Olfactory receptor and circuit evolution promote host specialisation

Thomas O. Auer[1,*], **Mohammed A. Khallaf**[2], **Ana F. Silbering**[1], **Giovanna Zappia**[1], **Kaitlyn Ellis**[3], **Raquel Álvarez-Ocaña**[1], **J. Roman Arguello**[4], **Bill S. Hansson**[2], **Gregory S.X.E. Jefferis**[5], **Sophie J.C. Caron**[3], **Markus Knaden**[2], **Richard Benton**[1,*]

[1]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, CH-1015, Lausanne, Switzerland [2]Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, D-07745, Jena, Germany [3]Department of Biology, University of Utah, Salt Lake City, Utah, 84112, USA [4]Department of Ecology and Evolution, Faculty of Biology and Medicine, University of Lausanne, CH-1015, Lausanne, Switzerland [5]Division of Neurobiology, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK

Summary

The evolution of animal behaviour is poorly understood[1,2]. Despite numerous correlations of behavioural and nervous system divergence, demonstration of the genetic basis of interspecific behavioural differences remains rare[3–5]. Here, we develop a novel neurogenetic model, *Drosophila sechellia*, a close cousin of *D. melanogaster*[6,7] that displays profound behavioural changes linked to its extreme specialisation on noni fruit[8–16]. Using calcium imaging, we identify *D. sechellia* olfactory pathways detecting host volatiles. Mutational analysis indicates roles for different olfactory receptors in long- and short-range attraction to noni. Cross-species allele transfer demonstrates that tuning of one of these receptors is important for species-specific host-seeking. We identify the molecular determinants of this functional change, and characterise their evolutionary origin and behavioural significance. Through circuit tracing in the *D. sechellia* brain,

*Corresponding authors: T: ++41 21 692 3932, Richard.Benton@unil.ch, Thomas.Auer@unil.ch.

we find that receptor adaptations are accompanied by increased sensory pooling onto interneurons and novel central projection patterns. This work reveals the accumulation of molecular, physiological and anatomical traits linked to behavioural divergence, and defines a powerful model for investigating nervous system evolution and speciation.

The genetic and neural basis by which animals adapt behaviourally to their ecological niche is largely unknown[1,2]. Insights have been gained from intraspecific variation in traditional model organisms, including anxiety behaviours in *Mus musculus*[17] and exploration/exploitation decisions in *Caenorhabditis elegans*[18]. Interspecific differences are more dramatic: *Peromyscus* mice species display variations in burrowing and parental care[3,5], while the predatory nematode *Pristionchus pacificus* exhibits distinct feeding behaviours to *C. elegans*[19]. Defining the molecular basis of interspecific differences is challenging as it requires that species are comparable molecularly and anatomically, and genetically manipulatable.

Drosophilid flies are attractive models to investigate behavioural evolution: *D. melanogaster* offers deep neurobiological knowledge in a numerically-simple brain, and closely-related drosophilid species show distinct behaviours linked to their diverse ecologies[20]. Several of these behavioural traits have been correlated to anatomical and/or physiological changes in sensory or central pathways[4,11,13,15,21,22]. One remarkable drosophilid is the Seychelles endemic, *D. sechellia*, which shares a recent common ancestor with the cosmopolitan, ecological generalists *D. melanogaster* and *D. simulans*[6,7] (Fig. 1a). *D. sechellia* has evolved extreme specialism for *Morinda citrifolia* "noni" fruit (Fig. 1a), and displays unique olfactory[11–13,15,16], gustatory[14], and reproductive behaviours[8–10]. Mapping approaches have located causal loci for some *D. sechellia*-specific traits within typically large genomic regions[8,10], while candidate approaches have correlated chemosensory phenotypes with changes in its peripheral sensory pathways[11,14,15].

Despite the potential *D. sechellia* presents for comparative neuroscience, investigations of *D. sechellia*'s behaviours have been limited by the lack of genetic tools. Here we develop *D. sechellia* into a genetic model system, moving from genotypic/phenotypic correlations to test the role of genetic changes in behavioural evolution.

Specific noni attraction of *D. sechellia*

Noni-derived volatiles are likely the initial cues that guide *D. sechellia* host-seeking[16]. We used two assays to compare attraction of wild-type strains of *D. sechellia*, *D. simulans* and *D. melanogaster* to noni at distinct spatial scales (Fig. 1b,c, Extended Data Fig. 1): In a long-range, wind tunnel assay[23], *D. sechellia* displayed higher attraction to noni than its sister species (Fig. 1b); in a short-range trap assay[15], only *D. sechellia* exhibits marked noni preference (Fig. 1c). The level of attraction/preference of this species to noni juice (a more reproducible odour stimulus) was comparable to that for ripe fruit (Fig. 1b,c), concordant with their qualitatively similar odour bouquets (Fig. 1d, Extended Data Fig. 2a-b, Methods). Assays with other natural odour sources, as previously described in field studies[24], confirmed the unique attractiveness of noni for *D. sechellia* (Extended Data Fig. 1a-e,g).

Noni-sensing olfactory pathways

Drosophilids detect odours using olfactory sensory neurons (OSNs) in sensilla on their antennae and maxillary palps[25]. Most OSNs express a single Odorant Receptor (OR) or Ionotropic Receptor (IR) – which defines odour-tuning properties – along with an obligate co-receptor[26–28]. Neurons expressing the same tuning receptor converge onto a discrete glomerulus in the antennal lobe[25]. Electrophysiological analyses in *D. sechellia* have identified several OSN populations responding to individual noni odours[11,13,15,29], but the global representation of the noni bouquet has not been examined.

We generated transgenic *D. sechellia* expressing GCaMP6f in the majority of OSNs under the control of *Gal4* inserted at the *Or co-receptor* (*Orco*) locus (Extended Data Fig. 3). Using widefield imaging to compare this and an equivalent *D. melanogaster* line (Extended Data Fig. 4a-b), we did not detect novel noni-responsive olfactory channels in *D. sechellia* but rather quantitative differences in individual glomerular responses between species (Extended Data Fig. 4a). Two-photon calcium imaging highlighted two glomeruli (DM2 and VM5d) distinguished by their very high sensitivity to noni compared to grape juice in *D. sechellia* (Fig. 1e-f; Extended Data Fig. 4c-d). These glomeruli are innervated by OSNs housed in a common antennal basiconic sensillum class (ab3) that responds electrophysiologically to individual noni odours[11,13].

Genetic targeting of olfactory receptors

For noni-sensitive olfactory channels, we determined their electrophysiological responses to noni odours (Fig. 2a-b, Extended Data Fig. 4e, 5-6) and mutated candidate olfactory receptors (Extended Data Fig. 5-6). In wild-type ab3 sensilla, the larger-spiking ab3A neuron responded most strongly to methyl esters while the smaller-spiking ab3B neuron was highly stimulated by *2*-heptanone and *1*-hexanol (Fig. 2a). In *DsecOrco* mutants, all of these responses were lost (Fig. 2a), indicating their dependence upon OR signalling.

D. melanogaster ab3A expresses the *Or22a* and *Or22b* genes[30] while *D. sechellia* possesses only *Or22a*[11]. Targeted mutation of this locus abolished odour-evoked responses of ab3A, but not ab3B (Fig. 2a). The receptor in *D. melanogaster* ab3B is thought to be OR85b[25,31], but *DsecOr85b* mutant neurons retained some sensitivity to noni odours (Fig. 2a). Deletion of *DsecOr85b* and the neighbouring *DsecOr85c*, whose transcripts were detected in an antennal transcriptome[32], led to complete loss of ab3B responses, arguing for partial receptor redundancy (Fig. 2a).

D. sechellia Ir75b neurons (in antennal coeloconic 3I (ac3I) sensilla) have evolved novel sensitivity to hexanoic acid[15]. Mutations in *DsecIr75b* or *DsecIr8a* (which encodes the acid-sensing IR co-receptor[27]) led to selective loss of hexanoic and butyric acid responses in ac3I (Fig. 2b, Extended Data Fig. 6). Mutation of *DsecOr35a* (expressed in the paired neuron) diminished responses to all odours except these acids, in-line with the broad tuning of this receptor in *D. melanogaster*[33].

ORs required for long-range attraction

We used the receptor mutants to determine the behavioural role of individual olfactory pathways. In the long-range assay, *DsecOrco* exhibited no attraction to the odour source (Fig. 2c). Strikingly, both *DsecOr22a* and *DsecOr85c/b* mutants displayed similarly strong defects (Fig. 2c). By contrast, *DsecOr35a* mutants were not impaired (Fig. 2c). Loss of IR8a also led to a significant decrease in long-range attraction in *D. sechellia* (Fig. 2c). This does not appear to be primarily due to defects in the hexanoic acid-sensing pathway, as *DsecIr75b* mutants had either no or milder defects than *DsecIr8a* mutants (Fig. 2c). Mutations in *DsecIr64a*, which is broadly-tuned to acids in *D. melanogaster*[34] and responded to noni in *D. sechellia* (Extended Data Fig. 4f-g), had no effect on this behaviour.

In the short-range assay, *DsecOrco* mutant flies displayed reduced, but not abolished, preference for noni (Fig. 2d). Individual Or pathway mutants had very slight (*DsecOr22a*) or no (*DsecOr85c/b*, *DsecOr35a*) defects in this behaviour (Fig. 2d, Extended Data Fig. 7). Mutants for *DsecIr8a* or *DsecIr75b*, but not *DsecIr64a*, displayed reduced preference, with notable frequent preference reversals in several trials (Fig. 2d). *Orco/Ir8a* double mutants, or antenna-less flies, displayed no noni preference (Fig. 2e, Extended Data Fig. 7c-d), indicating that this short-range behaviour depends upon multiple, partially redundant olfactory inputs. Consistent with these observations, individual noni odours promoted strong preference at short-range, while they triggered no or little flight attraction at long-range (Extended Data Fig. 7f-g)[11,13,15]. The relative contribution of individual channels to these behaviours may be related to their different detection thresholds (Extended Data Fig. 2c) and/or differential diffusion of cognate odours within each assay (Extended Data Fig. 2d-e).

Tuning of OR22a impacts behaviour

Given the crucial role of OR22a and OR85c/b in long-range attraction, we explored the evolution of these pathways. Or85c/b neurons displayed an indistinguishable sensitivity across species to their best agonist, *2*-heptanone. By contrast, both *D. sechellia* and *D. simulans* Or22a/(b) neurons exhibited increased sensitivity to methyl hexanoate compared to *D. melanogaster* (Fig. 3a, Extended Data Fig. 8a)[11,29]. Broader profiling of Or22a/(b) neurons in the *D. melanogaster* species subgroup (Fig. 3b, Extended Data Fig. 8g) revealed that *D. sechellia* was the only species with selective, high sensitivity to methyl esters; others, including *D. simulans*, also responded to ethyl esters (Fig. 3b). This suggests that changes in tuning sensitivity and/or breadth of OR22a – but not OR85c/b – contribute to the differential behaviour of *D. sechellia*.

We next reintroduced *DsecOr22a^wt*, *DsimOr22a^wt* or *DmelOr22a^wt* into the *DsecOr22a* endogenous locus (Extended Data Fig. 8b). *DsecOr22a^wt* and *DmelOr22a^wt* restored electrophysiological response profiles similar to native neuronal responses, indicating that the receptor is key for species-specific neuron tuning (Fig. 3c, upper part). *DsimOr22a^wt* conferred sensitivity to methyl but not ethyl esters (Fig. 3b-c, Extended Data Fig. 8h); genetic analysis in *D. simulans* indicated that detection of ethyl esters by the endogenous Or22a/b neurons depends upon the co-expressed OR22b (Extended Data Fig. 8c-f).

Concordant with their physiological properties, *Dsec*OR22a[wt] and *Dsim*OR22a[wt] but not *Dmel*OR22a[wt], rescued long-range behavioural responses to almost wild-type levels (Fig. 3d). Reciprocally, expression of *Dsec*OR22a[wt] in *Or22a/b* mutant neurons in *D. melanogaster* (Fig. 3e, upper part) conferred higher noni sensitivity and long-range attraction than *Dmel*OR22a[wt] (Fig. 3f, Extended Data Fig. 8i, 9a).

Molecular basis of OR22a tuning changes

We next sought the molecular basis of OR22a's tuning differences. Expression of chimeric versions of *Dsec*OR22a[wt] and *Dmel*OR22a[wt] in *D. melanogaster* Or22a/b neurons (Extended Data Fig. 9c) indicated that high-sensitivity and selectivity for methyl esters were determined by the N-terminal 100 amino acids of *Dsec*OR22a (chimera C) (Extended Data Fig. 9b-c,e). Within these, three positions (I45, I67 and M93) differ between *Dmel*OR22a and its orthologues in species displaying narrowed tuning for methyl esters (Extended Data Fig. 9b). Exchange of these residues (*Dmel*OR22a[I45V,I67M,M93I] (*Dmel*OR22a[triple])) narrowed responsiveness to methyl esters, similar to *Dsec*OR22a[wt] (Fig. 3c,e (lower parts), Extended Data Fig. 9d,f). Individual mutations revealed that *Dmel*OR22a[M93I] most closely recapitulated the higher sensitivity of this receptor to methyl esters over ethyl esters (Fig. 3c,e, (lower parts), Extended Data Fig. 9d-g). Conversely, *Dsec*OR22a[I93M] exhibited broadened sensitivity to both ester classes (Fig. 3c (lower part), Extended Data Fig. 9g).

In the long-range olfactory behaviour assay, expression of *Dmel*OR22a[triple] in *D. sechellia* Or22a neurons restored *D. sechellia*-like attraction to noni, while both *Dmel*OR22a[M93I] and *Dsec*OR22a[I93M] displayed levels of attraction intermediate between those of the wild-type receptor rescues (Fig. 3d). Similarly, expression of *Dmel*OR22a[triple] in *D. melanogaster* conferred noni attraction at equivalent levels to *Dsec*OR22a[wt], while *Dmel*OR22a[M93I] supported intermediate attraction levels (Fig. 3f). These results provide evidence that the molecular differences in OR22a orthologues contribute to species-specific olfactory behaviours.

Sensory representation of Or22a

The functional similarity of OR22a orthologues in *D. sechellia* and *D. simulans* (Fig. 3c-d) indicates that additional changes occurred during speciation of *D. sechellia*. Concordant with ab3 sensilla counts[11,29], *D. sechellia* exhibits a three-fold increase in the number of Or22a neurons (recapitulated in rescue experiments, Extended Data Fig. 10a) and the paired Or85c/b neurons, but not several other neuron classes (Fig. 4a-b, Extended Data Fig. 5d, 10b and [15]).

To analyse OSN projections in *D. sechellia*, we inserted *Gal4* at the corresponding receptor loci, and combined these with *UAS-GCaMP6f* as anatomical marker (Extended Data Fig. 3). Extending single-neuron dye-filling analyses[11,13,15], OSN glomerular innervation patterns were indistinguishable between *D. sechellia* and *D. melanogaster* (Fig. 4c, Extended Data Fig. 3c-d). However, the glomerular targets of Or22a and Or85c/b neurons (DM2 and VM5d, respectively) were nearly doubled in volume in *D. sechellia* compared to both other species (Fig. 4d)[11,13].

Differences in Or22a circuit wiring

To visualise higher-order elements of the Or22a pathway, we combined a pan-neuronal driver (Extended Data Fig. 10c) with a photoactivatable-GFP transgene to selectively photolabel DM2 projection neurons (PNs). Analysis with analogous genetic reagents in *D. melanogaster* – as well as targeted electroporation of a lipophilic dye[35] into this glomerulus in *D. simulans*, *D. sechellia* and *D. melanogaster* – permitted cross-species comparisons. Two DM2 projection neurons (PNs) were consistently labelled in all three drosophilids (Fig. 4e).

PNs innervate the mushroom body (required for learning/memory) and the lateral horn (implicated in innate olfactory responses)[36]. Within the former, the number and arrangement of PN axonal branches were similar between species (Fig. 4e). In the lateral horn, global anatomy was conserved, with the main tract bifurcating into dorsal and ventral branches. However, dorsal to the bifurcation, *D. sechellia* DM2 PNs had a prominent branch innervating an area not targeted by the homologous *D. melanogaster* or *D. simulans* neurons (Fig. 4e-f). Using successive photo- and dye-labelling to visualise single DM2 PNs in *D. sechellia* and *D. melanogaster*, we confirmed quantitatively the presence of a *D. sechellia*-specific branch (Fig. 4g); this was also detected in *DsecOr22a* mutant animals (Extended Data Fig. 10d-e), indicating its independence of sensory input. These data raise the possibility that *D. sechellia*-specific central circuit changes form part of its olfactory specialisation towards noni.

Discussion

We have developed *D. sechellia* as a model to link genetic and neural circuit changes to behaviours relevant for its ecology. The characterisation of the Or22a pathway and comparison of this circuit's functional and structural properties across closely-related species provides several insights into behavioural evolution (Fig. 4h).

The *Or22a* allele transfer experiments provide evidence that olfactory receptor tuning contributes to species-specific odour-evoked behaviour. Our definition of determinants of OR22A re-tuning also informs the molecular basis of odour/receptor interactions. When mapped onto a presumed homologous ORCO structure[37], the key change (M93I) falls within a putative ligand-binding pocket, and may be a "hot-spot" for functional evolution (Extended Data Fig. 11a-c).

Although functional differences in this receptor are important, they cannot explain the behavioural differences of *D. sechellia* and *D. simulans*, as these receptors are interchangeable for supporting noni attraction. We note that the native Or22a/(b) neuron responses of these species are not identical (Fig. 3a-b); loss of *Or22b* in *D. sechellia* led to a narrowed (and possibly slightly increased) sensitivity to methyl esters, which could be behaviourally relevant. The *D. sechellia*-specific expansion of this neuron population is a likely key additional evolutionary innovation, although it is insufficient alone to restore *D. sechellia*-like host attraction when expressing *Dmel*OR22a. The difference in *D. sechellia* PN axon projections suggests that central circuit connectivity changes form part of this

species' adaptation to noni. Future studies are necessary to understand the genetic bases and behavioural significance of these neuroanatomical differences.

The critical role of OR22a in host attraction in *D. sechellia* may account for the rapid molecular evolution of this locus (Extended Data Fig. 11d-h)[38–40]. *D. erecta*, a specialist on *Pandanus* fruit, also exhibits expansion of this OSN population[21]. However, a second noni-adapted drosophilid, *D. yakuba mayottensis*[41] does not share the same receptor or OSN number changes we describe (Extended Data Fig. 11i-m) implying an independent evolutionary solution to locate a common host fruit.

Finally, other olfactory channels are important for noni attraction. These include Or85c/b neurons, which have conserved physiology while increasing in number in *D. sechellia*, and Ir75b neurons, which have both changed in function and number while apparently preserving partner PN projections[15]. Future application of the *D. sechellia* genetic toolkit should offer further fundamental insights into how genes and neurons control behaviour and enable evolution of novel traits.

Methods

Data reporting

Preliminary experiments were used to assess variance and determine adequate sample sizes in advance of acquisition of the reported data. Several experiments were carried out repeatedly because they served as controls for different genetic manipulations. In particular, we ran wild-type controls in parallel with mutant analyses in behavioural assays and therefore replicated them multiple times. In the wind tunnel assay the number of possible samples per day is rather low, leading to testing of flies of different genotypes on different days. In these cases experiments were started at the same time of the day under the stringently controlled conditions of temperature, humidity, light, and age of flies. For electrophysiological recordings, data were collected from multiple flies on multiple days in randomised order interleaving wild-type and mutant genotypes. Within datasets the same odour dilutions were used for acquisition of the dataset. In all cases the results were reliable and robust over the course of the many years it took to complete this study. For olfactory trap assays, the experiments were conducted with the experimenter blinded to the genotype. The experimenter was not blinded to the genotype of flies in the wind tunnel assay or physiological experiments. All replicates are biological replicates.

