



Published in final edited form as:

*Nat Neurosci.* 2021 January ; 24(1): 93–104. doi:10.1038/s41593-020-00743-y.

## Auditory activity is diverse and widespread throughout the central brain of *Drosophila*

Diego A. Pacheco<sup>1</sup>, Stephan Y. Thiberge<sup>1,2</sup>, Eftychios Pnevmatikakis<sup>3</sup>, Mala Murthy<sup>1,2,✉</sup>

<sup>1</sup>Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA.

<sup>2</sup>Bezos Center for Neural Circuit Dynamics, Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA.

<sup>3</sup>Center for Computational Mathematics, Flatiron Institute, Simons Foundation, New York, NY, USA.

### Abstract

Sensory pathways are typically studied by starting at receptor neurons and following postsynaptic neurons into the brain. However, this leads to a bias in analyses of activity toward the earliest layers of processing. Here, we present new methods for volumetric neural imaging with precise across-brain registration to characterize auditory activity throughout the entire central brain of *Drosophila* and make comparisons across trials, individuals and sexes. We discover that auditory activity is present in most central brain regions and in neurons responsive to other modalities. Auditory responses are temporally diverse, but the majority of activity is tuned to courtship song features. Auditory responses are stereotyped across trials and animals in early mechanosensory regions, becoming more variable at higher layers of the putative pathway, and this variability is largely independent of ongoing movements. This study highlights the power of using an unbiased, brain-wide approach for mapping the functional organization of sensory activity.

---

A central problem in neuroscience is determining how information from the outside world is represented by the brain. Solving this problem requires knowing which brain areas and neurons encode particular aspects of the sensory world, how representations are transformed

---

✉ **Correspondence and requests for materials** should be addressed to M.M. mmurthy@princeton.edu.

Author contributions

D.A.P. and M.M. designed the study. D.A.P. collected and analyzed data. S.Y.T. designed and built the two-photon imaging microscope. E.P. generalized ROI segmentation from 2D to 3D datasets. D.A.P. and M.M. wrote the manuscript.

Code availability

The Matlab code for the neural imaging data processing pipeline is available at <https://github.com/murthylab/FlyCaImAn>.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-020-00743-y>.

Competing interests

The authors declare no competing interests.

Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41593-020-00743-y>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41593-020-00743-y>.

**Peer review information** *Nature Neuroscience* thanks Ioana Carcea, Elizabeth Hillman, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

from one layer to the next and the degree to which these representations vary across presentations and individuals. Most sensory pathways have been successfully studied by starting at the periphery—the receptor neurons—and then characterizing responses along postsynaptic partners in the brain. We therefore know the most about the earliest stages of sensory processing in most model organisms. However, even for relatively simple brains, there exists a significant amount of cross-talk between brain regions<sup>1,2</sup>. Thus, neurons and brain areas may be more multimodal than appreciated. In this article, we present methods to systematically characterize sensory responses throughout the *Drosophila* brain, for which abundant neural circuit tools facilitate moving from activity maps to the targeting of identifiable neurons<sup>3</sup>.

We focus on auditory coding in the *Drosophila* brain because, despite its relevance to courtship behavior, surprisingly little is known regarding how auditory information is processed downstream of primary mechanosensory neurons. In addition, in contrast to olfaction or vision, for which the range of odors or visual patterns that flies experience in naturalistic conditions is large, courtship song is the main auditory cue to which flies respond and comprises a narrow range of species-specific features that are easily probed in experiments<sup>4</sup>. *Drosophila* courtship behavior unfolds over many minutes, and females must extract features from male songs to inform mating decisions<sup>5,6</sup>, whereas males listen to songs to inform their courtship decisions in group settings<sup>7</sup>. Fly song has only three modes, two types of ‘pulse song’ and one ‘sine song’<sup>8</sup>, and males both alternate between these modes to form song bouts and modulate song intensity based on sensory feedback<sup>9,10</sup>.

Flies detect sound using a feathery appendage of the antenna, the arista, which vibrates in response to near-field sounds<sup>11</sup>. Antennal displacements activate mechanosensory Johnston’s organ neurons (JONs) housed within the antenna. Three major populations of JONs (A, B and D) respond to vibratory stimuli at frequencies found in natural courtship song<sup>12,13</sup>. These neurons project to distinct areas of the antennal mechanosensory and motor center (AMMC) in the central brain. Recent studies suggest that the auditory pathway continues from the AMMC to the wedge (WED), then to the ventrolateral protocerebrum (VLP) and to the lateral protocerebral complex (LPC)<sup>4,14–19</sup>. However, our knowledge of the fly auditory pathway remains incomplete, and the functional organization of regions downstream of the AMMC and WED are largely unexplored. Moreover, nearly all studies of auditory coding in *Drosophila* have been performed using female brains, even though both males and females process courtship song information<sup>4</sup>.

To address these issues, we developed methods to investigate the representation of behaviorally relevant auditory signals throughout the central brain of *Drosophila* and to make comparisons across animals. We use two-photon microscopy to sequentially target the entirety of the *Drosophila* central brain in vivo, combined with fully automated segmentation of regions of interest (ROIs)<sup>20</sup>. In contrast to recent brain-wide imaging studies of *Drosophila*<sup>1,21</sup>, we traded off temporal speed for enhanced spatial resolution. Imaging at high spatial resolution facilitates automated ROI segmentation, with each ROI covering subneuropil structures, including cell bodies and neurites. ROIs were accurately registered into an in vivo template brain to compare activity across trials, individuals and sexes, and to build comprehensive maps of auditory activity throughout the central brain. Our results

reveal that the representation of auditory signals is broadly distributed throughout 33 out of 36 major brain regions, including in regions known to process other sensory modalities, such as all levels of the olfactory pathway, or to drive various motor behaviors, such as the central complex. The representation of auditory stimuli is diverse across brain regions, but focused on conspecific features of courtship song. Auditory activity is more stereotyped (across trials and individuals) at early stages of the putative mechanosensory pathway, becoming more variable and more selective for particular aspects of the courtship song at higher stages. This variability cannot be explained by simultaneous measurements of ongoing fly behavior. Meanwhile, auditory maps are largely similar between male and female brains, despite extensive sexual dimorphisms in neuronal number and morphology. These findings provide the first brain-wide description of sensory processing and feature tuning in *Drosophila*.

## Results

### A new pipeline for mapping sensory activity throughout the central brain of *Drosophila*.

We developed a pipeline to volumetrically image sensory responses (via the pan-neuronal expression of GCaMP6s) and then precisely register sequentially imaged volumes into an in vivo template (Fig. 1a and Extended Data Fig. 1a). We presented three distinct auditory stimuli, with the first two representing the major modes of *Drosophila melanogaster* courtship song (pulse and sine). These stimuli should drive auditory neurons but not neurons sensitive to slow vibrations, such as those induced by wind or gravity<sup>12,13</sup>. For each fly, we imaged 14–17 subvolumes at 1 Hz (voxel size of  $1.2 \times 1.4 \times 2 \mu\text{m}^3$ ) until the whole extent of the *z* axis per fly (posterior to anterior) and one-quarter of the central brain in the *x–y* axis were covered, what we refer to hereafter as a ‘volume’. These volumes were time-series motion-corrected and stitched along the posterior–anterior axis using a secondary signal<sup>22</sup> (Fig. 1a, step (1)). We imaged volumes from either dorsal or ventral quadrants, providing full hemisphere coverage by imaging only two flies (Fig. 1b). We mirrored each volume such that all dorsal or ventral volumes imaged could be compared in the same reference brain coordinates (see below); this is why all activity maps appear on half of the brain (for example, see Extended Data Fig. 1b).

To automate the segmentation of ROIs throughout the brain, we used a constrained non-negative matrix factorization algorithm optimized for three-dimensional (3D) time-series data<sup>20,23</sup> (Methods and Fig. 1a, step (2)). This method groups contiguous pixels that are similar in their temporal responses; thus, each ROI is a functional unit. We remained agnostic as to the specific relationship between ROIs and neurons, whereby a single ROI could represent activity from many neurons or a single neuron could be segmented into multiple ROIs if different neuronal compartments differ in their activity. Importantly, this method enables the effective segmentation of responses from neuropil (axons, dendrites and processes), not only somas, and this is critical for capturing sensory responses in *Drosophila*, for which somas may not reflect activity in processes<sup>24</sup>. We typically extracted thousands of ROIs per volume and collectively sampled activity from all central brain regions (Fig. 1c and Extended Data Fig. 1b,c). These ROIs included both spontaneous and stimulus-driven activity (Fig. 1c), whereby stimulus-driven ROIs were identified using quantitative criteria (Methods, Fig. 1a, step (3), and Extended Data Fig. 1d,e). These criteria selected ROIs with

responses consistent across trials (Fig. 1d), but see below for the analysis of trial-to-trial variation even given these criteria. We found hundreds of auditory ROIs per fly ranging from 1.2% to 25.4% of all segmented ROIs per fly (Extended Data Fig. 1f).

To compare activity across individuals and sexes, we next registered functionally imaged brain segments into an in vivo intersex atlas (IVIA) using the computational morphometry (CMTK) toolkit (Methods, Fig. 1a, step (4), and Extended Data Figs. 1a and 2a). Registration of individual fly brains to our IVIA was at cellular accuracy (~4.3  $\mu\text{m}$  across all tracts; Extended Data Fig. 2b). Using this approach, we successfully registered volumes from 45 out of 48 flies (see examples in Extended Data Fig. 2c).

The Insect Brain Nomenclature Whole Brain (IBNWB) *Drosophila* fixed brain atlas<sup>25</sup> contains detailed brain neuropil and neurite tract segmentation information. By linking our IVIA template to the IBNWB atlas, we could map auditory activity onto known brain neuropils and processes (Fig. 1f,g and Supplementary Table 1) and thereby relate it to the entire network of *D. melanogaster* anatomical repositories<sup>3</sup>. We built a bridge registration between IVIA and the IBNWB atlas using a point set registration algorithm<sup>26</sup> with high accuracy (Extended Data Fig. 2d–g).

In summary, we developed a novel open-source pipeline for monitoring, anatomically annotating and directly comparing sensory-driven activity, at high spatial resolution, across trials, individuals and sexes.

### **Auditory activity is widespread across the central brain.**

We extracted 19,036 auditory ROIs from 33 female and male brains (Fig. 2a and Extended Data Fig. 3a). Within a single fly, auditory ROIs were broadly distributed across the anterior–posterior and dorsal–ventral axes (Fig. 2b and Supplementary Videos 1–4). Probability density maps across flies using the spatial components of the ROIs (Methods) revealed across-individual (Fig. 2c) and across-sex (Extended Data Fig. 3b) consistency. A total of 77% of all segmented ROIs were assigned to an identifiable neuropil or tract (Fig. 2d–h, Extended Data Fig. 3c,d and Supplementary Video 5). Surprisingly, 33 out of 36 central brain neuropils contained robust auditory activity (Methods). In addition, testing other song-relevant stimuli confirmed that we did not miss auditory responses that were narrowly selective to additional frequencies and patterns not present in our sine and pulse stimuli (Extended Data Fig. 3e–l).

We found auditory activity in all neuropils innervated by neurons previously described as auditory or mechanosensory<sup>4,6,14–16,18,27–29</sup>; that is, in the AMMC, the saddle (SAD), the WED, the anterior VLP (AVLP), the gnathal ganglion (GNG) and the inferior and superior clamp, which we herein refer to as ‘expected auditory’ neuropils (Fig. 2d,e). Similarly, we detected auditory ROIs in many tracts or commissures (the great commissure, the wedge commissure, and the inferior and superior saddle commissure) connecting expected auditory neuropils (Fig. 2g and Extended Data Fig. 3m,n).

Surprisingly, we found that a number of neuropils and tracts outside the expected auditory neuropils contained a large percentage of auditory responses (fraction of voxels with

auditory activity (Fig. 2e)). For example, neuropils of the olfactory pathway (the antennal lobe (AL), the medial antennal lobe tract containing olfactory projection neurons (PNs), the mushroom body (MB), including the peduncle, lobes and calyx, and the lateral horn (LH)), parts of the visual pathway (the posterior VLP (PVLP) and the posterior lateral protocerebrum (PLP)), and the central complex (the ellipsoid body, the fan-shaped body, the protocerebral bridge, the bulb and the lateral accessory lobe); such neuropils are referred to as ‘new auditory’ neuropils (Fig. 2e,h). We confirmed that activity within the olfactory pathway was present within the somas of individual AL PNs and MB Kenyon cells (Extended Data Fig. 4a–d). Some of this new activity was in neuropils known to be downstream targets of expected auditory neuropils, but never previously studied functionally (for example, the inferior posterior slope and the gorget<sup>17</sup>). These results demonstrate that auditory-evoked activity is widespread throughout the brain and extends beyond expected auditory neuropils, including in the olfactory, visual and pre-motor pathways.

### **Widespread auditory activity originates in mechanosensory receptor neurons.**

To determine how auditory information makes its way to so many neuropils, we used the new ‘hemibrain’ connectomic resource<sup>30</sup>. Mechanosensory receptor neurons (JONs) project to the AMMC/SAD, and AMMC/SAD neurons connect with the WED<sup>14,17</sup>. We determined the primary and secondary target neuropils for all AMMC/SAD and WED neurons identified in the hemibrain. This analysis revealed that neurons that originate in the AMMC/SAD or WED directly connect with neurons in 22 or 24 neuropils, and with one additional synapse, connect with neurons in 32 or 33 neuropils, respectively (Fig. 2i).

To confirm that this activity was in fact driven by mechanosensory receptor neurons, we imaged flies carrying the *jav1* mutation, which renders flies deaf<sup>31,32</sup>. This manipulation resulted in a 99% reduction in activity (Extended Data Fig. 4e,f). Response magnitudes of the small number of auditory ROIs found in *jav1* flies were substantially reduced compared with wild-type flies (Extended Data Fig. 4g). We therefore conclude that the overwhelming majority of the auditory activity we mapped with our stimuli are transduced through the mechanoreceptor neurons.

### **Auditory activity throughout the central brain is characterized by diverse temporal responses.**

We next hierarchically clustered responses from auditory ROIs based on their temporal profiles (Methods); this process identified 18 distinct stimulus-locked response types (Fig. 3a). These same 18 response types were present in both male and female brains (Extended Data Fig. 5a), with only minor differences between the sexes (Extended Data Fig. 5b,c)—we therefore pooled male and female auditory ROIs for subsequent analyses. We measured the frequency of response types for each neuropil to evaluate differences in auditory representations by neuropil (Fig. 3b). Auditory activity was diverse throughout the central brain, which consisted of both inhibitory (types 1–3) and excitatory (types 5–18) responses, along with diversity in response kinetics (Fig. 3d–f). Response types also differed in their selectivity for auditory stimuli (Fig. 3g).

The diversity of response types (Fig. 3b) suggests that our method samples activity from many different neurons per neuropil. In addition, each response type had a distinct spatial distribution (Extended Data Fig. 5d,e). When we measured the entropy of response types per neuropil as a proxy for the diversity of auditory activity, responses were less diverse in the earliest mechanosensory areas and became more diverse at higher levels of the putative pathway (Figs. 2d and 3c). Response diversity was highest in new auditory areas, such as the PLP, the LH and the superior lateral protocerebrum. Clustering the data in Fig. 3b revealed neuropils with highly similar response profiles, which suggests that there is functional interconnectivity (Extended Data Fig. 6a).

Within the expected auditory neuropils (Fig. 2d), we found that the AMMC contained predominantly nonselective (type 13) and sine-preferring (type 18) response types, but that brain regions downstream of the AMMC showed changes in the distribution of response types, such as a decrease in nonselective response types and an increase in both noise-preferring and pulse-preferring response types (Fig. 3b). Prior work on individual neurons innervating the AVLP suggested that this brain area mostly encodes pulse-like stimuli<sup>6,14,16,19</sup>. In contrast, we found strong representation of both sine-preferring and pulse-preferring response types, with responses becoming more narrowly tuned in this region (for example, response types 7 and 11). In the lateral junction (the inferior and superior clamp), as expected from innervations of pulse-tuned neurons<sup>4,19</sup>, there were a number of pulse-preferring response types (types 11 and 12). However, the dominant response types were sine-preferring.

In the new auditory neuropils, we found that the olfactory system was dominated by nonselective (responding equivalently to pulse and sine stimuli) and sine-preferring response types (Fig. 3b and Extended Data Fig. 4a–d); however, the response types became more diverse in higher-order olfactory regions (Fig. 3c). In the visual neuropils (the PVLP, the PLP and the anterior optic tubercle), auditory activity was highly compartmentalized. For example, in the PVLP, activity was concentrated in the vicinity of ventromedial optic glomeruli, while in the PLP, auditory ROIs formed a single discrete compartment—minimally overlapping with optic glomeruli—that runs from the anterior WED–PLP boundary to a more posterior PLP–LH boundary (Extended Data Fig. 6b). In the central complex (the ellipsoid body, the fan-shaped body and the protocerebral bridge) we found primarily nonselective and inhibitory response types (type 1), while the lateral complex (the bulb and the lateral accessory lobe) contained both sine-tuned (types 6 and 18) and nonselective excitatory response types (types 13 and 16) (Fig. 3b).

In summary, auditory activity across the central brain comprises at least 18 distinct response types, with both expected and new auditory neuropils showing a diversity of response types and with responses becoming more stimulus-selective in deeper brain neuropils.

### **Widespread auditory activity is centered on features of conspecific courtship songs.**

To investigate tuning for courtship song features (Fig. 4a), we focused on the neuropils with strongest auditory activity (Figs. 2e and Fig. 4b) and imaged subregions of each of these neuropils (Extended Data Fig. 7a) to sample ROIs with varying degrees of pulse or sine selectivity. Tuning curves were formed for each ROI and for each set of stimuli (Extended

Data Fig. 7b). These tuning curves could be clustered into seven tuning types (Fig. 4c,d), but auditory responses were still divided into three main categories (pulse-tuned (tuning type 1), sine-tuned (tuning type 5–7) or nonselective (tuning type 2–4); compare with Fig. 3), thereby validating our initial choice of stimuli.

Sine-tuned ROIs were either low pass (tuning type 5–6: strongest responses to 100 or 150 Hz sines, matching conspecific sine song frequencies) or had a broader distribution of preferred carrier frequencies (tuning type 7) (Fig. 4e). Pulse-tuned ROIs responded best to pulse frequencies, pauses and durations present in natural courtship song (Fig. 4e,f). Sine-tuned ROIs preferred continuous stimuli, which explains their preference for pulses with short pauses (below 20 ms) and longer pulse durations. Most tuning types were sensitive to the broad range of intensities tested ( $0.5\text{--}5\text{ mm s}^{-1}$ ), with proportional increases in response to stimulus magnitude, and responded best to longer pulse train or sine tone durations (4 s or longer) (Extended Data Fig. 7d,e). Although, during natural courtship, pulse or sine trains (stretches of each type of song) have a mode of  $\sim 360\text{ ms}$  and rarely last longer than  $4\text{ s}^9$ , preference for unnaturally long bouts have been recently reported<sup>4</sup>. Finally, the distribution of tuning types shifted along the pathway from the AMMC to the six additional neuropils examined here (Fig. 4d).

*Drosophila melanogaster* pulse song is a sine tone (carrier frequency of  $\sim 250\text{ Hz}$ ) with an amplitude modulation at  $\sim 27\text{ Hz}$ , whereas sine song is an unmodulated sinusoid of  $\sim 150\text{ Hz}$ . How do carrier frequency and envelope interact when it comes to tuning for pulse and sine stimuli? To address this question, we used a novel set of stimuli. We added an  $\sim 27\text{-Hz}$  envelope to pure tones of 150 or 250 Hz carrier frequency at different strengths (Methods). Responses tuned to pulses increased their response with the strength of the envelope (in other words, as a sine stimulus became more pulse-like), and the magnitude of this enhancement depended on the carrier frequency (Fig. 4g). Similarly, responses tuned to sine tones were strongly attenuated by the presence of an envelope, with different sensitivities by tuning type. These experiments suggest that there is a strong categorical boundary between pulse and sine stimuli among neural responses, with little evidence for selectivity for intermediate envelopes and frequencies.

### Evaluating auditory response stereotypy across both trials and individuals.

It is not known how reliable sensory responses are across trials, individuals and sexes for most sensory modalities and brain areas. Our method of precise registration of volumetrically imaged auditory activity enabled us to systematically evaluate response variability across different neuropils and tracts. We first measured the residual of the variance explained by the mean across trials for all auditory ROIs (Methods). Our criteria for selecting auditory ROIs (Extended Data Fig. 1d–f) excluded ROIs that did not show consistent responses across trials to at least one auditory stimulus. Nonetheless, we still observed a range of across-trial variabilities for auditory ROIs (Fig. 5a), with a pattern across neuropils. Early mechanosensory neuropils (the AMMC, the SAD, the WED and the AVL) contained the lowest across-trial response variance, while all other neuropils exhibited higher trial-to-trial variability (Fig. 5b). This variation was, to some degree,

correlated with the magnitude of auditory responses across neuropils (Extended Data Fig. 8a).

