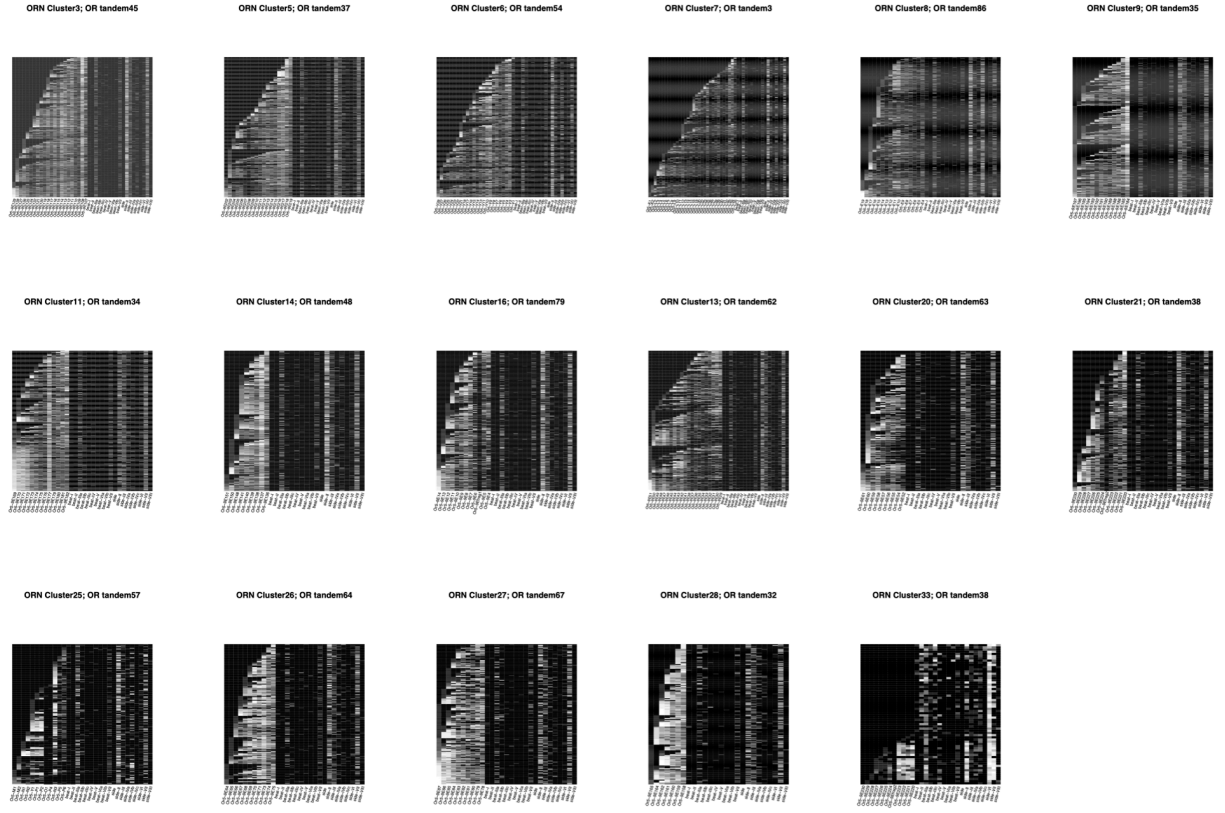
B *Aedes aegypti* single-cell RNA-seq of maxillary palp ORNs



C *Ooceraea biroi* bulk antennal RNA-seq during development



D *Ooceraea biroi* single-cell RNA-seq in each ORN cluster



850

Figure S5. Additional analyses of *beat/sides* expression in mosquitoes and ants.
(A) AlphaFold-predicted structures of example Beat and Side orthologs in fruit flies *Drosophila* melanogaster (DROME), yellow fever mosquitoes *Aedes aegypti* (AEDAE), and clonal raider ants *Ooceraea biroi* (OOCBI).
(B) Expression of *beats/sides* in the adult maxillary palp-housed ORN clusters of yellow mosquitoes based on the previously published single-cell RNA-seq data (Herre et al., 2022).
(C) Bulk RNA-seq showing the expression of *beats/sides* in developing antennal tissues of clonal raider ants, reanalyzed from the data previously published (Ryba et al., 2020).
(D) Heatmap showing the expression of *beats/sides* and *OR* genes in all ORN clusters of clonal raider ants that exhibit the ladder expression pattern of *OR* genes from 5' end to 3' end along a tandem, based on the single-cell RNA-seq data from Brahma et al., 2023. *OR* genes are ordered from 5' end to 3' end in tandem in each cluster. Two clusters shown in Figure 6G are also reproduced here for comparison. Data reanalyzed from (Brahma et al., 2023).

865 ## Materials and Methods

866 **Table 1. References and identifiers of the transgenic flies used in this study.**

Genotype	Reference	Identifier
hs-cre, vas-phiC31 (chrX)	(Diao et al., 2015)	BDSC 60299
Trojan-Gal4.0; Dr/TM3	(Diao et al., 2015)	BDSC 60301
Sp/CyO; Trojan-Gal4.0	(Diao et al., 2015)	BDSC 60302
Trojan-Gal4.1; Dr/TM3	(Diao et al., 2015)	BDSC 60304
Sp/CyO; Trojan-Gal4.1	(Diao et al., 2015)	BDSC 60305
Trojan-Gal4.2; Dr/TM3	(Diao et al., 2015)	BDSC 60307
Sp/CyO; Trojan-Gal4.2	(Diao et al., 2015)	BDSC 60308
10XUAS-IVS-myr.GFP attP40		BDSC 32198
10XUAS-IVS-myr.GFP attP2		BDSC 32197
beat-Ia-GAL4 [MI11181] chr2	This study	N.A.
beat-Ib-T2A-GAL4 [MI02527] chr2	This study	N.A.
beat-Ic-GAL4 [MI03347] chr2	This study	N.A.
beat-Ic-T2A-GAL4 [MI01467] chr2	This study	N.A.
beat-IIa-T2A-GAL4 [MI11485] chr3	(Lee et al., 2018)	BDSC 76215
beat-IIb-T2A-GAL4 [MI03102] chr3	This study	N.A.
beat-IIIa-T2A-GAL4 [MI14565] chr2	This study	N.A.
beat-IIIb-T2A-GAL4 [MI13179] chr2	(Lee et al., 2018)	BDSC 76227
beat-IIIc-T2A-GAL4 [MI03726] chr2	This study	N.A.
beat-IV-T2A-GAL4 [MI13984] chr3	This study	N.A.
beat-Va-T2A-GAL4 [MI00191] chr3	This study	N.A.
beat-Vc-T2A-GAL4 [MI08482] chr3	This study	N.A.
beat-VI-GAL4 [MI13252] chr3	This study	N.A.
beat-VII-GAL4 [MI07880] chr3	This study	N.A.
side-T2A-GAL4 [MI03741] chr3	This study	N.A.
side-II-T2A-GAL4 [MI07122] chr2	This study	N.A.
side-III-GAL4 [MI01491] chr3	This study	N.A.
side-IV-GAL4 [MI10049] chr3	This study	N.A.
side-V-GAL4 [MI04721] chr2	This study	N.A.
side-VI-T2A-GAL4 [MI03149] chr3	This study	N.A.
side-VII-GAL4 [MB10368] chr3		BDSC 29116
side-VIII-T2A-GAL4 [MI04810] chr2	This study	N.A.
ey-FLP (chrX)	(Newsome et al., 2000)	BDSC 5580
GH146-FLP (chr2)	(Hong et al., 2009)	BDSC 81291
UAS>STOP>mCD8.GFP (chr2)	(Hong et al., 2009)	BDSC 30125
UAS>STOP>mCD8.GFP (chr3)	(Hong et al., 2009)	BDSC 30032
peb-GAL4 (chrX)	(Sweeney et al., 2007)	N.A.
GH146-GAL4 (chrX)		BDSC 91812
UAS-Dcr (chrX)		BDSC 58726