Volatile collection, gas chromatography and mass spectrometry

Volatiles were collected from 1 ml of fruit juice or 13 g of noni fruit at different ripening stages in capped 15 ml glass vials with poly-tetra-fluoroethylene-lined silicone septa (Sigma, 23242-U). After penetrating the septum of the cap with a Solid Phase Microextraction (SPME) fibre holder, the SPME fibre (grey hub plain, coated with 50/30 μm divinylbenzene/carboxen on polydimethylsiloxane on a StableFlex fibre (Sigma, 57328-U)) was exposed to the headspace of each vial for 30 min at room temperature. For collection of headspaces in the trap assay, a single noni juice trap was placed in the arena and odours were collected with SMPE for 5 min at 0 h, 5 h and 10 h after placement. For collection of

headspaces in the wind tunnel assay, samples were captured with SPME for 10 min at the landing platform and the release platform directly after noni juice application. After each odour collection, the SPME fibre was retracted and immediately inserted into the inset of a Gas Chromatography-mass spectrometry (GC-MS) system (Agilent 7890B fitted with MS 5977A unit) for desorption at 260ºC in split mode (split ratio 100:1). The GC was operated with a HP-INNOWax column (Agilent 19091N-133UI). Samples were injected at an initial oven temperature of 50°C; this temperature was held for 1 min and gradually increased (3°C min^{-1}) to 150°C before holding for 1 min. Subsequently, the temperature was increased (20°C min^{-1}) to 260°C and held for 5 min. The MS-transfer-line was held at 260°C, the MS source at 230°C, and the MS quad at 150°C. MS spectra were taken in EI-mode (70 eV) in a 29-350 m/z range. Between different collections, the SPME fibre was conditioned at 270°C for 15 min. All chromatograms were processed using MSD ChemStation F.01.03.2357 software. Volatile compounds were identified using the NIST library and matched to standards of the Max-Planck-Institute for Chemical Ecology library. For quantification, peak areas were measured for 3 replicates for each sample. Note that SPME allows only for a qualitative analysis of odour compositions as well as for an estimate of changing ratios of odours across different samples. Vapour pressure values for hexanoic acid, methyl hexanoate and *2*-heptanone were described previously (www.thegoodscentscompany.com)[42,43].

Drosophila strains

Drosophila stocks were maintained on standard wheat flour/yeast/fruit juice medium under a 12 h light:12 h dark cycle at 25°C. For all *D. sechellia* strains, a few g of Formula 4-24® Instant *Drosophila* Medium, Blue (Carolina Biological Supply Company) soaked in noni juice (nu3 GmbH) were added on top of the standard food. Wild-type *Drosophila* strains are described in the corresponding figure legends or Supplementary Table 2. These strains do not show intraspecific sequence variation in OR22a or IR75b for known odour specificity-determining residues (Extended Data Fig. 12), suggesting that other polymorphisms (or non-genetic factors) underlie the observed minor intraspecific behavioural differences (Fig. 1b,c). The mutant and transgenic lines used and generated in this study are listed in Supplementary Table 2.

CRISPR/Cas9-mediated genome engineering

sgRNA expression vectors: for expression of single sgRNAs, oligonucleotide pairs (Supplementary Table 3) were annealed and cloned into *BbsI*-digested *pCFD3-dU6-3gRNA* (Addgene #49410), as described[44]. To express multiple sgRNAs from the same vector backbone, oligonucleotide pairs (Supplementary Table 4) were used for PCR and inserted into *pCFD5* (Addgene #73914) via Gibson Assembly, as described[45].

Donor vectors for homologous recombination: to generate an eGFP-expressing donor vector (*pHD-Stinger-attP*), the fluorophore was excised from *pStinger*[46] with *NcoI*/*HpaI* and used to replace the *DsRed* sequence in *NcoI*/*HpaI*-digested *pHD-DsRed-attP* (Addgene plasmid #51019)[47]. Homology arms (1-1.6 kb) for individual target genes were amplified from *D. sechellia* (*Drosophila* Species Stock Center [DSSC] 14021-0248.07), *D. simulans* (DSSC 14021-0251.195) or *D. melanogaster* (Research Resource Identifier Database:Bloomington *Drosophila* Stock Center [RRID:BDSC]_58492) genomic DNA and inserted either into

pHD-DsRed-attP or *pHD-Stinger-attP* via restriction cloning. Details and oligonucleotide sequences are available from the corresponding authors upon request.

Transgenic source of Cas9: *pBac(nos-Cas9,3XP3-YFP)* (gift of D. Stern) was integrated into *D. sechellia* (DSSC 14021-0248.07) via *piggyBac* transgenesis. The insertion was mapped to the fourth chromosome using TagMap[48].

Transgene construction

Oligonucleotides for each cloning step are listed in Supplementary Table 5.

attB-nSyb-Gal4,miniW: 1.9 kb upstream sequence of the *neuronal Synaptobrevin* (*nSyb*) gene were amplified from *D. sechellia* genomic DNA (DSSC 14021-0248.07) and inserted into *pGal4attB*[49] via restriction cloning using *NotI* and *KpnI*.

attB-Gal4,3XP3-Stinger: we first generated an *eGFPnls-SV40* fragment via PCR (using *pHD-Stinger-attP* as template) and fused it to a minimal *attB40* site[50,51] before insertion into *pCR-Blunt II-TOPO* (Thermo Fisher). We added a *3XP3-Stinger* fragment amplified from *pHD-Stinger* via restriction cloning using *EcoRV* and *SalI*. Subsequently, we placed a *loxP* site downstream of the initial *SV40* sequence via oligonucleotide annealing and *SpeI/KpnI* restriction cloning, to produce *pCR-TOPO-loxP-attB40-eGFPnlsSV40rev-3XP3:Stinger*. We replaced the *eGFPnls-SV40* sequence with an *hsp70-Gal4-SV40* fragment via PCR amplification of the vector backbone and *Gal4* from *pGal4attB*[49] and subsequent Gibson Assembly resulting in *attB-Gal4,3XP3-Stinger*.

attB-Or22a^wt,3XP3-Stinger: in the *pCR-TOPO-loxP-attB40eGFPnlsSV40rev-3XP3-Stinger* plasmid described above, the *eGFPnlsSV40* fragment was flanked by *EcoRV* and *SalI* sites, which were used to integrate the *D. sechellia* or *D. melanogaster Or22a* ORF+3'UTR after PCR amplification from cDNA, or the *D. simulans Or22a* ORF+3'UTR (synthesised by Eurofins Genomics), to produce *attB-DsecOr22a^wt,3XP3-Stinger*, *attB-DmelOr22a^wt,3XP3-Stinger*, and *attB-DsimOr22a^wt,3XP3-Stinger*, respectively.

Or22a chimeras: chimeric sequences of *D. sechellia* and *D. melanogaster Or22a* were generated by PCR amplification and fusion using the respective species' *Or22a* gene templates. After subcloning into *pCR-Blunt II-TOPO* and sequence confirmation, the chimeras were integrated into *pCR-TOPO loxP attB40eGFPnlsSV40rev-3XP3-Stinger* via restriction cloning.

Or22a site-directed mutant constructs: point mutations were introduced via site directed mutagenesis following standard procedures.

attB-UAS-constructs: *DsecOr22a^wt*, *DmelOr22a^wt*, *DmelOr22a^triple* and *DmelOr22a^M93I* were amplified by PCR incorporating flanking *EcoRI* and *SalI* restriction sites using the constructs described above as template, and integrated into the *EcoRI/XhoI*-digested *pUAST-attB*[52].

pDONR221-MCS: a *pDONR221* entry vector carrying a multiple cloning site (MCS) was generated by amplification of the MCS of *pCR-Blunt II-TOPO* incorporating flanking

attB1/2 sites and integration of the PCR fragment into *pDONR221* via a BP reaction (Gateway, Thermo Fisher Scientific).

pDONR221-DsimOr22a: a 5.7 kb promoter region upstream of the *Or22a* start codon was amplified from *D. simulans* (DSSC 14021-0251.195) genomic DNA, subcloned into *pCR-Blunt II-TOPO* and transferred to *pDONR221-MCS* via *BamHI/EcoRV* restriction cloning.

pDEST-Hemmar-eGFPnls: the *eGFPnls* fragment of *pStinger*[46] was amplified by PCR incorporating *XhoI* and *SpeI* restriction sites and integrated into the *XhoI/XbaI*-digested vector *pDEST-HemmarG*[53] to replace *eGFP*.

pDsimOr22a-eGFPnls: *pDONR221-DsimOr22a* and *pDEST-Hemmar-eGFPnls* were combined using LR recombination (Gateway, Thermo Fisher Scientific).

Drosophila microinjections

Transgenesis of *D. sechellia*, *D. simulans* and *D. melanogaster* was performed in-house following standard protocols (http://gompel.org/methods), except for *DsimOr22a-GFPnls* (generated by Rainbow Transgenic Flies Inc.). For the *D. sechellia* egg-laying agar plates, we replaced grape juice with noni juice and added on the surface a few g of Formula 4-24® Instant *Drosophila* Medium, Blue (**Carolina Biological Supply Company**) soaked in noni juice (nu3 GmbH). Embryos were manually selected for the appropriate developmental stage prior to alignment and injection. For *piggyBac* transgenesis, we co-injected *piggyBac* vector (300 ng μl^{-1}) and *piggyBac* helper plasmid[54] (300 ng μl^{-1}). For CRISPR/Cas9-mediated homologous recombination, we injected a mix of an sgRNA-encoding construct (150 ng μl^{-1}), donor vector (400 ng μl^{-1}) and *pHsp70-Cas9* (400 ng μl^{-1}) (Addgene #45945)[55]. The DsRed fluorescent marker was destroyed in *DsecnSyb-Gal4* and *DsecUAS-C3PA-GFP* via injection of an sgRNA construct targeting *DsRed* (150 ng μl^{-1}) and *pHsp70-Cas9* (400 ng μl^{-1}). Injections into *Dsecnos-Cas9* were of a mix of an sgRNA construct (150 ng μl^{-1}) and donor vector (500 ng μl^{-1}). Site-directed integration into *attP* sites was achieved by co-injection of an *attB*-containing vector (400 ng μl^{-1}) and either *p3xP3-EGFP.vas-int.NLS* (400 ng μl^{-1}) (Addgene #60948)[56] or *pBS130* (encoding phiC31 integrase under control of a heat shock promoter (Addgene #26290)[57]). All concentrations are given as final values in the injection mix.

Wind tunnel assay

Long-range attraction experiments were performed in a wind tunnel as described previously[23] with a flight arena of 30 cm width, 30 cm height and 100 cm length. The airstream in the tunnel (0.3 m s^{-1}) was produced by a fan (Fischbach GmbH, Neunkirchen, Germany), and filtered through an array of four activated charcoal cylinders (14.5 cm diameter × 32.5 cm length; Camfil, Trosa, Sweden). The wind tunnel was maintained within a climate chamber at 25°C and 50-55% relative humidity under white light. Flies were starved for approximately 20 h; to ensure the flight ability of assayed animals, flies were first released into a mesh cage (50 × 50 × 50 cm, maintained at the same conditions as the wind tunnel) and females escaping from the food vial were collected with an aspirator. For each assay, ten 4-6 day-old females were released from a plastic tube (with a mesh covering

one end and the open end facing the landing platform) fixed horizontally in the centre of the first 5-10 cm of the downwind end of the tunnel. The landing platform was built by a filter paper (3 × 3 cm) charged with either 100 µl of juice (noni (Nu3 GmbH), grape (Beutelsbacher Fruchtsaftkelterei), pineapple (Andros), mango (Migros)), apple cider vinegar (Migros) or ~100 µl of homogenised ripe fruit (noni, fig) and fixed on a metal holder. Fruit homogenisation was performed by blending 10 mg of ripe fruit in 20 ml of distilled water, fruit particles were pelleted by centrifugation and the supernatant harvested for experiments. In two-choice assays two identical landing platforms were positioned with equal distance (7.5 cm) from the centre of the air stream alternating the position of noni fruit and apple cider vinegar between assays. The fly tube was placed within the centre of the airstream and 85 cm downwind of the odour source. An experimenter observed the landing platform(s) for the entire duration of the assay; flies arriving and staying on the landing platform(s) within the first 10 min after release were counted.

Olfactory trap assay

The two-choice olfactory trap assay was performed essentially as described[15]. For each experiment, traps contained either 300 µl of juice (noni (Nu3 GmbH), grape (Beutelsbacher Fruchtsaftkelterei), pineapple (Andros), mango (Migros)) or apple cider vinegar (Migros). When using fruits as stimuli (noni, grape, papaya, banana, fig), ripe fruits were peeled (banana, papaya) or used whole, homogenised with pestles, and each trap was filled with a spatula of the mix (to ~300 µl). Single odours (see below for CAS numbers) were used at a 10^{-2} dilution in grape juice. Triton X-100 (final concentration 0.2%) was added to all traps containing single odours and respective control traps to drown trapped flies. 25 fed, mated, ice-anesthetised female flies (3-5 day-old) were used for each experiment. *D. sechellia* flies were transferred to standard food without noni supplement 24 h prior to the start of the assay, unless stated otherwise. The distribution of flies was scored after 24 h at 25°C under red light at 60% relative humidity; experiments with >25% dead flies in the arena after 24 h were discarded. The attraction index was calculated as follows: (number of flies in treatment (e.g., noni juice) trap - number of flies in control (e.g., grape juice) trap)/number of trapped and untrapped flies alive. For the trap assay quantifications in Fig. 2e, all untrapped flies (including those that died during the assay) were counted due to the high mortality rate in these experiments (Extended Data Fig. 7c).

The multiple-choice olfactory trap assay was performed using eight traps containing mango, pineapple, noni and grape juices, apple cider vinegar and fig, banana and papaya fruits. Traps were placed equidistantly in a circle in random order for each experiment and conditions were as described for the two-choice assay. The percentage of flies per trap was calculated as follows: (number of flies in trap/number of trapped and untrapped flies alive) ×100. Experiments with >25% dead flies in the arena after 24 h were discarded.

Two-photon calcium imaging

Flies were mounted and dissected as previously described[58], and images were acquired using a commercial upright two-photon microscope (Zeiss LSM 710 NLO). In detail, an upright Zeiss AxioExaminer Z1 was fitted with a Ti:Sapphire Chameleon Ultra II infrared laser (Coherent) as excitation source. Images were acquired with a 20× water-immersion

objective (W Plan-Apochromat 20×; NA 1.0, VIS-IR DIC), with a resolution of 128×128 pixels (1.1902 pixels μm^{-1}) and a scan speed of 12.6 μs pixel^{-1}. The excitation wavelength was set to 920 nm at a laser output of 64.1-70.2 mW measured at the exit of the objective. Emitted light was filtered with a 500-550 nm band-pass filter, and photons were collected by an external non-descanned detector. Each measurement consisted of 50 images acquired at 4.13 Hz, with stimulation starting ~5 s after the beginning of the acquisition and lasting for 1 s. Fly antennae were stimulated using a custom-made olfactometer as previously described[59], with minor modifications. In brief, the fly's antennae were permanently exposed to air flowing at a rate of 1.5 l min^{-1} and with 55% relative humidity obtained by combining a main stream of humidified room air (0.5 l min^{-1}) and a secondary stream (1 l min^{-1}) of normal room air. Both air streams were generated by vacuum pumps (KNF Neuberger AG) and the flow rate was controlled by two independent rotameters (Analyt). The secondary air stream was guided through either an empty 2 ml syringe or, to generate an odour pulse, a 2 ml syringe containing 20 μl of odour or solvent on a small cellulose pad (Kettenbach GmbH). Solvents were either double-distilled water (for noni juice and apple cider vinegar) or paraffin oil (for methyl hexanoate (CAS 106-70-7), *2*-heptanone (CAS 110-43-0), *2,3*-butanedione (CAS 431-03-8), ethyl propionate (CAS 105-37-3) and *1*-hexanol (CAS 111-27-3)). To switch between control air and odour stimulus delivery, a three-way magnetic valve (The Lee Company, Westbrook, CT) was controlled using Matlab via a VC6 valve controller unit (Harvard Apparatus). Data were processed using Fiji[60] and custom written scripts in Matlab and R as previously described[59]. Since bleaching was very strong at the beginning of each acquisition, the first 1.5 s were not considered for the analysis, and bleach correction was not applied. Colour-coded images and boxplots show the peak response calculated as the mean relative change in fluorescence (% F/F) of three frames around the maximum during frames 19-30.

Widefield calcium imaging

Flies were mounted and dissected as previously described[58]. Images were acquired with a CCD camera (CoolSNAP-HQ2 Digital Camera System) mounted on a fluorescence microscope (upright fixed stage Carl Zeiss Axio Examiner D1) equipped with a 20× water-immersion objective (W Plan-Apochromat 20×; NA 1.0, VIS-IR DIC, Extended Data Fig. 4a) or 40× water-immersion objective (W Plan-Apochromat 40×; NA 1.0 VIS-IR DIC, Extended Data Fig. 4f). Excitation light of 470 nm was produced with an LED light (Cool LED pE-100, VisiChrome, intensity 1.8-3.9%). Light was guided through a filter block consisting of a 450-490 nm excitation filter, a dichroic mirror (T495LP), and a 500-550 nm emission filter (Chroma ET). Binned image size was 400×300 pixels (Extended Data Fig. 4a) or 266×200 pixels (Extended Data Fig. 4f) on the chip, corresponding to 465×349 μm (Extended Data Fig. 4a) or 149×112 μm (Extended Data Fig. 4f) in the preparation. Exposure time varied between 30-100 ms to adjust for different basal fluorescence values across preparations. Films (12.5 s duration) were recorded with an acquisition rate of 4 Hz. Metafluor software (Visitron) was used to control the camera, light, data acquisition and onset of odour stimulation. Odour stimulation and data analysis were otherwise performed as described for two-photon calcium imaging.

Electrophysiology

Single sensillum electrophysiological recordings were performed as described previously[61]. Noni and grape juice were purchased from Nu3 (nu3 GmbH) and Beutelsbacher (Beutelsbacher Fruchtsaftkelterei GmbH) and chemicals of the highest purity available from Sigma-Aldrich. Odorants were used at 10^{-2} (v/v) in all experiments unless noted otherwise in the figures or figure legends. Solvents were either double-distilled water (for noni juice, butyric acid (CAS 107-92-6), hexanoic acid (CAS 1821-02-9)) or paraffin oil (for octanoic acid (CAS 124-07-2), methyl butanoate (CAS 623-42-7), methyl hexanoate (CAS 106-70-7), methyl octanoate (CAS 111-11-5), ethyl butanoate (CAS 105-54-4), ethyl hexanoate (CAS 123-66-0), ethyl octanoate (CAS 106-32-1), *2*-heptanone (CAS 110-43-0), *1*-hexanol (CAS 111-27-3) and *3*-buten-*1*-ol (CAS 627-27-0)). Odours for sensilla stimulation were used for a maximum of five consecutive trials. If noni juice or noni fruit extract (see Wind tunnel assay section for details) were used, odours were renewed after two stimulations or, for 10^{-4}, 10^{-3} and 10^{-2} dilutions, after each stimulation. Corrected responses were calculated as the number of spikes in a 0.5 s window at stimulus delivery (200 ms after stimulus onset to take account of the delay due to the air path) subtracting the number of spontaneous spikes in a 0.5 s window 2 s before stimulation, multiplied by two to obtain spikes s^{-1}. The solvent-corrected responses shown in the figures were calculated by subtracting from the response to each diluted odour the response obtained when stimulating with the corresponding solvent. Recordings were performed on a maximum of three sensilla per fly. Exact n and mean spike counts for all experiments are provided in Supplementary Data Table 7.

