bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title: Auditory Activity is Diverse and Widespread Throughout the Central Brain of *Drosophila*

Authors: Diego A Pacheco[1], Stephan Y Thiberge[1,2], Eftychios Pnevmatikakis[3], Mala Murthy[1,2,*]

Affiliations:
[1] Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08540, USA
[2] Bezos Center for Neural Circuit Dynamics, Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08540, USA
[3] Center for Computational Mathematics, Flatiron Institute, Simons Foundation, New York, NY, 10010, USA

*lead contact
Correspondence: mmurthy@princeton.edu

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 analysis of sensory activity towards the earliest layers of sensory processing or to brain regions containing the majority of postsynaptic neurons. Here, we present new methods for unbiased volumetric neural imaging with precise across-brain registration, to characterize auditory activity throughout the entire central brain of *Drosophila*, and to make comparisons across trials, individuals, and sexes. We discover that auditory activity is widespread across nearly all central brain regions and in neurons known to carry responses to other modalities. These auditory representations are diverse in their temporal profiles, but the majority of activity, regardless of brain region, is focused on aspects of the conspecific courtship song. We find that auditory responses are stereotyped across trials and animals in early mechanosensory regions, becoming more variable at higher layers of the putative pathway. This study highlights the power of using an unbiased, brain-wide approach for mapping the functional organization of sensory activity.

INTRODUCTION:

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 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 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 (Ahrens et al. 2013; Mann et al. 2017; Schrödel et al. 2013; Chen et al. 2018). Thus, neurons and brain areas may be more multi-modal than appreciated. Here, we present methods to systematically characterize sensory responses throughout the *Drosophila* brain, for which abundant neural circuit tools facilitate moving from activity maps to genetic reagents that target identifiable neurons (Jenett et al. 2012; Tirian & Dickson 2017).

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 sensory and second order mechanosensory neurons. In addition, in contrast with olfaction or vision, for which the range of odors or visual patterns flies experience in naturalistic conditions is large, courtship

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

song is the main auditory cue to which flies respond, and it is comprised of a narrow range of species-specific features that can be easily probed in experiments (Arthur et al. 2013; Jan Clemens, Deutsch, et al. 2018). This provides an opportunity to deliver a small number of stimuli to cover the full range of behaviorally-relevant sounds. During courtship, males chase females while producing an acoustic signal via unilateral wing vibration. This behavior unfolds over many minutes, and females must extract features from the songs males produce to inform their mating decisions (von Schilcher 1976; Jan Clemens, Girardin, et al. 2018) while males listen to songs to inform their courtship decisions in group settings (Eberl et al. 1997; Yoon et al. 2013; Vaughan et al. 2014). Fly song has only three modes, two types of brief sound impulses or 'pulse song' and one hum-like 'sine song' (Bennet-Clark & Ewing 1967; J. Clemens et al. 2018), and males both alternate between these modes to form song bouts and modulate song intensity based on sensory feedback (Coen et al. 2014; Coen et al. 2016). Contained within the pulse and sine songs are species-specific patterns (Riabinina et al. 2011; Arthur et al. 2013).

Flies detect sound using a feathery appendage of the antenna, the arista, which vibrates in response to near field sounds (Göpfert & Robert 2002; Göpfert & Robert 2001). Antennal displacements activate mechanosensory Johnston's Organ neurons (JONs) housed within the antenna (Ewing 1978; Göpfert & Robert 2002). Three major populations of JONs (A, B, and D) respond to vibratory stimuli at frequencies found in natural courtship song (Yorozu et al. 2009; Kamikouchi et al. 2009; Matsuo et al. 2014; Ishikawa et al. 2017). These neurons project to distinct areas of the antennal mechanosensory and motor center (AMMC) in the central brain (Kamikouchi et al. 2006). Downstream targets of the AMMC are being worked out (Matsuo et al. 2016), and 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) (Lai et al. 2012; Vaughan et al. 2014; Zhou et al. 2015; Matsuo et al. 2016; Jan Clemens, Deutsch, et al. 2018; Patella & Wilson 2018). 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 on female brains, even though both males and females process courtship song information (Jan Clemens, Deutsch, et al. 2018; Zhou et al. 2015; Vaughan et al. 2014). This hampers the evaluation of sex-specific differences in auditory processing.

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 and sexes, since the extent of across-animal and across-sex variability in sensory activity is not well characterized but has a major impact on how sensory neurons are defined. We image activity volumetrically combined with fully automated segmentation of regions of interest (Pnevmatikakis et al. 2016). We use GCaMP6s (Chen et al. 2013) and two-photon microscopy to sequentially target the entirety of the *Drosophila* central brain *in vivo* and at cellular resolution (1.4x1.2x2 μm^3 voxel size). In contrast with recent brain-wide imaging studies in *Drosophila* (Mann et al. 2017; Aimon et al. 2019), we have traded off temporal speed for enhanced spatial resolution. Imaging at cellular resolution facilitates automated region of interest (ROI) segmentation, each ROI covering sub-neuropil 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 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. However, auditory maps are largely similar between male and female

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

93 brains, despite extensive sexual dimorphisms in neuronal number and morphology. We then exhaustively
94 map song feature tuning within brain regions that carry the majority of auditory activity and find that
95 responses are focused on conspecific features of courtship song. These findings provide the first brain-
96 wide description of sensory processing and feature tuning in *Drosophila*.

