Targeted molecular profiling of rare cell populations identifies olfactory sensory neuron fate and wiring determinants

J. Roman Arguello\textsuperscript{1,2,3,#}, Liliane Abuin\textsuperscript{1,#}, Jan Armida\textsuperscript{1,&}, Kaan Mika\textsuperscript{1,&}, Phing Chian Chai\textsuperscript{1} and Richard Benton\textsuperscript{*}

\textsuperscript{1}Center for Integrative Genomics
Faculty of Biology and Medicine
University of Lausanne
CH-1015
Lausanne
Switzerland

\textsuperscript{2}Department of Ecology and Evolution
Faculty of Biology and Medicine
University of Lausanne
CH-1015
Lausanne
Switzerland

\textsuperscript{3}Swiss Institute of Bioinformatics
CH-1015 Lausanne
Switzerland

\textsuperscript{*}Corresponding author:

T: ++41 21 692 3932
E: Richard.Benton@unil.ch

\textsuperscript{#} These authors contributed equally to this work.
\textsuperscript{&} These authors contributed equally to this work.
Abstract

Determining the molecular properties of neurons is essential to understand their development, function and evolution. We used Targeted DamID (TaDa) to characterize RNA polymerase II occupancy and chromatin accessibility in selected Ionotropic Receptor (IR)-expressing sensory neurons in the Drosophila antenna. Although individual populations represent a minute fraction of cells, TaDa is sufficiently sensitive and specific to identify the expected receptor genes. Unique Ir expression is not linked to substantial differences in chromatin accessibility, but rather to distinct transcription factor profiles. Heterogeneously-expressed genes across populations are enriched for neurodevelopmental factors, and we identify functions for the POU-domain protein Pdm3 as a genetic switch of Ir neuron fate, and the atypical cadherin Flamingo in segregation of neurons into discrete glomeruli. Together this study reveals the effectiveness of TaDa in profiling rare neural populations, identifies new roles for a transcription factor and a neuronal guidance molecule, and provides valuable datasets for future exploration.

Introduction

Nervous systems are composed of vast numbers of neuron classes with specific structural and functional characteristics. Determining the molecular basis of these properties is essential to understand their emergence during development and contribution to neural circuit activity, as well as appreciating how neurons exhibit plasticity over both short and evolutionary timescales.

The olfactory system of adult Drosophila melanogaster is a powerful model to identify and link molecular features of neurons to their anatomy and physiology. The Drosophila nose (antenna) contains ~50 classes of olfactory sensory neurons (OSNs), which develop from sensory organ precursors cells specified in the larval antennal imaginal disc. Each OSN is characterized by the expression of a unique olfactory receptor (or, occasionally, receptors), of the Odorant Receptor (OR) or Ionotropic Receptor (IR) families. These proteins function – together with broadly-expressed, family-specific co-receptors – to define neuronal odor-response properties. Moreover, the axons of all neurons expressing the same receptor(s) converge to a discrete and stereotyped glomerulus within the primary olfactory center (antennal lobe) in the brain, where they synapse with both local interneurons (LNs) and projection neurons (PNs). While the unique properties of individual OSN classes are assumed to reflect distinct patterns of gene expression, it is unclear how many molecular differences exist between sensory neuron populations, which types of genes (beyond receptors) distinguish OSN classes, and whether gene expression differences reflect population-specific chromatin architecture.

Exploration of the molecular properties of specific neuron types has been revolutionized by single-cell RNA sequencing (scRNA-seq) technologies. However, the peripheral olfactory system of Drosophila poses several challenges to single-cell isolation: OSNs are very small (~2-3 µm soma diameter) and up to four neurons are tightly embedded under sensory sensilla (a porous cuticular hair that houses the neurons’ dendrites) along with several support cells that seal each neuronal cluster from neighboring sensilla. Furthermore, each OSN class represents only a very small fraction of cells of the whole antenna, as few
as 5-10 neurons of a total of ~3000 neurons and non-neuronal cells. Recent work – discussed below – successfully profiled RNA levels in OSNs at a mid-developmental time-point (42-48 h after puparium formation (APF)), before these cells become encased in the mature cuticle.

To characterize the molecular profile of specific populations of OSNs we sought to use the Targeted-DNA adenine methyltransferase identification (TaDa) method. TaDa entails expression of an RNA polymerase II subunit (Rp215, hereafter PolII) fused to the E. coli DNA adenine methyltransferase (Dam). This complex (or, as control, Dam alone) is expressed within a specific cell population (“targeted”) under the regulation of a cell type-specific promoter using the Gal4/UAS system. When bound to DNA, Dam:PolII methylates adenine within GATC motifs; enrichment of methylation at genes compared to free Dam control samples provides a measure of RNA PolII occupancy, which is correlated with genes’ transcriptional activity (Figure 1A). In addition to indicating RNA PolII occupancy within a given cell type, TaDa datasets can also provide information on chromatin accessibility (CATaDa) (Figure 1A). Analysis of the methylation patterns across the genome generated by untethered Dam is assumed to reflect the openness of chromatin, analogous to – and in good agreement with – ATAC-seq datasets.

The selectivity of expression of Dam:PolII or Dam avoids the requirement for cell isolation prior to genomic DNA extraction. Previous applications of TaDa in Drosophila have profiled relatively abundant cell populations in central brain tissue, the intestine or the male germline. Although OSN classes can be effectively labeled using transgenes containing promoters for the corresponding receptor genes, it was unclear whether TaDa would be sufficiently sensitive to measure transcriptional activity from an individual population of sensory neurons within the antenna. We therefore set out to profile several classes of IR-expressing OSNs to test the approach and identify novel factors that distinguish the properties of these evolutionarily closely-related neuronal subtypes.

Results

Targeted DamID of OSN populations

To test the feasibility of TaDa for profiling Drosophila OSNs, we first focused on the Ir64a population, which comprises ~16 neurons that are housed in sensilla in the sacculus (a chamber within the antenna), where they detect acidic odors.

We first verified that Ir64a promoter-Gal4 (hereafter, Ir64a-Gal4) induction of Dam or Dam:PolII did not impact the expression or localization of the IR64a receptor (Figure 1B) or the projection of these neurons to the antennal lobe (Figure 1C). We proceeded to collect triplicate samples of ~2000 adult antennae and processed these for TaDa (see Methods). This experimental design is expected to integrate PolII occupancy patterns from mid-pupae (when Ir-Gal4 expression initiates) to mature OSNs. As a positive control, we first examined the mapping of reads of methylated DNA fragments to the Ir64a gene. Dam:PolII consistently accessed this gene region at higher levels than Dam alone, particularly near the 5’ end of the locus (Figure 1D), consistent with previously-described genome-wide PolII occupancy patterns.

We extended the TaDa approach using drivers to target six additional individual OSN classes, including all acid-sensing Ir populations (Ir31a, Ir75a,
Ir75b, Ir75c, Ir84a)\textsuperscript{21,22}, as well as the hygroscopic Ir40a neurons\textsuperscript{23,24}; these populations range from \textasciitilde8 to \textasciitilde25 cells per antenna\textsuperscript{25}. As expected, the enrichment of Dam:PolII access to the Ir64a gene was specific to the Ir64a neurons (Figure 1E). We quantified “PolII occupancy scores” (OSs) by calculating the log\textsubscript{2} ratio of the (normalized) number of reads from the Dam:PolII samples compared to the (normalized) number of reads in control Dam samples across the entire gene body. In Ir64a neurons, the Ir64a gene OS was significantly positive (OS = 0.20; FDR = 4.26e-10), while it was not significantly different from zero in the other six neuron populations (Figure 1E, bottom).

Similarly, OSs for other Ir genes were significantly positive in the corresponding neuronal population, but not in those in which the receptors are not expressed (Figure 1F). We noted two exceptions: first, in Ir31a neurons, Ir40a had a positive OS (0.12; FDR = 1.93e-13), although it was much lower than that of Ir31a (OS = 0.41; FDR = 2.67e-6) (Figure 1F). Second, the Ir75a, Ir75b and Ir75c genes had positive OSs in all three of the corresponding neuron populations (Figure 1F). This latter result is most likely due to these tandemly-clustered genes having overlapping transcription units\textsuperscript{25}.

To further examine the specificity of TaDa in reporting on neuron population-specific gene expression, we plotted OSs for all members of the Ir repertoire, as well as Or and Gustatory receptor (Gr) families. The vast majority of these genes are expressed only in other antennal sensory neuron populations, distinct chemosensory organs and/or life stages\textsuperscript{26-28}. Concordantly, only very few genes show significant OSs in any of the seven populations of Ir neurons analyzed by TaDa (Figure 1F and Figure S1). Two prominent examples are Ir8a and Ir25a, which encode co-receptors for subsets of tuning IRs\textsuperscript{29}. Ir8a has a significant OS in all Ir populations except for Ir40a neurons, which is consistent with the selective expression and function of IR8a in acid-sensing OSNs but not hygroscopic neurons\textsuperscript{29} (Figure 1F). Ir25a is broadly-expressed in most antennal neurons\textsuperscript{29,30} – although it may have a role in only a subset of these – which is reflected in a significant OS for Ir25a in all seven populations (Figure 1F). Ir40a is co-expressed with another receptor, IR93a\textsuperscript{23,24}, but we did not detect a significant OS for Ir93a in these neurons (Figure S1 and Table S1). This negative result may reflect the observation that while IR93a protein is readily detected\textsuperscript{23}, Ir93a transcripts levels (and so potentially PolII occupancy) are extremely low\textsuperscript{30,31}.

Of the other chemosensory receptors displaying unexpected significant OSs in these Ir neurons, many of these genes are located in the introns of other genes (e.g., Ir47a, within the slowpoke 2 locus, which encodes a nervous-system expressed potassium channel\textsuperscript{32}) or have overlapping transcripts with other genes (e.g., Gr2a, which overlaps with futsch, a gene encoding a neuronal microtubule-binding protein\textsuperscript{33}) (Figure S1). In these cases, calculation of a specific OS for the chemosensory genes is impossible, and we suspect that in many cases the OS reflects transcription of the non-receptor gene.

Olfactory receptors have lower transcript levels in the antenna compared to other classes of genes with broader neuronal or non-neuronal expression patterns\textsuperscript{31,34}. We extended the examination of the TaDa datasets to other genes that are known to be expressed (or not expressed) in the analyzed Ir neuron populations (Figure 1G). Two broadly-expressed neuronal genes, elav and nSyb, display significant OSs across populations, as does pebbled (peb), a molecular marker for OSNs\textsuperscript{35}. By contrast, oaz, a broad marker of PNs\textsuperscript{36}, is not significantly occupied. Many of the most highly expressed genes in the antenna are those
encoding Odorant Binding Proteins (OBPs), which are transcribed in non-
neuronal cells, often at >10-fold higher levels than olfactory receptor genes, as
determined by bulk RNA-seq$^{31,34}$. Two of these, *Obp59a* and *Obp84a* are present
in sensilla housing Ir neurons$^{37,38}$, but neither of these genes display significant
OSs in any neuron population (Figure 1G). Similarly, markers for other support
cell types, such as *nompA* (thecogen cell)$^{37,39}$, and the glial marker *repo*
are not significantly occupied (Figure 1G).

Electrophysiological experiments indicate that antennal OSNs are
cholinergic$^{40}$. In line with this functional property, genes encoding the vesicular
acetylcholine transporter (*VACHT*) and choline acetyltransferase (*ChAT*) have
significant OSs in all populations, while glutamatergic and GABAergic neuron
markers (vesicular glutamate transporter (*VGLUT*) and glutamic acid decarboxylase
1 (*GAD1*)) do not (Figure 1G). We note, however, that a GABA receptor subunit
(encoded by *GABA-B-R2*) – which is heterogeneously-expressed and mediates
population-specific presynaptic gain control in different Or OSNs$^{41}$ – displays
significant, but variable, OSs in different Ir populations (Figure 1G). Finally, of the
multiple neuropeptides detected in the antennal lobe, only one, short
neuropeptide F (sNPF), originates from OSNs$^{42}$, where it has a role in glomerular-
specific presynaptic facilitation$^{43}$. Consistently, we observe significant and
heterogeneous OSs for sNPF, but not other neuropeptide precursor genes that
are expressed in central neuron populations (e.g., Tachykinin (*Tk*), Myoinhibiting
peptide (*MIP*))$^{42}$.

Together, these results indicate that TaDa is sufficiently sensitive and
specific to detect RNA PolII occupancy at transcriptionally-active genes in very
small populations of neurons embedded within the cellurally-complex antennal
tissue.

**Global analysis of TaDa datasets**

We next surveyed OSs at a genome-wide level, to gain insights into patterns of
transcriptional activity in different Ir neuron populations. The number of genes
significantly occupied by PolII ranged from 3603 to 4775 across the seven
datasets, representing 20-27% of all genes (Figure 2A; Dataset S1). To
investigate patterns of similarity in global OSs among these populations, we
carried out hierarchical clustering of the mean OSs for all genes in the genome
(Figure 2B). This analysis revealed an interesting correspondence in the
clustering of OSs with the phylogeny of the corresponding receptor genes$^{25}$:
neurons expressing the most recently-duplicated receptor genes, Ir75b and Ir75c,
were clustered together, with the next most similar neuron population, Ir75a,
expressing the likely ancestral member of this tandem gene array$^{26}$. Ir40a and
Ir64a neurons clustered together away from the other acid-sensing populations,
possibly reflecting the development of these neurons within the sacculus, rather
than in sensilla located on the antennal surface.