We showed above that particular response types (Fig. 3a) are found in consistent spatial locations across flies (Extended Data Fig. 5e), which suggests that there is a high degree of across-individual stereotypy of auditory responses. To quantify this, we used a similarity index (Methods) to compare activity profiles across individuals (Fig. 4c,d and Extended Data Fig. 9). Some neuropils, like the GNG, contained ROIs with temporal responses highly similar across individuals (for example, all individuals have a high frequency of response type 18 in the GNG), whereas other neuropils, like the PLP, showed more variable activity across individuals. Comparing the distribution of the similarity index across all neuropils revealed that early mechanosensory neuropils also had the highest across-individual similarity in auditory responses (Fig. 4d and Extended Data Fig. 8b). Similarity decreased in downstream neuropils, and further decreased in neuropils outside the canonical mechanosensory system (for example, visual, olfactory and pre-motor neuropils). These differences in similarity across neuropils were robust against changes in how auditory responses are clustered (Extended Data Fig. 8c) and did not arise from variations in GCaMP fluorescence values across animals (Extended Data Fig. 8d).

### **Spontaneous movements do not account for auditory response variability.**

Across-trial variability in auditory responses could arise from the time-varying behavioral state of the animal over the course of an experiment, as has been recently shown for cortical activity in mice<sup>33,34</sup>. Above, we did not track animal behavior; therefore, we could not correlate response variation with behavioral variation. To address this issue, we collected a new dataset in which we simultaneously recorded fly motion on a spherical treadmill while imaging pan-neuronally from brain regions with a range of across-trial variation from stereotyped (the AMMC and the SAD) to more variable (the AVL, the PVLP and the LH) (Figs. 4 and 6a–e). We observed strong auditory activity, with across-trial variability similar to our previous dataset (Extended Data Fig. 10a), whereby flies showed a range of walking speeds on the ball (Extended Data Fig. 10b) that matched natural speeds during fly courtship<sup>9</sup>. Fly movements in these experiments appeared to be spontaneous because they were not reliably predicted by the auditory stimuli (Extended Data Fig. 10c). In addition, despite observing auditory activity in pre-motor neuropils (Fig. 3b), imaging from the fly ventral nerve cord (VNC) in non-walking flies did not uncover much activity (only 247 out of 39,580 ROIs recorded from the VNC were significantly correlated with any of the stimuli; Extended Data Fig. 10d–g). This suggests that auditory ROIs in our study reflect sensory responses to stimuli and not motor responses.

Because flies spontaneously moved during the presentation of auditory stimuli, this allowed us to determine what fraction of the across-trial variance in responses could be explained by the stimulus, by fly behavior or by some combination of the two. We built a linear encoding model to predict neural responses (Fig. 6f), similar to ref. <sup>33</sup>, from a combination of the mean response to the stimulus ( $\mu(t)$ ) and behavior predictors ( $X(t)$ ), which consisted of all tracked motor variables (Fig. 6b–e and Methods). For this analysis, we used all ROIs and not just those that passed our criteria for being auditory (Extended Data Fig. 1d). Of 63,106

ROIs segmented from the central brain, each had a different amount of explained variance from either the stimulus or behavior (Fig. 6g,h). However, the stimulus component ( $\mu(t)$ ) explained the majority of variance, while the behavior component ( $\text{Beta} \times X(t)$ ) explained a much smaller fraction (Fig. 6i); the full model ( $\mu(t) + \text{Beta} \times X(t)$ ) was as good at explaining the variance across ROIs as the stimulus component alone (Fig. 6j). When plotting the explained variance of the stimulus component against the behavior component, we did not find a substantial trend for explained variance from the behavior predictors; that is, ROIs with little correlation to the stimulus had as much explained variance from the motor variables as ROIs with stronger correlation to the stimulus (Fig. 6k). These results suggest that the variability we observed across trials in auditory ROIs, at least for the brain regions imaged here, does not arise from fluctuations in the behavioral state of the animal.

## Discussion

Sensory systems are typically studied starting from the periphery and continuing to downstream partners guided by anatomy. This has limited our understanding of sensory processing to early stages of a given sensory pathway. Here, we used a brain-wide imaging method to unbiasedly screen for auditory responses beyond the periphery and, via precise registration of recorded activity, to compare auditory representations across brain regions, individuals and sexes (Fig. 1). We found that auditory activity is widespread, extending well beyond the canonical mechanosensory pathway, and is present in brain regions and tracts known to process other sensory modalities (that is, olfaction and vision) or to drive motor behaviors (Fig. 2). The representation of auditory stimuli diversified, in terms of both temporal responses to stimuli and tuning for stimulus features, from the AMMC to later stages of the putative pathway, becoming more selective for particular aspects of courtship song (that is, sine or pulse song, and their characteristic spectrotemporal patterns) (Figs. 3 and 4). Auditory representations were more stereotypic across trials and individuals in early stages of mechanosensory processing, and more variable at later stages (Fig. 5). By recording neural activity in behaving flies, we found that fly movements accounted for only a small fraction of the variance in neural activity, which suggests that across-trial auditory response variability stems from other sources (Fig. 6). These results have important implications for how the brain processes auditory information to extract salient features and guide behavior.

Our understanding of the *Drosophila* auditory circuit thus far has been built up from targeted studies of neural cell types that innervate particular brain regions close to the auditory periphery<sup>14–17</sup>. Altogether, these studies have delineated a pathway that starts in the Johnston's organ and extends from the AMMC to the WED, the VLP and the LPC. By imaging pan-neuronally, we found widespread auditory responses that spanned brain regions beyond the canonical pathway, which suggests that auditory processing is more distributed. However, for neuropil signals, it was challenging to determine the number of neurons that contribute to the ROI responses we describe. Although the diverse set of temporal and tuning types per neuropil (Fig. 3) suggests that we sampled many neurons per neuropil, restricting GCaMP to spatially restricted genetic enhancer lines (Extended Data Fig. 4c,d) will assist with linking broad functional maps with the cell types constituting them<sup>35</sup>.

Our findings of widespread auditory activity are likely not unique to audition. So far, in adult *Drosophila*, only taste processing has been broadly surveyed<sup>36</sup>. While that study did not map activity onto neuropils and tracts, nor did it make comparisons across individuals, it suggested that taste processing was distributed throughout the brain. Similarly, in vertebrates, widespread responses to visual and nociceptive stimuli have been observed throughout the brain<sup>37,38</sup>. Our findings are consistent with anatomical studies that found connections from the AMMC and the WED to several other brain regions<sup>17,39</sup>, and with our own analysis of the hemibrain connectome (Fig. 2). While we do not yet know what role this widespread auditory activity plays in behavior, we show that ROIs that respond to auditory stimuli do so with mostly excitatory (depolarizing) responses and that activity throughout the brain is predominantly tuned to features of the courtship song. During courtship, flies evaluate multiple sensory cues (olfactory, auditory, gustatory and visual) to inform mating decisions and to modulate their mating drive. Although integration of multiple sensory modalities has been described in higher-order brain regions<sup>40</sup>, our results suggest that song representations are integrated with olfactory and visual information at earlier stages (Figs. 2 and 3). In addition, song information may modulate the processing of non-courtship stimuli. Song representations in the MB may be useful for learning associations between song and general olfactory, gustatory or visual cues<sup>41</sup>, while diverse auditory activity throughout all regions of the LH may indicate an interaction between song processing and innate olfactory behaviors<sup>42</sup>. Finally, we found auditory activity in brain regions involved in locomotion and navigation (the central and lateral complex, and the superior and ventromedial neuropils). Auditory activity in these regions is diverse (Fig. 3), which suggests that pre-motor circuits receive information about courtship song patterns and could therefore underlie stimulus-specific locomotor responses<sup>4</sup>.

*D. melanogaster* songs are composed of pulses and sines that differ in their spectral and temporal properties; however, it is unclear how and where selectivity for the different song modes arises in the brain. Since neurons in the LPC are tuned for pulse song across all time scales that define that mode of song<sup>4</sup>, neurons upstream must carry the relevant information to generate such tuning. Here, we found many ROIs that are selective for either sine or pulse stimuli throughout the entire central brain (Fig. 3). A more detailed systematic examination of tuning in the neuropils that carry the most auditory activity (the AMMC, the SAD, the WED, the AVLP and PVLP, the PLP and the LH) revealed that most ROIs are tuned to either pulse or sine features, with few ROIs possessing intermediate tuning (Fig. 4). This suggests that sine and pulse information splits early in the pathway. We also found that sine-selective responses dominate throughout the brain. Although some of this selectivity may simply reflect preference for continuous versus pulsatile stimuli, our investigation of feature tuning revealed that many of these ROIs preferred frequencies that are specifically present in courtship songs (Fig. 4). Previous studies indicated that pulse song is more important for mating decisions, with sine song purported to play a role in only priming females<sup>5,43</sup>. However our results, in combination with the fact that males spend a greater proportion of time in courtship singing sine versus pulse song<sup>9</sup>, suggest a need for reevaluation of the importance and role of sine song in mating decisions. This study therefore lays the foundation for exploring how song selectivity arises in the brain.

Our results also revealed that early mechanosensory brain areas contain ROIs with less variable auditory activity across trials and animals (Fig. 5). Our results for across-animal variability have parallels to the *Drosophila* olfactory pathway, whereby third-order MB neurons are not stereotyped, while presynaptic neurons in the AL, the PNs, are<sup>44</sup>. Similarly, a lack of stereotypy beyond early mechanosensory brain areas may reflect stochasticity in synaptic wiring. The amount of variation we observed in some brain areas was large (Extended Data Fig. 9), and follow-up experiments with sparser driver lines will be needed to validate whether what we report here applies to variation across individual identifiable neurons.

We also observed a wide range of across-trial variability throughout neuropils with auditory activity (Fig. 5b). Imaging from a subset of brain regions in behaving flies revealed that trial-to-trial variance in auditory responses is not explained by spontaneous movements (Fig. 6), which suggests that variance is driven by internal dynamics. This result differs from recent findings in the mouse brain<sup>33,34</sup>, which showed that a large fraction of activity in sensory cortices corresponds to non-task-related or spontaneous movements. This may indicate an important difference between invertebrate and vertebrate brains and the degree to which ongoing movements shape activity across different brains. However, we should point out that while motor activity is known to affect sensory activity in flies<sup>45</sup>, this modulation is tied to movements that are informative for either optomotor responses or steering<sup>46,47</sup>. In our experiments, although flies walked abundantly, they did not produce reliable responses to auditory stimuli, although playback of the same auditory stimuli can reliably change walking speed in freely behaving flies<sup>4</sup>. Adjusting our paradigm to drive such responses<sup>48</sup> might uncover behavioral modulation of auditory activity. Alternatively, behavioral modulation of auditory responses may occur primarily in motor areas, such as the central complex, or areas containing projections of descending neurons<sup>49</sup>. Further dissection of the sources of this variability would require the simultaneous capture of more brain activity in behaving animals while not significantly compromising spatial resolution.

Here, we provide tools for characterizing sensory activity registered in common atlas coordinates for comparisons across trials, individuals and sexes. By producing maps for additional modalities and stimulus combinations and by combining these maps with information on connectivity between and within brain regions<sup>30,50</sup>, the logic of how the brain represents the myriad stimuli and their combinations present in the world should emerge.

## Methods

### Fly stocks.

Flies used to generate the IVIA and for functional-imaging experiments were of the following genotype: w/+; GMR57C10-LexA/+; 13xLexAop-GCaMP6s, 8xLexAop-mCD8tdTomato/+ (nsyb-LexA-G6s-tdtom), all '+' chromosomes came from the NM91 wild-type strain<sup>9</sup>. *Iav*<sup>1</sup> mutants had the following genotype: *iav*<sup>1</sup>/Y; GMR57C10-LexA/+; 13xLexAop-GCaMP6s, 8xLexAop-mCD8tdTomato/+. For imaging of olfactory PNs and Kenyon cells, flies had the w; GMR57C10-LexA, 8xLexAop-myr-tdTomato/GH146-Gal4; 20xUAS-GCaMP6s/+ genotype and w; GMR57C10-LexA, 8xLexAop-myr-tdTomato/+;

20xUAS-GCaMP6s/OK107-Gal4 genotype, respectively. For stalk segmentation of pC1, pC21 and pC2m *Dsx*-expressing neurons, we used the w; GMR57C10-LexA/20xUAS-GCaMP6s, 20xUAS-myr-tdTomato; 8xLexAop-mCD8tdTomato/*dsx*<sup>Gal4</sup> genotype. For stalk segmentation of P1 neurons and pMP7 neurons, we used the w; NP2631-Gal4, 20XUAS>STOP>CsChrimson.mVenus/GMR57C10-LexA; 8xLexAop-mCD8tdTomato/*fru*<sup>FLP</sup> genotype. Other flies were acquired as follows: 13xLexAop-GCaMP6s (44590), GH146-Gal4(30026) and OK107-Gal4 (854) from the Bloomington Stock Center; *iav*<sup>1</sup> from the Kyoto Stock Center; GMR57C10-LexA from G. Rubin at Janelia Research Campus; 20XUAS>STOP>CsChrimson.mVenus from V. Jayaraman at Janelia Research Campus; 8xLexAop-mCD8tdTomato from Y. N. Jan at the University of California San Francisco; NP2631-Gal4 from the Kyoto Stock Center; *fru*<sup>FLP</sup> from B. Dickson at Janelia Research Campus; *dsx*<sup>Gal4</sup> from S. Goodwin at University of Oxford; NM91 from the Andolfatto Group at Columbia University; and 20xUAS-myr-tdTomato from G. Turner at Janelia Research Campus.

### Sound delivery and acoustic stimuli.

Sound was delivered as described in ref. <sup>6</sup>. During experiments, the sound tube was positioned in front of the fly at a distance of 2 mm and an angle of ~17° from the midline (elevation of ~39°). In this configuration, we observed qualitatively similar responses from the AMMC in both hemispheres.

Acoustic stimuli were organized in two main protocols: (1) coarse-tuning protocol and (2) feature-tuning protocol (see Supplementary Table 2 for a list of all stimulus sets used). For both protocols, the intensity of auditory stimuli presented was within the dynamic range of JONs<sup>51</sup> (0.1–6 mm s<sup>-1</sup>) and within the distribution of song intensities produced during natural courtship<sup>52</sup> (pulse and sine song intensity of 5 mm s<sup>-1</sup>, and broadband white noise intensity of 2 mm s<sup>-1</sup>). Stimuli were preceded and followed by 10 or 4 s of silence for the coarse-tuning and feature-tuning protocol, respectively, to reduce cross-talk between responses to the subsequent stimulus. Stimuli were presented in blocks (for example, pulse, sine, white noise, repeated six times) or randomized for the coarse-tuning and feature-tuning protocol, respectively. We chose a long stimuli duration (10 or 4 s) to maximize the signal-to-noise ratio of auditory responses.

### Head-fixed walking setup and behavioral tracking.

We shaped 9-mm diameter balls from polyurethane foam (FR-7120, General Plastics) similar to that described in ref. <sup>53</sup>. The ball rested in an aluminum ball holder with a concave hemisphere 9.6 mm in diameter (Fig. 6a). The ball holder had a 1.27-mm channel drilled through the bottom of the hemisphere and connected to air flowing at ~75 ml min<sup>-1</sup> and mounted on a manipulator (MS3, Thorlabs) to adjust the position of the ball under each fly.

We imaged the fly and ball from the side (90° to the anterior–posterior axis of the fly), illuminated by a pair of infrared 850-nm light-emitting diodes (M850F2, Thorlabs) coupled to optic fibers and collimator lenses (F810SMA-780, Thorlabs). Imaging was done with a monochromatic camera (Flea3 FL3-U3-13Y3M-C 1/2, FLIR), externally triggered at 100 Hz, with a zoom lens (C-Mount 15.5–20.4 mm Varifocal, Computar). The lens had a 770–

790-nm narrow-band filter (BN785, MidOpt) attached to block a two-photon excitation laser (920 nm). This camera was used to position the fly and to track both the ball and fly movement.

To track the rotation of the ball, we used a Windows implementation of FicTrac software<sup>54</sup> (see <https://github.com/murthylab/fictrac#windows-installation>). FicTrac calculates the angular position of the ball for each frame and reconstructs the  $x$  and  $y$  trajectory on a fictive two-dimensional (2D) surface. The ball was marked with non-repetitive shapes using a black permanent marker (PK10, Sarstedt) to allow tracking of the rotation of the ball. We used the  $x$  and  $y$  trajectories to calculate instantaneous ground velocities (forward and lateral velocity, and the speed in the fictive heading direction (forward plus lateral velocity)) and the angular position to calculate the instantaneous angular velocities (yaw, pitch and roll) (Fig. 6b).

To extract a high-dimensional representation of the body motion of the fly (that is, legs, abdomen and wings), we used the toolbox facemap<sup>34</sup> (see <https://github.com/MouseLand/facemap>). Facemap applies singular value decomposition to the motion energy movie (the absolute difference between frame $_t$  and frame $_{t+1}$ ) (Fig. 6c), extracting the top 500 coefficients of the components (spatial weights) and corresponding scores (temporal profiles) of the motion energy movie (Fig. 6d,e). For this analysis, the field of view of each recorded video was reduced to the smallest box containing the whole fly body, masking out pixels that belonged to the ball, and the head plus anterior thorax (pixels of these body parts were removed due to its contamination with the calcium-imaging-related signal). The top 500 components of the motion energy movie per fly were used as a summary of motor behaviors for later analysis (see the section “Linear modeling of neural activity from stimulus and behavior”).

### Fly preparation and functional imaging.

Virgin female or male flies (3–7 days old) were mounted and dissected as previously described<sup>27,55</sup>, with minor differences for non-behaving animals. The angle of the thorax to the posterior side of the head was kept close to 90°, keeping the head as parallel as possible to the holder floor. Following dissection of the head cuticle, we removed the air sacks and tracheas using sharp forceps (Dumont number 5SF), and any additional fat or soft tissue was removed with suction using a sharp glass pipette. In addition, to enable imaging of ventral central brain regions close to the neck, we pushed the thorax posteriorly and fixed it using a tissue adhesive (3M vetbond) delivered using a sharp glass pipette. Proboscis or digestion-related motion artifacts were minimized by pulling the proboscis and waxing it at an extended configuration and removing muscle M16. In addition, during dissection, we kept the antennae dry and mobile for auditory stimulation. For VNC recordings, in addition to the previous steps, we removed thoracic tissue dorsal to the VNC (for example, the cuticle, indirect flight muscles and the digestive system), exposing the first and second segments of the VNC. For head-fixed behaving animals, flies were fixed to a custom holder (similar to ref. <sup>56</sup>) by gluing the lateral and dorsal anterior edges of the thorax with light cured glue (Bondic), followed by waxing of the posterior side of the head (the head was pushed forward to end up with a similar configuration as for dissection for non-behaving animals). The rest of the dissection was the same as for non-behaving animals.

Imaging experiments were performed on a custom-built two-photon laser scanning microscope equipped with 5-mm galvanometer mirrors (Cambridge Technology), an electro-optic modulator (M350–80LA-02 KD\*P, Conoptics) to control the laser intensity, a piezoelectric focusing device (P-725, Physik Instrumente), a Chameleon Ultra II Ti:sapphire laser (Coherent) and a water-immersion objective (Olympus XLPlan  $\times 25$ , NA = 1.05). The fluorescence signal collected by the objective was reflected by a dichroic mirror (FF685 Dio2, Semrock), filtered using a multiphoton short-pass emission filter (FF01–680/sp-25, Semrock), split by a dichroic mirror (FF555 Dio3, Semrock) into two channels (green (FF02–525/40–25, Semrock) and red (FF01–593/40–25, Semrock)) and detected by GaAsP photo-multiplier tubes (H10770PA-40, Hamamatsu). We used a laser power below 20 mW to minimize photodamage. The microscope was controlled in Matlab using ScanImage 5.1 (Vidrio). Dissection chambers were placed beneath the objective, and perfusion saline was continuously delivered directly to the meniscus. The temperature of the perfusion saline was kept at 24 °C using a miniature perfusion cooler and heater unit (TC-RD, TC2–80-150-C, Biosciencetools). We chose GCaMP6s<sup>57</sup> over faster sensors to maximize the signal-to-noise ratio of auditory responses. The experimenter was not blind to the animal sex or genotype, but males or females (of the same genotype) were randomly chosen for imaging experiments. For all flies not carrying the *iav*<sup>1</sup> mutation, we only recorded from animals that first exhibited auditory-evoked responses in the AMMC of both hemispheres. Recordings typically lasted for 2–3 h. Flies whose global baseline fluorescence saturated and became homogeneous were discarded. We interpreted this as a sign of an unhealthy fly preparation.