97 RESULTS:

98 **Mapping auditory activity throughout the central brain at cellular resolution**

99 To map auditory responses throughout the central brain of both male and females flies, we first developed
100 a pipeline to volumetrically image sensory responses (via GCaMP6s (Chen et al. 2013) expressed pan-
101 neuronally), and then to precisely register sequentially imaged volumes into an *in vivo* template (using
102 the structural marker membrane-tagged tdTomato (Shaner et al. 2004)) [Figure 1A]. We presented three
103 distinct auditory stimuli, 6 trials each: 1) sinusoids of 150 Hz, 2) impulses with 250 Hz carrier frequency
104 and 36 ms inter-pulse interval (IPI), and 3) white noise filtered between 80-1000 Hz (see Methods;
105 stimulus intensities were comparable to song intensities during natural courtship (Morley et al. 2018)).
106 The first two stimuli represent the two major modes of *Drosophila melanogaster* courtship song (Shorey
107 1962; von Schilcher 1976) - these stimuli were used to broadly map auditory responses - the last stimulus
108 (broadband noise) was used to identify auditory activity responsive to other frequencies present in
109 courtship song. These stimuli should insignificantly drive neurons that encode slow vibrations, such as
110 wind or gravity responses (Yorozu et al. 2009; Kamikouchi et al. 2009). For each fly, we imaged 14-17
111 subvolumes of 300x300x20 μm^3 at a temporal resolution of 1 Hz (1.4x1.2x2 μm^3 voxel size) until covering
112 the whole extent of the posterior-anterior axis per fly (the full volume scanned per fly was half of a
113 hemibrain of ~300x300x250 μm^3 - what we refer to hereafter as a "volume"). These volumes were time-
114 series motion corrected and stitched along the posterior-anterior axis using the tdTomato signal
115 (Pnevmatikakis & Giovannucci 2017) [Figure 1A-i]. We imaged volumes from dorsal and ventral quadrants,
116 providing full hemisphere coverage by imaging only two flies [Figure 1B].
117
118 To automate the segmentation of regions of interest (ROIs) throughout the brain, we used a constrained
119 non-negative matrix factorization algorithm (CNMF) optimized for 3D time series data (Pnevmatikakis et
120 al. 2016; Giovannucci et al. 2019) (see Methods) [Figure 1A-ii]. This method connects contiguous pixels
121 that are similar in their temporal responses; thus, each ROI is a functional unit. We remain agnostic as to
122 the specific relationship between ROIs and neurons - a single ROI could represent activity from many
123 neurons (e.g., a cluster of axon tracts with similar responses), or a single neuron can be segmented into
124 multiple ROIs if different neuronal compartments differ in their activity. Importantly, this method enables
125 effective segmentation of responses from neuropil (axons, dendrites, and processes), not only somas -
126 this is critical for capturing sensory responses in *Drosophila*, for which somas may not reflect activity in
127 processes (Yang et al. 2016; Gouwens & Wilson 2009). We typically extracted thousands of ROIs per
128 volume, and collectively sampled activity from all central brain regions [Figures 1C and S1A-B]. These ROIs
129 included both spontaneous and stimulus-driven activity [Figure 1C]; stimulus-driven ROIs were identified
130 using quantitative criteria (see Methods) [Figures 1A-iii and S1C-E]. These criteria bias towards detection
131 of ROIs with responses consistent across trials [Figure 1D]. We found hundreds of auditory ROIs per fly
132 which range from 1.2% to 25.4% of all segmented ROIs per fly [Figure S1F].
133
134 To compare activity across individuals and sexes, we next registered functionally imaged brain segments
135 into an *in vivo* intersex atlas using the computational morphometry toolkit (CMTK) (see Methods) [Figure
136 1A-iv; Figure S2A-B]. Recently, an *in vivo D. melanogaster* female atlas was built (Mann et al. 2017),

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

however it does not account for known anatomical differences between female and male brains (Cachero et al. 2010). Registration of individual fly brains to our *in vivo* intersex atlas (IVIA) is at cellular accuracy (~4.3 um, across all tracts), measured as the deviation from the mean trace of several identifiable neural tracts (e.g., pC1 neurons, Figure 1E, and pC2l, pC2m, pMP7 neurons, and two additional Dsx+ tracts, Figure S2C (Philipsborn et al. 2011; Rideout et al. 2010; Kimura et al. 2015; Zhou et al. 2014)). Using this approach, we successfully registered volumes from 45 out of 48 flies (see examples in Figure S2D).

The IBNWB *Drosophila* fixed brain atlas (Ito et al. 2014) contains detailed brain neuropil and neurite tract segmentation information. By linking our *in vivo* intersex template (IVIA) to the IBNWB atlas, we could map auditory activity onto known brain neuropils and processes [Figure 1F-G], and thereby link to the whole network of *D. melanogaster* anatomical repositories (Cachero et al. 2010; Chiang et al. 2011; Jenett et al. 2012; Manton et al. 2014; Tirian & Dickson 2017). We built a bridge registration between IVIA and the IBNWB atlas using a point set registration algorithm (Myronenko & Song 2010) with high accuracy (deviation of ~2.24 um for IVIA-to-IBNWB, and ~2.8 um for IBNWB-to-IVIA) (see Methods, [Figure S3A-C]).

In summary, we have developed a novel pipeline for monitoring, anatomically annotating, and directly comparing sensory-driven activity, at cellular resolution, across trials, individuals, and sexes. All code for the pipeline is available at https://github.com/dpacheco0921/CaImProPi.

Auditory activity is widespread across the central brain

We extracted 19,036 auditory ROIs from 33 female and male brains [Figure 2A]. Within a single fly, auditory ROIs were broadly distributed across the anterior-posterior and dorsal-ventral axes [Figure 2B]. To evaluate the across-individual consistency of this broad auditory-driven activity, we generated a probability density map across flies (both male and female) using the ROIs' spatial components (see Methods) [Figure 2C]. We then evaluated the distribution of auditory ROIs by brain neuropil or neurite tract (77% of all segmented ROIs were assigned to an identifiable neuropil or tract [Figures 2D-F and S4A-B]). Surprisingly, 33 out of 36 central brain neuropils and at least 5 out of 44 tracts contained robust auditory activity (see Methods).