Immunohistochemistry

Fluorescent RNA *in situ* hybridisation using digoxigenin- or fluorescein-labelled RNA probes and immunofluorescence on whole-mount antennae were performed essentially as described[59,62]. *D. sechellia* probe templates were generated by amplification of regions of genomic DNA (DSSC 14021-0248.07) using primer pairs listed in Supplementary Data Table 6; these were cloned into *pCR-Blunt II-TOPO* and sequenced. *D. sechellia* OR22a antibodies were raised in rabbits against the peptide epitope PHISKKPLSERVKSRD (amino acids 7-22), affinity-purified (Proteintech Group Inc) and diluted 1:250. Other antibodies used were: guinea pig α-IR75b (RRID:AB_2631093)[63] 1:200, rabbit α-IR64a (RRID:AB_2566854)[34] 1:100, rabbit α-ORCO[28] 1:200, guinea pig α-IR8a (RRID:AB_2566833)[27] 1:500, rabbit α-IR25a (RRID:AB_2567027) 1:500[64], rabbit α-GFP 1:500 (Invitrogen). Immunofluorescence on adult brains was performed as described[65] (except for the *D. sechellia* reference brain samples; see below) using mouse monoclonal antibody nc82 1:10 (Developmental Studies Hybridoma Bank), rat monoclonal α-Elav 1:10 (Developmental Studies Hybridoma Bank) and rabbit α-GFP 1:500 (Invitrogen). Alexa488-, Cy3- and Cy5-conjugated goat α-guinea pig, goat α-mouse, goat α-rat and goat α-rabbit IgG secondary antibodies (Molecular Probes; Jackson Immunoresearch) were used at 1:500.

D. sechellia reference brain

D. sechellia (DSSC 14021-0248.07) brains (2-7 day-old animals) were stained with nc82 and imaged as described[66]. From 88 female brains imaged, 26 high-quality confocal stacks of the midbrain were selected for averaging on a selected "seed" brain, essentially as

described[67,68]. Similarly, a male reference brain (not shown here) was constructed, using 20 high-quality confocal stacks (from 87 initially imaged). Reciprocal bridging registrations between *D. melanogaster* and *D. sechellia* references brains were also generated to permit comparison of homologous neurons within a common template, essentially as described[68]. The reference brains (*DsecF* and *DsecM*), bridging registrations and associated code are available for download via http://jefferislab.org/si/auer2019.

Image acquisition and processing

Confocal images of antennae and brains were acquired on an inverted confocal microscope (Zeiss LSM 710) equipped with an oil immersion 40× objective (Plan Neofluar 40× Oil immersion DIC objective; 1.3 NA), unless stated otherwise. Images were processed in Fiji[60]. *D. sechellia* brains were imaged and registered to a *D. sechellia* reference brain using the Fiji CMTK plugin (https://github.com/jefferis/fiji-cmtk-gui), as described[69]. For segmentation of individual glomeruli of the antennal lobe, glomerular identity was confirmed by location and labelling with *Gal4* reporters (Extended Data Fig. 3) and segmentation performed using Amira 6.5 (Thermo Fisher Scientific). Glomerular volumes were calculated following segmentation with the Segmentation Editor plugin of Fiji using the 3D Manager Plugin. OSN numbers were counted using the Cell Counter Plugin in Fiji or Imaris (Bitplane). Projection neuron morphologies were reconstructed and measured in neuTube 1.0z[70].

Projection neuron labelling

Photoactivation was performed as described[35] on 3-5 day-old female flies. Brains were dissected in saline[71] (low carbonate; 2 mM Mg^{2+} pH 7.2) and treated with collagenase (2 mg ml^{-1}, 45 s). Dissected brains were initially imaged at 925 nm to identify the DM2 glomerulus based on anatomical position. Photoactivation was achieved through multiple cycles of exposure to 710 nm laser light with a 15 min rest period between each cycle to allow diffusion of the photoactivated fluorophore within the neuron. Photoactivation and imaging was performed on an Ultima two-photon laser scanning microscope (Bruker) equipped with galvanometers driving a Chameleon XR laser (Coherent). Emitted photons were collected with a GaAsP photodiode detector (Bruker) or a PMT detector through a 60× objective (Olympus 60× water immersion; 0.9 NA).

PN dye-fillings were performed as described[35] with some modifications. Brains were dissected in saline, briefly treated with collagenase (2 mg ml^{-1}, 45 s), washed and pinned with fine tungsten wires to a Sylgard sheet (World Precision Instruments) in a 35 mm Petri dish (Falcon) filled with saline. Pulled glass electrodes were backfilled with Texas Red Dextran (3000, lysine fixable, Thermo Scientific). The electrode was targeted to the DM2 glomerulus using as a guide either basal expression of the pan-neuronal, photoactivatable GFP (*D. sechellia* and *D. melanogaster*) or DM2-specific labelling (*D. simulans*, i.e., the *Or22a-GFPnls* strain, in which trace GFP levels were detected in OSN axon termini). The dye was electroporated by applying voltage pulses (30 V) until it became visible in distal neural processes of the PN, and left to diffuse for 60 min. Brains were subsequently imaged by two-photon microscopy, as described above.

To label single PNs, the DM2 glomerulus was first subjected to one cycle of exposure to 710 nm laser light to identify the cell bodies of DM2 PNs. Subsequently, the filled glass electrode was placed in the centre of the soma of one DM2 PN and the dye was electroporated by applying voltage pulses (30 V) until it became visible in distal neural processes of the neuron. The dye was left to diffuse for 60 min, before the brains were fixed with 2% paraformaldehyde for 45 min and subjected to immunofluorescence using nc82 (1:20) and α-mouse Alexa488 (1:500), as described above. Images were acquired on a Zeiss LSM 880 Airy scan confocal microscope using a 40× objective (Plan Neofluar 40× oil immersion DIC objective; 1.3 NA).

Molecular evolution and polymorphism analyses

The *Or22a/b* gene tree was inferred using annotations after manual verification by BLASTing OR22a/b protein sequences (using tblastn with default settings, BLAST+ v2.7.1[72]) to the genomes of *D. simulans* (r2.01), *D. mauritiana* (r1), *D. yakuba* (r1.04), *D. sechellia* (r1.3) and *D. erecta* (r1.3). Genomic regions were annotated using Wise2 (v2.4.1)[73], with *D. melanogaster*'s protein sequences as guides. All genes appeared intact and consistent with existing models, with one exception: *D. mauritiana Or22b* is truncated by ~240 bp due to a gap in the reference genome. The cDNAs were frame-aligned using TranslatorX[74]. The gene tree based on these alignments was inferred using MrBayes (v3.2.6)[75] with the following settings: rate variation = invariable+gamma, number of substitutions = 6, substitution model = default, number of generations = 10, sample frequency = 10, burn-in = 250.

The *Or22a/b* topology outputted from MrBayes was used for PAML (v4.8)[76] CodeML branch tests of protein evolution rate change. Seven models were tested, one with a freely varying ω value and six other models with 1, 4, 5, 7, 8 and 9 ω values (CodeML parameters were all set to zero except for the following: seqtype = 1, CodonFreq = 2, ndata = 1, model = 2, kappa = 2, omega = 0.4, fix_alpha = 1, ncatG = 5, getSE = 1, Small_Diff = 5e-7, cleandata = 1). Nested comparisons between these models using likelihood ratio tests identified a model with 5 ω values to provide the best fit (Supplementary Data Table 8, Extended Data Fig. 11d).

Polymorphism data for *Or22a/b* was extracted from previously published data. The *D. simulans* and *D. sechellia* datasets were from[7,77] and the *D. melanogaster* datasets from[78,79]. VCFtools (v0.1.17)[80] was used to extract variable sites from the respective VCF files, as well as to output allele frequencies. Variant calling in[7] was based on a subversion of the *D. simulans* r2 genome, which is slightly different than r2 on FlyBase. This subversion is available at https://github.com/kern-lab/FILET/blob/master/simSechResults/dsimV2-Mar2012_chrsonly.fa.bz2.

Statistics and reproducibility

Data were analysed and plotted using Excel and R (v3.2.3; R Foundation for Statistical Computing, Vienna, Austria, 2005; R-project-org).

Extended Data
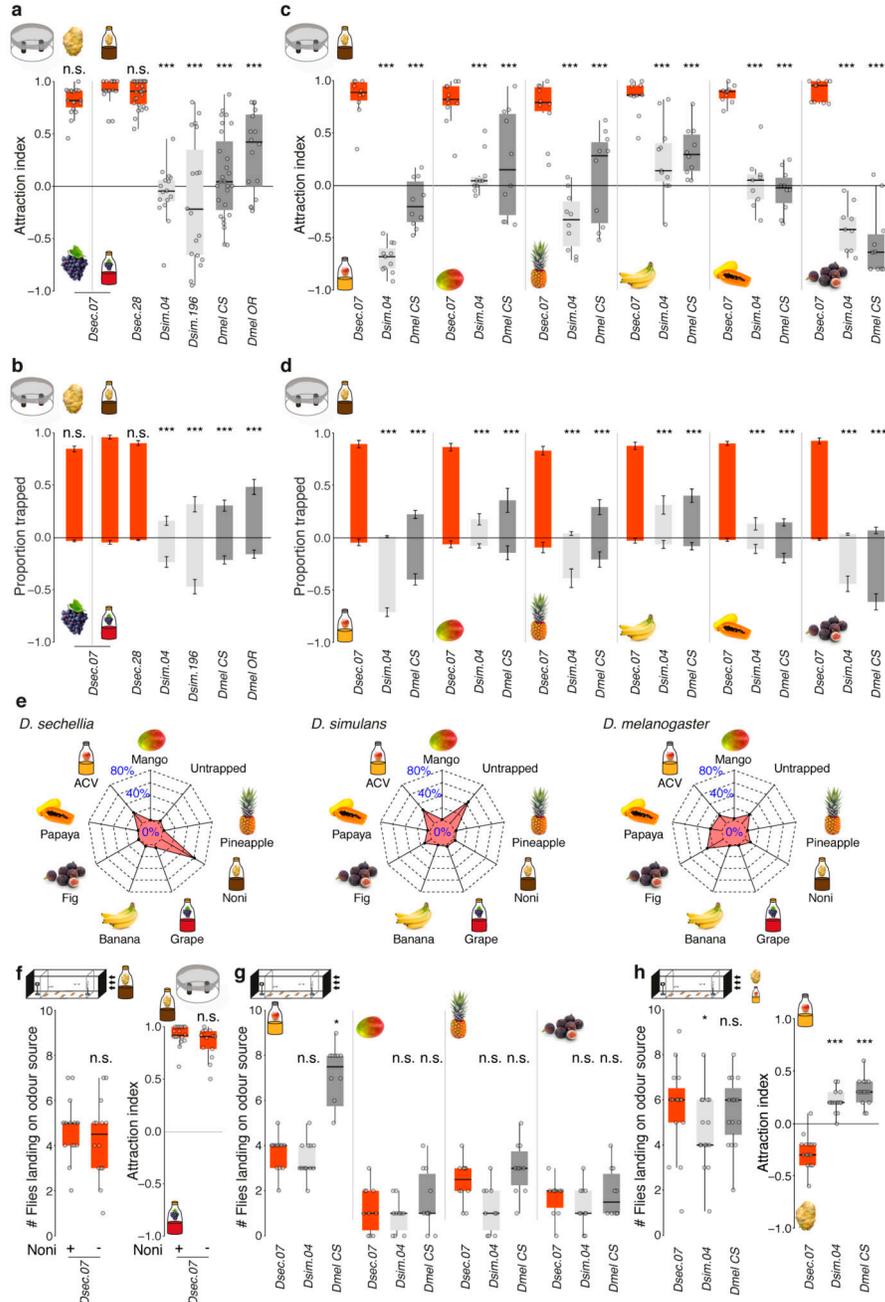
Extended Data Fig. 1. Species-specific short- and long-range behavioural responses to diverse fruit stimuli.

a, Data reproduced from Fig. 1c. Behavioural responses in a trap assay testing preferences between noni and grape, or between noni juice and grape juice (n = 15-27 experiments; 22-25 females/experiment). Comparisons to *Dsec.07* responses to noni juice are shown (**a-d** and **h** (right): pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** *P* < 0.001; n.s. *P* > 0.05.

b, Proportion of flies (mean ± SEM) in each stimulus trap for the assays shown in **a**. Comparisons to *Dsec.07* responses are shown.

c, Behavioural responses in a trap assay testing preferences between noni juice and diverse fruit juices or fruits for *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004) and *D. melanogaster* (Canton-S) (n = 10-11 experiments; 22-25 females/ experiment). Comparisons to *Dsec.07* responses are shown.

d, Proportion of flies (mean ± SEM) in each stimulus trap for the assays shown in **c**. Comparisons to *Dsec.07* responses are shown.

e, Radar plot showing the mean percentage of flies per trap in a multiple choice trap assay with eight different stimuli for *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004) and *D. melanogaster* (Canton-S) (n = 11 experiments; 22-25 females/ experiment). ACV = Apple cider vinegar.

f, Left: Behavioural responses to noni juice in a wind tunnel assay of *D. sechellia* (DSSC 14021-0248.07) reared on standard food with (+) and without (-) noni supplement (Kruskal-Wallis test: n.s. *P* > 0.05). Right: Behavioural responses in a trap assay testing preferences between noni juice and grape juice for *D. sechellia* (DSSC 14021-0248.07) reared on standard food with (+) and without (-) noni supplement (pairwise Wilcoxon rank-sum test: n.s. *P* > 0.05).

g, Behavioural responses to apple cider vinegar, mango juice, pineapple or fig in a wind tunnel assay of *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004) and *D. melanogaster* (Canton-S) (n = 10-12 experiments; 10 females/experiment). Comparisons to *Dsec.07* responses are shown (**g** and **h** (left): Kruskal-Wallis test with Dunn's post-hoc correction): * *P* < 0.05; n.s. *P* > 0.05.

h, Behavioural responses in a wind tunnel assay testing preference between noni fruit and apple cider vinegar of *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004) and *D. melanogaster* (Canton-S) (n = 15 experiments; 10 females/ experiment). Left: total number of flies landing on an odour source. Comparisons to *Dsec.07* responses are shown. Right: attraction index calculated as: (flies landing on apple cider vinegar – flies landing on noni)/flies landing on either source. Comparisons to *Dsec.07* responses are shown.
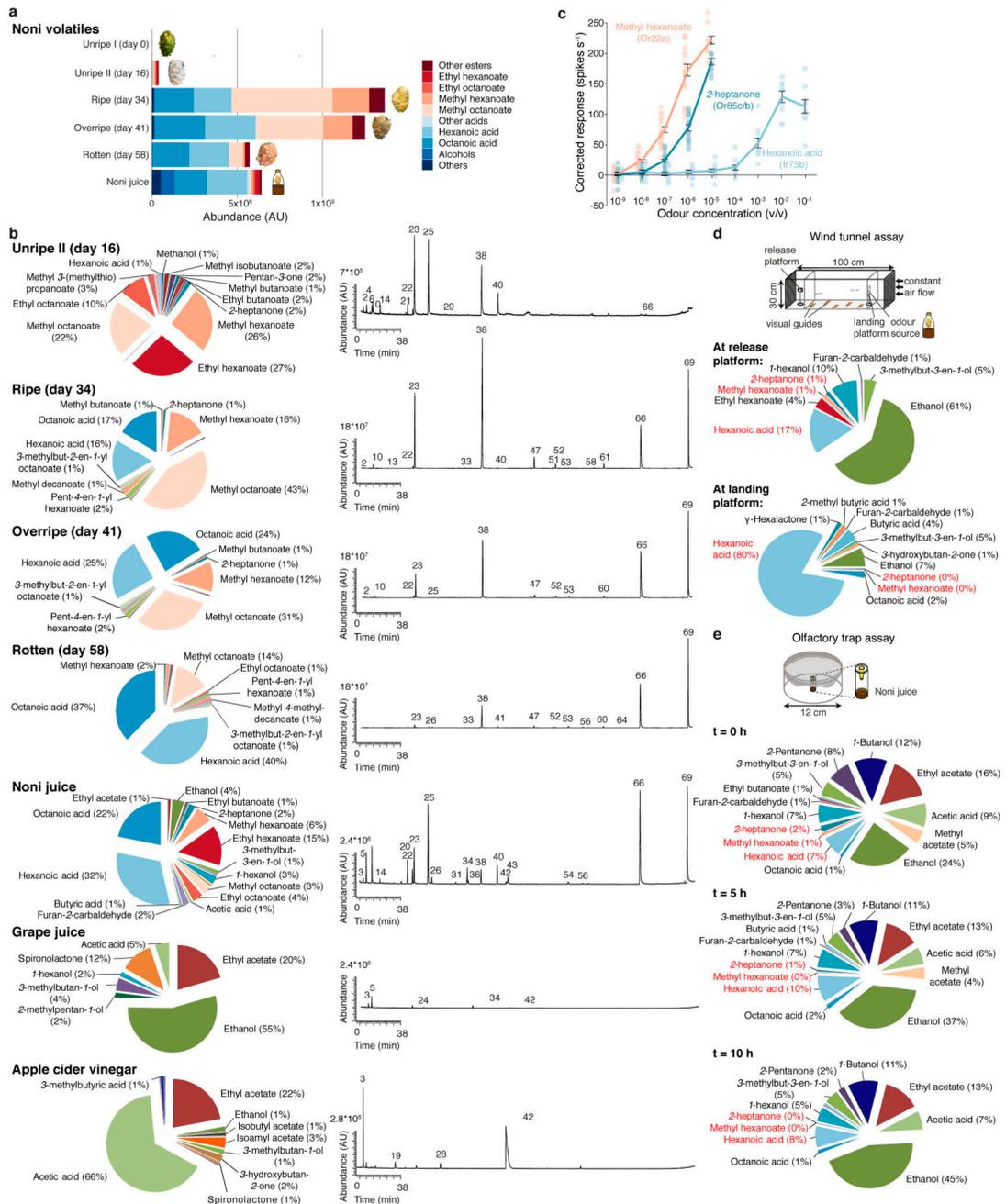
Extended Data Fig. 2. Chemicals emitted by natural odour sources and odour bouquet changes in behavioural assays.

a, Principal constituents of the odour bouquet of noni fruit at different ripening stages and commercial noni juice, as determined by gas chromatography/mass spectrometry. AU = arbitrary units.

b, Chemical composition of the odour bouquet of noni fruit at different stages of ripening, noni juice, grape juice and apple cider vinegar. Representative gas chromatograms are shown

on the right. Numbers correspond to compounds as listed in Supplementary Table 1 (not all identified peaks are shown).

c, Dose-dependent electrophysiological responses of Or22a, Or85c/b and Ir75b neurons in wild-type *D. sechellia* (DSSC 14021-0248.07) to their best odour agonists (mean ± SEM and individual data points, n = 7-20, females). The contribution of Or35a neurons (whose spiking is difficult to separate from Ir75b neurons in ac3I) to hexanoic acid responses is likely to be minimal (Fig. 2b). *D. sechellia* Or22a and Or85c/b dose response data are replotted from Fig. 3a.

d, Chemical profile of odours collected by SPME at the release and landing platforms in the wind tunnel assay within the first 10 min after noni juice application.

e, Chemical profile of odours collected by SPME in the trap assay arena within 5 min after placement of a trap (i.e., t = 0 h), and after 5 h and 10 h, using noni juice as stimulus. "0%" indicates only trace proportions of the compound were detected.