### Functional imaging analysis pipeline.

All steps from image preprocessing (step (1)) to registration of functional data to IVIA (step (4)) are schematized in Extended Data Fig. 1a, in which we take one example dataset step by step through the entire functional-imaging analysis pipeline.

**Image preprocessing.**—For all experiments, we recorded data from each animal using one of the following three protocols. For protocol one, 14–17 subvolumes at 1 Hz were recorded until the whole extent of the *z* axis per fly (posterior to anterior) was covered, with a  $1.2 \times 1.4 \times 2 \mu\text{m}^3$  voxel size and an  $\sim 542$  kHz pixel rate (Figs. 1–3 and 5 and Extended Data Figs. 3, 4 and 10d–g). The full volume scanned per fly was one-quarter of the central brain or  $\sim 300 \times 300 \times 250 \mu\text{m}^3$ . For protocol two, 4–6 subvolumes from selected brain regions (not necessarily contiguous) at 2 Hz were recorded, with a  $1 \times 1 \times 2 \mu\text{m}^3$ ,  $1 \times 1 \times 3 \mu\text{m}^3$  or  $1.2 \times 1.2 \times 5 \mu\text{m}^3$  voxel size and an  $\sim 542$  kHz pixel rate (Fig. 4). For protocol three, one subvolume that extends from the LH in the posterior brain to the AMMC in the anterior brain at 1 Hz was recorded, with a  $1.2 \times 1.4 \times 15 \mu\text{m}^3$  voxel size and an  $\sim 542$  kHz pixel rate (Fig. 6). To generate the private whole brain atlas (per fly), we imaged 1–4 *z*-stack volumes (in both GCaMP and tdTomato channels) with a  $0.75 \times 0.75 \times 1 \mu\text{m}^3$  voxel size to span the entire central brain (both hemispheres) in the *z* dimension.

For each volumetric time series or segment (*x–y–z* and time) (Extended Data Fig. 1a), we performed rigid motion correction on the *x–y–z* axes on the tdTomato signal using the NoRMCorre algorithm<sup>22</sup>, and this transformation was then applied to the GCaMP6s signal. Segments with an anisotropic *x–y* voxel size were spatially resampled to have an isotropic

$x$ - $y$  voxel size (bilinear interpolation on the  $x$  and  $y$  axes), and temporally resampled to correct for different slice timing across planes of the same volume. Finally, segments were upsampled (linear interpolation) to twice the original sampling rate (2 Hz and 4 Hz for segments recorded at 1 Hz and 2 Hz, respectively) and aligned relative to the start of the first stimulus (linear interpolation) before analysis.

Segments consecutively recorded along the  $z$  axis (for protocol one) were stitched along the  $z$  axis using NoRMCorre (similarly, we used the average tdTomato signal as the reference image and applied the transformation to the GCAMP6s signal), obtaining a volume. For protocols two and three, each segment was treated as an independent volume. Volumes were then mirrored to the right hemisphere. Average tdTomato images were saved as individual NRRD (<http://teem.sourceforge.net/nrrd>) files for registration to IVIA (using the Matlab `nrrdWriter` function).

For the private whole brain atlases (per fly),  $z$ -stacks ( $\sim 384 \times 384 \times 280 \mu\text{m}^3$  each) were stitched on the  $x$ - $y$  axes using a pairwise stitching algorithm in Fiji<sup>58,59</sup>. These whole-brain images were smoothed on the  $x$ - $y$  axes (Gaussian kernel size of [3, 3] voxels and Gaussian kernel standard deviation of [2, 2]), mirrored on the  $x$  axis as for the volumetric time series. Images were split into tdTomato and GCaMP6s channels and saved separately as NRRD files for registration to IVIA (using the Matlab `nrrdWriter` function).

**ROI segmentation.**—We used the constrained non-negative matrix factorization algorithm<sup>20</sup> for ROI segmentation. The algorithm was implemented and generalized to 3D data in the Matlab version of the CaImAn toolbox<sup>23</sup>. We used the greedy initialization (Gaussian kernel size of [9, 9, 5] voxels and Gaussian kernel standard deviation of [4, 4, 2], and 10th percentile baseline subtraction) with 600 components per substack of 11 planes and walked through the entirety of the  $z$  axis with a substack overlap of 3 planes. After initialization, spatial components were updated using the ‘dilate’ search method, and default settings were used for updating temporal components. Then, spatially overlapping components were iteratively merged if the Pearson’s correlation of temporal components between them was greater than 0.9 (this allows stitching of contiguous patches of neuropil with similar temporal profiles). The combination of these two steps segmented soma-like and neuropil-like ROIs accordingly. We wrote custom code to compile spatial and temporal components from all substacks per fly (similar to `run_CNMF_patches.m`) and to calculate  $F(t)/F_0(t)$  ROI. Background fluorescence  $F_0(t)$  and delta fluorescence  $F(t)$  (that is, ROI activity) per ROI are directly modeled by CaImAn and provided as output variables. To calculate the  $F(t)/F_0(t)$ , first, for each ROI, we detrended  $F(t)$  (we performed a 20th percentile filtering over time, calculated from overlapping windows of 60 s each). Second, for each ROI we added to  $F_0(t)$  its corresponding trend removed from  $F(t)$ .  $F(t)/F_0(t)$  was calculated as the detrended  $F(t)$  divided by  $F_0(t)$  (which is now the background fluorescence plus the trend from  $F(t)$ ) for each time point). See also <https://github.com/murthylab/FlyCaImAn>.

**Identification of stimulus-modulated ROIs.**—We linearly modeled each ROI signal as a convolution of the stimuli history and a set of three filters (one per stimulus) (Extended Data Fig. 1d). Filters were estimated using ridge regression<sup>60</sup>, with a filter length of 10 and 6

s for the coarse-tuning protocol (Figs. 1–3, 5 and 6) and the feature-tuning protocol (Fig. 4), respectively. For filter estimation, we partitioned each ROI raw signal ( $F(t)$ ) into training (80%) and testing (20%) datasets. Estimated filters ( $q(T)$ ) were then convolved with the stimulus history ( $f(t)$ ) to generate the predicted signal ( $g(t)$ ) for each ROI. The prediction goodness was measured as the Pearson's correlation coefficient between raw and predicted signals  $\rho(F, g)$  using 15-fold cross-validation. The statistical significance of correlation coefficients ( $\rho(F, g)$ ) was determined by bootstrapping. Each ROI raw signal was randomly shuffled in chunks of 10 s ( $sF(t)$ ), and a distribution of 10,000 correlation coefficients ( $\rho(sF, g)$ ) between each independent shuffle and the predicted signal was generated.  $P$  values for  $\rho(F, g)$  significance were calculated as the fraction of  $\rho(sF, g)$  with values greater than the 30th percentile of  $\rho(F, g)$ , and adjusted using Benjamini–Hochberg false discovery rate (FDR) correction (FDR = 0.01).

**Registration of functional data to IVIA.**—In vivo functional data from each fly were registered to the IVIA using the structural channel (mCD8–tdTomato signal) in a two-step fashion: (1) volumes to private whole brain registration and (2) private whole brain to IVIA atlas registration (Extended Data Fig. 1a).

For volume to private whole brain registration, for each fly, we registered (linear followed by nonlinear registration) the volume tdTomato image to its own whole brain tdTomato image (z-stack volume, see the section “Image preprocessing”). Linear transformation included rotation, translation and anisotropic scaling (metric: normalized mutual information). For the nonlinear transformation (metric: normalized mutual information), we used a small grid spacing (20  $\mu\text{m}$ , with 2 refinements), a Jacobian constraint weight of 0.01 and a smoothness constraint weight of 0.1.

For private whole brain to IVIA atlas registration, for each fly, we registered (linear followed by nonlinear registration) the private whole brain tdTomato image to the IVIA atlas. Linear transformation included rotation, translation, anisotropic scaling and shearing (metric: normalized mutual information). For the nonlinear transformation (metric: normalized mutual information), we used a bigger grid spacing (170  $\mu\text{m}$ , with 5 refinements), a Jacobian constraint weight of 0.001 and a smoothness constraint weight of 1.

These volume-to-whole brain and whole brain-to-IVIA atlas transformations per fly were concatenated to transform spatial components of segmented ROIs from the native volume coordinates to in vivo atlas coordinates.

### Construction of the average IVIA.

Recently, an in vivo *D. melanogaster* female atlas was built<sup>1</sup>; however, it does not account for known anatomical differences between female and male brains<sup>61</sup>. Therefore, we built an average IVIA using a pool of five male and eight female brains (expressing membranal tdTomato pan-neuronally (mCD8–tdTomato) (Extended Data Fig. 2a). In the first iteration, we picked a seed brain, half cropped on the  $x$  axis and stitched it to its mirror image to generate a symmetric initial brain seed. We then registered all the brains to the seed brain using the CMTK registration toolbox<sup>42,62</sup>. We used linear followed by nonlinear registration. We then generated a new average intensity and average deformation seed brain using an

active deformation model (implemented in CMTK as `avg_adm` function). We iterated this process and obtained the IVIA at the end of the fifth iteration.

### Construction of IVIA–IBNWB bridging registrations.

From all the fixed-brain atlases with anatomical labels (IBNWB and JFRC), we chose the IBNWB atlas based on the similar *nsyb*–GFP signal coverage (neuropil, fiber bundles and cell body rind) compared to the IVIA atlas (pan-neuronal *mCD8*–*tdTomato*). Despite this similarity, the two atlases have dissimilar distributions of fluorescence intensities per anatomical region and differences in signal-to-noise ratio (much higher for IBNWB), which limited the accuracy of intensity-based registration algorithms. Therefore, we used the rigid and non-rigid point set registration algorithm coherent point drift (CPD)<sup>26</sup>. We chose the surface of selected neuropils and neurite tracts and the entire central brain as sources of points to compare between atlases (Extended Data Fig. 2d). We converted segmented binary images from each of these brain regions to triangular surface meshes, generating a set of vertices and edges per brain region to be used for registration. We applied a rigid (translation, rotation and scaling) followed by a non-rigid CPD registration in both directions (IVIA-to-IBNWB and IBNWB-to-IVIA) between the two sets of vertices across all brain regions. For the non-rigid registration, we explored the space of the two main meta-parameters *beta* (Gaussian smoothing filter size) and *lambda* (regularization weight). We chose the non-rigid transformation (for either direction) with the highest Jaccard index across all segmented brain regions that had spatially smooth deformations (visually inspected) (Extended Data Fig. 2e).

Using the IVIA-to-IBNWB transformation, we mapped an anatomical segmentation of neuropils and neurite bundles<sup>25</sup> to the IVIA coordinates (Fig. 1g). In this segmentation, the AMMC would also include JON projections. In addition, we merged the MB accessory calyx to the MB calyx due to its small size relative to the average ROI volume.

### Registration accuracy.

To assess IVIA within-atlas accuracy, we segmented stalks from pC1, pC2l, pC2m and pM7 neurons ( $n = 20, 10, 12$  and  $7$  hemibrains for pC1, pC2l, pC2m and pM7, respectively), which were consistently identified across individuals in flies expressing *GCaMP6s* via *dsx*<sup>Gal4</sup> or *CsChrimson* via NP2631–GAL4 and *fru*<sup>FLP</sup> combo. These stalks covered different dorsoventral regions from the posterior half of the brain (Extended Data Fig. 2b). We measured the within-atlas registration accuracy as the standard deviation of each trace class from the mean trace (all traces across animals). The overall within-atlas accuracy was calculated as the standard deviation across all trace classes (Extended Data Fig. 2b). This estimate is the sum of biological variability and algorithm-associated error. Although we obtained cellular accuracy ( $\sim 4.3 \mu\text{m}$ ), this is slightly lower compared with values reported for fixed-brain atlases ( $2\text{--}3 \mu\text{m}$ )<sup>42,63–65</sup>, and this might be due to both a richer spatial pattern and higher signal-to-noise ratio of the BRP (*nc82*) antibody staining in fixed-brain experiments.

To assess IVIA and IBNWB between-atlas accuracy, we used segmented pC1 stalks from the IVIA (described above) and compared them to pC1 neurons originally collected in the

FCWB atlas (imaged via *fru*<sup>Gal4</sup> expression of GFP<sup>63</sup>), which were mapped to IBNWB (error of the transformation between FCWB and IBNWB is assumed to be negligible). Similar to the IVIA within-atlas accuracy, we calculated the mean pC1 trace for each atlas ( $\mu$ -pC1-IVIA and  $\mu$ -pC1-IBNWB) and the within-atlas accuracy ( $\sigma$ -pC1-IVIA and  $\sigma$ -pC1-IBNWB) (Extended Data Fig. 2f,g). We measured the IVIA-to-IBNWB registration accuracy as the difference between the  $\sigma$ -pC1-IBNWB and the standard deviation of pC1-IBNWB traces mapped to the IVIA relative to  $\mu$ -pC1-IVIA. Moreover, we measured the IBNWB-to-IVIA registration accuracy as the difference between the  $\sigma$ -pC1-IVIA and the standard deviation of pC1-IVIA traces mapped to IBNWB relative to  $\mu$ -pC1-IBNWB.

### Anatomical cataloging of central brain activity.

For each fly, we combined all auditory ROIs and binarized these volumes to obtain binary auditory voxel maps per fly. Auditory maps were then transformed to IVIA coordinates. To account for the known fly-to-atlas error, we dilated these transformed maps ( $[1.25, 1.25, 2]$   $\mu\text{m}^3$  in  $x$ - $y$ - $z$ ). We then summed these maps across flies to generate a density volume (Fig. 2c and Extended Data Figs. 1b, 3b,c,i, 4f and 5d,e). To spatially map auditory activity, we transformed neuropils and neurite bundle segmentation from IVIA coordinates to imaged volume coordinates. We assigned ROIs to a given neuropil or neurite bundle using the following criteria: ROIs must have at least a 15% volume overlap with its assigned neuropil or neurite bundle; if overlapping with multiple neuropils or neurite bundles, ROIs are assigned to the neuropil or neurite bundle with the highest overlap. For auditory ROIs, we then quantified the volume they represent relative to the neuropil or neurite bundle for each fly (we used ROI, neuropil and neurite bundle volume in IVIA space) (Fig. 2e and Extended Data Fig. 3m).

### Functional clustering of auditory responses to pulse, sine and white noise stimuli.

For all auditory-modulated ROIs from intact and *iav*<sup>l</sup> flies (19,389 ROIs from 21 male (4 male flies carried *iav*<sup>l</sup>) and 17 female flies expressing tdTomato and GCaMP pen-neuronally) presented with the coarse-tuning protocol (this includes pulse, sine, white noise, natural song and P<sub>fast</sub>-like pulse; see the section “Sound delivery acoustic stimuli”), we calculated the median response to each stimulus across trials, including 10 s before stimulus onset to 10 s after stimulus offset. We concatenated the median signal across pulse, sine and white noise stimuli only (183 time points) for each ROI and  $z$ -scored it. We then hierarchically clustered ROI signals (pooled from all intact and *iav*<sup>l</sup> flies) using a Euclidean distance metric and inner square distance metric between clusters (Ward’s method). This clustering split ROI responses into inhibitory and excitatory responses (first branching of the hierarchical tree), and we then chose a distance (distance of 136) that would split both inhibitory and excitatory responses into the smallest number of distinct responses types (18 response types). We also required that each cluster be present in at least 18 flies (half the entire dataset). Although we could further split these clusters, we stopped at 18 given that subclusters within each of the 18 clusters were similar to the parent cluster, which suggested the occurrence of oversegmentation (Fig. 3a).

### Response type kinetics and tuning.

For each mean trace per response type, we measured adaptation, half-time to peak and decay of auditory responses (Fig. 3d–f). Adaptation was measured as the percentage of the absolute baseline subtracted signal 2 s before stimulus offset with respect to the signal 2 s after stimulus onset ('adaptation index'). Therefore, adaptation indices below 100% indicate some degree of adaptation (the lower the adaptation index, the higher the adaptation), while adaptation indices above 100% indicate a sustained increase in response (the higher the adaptation index, the higher the increase). Half-time to peak was measured as half the time to reach the peak signal during stimuli presentation. The decay time constant of auditory responses was determined by fitting an exponential to the 20 values after stimulus offset (10 s), except for response types 2 and 3, for which we fit an exponential to the 20 values after positive rebound.

To determine the coarse tuning of response types to pulse, sine or white noise, we measured the magnitude of stimulus-evoked responses. The response magnitude was defined as the absolute 80th or 20th percentile with the greater absolute value of baseline subtracted signal from stimulus onset to 2 s after stimulus offset (the baseline was defined as the mean signal from –4 to –0.5 s relative to stimulus onset, and dividing the response magnitude by the standard deviation during baseline gave us the signal-to-noise ratio). Response magnitudes were normalized by the maximum value across stimuli and multiplied by 100 to have units as a percentage. We considered a response type as tuned to a particular stimulus if the response magnitude to the other stimuli was below 85% of the response to the preferred stimulus, otherwise it was considered nonselective (Fig. 3g).

### Clustering of tuning curves.

We generated tuning curves to the 'feature tuning' stimulus set (Fig. 4). For these experiments, we sampled a fraction of each neuropil selected (Extended Data Fig. 7a). For all auditory-modulated ROIs from intact flies (10,872 ROIs from 10 male and 11 female flies) presented with the feature-tuning protocol, we calculated the median response to each stimulus across trials, including 4 s before stimulus onset to 4 s after stimulus offset, and z-scored it. We measured the magnitude of responses to the auditory stimulus for each ROI. The baseline was defined as –2 to –0.25 s from stimuli onset, and we then calculated the mean  $\mu_b$  of the signal during this time. The response magnitude was defined as the absolute 80th or 20th percentile with the greater absolute value of  $\mu_b$  subtracted signal from stimulus onset to 2 s after stimulus offset. For each ROI, we generated a tuning curve by concatenating the absolute response magnitude across the presented sine song at different carrier frequencies, and the pulse song at different carrier frequencies, pulse pause and pulse durations (Extended Data Fig. 7b). We then hierarchically clustered ROI tuning curves using a Euclidean distance metric and inner square distance metric between clusters (Ward's method). Examination of the tree revealed that tuning curves divided into three main classes: tuning to sine, tuning to pulses and nonselective. We then reduced the distance threshold (distance of 60) until we split these three tuning types into the smallest number of visually distinct clusters (Fig. 4c).

### Measuring activity diversity across brain regions.

We collected ROIs belonging to the same neuropil across flies (separately for intact or *iav*<sup>1</sup> flies). These ROIs were sorted by their response type within a neuropil. Using the volume of each ROI, we calculated the total number of voxels for each response per neuropil to obtain a distribution of voxel count per response type for each neuropil (we use ROI volume in IVIA space). These distributions were then max-normalized within neuropils; these are the values shown in Fig. 3b. We then calculated the entropy for each response type distribution as a measure of activity diversity, whereby regions with sparse response types will have a lower entropy value than regions with more broad or uniform distribution.

### Measuring trial-to-trial variability.

To evaluate the variability of auditory responses across trials, we measured the residual of the ROI activity variance explained by the mean ROI activity across trials for pulse, sine and white noise stimuli ('variability index'). Each trial included 10 s before stimulus onset to 10 s after stimulus offset. We concatenated the signal across stimuli and trials (183 time points) for each ROI and *z*-scored it. We computed the variability index for each ROI and then sorted these indices by neuropil (Fig. 5b).

### Measuring across-individual variability.

We calculated response type distributions per neuropil (Fig. 3), but for each fly separately, obtaining an array of vectors of 18 dimensions per neuropil per fly. We normalized each vector to unit norm. We then used the cosine of the angle between normalized vectors as our measure of similarity of response distributions ('similarity index'). Response distributions contained values greater than or equal to 0, thereby restricting the distribution of similarity indices to the 0–1 range.

### Measuring sex-specific differences in activity.

Similar to the previous section, we calculated response-type distributions per neuropil for each fly separately. These distributions were then sorted by sex. We then evaluated differences in probability of each of the 18 response types across sexes for each neuropil (for significance of sex-specific differences see the section "Statistics").

### Identification of stimuli-modulated motor behaviors.

Similar to the previous section ("Identification of stimulus-modulated ROIs"), we linearly modeled each motor variable (velocity and motion energy component,  $X(t)$ ) as a convolution of the stimuli history and a set of three filters (one per stimulus). Filters were estimated using ridge regression<sup>60</sup>, with a filter length of 10 s. The Pearson's correlation coefficient between the raw motor variable ( $X(t)$ ) to its linear prediction ( $g(t)$ ), the Pearson's correlation coefficient between the shuffled motor variable ( $sX(t)$ ) to  $g(t)$  and the corrected *P* values were calculated as described in the section "Identification of stimulus-modulated ROIs".