As a validation of our method, we found that the AMMC, SAD, WED, AVLP, and GNG (previously referred as the subesophageal zone (SEZ) or ganglion (SOG)) all contained auditory activity [Figures 2D-E] - auditory or mechanosensory neurons innervating each of these neuropils were described previously (Lai et al. 2012; Patella & Wilson 2018; Tootoonian et al. 2012; Jan Clemens, Girardin, et al. 2018; Zhou et al. 2015; Vaughan et al. 2014; Azevedo & Wilson 2017; Chang et al. 2016). Some higher-order auditory neurons (e.g., pC2 (in males and females), pMN2 and pC1 (in females)) innervate sub compartments of the lateral protocerebral complex (LPC) (Jan Clemens, Deutsch, et al. 2018). As expected, we observed auditory ROIs in neuropils overlapping with the LPC (which includes the inferior (ICL), the superior clamp (SCL), and the anterior ventrolateral protocerebrum (AVLP)). Similarly, we detected auditory ROIs in many tracts or commissures (great commissure (GC), wedge commissure (WEDC), inferior (iSADC) and superior saddle commissure (sSADC)) connecting known mechanosensory neuropils (AMMC, WED, VLP, and SAD) [Figures 2Fiii and S4C-D]. To confirm that this activity was in fact driven by mechanosensory receptor neurons, we imaged flies carrying the *iav*^{1} mutation, which impairs Johnston's Organ (JO) neuron mechanotransduction, rendering flies deaf (Lehnert et al. 2013; Senthilan et al. 2012; Gong et al. 2004). This manipulation dramatically reduced auditory-driven activity in the central brain, and we found only 21 auditory ROIs from one out of four flies imaged, a 99.1 % reduction in activity [Figures S4E and S4F]. In addition, response magnitudes of these ROIs were significantly reduced compared with wild type flies

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[Figure S4G]. While we do not yet know the origin of this residual activity in the absence of mechanotransduction, we conclude that the overwhelming majority of auditory activity we map with our stimuli are transduced through the JO.

We found that a number of neuropils and tracts never previously described as auditory contained a large percentage of auditory responses (fraction of voxels with auditory activity [Figure 2E]) - for example, neuropils of the olfactory pathway (antennal lobe (AL), the mALT 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 (PVLP and PLP), and the central complex (ellipsoid body (EB), fan-shaped body (FB), protocerebral bridge (PB), bulb (BU), and lateral accessory lobe (LAL)) [Figures 2E and 2F]. Some of this new activity was in neuropils known to be downstream targets of the AMMC but never previously studied functionally (e.g., the inferior posterior slope (IPS) and gorget (GOR) (Matsuo et al. 2016)). We confirmed that activity within the olfactory pathway was present within the somas of individual sensory neurons, such as projection neurons and Kenyon cells [Figures S5A and S5B]. We also showed that auditory activity is spatially restricted within these newly-defined auditory neuropils [Figure S5C], indicating that only subregions of each of these areas receives and processes auditory signals. Below we characterize the temporal responses of all auditory activity throughout the central brain.

Diverse auditory activity throughout the central brain

We next hierarchically clustered auditory responses based on their temporal profiles (see Methods), and identified 18 distinct stimulus-locked response types [Figure 3A]. We measured the frequency of response types for each neuropil to evaluate differences in auditory representations by brain region [Figure 3B]. Auditory activity is diverse throughout the central brain, consisting of both inhibitory (types 1-3) and excitatory (types 5-18) responses, along with diversity in response kinetics [Figures 3C-E]. Response types differed in their tuning for auditory stimuli [Figure 3F]: responses showed a strong preference for pulse [Figure 3-F-i], sine [Figure 3-F-ii], white noise [Figure 3-F-iii], or were non-selective [Figure 3-F-iv]. We interpret white noise preferring responses as a preference for frequencies outside of the range of the two fly courtship song stimuli we tested (see also Figure 6). The diversity of response types found suggests that our method samples activity from many different neurons per neuropil [Figure 3B]. In addition, by comparing probability density maps, we found that each temporal response type had a distinct spatial distribution [Figure S6A-B]. Finally, 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 [Figures 2D and 3G]. Response diversity was highest in newly discovered auditory areas, such as the PLP, LH, and SLP. We also compared response types between neuropils and found strong similarity between groups of neuropils - for example, between the AMMC and SAD, between the GNG, WED, and GOR, or between the AVLP and PVLP [Figure 3H]. This similarity may indicate we are sampling neurons that innervate multiple neuropils.

Within the previously defined mechanosensory pathway [Figure 2D], we found that the AMMC contained predominantly non-selective (types 13 and 16) and sine-preferring (type 18) responses, but that brain regions downstream of the AMMC, such as the WED and GNG, showed changes in the distribution of response types, such as a decrease in non-selective responses and an increase in both white noise- and pulse-preferring responses [Figure 3B]. This result is similar to results from (Patella & Wilson 2018), which also found that auditory coding is transformed from the AMMC to the WED. Downstream of the AMMC and WED, we found that auditory representations changed more dramatically, with a decrease in non-selective responses and an overrepresentation of stimulus-selective responses. While recordings from some AVLP intrinsic neurons (vPN1, V1, and V2) and AMMC-to-AVLP projection neurons (AV1-6) have

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suggested that the VLP mostly encodes pulse-like stimuli (Lai et al. 2012; Zhou et al. 2015; Jan Clemens, Girardin, et al. 2018), we found, in contrast, strong representation of both sine- and pulse-preferring response types, with responses becoming more narrowly tuned in these regions (e.g., response types 7 and 11). In the SCL and ICL, as expected from innervations of pulse-tuned neurons (pC2, pMN2, and pC1), there are a number of pulse-preferring response types (types 11 and 12). However, the predominant responses are sine-preferring, suggesting that the lateral junction does not solely represent pulse song (Jan Clemens, Deutsch, et al. 2018). Our results thus reveal that pulse and sine representations are distributed across early and higher-order neuropils associated with the auditory pathway, with stimulus-tuned responses (whether sine or pulse) becoming more predominant in higher-order neuropils.