Extended Data Fig. 3. Olfactory sensory neuron Gal4 driver lines in *D. sechellia*.
a, Schematic of the *Gal4* reporter allele generation strategy, through CRISPR/Cas9-mediated integration of an *attP* site (marked by *3xP3:DsRed*) into the desired *Or* or *Ir* gene (see Extended Data Fig. 5-6 for details on specific alleles), followed by introduction of a *Gal4* open reading frame via phiC31-mediated transgenesis.
b, Co-expression of the indicated *Or^Gal4^*-driven, *Ir^Gal4^*-driven or control background GCaMP6f signal (detected by α-GFP) with the corresponding receptor protein or RNA in whole-mount antennae. Arrowheads point to examples of co-labelled cells. Scale bars, 25

μm; inset scale bars, 5 μm. While $Or22a^{Gal4}$ and $Ir64a^{Gal4}$ completely recapitulate endogenous receptor expression, $Orco^{Gal4}$ and $Or85c/b^{Gal4}$ lack expression in some receptor-expressing neurons. $Or35a^{Gal4}$ and $Ir75b^{Gal4}$ might be expressed in ectopic cells (see also **c**-**d**) or the protein/RNA signal for these receptor genes could be below the detection threshold.

c, Expression of the indicated Or^{Gal4}-driven, Ir^{Gal4}-driven or control background GCaMP6f signal (detected by α-GFP) in glomeruli of whole-mount antennal lobes. Three focal planes of the neuropil (visualised with nc82 (magenta)) are shown. Images were registered to a *D. sechellia* reference brain (see Methods) for better comparison of antennal lobe structure. Scale bar, 25 μm.

d, Summary of the glomerular labelling by Or^{Gal4} or Ir^{Gal4} drivers as characterised in **c** (dark blue indicates GCaMP6f signal was detected in at least 3/3 independent brains). Glomeruli are organised by the compartmentalisation of the corresponding OSN populations into different sensilla classes (based on data in *D. melanogaster*[81], ab: antennal basiconic; at: antennal trichoid; ai; antennal intermediate; ac: antennal coeloconic, pb = palp basiconic, sac = sacculus, ? = OSN population unknown). $Orco^{Gal4}$ is expressed in most but not all (e.g., Or67d/DA1) expected OSN populations; $Or35a^{Gal4}$ and $Or85c/b^{Gal4}$ display some ectopic expression as inferred by their labelling of more than one glomerulus.
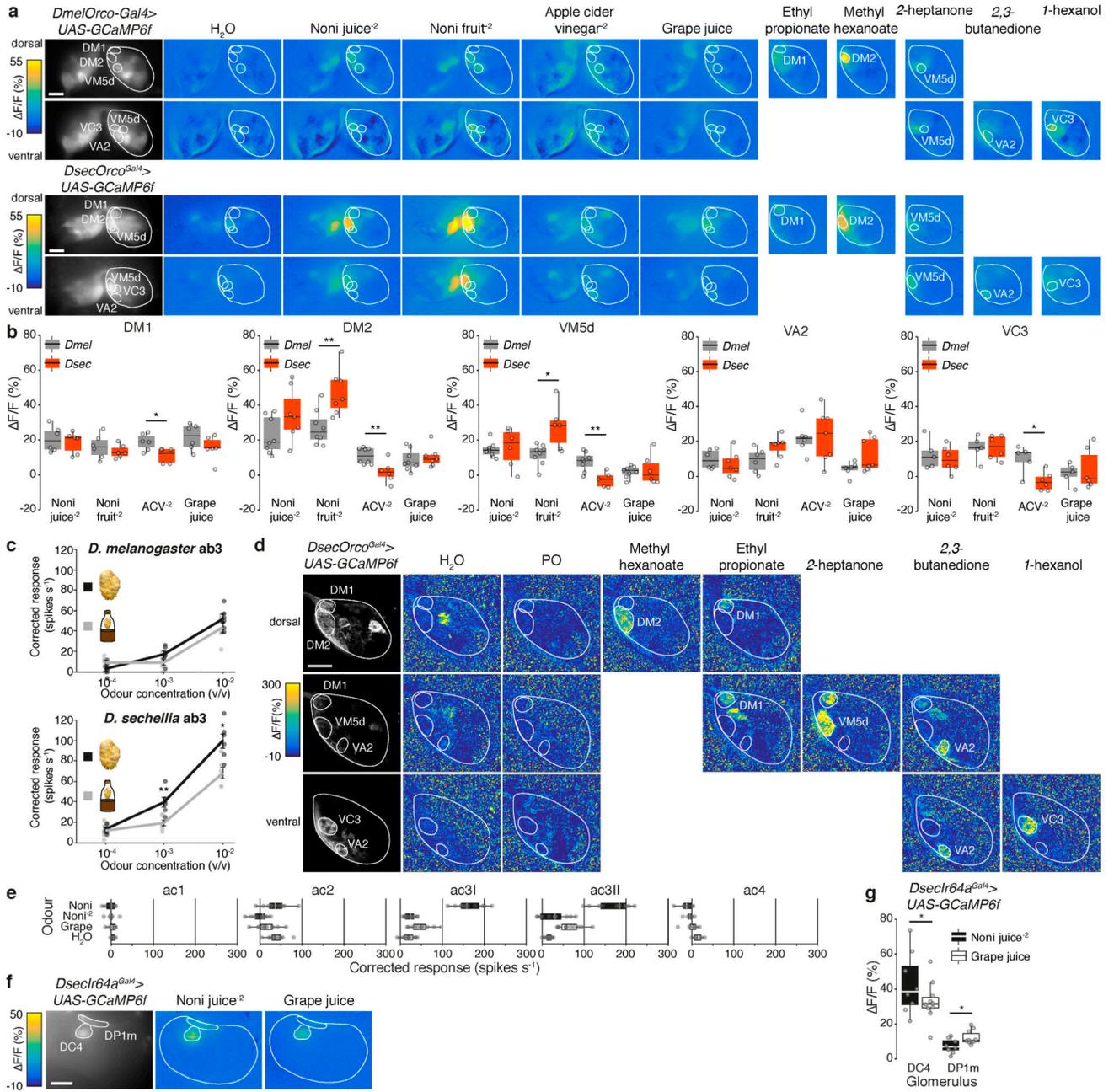
Extended Data Fig. 4. Comparative olfactory representations of noni in *D. melanogaster* and *D. sechellia*.

a, Representative odour-evoked calcium responses in the axon termini of Orco OSNs in the antennal lobes of *D. melanogaster* (*Orco-Gal4/Orco-Gal4;UAS-GCaMP6f/UAS-GCaMP6f*) and *D. sechellia* (*UAS-GCaMP6f/UAS-GCaMP6f;;DsecOrco^{Gal4}/+;*), acquired by widefield imaging. Left images: raw fluorescence signals. Right images: relative increase in GCaMP6f fluorescence (ΔF/F%) after stimulation with the indicated complex stimuli and single odours. Diagnostic odours: methyl hexanoate (10^{-6} (v/v)) for Or22a/(b)/DM2, ethyl propionate (10^{-4}) for Or42b/DM1, *2*-heptanone (10^{-5}) for Or85c/b/VM5d, *2,3*-butanedione

(10^{-4}) for Or92a/VA2 and *1*-hexanol (10^{-4}) for Or35a/VC3. Glomerular boundaries, and the entire antennal lobe, are outlined. Scale bars, 50 µm.

b, Quantification of odour-evoked calcium responses for the animals represented in **a**. Maximum calcium response amplitudes for each experiment are plotted (n = 5-8 females). Significantly different responses of species to the same stimulus are shown (Wilcoxon signed-rank test): ** $P < 0.01$; * $P < 0.05$.

c, Combined electrophysiological responses of neurons in the ab3 sensillum in *D. melanogaster* (top) and *D. sechellia* (bottom) upon stimulation with increasing concentrations of noni juice or noni fruit extract (mean ± SEM and individual data points, n = 6, females). Significant differences in responses are shown (pairwise Wilcoxon rank-sum test): ** $P < 0.01$; * $P < 0.05$. Note that responses of *D. sechellia* are stronger to noni fruit than noni juice, which may reflect the lower abundance of relevant ligands in the juice.

d, Representative odour-evoked calcium responses in the axon termini of Orco OSNs in the antennal lobe of *D. sechellia* (*UAS-GCaMP6f/UAS-GCaMP6f;;DsecOrco^{Gal4}/+*) acquired by two-photon imaging. Three focal planes are shown, revealing different glomeruli along the dorsoventral axis. Left column: raw fluorescence images. Other columns: relative increase in GCaMP6f fluorescence (F/F%) after stimulation with diagnostic odours. Scale bar, 25 µm.

e, Electrophysiological responses in the antennal coeloconic (ac) sensilla classes to the indicated stimuli (n = 6-11 sensilla, females) in *D. sechellia* (DSSC 14021-0248.07) representing the summed, solvent-corrected activities of the two or three neurons they house.

f, Representative odour-evoked calcium responses in the axon termini of Ir64a OSNs in the *D. sechellia* antennal lobe (genotype: *UAS-GCaMP6f/UAS-GCaMP6f;;DsecIr64a^{Gal4}/+*) acquired by widefield imaging. Left images: raw fluorescence signals. Right images: relative increase in GCaMP6f fluorescence (F/F%) after stimulation with noni juice (10^{-2}) or grape juice. Scale bar, 25 µm.

g, Quantification of odour-evoked calcium responses for the animals represented in **f**. Maximum calcium response amplitudes for each experiment are plotted (n = 7-10, females). Comparisons of responses to noni (10^{-2}) and grape juice are shown (Wilcoxon signed-rank test): * $P < 0.05$.
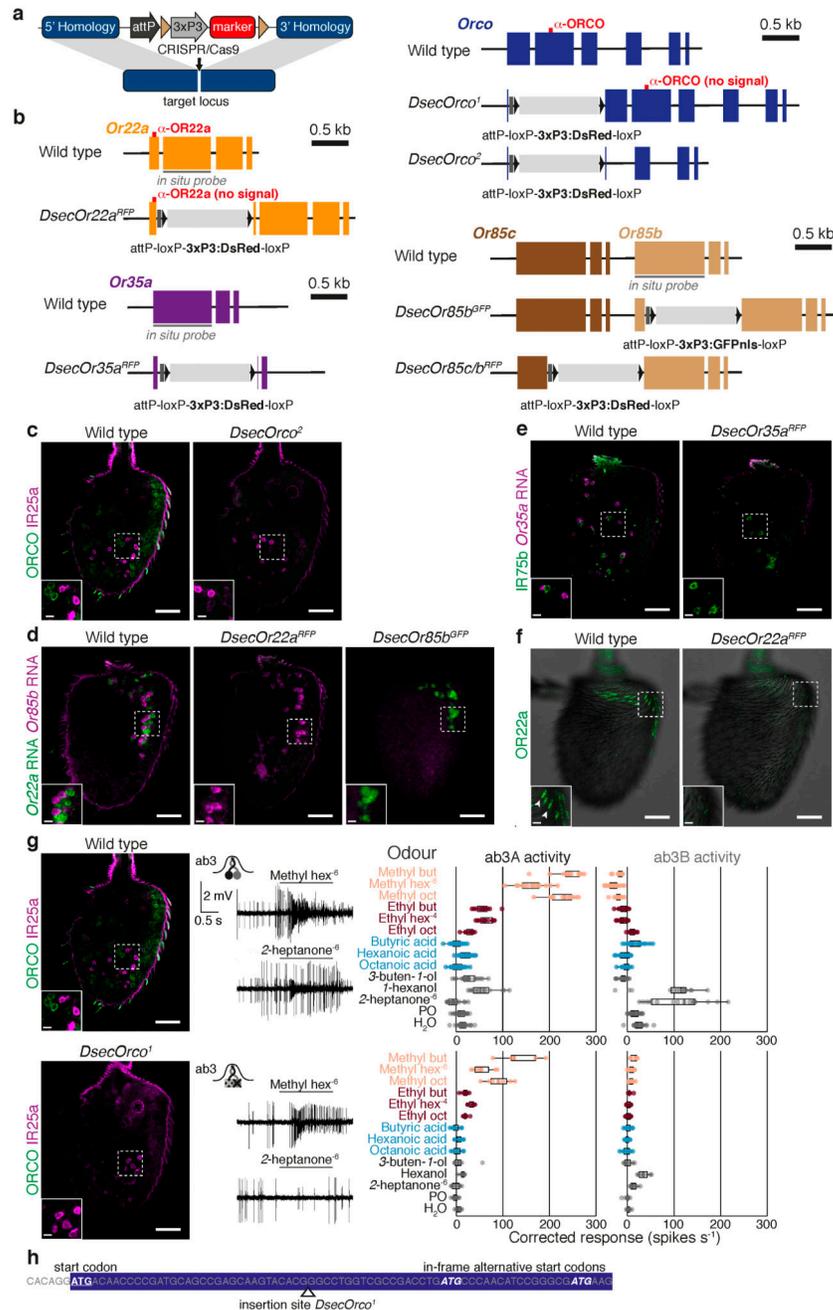
Extended Data Fig. 5. Generation and validation of loss-of-function alleles of *D. sechellia Or* genes.

a, Schematic of the strategy for generating olfactory receptor mutant alleles, through integration of an eye-expressed fluorescent marker (*3xP3:DsRed* or *3xP3:GFPnls*) into the desired locus via CRISPR/Cas9-cleavage induced homologous recombination. Brown triangles: *loxP* sites for removal of the fluorescent marker via Cre recombination.
b, Schematics depicting *Or* gene organisation, the structure of mutant alleles, and the location of the sequences encoding antibody epitopes. For *DsecOrco¹* the fluorescent marker

was integrated into the first coding exon; for *DsecOrco²* the marker replaces parts of exons 1 and 3 and the whole of exon 2. *DsecOr22a^{RFP}* carries the fluorescent marker in the first coding exon close to the start codon. *DsecOr35a^{RFP}* lacks most of exons 1 and 2. For *DsecOr85b^{GFP}* the marker was integrated into exon 1; for *DsecOr85c/b^{RFP}* the marker replaces most of the *Or85c* gene and part of exon 1 of *Or85b*.

c, Immunofluorescence for ORCO and IR25a (as internal staining control) on whole-mount antennae from wild-type and *DsecOrco²* animals. **c**-**g**: Scale bars, 25 μm; inset scale bars, 5 μm.

d, RNA FISH for *Or22a* and *Or85b* on whole-mount antennae from wild-type, *DsecOr22a^{RFP}* and *DsecOr85b^{GFP}* mutant animals.

e, Immunofluorescence for IR75b and RNA FISH for *Or35a* on whole-mount antennae from wild-type and *DsecOr35a^{RFP}* mutant animals. Arrowheads indicate *Or35a*-expressing cells. Note that Or35a neurons also pair with Ir75c neurons in ac3II sensilla[15], which is reflected in *Or35a*-positive cells that are not paired with IR75b-expressing cells in wild-type antennae.

f, Immunofluorescence for OR22a on whole-mount antennae from wild-type and *DsecOr22a^{RFP}* mutant animals. Arrowheads indicate sensilla housing Or22a neurons.

g, Left panel: immunofluorescence for ORCO and IR25a (as internal staining control) on whole-mount antennae from wild-type (same picture as shown in **c**) and *DsecOrco¹* animals. Central panel: electrophysiological responses in the two neurons of the ab3 sensillum (see Fig. 2a) to odours present in noni in *D. sechellia* (DSSC 14021-0248.07) and *DsecOrco¹* mutants (n = 5-20, females). Representative response traces to methyl hexanoate (10^{-6}) and *2*-heptanone (10^{-6}) are shown to the left. Data points represent the solvent-corrected activities per neuron. *D. sechellia* wild-type responses are replotted from Fig. 2a. Surprisingly, even though ORCO expression is undetectable by immunofluorescence, weak electrophysiological responses in ab3 sensilla (and other ORCO-dependent sensilla (data not shown)) can be detected. These observations suggest that trace levels of functional ORCO are produced from this allele, potentially through use of in-frame start codons downstream of the marker insertion site (see **h**).

h, A schematic depicting the location of the *Orco* start codon, the fluorescent marker insertion site of the *DsecOrco¹* allele and downstream potential alternative in-frame start codons.
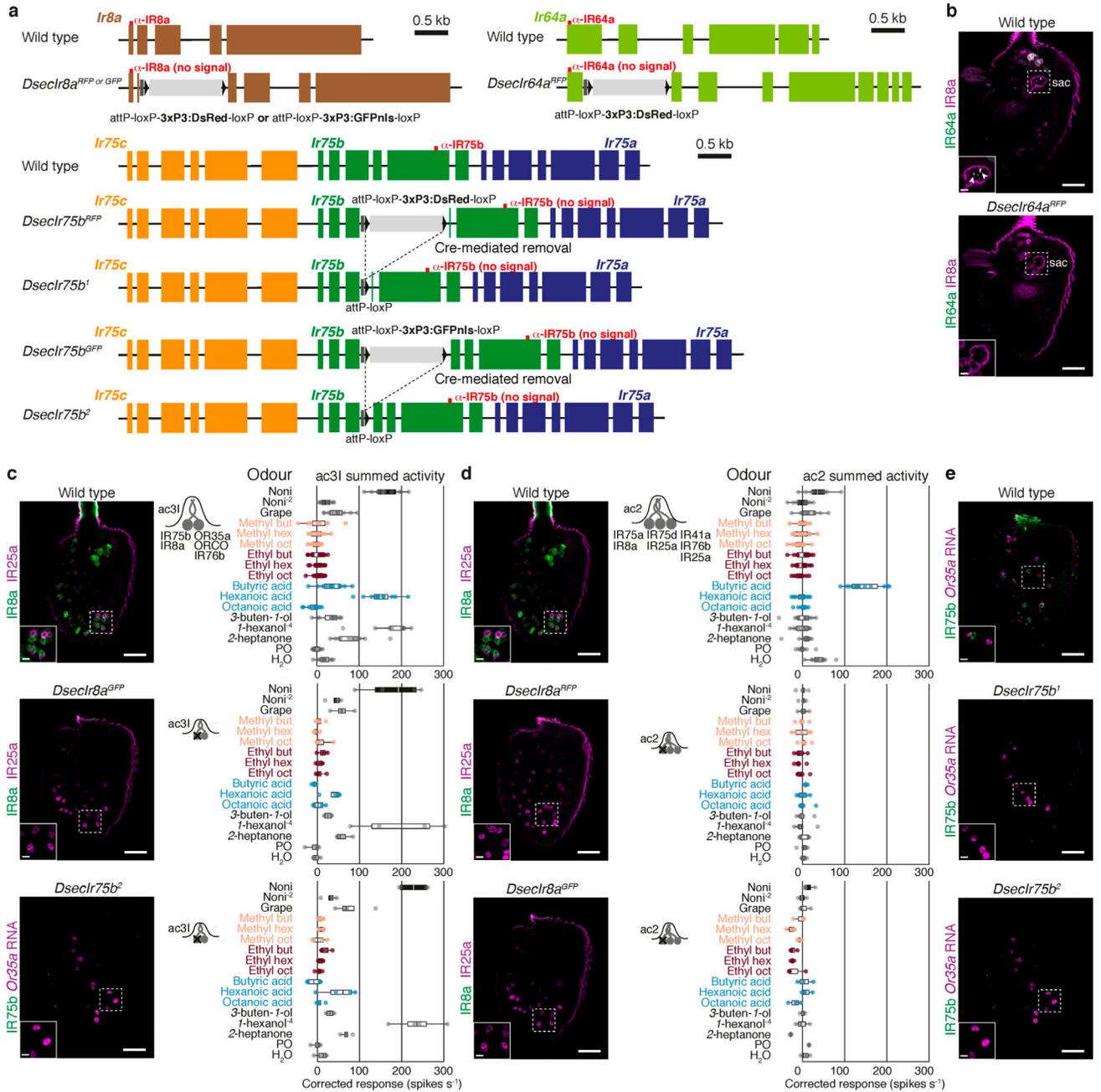
Extended Data Fig. 6. Generation and validation of loss-of-function alleles of *D. sechellia Ir* genes.