### Linear modeling of neural activity from stimulus and behavior.

To measure the fraction of neural variance that could be predicted from stimulus presentation or behavior of the fly, we constructed a linear model  $y(t) = \mu(t) + \text{Beta} \times X(t) +$

$\epsilon(t)$  (Fig. 6g). The raw signal for each ROI ( $y(t)$ ) was modeled as the sum of the stimulus component (the trial average across auditory stimuli pulse, sine and white noise,  $\mu(t)$ ), behavior component (behavior predictors—velocities and top 500—times a set of weights,  $\text{Beta} \times X(t)$ ) and noise ( $\epsilon(t)$ ). Behavior predictors were smoothed over time (Gaussian kernel size of ten time points and Gaussian kernel standard deviation of three) and downsampled to match the sampling rate of recorded activity (2 Hz). We computed the weights (Beta) of behavior predictors using ridge regression<sup>60</sup>. To quantify the explanatory power of the full model or reduced models (stimulus or behavior component only), we computed the explained variance using fivefold cross-validation.

### Analysis of the hemibrain connectome.

To find connections between early auditory neuropils—AMMC/SAD and WED—and downstream neuropils, we used the new hemibrain connectome<sup>30</sup> (Fig. 2g). The hemibrain project reconstructed 26,157 neurons and synaptic connections in an EM volume of approximately one-third of the *Drosophila* central brain. This volume includes a large portion of the AMMC, SAD and WED. We downloaded the connectivity matrix of all 26,157 neurons (<https://www.janelia.org/project-team/flyem/hemibrain>), and wrote a custom code in Python to find neuropils postsynaptic to a query neuropil, connected via one or two synapses. We considered a connection between two neurons if they had more than five synapses.

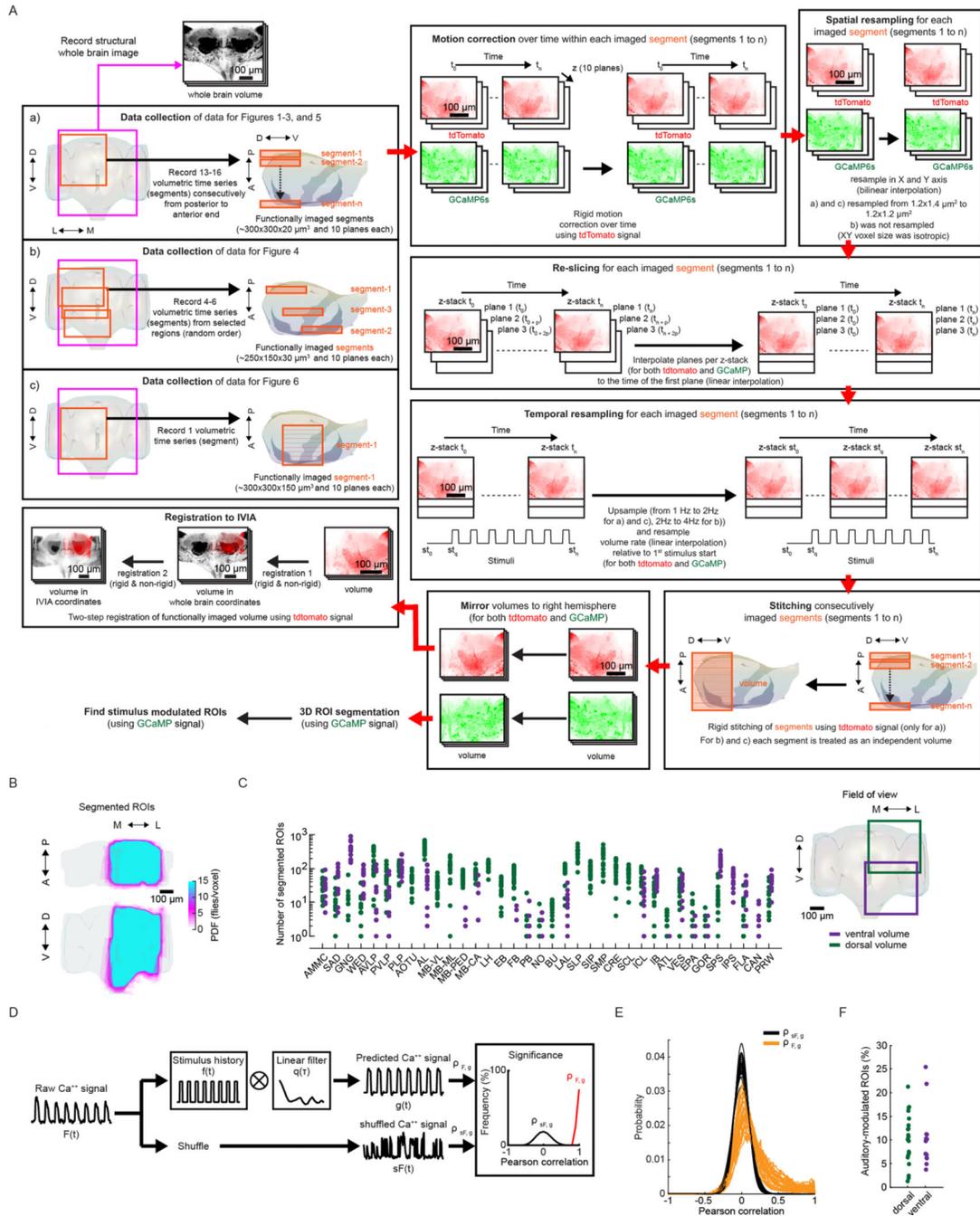
### Statistics.

All statistical analyses were performed using Matlab 2018a. Effect sizes of female-to-male differences were measured using Cohen's  $d$  (Extended Data Fig. 5c); data distribution was assumed to be normal, but this was not formally tested. Within a neuropil, we determined the significance of female-to-male differences for each response type using two-tailed two-sample  $t$ -tests with Benjamini–Hochberg FDR correction. For significance of stimuli modulation, see the section “Identification of stimulus-modulated ROIs”. No statistical methods were used to predetermine sample sizes, but our sample sizes were similar to those reported in previous publications<sup>4,18</sup>.

### Reporting Summary.

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

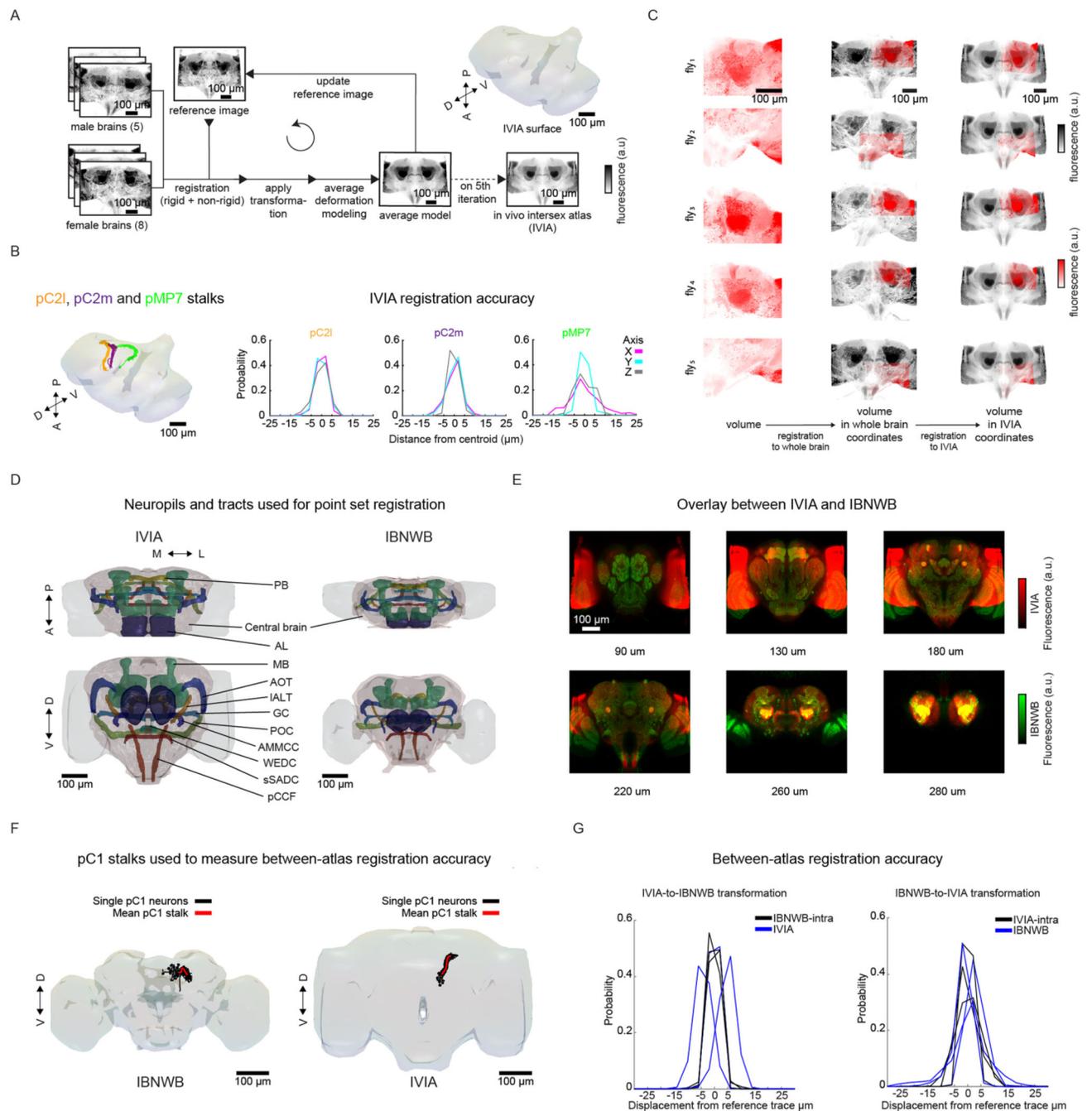
### Extended Data



**Extended Data Fig. 1 | Spatial coverage of imaged activity and measurement of auditory-evoked activity.**

**a.** Data collection and processing. We collected data using the following protocols: a) short volumetric recordings (~10 min) of consecutive segments (from anterior to posterior axis), b) long volumetric recordings (~30 min) of selected brain areas (non-consecutive), and c) short and coarse volumetric recording (~15 min) of a single segment per fly that spans a large volume. For all these datasets we record a private whole brain structural volume (used

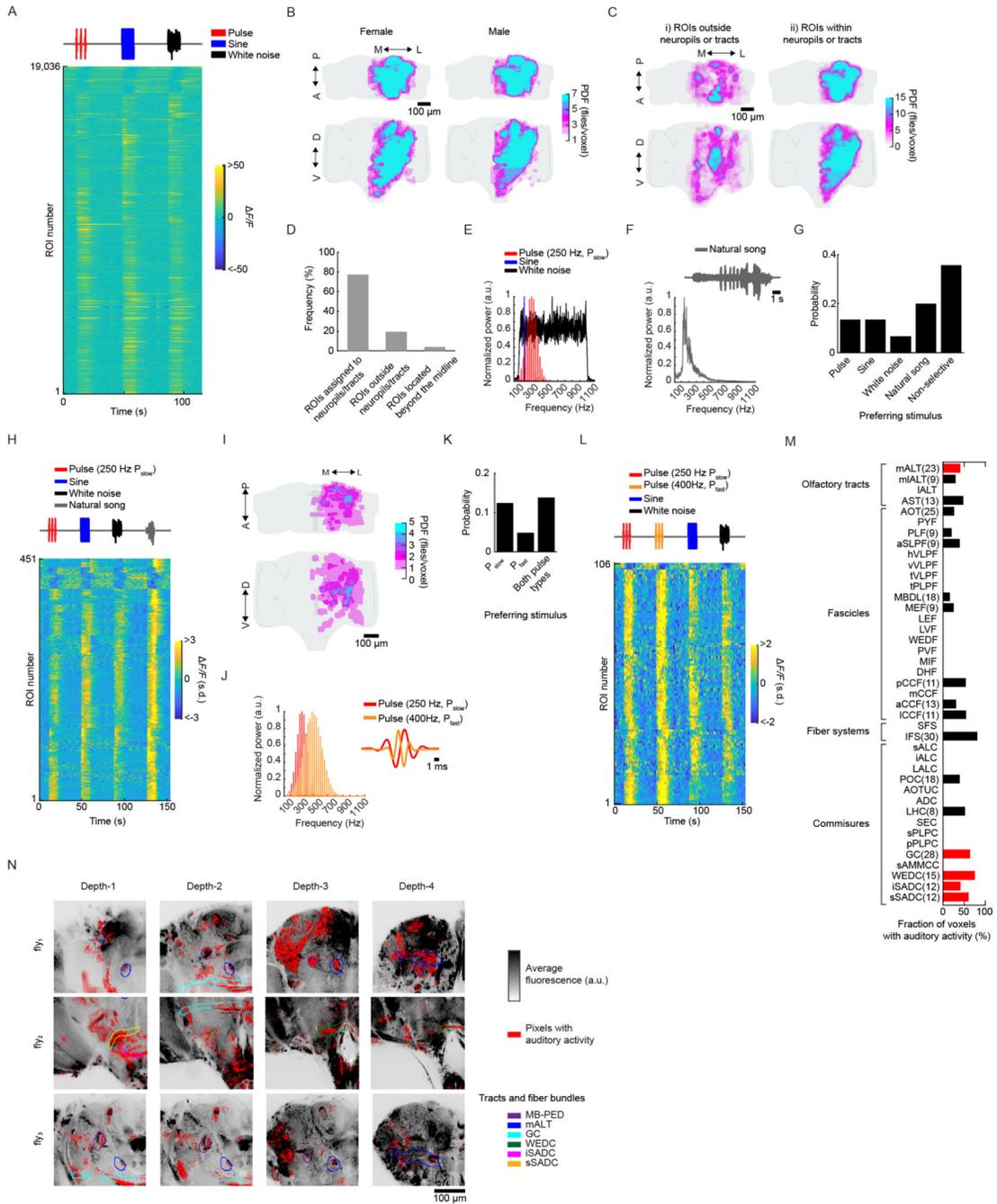
later for registration). For all the data sets we recorded both tdTomato and GCaMP signal, and performed motion correction (using NoRMCorre) on each imaged segment using the tdTomato signal. These segments were then spatially resampled to have isotropic XY pixel size. This was followed by re-slicing of each Z-stack per segment (align time of all planes per Z-stack to the first plane imaged), and temporal resampling (align Z-stack time relative to the start of the 1st stimulus and double the sampling rate). For protocol a), we stitched segments imaged consecutively (using NoRMCorre) obtaining a ‘volume’; for b) and c) each segment was treated as an independent volume. Volumes were mirrored to the right hemisphere, and the tdtomato signal was used for registration to the *in vivo* intersex atlas (IVIA). This is a two-step process, i) volumes per fly were registered to their own private whole brain volume (one per fly), and ii) whole brain volume was registered to the IVIA; registration i) and ii) were concatenated to map volumes to IVIA space. The GCaMP signal was used for ROI segmentation (via CaImAn), followed by identification of stimulus-modulated ROIs (see Extended Data Fig. 1d). **b**, Maximum projection (in each dimension) of segmented ROIs from all imaged volumes ( $n = 33$  flies, 185,395 ROIs) - ROIs from the left hemisphere are mirrored (see Extended Data Fig. 1a) such that all ROIs are projected onto right hemisphere. ROIs cover the entirety of the D-V and A-P axes. Color scale indicates the number of flies with an ROI in each voxel. **c**, Number of ROIs across neuropils (see Fig. 1f,g and Supplementary Movies 1–4) sampled by ventral volumes or dorsal volumes ( $n = 33$  flies, 185,395 ROIs). **d**, Method for identifying stimulus-modulated ROIs. Raw  $\text{Ca}^{++}$  signal ( $F(t)$ ) is convolved with the stimulus history ( $f(t)$ ) and a set of filters per stimulus type ( $q(\tau)$ ) to generate the predicted  $\text{Ca}^{++}$  signal. Auditory modulation is measured by the cross-validated correlation scores ( $\rho_{F,g}$ ) between raw and predicted  $\text{Ca}^{++}$  signals. Correlation of shuffled  $\text{Ca}^{++}$  signal ( $sF(t)$ ) to predicted signal ( $g(t)$ ) is used to generate the null-distribution of correlation scores ( $\rho_{sF,g}$ ), which is used to determine significance. **e**, Distribution of Pearson correlation coefficients of shuffled-vs-predicted signals ( $\rho_{sF,g}$ ) and raw-vs-predicted signal ( $\rho_{F,g}$ ) across all flies imaged ( $n = 33$  flies). Unlike the distribution of  $\rho_{sF,g}$ ,  $\rho_{F,g}$  has a distribution with a long tail of positive correlation scores. ROIs within the positive tail and outside the null distribution are considered to have significant stimuli modulation and selected as auditory ROIs. **f**, Frequency of auditory ROIs separated by whether the ROI was segmented from a dorsal or ventral volume ( $n = 33$  flies).



### Extended Data Fig. 2 | Building the *in vivo* intersex atlas (iViA) and registering it to fixed-brain atlas (iBNWB).

**a.** Generation of *in vivo* intersex atlas (IVIA): images of male ( $n = 5$ ) and female ( $n = 8$ ) brains expressing membranaral tdTomato pan-neuronally are registered to a seed brain (reference image). Images are then transformed to generate an average model image, which, after five iterations, produces the *in vivo* intersex atlas (IVIA). **b.** IVIA registration accuracy. Left: 3D-rendering of traced tracts of Dsx+ neurons (pC2m, pC2l), and Fru+ neuron pMP7. Right: Per-axis jitter (X, Y, and Z) between matched traced tracts of pC2m, pC2l, and pMP7

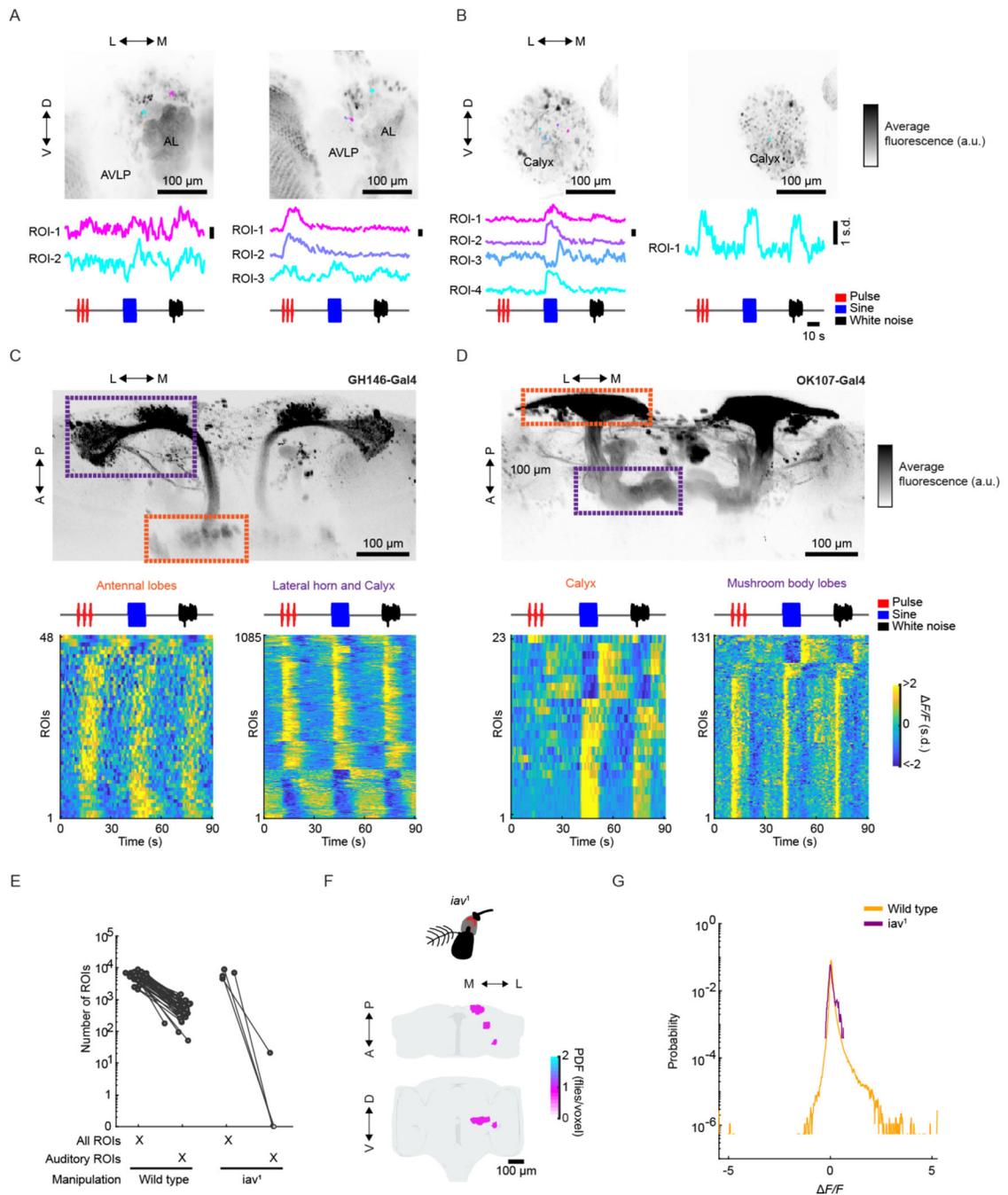
across flies (n = 10, 12, and 7 brains, respectively). **c**, Example imaged tdTomato volumes (red scale) from different flies registered to the IVIA (right-most column, gray scale). Middle column is the intermediate registration of volumes to their own private whole brain atlas (as in Extended Data Fig. 1a). 45 out of 48 flies were successfully registered to the IVIA. **d**, Central brain neuropils and tracts used for point set registration to morph the *in vivo* intersex atlas (IVIA) to IBNWB fixed brain atlas. Points from meshes of segmented antennal lobe (AL), mushroom body (MB) (includes mushroom body lobes, peduncle and calyx), protocerebral bridge (PB), antennal mechanosensory and motor center commissure (AMMCC), anterior optic tract (AOT), great commissure (GC), wedge commissure (WEDC), posterior optic commissure (POC), lateral antennal lobe tract (IALT), posterior cerebro-cervical fascicle (pCCF), superior saddle commissure (sSADC), and whole central brain were used to generate IVIA-to-IBNWB and IBNWB-to-IVIA transformations. **e**, Overlay of IVIA (red) and registered IBNWB (in IVIA space, green) at different depths (90, 130, 180, 220, 260, and 280  $\mu\text{m}$ ). 0  $\mu\text{m}$  is the most anterior section of the brain and 300  $\mu\text{m}$  is the most posterior. **f,g**, Atlas-to-atlas registration accuracy measured using pC1 stalks from IBNWB and IVIA. **(f)** pC1 traces from IBNWB and IVIA atlases; black traces are single pC1 neurons (from IBNWB or IVIA, n = 70 and 20 pC1 traces respectively) and red trace is the mean reference pC1 stalk. **(g)** Between-atlas registration accuracy; IVIA-to-IBNWB transformation increases the jitter across all axes from the reference mean pC1 trace by  $\sim 2.24 \mu\text{m}$ , while IBNWB-to-IVIA transformation increases the jitter by  $\sim 2.8 \mu\text{m}$ .