Or42b-mCD8.GFP (chr2)		BDSC 52648
Or92a-mCD8.GFP (chr2)		BDSC 52646
Or47a-syt.GFP, Or47b-syt.GFP, Gr21a-syt.GFP (chr2)	Volkan lab stock	N.A.
UAS-Luc RNAi attP2		BDSC 31603
attP40		BDSC 36304
UAS-beat-Ia RNAi attP40		BDSC 64938
UAS-beat-Ib RNAi attP40		BDSC 57157
UAS-beat-IIa RNAi attP2		BDSC 28072
UAS-beat-IIb RNAi attP40		BDSC 57157
UAS-beat-IIb RNAi KK (chr2)		VDRC 104935
UAS-beat-IIIa RNAi attP40		BDSC 64526
UAS-beat-IIIb RNAi attP40		BDSC 56984
UAS-beat-IIIc RNAi attP40		BDSC 50941
UAS-beat-IV RNAi attP40		BDSC 56981
UAS-beat-VII RNAi attP40		BDSC 60056
UAS-side attP2		BDSC 50642
UAS-side-II RNAi KK (chr2)		VDRC 107512
UAS-side-III RNAi KK (chr2)		VDRC 103669
UAS-side-IV RNAi GD (chr2)		VDRC 16636
UAS-side-V RNAi attP40		BDSC 61953
UAS-side-VI RNAi KK (chr2)		VDRC 103456
UAS-side-VIII RNAi attP40		BDSC 62897
UAS-SynLight attP2	(Aimino et al., 2023)	BDSC 602367
UAS-TNT (-) (chr2)	(Sweeney et al., 1995)	BDSC 28840
2xUAS-EKO (chr2)	(White et al., 2001)	BDSC 40974

867

868

869 **Table 2. References and identifiers of the sequencing datasets used in this study.**

Dataset	Reference	Identifier
Antennal WT bulk RNA-seq of *Drosophila melanogaster*	(Barish et al., 2018)	GSE75986
PN WT bulk RNA-seq of *Drosophila melanogaster*	(Li et al., 2020b)	GSE140093
ORN single-cell RNA-seq of *Drosophila melanogaster*	(Li et al., 2020a; McLaughlin et al., 2021)	GSE162121 GSE143038
PN single-cell RNA-seq of *Drosophila melanogaster*	(Li et al., 2017a; Xie et al., 2021)	GSE161228 GSE100058
Amos mutant vs WT antennal RNA-seq of *Drosophila melanogaster*	(Mohapatra and Menuz, 2019)	PRJNA532415
Atonal mutant vs WT antennal RNA-seq of *Drosophila melanogaster*	(Menuz et al., 2014)	N.A.
Or mutant vs WT antennal RNA-seq of *Drosophila melanogaster*	(Deanhardt et al., 2023)	GSE179213
ORN single-cell RNA-seq of *Aedes aegypti*	(Herre et al., 2022)	GSE192978
ORN single-cell RNA-seq of *Ooceraea biroi*	(Brahma et al., 2023)	SRX21504666, SRX21504667

870

871

872 ***Drosophila* stocks and genetics**

873 Flies were raised in classic molasses media provided by Archon Scientific. Most crosses were kept
874 at room temperature (23 °C), except the RNAi experiments performed at 28 °C to maximize the
875 knockdown efficiency. For RNAi experiments, male and virgin female flies were mixed at room
876 temperature for three days to facilitate mating. Then flies were raised at a 28 °C incubator until 5-7
877 days after eclosion before dissection. The transgenic *Drosophila melanogaster* strains used in this
878 study are listed in Table 1. BDSC, Bloomington Drosophila Stock Center; VDRC, Vienna Drosophila
879 Resource Center.

880

881 **Generating *T2A-GAL4* transgenic flies from *MiMIC* lines**

882 We used the *in vivo* genetic cross-based method (Diao et al., 2015) to swap the *T2A-GAL4*
883 construct with the *MiMIC* cassette. Briefly, by multiple crosses, the dual-recombinase (*hs-cre* +
884 *vas-phiC31*) helper component, the in-frame *T2A-GAL4* donor component, and the target *MiMIC*
885 locus in the genes of interest were introduced together into one parental animal, where the *T2A-*
886 *GAL4* donor sequence was excised out by Cre recombinase and inserted into the *MiMIC* docking
887 site by germline-expressed phiC31 integrase. When this parental animal was crossed to the *UAS-*
888 *myr.GFP* reporter line, the recombinant progeny would have GFP expression driven by the *MiMIC-*
889 *GAL4* if *T2A-GAL4* was inserted in the correct frame and orientation. The recombinant was crossed
890 with a double-balancer line to make a stable stock. Specifically, we also found that the *T2A-GAL4*
891 construct could be used in any 5' UTR-located *MiMIC* sites in a frame-independent manner. In this
892 case, the *GAL4* coding sequence with ATG start codon in the 5' UTR would hijack the translation of
893 the native gene as an upstream open reading frame, and the *linker-T2A* sequence could be
894 neglected. As the *GAL4* open reading frame is within the whole transcript of the native gene, it is
895 still expected to faithfully represent the transcription levels and patterns of the native gene. We
896 name the driver line *gene-GAL4* if *GAL4* is within 5' UTR of the host gene (Table 1).

897

898 **Whole animal fluorescence imaging of *beat/side* expression**

899 The whole animal fluorescence imaging was performed on an Olympus BX51WI upright scope
900 equipped with a C11440-36U camera. Larvae and adults were killed by 70% ethanol first and
901 mounted on slides without coverslips. Pupae were directly mounted on slides without coverslips.
902 Larvae and pupae were dorsal side up, whereas adults were mounted with the lateral side up.
903 Images were acquired under blue light and the GFP channel, with manually adjusted exposure
904 time.

905

906 **Immunohistochemistry**

907 Flies were first killed with 70% ethanol. Then, brains were dissected in PBST buffer (0.2% Triton X-
908 100 in 1X PBS), fixed in 200 μL centrifuge tubes with 4% paraformaldehyde for 30 minutes,
909 followed by three 10-minute washes in PBST. Primary antibody mix (150 μl) was then added to the
910 tubes, incubating brains on the orbital shaker at 4 °C overnight. Brains were washed three times
911 for 20 minutes each with PBST at room temperature before secondary antibody staining.