Different OSN populations are characterized both by the sensory receptors
they express and by their distinct morphological properties, notably their axonal
projections to discrete glomeruli within the antennal lobe. Genetic studies have
identified several secreted and transmembrane proteins that play roles in different
steps of the neuronal guidance process$^{3,8,9}$, although our understanding remains
incomplete and many key molecules likely remain to be discovered. To gain
insight into the potential complexity and neuron-type specificity of the cell surface
proteome, we examined OSs of ∼2500 cell surface/secreted proteins (CSSPs) annotated in the genome (from the FlyXCDB database\textsuperscript{44}) (Figure 2C and Table S2). Each neuronal population is predicted to express 244-389 CSSP genes (Table S3): these are likely to encode proteins involved in diverse structural and functional properties of OSNs, including sensory transduction, dendrite morphogenesis, axon guidance, and synaptic maturation and transmission. We did not attempt to further categorize the expressed genes, however, as most of them have unknown in vivo roles. While each neuron population had a unique combination of PolII-occupied CSSP genes, only a small number of these were unique to a single dataset (Figure 2C; Table S4). One exception is the Ir40a population (70 uniquely occupied genes); this observation could reflect the distinct sensory modality mediated by these neurons (hygrosensation) and/or their unusually distributed glomerular innervations in the antennal lobe\textsuperscript{22}.

The expression of both chemosensory receptors and guidance molecules depends upon regulatory networks of transcription factors\textsuperscript{2,45}. We therefore assessed global patterns of PolII occupancy of the set of 754 predicted Drosophila transcription factors across the seven Ir populations (using the FlyTF database\textsuperscript{46}; Table S5). Half of these genes (379/754) displayed significant occupancy in at least one population of Ir neuron, with individual OSN classes displaying 190-286 occupied transcription factor genes (Figure 2D and Table S6-S7). There was considerable variation in Dam:PolII occupancy of transcription factor genes across the neuron populations (Figure 2D), consistent with each neuronal subtype possessing a unique and complex gene regulatory network to support its specific differentiation properties.

Many of the transcription factors with known function in OSN development are broadly-expressed across OSNs, such as Acj6, Fer1 and Onecut\textsuperscript{45,47}. Consistently, these genes display positive OS in all seven neuron populations (Table S6). One exception is Engrailed (En), which is expressed and required in just a subset of OSNs\textsuperscript{48,49}. In our TaDa datasets, we observed that en displays significant OSs in Ir31a, Ir40a and Ir64a populations (Figure S2A), consistent with the robust detection of En protein in only these three Ir neuron classes (Figure S2B-C).

**Comparative chromatin accessibility in OSNs**

We next performed CATaDa analysis through investigation of the Dam-alone datasets, by using these to generate chromatin accessibility maps similar to ATAC-seq and FAIRE-seq\textsuperscript{15}. Genome-wide visualization of the peaks of methylated sequences across the genome (see Methods) revealed heterogeneous patterns of chromatin accessibility, with the expected overall decrease in open chromatin towards the heterochromatic centromeric regions of the chromosome arms (Figure 3A).

To examine chromatin accessibility patterns at a gene level, we focused on CATaDa signals at the Ir loci that distinguish different populations – i.e., the seven tuning Ir genes and Ir8a – whose cell type-specific transcriptional activity is well-established\textsuperscript{25,29,30}. For each gene, peaks were present at the presumed promoter region, as well as in upstream regions, which may represent enhancers. Comparisons across neuron populations revealed that peak distribution was very similar across all eight genes (Figure 3B). This observation indicates that neuron-specific differences in Ir gene transcription are not due to differences in promoter
accessibility and also not reflecting dramatic differences in enhancer accessibility. The only exceptions were the presence of peaks upstream of \textit{Ir75a} and \textit{Ir75c} that were unique to their respective neuron populations (q-value < 0.05). Notably, these peaks lie within sequences that recapitulate receptor expression patterns in reporter transgenes\textsuperscript{20,22,25} (Figure 3B), raising the possibility that they reflect enhancers that are accessible/functional only in these Ir neuron populations.

**Identifying genes displaying differential expression in Ir populations**

The analyses of genes encoding chemosensory receptors, transcription factors and CSSPs demonstrate the use of TaDa OSs to identify genes that are expressed (or not expressed) within specific OSN populations. While differences in OS for a given gene across populations may suggest heterogeneous expression levels, we sought to test for PolII occupancy differences between populations. We implemented a different analysis method that used variation in read-depth in the triplicate paired Dam-alone/PolII:Dam TaDa datasets along the gene body (Figure 4A and Methods). We identified candidates that had at least two GATC motifs for which the excess of PolII:Dam read-depth compared to the Dam-alone read-depth significantly varied in the same direction between two or more population datasets (see Methods). With this criterion, 1694 genes exhibited significant variation across the seven populations, indicating that only a relatively small proportion of genes (~8%; Dataset S2) differ among these neuron types. This number is likely to be an overestimate as it includes overlapping and/or intronic genes that, as described above, cannot be parsed further using TaDa data. Importantly this subset encompasses all of the expected \textit{Ir} genes (see Figure 4D and Dataset S2). To examine how each individual Ir neuron dataset varied in comparison to the six others, we also quantified significant differences between each pair of neural populations (Figure 4B; Dataset S3). Despite the methodological differences with the TaDa analysis that produces OSs, the proportion of genes contributing to the total number of pairwise difference approximates the clustering of global OSs (Figure 2B). For example, comparisons among the closely clustered Ir75a/Ir75b/Ir75c datasets result in very few of the total differences (0.4-1.7%). Conversely, slightly more than half (51%) of the total differences in the Ir84a dataset arise from comparisons with the distantly clustered Ir40a dataset.

We investigated the general characteristics of the 1694 differentially-occupied genes by performing Gene Ontology analyses, finding highly significant enrichment in neural, neural development, and signaling categories (Figure 4C). This enrichment is consistent with these genes underlying the known unique physiological and anatomical properties of these neuron types, and indicates that our comparative TaDa approach would be valuable in identifying candidates that contribute functionally to these unique properties.

To establish a priority list for follow-up experiments, we ordered these genes by the total number of differentially-occupied GATC motifs, reasoning that this would highlight those displaying differential occupancy along a broad region of the gene body (Figure 4D and Table S8). While this “ranking” is biased towards longer genes (which on average will have more GATC motifs), we emphasize that this does not exclude interest in genes with fewer GATC motifs; for example, \textit{Ir} genes are distributed widely in this list (Figure 4D). We subsequently focused on experimental validation of the top-ranked transcription factor gene, \textit{pou domain}
motif 3 (pdm3), and the top-ranked CSSP gene, flamingo (fmi; also known as starry night (stan), flybase.org/reports/FBgn0024836), as described below.

**Pdm3 acts as a genetic switch to distinguish Ir75a and Ir75b neuron fate**

The robust signal displayed by *pdm3* for between-population differences is consistent with it having a significant OS only in Ir75a neurons (Figure 4E).

Notably, from the pan-transcription factor survey, *pdm3* is the sole gene that is uniquely occupied in this neuron population (Figure 2D and Table S4). Pdm3 is a nervous-system enriched transcription factor, although its precise *in vivo* binding specificity is unknown. In another olfactory organ, the maxillary palp, Pdm3 is expressed in multiple classes of Or neurons and functions in controlling receptor expression and axon guidance, but its role in Ir neurons is unknown.

Using Pdm3 antibodies, we first analyzed protein expression in antennae in which we co-labeled each of the Ir populations with the driver lines used in the TaDa experiments. Only Ir75a neurons displayed robust nuclear Pdm3 immunofluorescence, although a subset of Ir40a neurons had above-background signals (Figure 5A-B).

To examine the role of *pdm3* in Ir neurons we performed transgenic RNA interference (RNAi) using *peb-Gal4*. This driver begins to be expressed very broadly in OSNs after their terminal divisions (∼16 h APF) and throughout the initiation of olfactory receptor expression (∼40-48 h APF). *pdm3RNAi* led to loss of most IR75a expression, with the remaining detectable neurons expressing very low levels of this receptor (Figure 5C-D). By contrast, robust expression of all other Irs was maintained (Figure 5C-D). We noted, however, that the number of cells expressing IR75b increased in *pdm3RNAi* antennae (Figure 5D). These observations suggest that Pdm3 functions (directly or indirectly) to promote IR75a expression and repress IR75b expression.

The novel IR75b-expressing cells in *pdm3RNAi* antennae are mainly located in the proximal region of the antenna near the sacculus, where Ir75a neurons are normally found (Figure 6A). This distribution suggested that loss of Pdm3 results in a switch of receptor expression from IR75a to IR75b. We investigated this possibility by examining the activity of transcriptional reporters for these receptors, encompassing minimal regulatory DNA sequences fused to CD4:tdGFP (hereafter, GFP). Both *Ir75a-GFP* and *Ir75b-GFP* transgenes are faithfully expressed in IR75a and IR75b protein-expressing neurons (Figure 6B-C). In *pdm3RNAi* antennae *Ir75a-GFP* expression is lost, while *Ir75b-GFP* is expressed in many ectopic cells (Figure 6B-D).

We further used these *Ir-GFP* transgenes to examine the glomerular projection of Ir75a and Ir75b neurons in *pdm3RNAi* animals. In controls, *Ir75a-GFP* and *Ir75b-GFP* label exclusively the DP1l and DL2d glomeruli, respectively (Figure 6E), concordant with previous analyses. In *pdm3RNAi* animals the *Ir75a-GFP* DP1l signal is almost completely lost (Figure 6E), consistent with the lack of reporter expression in the antennae. By contrast, the *Ir75b-GFP* signal in the DL2d glomerulus is greatly intensified, and the glomerulus is enlarged (Figure 6E), presumably reflecting the increased number of Ir75b neurons (Figure 6C-D).

This simplest interpretation of these observations is that in the absence of Pdm3, Ir75a neurons adopt Ir75b neuron fate encompassing both its receptor identity and glomerular innervation pattern.
While OSN fate is established in early pupae\textsuperscript{1-3}, the expression of Pdm3 in Ir75a neurons in adult antennae (Figure 5A-B) suggests that it has a persistent role in these cells. To test this idea, we drove pdm3\textsuperscript{RNAi} using Ir8a-Gal4\textsuperscript{29}, which is active only from the second half of pupal development (\~48 h APF) by which time OSN axonal convergence on glomeruli has occurred. In adult Ir8a-pdm3\textsuperscript{RNAi} antennae, we found that the number of IR75a-expressing cells was unchanged compared to controls, although protein levels were lower in several cells, arguing that IR75a neuron fate had been largely correctly established (Figure 6F-G). However, as observed with peb-Gal4-induced pdm3\textsuperscript{RNAi}, IR75b was still detected in many ectopic cells (Figure 6F). Notably, we detected several cells expressing both IR75b and IR75a, which are never observed in control antennae (Figure 6G-H). These observations indicate that Pdm3 maintains its role in repressing IR75b expression in mature IR75a neurons.

Dynamic, heterogeneous expression of Fmi in OSNs

fmi attracted our attention both for its heterogeneous occupancy across Ir neuron populations (Figure 4D-E), and because this gene encodes an atypical cadherin – comprising an extracellular cadherin-like domain and a transmembrane G protein-coupled receptor (GPCR)-like domain – with diverse roles in axonal and dendritic targeting in other regions of the nervous system\textsuperscript{53-58}. To characterise the endogenous expression of Fmi in the olfactory system, we used an antibody to profile protein levels in OSNs. We first examined antennae from mid-pupal stages (48 h APF) but observed only trace levels of immunoreactivity across OSN soma (Figure S3A). At 24 h APF this signal was slightly stronger, but much more prominent in the OSN axons as they left the antenna (Figure S3A), suggesting that this protein is predominantly transported to the axonal compartment of these neurons. Indeed, Fmi protein was strongly detected in the OSN axons as they enter the antennal lobe at 24 h APF (Figure 7A). Fmi was still robustly expressed at 48 h APF as these OSNs coalesced onto individual glomeruli with broad, but slightly heterogeneous, expression in glomeruli across the antennal lobe (Figure 7A). Subsequently, Fmi levels began to decrease to undetectable levels in some glomeruli in the adult stage, while reproducibly persisting in others (Figure 7B). Most of these glomeruli are innervated by Or neurons, but we could detect Fmi in DL2d (Ir75b) and DL2v (Ir75c), but not the neighboring DP1l (Ir75a) (Figure 7B), in partial concordance with the TaDa data (Figure 4E).