**Extended Data Fig. 3 | Auditory responses in neuropils and neurite tracts, and to additional song stimuli: natural song and fast pulses.**

**a**, Median ROI responses (across 6 trials) to pulse, sine, and white noise stimuli ( $n = 33$  flies, 19,036 ROIs) as in Fig. 2a, but without z-scoring  $FF$  signal. **b**, Spatial distribution of auditory activity across sexes. Maximum projections (from two orthogonal views) of the density of auditory ROIs in female ( $n = 17$ ) or male flies ( $n = 16$ ). **c**, Maximum projection (from two orthogonal views) of the density of i) auditory ROIs outside neuropils or neurite tracts ( $n = 33$  flies, 4,346 ROIs), and ii) auditory ROIs within neuropils or neurite tracts ( $n =$

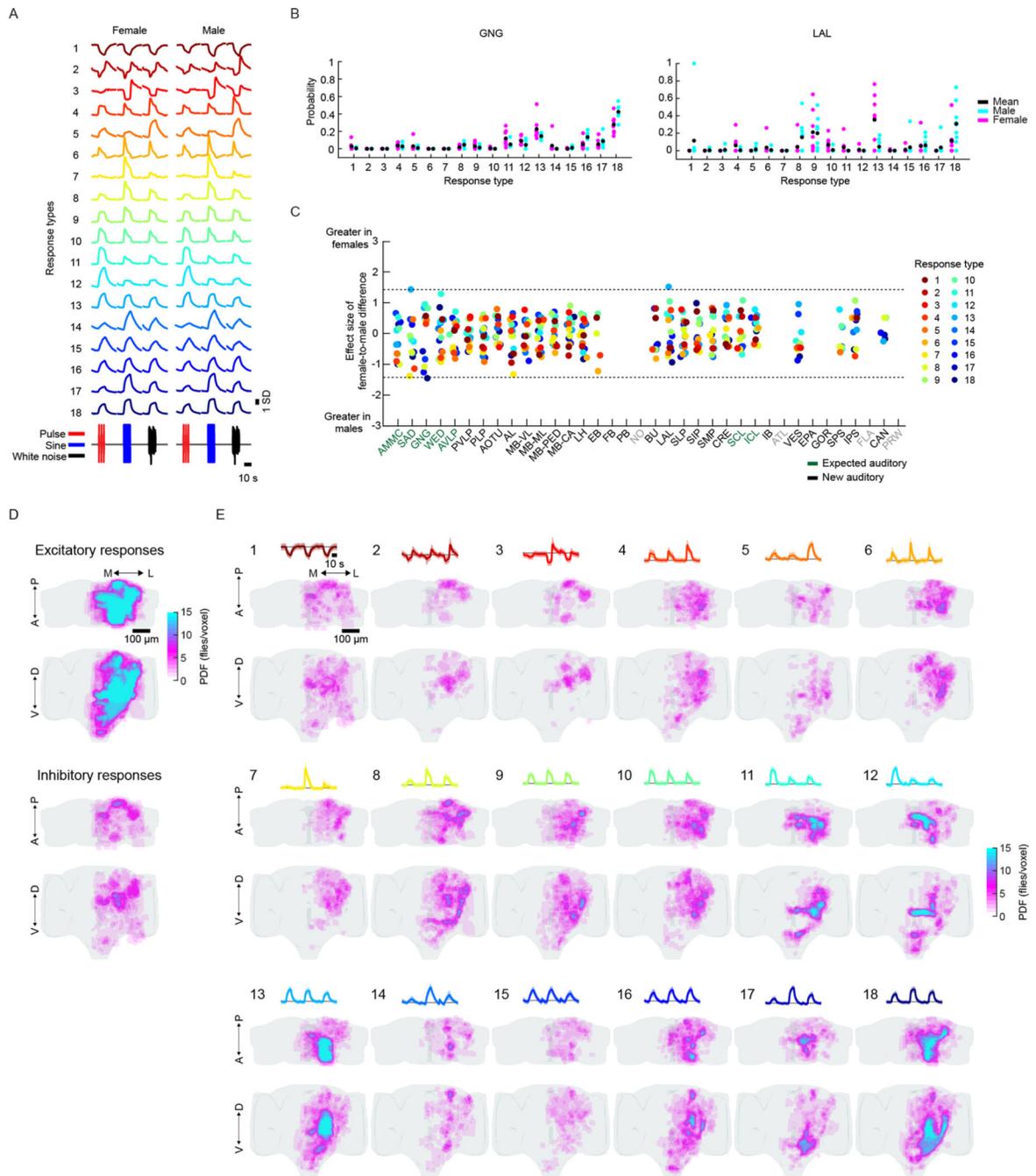
33 flies, 14,658 ROIs). Color scale for (B) and (C) is the number of flies with an auditory ROI per voxel. **d**, Percentage of auditory ROIs within and outside neuropils and tracts, and beyond the midline. **e**, Spectral profile of auditory stimuli - pulse (Pslow), sine, and white noise - used to classify response types in Fig. 3a, and their spectral features. **f**, Spectral profile of natural song stimulus used. **g**, Distribution of stimuli preference (to pulse, sine, white noise, and natural song) across auditory ROIs (n = 5 flies, 2,258 ROIs). Only ~20% of auditory ROIs prefer natural song. Preference is defined by the stimulus that drives the maximum absolute response (at least 15% greater than the second highest response), as in Fig. 3f. **h**, Responses from auditory ROIs that prefer natural song (n = 451 ROIs). Each row is the median z-scored  $F/F$  response across 6 across trials. Activity is plotted as the change in the s.d. of the  $F/F$  signal. **i**, Spatial distribution of natural song preferring ROIs. Images are the maximum projection (from two orthogonal views) of the density of auditory ROIs with preference for natural song throughout the central brain (n = 5, 451 ROIs). **j**, Spectral profile of Pfast and Pslow stimuli. **k**, Distribution of stimuli preference to Pslow, Pfast, or broad preference for both pulse types across auditory ROIs (n = 2 flies, 2,193 ROIs). Only ~4% of auditory ROIs prefer Pfast. Preference is defined as in (G). **l**, Responses from auditory ROIs that prefer Pfast (n = 106 ROIs). Pfast preferring ROIs also show strong responses to Pslow. Conventions same as in (H). **m**, Fraction of voxels with auditory activity by central brain neurite tracts; percentages averaged across 33 flies (a minimum of 4 flies with auditory activity in a given tract was required for inclusion). Red represents tracts that were clearly distinguishable from neuropil by visual inspection (see (N)). **n**, For three flies, pixels with auditory activity (red) overlaid on time-averaged GCaMP6 fluorescence (grayscale). Several neurite tracts are indicated (planes from different depths are arbitrarily selected for each fly to highlight ROIs contained within neurite tracts).



**Extended Data Fig. 4 | Auditory activity is present in olfactory pathway neurons and absent in deaf flies.**

**a,b**, Auditory responses from (A) putative individual antennal lobe projection neurons (PNs), and (B) putative individual mushroom body Kenyon cells (KCs) from pan-neuronal recordings. Time-averaged GCaMP6s signal is shown in grayscale (over the entire experiment). Pixels belonging to individual ROIs and their corresponding time traces (median z-scored  $F/F$  responses across 6 trials to pulse, sine, and white noise stimuli) are indicated in different colors. ROIs were drawn manually over the location of PN or KC

somas from two independently imaged flies. Activity scale bar unit is s.d. **c,d**, Auditory responses from (C) antennal lobe projection neurons, and (D) mushroom body Kenyon cells using cell type specific genetic lines (GH146-Gal4 and OK107-Gal4, respectively). Time-averaged GCaMP6s signal is shown in grayscale (over the entire experiment) - the compartments where activity was recorded from are indicated (data was collected and processed as we did for pan-neuronal data (protocol a) in Extended Data Fig. 1a) but the field of view was restricted to the region defined by the orange and purple boxes). Auditory ROIs are detected in all imaged compartments of antennal lobe projection neurons (n = 8 flies, 1,133 ROIs) and Kenyon cells (n = 4 flies, 154 ROIs). Each row is the median across 6 trials - all responses are z-scored and therefore plotted as the change in the s.d. of the  $F/F$  signal. **e**, Number of detected auditory ROIs per fly in wild type (n = 33 flies) and *iav<sup>1</sup>* flies (n = 4 flies). **f**, Maximum projection (from two orthogonal views) of the density of auditory ROIs in *iav<sup>1</sup>* flies (all 21 ROIs come from 1 out of 4 flies imaged). These few ROIs were located in the mushroom body calyx (MB-CA), posterior VLP (PVLP), and the posterior lateral protocerebrum (PLP). Color scale is the number of flies with an auditory ROI per voxel. **g**, Distribution of mean  $F/F$  values (during stimuli) for auditory ROIs in wild type (n = 33 flies) and *iav<sup>1</sup>* flies (n = 4 flies).



**Extended Data Fig. 5 | Auditory responses are not sexually dimorphic and Auditory response types have distinct spatial distributions throughout the central brain.**

**a.** Hierarchical clustering of auditory responses into 18 distinct response types, and divided by sex. The mean response across ROIs belonging to each response type is shown. All responses are z-scored (as in Fig. 3a), and therefore plotted as the s.d. of the responses over time. Color code is the same as in Fig. 3a. **b.** Probabilities of neuropil voxels with a given response type (separated by sex) in two neuropils, the gnathal ganglion (GNG) and the lateral accessory lobe (LAL). These are the only two neuropils with sex differences - males

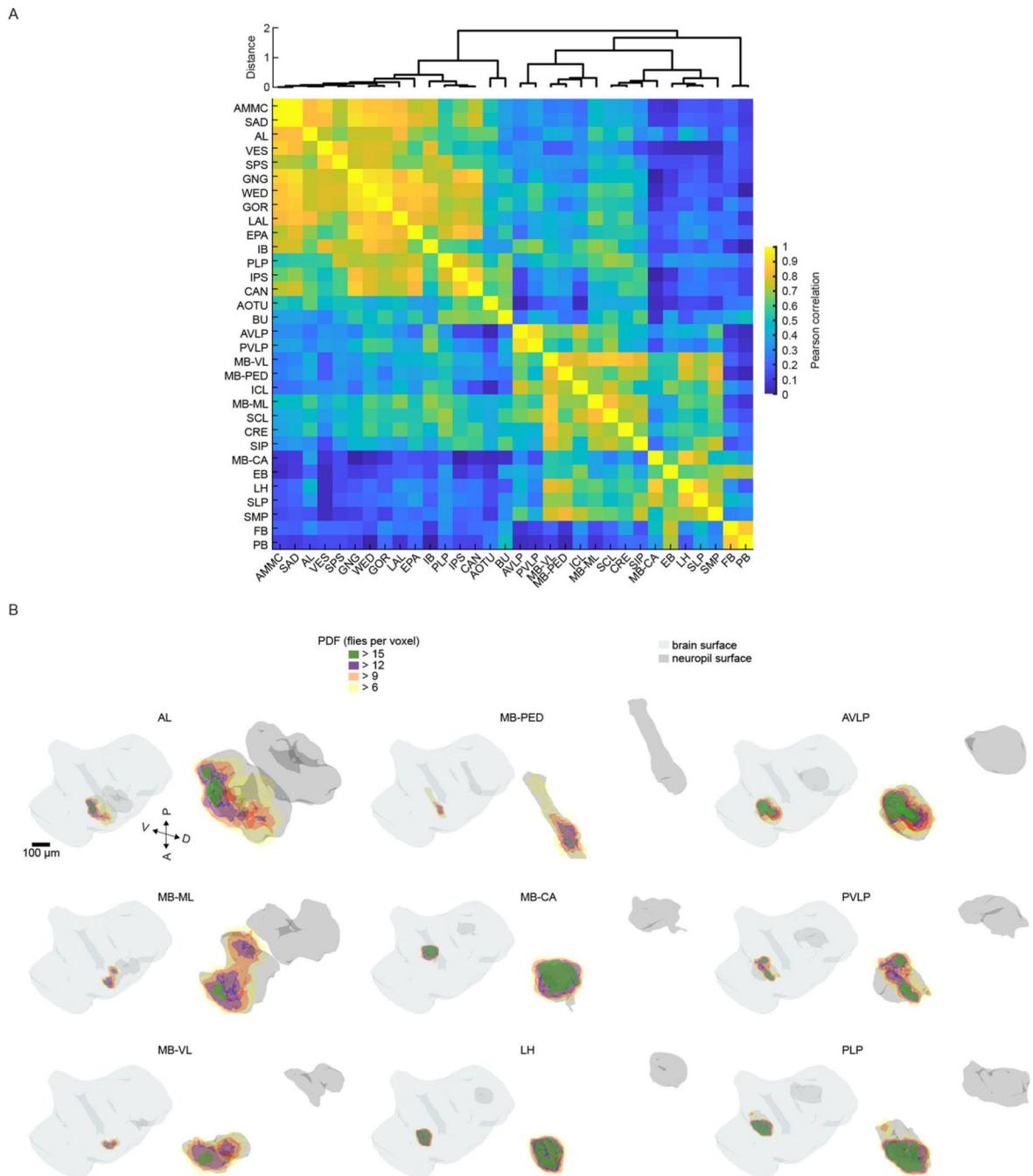
have a slightly higher probability of response type 18 activity in the GNG, while females have a slightly higher probability of response type 13 activity in the LAL. Each dot is the probability for one fly. **c**, Sex-related differences in all response types across all neuropils. Each dot is the effect size (see Methods) of the difference in probability for each response type across sexes, color code is the same as (A). Neuropils with the greatest effect size are the GNG (response type 18) and LAL (response type 13). However all differences are not significant ( $p > 0.05$ ). Statistical significance was determined using two-tailed two-sample t-test with Benjamini-Hochberg FDR correction. Neuropils with no auditory activity are indicated in gray font. **d**, Spatial distribution of excitatory vs inhibitory auditory ROI responses across the central brain. Images are the maximum projection (from two orthogonal views) of the density of excitatory and inhibitory responses across flies ( $n = 33$  flies). **e**, Spatial distribution of auditory activity belonging to each response type (see Fig. 3a) across the central brain. Images are the maximum projection (from two orthogonal views) of the density of responses across flies for each response type ( $n = 33$  flies). Color scale for (D) and (E) is the number of flies with an auditory ROI belonging to each category (excitatory, inhibitory, or response type) per voxel.

Author Manuscript

Author Manuscript

Author Manuscript

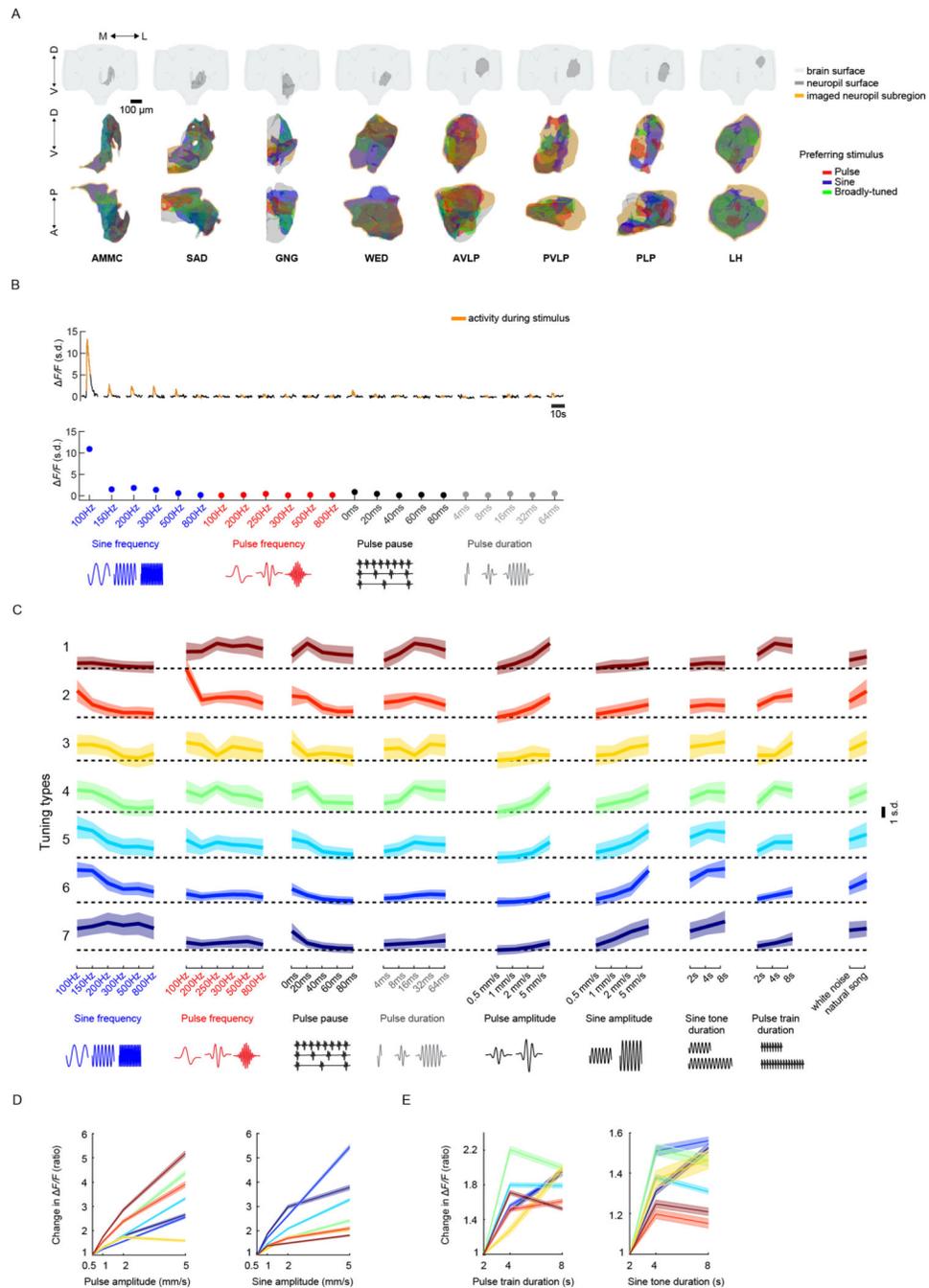
Author Manuscript



**Extended Data Fig. 6 | Similarity of auditory activity between neuropils, and compartmentalization within neuropils.**

**a**, Pairwise Pearson's correlation coefficients of response type distributions (as in Fig. 3b) between neuropils. Neuropils are ordered based on the hierarchical clustering of the correlation matrix. Only positive correlation values are plotted for clarity. **b**, Spatial distribution of auditory responses within selected olfactory (AL, MB-ML, MB-VL, MB-PED, MB-CA, and LH), visual (PVL, and PLP) and mechanosensory neuropils (AVLP). Auditory activity is spatially restricted within each neuropil. Voxel density was calculated