912 Secondary antibody mix (150 μL) was then added to the tubes, incubating brains on the orbital
913 shaker at 4 °C overnight. Brains were washed three times for 20 minutes each with PBST at room
914 temperature again before being mounted on the imaging slides. Natural goat serum (1%) was
915 added to the primary and secondary antibody mix for blocking. The following primary antibodies
916 with dilution ratio were used: rabbit anti-GFP (Invitrogen, 1:1000), rat anti-Ncad (DSHB, 1:20),
917 rabbit anti-DsRed (TaKaRa Bio, 1:250); the following secondary antibodies with dilution ratio were
918 used: Alexa Fluor 488 goat anti-rabbit IgG (Invitrogen, 1:1000), Alexa Fluor 647 goat anti-rat IgG
919 (Invitrogen, 1:200), Alexa Fluor cy3 goat anti-rabbit IgG (Invitrogen, 1:200). Both primary and
920 secondary antibody cocktails were diluted in PBST.

921 Specifically, for Brp synaptic analysis, we only stained RFP mStraw and Ncad without staining GFP
922 for two reasons. First, the transgenic mStraw fluorescence of the synaptic marker is very low and
923 needs a very high laser power to detect the signal, while the natural GFP fluorescence is bright
924 enough. Second, as we later normalized the RFP intensity by the GFP intensity as the proxy for
925 synaptic density in the region of interest, no staining of GFP can reduce the variability introduced
926 during immunohistochemistry.

927

928 **Confocal Imaging**

929 To prepare brain samples for imaging, mounting solution (Flouromount-G, SouthernBiotech) was
930 added to the brains on the glass slide before mounting the cover slip. Confocal images were
931 acquired by either Olympus Fluoview FV1000 microscope or Zeiss 880 microscope with 40X or
932 60X objective lens. Brains were scanned through the Z-axis from the posterior side to the most
933 anterior side of the antennal lobes. For *beat/side-GAL4*-based glomerular innervation analysis,
934 imaging parameters varied between different *GAL4* drivers to highlight the lowly expressing
935 glomeruli. For phenotypical determination, imaging parameters were kept consistent between
936 experimental and control groups.

937

938 ***beat/side* glomerular expression pattern analysis of *beats/sides***

939 For *ey-FLP* induced labeling of each *beat/side*-expressing ORN class, we found *ey-FLP* expression
940 appears leaky in germline and crossing *ey-FLP* first with *UAS>STOP>mCD8.GFP* may lead to the
941 removal of the *STOP* sequence constitutively and thereafter fail to restrict the GAL4-driven GFP
942 expression solely in ORNs. To overcome this issue, we first crossed each *GAL4* driver line to the
943 *ey-FLP* line and then crossed to *UAS>STOP>mCD8.GFP* reporter line. For *GH146-FLP* induced
944 labeling of each *beat/side*-expressing PN class, each *GAL4* driver line can be directly crossed to
945 the *GH146-FLP*; *UAS>STOP>mCD8.GFP* line. Only the results from female animals were reported
946 in this study.

947 As we aimed to determine the ORN or PN expression pattern of *beats/sides* in a binary manner,
948 i.e., positive or negative, we didn't use the constant parameters to acquire the images for different
949 *GAL4* driver lines. We adjusted the acquisition parameters in order to clearly show even lowly
950 expressing glomeruli. After scanning each brain from the most posterior end to the most anterior
951 end of the antennal lobe, we referred to the previously characterized glomerular map (Task et al.,

952 2022) and manually determined the glomerular identity and the corresponding *beat/side*
953 expression, assigning 0 to negative expression and 1 to positive expression. At least three brains
954 were examined for each *GAL4* driver line. Notably, for the *GH146-FLP*-based labeling, we
955 occasionally found inconsistent glomerular innervation patterns between individuals and even
956 right and left lobes of the same brain (data not shown). This is likely due to the incomplete excision
957 of the *STOP* cassette in a subset of PNs. As each glomerulus could be innervated by as few as one
958 PN, it is possible that some glomeruli are not labeled in an antennal lobe. We thus determined
959 these glomeruli to be positive if we observed the GFP signal for that glomerulus in two or more
960 antennal lobes. The expression matrix, expression value (0,1) of each gene in each glomerulus,
961 was then input to the downstream clustering analysis.

962

963 **Phenotype quantification**

964 The measurement of the glomerular phenotype was determined by the proportion of antennal
965 lobes showing morphological or positional abnormalities across all studied brains within each
966 group, compared to the controls. The P-value was determined using the two-tailed Fisher's exact
967 test, utilizing the integrated features of the GraphPad Prism 9 software.

968 To quantify synaptic density at the gross antennal lobe level, we first obtained the Z-stack of
969 confocal sections by maximum projection. Then, we manually selected the regions of interest
970 (ROI) according to the Ncad channel, which are the right and left antennal lobes. We also selected
971 a control region in the dorsal part of the brain where the driver *peb-GAL4* has no expression. We
972 then subtracted the average intensity of the RFP channel of the control region from that of the ROI.
973 This is the adjusted mean RFP intensity of the ROI. We did the same thing to get the adjusted
974 mean GFP intensity of the ROI. We finally calculated the ratio between the adjusted mean
975 intensity of RFP and the adjusted mean intensity of GFP as the readout for each antennal lobe.
976 Each group has 30-40 antennal lobes. An unpaired t-test was used to compare the average
977 RFP/GFP ratio between groups.

978

979 **Bulk tissue RNA-seq and single-cell RNA-seq analysis from previously published datasets for**
980 ***Drosophila melanogaster***

981 The publicly available datasets used are listed in Table 2. For bulk tissue RNA-seq datasets, the
982 author-processed datasets (expression matrix and differentially expressed gene analysis by either
983 EdgeR or DESeq) were directly used for customized analyses and visualization in this study. For
984 single-cell RNA-seq datasets, the author-annotated expression matrices were directly used for
985 similarity analysis (see details below).

986

987 **Pairwise similarity analysis for gene combinatorial expression in ORNs and PNs**

988 The ORN single-cell RNA-seq dataset was first filtered to keep cells annotated to a single ORN
989 class. Then, for cells in each stage, we extracted their *beat/side* expression profile, i.e., the vector
990 of which each element is the expression value of each *beat/side* gene. We further filtered out the
991 cells with zero expression for all *beat/side* genes. The remaining cells were used to calculate the

992 expression correlation for *beat/side* genes and other gene combinations, e.g., IgSF-encoding
993 genes or CSM-encoding genes. We first calculated the Spearman's correlation matrix for *beat/side*
994 expression across the selected cells in each stage (Lv et al., 2022). Then, we assigned the kinship
995 for each cell-cell pair based on the annotated ORN/glomerular identity, sensillum where they are
996 housed, and the sensillar type (Couto et al., 2005; Fishilevich and Vosshall, 2005; Grabe et al.,
997 2016; Marin et al., 2020; McLaughlin et al., 2021; Silbering et al., 2011; Task et al., 2022).
998 Specifically, if two cells are annotated as the same ORN class, this pair is defined as "within ORN
999 class"; if two cells are annotated as different ORN classes but both in the same sensillar subtype,
000 this pair is defined as "within sensillar subtype"; if two cells are annotated as different ORN
001 classes from different sensillar subtypes, but belong to the common sensillar type, this pair is
002 defined as "within sensillar type"; if two cells are annotated as different ORN classes from
003 different sensillar types, the pair is defined as "between sensillar types".