In the olfactory system, activity was dominated by non-selective and sine-preferring responses, while pulse-preferring responses were mostly absent [Figures 3B, S5A and S5B]). Interestingly, while auditory responses were mainly excitatory in the AL, MB-ML, MB-VL, and MB-PED, inhibitory responses dominated at the MB-CA and LH; in addition, response types became more diverse in higher-order olfactory regions [Figure 3G]. We explore tuning in the lateral horn (LH) further below.

Although the PVLP, PLP and AOTU have been considered to be primarily visual neuropils (Otsuna & Ito 2006), we found that they also contained strong and diverse auditory responses [Figure 3B]. Auditory activity in the PVLP (which is overall similar to the AVLP [Figure 3H]) is concentrated in two compartments [Figure S5C]. The first compartment is located in the vicinity of ventro-medial optic glomeruli (LC4, LPLC1 and LPLC2), which is consistent with reported projections from AMMC neurons to glomerulus LC4 (Matsuo et al. 2016; Pézier et al. 2014). The second compartment continues from the AVLP auditory ROIs in a seemingly continuous tract. On the other hand, auditory ROIs in the PLP 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 [Figure S5C]. The PLP was dominated by sine-preferring and non-selective responses, and its activity profile differed from both the WED and the LH [Figure 3H].

We also found robust auditory activity in several pre-motor neuropils. For example, the central complex (EB, FB and PB) contained primarily non-selective and inhibitory responses (type 1), while the lateral complex (BU and LAL) contained both sine-tuned (types 18 and 6) and non-selective excitatory responses (types 16 and 13) [Figure 3B]. Similarly, neuropils with innervation from descending neurons that modulate locomotion contained a more diverse set of responses, relative to the central complex [Figure 3G].

Courtship song contains additional frequencies and patterns not present in the synthetic sine and pulse song we used as stimuli [Figure S7A and S7B]. To determine if we missed auditory responses sensitive to these additional parameters, we recorded auditory responses from 10 flies presented with the three stimuli used above in addition to recorded courtship song [Figure 3I] We observed that all auditory ROIs with responses to natural song, also responded to one of the other three stimuli, and that the majority of auditory responses (78 %) to natural song were similar or weaker compared with responses to pulse, sine or white noise stimuli (measured as the maximum signal from stimulus onset till 5 seconds after stimulus offset). However, ~12 % of auditory response subtypes responded more strongly to natural song [Figures S7C-E], and this activity was consistent across flies in the PLP, LH and MB-CA [Figure S7D]. Finally, a second pulse type (Pfast, with ~400 Hz carrier frequency) has been reported in *D. melanogaster* and other sibling species (J. Clemens et al. 2018). It is therefore possible that we missed auditory activity with narrow tuning to pulses with higher carrier frequencies. To address this possibility, we recorded auditory responses from flies presented with the three standard stimuli plus a synthetic P_{fast} stimulus [Figure S7F]. Despite the significant difference in carrier frequency (delta of ~150 Hz) between the two pulse stimuli, we observed

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that subclusters with strong responses to P_{slow}, also responded to P_{fast}, with some differences in strength of response [Figure S7G and S7H]. In other words, P_{fast} also drives the majority of pulse-responsive ROIs. Therefore our three initial stimuli [Figure 1A] were sufficient to sample auditory activity driven by the courtship song.

Comparing auditory activity across trials, individuals, and sexes

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 variance in response magnitude across trials for all auditory ROIs and across brain neuropils (see Methods for definition of variability index) - some ROIs were stereotyped across trials, whereas others were more variable [Figure 4A]. We found that early mechanosensory neuropils AMMC, SAD, WED, and AVLP [Figure 2D] contained the lowest across-trial response variance [Figure 4B]. All other neuropils exhibited higher trial-to-trial variability. Interestingly, the magnitude of auditory responses (80^{th} z-scored DF/F) is inversely proportional to trial-to-trial variability [Figure S8A]. Therefore, some of the differences in trial-to-trial variability could be explained by differences in signal-to-noise ratio of auditory responses across neuropils.

That responses of a particular type are found in the same spatial locations across flies [Figure S6B] already suggests a high degree of across-individual stereotypy of auditory responses. To quantify this, we next used a similarity index (see Methods) to compare activity profiles per neuropil (as in Figure 3B), but now across individuals [Figures 4C-D and S9]. Some neuropils, like the GNG, contain ROIs with temporal responses that are highly similar across individuals (e.g., all individuals have a high frequency of response type 18 in the GNG), whereas other neuropils, like the PLP, have more variable activity across individuals. Comparing the distribution of the similarity index across all neuropils revealed that early mechanosensory neuropils (AMMC, SAD, GNG, and WED) have the highest similarity in auditory responses across individuals [Figure 4D]. Similarity decreases in downstream neuropils, (AVLP and PVLP), and further decreases in neuropils outside the canonical mechanosensory system (e. g., visual, olfactory, and pre-motor neuropils). These differences in similarity across neuropils are robust to changes in how auditory responses are clustered [Figure S8B]. We find that across-trial variability and across-fly stereotypy are strongly correlated specifically in early mechanosensory neuropils [Figure S8C].

Large anatomical differences between female and male fly brains have been reported for fixed brains (Cachero et al. 2010). Similarly, using deformation-based morphometry, we found anatomical differences for *in vivo* brains between sexes [Figures 5A and S10A-B]. While male-enlarged regions (MER) were consistent with previous work (Cachero et al. 2010; Manton et al. 2014), we found that female enlarged regions (FER), mostly located in the GNG, do not overlap with previously identified regions in fixed brains. We observed that the GNG has an average 5% increase of its normalized volume (relative to central brain volume) in females compared with males, while the AOTU has an average ~3.7% decrease (smaller, but still significant, differences were also found in the MB-VL, MB-PED and ICL) [Figures S10C-D].