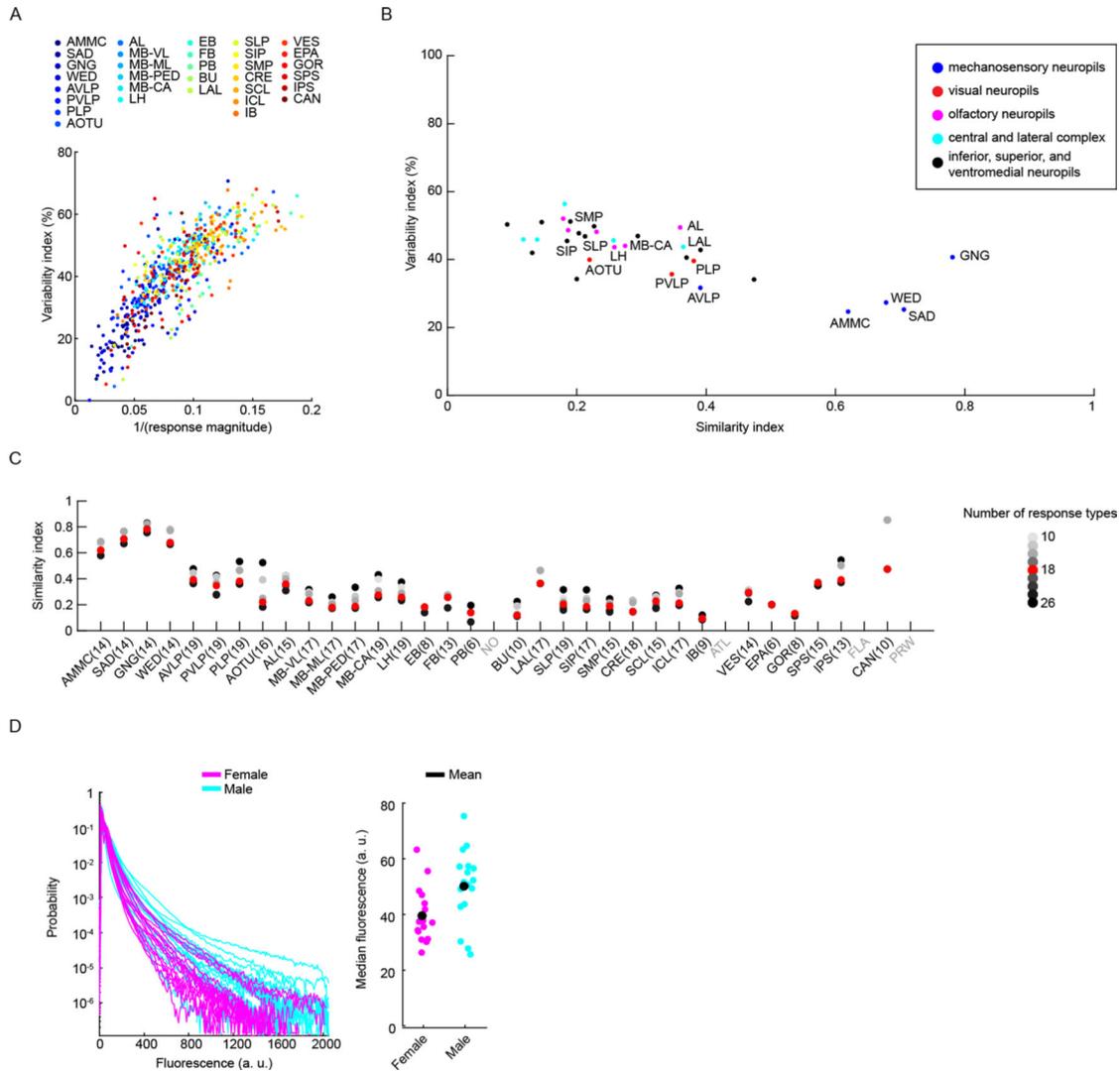
from 19, 15, 17, 17, 22, 19, 18, 30, and 25 flies for the AL, MB-ML, MB-VL, MB-PED, MB-CA, LH, PVLV, PLP, and AVLP respectively.



**Extended Data Fig. 7 | Neuropil subregions imaged, definition of stimulus tuning, and tuning for additional song features.**

**a**, Neuropil volume and subregion of neuropil imaged in Fig. 4. Top row, location of the neuropil imaged (dark grey surface) relative to the central brain (light grey surface). Middle and bottom rows show maximum projections (from two orthogonal views) of imaged neuropil subvolume (orange surface) and the distribution of ROI stimulus selectivity (based

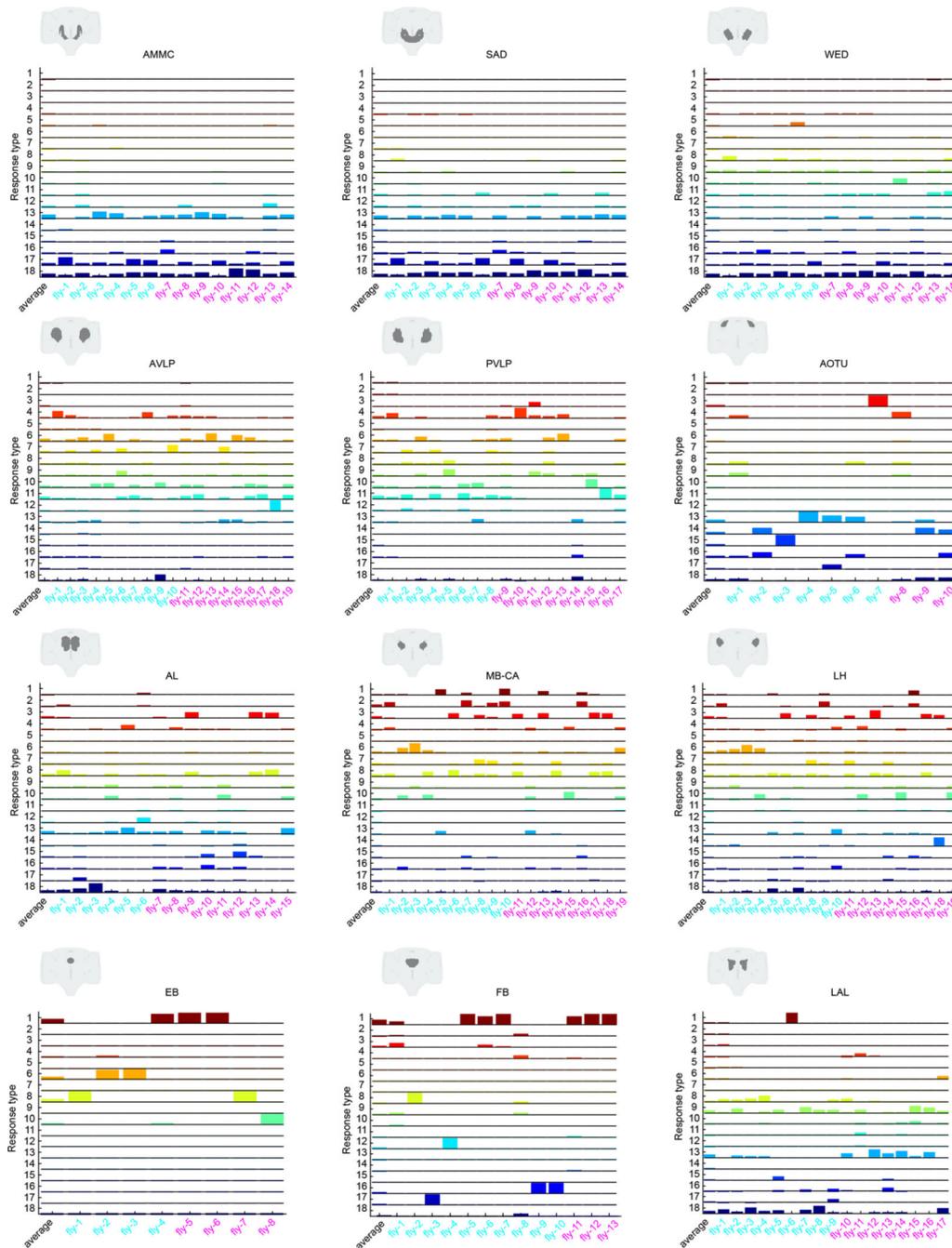
on data in Fig. 3b). The neuropil subvolume (orange) is the volume imaged in at least 2 flies. This volume represents 24.4, 46.2, 10.5, 83, 39.8, 99.3, 63.6, and 91.8 % of the AMMC, SAD, GNG, WED, AVL, PVL, PLP and LH, respectively. **b**, Example responses from one ROI to song feature stimuli (see Fig. 6a). Top panel, Median (across trials) and z-scored responses to each song feature (responses are baseline subtracted, and baseline is defined as activity  $-4$  to  $-0.25$  seconds before stimulus onset). Orange timepoints correspond to activity during the stimulus. Bottom panel, response magnitudes (80th or 20th percentile of activity - for excitatory or inhibitory responses - during stimulus plus 2 seconds after) to each song feature calculated from the top trace (that is the tuning curve for this ROI). All responses are z-scored, so responses are plotted as the s.d. of  $F/F$  value. **c**, Tuning types as in Fig. 6c, but plotting additional responses to pulses and sines of different amplitudes, varying pulse and sine train durations, and also to white noise and natural song (these additional responses were not used to cluster responses). Thick traces are the mean response magnitudes (calculated as in (B)) across all ROIs within each tuning type and shading is the standard deviation (1,783, 513, 739, 1,682, 2,410, 2,321 and 1424 ROIs for each tuning type 1–7 respectively,  $n = 21$  flies). Responses are plotted as the s.d. of  $F/F$  value. **d**, Auditory responses to different sine and pulse intensities, sorted by tuning type. Response magnitudes (calculated as in (B), but 80th or 20th percentile of the activity is measured during stimuli only) per ROI are normalized to the response to 0.5 mm/s stimuli. Thick traces are the mean normalized response magnitude, and shading is the s.e.m (ROI number per tuning type is the same as in (C)). **e**, Auditory responses to different sine and pulse train durations. Response magnitudes (calculated as in (D)) per ROI are normalized to the response to 2 seconds stimuli. Conventions same as in (D).



**Extended Data Fig. 8 |. Across-trial and cross-individual comparisons of auditory activity by neuropil.**

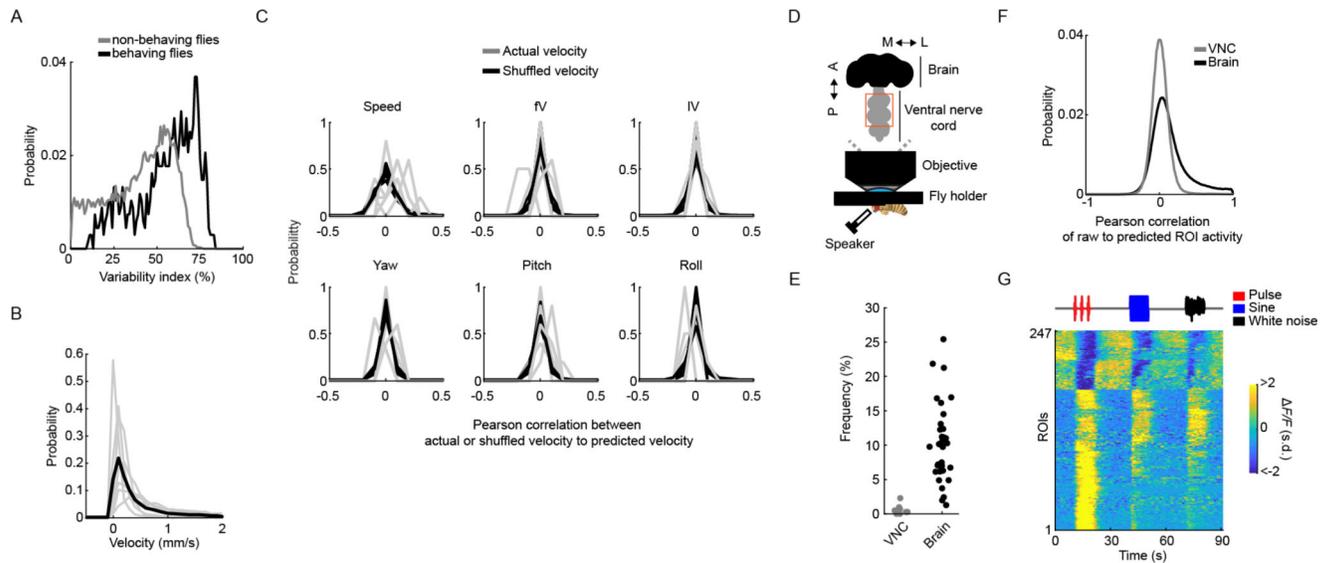
**a.** Comparison of across-trial variability index (see Fig. 5b) and response magnitude. Each dot is the mean variability index and response magnitude (80<sup>th</sup> or 20<sup>th</sup> percentile - for excitatory or inhibitory responses - of z-scored  $F/F$  from stimuli onset until 5 seconds after stimulus offset) per neuropil for each fly (across all ROIs). Neuropils are color coded according to the legend. Variability index is inversely correlated with response magnitude. **b.** Comparison of across-trial variability (see Fig. 5b) and across-individual similarity (see Fig. 5d). Each dot is the mean across-trial variability index and mean across-individual similarity index per neuropil. Groups of neuropils are color coded according to the legend. Early mechanosensory neuropils have low across-trial variability and high across-individual similarity. **c.** Robustness of differences in similarity index across neuropils to the number clusters selected for hierarchical clustering of response types (see Fig. 3a). Similarity index per neuropil (as in Fig. 5d) is calculated for different numbers of clusters (from 10 to 26 - in Fig. 3 we used 18 types (red)). **d.** No systematic difference in distribution of time-average fluorescence (over the entire experiment) across individuals. Left: Histograms of time-

averaged fluorescence per individual, cyan and magenta correspond to male (n = 16) and female (n = 17) flies. Right: Median fluorescence per individual. Black dot corresponds to the mean (of median fluorescence) across flies.



**Extended Data Fig. 9 | Distribution of 18 auditory response types per neuropil, separated by flies.**

Response type distributions, similar to Fig. 5c, for mechanosensory, visual, olfactory, central and lateral complex neuropils. For each neuropil, male and female flies are indicated in cyan or magenta, respectively.



### Extended Data Fig. 10 | Behavior of head-fixed flies and minimal auditory responses in the *Drosophila* ventral nerve cord.

**a.** Distribution of across-trial variability of auditory ROIs in behaving (data used in Fig. 6,  $n = 7$  flies, 4,560 ROIs) and non-behaving flies (data used in Fig. 5b,  $n = 33$  flies, 19,036 ROIs). **b.** Distribution of speed across flies. Each gray line is the speed distribution per fly ( $n = 7$  flies), and black line is the mean across animals. **c.** Correlation between linearly predicted velocities (based on stimuli history) and actual or shuffled (see Methods) velocities. Each black or gray line is the probability per fly ( $n = 7$  flies). **d.** Schematic of VNC functional imaging during auditory stimulation. Similar to Fig. 1a, but the dorsal side of the thorax is dissected to expose the dorsal side of the first and second segment of the VNC depicted inside the orange rectangle. **e.** Frequency of auditory ROIs detected in the brain ( $n = 33$  flies, 185,395 ROIs) and the VNC ( $n = 8$  flies, 39,580 ROIs) using the same criteria as in Extended Data Fig. 1d. **f.** Distribution of Pearson correlation coefficients of raw ROI activity to predicted ROI activity (based on stimuli history, as in Extended Data Fig. 1e) for ROIs recorded from the brain ( $n = 33$  flies, 185,395 ROIs) and the VNC ( $n = 8$  flies, 39,580 ROIs). **g.** Responses of VNC auditory ROIs to pulse, sine, and white noise stimuli ( $n = 8$  flies, 39,580 ROIs). Each row is the across-trial median and z-scored  $\Delta F/F$  response to each stimulus (6 trials per stimulus).  $\Delta F/F$  units are in s.d.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank J. Clemens for assistance on auditory stimuli delivery and linear modeling of calcium responses, as well as helpful discussions on auditory coding; A. Giovannucci for help with ROI segmentation of volumetric calcium signals and guidance with CalmAn toolbox usage; G. Jefferis and T. Rohlfing for help using the image registration toolbox CMTK and the neuroanatomy toolbox natverse; B. Cowley for help with linear modeling of calcium responses in behaving flies; A. Calhoun for help with analysis of the hemibrain connectome dataset; and T. Pereira for assistance with the CPD algorithm. We also thank G. Jefferis, S. Ahmed, C. Baker and J. Clemens for comments on the manuscript. D.A.P. was supported in part by a NSF Physics Frontier Center grant, and M.M. was supported

by a NIH NINDS New Innovator award, NIH BRAIN Initiative R01s NS104899 and NS110060, and a Howard Hughes Medical Institute Faculty Scholar award.

## Data availability

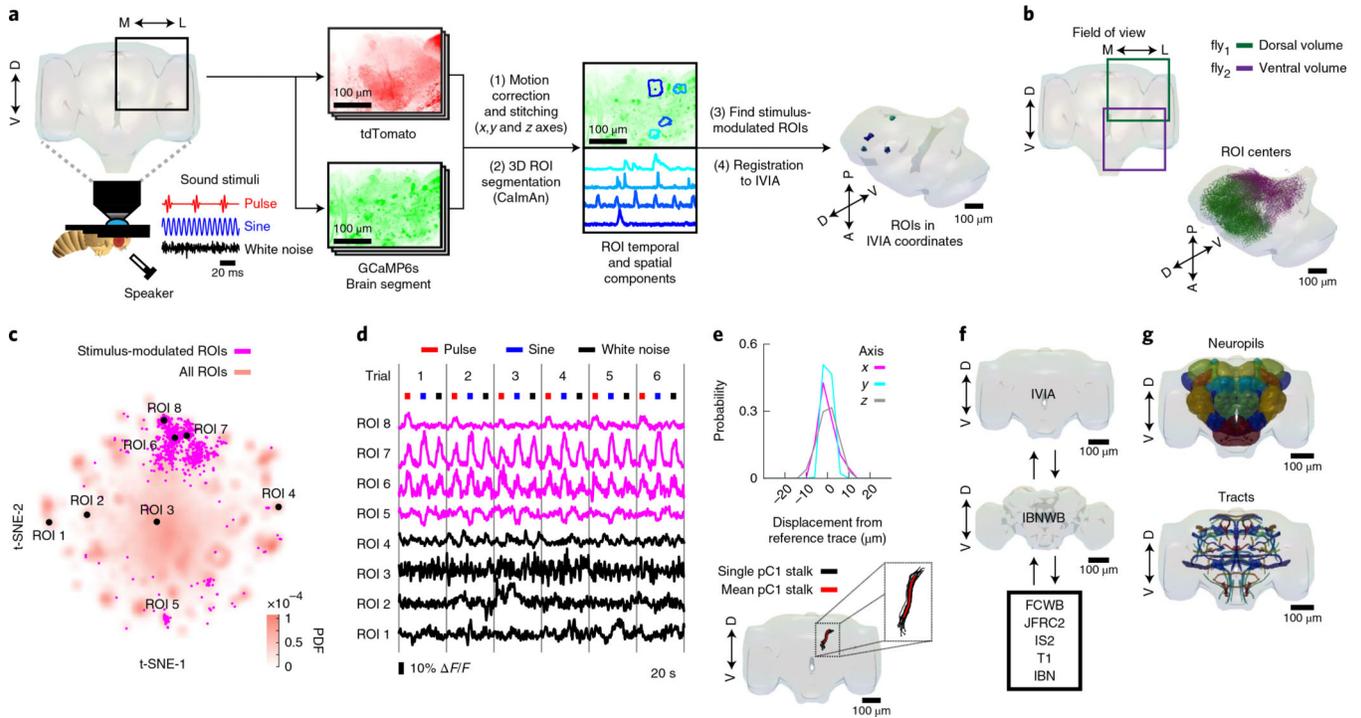
Calcium imaging data from this study (all segmented ROIs) are publicly available at <https://doi.org/10.34770/gv6w-5351>.