004 The PN single-cell dataset was first filtered to keep cells annotated to a single PN class (excluding
005 a small portion of anterior paired lateral (APL) neurons) and corresponding lineage identity. We
006 only kept cells annotated to either adPN or lPN lineage (removed multiglomerular vPNs). Then, we
007 extracted their *beat/side* expression vector for cells in each stage. We further filtered out the cells
008 with zero expression for all *beat/side* genes. The remaining cells were used to analyze different
009 gene combinations at the same stage. We first computed the Spearman's correlation matrix of
010 *beat/side* expression across the selected cells in each stage. Then, we assigned the kinship for
011 each cell-cell pair based on the annotated PN/glomerular identity and lineage identity.
012 Specifically, if two cells are annotated as the same PN class, this pair is defined as "same
013 glomerulus, same lineage"; if two cells are annotated as different PN classes but in the same PN
014 lineage, this pair is defined as "different glomeruli, same lineage"; if two cells are annotated as
015 different PN classes from different PN lineages, this pair is defined as "different glomeruli,
016 different lineages". Cell-cell pairs belonging to "same glomerulus, different lineages" were scarce
017 and removed before statistical tests.

018 We finally plotted the distribution of the correlation coefficient of each kinship category and
019 compared the mean between the two categories using the Mann–Whitney U test. This analysis
020 was expanded to the correlation of IgSF and CSM genes, randomly selected genes, or the whole
021 transcriptome (all genes) at each stage. As the numbers of pairs (n) in each kinship category vary
022 dramatically, we also randomly selected the same number of pairs in each kinship category. Then,
023 we performed statistical tests on the resampled data and found the results to be true. Therefore,
024 all related figures in this paper were plotted based on the original data without resampling.

025 Shuffled control panels were calculated under the same pipeline as the abovementioned, except
026 the annotations of each cell, ORN/PN class name, and corresponding lineage information were
027 shuffled. We shuffled 1000 times with different random seeds and plotted the average mean of
028 each shuffling run.

029 A similar analysis was also applied to the *beat/side* genetic labeling results, where the binary
030 expression matrices (with only 0 and 1) were used as the input. In this case, as each glomerulus
031 represents the whole ORN/PN classes, we would not have the "within ORN class" or "same
032 glomerulus, same lineage" category.

033

Ortholog identification and phylogenetic analysis of *beats* and *sides* in yellow fever mosquitoes and clonal raider ants

To identify the orthologs of *Drosophila melanogaster beat/side* genes in *Aedes aegypti* (yellow fever mosquitoes) and *Ooceraea biroi* (clonal raider ants), we used two ways to cross-validate the candidate list. For the biased way, we first queried the *Drosophila* genes one by one to get the putative orthologs in the other two species on www.orthodb.org (Tegenfeldt et al., 2025). Then, we queried the mosquito and ant genes to get the fruit fly orthologs. This reciprocal approach gave rise to a primary list of candidate *beats/sides* in mosquitoes and ants. For an unbiased way, we used the blastp tool on NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) to find the orthologous sequences encoded by the mosquito and ant genomes of *Drosophila* Beat-Ia and Side proteins. We applied additional filters, including protein size and AlphaFold-predicted structures (https://alphafold.ebi.ac.uk). We only kept candidates with comparable amino acid lengths to *Drosophila* homologs and similar protein structures (two Ig domains for Beats; five Ig domains and one fibronectin domain for Sides). This gave rise to a longer list of candidates. We compared these two lists and found that the first is contained in the second one. We then ran multiple sequence alignments of all Beat candidate and Side candidate proteins, together with the fruit fly ones, to build the phylogenetic trees, respectively, through the Clustal Omega program (Madeira et al., 2024) using the default parameters on https://www.uniprot.org.

Single-cell RNA-seq analysis of *beat/side* ortholog expression in the ORNs of yellow fever mosquitoes and clonal raider ants

Author-annotated single-cell RNA-seq datasets of *Aedes aegypti* and *Ooceraea biro* from two prior studies (Table 2) were directly used for downstream analyses and visualization. We extracted the mosquito and ant *beat/sides* genes from the original expression matrices. Based on the phylogenetic reconstruction described earlier, we renamed them to reflect their phylogenetic relationships with analogous genes in fruit flies.

Statistics and code availability

R 4.3.1 and the necessary packages were used for all the abovementioned analyses and visualization. Statistical tests were performed with the built-in functions of R or through GraphPad Prism 9 software. The customized codes to replicate all illustrations and statistical graphs of this study will be uploaded to https://github.com/volkanlab.

Author contributions

QD and PCV conceptualized the project and designed the experiments. QD did experiments, analyzed data, and prepared figures. Y-CDC, SO, QD, & LQR generated reagents. SO, RE, CY, and KMV helped with experiments. QD and PCV wrote and edited the manuscript.

Acknowledgement

The authors would also like to thank the Bloomington Drosophila Stock Center and the Vienna Drosophila Resource Center for providing the fruit fly strains and the Duke Light Microscopy Core Facility for technical support for imaging. The authors would also like to thank Drs. Claude Desplan and Richard S. Mann for sharing reagents, and members of the Volkan lab for discussions and comments on the manuscript. The authors particularly thank Allison Carson and Nick Manfred for assistance with the experiments. The authors also thank Drs. Hiroaki Matsunami, Corbin Jones, David McClay, and Amy Bejsovec for stimulating suggestions.

Funding

This study was supported by the U.S. National Science Foundation award 2006471 to PCV, National Institutes of Health NRSA F32EY032750 and NEI K99EY035757 to Y-CDC, and NINDS award NS070644 to RSM.

References

Aimino, M.A., J. Humenik, M.J. Parisi, J.C. Duhart, and T.J. Mosca. 2023. SynLight: a bicistronic strategy for simultaneous active zone and cell labeling in the Drosophila nervous system. *G3 (Bethesda)*. 13.

Arguello, J.R., L. Abuin, J. Armida, K. Mika, P.C. Chai, and R. Benton. 2021. Targeted molecular profiling of rare olfactory sensory neurons identifies fate, wiring, and functional determinants. *Elife*. 10.

Ashley, J., V. Sorrentino, M. Lobb-Rabe, S. Nagarkar-Jaiswal, L. Tan, S. Xu, Q. Xiao, K. Zinn, and R.A. Carrillo. 2019. Transsynaptic interactions between IgSF proteins DIP-alpha and Dpr10 are required for motor neuron targeting specificity. *Elife*. 8.

Barish, S., S. Nuss, I. Strunilin, S. Bao, S. Mukherjee, C.D. Jones, and P.C. Volkan. 2018. Combinations of DIPs and Dprs control organization of olfactory receptor neuron terminals in Drosophila. *PLoS Genet*. 14:e1007560.

Barish, S., and P.C. Volkan. 2015. Mechanisms of olfactory receptor neuron specification in Drosophila. *Wiley Interdiscip Rev Dev Biol*. 4:609-621.

Blagburn, J.M. 2008. Engrailed expression in subsets of adult Drosophila sensory neurons: an enhancer-trap study. *Invert Neurosci*. 8:133-146.