We then compared auditory activity across sexes, and, despite sex differences in brain morphology, observed extensive response similarity [Figures 5B and 5C]. We evaluated differences in response type distribution between sexes and across neuropils [Figures 5D and 5E] (see Methods). This comprehensive analysis revealed that although there were differences in the GNG, SAD, and LAL (enrichment of sine tuned responses in males (response type 18) or non-selective responses in females (response type 13)), these differences were not significant ($p > 0.05$) [Figures 5D and 5E]. Even though the AVLP has a male-specific

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

neuron that is pulse-tuned (vPN1; (Zhou et al. 2015)), we did not observe sex-related differences in this neuropil, suggesting there may be female-specific neurons with similar tuning. Sex differences in activity profiles in the GNG and LAL, which are innervated by descending neurons (Namiki et al. 2018; Cande et al. 2018), could mediate sex-specific auditory-driven locomotor behaviors (Jan Clemens, Deutsch, et al. 2018).

Auditory tuning to song features in mechanosensory neuropils

Our central brain-wide auditory activity map revealed that individual brain neuropils differ in their temporal responses to sine versus pulse stimuli [Figure 3B and 3F]. To further probe these differences, we more systematically evaluated the courtship song stimulus features [Figure 6A] that drive responses in the neuropils with strong auditory activity [Figure 2E]. We focused on a putative pathway (Frechter et al. 2018; Lai et al. 2012; Matsuo et al. 2016; Patella & Wilson 2018; Suver et al. 2019) starting in the AMMC [Figure 6B], and generated tuning curves; given the large size of our stimulus set, we were able to cover only a fraction of each neuropil (see Methods). These tuning curves could be clustered into 7 tuning types [Figures 6C and 6D], and this analysis revealed that auditory responses are still divided into three main categories: pulse-tuned (tuning type 1), sine-tuned (tuning type 5-7), or non-selective (tuning type 2-4) [compare with Figure 3]. Pulse-tuned ROIs responded best to pulse frequencies, pauses and durations present in natural courtship song [Figure 6C and 6Eii] - that is, pulses with 250 Hz carrier frequency, 16ms duration, and 20 ms pulse pause (Jan Clemens, Deutsch, et al. 2018). Sine-tuned ROIs were either low pass (strongest responses to 100 Hz or 150 Hz sines; tuning type 5-6) or had a broader distribution of preferred carrier frequencies (tuning type 7) [Figures 6C and 6Ei]. These sine-tuned ROIs preferred continuous stimuli, which explains their preference for pulses with short pauses (below 20 ms) and longer pulse durations [Figures 6C and 6E-ii]. Similar to pulse-tuned responses, frequency tuning of low pass sine-tuned responses overlapped the distribution of conspecific sine frequencies. Finally, non-selective auditory units were low pass (strongest responses at frequencies below 250 Hz), and similar to sine-tuned units, they preferred short pulse pauses (below 20 ms) [Figure 6C, and 6Ei-ii]. Most tuning types were sensitive to the broad range of intensities tested (0.5-5 mm/s) with proportional increases in response with stimulus magnitude [Figures S11A-B]. At longer timescales, most tuning types preferred longer pulse train or sine tone durations (4s or longer) [Figure S11A and S11C]. Although, during natural courtship, pulse or sine trains (stretches of each type of song) have a mode of ~360 ms and rarely last longer than 4s (Coen et al. 2014), preference for unnaturally long bouts have been recently reported (Jan Clemens, Deutsch, et al. 2018).

We next evaluated auditory tuning to song features by neuropil [Figure 6D]. The AMMC and SAD were dominated by narrowly tuned sine preferring responses (tuning type 6). The WED had a more diverse set of responses, although biased towards sine tuned responses (tuning types 5-7, which includes both narrowly and broadly tuned responses). Consistent with previous coarse tuning results [Figure 3B], at the GNG, AVLP and PVLP we observed a dramatic shift towards pulse tuned responses (type 1), although the AVLP and PVLP contained also sine-tuned responses (type 5) [Figure 6D]. Similar to the AMMC, the PLP and LH contained an over-representation of narrowly tuned sine preferring responses (types 5 and 6), accompanied by an increase in broadly tuned responses (types 3-4). These results validate our choice of a single pulse and sine stimulus [Figures 2-5] for probing the representation of song throughout the brain.

Although, pulse song is in essence an amplitude modulated sine tone (amplitude modulation at ~27Hz), responses to sine tones and pulses at various carrier frequencies correlated weakly across tuning types [Figure 6C] - that is, responses tuned for pulse stimuli (e.g., tuning type 1) prefer pulses at 250Hz, but do not prefer 250Hz sines. This suggests that the slow frequency of the amplitude envelope has nonlinear

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interactions with carrier frequencies. In order to evaluate the categorical boundary between sine and pulse responses (in other words, the interaction between stimulus envelope and carrier frequency), we used a new set of stimuli. We added a ~27 Hz envelope (pulse rate frequency) to pure tones of 150 or 250 Hz carrier frequency at different strengths (see Methods). As expected from the selectivity to pulse rate, 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 was carrier frequency dependent (i.e. greater increase for 250 Hz vs 150 Hz stimuli) [Figure 6F]. Similarly, responses tuned to sine tones were strongly attenuated by the presence of an envelope, with different sensitivities (tuning type 7 had a weak decrease at low envelope strength while tuning type 6 was strongly diminished) [Figure 6F].