a, Schematics depicting *Ir* gene organisation, the structure of mutant alleles, and the sequences encoding antibody epitopes. For *DsecIr8a^RFP/GFP* the fluorescent marker was integrated into the first coding exon. For *DsecIr64a^RFP* the marker replaces parts of exon 2. *DsecIr75b^RFP* lacks parts of exons 3 and 4, while for *DsecIr75b^GFP* the marker was integrated into exon 3. For both alleles of *Ir75b*, the fluorophore was removed via Cre-mediated recombination to produce *Ir75b^1* and *Ir75b^2*.

b, Immunofluorescence for IR64a and IR8a (as internal staining control) on whole-mount antennae from wild-type and $DsecIr64a^{RFP}$ mutant animals. Arrowheads indicate the Ir64a neuron dendrites innervating sensilla in the sacculus (sac). Scale bars, 25 μm; inset scale bars, 5 μm.

c, Left panel (top): immunofluorescence for IR8a and IR25a (as internal staining control) on whole-mount antenna from wild-type and $DsecIr8a^{GFP}$ animals. Left panel (bottom): immunofluorescence for IR75b and RNA FISH for $Or35a$ on whole-mount antennae from $DsecIr75b^2$ mutant animals. Scale bars, 25 μm; inset scale bars, 5 μm. Right panel: electrophysiological responses in the ac3I sensillum (neurons housed are indicated in the cartoon) to noni juice, grape juice and odours present in noni (n = 4-11, females) in *D. sechellia* (DSSC 14021-0248.07) and olfactory receptor mutants affecting the Ir75b neuron ($DsecIr8a^{GFP}$, $DsecIr75b^2$). Data points represent the summed, solvent-corrected activities of the sensillum. *D. sechellia* wild-type responses are replotted from Fig. 2b. Note that the Or35a neuron exhibits residual responses to hexanoic acid in the $DsecIr8a$ and $DsecIr75b$ olfactory receptor mutants (see also Fig. 2b).

d, Left panel: immunofluorescence for IR8a and IR25a (as internal staining control) on whole-mount antennae from wild-type (same picture as shown in **c**) and $DsecIr8a^{RFP}$ and $DsecIr8a^{GFP}$ (same picture as shown in **c**) animals. Scale bars, 25 μm; inset scale bars, 5 μm. Right panel: electrophysiological responses in the ac2 sensillum (neurons housed are indicated in the cartoon) to noni juice, grape juice and odours present in noni (n = 3-11, females) in *D. sechellia* (DSSC 14021-0248.07) and receptor mutants affecting the Ir75a neuron ($DsecIr8a^{RFP}$, $DsecIr8a^{GFP}$). Data points represent the summed, solvent-corrected neuronal activities of the sensillum. *D. sechellia* wild-type responses to noni and grape juice are as shown in Extended Data Fig. 4e.

e, Immunofluorescence for IR75b and RNA FISH for $Or35a$ on whole-mount antennae from wild-type and $DsecIr75b^1$ and $DsecIr75b^2$ (same picture as shown in **c**) animals. Scale bars, 25 μm; inset scale bars, 5 μm.
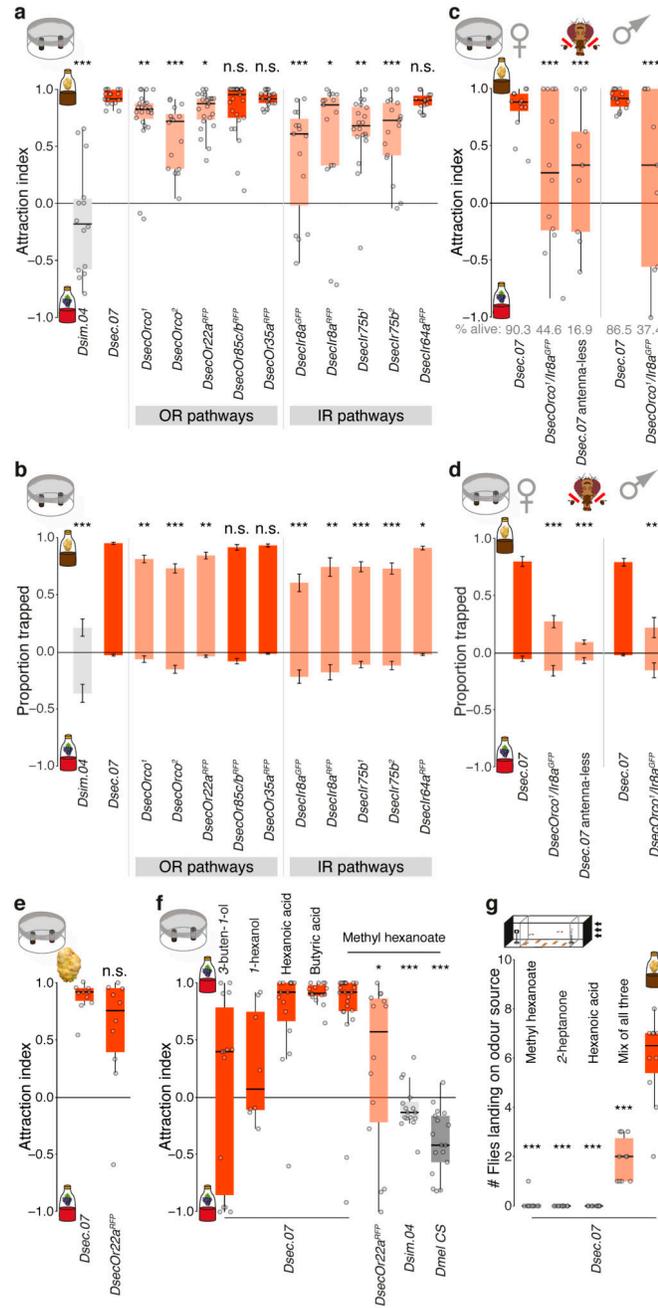
Extended Data Fig. 7. Genetic and chemical contributions promoting attraction of *D. sechellia* to noni.

a, Data reproduced from Fig. 2d. Behavioural responses in a trap assay testing preference of the indicated genotypes for noni juice or grape juice (n = 13-25 experiments; 22-25 females/experiment). Comparisons to *Dsec.07* responses are shown (**a-d** and **f**: pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): red bars, no significant difference; salmon bars, significantly different response; *** *P* < 0.001; ** *P* < 0.01; * *P* < 0.05; n.s. *P* > 0.05.

b, Proportion of flies (mean ± SEM) in each stimulus trap for the assays shown in **a**. Comparisons to *Dsec.07* responses are shown.

c, Olfactory responses in a trap assay testing preferences between noni juice and grape juice of *D. sechellia* (DSSC 14021-0248.07), *DsecOrco^1/Ir8a^{GFP}* double mutants and *D. sechellia* (DSSC 14021-0248.07) whose third antennal segments were removed (antenna-less) (n = 9-15 experiments; 22-25 females or males, as indicated/experiment). These data represent the same experiments shown in Fig. 2e, but attraction indices were calculated here taking only alive flies into account. The percentages of flies alive at the end of the assay are indicated below, revealing the high mortality rate of antenna-less flies and *DsecOrco^1/Ir8a^{GFP}* double mutants (*DsecOrco^2/Ir8a^{GFP}* mutants appeared to be non-viable). Normally, trap assay experiments with >25% animal mortality were discarded; see Methods. Comparisons to *Dsec.07* responses are shown.

d, Proportion of flies (mean ± SEM) in each stimulus trap for the assays shown in **c**. Comparisons to *Dsec.07* responses are shown.

e, Behavioural responses in a trap assay testing preferences between noni fruit and grape juice of *Dsec.07* and *DsecOr22a^{RFP}* flies. Comparisons to *Dsec.07* responses are shown (pairwise Wilcoxon rank-sum test): n.s. *P* > 0.05.

f, Behavioural responses in a trap assay testing preferences of *D. sechellia* (DSSC 14021-0248.07) between grape juice and 10^{-2} dilutions of the indicated odours in grape juice of *D. sechellia* (DSSC 14021-0248.07), *DsecOr22a^{RFP}*, *D. simulans* (DSSC 14021-0251.004) and *D. melanogaster* (Canton-S). Comparisons to *Dsec.07* responses to methyl hexanoate are shown.

g, Behavioural responses in a wind tunnel assay testing attraction of *D. sechellia* (DSSC 14021-0248.07) to three single noni odours (10^{-2} in water), a mix of all three in the approximate proportions of ripe noni fruit (1:0.04:1, methyl hexanoate:*2*-heptanone:hexanoic acid) and noni juice (n = 10 experiments; 10 females/experiment). Comparisons to *Dsec.07* responses to noni juice are shown (Kruskal-Wallis test with Dunn's post-hoc correction): *** *P* < 0.001.
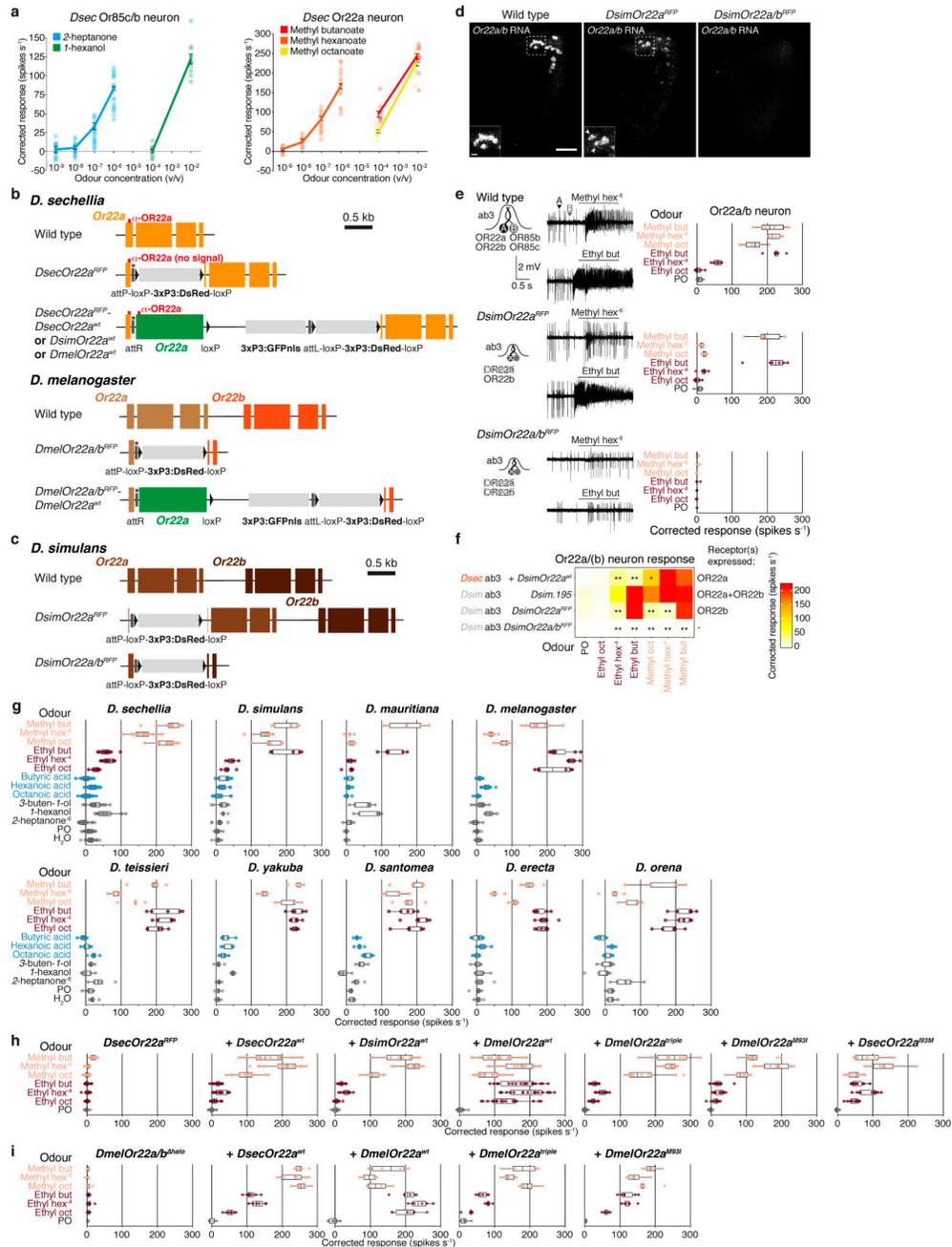
Extended Data Fig. 8. Odour-tuning properties of drosophilid Or85c/b and Or22a/(b) neurons and genomic modifications of the *Or22a/(b)* loci.

a, Left: dose-dependent electrophysiological responses of Or85c/b neurons (ab3B) in wild-type *D. sechellia* (DSSC 14021-0248.07) to *2*-heptanone and *1*-hexanol (mean ± SEM and individual data points; n = 10-20, females). Right: dose-dependent electrophysiological responses of Or22a neurons (ab3A) in wild-type *D. sechellia* (DSSC 14021-0248.07) to methyl butanoate, methyl hexanoate and methyl octanoate (mean ± SEM and individual data points; n = 10-20, females). The dose response curves for *2*-heptanone and methyl hexanoate are replotted from Fig. 3a.

b, Schematics depicting the arrangement of wild-type, mutant and rescue allele versions of *DsecOr22a* (top) and *DmelOr22a/Or22b* (bottom). Asterisk = stop codon preventing read through from the endogenous *Or22a* open reading frame.

c, Schematics depicting the arrangement of wild-type and mutant alleles of *DsimOr22a/Or22b*.

d, RNA FISH for *Or22a* on whole-mount antennae from wild-type *D. simulans* (DSSC 14021-0251.195), *DsimOr22a^{RFP}* and *DsimOr22a/b^{RFP}* mutant animals. As *Or22a* shares 85% sequence similarity with *Or22b*, the *Or22a* probe hybridises with transcripts from both genes. Arrowheads indicate *Or22b*-expressing cells in *DsimOr22a^{RFP}*. Scale bar, 25 μm; inset scale bar, 5 μm.

e, Electrophysiological responses of Or22a/b neurons to different esters in wild-type *D. simulans* (DSSC 14021-0251.195) and receptor mutants (*DsimOr22a^{RFP}*, *DsimOr22a/b^{RFP}*) (n = 6-10, females). Representative response traces to methyl hexanoate (10^{-6}) and ethyl butanoate (10^{-2}) are shown to the left.

f, Heat map representation of the data shown in **e**, together with the data of the *DsimOr22a^{wt}* response profile when expressed in *DsecOr22a^{RFP}* (replotted from Fig. 3c). The receptors expressed in the analysed neurons are listed to the right. Significant differences to *D. simulans* wild-type responses are shown (pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; n.s. $P > 0.05$. The equivalent responses to ethyl butanoate of *D. simulans* wild-type and *Or22a* mutant neurons (but complete loss in *Or22a/b* mutant neurons) suggests that this odour is detected principally by OR22b.

g, Boxplots with individual data points of the electrophysiological data presented in Fig. 3b.

h, Boxplots with individual data points of the electrophysiological data presented in Fig. 3c.

i, Boxplots with individual data points of the electrophysiological data presented in Fig. 3e.
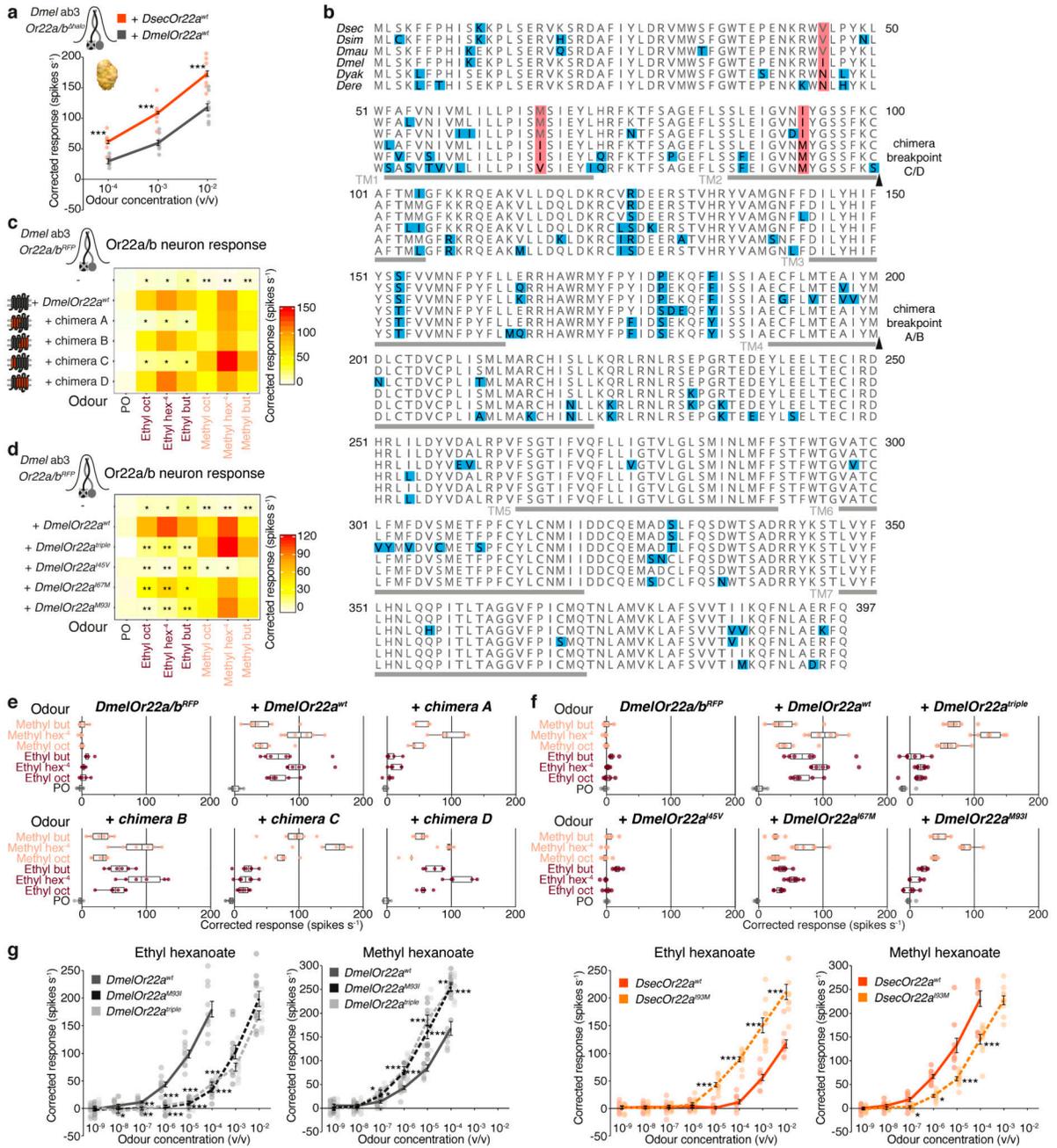
Extended Data Fig. 9. Mapping of odour-specificity determinants of OR22a.