Fmi was also detected in a group of \~16 cells located on the antero-dorsal region of the antennal lobe from later developmental stages (\~48 h APF) (Figure 7A). These cells are likely to be LNs, based upon their expression of Elav (Figure S3B) and presence of stained processes that innervate the antennal lobe but do not project to higher olfactory centers. Numerous LN subtypes exist, and these often have broad innervation patterns across many glomeruli\textsuperscript{59,60}. To examine the specific contribution of OSNs to the Fmi immunoreactivity observed in the antennal lobe, we depleted Fmi in LNs by RNAi using a combination of Gal4 drivers (1081-Gal4 and 449-Gal4\textsuperscript{60}) that cover the vast majority of these interneurons (Figure S3C). At mid-pupal development (48 h APF), differential glomerular expression of Fmi was much more marked compared to control animals (Figure 7C). Although we could not confidently define all glomerular identities at this earlier developmental stage, prior to full maturation of the antennal lobe, Fmi protein levels did not appear to correspond precisely with the
TaDa OSs (e.g., the VL2a glomerulus, innervated by Ir84a neurons, was not devoid of Fmi signal). This observation suggests that Fmi is dynamically expressed in different populations of OSNs during development. We note that the protein observed at 48 h APF must be due to transcription of \textit{fmi} at an earlier time point than captured by our TaDa analysis. Additionally, different post-transcriptional regulation within OSN populations could also contribute to lack of strict correlation between transcript and protein levels with these cells, as noted in other tissues\textsuperscript{61}. Nevertheless, this analysis reveals developmentally dynamic and population-specific control of Fmi expression in OSNs.

\textbf{Fmi is required for segregation of OSNs into distinct glomeruli}

To determine the function of Fmi in the developing olfactory system, we induced \textit{fmi} RNAi in developing antennal tissue first using a constitutive driver (\textit{ey-Flip,act-stop}\textit{Gal4}\textsuperscript{62}), which should deplete Fmi expression from the earliest stages of development, and examined adult brains stained with the general neuropil marker nc82 (Bruchpilot). The antennal lobes of these animals showed fully penetrant, severe morphological defects: the stereotyped glomerular boundaries observed in the lobes of control animals were largely lacking, with only a few, large glomerulus-like subregions detected (Figure 8A). This phenotype was reproduced with an independent RNAi line that targets a distinct region of the coding sequence (Figure 8A). We further examined this phenotype by visualizing the projection patterns of specific classes of OSNs labeled by receptor promoter-driven fluorescent reporters (Figure 8B-C). Neurons labeled by \textit{Ir75a-GFP}, \textit{Ir75b-GFP} or \textit{Ir75c-GFP} were no longer constrained to discrete glomeruli (Figure 8B), although the mistargeting was relatively mild within the antennal lobe as a whole. The limited defects of individual neuron populations suggests that the lack of Fmi does not disrupt long-range targeting properties of OSNs, but rather impacts local segregation of OSNs to discrete glomerular targets. Consistently, double labeling of neurons that normally project to adjacent glomeruli (\textit{Ir75a-GFP} and \textit{Ir75b-Tomato} (\textit{Tom}); \textit{Ir84a-GFP} and \textit{Ir31a-Tom}) revealed frequent overlap between these normally-segregated populations (Figure 8C).

The constitutive driver is expressed broadly across the antennal disc\textsuperscript{62}. The observed phenotypes could therefore potentially be due to a requirement for Fmi in disc development rather than in OSNs. To limit \textit{fmi} RNAi to OSNs, we used \textit{peb-Gal4}, which produced an equally strong defect in antennal lobe glomerular segregation (Figure 8D). In these animals the LN-expressed Fmi was still detectable but, in contrast to the expression in OSNs, Fmi was apparently homogeneously distributed across antennal lobe glomeruli (Figure S3D). This LN source of Fmi does not appear to contribute to glomerulus formation as \textit{LN>fmi\textsc{RNAi}} animals did not exhibit the same antennal lobe morphological defects (Figure 8D). Similarly, when we targeted \textit{fmi} to the synaptic partners of OSNs, the PNs – which pre-pattern the glomerular organization of the antennal lobe\textsuperscript{3} – using the \textit{GH146} driver, antennal lobe morphology appeared normal (Figure 8D). The absence of defects is consistent with the lack of detectable expression of Fmi in these second-order neurons (Figure 7A). Together, these results indicate that Fmi functions principally, if not exclusively, within OSNs to correctly pattern glomerular formation.
To understand the genesis of this defect during OSNs development we used a temperature-sensitive repressor of Gal4 (Gal80\textsuperscript{ts}; inactivated at 30°C) to temporally-control the onset of peb\textsuperscript{>fmRNAi}. When animals were shifted to 30°C 10 h APF (prior to initial expression of peb-Gal4), all animals displayed antennal lobe morphological defects of varying degrees, and Ir75b-GFP-labeled neurons did not coalesce to a discrete glomerular unit (Figure 8E). When shifted at 20 h APF, the severity of the phenotype was diminished (Figure 8E). When shifted at 40 h APF, the antennal lobe and Ir75b neuron targeting were indistinguishable from controls (Figure 8E). These experiments indicate an early developmental function for Fmi.

We examined this early requirement in more detail by imaging the OSN axons as they invade the antennal lobe. During 20-25 h APF, pioneer OSNs reach the antennal lobe in a unique bundle that separates into ventromedial and dorsolateral bundles, which project around the lobe’s border\textsuperscript{9,63} (Figure 8F). In fmi\textsuperscript{RNAi} animals, OSNs reach the antennal lobe, but bundle organization is highly disrupted, with subsets of axons apparently defasciculating prematurely in the central area of the lobe (Figure 8F).

Fmi (and its mammalian homologs) have roles in multiple biological processes in the nervous system as well as in the establishment of epithelial planar cell polarity\textsuperscript{58,64,65}. Genetic and biochemical studies have identified a number of interaction partners of Fmi in different biological contexts. We tested mutations and/or multiple independent RNAi lines for several of these genes to determine whether they function with this cadherin in OSNs (Figure S4). Mutations in two core components of the planar cell polarity pathway, frizzled and dishevelled\textsuperscript{66}, had no apparent impact on antennal lobe formation. Similarly, loss of frazzled (encoding a Netrin homolog which function with Fmi in midline axon guidance\textsuperscript{67}), golden goal (with which Fmi collaborates in photoreceptor axon guidance\textsuperscript{68}), or espinas (which encodes a LIM domain protein that binds Fmi intracellularly and functions in dendrite self-avoidance\textsuperscript{57}) did not reproduce the fmi phenotype. Finally, although the intracellular signaling mechanism of Fmi is not well understood, the Fmi GPCR-domain has been suggested to act through the G\textgreekalpha subunit to regulate dendrite growth\textsuperscript{69}; however, we found no evidence that this signaling mechanism operates in OSNs (Figure S4). These observations suggest that Fmi functions through other cellular mechanisms to control OSN axon segregation.

**Discussion**

Determination of the molecular composition of neurons is essential to understand the development and function of the nervous system. While scRNA sequencing is an undoubtedly powerful approach to address this challenge\textsuperscript{70-74}, it requires often technically-difficult cell isolation, and can be biased towards abundant cell types and highly-expressed genes\textsuperscript{75}. Moreover, it is unclear to what extent neurons change their expression properties during tissue dissection (as reported for other types of cell\textsuperscript{76}) or through inevitable loss of dendritic and axonal processes (where certain transcripts may be enriched\textsuperscript{77}). Our study was initially motivated by the desire to test the TaDa method for directed, genome-wide molecular profiling of very small populations of neurons that are tightly embedded within a highly heterogeneous tissue. Moreover, by applying TaDa to a set of functionally-related OSNs, we aimed to identify cell type-specific differences that would point towards
mechanisms underlying the development and evolution of novel neuronal populations.

The unique olfactory receptor expression pattern of OSNs (together with other chemosensory genes) offered several positive- and negative-control loci that allowed us to validate the specificity and sensitivity of TaDa. One caveat emerging from this analysis is that TaDa cannot easily discriminate PolII occupancy of interleaved genes, an issue that may be most pronounced in the relatively gene-dense genome of *Drosophila*. Nevertheless, our results indicate that this method could be effective in defining candidate receptors in neuronal classes for which a genetic driver is available. This possibility is of particular interest in the *Drosophila* gustatory system where many populations of neurons can be labeled and physiologically profiled, but incomplete knowledge of the combinatorial receptor expression properties that are characteristic of this sensory system hampers identification of the cognate sensory receptor(s).

Recent work described an scRNA-seq analysis of antennal OSNs at a mid-pupal stage (42-48 h APF), leading to the identification of transcriptional clusters for 20 molecularly-defined classes of OSNs (a further 13 clusters were not identified). Detailed comparison of these scRNA-seq data and our TaDa datasets is currently difficult, as only two of the scRNA-seq clusters correspond to populations we analyzed by TaDa (Ir75a and Ir84a neurons). Moreover, while the scRNA-seq captures the transcriptional profile during a narrow developmental window, our TaDa experiments integrated PolII occupancy patterns at an OSN population level over several days from the initiation of Ir-Gal4 expression. Perhaps reflecting this difference in transcriptomic profiling, scRNA-seq detected on average \(~1500\) expressed genes (including \(~150\) encoding transcription factors) per OSN type, which is lower than the number of occupied genes detected by TaDa (on average \(~4000\) per Ir population, including \(~225\) transcription factor genes). However, many other experimental and/or analytical differences in the two methods are likely to impact these global numbers.

Genome-wide analysis of TaDa datasets from different Ir populations revealed an intriguing positive relationship between the similarity of global patterns of PolII occupancy between neuronal populations and the phylogenetic distance of the olfactory receptors they express, notably the tandem array of Ir75a, Ir75b and Ir75c genes. Similar analysis of Or OSN populations is limited by the broad phylogenetic distribution of the ORs whose corresponding neuronal transcriptomes have been identified. However, through examination of these data we identified two pairs of closely-related (though not tandemly-arranged) Or genes whose neuronal transcriptional profiles also cluster (Or42b/Or59b and Or9a/Or47a). The simplest interpretation of these observations is that neuronal populations expressing distinct, relatively-recent receptor gene duplicates derive from a common ancestral neuronal population (which may have originally co-expressed the duplicated receptors); these extant populations therefore retain correspondingly close proximity in their transcriptomic profile. The principle of OSN “sub-functionalization” – which has not, to our knowledge, previously been recognized empirically – has important implications for understanding the evolution of novel olfactory pathways.

One important advantage of the TaDa profiling method over scRNA-seq is that it can also provide information on chromatin state, which remains difficult to characterize through standard chromatin immunoprecipitation-based methods. We recognize that CATaDa provides only a relatively coarse-grained view of
chromatin state (similar to ATAC-seq\textsuperscript{15}), and variant approaches of TaDa (e.g., Dam fusions to chromatin-binding proteins\textsuperscript{32}) are very likely to reveal finer-scale differences between populations. Nevertheless, our findings of similar chromatin accessibility at \textit{Ir} genes across \textit{Ir} populations are in-line with studies in the developing \textit{Drosophila} embryo where chromatin accessibility is comparable between anterior and posterior region, despite substantial differences in patterns of gene transcription across this body axis\textsuperscript{83}. We suggest that the specific developmental properties of OSNs are more likely to be driven by the unique combination of TFs they express. Indeed many of the differentially occupied genes between \textit{Ir} populations encode known or predicted TFs. In agreement, we validate the selective expression and function of Pdm3 in \textit{Ir75a} neurons, revealing it to be a key factor – and potentially the sole transcription factor – that distinguishes the closely-related \textit{Ir75a} and \textit{Ir75b} neuron populations. Future identification of Pdm3 targets in these neurons will be necessary to determine whether it regulates these \textit{Ir}s directly, and to identify other downstream genes that control the innervation of the corresponding neurons to distinct glomeruli.

Our TaDa datasets also revealed a rich diversity in the predicted cell surface proteome of \textit{Ir} neurons. Because of the temporal breadth of TaDa profiling, these proteins are likely to contribute to differences of these neurons in cilia/dendritic morphogenesis, axon targeting and synapse formation/plasticity. Our initial characterization of one of these, the atypical cadherin Fmi, reveal it to be a dynamically and heterogeneously-expressed, axon-localized protein. Importantly, Fmi is not restricted to \textit{Ir} neurons, but found throughout the peripheral olfactory system, concordant with the dramatic loss of glomerular architecture in the antennal lobe in the absence of this protein. We have delimited the role of Fmi to initial sorting of OSNs within the antennal lobe; unfortunately, this early requirement makes it difficult to assess the significance of the heterogeneous expression of this cadherin as levels of Fmi cannot be easily discerned in specific OSN populations within the bundle. We have attempted to overexpress Fmi to assess the consequence of this alternative perturbation of Fmi levels on OSN axonal projections but this did not yield overt defects (data not shown). This negative result may simply reflect technical limitation in achieving adequate high-level Fmi overexpression, as reported in studies of this protein in other systems\textsuperscript{55}. Differential expression of Fmi has been proposed to be important for its role in photoreceptor axon guidance\textsuperscript{54}. As Fmi is thought to function as a homophilic adhesion molecule\textsuperscript{66}, we speculate that heterogeneous Fmi-dependent adhesion between OSN populations enables them to sort themselves within the axon bundle and/or defasciculate at specific places, before local, instructive guidance cues direct them to particular locations within the lobe.

An adhesion-dependent role for Fmi is consistent with the \textit{fmi} loss-of-function phenotype most closely-resembling that of mutations in \textit{CadN}\textsuperscript{84}. \textit{CadN} is however homogeneously expressed across the antennal lobe, and also present in their PN synaptic partners\textsuperscript{84,85}. Future work will be required to uncover the special role of the atypical cadherin Fmi in correct segregation of OSNs axons into distinct glomerular subunits.