## References

1. Mann K, Gallen CL & Clandinin TR Whole-brain calcium imaging reveals an intrinsic functional network in *Drosophila*. *Curr. Biol* 27, 2389–2396.e4 (2017). [PubMed: 28756955]
2. Chen X et al. Brain-wide organization of neuronal activity and convergent sensorimotor transformations in larval zebrafish. *Neuron* 100, 876–890.e5 (2018). [PubMed: 30473013]
3. Bates AS et al. The natverse, a versatile toolbox for combining and analysing neuroanatomical data. *eLife* 9, e53350 (2020). [PubMed: 32286229]
4. Deutsch D, Clemens J, Thiberge SY, Guan G & Murthy M Shared song detector neurons in *Drosophila* male and female brains drive sex-specific behaviors. *Curr. Biol* 29, 3200–3215.e5 (2019). [PubMed: 31564492]
5. von Schilcher F The role of auditory stimuli in the courtship of *Drosophila melanogaster*. *Anim. Behav* 24, 18–26 (1976).
6. Clemens J et al. Connecting neural codes with behavior in the auditory system of *Drosophila*. *Neuron* 87, 1332–1343 (2015); erratum 97, 475 (2018). [PubMed: 26365767]
7. Yoon J et al. Selectivity and plasticity in a sound-evoked male–male interaction in *Drosophila*. *PLoS ONE* 8, e74289 (2013). [PubMed: 24086330]
8. Clemens J et al. Discovery of a new song mode in *Drosophila* reveals hidden structure in the sensory and neural drivers of behavior. *Curr. Biol.* 28, 2400–2412.e6 (2018). [PubMed: 30057309]
9. Coen P et al. Dynamic sensory cues shape song structure in *Drosophila*. *Nature* 507, 233–237 (2014). [PubMed: 24598544]
10. Coen P, Xie M, Clemens J & Murthy M Sensorimotor transformations underlying variability in song intensity during *Drosophila* courtship. *Neuron* 89, 629–644 (2016). [PubMed: 26844835]
11. Göpfert MC & Robert D The mechanical basis of *Drosophila* audition. *J. Exp. Biol* 205, 1199–1208 (2002). [PubMed: 11948197]
12. Yorozu S et al. Distinct sensory representations of wind and near-field sound in the *Drosophila* brain. *Nature* 458, 201–205 (2009). [PubMed: 19279637]
13. Kamikouchi A et al. The neural basis of *Drosophila* gravity-sensing and hearing. *Nature* 458, 165–171 (2009). [PubMed: 19279630]
14. Lai JS-Y, Lo S-J, Dickson BJ & Chiang A-S Auditory circuit in the *Drosophila* brain. *Proc. Natl Acad. Sci. USA* 109, 2607–2612 (2012). [PubMed: 22308412]
15. Vaughan AG, Zhou C, Manoli DS & Baker BS Neural pathways for the detection and discrimination of conspecific song in *D. melanogaster*. *Curr. Biol* 24, 1039–1049 (2014). [PubMed: 24794294]
16. Zhou C et al. Central neural circuitry mediating courtship song perception in male *Drosophila*. *eLife* 4, e08477 (2015).
17. Matsuo E et al. Organization of projection neurons and local neurons of the primary auditory center in the fruit fly *Drosophila melanogaster*. *J. Comp. Neurol* 524, 1099–1164 (2016). [PubMed: 26762251]
18. Patella P & Wilson RI Functional maps of mechanosensory features in the *Drosophila* brain. *Curr. Biol* 28, 1189–1203.e5 (2018). [PubMed: 29657118]
19. Wang K et al. Neural circuit mechanisms of sexual receptivity in *Drosophila* females. Preprint at bioRxiv 10.1101/2020.08.07.241919 (2020).
20. Pnevmatikakis EA et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* 89, 285–299 (2016). [PubMed: 26774160]

21. Aimon S et al. Fast near-whole-brain imaging in adult *Drosophila* during responses to stimuli and behavior. *PLoS Biol* 17, e2006732 (2019). [PubMed: 30768592]
22. Pnevmatikakis EA & Giovannucci A NoRMCorre: an online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* 291, 83–94 (2017). [PubMed: 28782629]
23. Giovannucci A et al. CaImAn an open source tool for scalable calcium imaging data analysis. *eLife* 8, e38173 (2019). [PubMed: 30652683]
24. Yang HH et al. Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell* 166, 245–257 (2016). [PubMed: 27264607]
25. Ito K et al. A systematic nomenclature for the insect brain. *Neuron* 81, 755–765 (2014). [PubMed: 24559671]
26. Myronenko A & Song X Point set registration: coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell* 32, 2262–2275 (2010). [PubMed: 20975122]
27. Tootoonian S, Coen P, Kawai R & Murthy M Neural representations of courtship song in the *Drosophila* brain. *J. Neurosci* 32, 787–798 (2012). [PubMed: 22262877]
28. Azevedo AW & Wilson RI Active mechanisms of vibration encoding and frequency filtering in central mechanosensory neurons. *Neuron* 96, 446–460. e9 (2017). [PubMed: 28943231]
29. Chang AEB, Vaughan AG & Wilson RI A mechanosensory circuit that mixes opponent channels to produce selectivity for complex stimulus features. *Neuron* 92, 888–901 (2016). [PubMed: 27974164]
30. Scheffer LK et al. A connectome and analysis of the adult *Drosophila* central brain. *Elife* 9, e57443 (2020). [PubMed: 32880371]
31. Lehnert BP, Baker AE, Gaudry Q, Chiang A-S & Wilson RI Distinct roles of TRP channels in auditory transduction and amplification in *Drosophila*. *Neuron* 77, 115–128 (2013). [PubMed: 23312520]
32. Senthilan PR et al. *Drosophila* auditory organ genes and genetic hearing defects. *Cell* 150, 1042–1054 (2012). [PubMed: 22939627]
33. Musall S, Kaufman MT, Juavinett AL, Gluf S & Churchland AK Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* 22, 1677–1686 (2019). [PubMed: 31551604]
34. Stringer C et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science* 364, 255 (2019). [PubMed: 31000656]
35. Baker CA et al. Neural network organization for courtship song feature detection in *Drosophila*. Preprint at bioRxiv 10.1101/2020.10.08.332148 (2020).
36. Harris DT, Kallman BR, Mullaney BC & Scott K Representations of taste modality in the *Drosophila* brain. *Neuron* 86, 1449–1460 (2015). [PubMed: 26051423]
37. Naumann EA et al. From whole-brain data to functional circuit models: the zebrafish optomotor response. *Cell* 167, 947–960. e20 (2016). [PubMed: 27814522]
38. Haesemeyer M, Robson DN, Li JM, Schier AF & Engert F A brain-wide circuit model of heat-evoked swimming behavior in larval zebrafish. *Neuron* 98, 817–831. e6 (2018). [PubMed: 29731253]
39. Frechter S et al. Functional and anatomical specificity in a higher olfactory centre. *eLife* 8, e44590 (2019). [PubMed: 31112127]
40. Auer TO & Benton R Sexual circuitry in *Drosophila*. *Curr. Opin. Neurobiol* 38, 18–26 (2016). [PubMed: 26851712]
41. Vogt K et al. Shared mushroom body circuits underlie visual and olfactory memories in *Drosophila*. *eLife* 3, e02395 (2014). [PubMed: 25139953]
42. Jefferis GSXE et al. Comprehensive maps of *Drosophila* higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell* 128, 1187–1203 (2007). [PubMed: 17382886]
43. Bennet-Clark HC & Ewing AW Pulse interval as a critical parameter in the courtship song of *Drosophila melanogaster*. *Anim. Behav* 17, 755–759 (1969).
44. Murthy M, Fiete I & Laurent G Testing odor response stereotypy in the *Drosophila* mushroom body. *Neuron* 59, 1009–1023 (2008). [PubMed: 18817738]

45. Chiappe ME, Seelig JD, Reiser MB & Jayaraman V Walking modulates speed sensitivity in *Drosophila* motion vision. *Curr. Biol* 20, 1470–1475 (2010). [PubMed: 20655222]
46. Fujiwara T, Cruz TL, Bohoslav JP & Chiappe ME A faithful internal representation of walking movements in the *Drosophila* visual system. *Nat. Neurosci* 20, 72–81 (2017). [PubMed: 27798632]
47. Kim AJ, Fitzgerald JK & Maimon G Cellular evidence for efference copy in *Drosophila* visuomotor processing. *Nat. Neurosci* 18, 1247–1255 (2015). [PubMed: 26237362]
48. Batchelor AV & Wilson RI Sound localization behavior in *Drosophila melanogaster* depends on inter-antenna vibration amplitude comparisons. *J. Exp. Biol* 222, jeb191213 (2019). [PubMed: 30733260]
49. Rayshubskiy A et al. Neural control of steering in walking *Drosophila*. Preprint at bioRxiv 10.1101/2020.04.04.024703 (2020).
50. Zheng Z et al. A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* 174, 730–743.e22 (2018). [PubMed: 30033368]
51. Effertz T, Wiek R & Göpfert MC NompC TRP channel is essential for *Drosophila* sound receptor function. *Curr. Biol* 21, 592–597 (2011). [PubMed: 21458266]
52. Morley EL, Jonsson T & Robert D Auditory sensitivity, spatial dynamics, and amplitude of courtship song in *Drosophila melanogaster*. *J. Acoust. Soc. Am* 144, 734 (2018). [PubMed: 30180716]
53. Seelig JD et al. Two-photon calcium imaging from head-fixed *Drosophila* during optomotor walking behavior. *Nat. Methods* 7, 535–540 (2010). [PubMed: 20526346]
54. Moore RJD et al. FicTrac: a visual method for tracking spherical motion and generating fictive animal paths. *J. Neurosci. Methods* 225, 106–119 (2014). [PubMed: 24491637]
55. Murthy M & Turner G Whole-cell in vivo patch-clamp recordings in the *Drosophila* brain. *Cold Spring Harb. Protoc* 2013, 140–148 (2013). [PubMed: 23378646]
56. Maimon G, Straw AD & Dickinson MH Active flight increases the gain of visual motion processing in *Drosophila*. *Nat. Neurosci* 13, 393–399 (2010). [PubMed: 20154683]
57. Chen T-W et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300 (2013). [PubMed: 23868258]
58. Preibisch S, Saalfeld S & Tomancak P Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* 25, 1463–1465 (2009). [PubMed: 19346324]
59. Schindelin J et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682 (2012). [PubMed: 22743772]
60. Park M & Pillow JW Receptive field inference with localized priors. *PLoS Comput. Biol* 7, e1002219 (2011). [PubMed: 22046110]
61. Cachero S, Ostrovsky AD, Yu JY, Dickson BJ & Jefferis GSXE Sexual dimorphism in the fly brain. *Curr. Biol* 20, 1589–1601 (2010). [PubMed: 20832311]
62. Rohlfing T & Maurer CR Jr. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. *IEEE Trans. Inf. Technol. Biomed* 7, 16–25 (2003). [PubMed: 12670015]
63. Chiang A-S et al. Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol* 21, 1–11 (2011). [PubMed: 21129968]
64. Peng H et al. BrainAligner: 3D registration atlases of *Drosophila* brains. *Nat. Methods* 8, 493–500 (2011). [PubMed: 21532582]
65. Yu JY, Kanai MI, Demir E, Jefferis GSXE & Dickson BJ Cellular organization of the neural circuit that drives *Drosophila* courtship behavior. *Curr. Biol* 20, 1602–1614 (2010). [PubMed: 20832315]



**Fig. 1 | A new pipeline for mapping sensory activity throughout the central brain.**  
**a**, Overview of data collection and processing pipeline (for more details see Extended Data Fig. 1a). Step (1): the tdTomato signal is used for motion correction of a volumetric time-series and to stitch serially imaged overlapping brain segments in the  $x$ ,  $y$  and  $z$  axes. Step (2): 3D ROI segmentation (via CalmAn) is performed on GCaMP6s signals. Step (3): auditory ROIs are selected. Step (4): ROIs are mapped to the in vivo intersex atlas (IVIA) space. D, dorsal; L, lateral; M, medial; V, ventral. **b**, Top: an example of 11,225 segmented ROIs combined from two flies (dorsal (green) and ventral (purple) volumes from each fly). Bottom: ROI centers (dots) span the entire anterior–posterior (A–P) and D–V axes. These ROIs have not yet been sorted for those that are auditory (Extended Data Fig. 1d,e). **c**, 2D t-SNE embedding of activity from all ROIs in **b**. PDF, probability density function. ROIs modulated by auditory stimuli (1,118 out of 11,225 ROIs) are shown in magenta. **d**,  $F/F$  from ROIs indicated in **c**. Magenta traces correspond to stimulus-modulated ROIs, while black traces are non-modulated ROIs. Time of individual auditory stimulus delivery is indicated (pulse, sine or band-limited white noise). **e**, Top: IVIA registration accuracy showing the per-axis jitter ( $x$ ,  $y$  and  $z$ ) between traced pC1 stalk values across flies relative to mean pC1 stalk values (see Extended Data Fig. 2a–c for more details). Bottom: 3D rendering of traced stalks of  $Dsx^+$  pC1 neurons (black traces,  $n = 20$  brains from  $Dsx$ – $GAL4/UAS$ – $GCaMP6s$  flies) and mean pC1 stalk (red trace). **f**, Schematic of the bridging registration between the IVIA and the Insect Brain Nomenclature Whole Brain (IBNWB) atlas<sup>25</sup>. This bidirectional interface provides access to a network of brain atlases (FCWB, JFRC2, IS2, T1 and IBN) associated with different *Drosophila* neuroanatomy resources<sup>3</sup>. See Extended Data Fig. 2d–g for more details. **g**, Anatomical annotation of the IVIA. Neuropil and neurite tract segmentations from the IBN atlas were mapped to the IVIA (for a

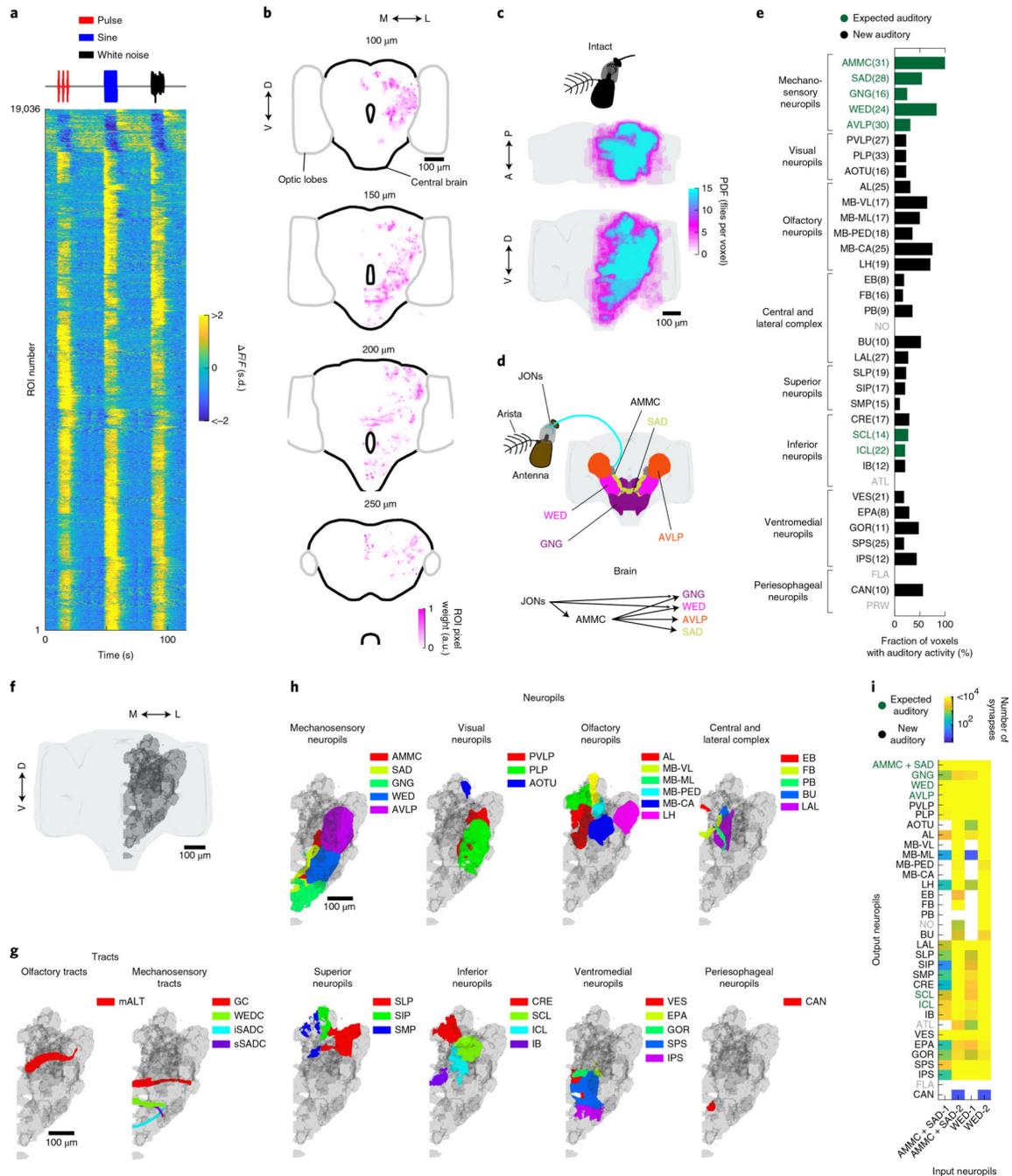
full list of neuropil and neurite tract names, see Supplementary Table 1). See also Supplementary Videos 1–4.

Author Manuscript

Author Manuscript

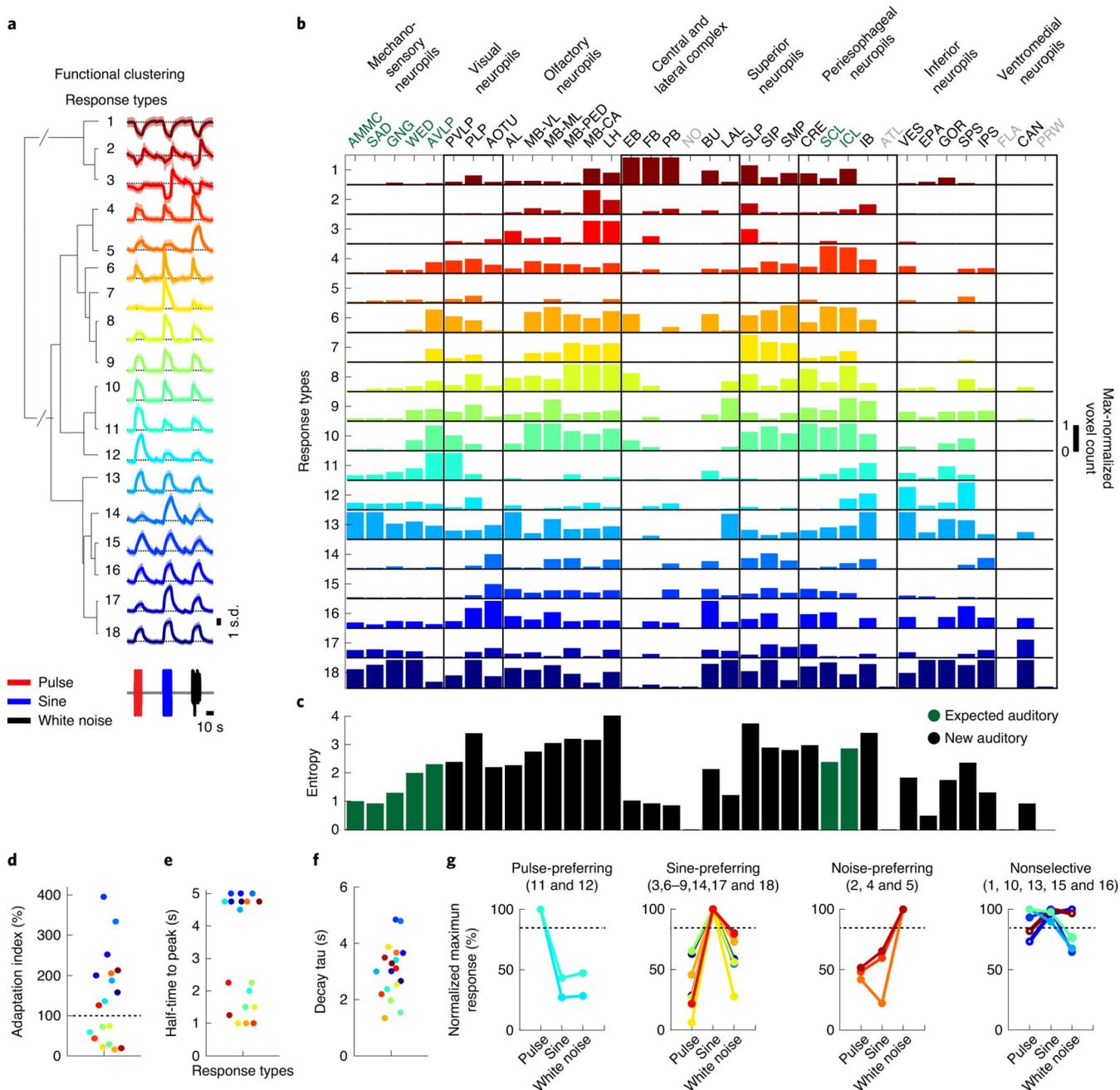
Author Manuscript

Author Manuscript



**Fig. 2 | Auditory activity is widespread throughout the central brain of *Drosophila*.**  
**a.** Auditory ROI responses to pulse, sine and white noise stimuli ( $n = 33$  flies, 19,036 ROIs). Each row is the median across six trials, with all responses  $z$ -scored and therefore plotted as the change in the s.d. of the  $F/F$  signal. ROIs are sorted on the basis of clustering of temporal profiles (Fig. 3a). **b.** The spatial distribution of auditory ROIs (ROI pixel weights (arbitrary units (a.u.)) in magenta) combined from two flies (one ventral and one dorsal volume) (ROIs are in IVIA coordinates). Four different depths are shown (0  $\mu\text{m}$  is the most anterior section of the brain and 300  $\mu\text{m}$  the most posterior). The gray contour depicts the