a, Electrophysiological responses of *D. melanogaster Or22a/b* mutant neurons expressing *DsecOr22a^wt* or *DmelOr22a^wt* upon stimulation with increasing concentrations of noni fruit extract (mean ± SEM and individual data points; n = 9, females). Significantly different values are indicated (pairwise Wilcoxon rank-sum test): *** *P* < 0.001.

b, Protein sequence alignment of OR22a orthologues of six species within the *D. melanogaster* species subgroup. Red shading: amino acid differences between *D. melanogaster* and *D. sechellia*, *D. simulans* and *D. mauritiana* that were analysed by

mutagenesis in this study; blue shading: all other sequence differences. Arrowheads: chimera breakpoints (see **c**). Predicted transmembrane (TM) domains are indicated with grey lines (location as in[11]).

c, Electrophysiological responses to a panel of noni odours conferred by chimeric OR22a proteins encoded by transgenes integrated at the *Or22a/b* locus of *D. melanogaster* (n = 5-6, females). Schematics on the left indicate the relative proportions of *D. sechellia* (red) and *D. melanogaster* (dark grey) sequences in each chimera (precise chimera breakpoints are shown in **b**). Significant differences to *DmelOr22a^wt* responses are shown (**c**, **d** and **g**: pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; n.s. $P > 0.05$.

d, Electrophysiological responses of *D. melanogaster Or22a/b* mutant neurons expressing different OR22a variants (n = 5-7, females). The location of each mutated residue is indicated in **b**. Data for *Or22a/b* mutant and *DmelOr22a^wt* responses are replotted from **c**. Significant differences to *DmelOr22a^wt* responses are shown.

e, Boxplots with individual data points showing the same data as in **c**.

f, Boxplots with individual data points showing the same data as in **d**.

g, Dose-dependent electrophysiological responses of *D. sechellia* Or22a neurons expressing the indicated transgenes to ethyl or methyl hexanoate (mean ± SEM and individual data points; n = 10, females). Significant comparisons to either the *DmelOr22a^wt* (left graphs) or the *DsecOr22a^wt* (right graphs) transgene responses are shown.
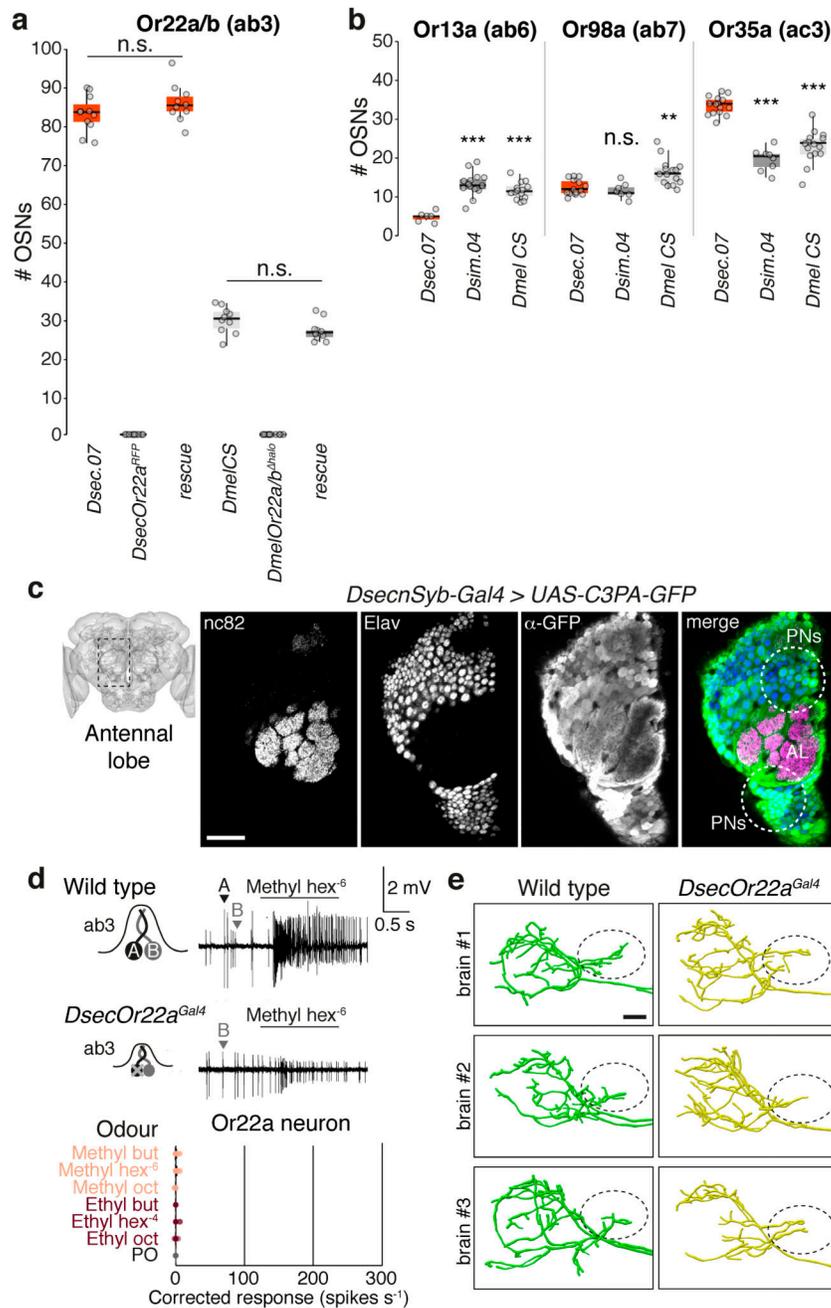
Extended Data Fig. 10. Peripheral and central olfactory circuit changes in *D. sechellia*.
a, Quantification of the number of OSNs expressing *Or22a/(b)* in antennae of wild-type *D. sechellia* (DSSC 14021-0248.07) and *D. melanogaster* (Canton-S) (data as shown in Fig. 4b), *Or22a/(b)* mutants in both species, and rescue lines expressing *DsecOr22a^wt^* (n = 9-11, females). Comparisons of rescue and wild-type genotypes for each species are shown (pairwise Wilcoxon rank-sum test). n.s. *P* > 0.05. No significant differences in *Or22a* cell number were observed for different rescue transgenes (data not shown).

b, Quantification of the number of OSNs expressing *Or13a* (ab6), *Or98a* (ab7) or *Or35a* (ac3I/II) in *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004), and *D. melanogaster* (Canton-S) (n = 10-15, females). Comparisons to *Dsec.07* cell number counts are shown (pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** *P* < 0.001; ** *P* < 0.01; n.s. *P* > 0.05.

c, Immunofluorescence with nc82 (neuropil), α-Elav (neurons) and α-GFP in a *DsecnSyb-Gal4/UAS-C3PA-GFP* transgenic line, which expresses photoactivatable GFP pan-neuronally. The schematic on the left indicates the region of image acquisition. An anterior section through the antennal lobe (AL) is shown to reveal the position of the labelled projection neuron (PN) soma (circled in the right panel). Scale bar, 25 μm.

d, Electrophysiological responses of the Or22a neuron to odours present in noni in homozygous *DsecOr22a^{Gal4}* (mutant) transgenic animals (n = 6, females). Data points represent the solvent-corrected activities. Representative response traces to methyl hexanoate (10^{-6}) in wild-type and transgenic animals are shown on top.

e, Tracing of axonal branches in the lateral horn of dye-filled DM2 PNs in *D. sechellia* wild-type and homozygous *DsecOr22a^{Gal4}* mutant flies. Three representative samples are shown. The circles depict the position of the *D. sechellia*-specific axonal branch. Scale bar, 10 μm. Samples could not be discriminated by genotype when presented to six independent researchers blindly.
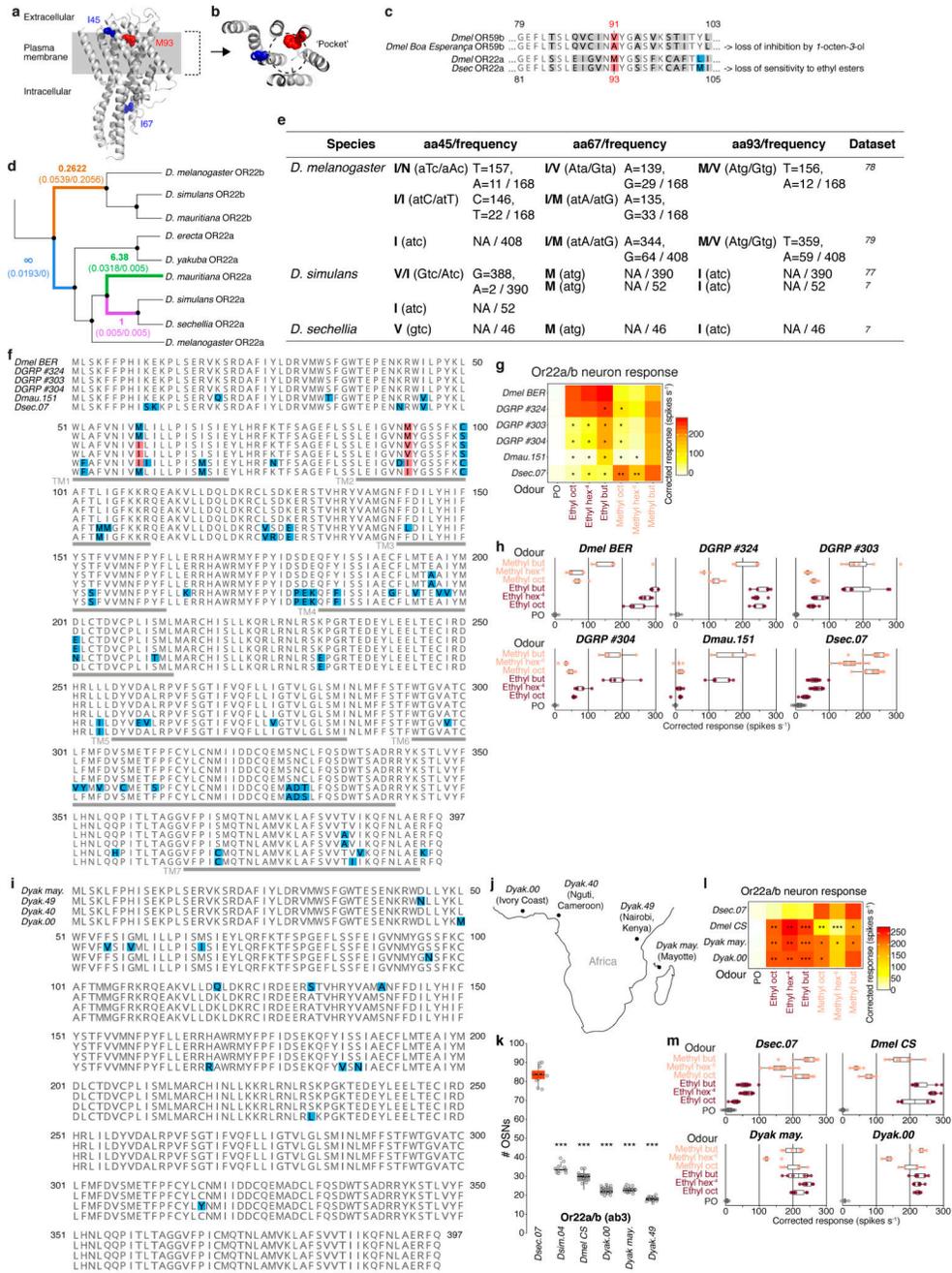
Extended Data Fig. 11. Phylogenetic and functional analysis of odour-specificity determinants in OR22a.

a, Side view of the ORCO monomer structure (determined by cryo-electronic microscopy[37]); the approximate location of the plasma membrane is indicated. The location of the residues corresponding to the odour-specificity determinants of OR22a analysed in this study (based on alignments generated in[37]) are highlighted as spheres.

b, Top view of a cross-section through the putative ligand-binding pocket of the ORCO structure shown in **a**.

c, Partial protein sequence alignment of OR22a and OR59b. The equivalent residue to *D. melanogaster* OR22a M93 in OR59b is V91, which exhibits intraspecific sequence variation impacting odour sensitivity[82].

d, Results of branch-based models of molecular evolution that tested for changes in the rates of protein evolution among OR22a and OR22b orthologues (see Methods and Supplementary Table 8): the rate of protein changes within the OR22a/OR22b phylogenetic tree highlights dN/dS ratios (ω) that differ from the "background rate" ($\omega = 0.1772$). Most branches exhibited low ω, arguing for strong purifying selection to maintain protein function over much of the tree. The two ω values that are >1 indicate an excess of protein changes, consistent with positive selection. The branch leading to *D. simulans* and *D. sechellia* OR22a displays nearly equal rates of silent and replacement substitutions, consistent with relaxed constraint during this period.

e, Allele frequencies within population datasets for *D. melanogaster*[78,79], *D. simulans*[7,77] and *D. sechellia*[7] at the three sites of OR22a that were functionally characterised in this study. The table displays amino acid (aa) positions 45, 67 and 93 of OR22a and the frequencies at which variants within the corresponding codons are segregating (number of alleles with respective variant/number of alleles analysed). "NA" (not applicable) indicates that positions within the codon are invariant. Datasets analysed are referenced on the right. Selected *Or22a* variants from the *Drosophila melanogaster* Genetic Reference Panel (DGRP)[79] were confirmed by sequencing (**f**) and Or22a neuron physiology analysed (**g**,**h**).

f, Protein sequence alignment of OR22a orthologues of *D. melanogaster*[83], three lines of the DGRP, *D. mauritiana* (DSSC 14021-0241.151) and *D. sechellia* (DSSC 14021-0248.07). Red shading: shared amino acid differences of *DGRP #303*, *DGRP #304* and *D. mauritiana* (compared to other sequences) at position 59 and the key odour specificity-determinant at residue 93; blue shading: all other sequence differences. No line within the DGRP with a polymorphism only at position 93 was identified.

g, Electrophysiological responses of the Or22a/b neuron to odours present in noni (n = 5-20, females) in the strains shown in **f**. The similarity between the response profiles of *DGRP #303*, *DGRP #304* and *D. mauritiana* suggests that their only shared polymorphism (at position 59) modifies OR22a response properties in these strains. Comparisons to *Dmel BER* responses are shown (pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** $P < 0.001$; ** $P < 0.01$; n.s. $P > 0.05$. *D. mauritiana* and *D. sechellia* data are replotted from Fig. 3b.

h, Boxplots with individual data points showing the same data as in **g**. *D. mauritiana* and *D. sechellia* data are replotted from Extended Data Fig. 8g.

i, Protein sequence alignment of OR22a orthologues of the noni-specialised *D. yakuba mayottensis*[41] (*Dyak may.*) and three other *D. yakuba* strains (DSSC 14021-0261.00, 14021-0261.40, 14021-0261.49). Blue shading: differences between these sequences.

j, Collection sites of *D. yakuba* strains shown in **i**.

k, Quantification of the number of OSNs expressing *Or22a/(b)* in *D. sechellia* (DSSC 14021-0248.07), *D. simulans* (DSSC 14021-0251.004), *D. melanogaster* (Canton-S) (data as shown in Fig. 4b), *D. yakuba* (DSSC 14021-0261.00, 14021-0261.49) and *D. yakuba mayottensis* (n = 10-12, females). Comparisons to *Dsec.07* cell number counts are shown (**k** and **l**: pairwise Wilcoxon rank-sum test and *P* values adjusted for multiple comparisons using the B&H method): *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; n.s. $P > 0.05$.

l, Electrophysiological responses to odours present in noni of the Or22a/(b) neurons in *D. sechellia* (DSSC 14021-0248.07), *D. melanogaster* (Canton-S), *D. yakuba mayottensis* (n = 5-20, females) and *D. yakuba* (DSSC 14021-0261.00). Comparisons to *Dsec.07* responses are shown. *D. sechellia*, *D. melanogaster* and *D. yakuba.00* data are replotted from Fig. 3b. **m**, Boxplots with individual data points showing the same data as in **l**. *D. sechellia*, *D. melanogaster* and *D. yakuba.00* data are replotted from Extended Data Fig. 8g.
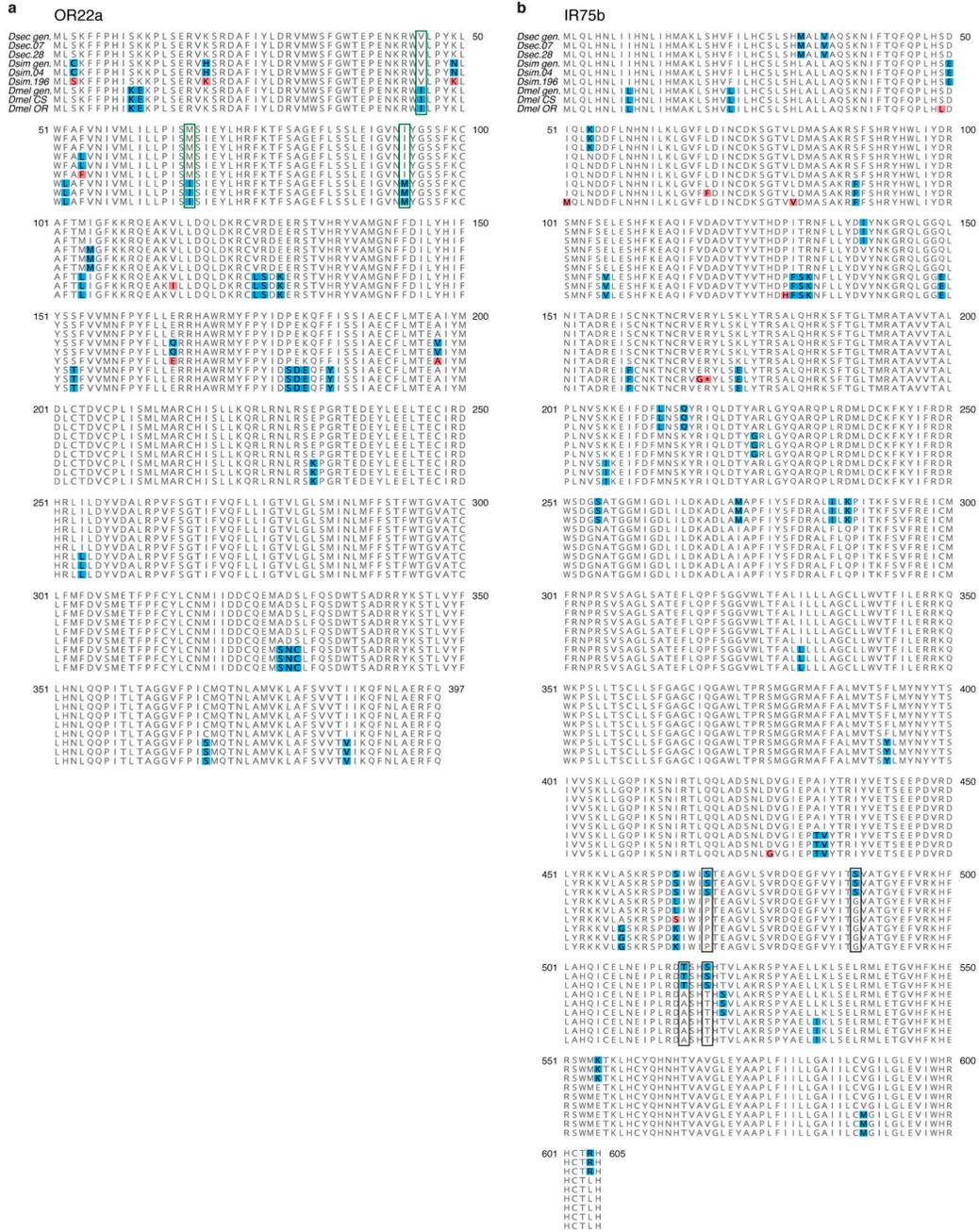
Extended Data Fig. 12. Protein sequence alignments of OR22a and IR75b.