Beyond further functional characterization of differentially-occupied genes identified in this study, several future applications of TaDa (and CATaDa) in the olfactory system, and other sensory systems, can be envisaged. For example, temporal analysis of PolII occupancy could be refined by restricting Gal4 activity (with Gal80\textsuperscript{55}) to a particular developmental time-window. Mining of our TaDa datasets also revealed a rich diversity in the predicted cell surface proteome of \textit{Ir} neurons. Because of the temporal breadth of TaDa profiling, these proteins are likely to contribute to differences of these neurons in cilia/dendritic morphogenesis, axon targeting and synapse formation/plasticity. Our initial characterization of one of these, the atypical cadherin Fmi, reveal it to be a dynamically and heterogeneously-expressed, axon-localized protein. Importantly, Fmi is not restricted to \textit{Ir} neurons, but found throughout the peripheral olfactory system, concordant with the dramatic loss of glomerular architecture in the antennal lobe in the absence of this protein. We have delimited the role of Fmi to initial sorting of OSNs within the antennal lobe; unfortunately, this early requirement makes it difficult to assess the significance of the heterogeneous expression of this cadherin as levels of Fmi cannot be easily discerned in specific OSN populations within the bundle. We have attempted to overexpress Fmi to assess the consequence of this alternative perturbation of Fmi levels on OSN axonal projections but this did not yield overt defects (data not shown). This negative result may simply reflect technical limitation in achieving adequate high-level Fmi overexpression, as reported in studies of this protein in other systems\textsuperscript{55}. Differential expression of Fmi has been proposed to be important for its role in photoreceptor axon guidance\textsuperscript{54}. As Fmi is thought to function as a homophilic adhesion molecule\textsuperscript{66}, we speculate that heterogeneous Fmi-dependent adhesion between OSN populations enables them to sort themselves within the axon bundle and/or defasciculate at specific places, before local, instructive guidance cues direct them to particular locations within the lobe.

An adhesion-dependent role for Fmi is consistent with the \textit{fmi} loss-of-function phenotype most closely-resembling that of mutations in \textit{CadN}\textsuperscript{84}. \textit{CadN} is however homogeneously expressed across the antennal lobe, and also present in their PN synaptic partners\textsuperscript{84,85}. Future work will be required to uncover the special role of the atypical cadherin Fmi in correct segregation of OSNs axons into distinct glomerular subunits.

Beyond further functional characterization of differentially-occupied genes identified in this study, several future applications of TaDa (and CATaDa) in the olfactory system, and other sensory systems, can be envisaged. For example, temporal analysis of PolII occupancy could be refined by restricting Gal4 activity (with Gal80\textsuperscript{55}) to a particular developmental time-window. Mining of our TaDa...
datasets for population-specific genes may help to design novel OSN lineage-specific driver lines that are expressed earlier in development (of which few are currently known\textsuperscript{62}) to profile gene expression patterns prior to onset of receptor expression. Finally, in adult flies, comparison of PolII occupancy patterns in a specific neuron population before, during and after odor exposure could provide insight into neural activity-regulated gene expression and its relevance for sensory adaptation and plasticity.

Acknowledgements

We thank Cheng-Ting Chien, Ya-Hui Chou, Fisun Hamaratoglu, Vladimir Katanaev, Tadashi Uemura, Liqun Luo, the Bloomington Drosophila Stock Center (NIH P40OD018537), the Vienna Drosophila Resource Center, and the Developmental Studies Hybridoma Bank (NICHD of the NIH, University of Iowa) for reagents. Sequencing was performed at the Lausanne Genomic Technologies Facility, and the Swiss Institute of Bioinformatics Vital-IT group provided computational resource support. We are grateful to Andrea Brand’s lab for assistance in establishing the antennal TaDa protocol. Tony Southall and Owen Marshall for advice on the DamID analysis pipeline, Simon Anders for suggestions on DESeq2 usage, and members of the Benton laboratory for comments on the manuscript. J.R.A. was supported by a post-doctoral fellowship from the Novartis Foundation for medical-biological Research (12A14), and is currently supported by a Swiss National Science Foundation Assistant Professorship (PP00P3 176956). Research in R.B.’s laboratory was supported by the University of Lausanne, an ERC Consolidator and Advanced Grants (615094 and 833548, respectively) and the Swiss National Science Foundation.

Author contributions

All authors contributed to experimental design, data analysis and interpretation of data. Figure contributions were as follows: J.R.A. (Figures 1F-G, 2A-D, 4B-E, 5B, 7C, 8E, S1, S2A,C); L.A. (Figures 1B-C, 5A-C, 7C, 8A-E, S2B-C, S3B-C, S4); J.A. (Figures 7A-B, 8A-C, 8E-F, S3D); K.M. (Figure 5C-D, 6, S3A); P.C.C. (Figure 3, 4C). R.B. wrote the paper with contributions and feedback from all other authors.

Competing interests

The authors declare no competing interests
Materials and Methods

Drosophila culture

Flies were maintained on a standard wheat flour/yeast/fruit juice diet at 25°C in 12 h light:12 h dark conditions. Published mutant and transgenic D. melanogaster strain are described in Table S9. For the temporal control of fmi RNAi with Gal80<sup>ts</sup>, flies of the desired genotype were cultivated at 19°C; animals were staged by selecting white pupae (designated as 0 h after puparium formation (APF)) and shifting these animals to 29°C after 10, 20 or 40 h, to permit induction of RNAi at different time-points during pupal development. Unless otherwise indicated for specific experiments, antennae or brains were collected from both sexes.

Transgenic flies

New transgenes were constructed using standard molecular biological procedures and transgenesis was performed by BestGene Inc. using the phiC31 site-specific integration system. Ir75a-CD4:tdGFP (Ir75a-GFP) was generated by subcloning the 5’ (7833 bp) and 3’ (1921 bp) genomic sequences from the Ir75a-Gal4 transgene<sup>22</sup> to flank CD4:tdGFP in pDESTHemmarG (Addgene 31221)<sup>36</sup>, and integrated into the attP40 landing site. Ir75c-CD4:tdGFP (Ir75a-GFP) was generated by subcloning the 5’ (3072 bp) and 3’ (2545 bp) genomic sequences from the Ir75c-Gal4 transgene<sup>25</sup> to flank CD4:tdGFP in pDESTHemmarG and integrated into attP40. Ir75b-CD4:tdTom (Ir75b-Tom) was generated by cloning the 299 bp genomic sequences immediately upstream of the Ir75b start codon (as in the previously-described Ir75b-GFP transgene<sup>25</sup>) into pDESTHemmarR (Addgene 31222) and integrated independently into attP5 and attP2 (on chromosome II and III, respectively). Ir31a-CD4:tdTom (Ir31a-Tom) was generated by cloning the 2000 bp genomic sequences immediately upstream of the Ir31a start codon, via Gateway recombination, into pDESTHemmarR, and integrated into attP2. Ir84a-CD4:tdGFP (Ir84a-GFP) was generated by cloning the 164 bp genomic sequences immediately upstream of the Ir84a start codon, via Gateway recombination, into pDESTHemmarG and integrated into VK00027.

TaDa sample preparation

Antennae from ~2000 flies (~1-8 days old) of the desired genotype were harvested by snap-freezing flies in a mini-sieve (Scienceware, Bel-Art Products) with liquid nitrogen, tapping flies to detach and collect appendages in a Petri dish under the sieve containing Triton X-100 (0.1%). Third antennal segments were selected under a binocular microscope by pipetting, and transferred to a 1.5 ml Eppendorf tube. After brief rinsing in PBS to remove detergent, antennae (occupying a volume of ~50-70 μl) were resuspended in 50 μl PBS and homogenized manually with a blue pestle (Sigma Aldrich Z359947-100EA) for at least 2 min. Antennal fragments were rinsed off the pestle with 100 μl PBS and 40 μl EDTA (100 mM) and 20 μl RNase (12.5 μg/μl) (Qiagen 19101) were added to the antennal lysate, mixed and allowed to sit for 3-5 min. 20 μl Proteinase K (Qiagen DNeasy Blood & Tissue Kit 69504) and 200 μl Buffer AL were added, and mixed by gently pipetting up and down with a blue tip ~50 times, before
incubation in a 56°C heat block for 30 min. The digested lysate was centrifuged at
10,000 rpm for 30 s and the supernatant transferred to a new tube and allowed to
cool before proceeding with DNA extraction.

Subsequent sample processing steps for TaDa – i.e., DpnI treatment (which cuts at adenine-methylated GATC sites), DamID adaptor ligation, DpnII treatment (which cuts at non-methylated GATC sites, thereby digesting unmethylated DNA fragments), and PCR amplification and purification of methylated GATC fragments – were performed essentially as described\(^87\). PCR products were sonicated using a Covaris S220 to obtain ~300 bp average fragment size; DNA size and quality were verified by Qubit quantification. DamID adaptors were removed by Sau3AI digestion. Truseq Nano libraries were prepared at the Lausanne Genomic Technologies Facility and sequenced (SR 100 bp) on a HiSeq 2500.

**TaDa data analysis**

To test for an enrichment of the genome occupancy of Dam:PolII relative to Dam alone, we used the “damidseq pipeline” v1.4\(^88\), with release 6 of the *D. melanogaster* genome as the reference. Within this pipeline, Bowtie v2.2.6\(^89\) was implemented to align the Illumina sequence reads to the reference, and data processing called on SAMtools v1.2\(^90\) and BEDTools v2\(^91\). This pipeline outputs log2 ratio files in bedgraph format. We calculated Pearson’s Correlation Coefficient among triplicate datasets based on the bedgraph files, and also used them to generate a single averaged occupancy scores file for each of the seven neuron data sets. Reproducibility among all seven triplicate datasets (average \(r^2 = 0.85\)) was consistent with previous TaDa studies\(^14\). The files containing the averaged OSs were then inputted into the “polii.gene.call” script (https://github.com/owenjm/polii.gene.call) for estimating an FDR for each gene’s occupancy score. A file detailing the workflow and corresponding code is available on GitLab (https://gitlab.com/roman.arguello/ir-tada).

The set of genes encoding cell surface and secreted proteins (CSSPs) were based on a custom filtered set from the *D. melanogaster* extracellular domain database (FlyXCDB, vFB2015_03\(^44\)) (Table S2). The set of genes encoding transcription factors used in the study were based on a custom filtered set from the FlyTF.org database v2\(^46\) (Table S5).

For testing the hypothesis that OSs are equal among datasets (neuron populations), we used DESeq2\(^92\). Unlike the DamID pipeline, this approach focused on individual GATC fragments within each gene. To quantify read coverage of each fragment, we used the same bam files that were outputted by the DamID pipeline (above). We converted these to bed files using BEDTools v2 “bamToBed” utility and calculated the coverage of GATC fragments using the Bedtools’ “coverage” utility (Dataset S4). Differential occupancy was tested over the full dataset using a Likelihood Ratio Test, where the full model specified neuron populations and DamID condition (Dam:PolII vs. Dam-alone) as factors, and an interaction term (between DamID condition and cell population): ~ cell.pop + cond + cond:cell.pop. The reduced model removed the interaction term: ~ cell.pop + cond. We also carried out pairwise tests using a Wald Test within DESeq2. Genes that were identified as candidates for differential expression were required to have \(\geq 2\) GATC fragments contributing to the difference (that were also significantly PolII occupied).
Gene Ontology (GO) enrichment analyses were performed on the subset of genes that showed differential occupancy using clusterProfiler (v3.10.1)\(^3\). We compared these genes against the three independent, controlled vocabularies provided by GO Consortium that model Biological Processes, Molecular Function and Cellular Component\(^4\).

**CATaDa analysis**

CATaDa was performed on the seven populations of neurons using the same sets of sequencing reads (i.e., Dam-alone data) acquired for TaDa analysis. The bioinformatics pipeline followed the general concepts of a previously described pipeline\(^5\), but with modifications. For each neuron population, the genomic sequencing reads from all three biological replicates were aligned to release 6.21 of *D. melanogaster* genome with “very-sensitive” settings (Bowtie2 v2.3.0). Following pooling of aligned reads from the replicates (samtools v1.7), the chromatin accessibilities were represented as the coverage of reads per bin of 10 bases, scaled to \(1 \times\) average genomic coverage (~142.57 Mb) and averaged over a sliding window of 100 bases (bamCoverage, version 3.1.3). The chromatin accessibility profile was plotted on genome browser tracks with pyGenomeTracks v2.1. Significant accessible regions were identified with peak-caller (MACS2, v2.1.1.20160309) at a FDR cutoff (q-value) of 0.05. The correlation clustering of peaks among all assayed neuron populations were analyzed with multiBamSummary v3.1.3 on Pearson method and bin size of 3000 bases.

**Histology**

Immunofluorescence and RNA FISH on whole-mount antennae or antennal cryosections were performed as described\(^6\). Immunofluorescence on whole-mount brains was performed essentially as described\(^7,8\). Antibodies used are listed in Table S10. RNA probes for *Ir31a* and *Ir84a* were previously described\(^9\).