Brahma, A., D.D. Frank, P.D.H. Pastor, P.K. Piekarski, W. Wang, J.D. Luo, T.S. Carroll, and D.J.C. Kronauer. 2023. Transcriptional and post-transcriptional control of odorant receptor choice in ants. *Curr Biol*. 33:5456-5466 e5455.

Brochtrup, A., and T. Hummel. 2011. Olfactory map formation in the Drosophila brain: genetic specificity and neuronal variability. *Curr Opin Neurobiol*. 21:85-92.

Brovero, S.G., J.C. Fortier, H. Hu, P.C. Lovejoy, N.R. Newell, C.M. Palmateer, R.Y. Tzeng, P.T. Lee, K. Zinn, and M.N. Arbeitman. 2021. Investigation of Drosophila fruitless neurons that express Dpr/DIP cell adhesion molecules. *Elife*. 10.

111 Carrier, Y., L. Quintana Rio, N. Formicola, V. de Sousa-Xavier, M. Tabet, Y.D. Chen, A.H. Ali, M.
112 Wislez, L. Orts, A. Borst, and F. Pinto-Teixeira. 2025. Biased cell adhesion organizes the
113 Drosophila visual motion integration circuit. *Dev Cell*. 60:762-779 e767.
114 Carrillo, R.A., E. Ozkan, K.P. Menon, S. Nagarkar-Jaiswal, P.T. Lee, M. Jeon, M.E. Birnbaum, H.J.
115 Bellen, K.C. Garcia, and K. Zinn. 2015. Control of Synaptic Connectivity by a Network of
116 Drosophila IgSF Cell Surface Proteins. *Cell*. 163:1770-1782.
117 Chai, P.C., S. Cruchet, L. Wigger, and R. Benton. 2019. Sensory neuron lineage mapping and
118 manipulation in the Drosophila olfactory system. *Nat Commun*. 10:643.
119 Cheng, S., J. Ashley, J.D. Kurleto, M. Lobb-Rabe, Y.J. Park, R.A. Carrillo, and E. Ozkan. 2019a.
120 Molecular basis of synaptic specificity by immunoglobulin superfamily receptors in
121 Drosophila. *Elife*. 8.
122 Cheng, S., Y. Park, J.D. Kurleto, M. Jeon, K. Zinn, J.W. Thornton, and E. Ozkan. 2019b. Family of
123 neural wiring receptors in bilaterians defined by phylogenetic, biochemical, and structural
124 evidence. *Proc Natl Acad Sci U S A*. 116:9837-9842.
125 Cosmanescu, F., P.S. Katsamba, A.P. Sergeeva, G. Ahlsen, S.D. Patel, J.J. Brewer, L. Tan, S. Xu, Q.
126 Xiao, S. Nagarkar-Jaiswal, A. Nern, H.J. Bellen, S.L. Zipursky, B. Honig, and L. Shapiro. 2018.
127 Neuron-Subtype-Specific Expression, Interaction Affinities, and Specificity Determinants
128 of DIP/Dpr Cell Recognition Proteins. *Neuron*. 100:1385-1400 e1386.
129 Courgeon, M., and C. Desplan. 2019. Coordination between stochastic and deterministic
130 specification in the Drosophila visual system. *Science*. 366.
131 Couto, A., M. Alenius, and B.J. Dickson. 2005. Molecular, anatomical, and functional organization
132 of the Drosophila olfactory system. *Curr Biol*. 15:1535-1547.
133 de Jong, S., J.A. Cavallo, C.D. Rios, H.A. Dworak, and H. Sink. 2005. Target recognition and
134 synaptogenesis by motor axons: responses to the sidestep protein. *Int J Dev Neurosci*.
135 23:397-410.
136 Deanhardt, B., Q. Duan, C. Du, C. Soeder, A. Morlote, D. Garg, A. Saha, C.D. Jones, and P.C.
137 Volkan. 2023. Social experience and pheromone receptor activity reprogram gene
138 expression in sensory neurons. *G3 (Bethesda)*. 13.
139 Diao, F., H. Ironfield, H. Luan, F. Diao, W.C. Shropshire, J. Ewer, E. Marr, C.J. Potter, M. Landgraf,
140 and B.H. White. 2015. Plug-and-play genetic access to drosophila cell types using
141 exchangeable exon cassettes. *Cell Rep*. 10:1410-1421.
142 Dombrovski, M., Y. Zang, G. Frighetto, A. Vaccari, H. Jang, P.S. Mirshahidi, F. Xie, P. Sanfilippo, B.W.
143 Hina, A. Rehan, R.H. Hussein, P.S. Mirshahidi, C. Lee, A. Morris, M.A. Frye, C.R. von Reyn,
144 Y.Z. Kurmangaliyev, G.M. Card, and S.L. Zipursky. 2025. Gradients of Cell Recognition
145 Molecules Wire Visuomotor Transformation. *bioRxiv*.
146 Duan, Q., R. Estrella, A. Carson, Y. Chen, and P.C. Volkan. 2023. The effect of Drosophila attP40
147 background on the glomerular organization of Or47b olfactory receptor neurons. *G3*
148 *(Bethesda)*.
149 Endo, K., T. Aoki, Y. Yoda, K. Kimura, and C. Hama. 2007. Notch signal organizes the Drosophila
150 olfactory circuitry by diversifying the sensory neuronal lineages. *Nat Neurosci*. 10:153-160.
151 Fambrough, D., and C.S. Goodman. 1996. The Drosophila beaten path gene encodes a novel
152 secreted protein that regulates defasciculation at motor axon choice points. *Cell*. 87:1049-
153 1058.