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 with precise registration to unbiasedly screen for auditory responses beyond the periphery, and to compare auditory representations across brain regions, individuals and sexes [Figure 1]. We found that auditory activity is widespread, extending well beyond the canonical mechanosensory pathway, and present in brain regions and tracts known to process other sensory modalities (i.e. olfaction and vision), or to drive motor behaviors [Figure 2]. The representation of auditory stimuli diversified from the AMMC to later stages of the putative pathway, becoming more selective for particular aspects of courtship song (i.e. sine or pulse song, and their characteristic spectral and temporal features) [Figures 3 and 6]. These auditory representations were more stereotypic across trials and individuals in early stages of mechanosensory processing, becoming more variable at later stages [Figure 4]. Nonetheless, auditory representations are to a surprising extent similar across sexes [Figure 5]. By mapping tuning to song features in a subset of neuropils with strong auditory activity (AMMC, SAD, WED, GNG, VLP, PLP and LH), we found an overrepresentation of conspecific features of courtship song and elucidated the categorical boundary between sine and pulse responses [Figure 6]. The diversity of auditory representations, the variability across renditions and across individuals, and the tuning to song features at different processing stages have important implications for how the brain processes auditory information to extract salient features and to ultimately guide behavior.

C. elegans, larval *Drosophila,* and larval zebrafish are transparent, permitting non-invasive imaging of the whole brain through the cuticle or skin with single-photon imaging methods (Ahrens et al. 2013; Lemon et al. 2015; Nguyen et al. 2016). Here, we used two-photon imaging for effective imaging deep into the adult fly brain (~250 um along the anterior-to-posterior axis) - achieving cellular resolution necessitated imaging sub-volumes of the brain. We therefore developed a method to precisely stitch contiguously imaged volumes together and then to register them into an *in vivo* intersex template brain. The enhanced spatial resolution facilitated segmentation of subneuropil structures and even somas, and thereby comparison of ROIs across trials, individuals, and sexes (in contrast with (Mann et al. 2017; Aimon et al. 2019)). However, because we imaged activity sequentially (although relative to the same, repeatable sensory stimuli) versus simultaneously, it is possible that varying internal dynamics of the brain contributed to variability across trials and animals. New methods for fast imaging of large regions with two-photon microscopy by improving lateral and axial speed (Akemann et al. 2015; Piazza et al. 2018; Sofroniew et al. 2016), beam shaping (vTwins (Song et al. n.d.), Bessel focus scanning (Lu et al. 2017), and SCAPE (Bouchard et al. 2015)), are likely to be powerful for capturing more brain activity simultaneously, while not significantly compromising on spatial resolution. Nonetheless, our preprocessing and analysis

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pipeline for motion correction, ROI segmentation, registration, and profiling of temporal activity patterns, will facilitate analysis of any new data sets.

Our understanding of the *Drosophila* auditory circuit thus far has been built up from targeted studies of single brain regions and neural cell types composing them (Kamikouchi et al. 2006; Matsuo et al. 2016; Yorozu et al. 2009; Tootoonian et al. 2012; Jan Clemens, Girardin, et al. 2018; Vaughan et al. 2014; Zhou et al. 2015; Ishikawa et al. 2017; Matsuo et al. 2014; Yamada et al. 2018). Altogether, these studies have delineated an auditory pathway that starts in the Johnston's Organ (JO), and extends from the AMMC, to the WED, VLP, and LPC (Lai et al. 2012; Vaughan et al. 2014; Zhou et al. 2015; Matsuo et al. 2016; Jan Clemens, Deutsch, et al. 2018; Patella & Wilson 2018). However, the widespread auditory responses we describe here, spanning brain regions beyond the canonical pathway, suggests auditory processing is more distributed than previously thought. This is consistent with anatomical studies that find connections from the AMMC and WED to several other brain regions (e.g. GNG, SAD, PLP, SCL, and ATL) (Suver et al. 2019; Matsuo et al. 2016; Frechter et al. 2018). However, this is likely not unique to audition. So far, in adult *Drosophila*, only taste processing has been surveyed brain-wide (Harris et al. 2015). While that study did not map activity onto neuropils and tracts, nor made comparisons across individuals, it suggested that taste processing was distributed throughout the brain. Similarly in vertebrates, widespread responses to visual, nociceptive and vestibular stimuli have been observed throughout the brain (Naumann et al. 2016; Haesemeyer et al. 2018; Favre-Bulle et al. 2018; Migault et al. 2018)). Brain-wide mapping of neural activity during behavior has also uncovered that neural signals related to animal movements are widely distributed throughout the brain, including in sensory areas ((Musall et al. 2019; Stringer et al. 2019; Aimon et al. 2019; Scholz et al. 2018)). The methods we develop here will enable systematic mapping of responses to additional sensory modalities, all registered in the same atlas coordinates for comparisons. By combining such maps with information on connectivity between and within brain regions (e.g., using trans-synaptic labeling (Talay et al. 2017), or electron microscopy (Zheng et al. 2018)), the logic of multisensory coding in the *Drosophila* brain should emerge.

During the courtship ritual, female and male flies evaluate multiple sensory cues (olfactory, auditory, gustatory and visual) to inform their mating decisions and modulate their mating drive respectively (Yamamoto & Koganezawa 2013). Although multisensory integration has been described in higher order brain regions (e.g., pC1 neurons in the LPC, (Clowney et al. 2015; Kallman et al. 2015)), our results suggest that song representations could be integrated with olfactory and visual information at earlier stages. In the olfactory system, auditory responses occur at both early and later stages of olfactory processing (AL, PN, MB, and LH), suggesting that integration of olfactory and auditory stimuli could be a serial and hierarchical process, which is supported by a diversification of auditory responses from AL-to-MB or AL-to-LH. Auditory modulation could be relayed directly from early mechanosensory neuropils to the AL and LH (Matsuo et al. 2016; Jan Clemens, Girardin, et al. 2018; Tanaka et al. 2004; Frechter et al. 2018; Yu et al. 2013), or indirectly through modulatory input neurons (e.g., dopaminergic neurons PAM and PPL1 which innervate the SLP (Aso et al. 2014), a region that herein showed strong auditory activity). Representation of auditory stimuli in the MB and Kenyon cells (KCs), the locus of associative learning, is consistent with the idea that KCs represent multiple sensory modalities (olfactory and visual) (Vogt et al. 2014; Vogt et al. 2016). Indeed, auditory activity in the MB could underlie appetitive conditioning to auditory stimuli (Menda et al. 2011). Similarly, innate olfactory behaviors may be modulated by sounds via auditory-driven changes in the percept of olfactory categories in the LH (Frechter et al. 2018; Jefferis et al. 2007). Auditory modulation at early stages of olfactory processing in *Drosophila* may contrast with mammals, where it is thought to occur in cortex (Varga & Wesson 2013; Wesson & Wilson 2010). Auditory activity in visual (AOTU, PVLP and PLP) neuropils mostly occupies compartments with little overlap to optic glomeruli (Wu et al. 2016). This, however, does not rule out the possibility that local interneurons