Quantification of En and Pdm3 immunofluorescence signals in OSN soma nuclei (Figure S2B-C and Figure 5A-B) was performed in Fiji\(^10\). In brief, regions-of-interest (ROIs) were manually defined using the GFP signal (which circumscribes the nucleus) in a single optical slice where the cross-section of a given nucleus has the largest area, followed by automated measurement of the average En or Pdm3 immunofluorescence signal within this ROI. Background fluorescence signals in the tissue were measured in the same way within equivalently-sized, arbitrarily-chosen ROIs that did not overlap with GFP signals or any other OSN nuclei (visualized with DAPI).

**Statistics statement**

Unless stated otherwise, statistical significance thresholds were \(p < 0.05\) and the tests carried were two-tailed. Measurements were taken from distinct samples, unless otherwise stated, along with a corresponding correction for multiple tests.

**Data availability**

TaDa sequencing data and experimental information have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) and will be
available under accession number E-MTAB-8935. Raw image data for Figures 1, 5-8, S2-S4 are available upon request.

**Code availability**

A detailed workflow for the TaDa analyses and associated code are available at GitLab (https://gitlab.com/roman.arguello/ir-tada).

**Biological material availability**

All unique biological material generated in this work is available from the corresponding author upon request.
Figure Legends

Figure 1. Targeted DamID of OSN populations

(A) Schematic of the principles of Targeted DamID (TaDa) and Chromatin accessibility TaDa (CATaDa). TaDa reports on the enrichment of GATC methylation (Me; blue lines) by an RNA PolII:Dam fusion (top) relative to Dam alone (bottom), thereby identifying PolII-transcribed genes. In CATaDa, analysis of methylation patterns of free Dam provides information on chromatin accessibility.

(B) Immunofluorescence for IR64a and IR8a on antennal sections of animals expressing in Ir64a neurons either Dam:PolII (Ir64a-Gal4/+;UAS-LT3-Dam:Rp215/) or Dam alone (Ir64a-Gal4/+;UAS-LT3-Dam/+). The approximate field of view shown is indicated by the green box on the antennal schematic on the left. The merged fluorescent channels are overlaid on a brightfield background to reveal anatomical landmarks. The arrowheads point to IR64a that is localized in the neuron sensory endings with the sensillar hairs. Scale bar = 10 μm.

(C) Immunofluorescence for GFP and the neuropil marker nc82 on whole mount brains of animals expressing in Ir64a neurons a CD8:GFP reporter together with Dam:PolII (Ir64a-Gal4/+;UAS-LT3-Dam:Rp215/) or free Dam (Ir64a-Gal4/+;UAS-LT3-Dam/UAS-mCD8:GFP). Two focal planes are shown to reveal the two glomeruli (DC4 and DP1m) innervated by Ir64a neurons. Background fluorescence in the GFP channel was slightly higher across the brain in the Dam-alone samples. Scale bar = 20 μm.

(D) Quantification of sequence reads of methylated GATC fragments mapping to the Ir64a gene (transcribed left-to-right, as indicated by the arrowhead; gray dots under the gene model indicate the approximate position of GATC motifs). The transcription start site of Ir64a is unknown. For each of three replicate experiments, the top row is the Dam:PolII sample, which has higher read counts compared to the Dam-alone sample below. Numbers on the right of each row indicate the read depth. The OS is based on data across the full gene body, and the FDR takes into consideration the number of GATC fragments. The range of the y-axis is 0-1000 for all rows.

(E) Top: comparisons among the seven Ir neuron datasets of the sequence reads mapping to the Ir64a gene for the Dam:PolII and Dam samples (a single representative replicate is shown for each). Genotypes are of the form shown in (B), except for Ir75a and Ir75c neuron populations where third chromosome Ir-Gal4 driver transgenes were used. Only the Ir64a population displays a higher read count for Dam:PolII samples compared to Dam-alone samples. The range of the y-axis is 0-1000 for all rows. Bottom: heatmap representation of the occupancy scores (OSs; color scale on the right) for Ir64a calculated across triplicates for each of the seven neuron populations (see Methods).

(F) OS heatmap for the seven target Ir genes, as well as the main Ir co-receptor genes, in the seven Ir neuron populations. OS and FDR values associated with each Ir gene within the corresponding neuron population are shown in Table S1.

(G) OS heatmap of diverse positive- and negative-control genes in the seven Ir neuron populations (i.e., with known or expected expression/lack of expression).

See text for definitions of abbreviations.
Figure 2. Global analyses of TaDa datasets

(A) Bar plot summarizing the number and percentage of the total genes in the genome that have significantly positive OSs within each Ir neuron population (out of 17,611 genes in the D. melanogaster genome).

(B) Hierarchical clustering of the seven Ir neuron populations based upon the OSs of all genes. All nodes had strong support except for the one denoted with an Approximately Unbiased $p$-value of 60.

(C) Heatmap displaying the hierarchical clustering of OSs for the 2509 genes encoding cell surface/secreted proteins (CSSPs). For each population, the total number of genes with a significant OS, and the number of those that are unique to a given population, are shown at the bottom (see also Tables S2-4).

(D) Heatmap displaying the hierarchical clustering of OSs for the set of 754 putative transcription factor genes. For each population, the total number of genes with a significant OS, and the number of those that are unique to a given population, are shown at the bottom (see also Tables S5-7).

Figure 3. CATaDa reveal global similarity in chromatin architecture at Ir genes across Ir neuron populations

(A) Comparison of chromatin accessibility peak locations ($q < 0.05$) over the D. melanogaster genome for two example neuron populations, Ir75c and Ir40a. The y-axis represents the fold enrichment for the peak summit against random Poisson distribution of the small local region (1000 bp). 2L/3L and 2R/3R refer to left and right arms of chromosomes 2/3; “c” at the end of each arm indicates the centromeric (heterochromatic) end of the arm (the Y chromosome is mainly heterochromatic).

(B) Plots of Ir gene exons are shown in black, and UTRs in white; exons of flanking genes are shown in gray. The arrowheads indicate the direction of Ir transcription. The minimal defined regulatory sequence to recapitulate Ir expression is indicated with blue bars (color-coordinated with the arrowheads for Ir75a, Ir75b and Ir75c). The y-axis scale shows the CATaDa accessibility at the genomic region with RPGC normalization method. The red bars are the significant peaks with $q < 0.05$. In the Ir75a/Ir75b/Ir75c genomic region, the arrows mark significant peaks that are unique to the corresponding Ir neuron populations, and which are contained with known regulatory regions of these Ir genes.

Figure 4. Differentially-occupied genes across Ir populations

(A) Schematic illustrating the identification of a differentially-occupied gene between two Ir populations by using variation in read depth at GATC motifs (see Methods). In this example, occupancy of "gene X" is compared between Ir64a neurons and Ir31a neurons. At GATC motifs (black bars), a test is applied to determine if significant variation exists in the relative read depths (in triplicate Dam:PolII versus Dam-alone experiments) between these two neuron populations. Fictitious data for two GATC motifs are depicted: for both, the relative read depth in the Dam:PolII experiments (compared to the Dam-alone experiments) is greater in the Ir64a neurons.

(B) Stacked bar plots of the numbers of genes in the genome that are differentially occupied among Ir neuron populations based on pairwise tests (see Methods). The colors indicate proportions of differentially occupied genes that
arise from comparisons with the other six neuron populations. The figures above the bars indicate the numbers of genes across all comparisons.

(C) Gene Ontology (GO) analysis of the 1694 differentially-occupied genes, showing the top ten over-represented GO terms. The x-axis represents the number of genes annotated to a particular GO term in the input subset. The $p$-value indicated by the colored bar is the probability of observing at least the same number of genes associated to that GO term compared to what would have occurred by chance.

(D) Ranking of the 1694 differentially-occupied genes ordered by the number of GATC motifs contributing to their between-neuron differences (uncorrected for gene length).

(E) OS heatmaps (calculated as in Figures 1-2) of $pdm3$ and $fmi$ in the seven Ir neuron populations.

Figure 5. Heterogeneous expression and function of Pdm3 in Ir neurons

(A) Immunofluorescence for Pdm3 and GFP on antennal sections of animals in which the indicated Ir neuron populations are labeled with a CD8:GFP reporter. Genotypes are of the form: $\text{Ir}xx\text{-Gal4/+;UAS-mCD8:GFP/+}$ or, for Ir75a and Ir75c neurons, $\text{UAS-mCD8:GFP/+;Irxx-Gal4/+}$. Scale bar = 10 µm.

(B) Density plots for Pdm3 immunofluorescence signals (arbitrary units, a.u.) quantified from Ir neuron nuclei (each represented by a vertical line) in the genotypes in (A) and from background signals pooled across genotypes (see Methods). Most signals from Ir neuron nuclei fall within the range of background signals; Ir75a neurons have consistently higher levels of immunofluorescence. Sample sizes are shown to the right of the plot.

(C) Immunofluorescence or RNA FISH for the indicated IRs on whole-mount antennae (or antennal sections for IR40a and IR64a, due to poor antibody penetration of whole-mount tissue) of control animals ($\text{peb-Gal4,UAS-Dcr-2/+;UAS-Dcr-2/+;CyO}$) or two independent $pdm3$ RNAi lines ($\text{peb-Gal4,UAS-Dcr-2/+;UAS-pdm3^{JF02312}/CyO}$ (RNAi#1) and $\text{peb-Gal4,UAS-Dcr-2/+;UAS-pdm3^{HMJ21205}}$/CyO (RNAi#2)). Scale bar = 10 µm.

(D) Quantification of neuron number for the indicated Ir neuron populations for the genotypes shown in (C). In this and other panels, boxplots show the median, first and third quartile of the data, overlaid with individual data points. Comparisons to controls are shown for each neuron type (pairwise Wilcoxon rank-sum two-tailed test and $P$ values adjusted for multiple comparisons with Bonferroni method, ** $P < 0.001$; NS $P > 0.05$). Sample sizes are shown below the below. Ir40a and Ir64a OSN population sizes could not be quantified because the sections visualized necessarily contain a variable number of neurons; similarly, we could not confidently count Ir84a OSN number because of the weak signal. Nevertheless, when visualized blindly, the control and RNAi genotypes were not distinguishable for any of these neuron populations (assessing the phenotype in antennae of at least 10 animals from two independent genetic crosses).

Figure 6. Pdm3 is required to distinguish Ir75a and Ir75b neuron fate

(A) Immunofluorescence for IR75a and IR75b on whole-mount antennae of control ($\text{peb-Gal4,UAS-Dcr-2/+}$) and $pdm3^{\text{RNAi#2}}$ ($\text{peb-Gal4,UAS-Dcr-2/+;UAS-pdm3^{HMJ21205}}$/+) animals. Scale bar = 10 µm. The schematics on the right summarize the distribution of labeled neurons.
(B) Immunofluorescence for GFP and IR75a on whole-mount antennae of control 
(peb-Gal4, UAS-Dcr-2/+; Ir75a-CD4:tdGFP/+ and pdm3RNAi#2 (peb-Gal4, UAS-Dcr-
2/+; Ir75a-CD4:tdGFP/UAS-pdm3HMJ21205) animals. Scale bar = 10 µm.

(C) Immunofluorescence for GFP and IR75b on whole-mount antennae of control 
(peb-Gal4, UAS-Dcr-2/+; Ir75b-CD4:tdGFP/+ and pdm3RNAi#2 (peb-Gal4, UAS-
Dcr2Dcr-2/+; Ir75b-CD4:tdGFP/UAS-pdm3HMJ21205) animals. Scale bar = 10 µm.

(D) Quantifications of the number of neurons that express IR75a-GFP or IR75b-
GFP in the genotypes shown in (B-C). Comparisons to the controls are shown 
(pairwise Wilcoxon rank-sum two-tailed test and $P$ values adjusted for multiple 
comparisons with Bonferroni method, ** $P < 0.001$). The increase in number of 
Ir75b-GFP labelled neurons in pdm3RNAi is lower than the increase in IR75b-
expressing neurons, potentially because the transgenic reporter is not fully faithful 
in this genetic background.

(E) Antennal lobe projections of Ir75a and Ir75b neurons in control and in 
pdm3RNAi animals. Genotypes are as in (B). Scale bar = 20 µm. Quantifications of 
projection pattern phenotypes are shown on the right.

(F) Quantification of numbers of Ir75a and Ir75b neurons in control (Ir8a-Gal4+) 
and pdm3RNAi#2 (Ir8a-Gal4/UAS-pdm3HMJ21205) animals. Comparisons to the 
control are shown for each neuron type (pairwise Wilcoxon rank-sum two-tailed 
test, ** $P < 0.001$, NS $P > 0.05$).

(G) Immunofluorescence for IR75a and IR75b in whole-mount antennae of control 
and pdm3RNAi#2 animals. Single optical sections are shown, to reveal the weak co-
expression of IR75a and IR75b in a subset of cells (arrowheads). Scale bar = 10 
µm.

(H) Quantifications of number of neurons that co-express IR75a and IR75b. 
Comparison to the control is shown (pairwise Wilcoxon rank-sum two-tailed test).

Figure 7. Expression analysis of Fmi in OSNs

(A) Immunofluorescence for Fmi and the ubiquitously-expressed neuronal 
cadherin (CadN) on whole mount antennal lobes of wild-type (w1118) animals of 
the indicated age. Three optical sections are shown to reveal the morphology of 
most glomeruli. Bottom row: Fmi-positive glomeruli were identified in adult 
antennal lobes based upon stereotyped position and morphology. Dorsal-ventral 
(D-V) and lateral-medial (L-M) axes are indicated. Scale bar = 20 µm.