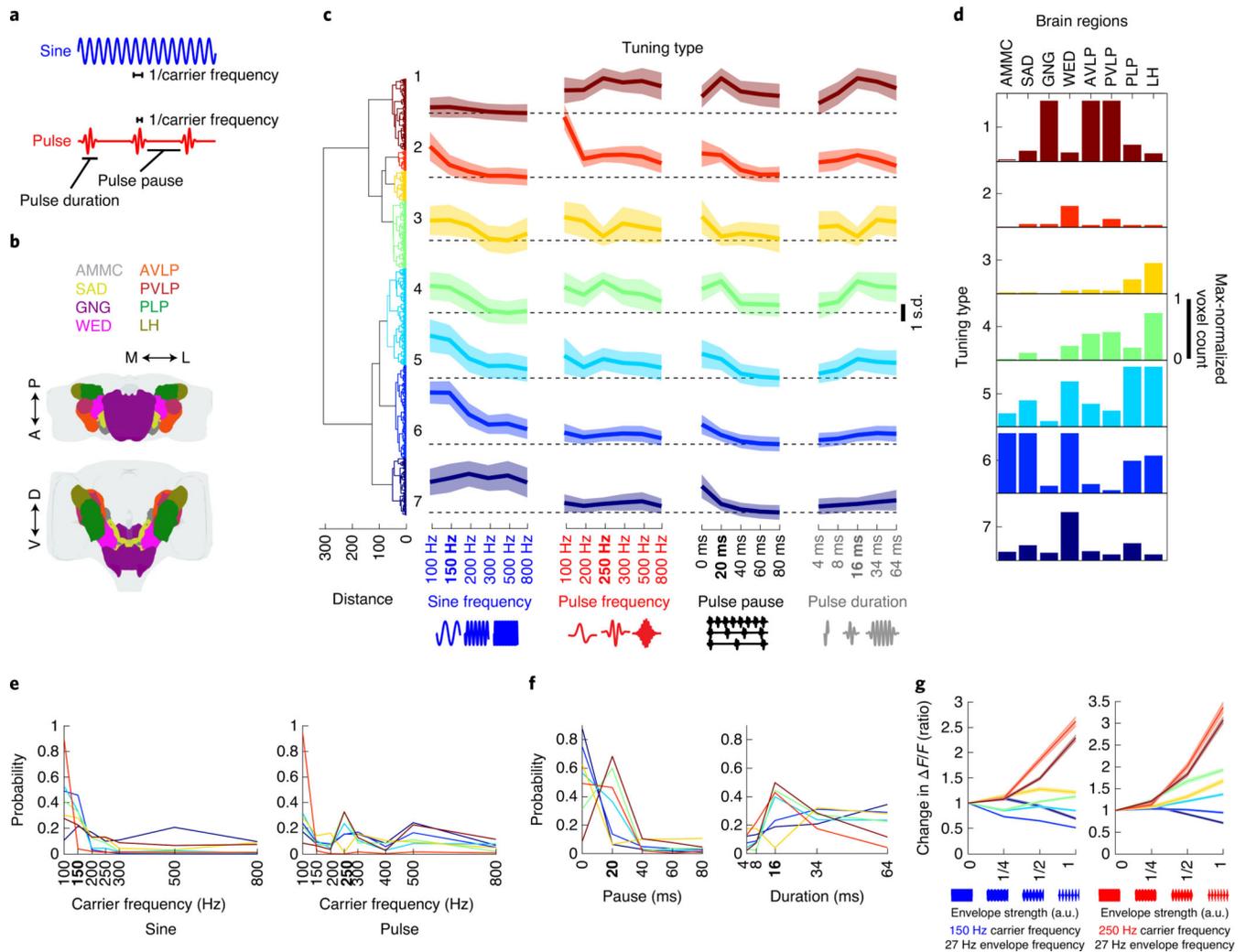
optic lobes, while the black contour depicts the central brain. **c**, Maximum projection (from two orthogonal views) of the density of auditory ROIs ( $n = 33$  flies, 19,036 ROIs). The color scale represents the number of flies with an auditory ROI per voxel. **d**, Schematic of the canonical mechanosensory pathway in *D. melanogaster*. JONs in the antenna project to the AMMC, the GNG and the WED. AMMC neurons connect with the GNG, the WED, the SAD and the AVLP. **e**, The fraction of voxels with auditory activity by central brain neuropil. Percentages are averaged across 33 flies (a minimum of 4 flies with auditory activity in a given neuropil was required for inclusion). The number of flies with auditory responses in each neuropil are indicated in parentheses. Neuropils with no auditory activity are indicated in gray font (also for **i**). **f–h**, 3D rendering of the volume that contains all auditory voxels across all flies (gray) (**f**) and the overlap of this volume with tracts (**g**) and brain neuropils (**h**). **i**, Using the hemibrain connectomic dataset<sup>30</sup>, a heatmap was generated of neuropils in which all identified AMMC/SAD or WED neurons (1,079 and 3,556, respectively) have direct synaptic connections (AMMC+SAD-1 or WED-1) and neuropils with connections to AMMC/SAD or WED neurons via one intermediate synaptic connection (AMMC+SAD-2 or WED-2). Neuropils with fewer than five synapses are shaded white. See also Extended Data Figs. 3 and 4 and Supplementary Video 5. For a full list of neuropil and neurite tract names, see Supplementary Table 1.



**Fig. 3 | Brain-wide auditory activity is characterized by a diversity of temporal response profiles across neuropils.**

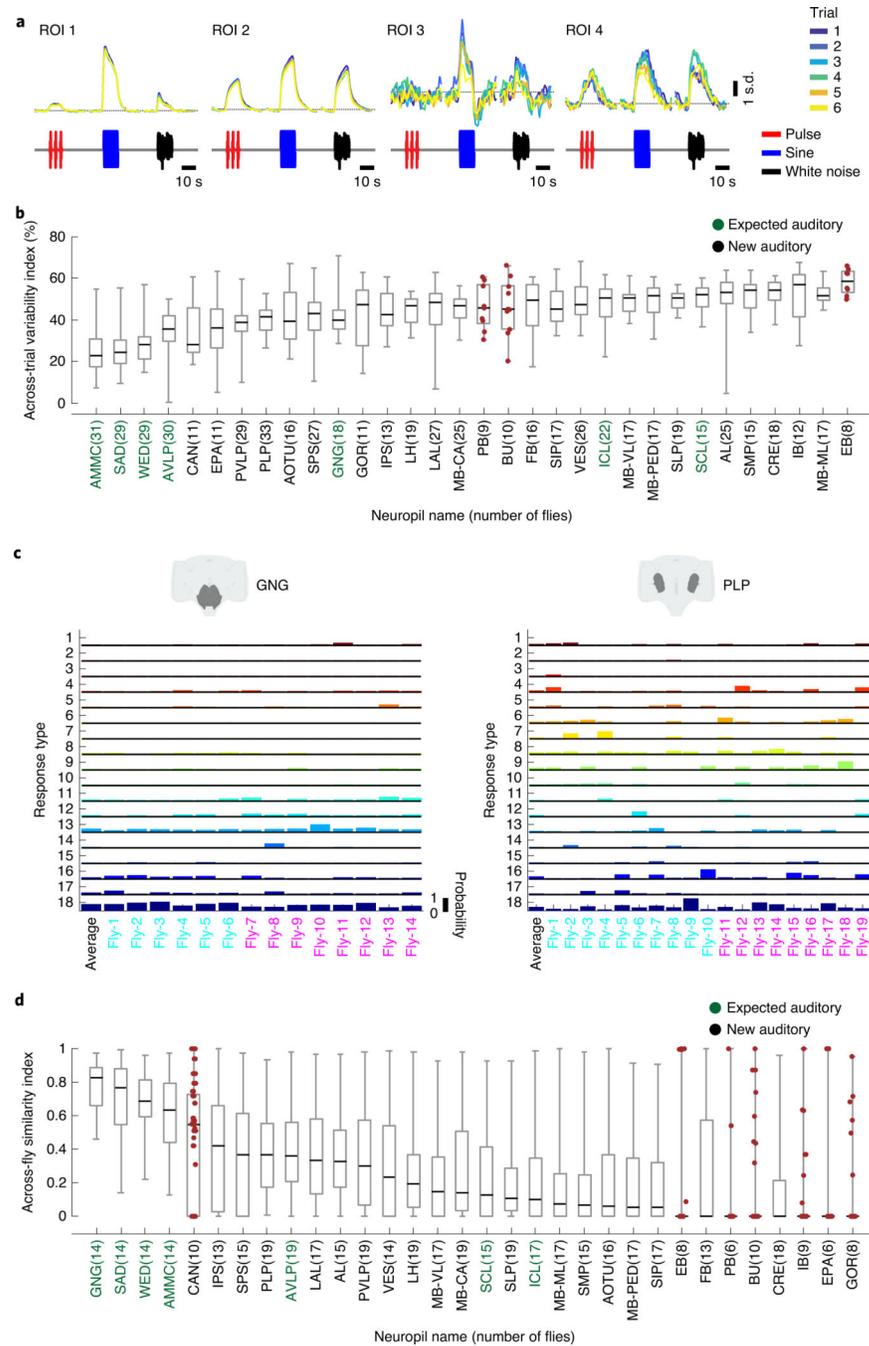
**a**, Hierarchical clustering of auditory responses into 18 distinct response types (Methods). Thick traces are the mean response (across ROIs) to pulse, sine and white noise, and shading is the s.d. ( $n = 33$  flies, 19,036 ROIs). All responses are z-scored, and therefore plotted as the s.d. of the responses over time, with the black dash line corresponding to the mean baseline across stimuli for each response type. **b**, Distribution of 18 response types across 36 central brain neuropils. The histogram of voxel count (how many voxels per neuropil with auditory activity of a particular response type) was max-normalized for each neuropil (normalized per

column). The color code is the same as in **a** ( $n = 33$  flies). Neuropils with no auditory activity are indicated in gray font. **c**, The diversity (measured as the entropy across response-type distributions shown in **b**) of auditory responses across central brain neuropils (Methods). **d–f**, Kinetics of all auditory responses by response type. Adaptation (**d**), half-time to peak (**e**) and decay time tau (**f**). **g**, Diversity in tuning for response types showing pulse-preferring, sine-preferring, noise-preferring and nonselective response types (preference is defined by the stimulus that drives the maximum absolute response (at least 15% greater than the second highest response). Nonselective responses are divided into sine-and-pulse-preferring (filled circles) and sine-and-noise-preferring (open circles). See also Extended Data Figs. 5 and 6.



**Fig. 4 | Widespread auditory activity is tuned to features of conspecific courtship songs.**  
**a**, Spectral and temporal features of the two main modes of *Drosophila* courtship song. **b**, Neuropils imaged. **c**, Hierarchical clustering of auditory tuning curves (see Extended Data Fig. 7a for how tuning curves were generated) into seven distinct tuning types. Tuning types 1–7 comprise 1,783; 513; 739; 1,682; 2,410; 2,321 and 1,424 ROIs, respectively ( $n = 21$  flies). Thick traces are the mean  $z$ -scored response magnitudes (80th or 20th percentile of activity—for excitatory or inhibitory responses—during stimuli plus 2 s after. Activity is plotted as the s.d. of  $F/F$  values to the stimuli indicated on the  $x$  axis (across ROIs and within each tuning type) and shading is the s.d. (across ROIs within tuning type). **d**, *D. melanogaster* conspecific values of each courtship song feature are indicated in bold on the  $x$  axis (same for **e** and **f**). **d**, Distribution of tuning types across regions of selected brain neuropils imaged. The histogram of voxel count per type (number of voxels with activity that falls into each tuning type) was max-normalized for each neuropil (normalized per column). **e**, Probability distributions of best frequency for sine and pulse, separated by tuning type. The color code is the same as in **c**. **f**, Probability distributions of best pulse duration and pulse pause. Conventions are the same as in **e**. **g**, Effects of the amplitude

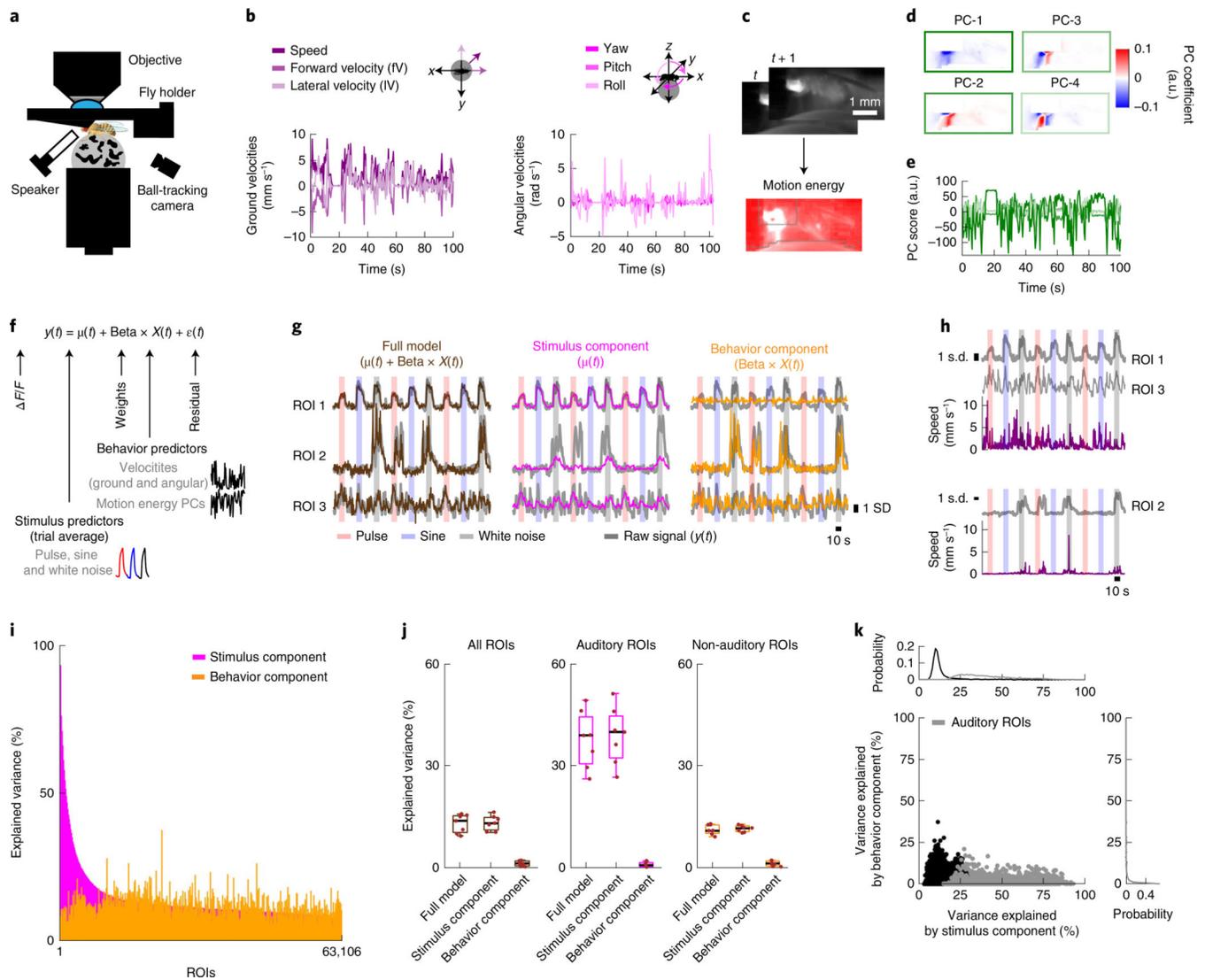
envelope on responses to sine tones. Envelopes (~27-Hz envelope) of different amplitudes were added to sine tones (as depicted on the  $x$  axis; see Methods for more details) of 150 or 250 Hz. ROI response magnitudes to different envelope strengths were normalized to responses to envelope strength 0 (that is, a sine tone of 150 or 250 Hz, with no envelope modulation) and sorted by tuning type from **c**. Thick traces are the mean-normalized response magnitudes per tuning type, and shading is the s.e.m. (ROI number per tuning type is the same as in **c**). See also Extended Data Fig. 7.



**Fig. 5 | Auditory activity is more similar across trials and individuals in early mechanosensory areas.**

**a**, Example auditory responses from five different ROIs to pulse, sine and white noise stimuli. Individual trials are plotted in different colors. All responses are z-scored, and therefore plotted as the s.d. of the responses over time, with the black dashed line corresponding to the mean baseline across stimuli for each ROI. **b**, The across-trial variability index of individual ROIs was computed as the residual of the variance explained by the mean across trials (Methods). The gray box shows the 25th and 75th percentile, the

inner black line is the median variability across flies (number of flies per neuropil are shown in parentheses), and the whiskers correspond to minimum and maximum values. If  $n < 11$ , we plot means per fly with brown dots. **c**, Stereotypy of auditory response types (Fig. 3a,b) across individuals for the GNG and the PLP. Male flies are indicated in cyan and female flies in magenta on the  $x$  axis. Auditory activity in the GNG is similar across flies, while auditory activity in the PLP is more variable, in terms of response type diversity. Average response across all flies is shown in the left-most column. **d**, The across-individual similarity index was computed by measuring the cosine between response type distributions (Fig. 3b) per neuropil across individuals (Methods). Conventions are the same as in **b**. For each neuropil, we computed the similarity index for all possible pairs of flies. Responses in the GNG, SAD, WED and AMMC are the most stereotyped across individuals. See also Extended Data Figs. 8 and 9.



**Fig. 6 | Spontaneous movements do not account for trial-to-trial variability in auditory responses.**

**a**, Schematic of the functional-imaging protocol of a head-fixed fly walking on a ball. **b**, Instantaneous ground and angular velocities measured from ball tracking of a representative tethered fly walking. **c**, Motion energy (absolute value of the difference of consecutive frames) from a video recording of tethered fly. Gray contours correspond to pixels excluded from analysis (Methods). **d**, Coefficients of the top four principal components (PCs) of the motion energy movie (image scale represents low-motion (blue) to high-motion energy (red)) of a representative tethered fly (same as **b**). **e**, Scores of the top four PCs of the motion energy movie (the color of each PC is the same as in box color in **d**) of a representative tethered fly (same as **b**). **f**, Schematic of the linear model used to predict neural activity related to either stimuli or behavior. ROI activity ( $y(t)$ ) is modeled as the sum of the stimulus component (trial average,  $\mu(t)$ ), behavior component ( $\text{Beta} \times X(t)$ ), plus noise ( $\epsilon(t)$ ). **g**, ROI activity along with predictions by stimulus component ( $\mu(t)$ ), behavior component ( $\text{Beta} \times X(t)$ ) or the full model for three example ROIs most strongly modulated by auditory stimuli

(ROI 1), by behavior (ROI 2) or by the combination of both (ROI 3). The ROI activity values were  $z$ -scored, and  $FF$  units are in s.d. (also for **h**). **h**, Activity of ROIs from **g** along with speed. ROIs from the same fly are plotted together. **i**, Explained variance by stimulus component or behavior component for individual ROIs, sorted by stimuli-only model performance ( $n = 7$  flies, 63,106 ROIs). **j**, Mean explained variance by stimulus component ( $\mu(t)$ ), behavior component ( $\text{Beta} \times X(t)$ ) or both ( $\mu(t) + \text{Beta} \times X(t)$ ) across flies, for all ROIs (left), auditory ROIs only (middle) and non-auditory ROIs (right). The box shows the 25th and 75th percentile, the inner black line is the median explained variance across flies, and whiskers correspond to minimum and maximum values. Means per fly are plotted with brown dots ( $n = 7$  flies). **k**, Explained variance by stimulus component versus behavior component for individual ROIs. Gray dots correspond to auditory ROIs (using the same criteria as in Extended Data Fig. 1d). See also Extended Data Fig. 10.