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

spanning visual and auditory compartments could integrate both modalities, or that integration occurs in circuits downstream of optic glomeruli, yet to be described. Finally, we find auditory activity in brain regions involved in locomotion and navigation (central and lateral complex, superior and ventromedial neuropils). While some of this activity may represent startle responses to sound (Batchelor and Wilson 2019), collectively, activity in these regions is temporally diverse (Figure 3), suggesting that pre-motor circuits receive information about courtship song patterns, and could therefore underlie stimulus-specific locomotor responses (Jan Clemens, Deutsch, et al. 2018). In addition, the higher across-trial variability in these brain regions compared to early mechanosensory neuropils may suggest that sensory responses are modulated by the state of the animal. Future experiments that combine brain-wide mapping of sensory activity with simultaneous recordings of locomotor behavior will allow us to determine the relationship between sensory-driven activity and ongoing behavior.

Drosophila melanogaster songs are composed of pulses and sines, which differ in their spectral and temporal properties. These two distinct song modes elicit stimuli-specific motor behaviors (Jan Clemens, Deutsch, et al. 2018; Yoon et al. 2013; Zhou et al. 2015), but it is unclear how and where selectivity for the different song modes arises in the brain. Previous work has shown that the tuning of receptor neurons (JONs) and early mechanosensory neurons in the AMMC and WED, span the frequencies present in song (Patella & Wilson 2018; Ishikawa et al. 2017; Matsuo et al. 2014; Yorozu et al. 2009; Kamikouchi et al. 2009). Here, we find that at the AMMC and WED (along with the SAD, whose auditory representations are nearly identical to the AMMC) there is an overrepresentation of frequencies present in conspecific sine song [Figure 6D], but the WED also codes for higher frequencies (consistent with findings from (Patella & Wilson 2018)). We find sine-selective responses also in the AVLP, PVLP, PLP, and LH, suggesting further processing downstream of the WED. In addition, as we show in Figure 6, temporal features of the song envelope (for example, the inter-pulse interval or IPI) can be used to discriminate pulses from sines. At the periphery, many auditory neurons (JONs and AMMC) constitute low-pass filters for pulse interval (Jan Clemens, Girardin, et al. 2018; Tootoonian et al. 2012). However, from the AMMC to the VLP and LPC, selectivity for conspecific IPIs increases (Vaughan et al. 2014; Zhou et al. 2015; Jan Clemens, Girardin, et al. 2018). In fact, pC2 neurons innervating the LPC serve as feature detectors for all features of the pulse song (Jan Clemens, Deutsch, et al. 2018). Consistent with previous studies, here we find strong representation of pulse song in the VLP (both AVLP and PVLP) and to a lesser extent in the LPC (SCL and ICL). Interestingly, the GNG also shows strong pulse-selective activity. Auditory tuning of the VLP and GNG pulse-selective responses matches the distribution of pulse features present in conspecific pulse song. Although it is as of yet unclear how this selectivity arises (but see (Yamada et al. 2018)), high-pass interval filters at the level of the WED could be used to generate band-pass interval tuning at later stages of the pathway. This study therefore lays the foundation for exploring how song selectivity arises in the brain.

Effective sensory encoding requires representation of sensory stimuli to be reliable across presentations. Here, we find that auditory activity is more reliable close to the periphery (AMMC and WED), but becomes variable in downstream regions. Although, this increase in variability is a phenomena also observed in the visual system (Kara et al. 2000; Scholvinck et al. 2015), it might not be generalizable across sensory pathways (e.g., in the olfactory system trial-to-trial variability decreases from 1st to 2nd order neurons (Bhandawat et al. 2007)). Because we recorded brain-wide activity sequentially rather than simultaneously, we could not evaluate how much of this variability is explained by activity from other connected brain regions (Mann et al. 2017; Chen et al. 2018)) or possibly behavioral state. In addition, to allow for flexibility in sensory processing or behavior, some sensory circuits may be more or less stereotyped across individuals. For example, in the *Drosophila* olfactory system, PNs (2nd layer) are functionally stereotyped (Wilson et al. 2004), while KCs (3rd layer) and MBONs (4th layer) are more variable (Murthy et al. 2008; Hige et al. 2015). In this study, we systematically compared stereotypy across brain

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regions, finding that auditory activity is stereotyped across individuals only in early mechanosensory regions (AMMC, WED, SAD, and GNG) but became more variable at higher stages. The factors underlying these differences in stereotypy across brain regions remain to be determined.