(B) Histogram of frequency of detectable Fmi immunoreactivity in glomeruli of the 
adult antennal lobe (n = 14 antennal lobes).

(C) Immunofluorescence for Fmi and CadN on control (1081-Gal4/++;449-
Gal4/UAS-Dcr-2) and LN>fmiRNAi1 (1081-Gal4/UAS-fmiKK100512;449-Gal4/UAS-
Dcr-2) animals. Open and filled arrowheads highlight adjacent glomeruli with low 
and high levels of Fmi staining. Scale bar = 20 µm.

Quantifications of wild-type or defective antennal lobe glomerular architectures 
are shown on the right; the n for each phenotypic category are indicated on the 
plots.

Figure 8. Fmi is required in OSNs for glomerular segregation

(A) Immunofluorescence for nc82 on whole mount antennal lobes of control 
(w;ey-Flp/++;act>stop>Gal4/+) and fmiRNAi1 (w;ey-Flp/UAS-
fmiKK100512;act>stop>Gal4/+) and fmiRNAi2 (w;ey-Flp/++;act>stop>Gal4/UAS-
fmiIF02047) animals. Three optical sections are shown. Scale bar = 20 µm.

Quantifications of wild-type or defective antennal lobe glomerular architectures 
are shown on the right; the n for each phenotypic category are indicated on the 
plots.
(B) Immunofluorescence for nc82 and GFP on whole mount antennal lobes of control and fmiRNAi#1 animals. Genotypes: Ir75a-GFP (w;ey-Flp, Ir75a-GFP/+ or UAS-fmiKK100512;act>stop>Gal4/+), Ir75b-GFP (w;ey-Flp, Ir75b-GFP/+ or UAS-fmiKK100512;act>stop>Gal4/+), Ir75c-GFP (w;ey-Flp, Ir75c-GFP/+ or UAS-fmiKK100512;act>stop>Gal4/+). Scale bar = 20 µm.

(C) Immunofluorescence for nc82, GFP, RFP (detects Tomato (Tom)) on whole mount antennal lobes of control and fmiRNAi#1 animals. Genotypes: w;Ir75a-GFP, Ir75b-RFP (w;ey-Flp, Ir75a-GFP/+ or UAS-fmiKK100512;act>stop>Gal4/Ir75b-RFP (top panels) and eyFlp,act>stop>Gal4/Ir75b-RFP (bottom panels). Scale bar = 20 µm.

(D) Immunofluorescence for CadN on whole mount antennal lobes of control (peb-Gal4), OSN>fmRNAi#1 (peb-Gal4/+;UAS-fmiKK100512/+;UAS-Dcr-2/+), LN>fmRNAi#1 (1081-Gal4/UAS-fmiKK100512;449-Gal4/UAS-Dcr-2), PN>fmRNAi#1 (GH146-Gal4/+;UAS-fmiKK100512/+ ) animals. Scale bar = 20 µm.

(E) Immunofluorescence for nc82 and GFP on whole-mount antennal lobes of peb-Gal4, UAS-Dcr-2/+;Ir75b-GFP/tub-Gal80ts;UAS-fmiJP02047/+ animals in which Gal4-driven RNAi was suppressed until shifting from the permissive to restrictive temperature (19°C > 29°C) for Gal80ts at the indicated time points. Full confocal-stacks are shown. Scale bar = 20 µm.

(F) Immunofluorescence for GFP, Fmi and CadN on whole mount pupal antennal lobes (24 h APF) of control (peb-Gal4/+;UAS-mCD8::GFP/+ ) and OSN>fmRNAi#1 (peb-Gal4/+;UAS-fmiKK100512/+;UAS-mCD8::GFP/+ ) animals. The merged channels are shown on the right. The dorsolateral (DL) and ventromedial (VM) bundles of OSN axons are indicated. Scale bar = 20 µm.
References

1. Yan H, Jafari S, Pask G, Zhou X, Reinberg D, Desplan C. Evolution, developmental expression and function of odorant receptors in insects. *J Exp Biol.* 223, jeb208215 (2020).

2. Barish S, Volkan PC. Mechanisms of olfactory receptor neuron specification in *Drosophila*. *Wiley interdisciplinary reviews. Developmental biology.* 4, 609-621 (2015).

3. Jefferis GS, Hummel T. Wiring specificity in the olfactory system. *Semin Cell Dev Biol.* 17, 50-65 (2006).

4. Benton R. Multigene Family Evolution: Perspectives from Insect Chemoreceptors. *Trends Ecol Evol.* 30, 590-600 (2015).

5. Rytz R, Croset V, Benton R. Ionotropic Receptors (IRs): Chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem Mol Biol.* 43, 888-897 (2013).

6. Robertson HM. Molecular Evolution of the Major Arthropod Chemoreceptor Gene Families. *Annu. Rev. Entomol.* 64, 227-242 (2019).

7. Gomez-Diaz C, Martin F, Garcia-Fernandez JM, Alcorta E. The Two Main Olfactory Receptor Families in *Drosophila*, ORs and IRs: A Comparative Approach. *Front Cell Neurosci.* 12, 253 (2018).

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

9. Hong W, Luo L. Genetic control of wiring specificity in the fly olfactory system. *Genetics.* 196, 17-29 (2014).

10. Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc.* 13, 599-604 (2018).

11. Shanbhag SR, Muller B, Steinbrecht RA. Atlas of olfactory organs of *Drosophila melanogaster*. 1. Types, external organization, innervation and distribution of olfactory sensilla. *Int J Insect Morphol Embryol.* 28, 377-397 (1999).

12. Shanbhag SR, Muller B, Steinbrecht RA. Atlas of olfactory organs of *Drosophila melanogaster*. 2. Internal organization and cellular architecture of olfactory sensilla. *Arthropod Struct Dev.* 29, 211-229 (2000).

13. Li H, *et al.* Single-Cell Transcriptomes Reveal Diverse Regulatory Strategies for Olfactory Receptor Expression and Axon Targeting. *Curr Biol.* 30, 1189-1198 e1185 (2020).
14. Southall TD, et al. Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II occupancy in neural stem cells. *Dev Cell.* **26**, 101-112 (2013).

15. Aughey GN, Estacio Gomez A, Thomson J, Yin H, Southall TD. CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. *Elife.* **7**, e32341 (2018).

16. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* **10**, 1213-1218 (2013).

17. Doupe DP, Marshall OJ, Dayton H, Brand AH, Perrimon N. *Drosophila* intestinal stem and progenitor cells are major sources and regulators of homeostatic niche signals. *Proceedings of the National Academy of Sciences of the United States of America.* **115**, 12218-12223 (2018).

18. Widmer YF, Bilican A, Bruggmann R, Sprecher SG. Regulators of Long-Term Memory Revealed by Mushroom Body-Specific Gene Expression Profiling in *Drosophila melanogaster.* *Genetics.* **209**, 1167-1181 (2018).

19. Tamirisa S, et al. Decoding the Regulatory Logic of the *Drosophila* Male Stem Cell System. *Cell Rep.* **24**, 3072-3086 (2018).

20. Ai M, et al. Acid sensing by the *Drosophila* olfactory system. *Nature.* **468**, 691-695 (2010).

21. Grosjean Y, et al. An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. *Nature.* **478**, 236-240 (2011).

22. Silbering AF, et al. Complementary Function and Integrated Wiring of the Evolutionarily Distinct *Drosophila* Olfactory Subsystems. *The Journal of Neuroscience.* **31**, 13357-13375 (2011).

23. Knecht ZA, et al. Distinct combinations of variant ionotropic glutamate receptors mediate thermosensation and hygrosensation in *Drosophila*. *Elife.* **5**, e17879 (2016).

24. Enjin A, et al. Humidity Sensing in *Drosophila*. *Curr Biol.* **26**, 1352-1358 (2016).

25. Prieto-Godino LL, et al. Evolution of acid-sensing olfactory circuits in drosophilids. *Neuron.* **93**, 661-676 e666 (2017).

26. Joseph RM, Carlson JR. *Drosophila* chemoreceptors: a molecular interface between the chemical world and the brain. *Trends Genet.* **31**, 683-695 (2015).
Sanchez-Alcaniz JA, et al. An expression atlas of variant ionotropic glutamate receptors identifies a molecular basis of carbonation sensing. *Nat Commun.* **9**, 4252 (2018).

Chen YD, Dahanukar A. Recent advances in the genetic basis of taste detection in *Drosophila*. *Cell Mol Life Sci.* **77**, 1087-1101 (2020).

Abuin L, Bargeton B, Ulbrich MH, Isacoff EY, Kellenberger S, Benton R. Functional architecture of olfactory ionotropic glutamate receptors. *Neuron.* **69**, 44-60 (2011).

Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell.* **136**, 149-162 (2009).

Menuz K, Larter NK, Park J, Carlson JR. An RNA-Seq Screen of the *Drosophila* Antenna Identifies a Transporter Necessary for Ammonia Detection. *PLOS Genet.* **10**, e1004810 (2014).

Budelli G, Sun Q, Ferreira J, Butler A, Santi CM, Salkoff L. SLO2 Channels Are Inhibited by All Divalent Cations That Activate SLO1 K+ Channels. *J. Biol. Chem.* **291**, 7347-7356 (2016).

Hummel T, Krukkert K, Roos J, Davis G, Klambt C. *Drosophila* Futsch/22C10 is a MAP1B-like protein required for dendritic and axonal development. *Neuron.* **26**, 357-370 (2000).

Shiao MS, et al. Expression Divergence of Chemosensory Genes between *Drosophila sechellia* and Its Sibling Species and Its Implications for Host Shift. *Genome Biol. Evol.* **7**, 2843-2858 (2015).

Sweeney LB, et al. Temporal target restriction of olfactory receptor neurons by Semaphorin-1a/PlexinA-mediated axon-axon interactions. *Neuron.* **53**, 185-200 (2007).

Hong W, et al. Leucine-rich repeat transmembrane proteins instruct discrete dendrite targeting in an olfactory map. *Nat. Neurosci.* **12**, 1542-1550 (2009).

Larter NK, Sun JS, Carlson JR. Organization and function of *Drosophila* odorant binding proteins. *Elife.* **5**, e20242 (2016).

Sun JS, Larter NK, Chahda JS, Rioux D, Gumaste A, Carlson JR. Humidity response depends on the small soluble protein Obp59a in *Drosophila*. *Elife.* **7**, e39249 (2018).

Chung YD, Zhu J, Han Y, Kernan MJ. *nompA* encodes a PNS-specific, ZP domain protein required to connect mechanosensory dendrites to sensory structures. *Neuron.* **29**, 415-428 (2001).
40. Kazama H, Wilson RI. Homeostatic matching and nonlinear amplification at identified central synapses. *Neuron*. **58**, 401-413 (2008).

41. Root CM, *et al.* A presynaptic gain control mechanism fine-tunes olfactory behavior. *Neuron*. **59**, 311-321 (2008).

42. Carlsson MA, Diesner M, Schachtner J, Nassel DR. Multiple neuropeptides in the *Drosophila* antennal lobe suggest complex modulatory circuits. *J Comp Neurol.* **518**, 3359-3380 (2010).

43. Root CM, Ko KI, Jafari A, Wang JW. Presynaptic facilitation by neuropeptide signaling mediates odor-driven food search. *Cell*. **145**, 133-144 (2011).

44. Pei J, Kinch LN, Grishin NV. FlyXCDB-A Resource for *Drosophila* Cell Surface and Secreted Proteins and Their Extracellular Domains. *J. Mol. Biol.* **430**, 3353-3411 (2018).

45. Jafari S, Alkhori L, Schleiffer A, Brochtrup A, Hummel T, Alenius M. Combinatorial activation and repression by seven transcription factors specify *Drosophila* odorant receptor expression. *PLOS Biology*. **10**, e1001280 (2012).

46. Pfreundt U, James DP, Tweedie S, Wilson D, Teichmann SA, Adryan B. FlyTF: improved annotation and enhanced functionality of the *Drosophila* transcription factor database. *Nucleic Acids Res.* **38**, D443-447 (2010).

47. Clyne PJ, Certel SJ, de Bruyne M, Zaslavsky L, Johnson WA, Carlson JR. The odor specificities of a subset of olfactory receptor neurons are governed by Acj6, a POU-domain transcription factor. *Neuron*. **22**, 339-347 (1999).

48. Song E, de Bivort B, Dan C, Kunes S. Determinants of the *Drosophila* odorant receptor pattern. *Developmental Cell*. **22**, 363-376 (2012).

49. Chou YH, Zheng X, Beachy PA, Luo L. Patternning axon targeting of olfactory receptor neurons by coupled Hedgehog signaling at two distinct steps. *Cell*. **142**, 954-966 (2010).

50. Brown JB, *et al.* Diversity and dynamics of the *Drosophila* transcriptome. *Nature*. **512**, 393-399 (2014).

51. Chen CK, Chen WY, Chien CT. The POU-domain protein Pdm3 regulates axonal targeting of R neurons in the *Drosophila* ellipsoid body. *Dev Neurobiol.* **72**, 1422-1432 (2012).