154 Fishilevich, E., and L.B. Vosshall. 2005. Genetic and functional subdivision of the Drosophila
155 antennal lobe. *Curr Biol*. 15:1548-1553.
156 Goulding, S.E., P. zur Lage, and A.P. Jarman. 2000. amos, a proneural gene for Drosophila olfactory
157 sense organs that is regulated by lozenge. *Neuron*. 25:69-78.
158 Goyal, G., A. Zierau, M. Lattemann, B. Bergkirchner, D. Javorski, R. Kaur, and T. Hummel. 2019.
159 Inter-axonal recognition organizes Drosophila olfactory map formation. *Sci Rep*. 9:11554.
160 Grabe, V., A. Baschwitz, H.K.M. Dweck, S. Lavista-Llanos, B.S. Hansson, and S. Sachse. 2016.
161 Elucidating the Neuronal Architecture of Olfactory Glomeruli in the Drosophila Antennal
162 Lobe. *Cell Rep*. 16:3401-3413.
163 Gupta, B.P., G.V. Flores, U. Banerjee, and V. Rodrigues. 1998. Patterning an epidermal field:
164 Drosophila lozenge, a member of the AML-1/Runt family of transcription factors, specifies
165 olfactory sense organ type in a dose-dependent manner. *Dev Biol*. 203:400-411.
166 Gupta, B.P., and V. Rodrigues. 1997. Atonal is a proneural gene for a subset of olfactory sense
167 organs in Drosophila. *Genes Cells*. 2:225-233.
168 Hattori, D., E. Demir, H.W. Kim, E. Viragh, S.L. Zipursky, and B.J. Dickson. 2007. Dscam diversity is
169 essential for neuronal wiring and self-recognition. *Nature*. 449:223-227.
170 Herre, M., O.V. Goldman, T.C. Lu, G. Caballero-Vidal, Y. Qi, Z.N. Gilbert, Z. Gong, T. Morita, S.
171 Rahiel, M. Ghaninia, R. Ignell, B.J. Matthews, H. Li, L.B. Vosshall, and M.A. Younger. 2022.
172 Non-canonical odor coding in the mosquito. *Cell*. 185:3104-3123 e3128.
173 Heymann, C., C. Paul, N. Huang, J.C. Kinold, A.C. Dietrich, and H. Aberle. 2022. Molecular insights
174 into the axon guidance molecules Sidestep and Beaten path. *Frontiers in Physiology*. 13.
175 Hong, W., and L. Luo. 2014. Genetic control of wiring specificity in the fly olfactory system.
176 *Genetics*. 196:17-29.
177 Hong, W., T.J. Mosca, and L. Luo. 2012. Teneurins instruct synaptic partner matching in an
178 olfactory map. *Nature*. 484:201-207.
179 Hong, W., H. Zhu, C.J. Potter, G. Barsh, M. Kurusu, K. Zinn, and L. Luo. 2009. Leucine-rich repeat
180 transmembrane proteins instruct discrete dendrite targeting in an olfactory map. *Nat*
181 *Neurosci*. 12:1542-1550.
182 Honig, B., and L. Shapiro. 2020. Adhesion Protein Structure, Molecular Affinities, and Principles of
183 Cell-Cell Recognition. *Cell*. 181:520-535.
184 Hummel, T., M.L. Vasconcelos, J.C. Clemens, Y. Fishilevich, L.B. Vosshall, and S.L. Zipursky. 2003.
185 Axonal targeting of olfactory receptor neurons in Drosophila is controlled by Dscam.
186 *Neuron*. 37:221-231.
187 Hummel, T., and S.L. Zipursky. 2004. Afferent induction of olfactory glomeruli requires N-cadherin.
188 *Neuron*. 42:77-88.
189 Imai, T., H. Sakano, and L.B. Vosshall. 2010. Topographic mapping--the olfactory system. *Cold*
190 *Spring Harb Perspect Biol*. 2:a001776.
191 Jefferis, G.S., and T. Hummel. 2006. Wiring specificity in the olfactory system. *Semin Cell Dev Biol*.
192 17:50-65.
193 Jefferis, G.S., E.C. Marin, R.F. Stocker, and L. Luo. 2001. Target neuron prespecification in the
194 olfactory map of Drosophila. *Nature*. 414:204-208.
195 Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R.
196 Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B.
197 Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M.

198 Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals,
199 A.W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis. 2021. Highly accurate protein
200 structure prediction with AlphaFold. *Nature*. 596:583-589.
201 Kinold, J.C., M. Brenner, and H. Aberle. 2021. Misregulation of Drosophila Sidestep Leads to
202 Uncontrolled Wiring of the Adult Neuromuscular System and Severe Locomotion Defects.
203 *Front Neural Circuits*. 15:658791.
204 Kinold, J.C., C. Pfarr, and H. Aberle. 2018. Sidestep-induced neuromuscular miswiring causes
205 severe locomotion defects in Drosophila larvae. *Development*. 145.
206 Komiyama, T., J.R. Carlson, and L. Luo. 2004. Olfactory receptor neuron axon targeting: intrinsic
207 transcriptional control and hierarchical interactions. *Nat Neurosci*. 7:819-825.
208 Komiyama, T., W.A. Johnson, L. Luo, and G.S. Jefferis. 2003. From lineage to wiring specificity. POU
209 domain transcription factors control precise connections of Drosophila olfactory
210 projection neurons. *Cell*. 112:157-167.
211 Lee, P.T., J. Zirin, O. Kanca, W.W. Lin, K.L. Schulze, D. Li-Kroeger, R. Tao, C. Devereaux, Y. Hu, V.
212 Chung, Y. Fang, Y. He, H. Pan, M. Ge, Z. Zuo, B.E. Housden, S.E. Mohr, S. Yamamoto, R.W.
213 Levis, A.C. Spradling, N. Perrimon, and H.J. Bellen. 2018. A gene-specific T2A-GAL4 library
214 for Drosophila. *Elife*. 7.
215 Li, H., F. Horns, B. Wu, Q. Xie, J. Li, T. Li, D.J. Luginbuhl, S.R. Quake, and L. Luo. 2017a. Classifying
216 Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell*.
217 171:1206-1220 e1222.
218 Li, H., T. Li, F. Horns, J. Li, Q. Xie, C. Xu, B. Wu, J.M. Kebschull, C.N. McLaughlin, S.S. Kolluru, R.C.
219 Jones, D. Vacek, A. Xie, D.J. Luginbuhl, S.R. Quake, and L. Luo. 2020a. Single-Cell
220 Transcriptomes Reveal Diverse Regulatory Strategies for Olfactory Receptor Expression
221 and Axon Targeting. *Curr Biol*. 30:1189-1198 e1185.
222 Li, H., S.A. Shuster, J. Li, and L. Luo. 2018. Linking neuronal lineage and wiring specificity. *Neural*
223 *Dev*. 13:5.
224 Li, H., A. Watson, A. Olechwier, M. Anaya, S.K. Sorooshyari, D.P. Harnett, H.P. Lee, J. Vielmetter,
225 M.A. Fares, K.C. Garcia, E. Ozkan, J.P. Labrador, and K. Zinn. 2017b. Deconstruction of the
226 beaten Path-Sidestep interaction network provides insights into neuromuscular system
227 development. *Elife*. 6.
228 Li, J., S. Han, H. Li, N.D. Udeshi, T. Svinkina, D.R. Mani, C. Xu, R. Guajardo, Q. Xie, T. Li, D.J.
229 Luginbuhl, B. Wu, C.N. McLaughlin, A. Xie, P. Kaewsapsak, S.R. Quake, S.A. Carr, A.Y. Ting,
230 and L. Luo. 2020b. Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring
231 Regulators. *Cell*. 180:373-386 e315.
232 Li, Q., S. Barish, S. Okuwa, A. Maciejewski, A.T. Brandt, D. Reinhold, C.D. Jones, and P.C. Volkan.
233 2016. A Functionally Conserved Gene Regulatory Network Module Governing Olfactory
234 Neuron Diversity. *PLoS Genet*. 12:e1005780.
235 Li, Q., T.S. Ha, S. Okuwa, Y. Wang, Q. Wang, S.S. Millard, D.P. Smith, and P.C. Volkan. 2013.
236 Combinatorial rules of precursor specification underlying olfactory neuron diversity. *Curr*
237 *Biol*. 23:2481-2490.
238 Liang, L., Y. Li, C.J. Potter, O. Yizhar, K. Deisseroth, R.W. Tsien, and L. Luo. 2013. GABAergic
239 projection neurons route selective olfactory inputs to specific higher-order neurons.
240 *Neuron*. 79:917-931.