a, OR22a orthologues of strains used for behavioural assays as well as genome-sequenced strains ("*gen.*") of each species (version: *D. sechellia* (r1.3), *D. simulans* (r2.01), *D. melanogaster* (r6.28)). Blue/red shading: differences between species/strains. Green boxes: residues tested in this study for their role in defining ester tuning specificity.

b, IR75b orthologues of same strains as in **a**. Blue/red shading: differences between species/strains. Black boxes: residues predicted to be located within the ligand-binding domain that contribute to odour tuning specificity[63]. Note that the premature stop codon of the Canton-S

strain (position 169, marked by an asterisk) does not impair receptor function, as shown in other strains[63].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Arguello JR, Benton R. Open questions: Tackling Darwin's "instincts": the genetic basis of behavioural evolution. BMC Biol. 2017; 15:26. [PubMed: 28372547]

2. Bendesky A, Bargmann CI. Genetic contributions to behavioural diversity at the gene-environment interface. Nat Rev Genet. 2011; 12:809–820. [PubMed: 22064512]

3. Bendesky A, et al. The genetic basis of parental care evolution in monogamous mice. Nature. 2017; 544:434–439. [PubMed: 28424518]

4. Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. Nature. 2016; 536:329–332. [PubMed: 27509856]

5. Weber JN, Peterson BK, Hoekstra HE. Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. Nature. 2013; 493:402–405. [PubMed: 23325221]

6. Garrigan D, et al. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. Genome Res. 2012; 22:1499–1511. [PubMed: 22534282]

7. Schrider DR, Ayroles J, Matute DR, Kern AD. Supervised machine learning reveals introgressed loci in the genomes of *Drosophila simulans* and *D. sechellia*. Plos Genetics. 2018; 14

8. Amlou M, Moreteau B, David JR. Genetic analysis of *Drosophila sechellia* specialization: Oviposition behaviour toward the major aliphatic acids of its host plant. Behav Genet. 1998; 28:455–464. [PubMed: 9926614]

9. Cobb M, Burnet B, Blizard R, Jallon JM. Courtship in *Drosophila sechellia* - its structure, functional aspects, and relationship to those of other members of the *Drosophila melanogaster* species subgroup. J Insect Behav. 1989; 2:63–89.

10. Coyne JA. Genetics of sexual isolation in females of the *Drosophila simulans* species complex. Genet Res. 1992; 60:25–31. [PubMed: 1452013]

11. Dekker T, Ibba I, Siju KP, Stensmyr MC, Hansson BS. Olfactory shifts parallel superspecialism for toxic fruit in *Drosophila melanogaster* sibling, *D. sechellia*. Curr Biol. 2006; 16:101–109. [PubMed: 16401429]

12. Higa I, Fuyama Y. Genetics of Food Preference in *Drosophila sechellia*. 1. Responses to Food Attractants. Genetica. 1993; 88:129–136. [PubMed: 8224853]

13. Ibba I, Angioy AM, Hansson BS, Dekker T. Macroglomeruli for fruit odours change blend preference in *Drosophila*. Die Naturwissenschaften. 2010; 97:1059–1066. [PubMed: 20972770]

14. Matsuo T, Sugaya S, Yasukawa J, Aigaki T, Fuyama Y. Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. PLoS Biol. 2007; 5:e118. [PubMed: 17456006]

15. Prieto-Godino LL, et al. Evolution of acid-sensing olfactory circuits in *Drosophilids*. Neuron. 2017; 93:661–676. [PubMed: 28111079]

16. R'Kha S, Capy P, David JR. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioural, and genetical analysis. Proc Natl Acad Sci U S A. 1991; 88:1835–1839. [PubMed: 1900368]

17. Yalcin B, et al. Genetic dissection of a behavioural quantitative trait locus shows that Rgs2 modulates anxiety in mice. Nat Genet. 2004; 36:1197–1202. [PubMed: 15489855]

18. Bendesky A, Tsunozaki M, Rockman MV, Kruglyak L, Bargmann CI. Catecholamine receptor polymorphisms affect decision-making in *C. elegans*. Nature. 2011; 472:313–318. [PubMed: 21412235]

19. Bumbarger DJ, Riebesell M, Rodelsperger C, Sommer RJ. System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes. Cell. 2013; 152:109–119. [PubMed: 23332749]

20. Markow TA, O'Grady P. Reproductive ecology of *Drosophila*. Functional Ecology. 2008; 22:747–759.

21. Linz J, et al. Host plant-driven sensory specialization in *Drosophila erecta*. Proc Biol Sci. 2013; 280

22. Seeholzer LF, Seppo M, Stern DL, Ruta V. Evolution of a central neural circuit underlies *Drosophila* mate preferences. Nature. 2018; 559:564–569. [PubMed: 29995860]

23. Dweck HK, et al. Olfactory channels associated with the *Drosophila* maxillary palp mediate short- and long-range attraction. Elife. 2016; 5

24. Matute DR, Ayroles JF. Hybridization occurs between *Drosophila simulans* and *D. sechellia* in the Seychelles archipelago. J Evol Biol. 2014; 27:1057–1068. [PubMed: 24773151]

25. Vosshall LB, Stocker RF. Molecular architecture of smell and taste in *Drosophila*. Annu Rev Neurosci. 2007; 30:505–533. [PubMed: 17506643]

26. Larsson MC, et al. *Or83b* encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. Neuron. 2004; 43:703–714. [PubMed: 15339651]

27. Abuin L, et al. Functional architecture of olfactory ionotropic glutamate receptors. Neuron. 2011; 69:44–60. [PubMed: 21220098]

28. Benton R, Sachse S, Michnick SW, Vosshall LB. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors *in vivo*. PLoS Biol. 2006; 4:e20. [PubMed: 16402857]

29. Stensmyr MC, Dekker T, Hansson BS. Evolution of the olfactory code in the *Drosophila melanogaster* subgroup. Proc Biol Sci. 2003; 270:2333–2340. [PubMed: 14667348]

30. Dobritsa AA, van der Goes van Naters W, Warr CG, Steinbrecht RA, Carlson JR. Integrating the molecular and cellular basis of odour coding in the *Drosophila* antenna. Neuron. 2003; 37:827–841. [PubMed: 12628173]

31. Couto A, Alenius M, Dickson BJ. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. Curr Biol. 2005; 15:1535–1547. [PubMed: 16139208]

32. 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. 2015; 7:2843–2858. [PubMed: 26430061]

33. Yao CA, Ignell R, Carlson JR. Chemosensory coding by neurons in the coeloconic sensilla of the *Drosophila* antenna. J Neurosci. 2005; 25:8359–8367. [PubMed: 16162917]

34. Ai M, et al. Acid sensing by the *Drosophila* olfactory system. Nature. 2010; 468:691–695. [PubMed: 21085119]

35. Ruta V, et al. A dimorphic pheromone circuit in *Drosophila* from sensory input to descending output. Nature. 2010; 468:686–690. [PubMed: 21124455]

36. Grabe V, Sachse S. Fundamental principles of the olfactory code. Bio Systems. 2018; 164:94–101.

37. Butterwick JA, et al. Cryo-EM structure of the insect olfactory receptor ORCO. Nature. 2018; 560:447–452. [PubMed: 30111839]

38. Aguade M. Nucleotide and copy-number polymorphism at the odorant receptor genes *Or22a* and *Or22b* in *Drosophila melanogaster*. Mol Biol Evol. 2009; 26:61–70. [PubMed: 18922763]

39. Goldman-Huertas B, et al. Evolution of herbivory in *Drosophilidae* linked to loss of behaviours, antennal responses, odorant receptors, and ancestral diet. Proc Natl Acad Sci U S A. 2015; 112:3026–3031. [PubMed: 25624509]

40. Nozawa M, Nei M. Evolutionary dynamics of olfactory receptor genes in *Drosophila* species. Proc Natl Acad Sci U S A. 2007; 104:7122–7127. [PubMed: 17438280]

41. Yassin A, et al. Recurrent specialization on a toxic fruit in an island *Drosophila* population. Proc Natl Acad Sci U S A. 2016; 113:4771–4776. [PubMed: 27044093]

42. Ambrose D, Ellender JH, Lee EB, Sprake CHS, Townsend R. Thermodynamic properties of organic oxygen compounds XXXVIII. Vapour pressures of some aliphatic ketones. The Journal of Chemical Thermodynamics. 1975; 7:453–472.

43. Daubert TE, Danner RP. Physical and thermodynamic properties of pure chemicals data compilation. Data compilation. 1997

44. Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. Proc Natl Acad Sci U S A. 2014

45. Port F, Bullock SL. Augmenting CRISPR applications in *Drosophila* with tRNA-flanked sgRNAs. Nat Methods. 2016; 13:852–854. [PubMed: 27595403]

46. Barolo S, Carver LA, Posakony JW. GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. Biotechniques. 2000; 29:726–732. [PubMed: 11056799]

47. Gratz SJ, et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. Genetics. 2014; 196:961–971. [PubMed: 24478335]

48. Stern DL. Tagmentation-Based Mapping (TagMap) of Mobile DNA Genomic Insertion Sites. bioRxiv. 2016; doi: 10.1101/037762

49. Croset V, et al. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. PLoS Genet. 2010; 6:e1001064. [PubMed: 20808886]

50. Bateman JR, Wu CT. A simple polymerase chain reaction-based method for the construction of recombinase-mediated cassette exchange donor vectors. Genetics. 2008; 180:1763–1766. [PubMed: 18791223]

51. Groth AC, Olivares EC, Thyagarajan B, Calos MP. A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A. 2000; 97:5995–6000. [PubMed: 10801973]

52. Bischof J, Maeda RK, Hediger M, Karch F, Basler K. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A. 2007; 104:3312–3317. [PubMed: 17360644]

53. Han C, Jan LY, Jan YN. Enhancer-driven membrane markers for analysis of nonautonomous mechanisms reveal neuron-glia interactions in *Drosophila*. Proc Natl Acad Sci U S A. 2011; 108:9673–9678. [PubMed: 21606367]

54. Arnoult L, et al. Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. Science. 2013; 339:1423–1426. [PubMed: 23520110]

55. Gratz SJ, et al. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. Genetics. 2013; 194:1029–1035. [PubMed: 23709638]

56. Zhang X, Koolhaas WH, Schnorrer F. A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in *Drosophila*. G3 (Bethesda). 2014; 4:2409–2418. [PubMed: 25324299]

57. Gohl DM, et al. A versatile *in vivo* system for directed dissection of gene expression patterns. Nat Methods. 2011; 8:231–237. [PubMed: 21473015]

58. Knecht ZA, et al. Distinct combinations of variant ionotropic glutamate receptors mediate thermosensation and hygrosensation in *Drosophila*. Elife. 2016; 5

59. Silbering AF, et al. Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. J Neurosci. 2011; 31:13357–13375. [PubMed: 21940430]

60. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012; 9:676–682. [PubMed: 22743772]

61. Benton R, Dahanukar A. Electrophysiological recording from *Drosophila* olfactory sensilla. Cold Spring Harb Protoc. 2011; 2011:824–838. [PubMed: 21724819]

62. Saina M, Benton R. Visualizing olfactory receptor expression and localization in *Drosophila*. Methods Mol Biol. 2013; 1003:211–228. [PubMed: 23585045]

63. Prieto-Godino LL, et al. Olfactory receptor pseudo-pseudogenes. Nature. 2016; 539:93–97. [PubMed: 27776356]

64. Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. Cell. 2009; 136:149–162. [PubMed: 19135896]

65. Sanchez-Alcaniz JA, Zappia G, Marion-Poll F, Benton R. A mechanosensory receptor required for food texture detection in *Drosophila*. Nat Commun. 2017; 8

66. Ostrovsky A, Cachero S, Jefferis G. Clonal analysis of olfaction in *Drosophila*: immunochemistry and imaging of fly brains. Cold Spring Harb Protoc. 2013; 2013:342–346. [PubMed: 23547149]

67. Jefferis GS, et al. Comprehensive maps of *Drosophila* higher olfactory centers: spatially segregated fruit and pheromone representation. Cell. 2007; 128:1187–1203. [PubMed: 17382886]

68. Manton JD, et al. Combining genome-scale *Drosophila* 3D neuroanatomical data by bridging template brains. bioRxiv. 2014; doi: 10.1101/006353

69. Cachero S, Ostrovsky AD, Yu JY, Dickson BJ, Jefferis GS. Sexual dimorphism in the fly brain. Curr Biol. 2010; 20:1589–1601. [PubMed: 20832311]

70. Feng L, Zhao T, Kim J. neuTube 1.0: A New Design for Efficient Neuron Reconstruction Software Based on the SWC Format. eNeuro. 2015; 2

71. Caron SJ, Ruta V, Abbott LF, Axel R. Random convergence of olfactory inputs in the *Drosophila* mushroom body. Nature. 2013; 497:113–117. [PubMed: 23615618]

72. Camacho C, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009; 10:421. [PubMed: 20003500]

73. Birney E, Clamp M, Durbin R. GeneWise and Genomewise. Genome Res. 2004; 14:988–995. [PubMed: 15123596]

74. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 2010; 38:W7–13. [PubMed: 20435676]

75. Ronquist F, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61:539–542. [PubMed: 22357727]

76. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007; 24:1586–1591. [PubMed: 17483113]

77. Signor SA, New FN, Nuzhdin S. A large panel of *Drosophila simulans* reveals an abundance of common variants. Genome Biol Evol. 2018; 10:189–206. [PubMed: 29228179]

78. Grenier JK, et al. Global diversity lines - a five-continent reference panel of sequenced *Drosophila melanogaster* strains. G3 (Bethesda). 2015; 5:593–603. [PubMed: 25673134]

79. Huang W, et al. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. Genome Res. 2014; 24:1193–1208. [PubMed: 24714809]

80. Danecek P, et al. The variant call format and VCFtools. Bioinformatics. 2011; 27:2156–2158. [PubMed: 21653522]

81. Grabe V, et al. Elucidating the neuronal architecture of olfactory glomeruli in the *Drosophila* antennal lobe. Cell Rep. 2016; 16:3401–3413. [PubMed: 27653699]

82. Pellegrino M, Steinbach N, Stensmyr MC, Hansson BS, Vosshall LB. A natural polymorphism alters odour and DEET sensitivity in an insect odorant receptor. Nature. 2011; 478:511–514. [PubMed: 21937991]

83. Adams MD, et al. The genome sequence of *Drosophila melanogaster*. Science. 2000; 287:2185–2195. [PubMed: 10731132]
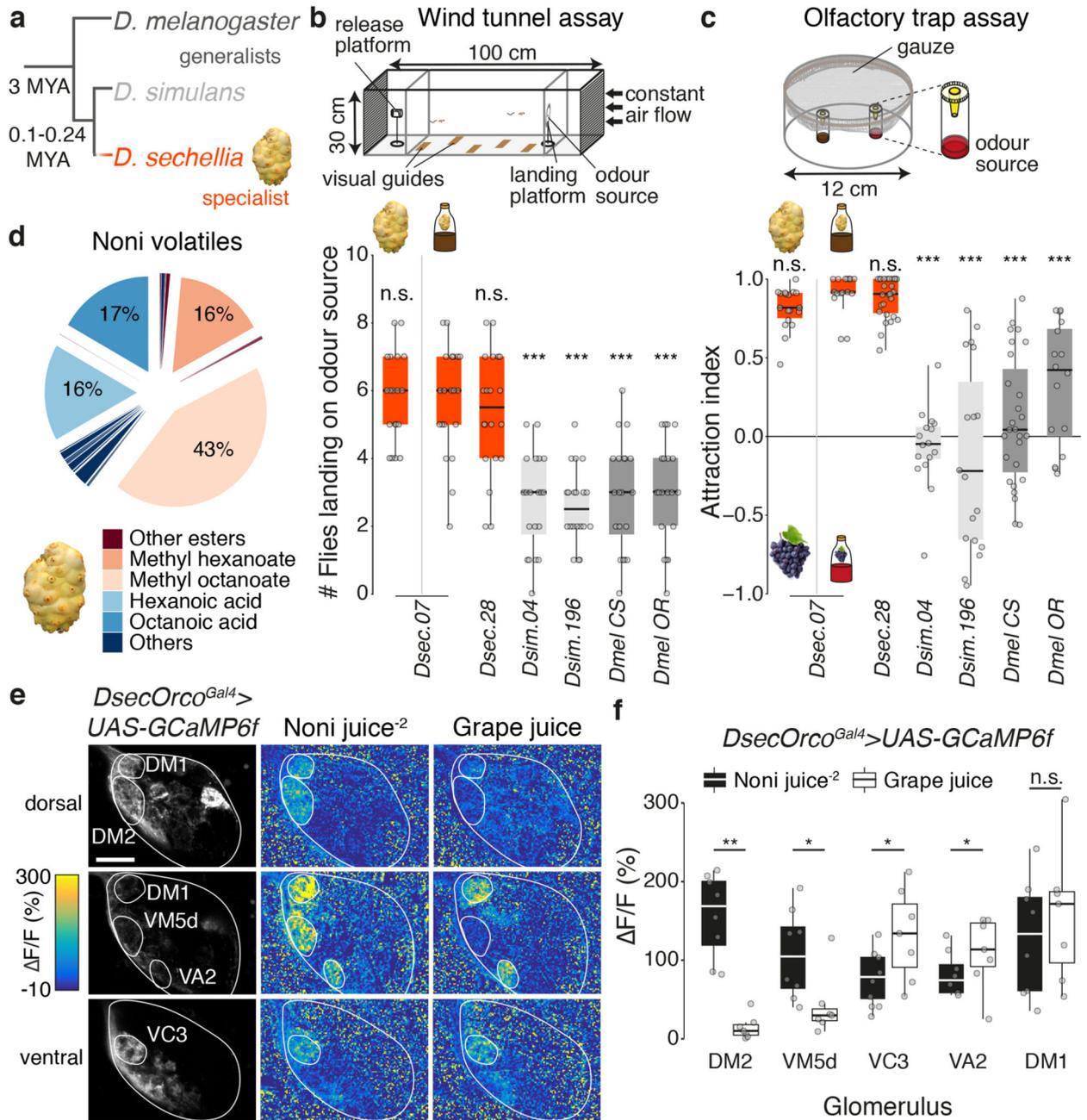
Fig. 1. Behavioural and physiological responses of *D. sechellia* to noni.