52. Tichy AL, Ray A, Carlson JR. A new *Drosophila* POU gene, *pdm3*, acts in odor receptor expression and axon targeting of olfactory neurons. *J Neurosci*. **28**, 7121-7129 (2008).
53. Schwabe T, Neuert H, Clandinin TR. A network of cadherin-mediated interactions polarizes growth cones to determine targeting specificity. *Cell.* 154, 351-364 (2013).

54. Chen PL, Clandinin TR. The cadherin Flamingo mediates level-dependent interactions that guide photoreceptor target choice in *Drosophila.* *Neuron.* 58, 26-33 (2008).

55. Gao FB, Kohwi M, Brenman JE, Jan LY, Jan YN. Control of dendritic field formation in *Drosophila*: the roles of flamingo and competition between homologous neurons. *Neuron.* 28, 91-101 (2000).

56. Lee RC, Clandinin TR, Lee CH, Chen PL, Meinertzhagen IA, Zipursky SL. The protocadherin Flamingo is required for axon target selection in the *Drosophila* visual system. *Nat. Neurosci.* 6, 557-563 (2003).

57. Matsubara D, Horiuchi SY, Shimono K, Usui T, Uemura T. The seven-pass transmembrane cadherin Flamingo controls dendritic self-avoidance via its binding to a LIM domain protein, Espinas, in *Drosophila* sensory neurons. *Genes Dev.* 25, 1982-1996 (2011).

58. Berger-Muller S, Suzuki T. Seven-pass transmembrane cadherins: roles and emerging mechanisms in axonal and dendritic patterning. *Mol Neurobiol.* 44, 313-320 (2011).

59. Chou YH, Spletter ML, Yaksi E, Leong JC, Wilson RI, Luo L. Diversity and wiring variability of olfactory local interneurons in the *Drosophila* antennal lobe. *Nat. Neurosci.* 13, 439-449 (2010).

60. Liou NF, et al. Diverse populations of local interneurons integrate into the *Drosophila* adult olfactory circuit. *Nat Commun.* 9, 2232 (2018).

61. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell.* 165, 535-550 (2016).

62. Chai PC, Cruchet S, Wigger L, Benton R. Sensory neuron lineage mapping and manipulation in the *Drosophila* olfactory system. *Nat Commun.* 10, 643 (2019).

63. Joo WJ, Sweeney LB, Liang L, Luo L. Linking Cell Fate, Trajectory Choice, and Target Selection: Genetic Analysis of Sema-2b in Olfactory Axon Targeting. *Neuron.* 78, 673-686 (2013).

64. Takeichi M. The cadherin superfamily in neuronal connections and interactions. *Nat Rev Neurosci.* 8, 11-20 (2007).

65. Keeler AB, Molumby MJ, Weiner JA. Protocadherins branch out: Multiple roles in dendrite development. *Cell adhesion & migration.* 9, 214-226 (2015).
66. Usui T, et al. Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell.* **98**, 585-595 (1999).

67. Organisti C, Hein I, Grunwald Kadow IC, Suzuki T. Flamingo, a seven-pass transmembrane cadherin, cooperates with Netrin/Frazzled in *Drosophila* midline guidance. *Genes Cells.* **20**, 50-67 (2015).

68. Hakeda-Suzuki S, et al. Golden Goal collaborates with Flamingo in conferring synaptic-layer specificity in the visual system. *Nat. Neurosci.* **14**, 314-323 (2011).

69. Wang Y, Wang H, Li X, Li Y. Epithelial microRNA-9a regulates dendrite growth through Fmi-Gq signaling in Drosophila sensory neurons. *Dev Neurobiol.* **76**, 225-237 (2016).

70. Davie K, et al. A Single-Cell Transcriptome Atlas of the Aging *Drosophila* Brain. *Cell.* **174**, 982-998 e920 (2018).

71. Croset V, Treiber CD, Waddell S. Cellular diversity in the *Drosophila* midbrain revealed by single-cell transcriptomics. *Elife.* **7**, e34550 (2018).

72. Li H, et al. Classifying *Drosophila* Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell.* **171**, 1206-1220 e1222 (2017).

73. Tosches MA, Yamawaki TM, Naumann RK, Jacobi AA, Tushev G, Laurent G. Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science.* **360**, 881-888 (2018).

74. Zeisel A, et al. Molecular Architecture of the Mouse Nervous System. *Cell.* **174**, 999-1014 e1022 (2018).

75. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16**, 133-145 (2015).

76. van den Brink SC, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods.* **14**, 935-936 (2017).

77. Middleton SA, Eberwine J, Kim J. Comprehensive catalog of dendritically localized mRNA isoforms from sub-cellular sequencing of single mouse neurons. *BMC Biol.* **17**, 5 (2019).

78. Croset V, Schleyer M, Arguello JR, Gerber B, Benton R. A molecular and neuronal basis for amino acid sensing in the *Drosophila* larva. *Sci Rep.* **6**, 34871 (2016).

79. Lee Y, Poudel S, Kim Y, Thakur D, Montell C. Calcium Taste Avoidance in *Drosophila*. *Neuron.* **97**, 67-74 e64 (2018).
80. Ramdya P, Benton R. Evolving olfactory systems on the fly. *Trends Genet.* **26**, 307-316 (2010).

81. Ludwig CH, Bintu L. Mapping chromatin modifications at the single cell level. *Development.* **146**, dev.170217 (2019).

82. van den Ameele J, Krautz R, Brand AH. TaDa! Analysing cell type-specific chromatin in vivo with Targeted DamID. *Curr. Opin. Neurobiol.* **56**, 160-166 (2019).

83. Haines JE, Eisen MB. Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo. *PLOS Genet.* **14**, e1007367 (2018).

84. Hummel T, Zipursky SL. Afferent induction of olfactory glomeruli requires N-cadherin. *Neuron.* **42**, 77-88 (2004).

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

86. Han C, Jan LY, Jan YN. Enhancer-driven membrane markers for analysis of nonautonomous mechanisms reveal neuron-glia interactions in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America.* **108**, 9673-9678 (2011).

87. Marshall OJ, Southall TD, Cheetham SW, Brand AH. Cell-type-specific profiling of protein-DNA interactions without cell isolation using targeted DamID with next-generation sequencing. *Nat Protoc.* **11**, 1586-1598 (2016).

88. Marshall OJ, Brand AH. damidseq_pipeline: an automated pipeline for processing DamID sequencing datasets. *Bioinformatics.* **31**, 3371-3373 (2015).

89. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* **9**, 357-359 (2012).

90. Li H, *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**, 2078-2079 (2009).

91. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**, 841-842 (2010).

92. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

93. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology.* **16**, 284-287 (2012).
94. Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 25, 25-29 (2000).

95. Saina M, Benton R. Visualizing olfactory receptor expression and localization in *Drosophila*. *Methods in Molecular Biology.* 1003, 211-228 (2013).

96. Ostrovsky A, Cachero S, Jefferis G. Clonal analysis of olfaction in *Drosophila*: immunochemistry and imaging of fly brains. *Cold Spring Harb Protoc.* 2013, 342-346 (2013).

97. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 9, 676-682 (2012).
Figure 1

B. **Ir64a > Dam:PolII**

C. **Ir64a > Dam + CD8:GFP**

D. **Ir64a neuron triplicates**

E. **Ir64a neuron population**

F. **Query genes**

G. **Neuron population**
A. Bar chart showing the number of genes with significant OS for different neuron populations.

B. Hierarchical clustering of OSs, with distance between clusters in a.u.

C. Heatmap of neuron population showing the distribution of CSSPs (n = 2509).

D. Heatmap of neuron population showing the distribution of transcription factors (n = 754).
Figure 3

Ir75c neurons and Ir40a neurons showing chromatin accessibility (coverage) and chromatin accessibility (fold change) for neuron population and distance along chromosome (Mb).

Ir8a neurons and Ir64a neurons showing chromatin accessibility (coverage) for neuron population and distance along chromosome (Mb).
Figure 4

A. Dam:PolII replicates
- Dam replicates

B. Comparison dataset
- Ir31a
- Ir40a
- Ir64a
- Ir75a
- Ir75b
- Ir75c
- Ir84a

C. Molecular Function
- signal transducer activity
- signaling receptor activity
- molecular transducer activity
- phosphotransferase activity, alcohol group as acceptor
- protein kinase activity
- calcium ion binding
- neurotransmitter receptor activity
- kinase activity
- cytoskeletal protein binding
- protein serine/threonine kinase activity

D. Gene X
- no. of GATC motifs with differential read-depth
- pdm3 (67)
- ir40a (28)
- fmi (28)

E. Neuron population
- pdm3
- ir31a
- ir40a
- ir64a
- ir75a
- ir75b
- ir75c
- ir84a

OS
### Figure 5

**A**

| Gene | Control | RNAi#1 | RNAi#2 |
|------|---------|--------|--------|
| Ir31a | ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) |
| Ir40a | ![Image](image4.png) | ![Image](image5.png) | ![Image](image6.png) |
| Ir64a | ![Image](image7.png) | ![Image](image8.png) | ![Image](image9.png) |
| Ir75a | ![Image](image10.png) | ![Image](image11.png) | ![Image](image12.png) |
| Ir75b | ![Image](image13.png) | ![Image](image14.png) | ![Image](image15.png) |
| Ir75c | ![Image](image16.png) | ![Image](image17.png) | ![Image](image18.png) |
| Ir84a | ![Image](image19.png) | ![Image](image20.png) | ![Image](image21.png) |

**B**

Mean fluorescence (a.u.)

**C**

| Gene | Control | peb-Gal4 | pdm3 |
|------|---------|----------|------|
| Ir31a RNA | ![Image](image22.png) | ![Image](image23.png) | ![Image](image24.png) |
| IR40a | ![Image](image25.png) | ![Image](image26.png) | ![Image](image27.png) |
| IR64a | ![Image](image28.png) | ![Image](image29.png) | ![Image](image30.png) |
| IR75a | ![Image](image31.png) | ![Image](image32.png) | ![Image](image33.png) |
| IR75b | ![Image](image34.png) | ![Image](image35.png) | ![Image](image36.png) |
| IR75c | ![Image](image37.png) | ![Image](image38.png) | ![Image](image39.png) |
| IR84a RNA | ![Image](image40.png) | ![Image](image41.png) | ![Image](image42.png) |

**D**

Number of OSMs

*NS* indicates no significant difference.
Figure 6

A. IR75a and IR75b expression in wild-type and defective conditions. Brightfield images show the expression patterns of IR75a and IR75b-GFP.

B. Comparison of wild-type and defective conditions for IR75a-GFP expression.

C. Comparison of wild-type and defective conditions for IR75b-GFP expression.

D. Bar graph showing the number of OSNs (olfactory sensory neurons) in control and RNAi#2 conditions.

E. nc82 and peb-Gal4 expression in wild-type and defective conditions. Images show the expression patterns of nc82 and peb-Gal4-GFP.

F. Comparison of wild-type and defective conditions for IR75a and IR75b expression.

G. Comparison of wild-type and defective conditions for Ir8a-Gal4 expression.

H. Merge images showing the expression of IR75a and IR75b-GFP in wild-type and defective conditions.
Figure 7

A

pupal antennal lobe (h APF)

24

48

62

81

adult Fmi only

Fmi

CadN

anterior

central

posterior

B

proportion Fmi-positive in adults (%)

C

pupal antennal lobe (48 h APF)

control pupal antennal lobe (48 h APF)

LN > mfn>nu

pupal antennal lobe (h APF)
Figure 8

A. Anterior, central, and posterior sections of the developing eye imaginal disc stained with a control antibody (top) and the constitutive-Gal4 driver (bottom). nc82 marks the anterior compartment.

B. Ir75a-GFP, Ir75b-GFP, and Ir75c-GFP expression in the eye imaginal disc. nc82 marks the anterior compartment.

C. Ir75a-GFP and Ir75b-Tom expression in the eye imaginal disc. nc82 marks the anterior compartment.

D. CadN expression in the eye imaginal disc. nc82 marks the anterior compartment.

E. Ir75b-GFP expression in the eye imaginal disc. nc82 marks the anterior compartment.

F. pbe-GFP, Fmi, and CadN expression in the eye imaginal disc. nc82 marks the anterior compartment.
Targeted molecular profiling of rare cell populations identifies olfactory sensory neuron fate and wiring determinants

J. Roman Arguello, Liliane Abuin, Jan Armida, Kaan Mika, Phing Chian Chai and Richard Benton

Supplementary Information

Supplementary Figure Legends

Figure S1. Chemosensory receptor gene Occupancy Scores across Ir populations
Heatmap of OSs for members of Ir, Or and Gr gene families across the seven Ir neuron populations. Information on individual genes displaying significant OSs in one or more populations is shown with colored bars to the left (key is at the bottom). See also Table S1.

Figure S2. Heterogeneous expression of Engrailed in Ir neuron populations
(A) OS heatmap of en in the seven Ir neuron populations.
(B) Immunofluorescence for En (magenta) and GFP (green) on antennal sections of animals in which the indicated Ir neuron populations are labeled with a CD8:GFP reporter. Genotypes are of the form: Irxx-Gal4/+;UAS-mCD8:GFP/+ or, for Ir75a and Ir75c neurons, UAS-mCD8:GFP/+;Irxx-Gal4/+.
Scale bar = 10 µm.
(C) Density plots for En immunofluorescence signals (arbitrary units, a.u.) quantified from Ir neuron nuclei (as in Figure 5B). Most signals from Ir75a, Ir75b, Ir75c and Ir84a neuron nuclei fall within the range of background signals; Ir31a, Ir40a and Ir64a neuron nuclei display shifted distributions.