241 Lin, S., C.F. Kao, H.H. Yu, Y. Huang, and T. Lee. 2012. Lineage analysis of Drosophila lateral
242 antennal lobe neurons reveals notch-dependent binary temporal fate decisions. *PLoS Biol*.
243 10:e1001425.
244 Lin, S., S.L. Lai, H.H. Yu, T. Chihara, L. Luo, and T. Lee. 2010. Lineage-specific effects of
245 Notch/Numb signaling in post-embryonic development of the Drosophila brain.
246 *Development*. 137:43-51.
247 Lobb-Rabe, M., W.I. Nawrocka, R. Zhang, J. Ashley, R.A. Carrillo, and E. Ozkan. 2024. Neuronal
248 Wiring Receptors Dprs and DIPs Are GPI Anchored and This Modification Contributes to
249 Their Cell Surface Organization. *eNeuro*. 11.
250 Lv, X., S. Li, J. Li, X.Y. Yu, X. Ge, B. Li, S. Hu, Y. Lin, S. Zhang, J. Yang, X. Zhang, J. Yan, A.L. Joyner, H.
251 Shi, Q. Wu, and S.H. Shi. 2022. Patterned cPCDH expression regulates the fine organization
252 of the neocortex. *Nature*. 612:503-511.
253 Madeira, F., N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A.R.N. Tivey, R. Lopez, and S.
254 Butcher. 2024. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024.
255 *Nucleic Acids Res*. 52:W521-W525.
256 Marin, E.C., L. Buld, M. Theiss, T. Sarkissian, R.J.V. Roberts, R. Turnbull, I.F.M. Tamimi, M.W.
257 Pleijzier, W.J. Laursen, N. Drummond, P. Schlegel, A.S. Bates, F. Li, M. Landgraf, M. Costa,
258 D.D. Bock, P.A. Garrity, and G. Jefferis. 2020. Connectomics Analysis Reveals First-,
259 Second-, and Third-Order Thermosensory and Hygrosensory Neurons in the Adult
260 Drosophila Brain. *Curr Biol*. 30:3167-3182 e3164.
261 Marin, E.C., G.S. Jefferis, T. Komiyama, H. Zhu, and L. Luo. 2002. Representation of the glomerular
262 olfactory map in the Drosophila brain. *Cell*. 109:243-255.
263 Marin, E.C., R.J. Watts, N.K. Tanaka, K. Ito, and L. Luo. 2005. Developmentally programmed
264 remodeling of the Drosophila olfactory circuit. *Development*. 132:725-737.
265 McLaughlin, C.N., M. Brbic, Q. Xie, T. Li, F. Horns, S.S. Kolluru, J.M. Kebschull, D. Vacek, A. Xie, J. Li,
266 R.C. Jones, J. Leskovec, S.R. Quake, L. Luo, and H. Li. 2021. Single-cell transcriptomes of
267 developing and adult olfactory receptor neurons in Drosophila. *Elife*. 10.
268 Menon, K.P., V. Kulkarni, S.Y. Takemura, M. Anaya, and K. Zinn. 2019. Interactions between Dpr11
269 and DIP-gamma control selection of amacrine neurons in Drosophila color vision circuits.
270 *Elife*. 8.
271 Menuz, K., N.K. Larter, J. Park, and J.R. Carlson. 2014. An RNA-seq screen of the Drosophila
272 antenna identifies a transporter necessary for ammonia detection. *PLoS Genet*.
273 10:e1004810.
274 Mohapatra, P., and K. Menuz. 2019. Molecular Profiling of the Drosophila Antenna Reveals
275 Conserved Genes Underlying Olfaction in Insects. *G3 (Bethesda)*. 9:3753-3771.
276 Morano, N.C., D.H. Lopez, H. Meltzer, A.P. Sergeeva, P.S. Katsamba, K.D. Rostam, H.P. Gupta, J.E.
277 Becker, B. Bornstein, F. Cosmanescu, O. Schuldiner, B. Honig, R.S. Mann, and L. Shapiro.
278 2025. Members of the DIP and Dpr adhesion protein families use cis inhibition to shape
279 neural development in Drosophila. *PLoS Biol*. 23:e3003030.
280 Mosca, T.J., and L. Luo. 2014. Synaptic organization of the Drosophila antennal lobe and its
281 regulation by the Teneurins. *Elife*. 3:e03726.
282 Newsome, T.P., B. Asling, and B.J. Dickson. 2000. Analysis of Drosophila photoreceptor axon
283 guidance in eye-specific mosaics. *Development*. 127:851-860.

Osaka, J., A. Ishii, X. Wang, R. Iwanaga, H. Kawamura, S. Akino, A. Sugie, S. Hakeda-Suzuki, and T. Suzuki. 2024. Complex formation of immunoglobulin superfamily molecules Side-IV and Beat-IIb regulates synaptic specificity. *Cell Rep*. 43:113798.

Özkan, E., R.A. Carrillo, C.L. Eastman, R. Weiszmann, D. Waghray, K.G. Johnson, K. Zinn, S.E. Celniker, and K.C. Garcia. 2013. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell*. 154:228-239.

Pipes, G.C., Q. Lin, S.E. Riley, and C.S. Goodman. 2001. The Beat generation: a multigene family encoding IgSF proteins related to the Beat axon guidance molecule in Drosophila. *Development*. 128:4545-4552.

Rodrigues, V., and T. Hummel. 2008. Development of the Drosophila olfactory system. *Adv Exp Med Biol*. 628:82-101.

Royet, J., and R. Finkelstein. 1997. Establishing primordia in the Drosophila eye-antennal imaginal disc: the roles of decapentaplegic, wingless and hedgehog. *Development*. 124:4793-4800.

Ryba, A.R., S.K. McKenzie, L. Olivos-Cisneros, E.J. Clowney, P.M. Pires, and D.J.C. Kronauer. 2020. Comparative Development of the Ant Chemosensory System. *Curr Biol*. 30:3223-3230 e3224.

Sanes, J.R., and S.L. Zipursky. 2020. Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. *Cell*. 181:1434-1435.

Siebert, M., D. Banovic, B. Goellner, and H. Aberle. 2009. Drosophila motor axons recognize and follow a Sidestep-labeled substrate pathway to reach their target fields. *Genes Dev*. 23:1052-1062.

Silbering, A.F., R. Rytz, Y. Grosjean, L. Abuin, P. Ramdya, G.S. Jefferis, and R. Benton. 2011. Complementary function and integrated wiring of the evolutionarily distinct Drosophila olfactory subsystems. *J Neurosci*. 31:13357-13375.

Sink, H., E.J. Rehm, L. Richstone, Y.M. Bulls, and C.S. Goodman. 2001. sidestep encodes a target-derived attractant essential for motor axon guidance in Drosophila. *Cell*. 105:57-67.

Song, E., B. de Bivort, C. Dan, and S. Kunes. 2012. Determinants of the Drosophila odorant receptor pattern. *Dev Cell*. 22:363-376.

Sweeney, L.B., A. Couto, Y.H. Chou, D. Berdnik, B.J. Dickson, L. Luo, and T. Komiyama. 2007. Temporal target restriction of olfactory receptor neurons by Semaphorin-1a/PlexinA-mediated axon-axon interactions. *Neuron*. 53:185-200.

Sweeney, S.T., K. Broadie, J. Keane, H. Niemann, and C.J. O'Kane. 1995. Targeted expression of tetanus toxin light chain in Drosophila specifically eliminates synaptic transmission and causes behavioral defects. *Neuron*. 14:341-351.