a, *D. sechellia* specialises on noni fruit (*Morinda citrifolia*) while *D. simulans* and *D. melanogaster* are food generalists (MYA = million years ago).

b, Behavioural responses to noni fruit or juice in a wind tunnel assay of *D. sechellia*, *D. simulans* and *D. melanogaster* wild-type strains (n = 20 experiments (10 females/experiment)). In this and other panels, boxplots show the median, first and third quartile of the data, overlaid with individual data points. Comparisons to *Dsec.07* noni juice responses are shown (Kruskal-Wallis test, Dunn's post-hoc correction): *** $P < 0.001$; n.s. $P > 0.05$.

c, Behavioural responses in a trap assay testing preferences between noni and grape, or between noni juice and grape juice (same strains as in **b**, n = 15-27 experiments, exact n in Source Data, (22-25 females/experiment)). Comparisons to *Dsec.07* noni juice responses are shown (pairwise Wilcoxon rank-sum test, *P* values adjusted for multiple comparisons using the Benjamini and Hochberg (B&H) method): *** *P* < 0.001; n.s. *P* > 0.05.

d, Odour bouquet of a ripe noni fruit determined by gas chromatography/mass spectrometry (see Methods, Extended Data Fig. 2, Supplementary Table 1).

e, Representative odour-evoked calcium responses in the axon termini of Orco OSNs in the *D. sechellia* antennal lobe (genotype: *UAS-GCaMP6f/UAS-GCaMP6f;;DsecOrco^{Gal4}/+*) acquired by two-photon imaging. Three focal planes are shown, revealing different glomeruli (outlined) along the dorsoventral axis. Left column: raw fluorescence images. Right columns: relative increase in GCaMP6f fluorescence (F/F%) after stimulation with noni juice (10^{-2} in H_2O) or grape juice. Scale bar, 25 μm.

f, Quantification of responses for the animals represented in **e**. Maximum response amplitudes for each experiment are plotted (n = 7-10 females). Wilcoxon signed-rank test: ** *P* < 0.01; * *P* < 0.05; n.s. *P* > 0.05.
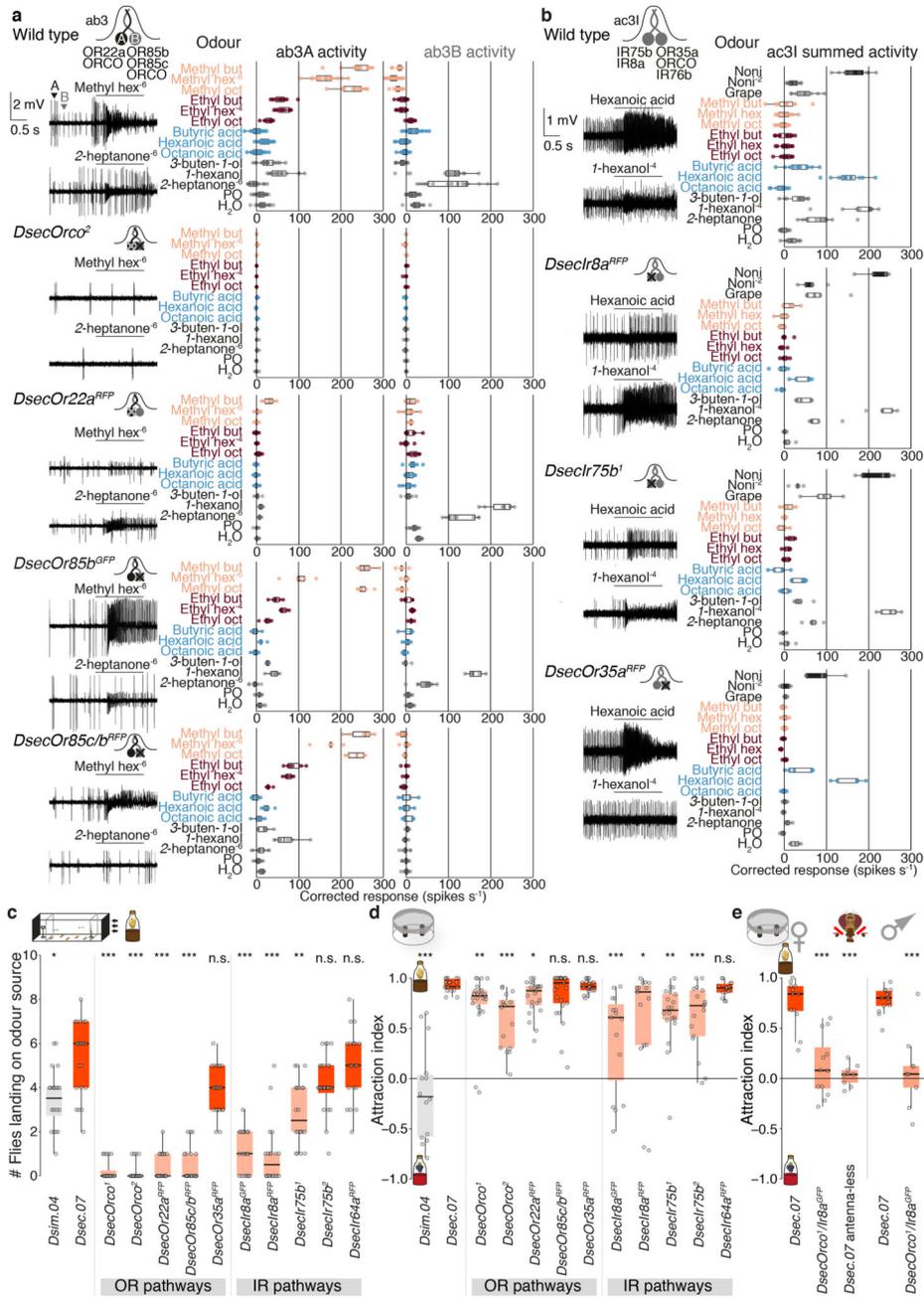
Fig. 2. Olfactory receptors contribution to noni-sensing.

a, Electrophysiological responses of ab3 sensillum neurons to noni odours (n = 5-20, females, see Supplementary Table 7 for exact n and mean spike counts) in wild-type and receptor mutant *D. sechellia* (schematised in the cartoons), with representative traces for methyl hexanoate and *2*-heptanone. Data points are the solvent-corrected activities of individual neurons (arrowheads in wild-type traces). Odours (oct, octanoate; hex, hexanoate; but, butanoate) are coloured according to chemical class: methyl esters (salmon), ethyl esters

(dark red), acids (light blue), others (black), used at 10^{-2} v/v unless indicated otherwise. PO = paraffin oil.

b, Responses of ac3I neurons to noni juice, grape juice and noni odours (n = 5-11, females) in wild-type and receptor mutant *D. sechellia*, with representative traces for hexanoic acid and *1*-hexanol. Data points are the summed, solvent-corrected activities of both neurons.

c, Behavioural responses to noni juice in the wind tunnel assay (n = 20 experiments). Comparisons to *Dsec.07* responses are shown (Kruskal-Wallis test, Dunn's post-hoc correction). **c, d, e**: red = no significant difference; salmon = significantly different response; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; n.s. $P > 0.05$.

d, Behavioural responses in the trap assay testing preference for noni juice or grape juice (n = 13-25 experiments). Comparisons to *Dsec.07* responses are shown (**d** and **e**, pairwise Wilcoxon rank-sum test, *P* values adjusted for multiple comparisons using the B&H method).

e, Behavioural responses in the trap assay testing preference for noni juice or grape juice using wild-type *D. sechellia*, *DsecOrco¹/Ir8a^{GFP}* double mutants and antenna-less *D. sechellia* (n = 9-15 experiments (22-25 females or males, as indicated/experiment)). Average attraction indices for *DsecOrco¹/Ir8a^{GFP}* and antenna-less flies are not significantly different from zero (pairwise Wilcoxon rank-sum test).
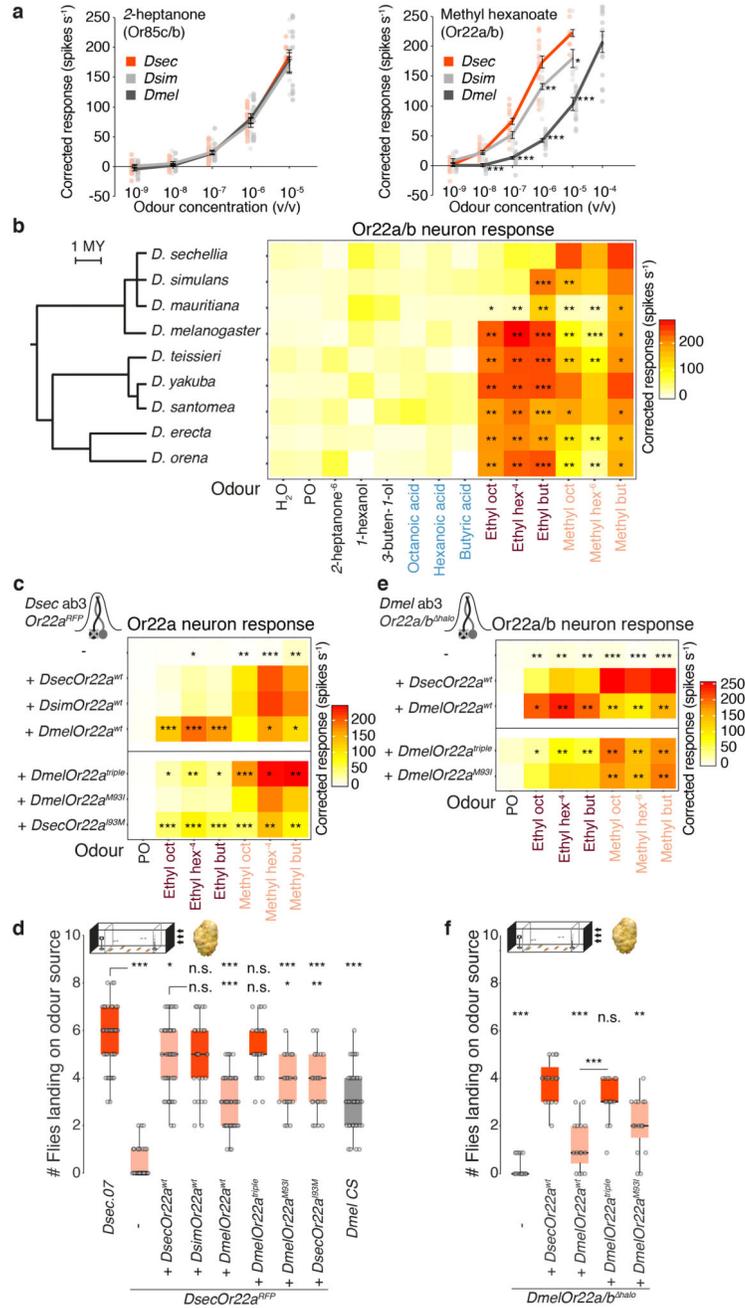
Fig. 3. Tuning of OR22a is important for noni attraction.

a, Dose-dependent responses of Or85c/b (left) and Or22a/(b) neurons (right) in the indicated species to *2*-heptanone and methyl hexanoate, respectively (mean ± SEM and individual data points, n = 7-20, females). Significant differences to *D. sechellia* responses are shown (**a**, **b**, **c** and **e**: pairwise Wilcoxon rank-sum test, *P* values adjusted for multiple comparisons using the B&H method): *** *P* < 0.001; ** *P* < 0.01; * *P* < 0.05.

b, Responses of Or22a/(b) neurons to noni odours across the *D. melanogaster* species subgroup; MY = million years (n = 5-20, females). *D. sechellia* responses as in Fig. 2a. Significant differences to *D. sechellia* ester responses are shown.

c, Responses of *D. sechellia* Or22a neurons expressing wild-type (top) or mutant (bottom) versions of *Or22a* at the *Or22a* locus (n = 10-11, females). Significant differences to *Dsec*OR22a^{wt} responses are shown (**c** and **e**). Responses to methyl hexanoate at 10^{-4} here compared to 10^{-6} in **b** and **e**.

d, Behavioural responses to noni fruit in the wind tunnel assay (n = 25-45 experiments). Comparisons to *Dsec.07* (top) and *Dsec*OR22a^{wt} (below) responses are shown (**d** and **f**, Kruskal-Wallis test, Dunn's post-hoc correction). Salmon = *D. sechellia* genotypes with significantly different response to *Dsec.07*; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$.

e, Responses of *D. melanogaster Or22a/b* mutant neurons expressing wild-type (top) or mutant (bottom) versions of *Or22a* (n = 5-7, females). Boxplot representations of **b**, **c** and **e** are shown in Extended Data Fig. 8g, h and i, respectively.

f, Behavioural responses to noni fruit in the wind tunnel assay (n = 20 experiments). Comparisons to *DsecOr22a^{wt}* (top) and *Dmel*OR22a^{wt} (below) responses are shown. Salmon = genotypes with significantly different responses.
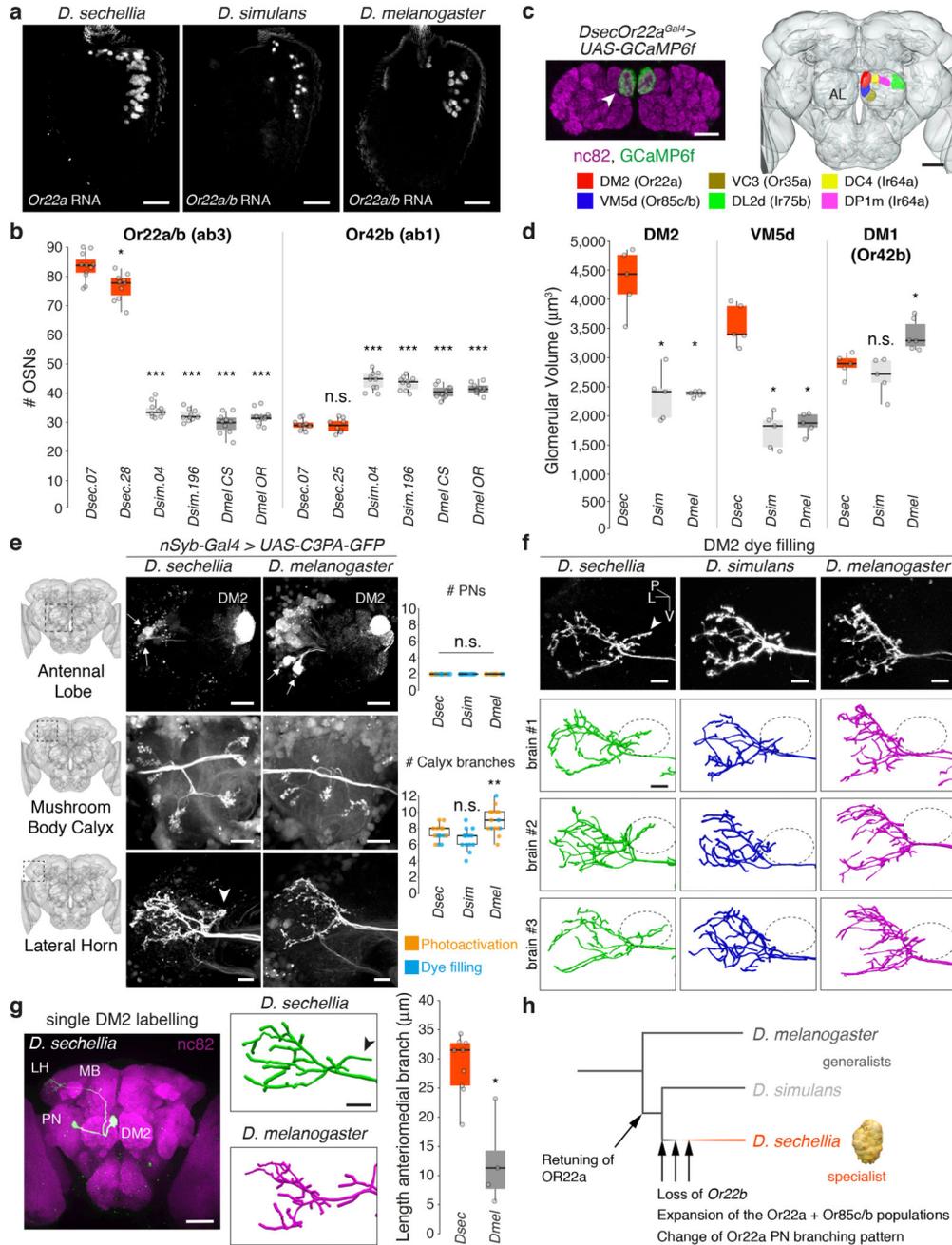
Fig. 4. Neuroanatomy of noni-sensing olfactory pathways.

a, Antennal *Or22a/b* RNA expression in different species. Scale bars, 25 μm.

b, Quantification of Or22a/(b) or Or42b OSNs (n = 8-11, females). Comparisons to *Dsec.07* are shown (pairwise Wilcoxon rank-sum test, *P* values adjusted for multiple comparisons using the B&H method, **b**, **d**, **e**): *** *P* < 0.001; ** *P* < 0.01; * *P* < 0.05; n.s. *P* > 0.05.

c, Left: *Or22a^Gal4*-driven GCaMP6f expression in DM2 (arrowhead); neuropil visualised with nc82 (magenta). Scale bar, 25 μm. Right: Antennal lobe (AL) glomerular segmentation in *D. sechellia* (Extended Data Fig. 3). Scale bar, 50 μm.

d, Quantification of DM2, VM5d and DM1 volumes (n = 5 females).

e, DM2 projection neurons (PNs) labelled via photoactivation in *D. sechellia* (*DsecnSyb-Gal4/UAS-C3PA-GFP*) and *D. melanogaster* (*UAS-SPA-GFP/UAS-C3PA-GFP;nSyb-Gal4/UAS-C3PA-GFP*). Left: Image acquisition site. Top: AL with labelled PNs (arrows) and DM2 glomerulus; scale bar, 20 μm. Middle and bottom: PN innervation of mushroom body calyx and lateral horn (LH); scale bars, 10 μm. Arrowhead (**e**, **f**, **g**): extra anterio-medial branch in *D. sechellia*. Right: quantification of DM2 PNs (top) and calyx branches (bottom) (n = 14-17 females).

f, LH arbours of dye-filled DM2 PNs. Genotypes: *D. sechellia*, *D. melanogaster* as in **e**, *D. simulans*: *DsimOr22aGFP*. Below: representative LH DM2 arbour traces. Ovals: location of *D. sechellia*-specific branch. P, posterior; L, lateral; V, ventral. Scale bars, 10 μm.

g, Left: single dye-filled DM2 PN in *D. sechellia*; scale bar, 50 μm. Middle: representative LH arbour traces of DM2 PNs in *D. sechellia* and *D. melanogaster*; scale bar, 10 μm. Right: quantification of anteriomedial branch length (n = 4-9 females) (pairwise Wilcoxon rank-sum test): * $P < 0.05$.

h, Evolution of structural and physiological changes in the Or22a pathway.