Figure S3. Characterization of Fmi expression in antennae and local interneurons
(A) Immunofluorescence for Fmi and Elav on whole-mount antennae of control animals (w;Ir8a-Gal4,UAS-RedStinger/CyO; note the RedStinger reporter is not shown in these images) at 24 and 48 h APF. Partial optical stacks are shown. The arrowhead points to the antennal nerve. Scale bar = 20 µm.
(B) Immunofluorescence for Fmi and Elav on whole-mount antennal lobes of 48 h APF control (w^{1186}) animals. A partial optical stack is shown, revealing expression of Elav in the Fmi-positive soma adjacent to the antennal lobe. Scale bar = 20 µm.
(C) Immunofluorescence for Fmi, GFP and CadN on a whole-mount antennal lobe of a 48 h APF animal in which Fmi-expressing LNs are labeled with the combined 49-Gal4 and 1081-Gal4 drivers (1081-Gal4/+;449-Gal4/UAS-mCD8:GFP), which were identified by screening a panel of LN drivers. A partial optical stack is shown. We distinguish two types of Fmi-positive neurons: 100% of strongly-expressing neurons (n = 194, 7 brains) are GFP-positive; 75% of weakly-expressing neurons (e.g., the neuron indicated by the arrowhead) are GFP-positive (n = 213 neurons, 7 brains). Note the drivers are expressed in several additional Fmi-negative cells. Scale bar = 20 µm.
(D) Immunofluorescence for Fmi and CadN on whole-mount antennal lobes of 48 h APF animals with OSN>fmiRNAi#1 (peb-Gal4, UAS-Dcr-2/+; UAS-fmiKK100512/+) or pan-neuronal>fmiRNAi#1 (elav-Gal4, UAS-Dcr-2/+; UAS-fmiKK100512/+). In OSN>fmiRNAi#1 animals the remaining, relatively homogeneous glomerular signal is derived from its expression in LN (soma labeled with an arrowhead); we cannot rule out that the homogeneity is a secondary consequence of the glomerular segregation defects. Pan-neuronal>fmiRNAi#1 is semi-lethal, but we recovered a few adult escapers in which all Fmi immunoreactivity in the antennal lobe is lost. Scale bar = 20 µm.

Figure S4. Fmi does not act with known genetic and physical interaction partners in OSNs
Representative images of nc82 immunofluorescence on whole-mount antennal lobes (three optical sections) of animals of the following genotypes (top-to-bottom; other loss-of-function alleles or RNAi lines tested, but not shown, for certain genes are also listed): control (peb-Gal4, UAS-Dcr-2/+) , frizzled+/- (fzP21/fzKD4A; also tested fzP21/fzH51) , dishevelled+/- (dsh1) , frazzledRNAi (peb-Gal4, UAS-Dcr-2; Ir75b-GFP/+; fraJF01457/++; also tested fraHMS01147 and fraJF01231) , golden goalRNAi (peb-Gal4, UAS-Dcr-2; Ir75b-GFP/gogoGD3616, also tested gogoHMC0537, espinas+/- (esnKO5) , GαqRNAi (peb-Gal4, UAS-Dcr-2/++; Gαq35RNAI.UAS.1f1/+; also tested GαqGL01048 and GαqJF01209). The Ir75b-GFP reporter, incorporated into some of these genotypes, is not shown in these images, but was indistinguishable from controls. Scale bar = 20 µm.

Supplementary Tables
Table S1. Chemosensory receptor genes with significant OSs in Ir neuron populations.
Only significant OSs are listed. Values correspond to those plotted in Figure S1.
Table S2. Set of CSSP genes.
The gene list is from the D. melanogaster extracellular domain database (FlyXCDB) database2, with additional gene symbols appended.
Table S3. CSSP genes with significant OSs in Ir neuron populations.
OSs correspond to those plotted in Figure 2C.
Table S4. CSSP genes with significant OSs unique to each population.
Table S5 Filtered set of transcription factors genes.
The gene list was extracted from the FlyTF.org database3.
Table S6. Transcription factor genes with significant OSs in Ir neuron populations
OSs correspond to those plotted in Figure 2D.
Table S7. Transcription factor genes with significant OSs unique to each population.
Table S8. Number of GATC motifs within candidate differentially-expressed genes contributing to significant differences.
Values corresponds to those plotted in Figure 4D.

Table S9. *D. melanogaster* stocks

| Stock             | Source                  |
|-------------------|-------------------------|
| wild-type         | RRID:BDSC_41731         |
| Ir8a-Gal4         | RRID:BDSC_41726         |
| Ir31a-Gal4        | RRID:BDSC_41727         |
| Ir40a-Gal4        | RRID:BDSC_41732         |
| Ir64a-Gal4        | RRID:BDSC_41748         |
| Ir75a-Gal4        | RRID:BDSC_41748         |
| Ir75b-Gal4        | RRID:BDSC_41748         |
| Ir75c-Gal4        | RRID:BDSC_41748         |
| Ir84a-Gal4        | RRID:BDSC_41734         |
| peb-Gal4          | RRID:BDSC_80570         |
| ey-Flip (II)      | RRID:BDSC_5576          |
| ey-Flip (III)     | RRID:BDSC_5577          |
| act>stop>Gal4 (II)| RRID:BDSC_3953          |
| act>stop>Gal4 (III)| RRID:BDSC_4780          |
| GH146-Gal4        | RRID:BDSC_30026         |
| 449-Gal4          | RRID:BDSC_458           |
| 1081-Gal4         | RRID:BDSC_458           |
| elav-Gal4         | RRID:BDSC_458           |
| tub-Gal80         | RRID:BDSC_7018          |
| UAS-Dcr-2 (X)     | RRID:BDSC_24648         |
| UAS-Dcr-2 (II)    | RRID:BDSC_24650         |
| UAS-Dcr-2 (III)   | RRID:BDSC_24651         |
| UAS-LT3-Dam (III) | RRID:BDSC_5130          |
| UAS-LT3-Dam:Rp215 (III)| RRID:BDSC_5130         |
| UAS-mCD8:GFP      | RRID:BDSC_8546          |
| UAS-RedStinger    | RRID:BDSC_8546          |
| Irsb-CD4:tdGFP    | RRID:BDSC_8546          |
| UAS-fmi[KK106512](RNAi #1)| VDRC v107993          |
| UAS-fmi[JF02047](RNAi #2)| RRID:BDSC_26022         |
| UAS-fra[HM30117](RNAi)| RRID:BDSC_40826         |
| UAS-fra[JF01267](RNAi)| RRID:BDSC_31469         |
| UAS-fra[JF01267](RNAi)| RRID:BDSC_31664         |
| UAS-Gaq[GL0108](RNAi)| RRID:BDSC_30735         |
| UAS-Gaq[JF01267](RNAi)| RRID:BDSC_31268         |
| UAS-Gaq[SL01058](RNAi)| RRID:BDSC_36820         |
| UAS-gogo[GU3676](RNAi)| VDRC v43928            |
| Antibody                                | Dilution | Source/Reference         |
|-----------------------------------------|----------|--------------------------|
| Mouse mAb anti-En                       | 1:100    | DSHB 4D9                 |
| Guinea pig anti-Pdm3                    | 1:200    |                          |
| Mouse mAb anti-Fmi                      | 1:20     | DSHB #74                 |
| Guinea pig anti-IR40a                   | 1:200    |                          |
| Rabbit anti-IR64a                       | 1:1000   |                          |
| Rabbit anti-IR75a                       | 1:200    |                          |
| Guinea pig anti-IR75b                   | 1:500    |                          |
| Rabbit anti-IR75c                       | 1:200    |                          |
| Mouse mAb anti-Bruchpilot               | 1:10     | DSHB nc82                |
| Rat anti-Cadherin-N                     | 1:25     | DSHB Ex#8-2              |
| Rat mAb anti-Elav                       | 1:10     | DSHB 7E8A10              |
| Chicken anti-GFP                        | 1:2000   | Abcam 13970              |
| Mouse anti-GFP                          | 1:1000   | Invitrogen A11120        |
| Anti-RFP                                | 1:1000   | Abcam 62341              |
| Alexa488 anti-Guinea pig                | 1:500    | Invitrogen A11073        |
| Alexa488 anti-Chicken                   | 1:1000   | Abcam 150169             |
| Cy3 anti-Rabbit                         | 1:1000   | Milan Analytica AG 111-165-1440 |
| Cy3 anti-Mouse                          | 1:1000   | Milan Analytica AG 115-165-166 |
| Cy5 anti-Rat                            | 1:250    | Jackson ImmunoResearch 712-175-153 |
| anti-DIG-POD                            | 1:300    | Roche Diagnostics 11 207 733 910 |
Supplementary Datasets

Supplementary Data 1. Full set of genes with significant OSs.

Supplementary Data 2. Candidate genes for differential occupancy based on the Likelihood Ratio Test within DESeq2. One file contains only the list of genes, the second file contains the list of GATC fragments and the associated annotations and statistics.

Supplementary Data 3. Candidate genes for pairwise differential occupancy based on the Wald Test with DESeq2. 21 Bed-formatted files (representing all possible pairwise comparisons between the 7 datasets) containing statistics and annotations.

Supplementary Data 4. Read coverage files for GATC fragments used for DESeq2 analyses.
Supplementary References

1. Liou NF, et al. Diverse populations of local interneurons integrate into the Drosophila adult olfactory circuit. Nat Commun. 9, 2232 (2018).

2. Pei J, Kinch LN, Grishin NV. FlyXCDB-A Resource for Drosophila Cell Surface and Secreted Proteins and Their Extracellular Domains. J. Mol. Biol. 430, 3353-3411 (2018).

3. Pfreundt U, James DP, Tweedie S, Wilson D, Teichmann SA, Adryan B. FlyTF: improved annotation and enhanced functionality of the Drosophila transcription factor database. Nucleic Acids Res. 38, D443-447 (2010).

4. Prieto-Godino LL, et al. Evolution of acid-sensing olfactory circuits in drosophilids. Neuron. 93, 661-676 e666 (2017).

5. Southall TD, et al. Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II occupancy in neural stem cells. Dev Cell. 26, 101-112 (2013).

6. Matsubara D, Horiuchi SY, Shimono K, Usui T, Uemura T. The seven-pass transmembrane cadherin Flamingo controls dendritic self-avoidance via its binding to a LIM domain protein, Espinas, in Drosophila sensory neurons. Genes Dev. 25, 1982-1996 (2011).

7. Chen CK, Chen WY, Chien CT. The POU-domain protein Pdm3 regulates axonal targeting of R neurons in the Drosophila ellipsoid body. Dev Neurobiol. 72, 1422-1432 (2012).

8. Silbering AF, et al. Ir40a neurons are not DEET detectors. Nature. 534, E5-7 (2016).

9. Ai M, et al. Acid sensing by the Drosophila olfactory system. Nature. 468, 691-695 (2010).

10. Prieto-Godino LL, et al. Olfactory receptor pseudo-pseudogenes. Nature. 539, 93-97 (2016).
Figure S1

neuron population

Ir31a  Ir40a  Ir64a  Ir75a  Ir75b  Ir75c  Ir84a

Ir

Or

Gr

OS

targeted IRs
IR co-receptors
significant OS
significant OS (overlaps with the exon(s) of another gene)
significant OS (located within an intron of another gene)
Figure S3

A. 24 and 48 h APF pupal antennal lobe

B. 48 h APF pupal antennal lobe

C. 449-Gal4 + 1081-Gal4 > CD8:GFP

D. pan-neuronal (elav) > fmi RNAi#1

Fmi, CD8:GFP, Fmi, CadN, CD8:GFP
Figure S4

| nc82   | anterior | medial | posterior |
|--------|----------|--------|-----------|
| control | ![control](image) | ![control](image) | ![control](image) |
| frizzled<sup>−/−</sup> | ![frizzled<sup>−/−</sup>](image) | ![frizzled<sup>−/−</sup>](image) | ![frizzled<sup>−/−</sup>](image) |
| dishevelled<sup>−/−</sup> | ![dishevelled<sup>−/−</sup>](image) | ![dishevelled<sup>−/−</sup>](image) | ![dishevelled<sup>−/−</sup>](image) |
| frazzled<sup>RNAi</sup> | ![frazzled<sup>RNAi</sup>](image) | ![frazzled<sup>RNAi</sup>](image) | ![frazzled<sup>RNAi</sup>](image) |
| golden goal<sup>RNAi</sup> | ![golden goal<sup>RNAi</sup>](image) | ![golden goal<sup>RNAi</sup>](image) | ![golden goal<sup>RNAi</sup>](image) |
| espinas<sup>−/−</sup> | ![espinas<sup>−/−</sup>](image) | ![espinas<sup>−/−</sup>](image) | ![espinas<sup>−/−</sup>](image) |
| Ga<sub>q</sub><sup>RNAi</sup> | ![Ga<sub>q</sub><sup>RNAi</sup>](image) | ![Ga<sub>q</sub><sup>RNAi</sup>](image) | ![Ga<sub>q</sub><sup>RNAi</sup>](image) |