Tan, L., K.X. Zhang, M.Y. Pecot, S. Nagarkar-Jaiswal, P.T. Lee, S.Y. Takemura, J.M. McEwen, A. Nern, S. Xu, W. Tadros, Z. Chen, K. Zinn, H.J. Bellen, M. Morey, and S.L. Zipursky. 2015. Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. *Cell*. 163:1756-1769.

Task, D., C.C. Lin, A. Vulpe, A. Afify, S. Ballou, M. Brbic, P. Schlegel, J. Raji, G. Jefferis, H. Li, K. Menuz, and C.J. Potter. 2022. Chemoreceptor co-expression in Drosophila melanogaster olfactory neurons. *Elife*. 11.

Tegenfeldt, F., D. Kuznetsov, M. Manni, M. Berkeley, E.M. Zdobnov, and E.V. Kriventseva. 2025. OrthoDB and BUSCO update: annotation of orthologs with wider sampling of genomes. *Nucleic Acids Res*. 53:D516-D522.

328 Tichy, A.L., A. Ray, and J.R. Carlson. 2008. A new Drosophila POU gene, pdm3, acts in odor
329 receptor expression and axon targeting of olfactory neurons. *J Neurosci*. 28:7121-7129.
330 Venkatasubramanian, L., Z. Guo, S. Xu, L. Tan, Q. Xiao, S. Nagarkar-Jaiswal, and R.S. Mann. 2019.
331 Stereotyped terminal axon branching of leg motor neurons mediated by IgSF proteins DIP-
332 alpha and Dpr10. *Elife*. 8.
333 Venken, K.J., K.L. Schulze, N.A. Haelterman, H. Pan, Y. He, M. Evans-Holm, J.W. Carlson, R.W.
334 Levis, A.C. Spradling, R.A. Hoskins, and H.J. Bellen. 2011. MiMIC: a highly versatile
335 transposon insertion resource for engineering Drosophila melanogaster genes. *Nat*
336 *Methods*. 8:737-743.
337 Vien, K.M., Q. Duan, C. Yeung, and P.C. Volkan. 2023. Conserved atypical cadherin, Fat2, regulates
338 axon terminal organization in the developing Drosophila olfactory sensory neurons.
339 *bioRxiv*:2023.2007. 2028.550990.
340 Wang, Y., M. Lobb-Rabe, J. Ashley, P. Chatterjee, V. Anand, H.J. Bellen, O. Kanca, and R.A. Carrillo.
341 2022. Systematic expression profiling of Dpr and DIP genes reveals cell surface codes in
342 Drosophila larval motor and sensory neurons. *Development*. 149.
343 Ward, A., W. Hong, V. Favaloro, and L. Luo. 2015. Toll receptors instruct axon and dendrite
344 targeting and participate in synaptic partner matching in a Drosophila olfactory circuit.
345 *Neuron*. 85:1013-1028.
346 White, B.H., T.P. Osterwalder, K.S. Yoon, W.J. Joiner, M.D. Whim, L.K. Kaczmarek, and H.
347 Keshishian. 2001. Targeted attenuation of electrical activity in Drosophila using a
348 genetically modified K(+) channel. *Neuron*. 31:699-711.
349 Xie, Q., M. Brbic, F. Horns, S.S. Kolluru, R.C. Jones, J. Li, A.R. Reddy, A. Xie, S. Kohani, Z. Li, C.N.
350 McLaughlin, T. Li, C. Xu, D. Vacek, D.J. Luginbuhl, J. Leskovec, S.R. Quake, L. Luo, and H. Li.
351 2021. Temporal evolution of single-cell transcriptomes of Drosophila olfactory projection
352 neurons. *Elife*. 10.
353 Xie, Q., J. Li, H. Li, N.D. Udeshi, T. Svinkina, D. Orlin, S. Kohani, R. Guajardo, D.R. Mani, C. Xu, T. Li,
354 S. Han, W. Wei, S.A. Shuster, D.J. Luginbuhl, S.R. Quake, S.E. Murthy, A.Y. Ting, S.A. Carr,
355 and L. Luo. 2022. Transcription factor Acj6 controls dendrite targeting via a combinatorial
356 cell-surface code. *Neuron*. 110:2299-2314 e2298.
357 Xie, Q., B. Wu, J. Li, C. Xu, H. Li, D.J. Luginbuhl, X. Wang, A. Ward, and L. Luo. 2019. Transsynaptic
358 Fish-lips signaling prevents misconnections between nonsynaptic partner olfactory
359 neurons. *Proc Natl Acad Sci U S A*. 116:16068-16073.
360 Xu, C., E. Theisen, R. Maloney, J. Peng, I. Santiago, C. Yapp, Z. Werkhoven, E. Rumbaut, B. Shum,
361 D. Tarnogorska, J. Borycz, L. Tan, M. Courgeon, T. Griffin, R. Levin, I.A. Meinertzhagen, B. de
362 Bivort, J. Drugowitsch, and M.Y. Pecot. 2019. Control of Synaptic Specificity by Establishing
363 a Relative Preference for Synaptic Partners. *Neuron*. 103:865-877 e867.
364 Xu, S., A.P. Sergeeva, P.S. Katsamba, S. Mannepalli, F. Bahna, J. Bimela, S.L. Zipursky, L. Shapiro,
365 B. Honig, and K. Zinn. 2022. Affinity requirements for control of synaptic targeting and
366 neuronal cell survival by heterophilic IgSF cell adhesion molecules. *Cell Rep*. 39:110618.
367 Xu, S., Q. Xiao, F. Cosmanescu, A.P. Sergeeva, J. Yoo, Y. Lin, P.S. Katsamba, G. Ahlsen, J. Kaufman,
368 N.T. Linaval, P.T. Lee, H.J. Bellen, L. Shapiro, B. Honig, L. Tan, and S.L. Zipursky. 2018.
369 Interactions between the Ig-Superfamily Proteins DIP-alpha and Dpr6/10 Regulate
370 Assembly of Neural Circuits. *Neuron*. 100:1369-1384 e1366.

Yoo, J., M. Dombrovski, P. Mirshahidi, A. Nern, S.A. LoCascio, S.L. Zipursky, and Y.Z. Kurmangaliyev. 2023. Brain wiring determinants uncovered by integrating connectomes and transcriptomes. *Curr Biol*. 33:3998-4005 e3996.

Yu, H.H., C.F. Kao, Y. He, P. Ding, J.C. Kao, and T. Lee. 2010. A complete developmental sequence of a Drosophila neuronal lineage as revealed by twin-spot MARCM. *PLoS Biol*. 8.

Zhu, H., T. Hummel, J.C. Clemens, D. Berdnik, S.L. Zipursky, and L. Luo. 2006. Dendritic patterning by Dscam and synaptic partner matching in the Drosophila antennal lobe. *Nat Neurosci*. 9:349-355.

Zhu, H., and L. Luo. 2004. Diverse functions of N-cadherin in dendritic and axonal terminal arborization of olfactory projection neurons. *Neuron*. 42:63-75.

zur Lage, P.I., D.R. Prentice, E.E. Holohan, and A.P. Jarman. 2003. The Drosophila proneural gene amos promotes olfactory sensillum formation and suppresses bristle formation. *Development*. 130:4683-4693.