Male and female flies exhibit distinct responses to courtship songs ((Jan Clemens, Deutsch, et al. 2018; Yoon et al. 2013), which could be due to sex-specific differences in sensory processing of sounds. While sexual dimorphisms in anatomy of the auditory system has been studied in the context of already defined dimorphic circuits (i.e. dsx+ or fru+ neurons (Rideout et al. 2010; Yu et al. 2010)), we know little about functional differences in auditory processing between the sexes. While some auditory neurons are present only in males (Zhou et al. 2015) or females (Kimura et al. 2015; Jan Clemens, Deutsch, et al. 2018)), most auditory responses have been characterized in only the female nervous system (with the exception of pC2 neuron characterization (Jan Clemens, Deutsch, et al. 2018)). As behavioral tuning to song features is similar in males and females, we might expect auditory responses to also be similar. Here, we carried out the first systematic brain-wide comparison of auditory processing across sexes. At neuropil resolution, we indeed find extensive similarity of activity between sexes throughout both early and later stages of the auditory circuit, even in neuropils innervated by known sexually dimorphic neurons (i.e. AVLP, ICL and SCL) [Figure 5]. The few neuropils with modest sex differences - GNG, SAD and LAL, although not significant - are of particular interest given that they are innervated by descending neurons that modulate locomotion (Namiki et al. 2018; Robie et al. 2017; Hsu & Bhandawat 2016). These sex differences could underlie stimulus and sex-specific motor behaviors. The robustness of sensory representation across sexes could be due to reciprocal dimorphisms (i.e. there could be female-specific neurons with similar tuning as male-specific neurons) or due to a greater contribution from non-dimorphic neurons to neuropil signal (i.e. neurons with similar tuning across sexes). Therefore, broad population imaging must ultimately be combined with cell-type specific approaches to evaluate dimorphisms at finer scale and at different stages of sensory processing.

Brain-wide imaging from neuropils, neurite tracts and somas allowed us to capture all sound-evoked activity. However, for neuropil signals it is challenging to determine how many neurons are sampled. On the one hand, neuropil activity could come from many neurons with different patterns of innervation tiling the entire neuropil. Conversely, activity could also come from a small number of neurons, each with broad innervation patterns. Although, a diverse set of temporal and tuning types per neuropil [Figure 3] suggests we have sampled many neurons per neuropil, this diversity could also arise from compartmentalized processing of fewer cell types (Yang et al. 2016; Meier & Borst 2019). One step towards further dissecting the circuits underlying this broad and diverse auditory activity is to restrict GCaMP expression to neuronal populations defined by spatially-restricted genetic enhancer lines available for *Drosophila* (Jenett et al. 2012; Kvon et al. 2014; Tirian & Dickson 2017). Such experiments should link broad functional maps with the cell-types constituting them. Here we provide the tools for unbiased screening and comprehensive mapping of sensory-driven activity throughout the brain. This is an important first step for making new discoveries about how the brain represents the myriad stimuli and stimuli combinations present in the world.

SUPPLEMENTAL INFORMATION:

Supplemental information includes 11 figures and 5 movies.

ACKNOWLEDGEMENTS:

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We thank Greg Jefferis, Sama Ahmed, Christa Baker, and Jan Clemens for comments on the manuscript. Jan Clemens for assistance on auditory stimuli delivery and linear modeling of calcium responses, and helpful discussions on auditory coding, Andrea Giovannucci for help with ROI segmentation of volumetric calcium signals and guidance with CaImAn toolbox usage, Gregory Jefferis and Torsten Rohlfing for help using the image registration toolbox CMTK and the neuroanatomy toolboxes NAT and NBLAST, and Talmo Pereira for assistance with the Coherent Point Drift algorithm. DAP was supported in part by an NSF Physics Frontier Center grant, and MM was supported by an NIH NINDS New Innovator award, NIH BRAIN Initiative R01s NS104899 and NS110060, and a Howard Hughes Medical Institute Faculty Scholar award.

AUTHOR CONTRIBUTIONS:

D.A.P. and M.M. designed the study. D.A.P. collected and analyzed data. ST 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.

DECLARATION OF INTERESTS:

The authors declare no competing interests.

REFERENCES:

Ahrens, M.B. et al., 2013. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nature methods*, 10(5), pp.413–420.

Aimon, S. et al., 2019. Fast near-whole-brain imaging in adult Drosophila during responses to stimuli and behavior. *PLoS biology*, 17(2), p.e2006732.

Akemann, W. et al., 2015. Fast spatial beam shaping by acousto-optic diffraction for 3D non-linear microscopy. *Optics express*, 23(22), pp.28191–28205.

Arthur, B.J. et al., 2013. Multi-channel acoustic recording and automated analysis of Drosophila courtship songs. *BMC biology*, 11, p.11.

Aso, Y. et al., 2014. The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife*, 3, p.e04577.

Azevedo, A.W. & Wilson, R.I., 2017. Active Mechanisms of Vibration Encoding and Frequency Filtering in Central Mechanosensory Neurons. *Neuron*, 96(2), pp.446–460.e9.

Bennet-Clark, H.C. & Ewing, A.W., 1967. Stimuli provided by Courtship of Male Drosophila melanogaster. *Nature*, 215, p.669.

Bhandawat, V. et al., 2007. Sensory processing in the Drosophila antennal lobe increases reliability and separability of ensemble odor representations. *Nature neuroscience*, 10(11), pp.1474–1482.

Bouchard, M.B. et al., 2015. Swept confocally-aligned planar excitation (SCAPE) microscopy for high speed volumetric imaging of behaving organisms. *Nature photonics*, 9(2), pp.113–119.

Cachero, S. et al., 2010. Sexual dimorphism in the fly brain. *Current biology: CB*, 20(18), pp.1589–1601.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

587 Cande, J. et al., 2018. Optogenetic dissection of descending behavioral control in Drosophila. *eLife*, 7.
588 Available at: http://dx.doi.org/10.7554/eLife.34275.

589 Chang, A.E.B., Vaughan, A.G. & Wilson, R.I., 2016. A Mechanosensory Circuit that Mixes Opponent
590 Channels to Produce Selectivity for Complex Stimulus Features. *Neuron*, 92(4), pp.888–901.

591 Chen, T.-W. et al., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*,
592 499(7458), pp.295–300.

593 Chen, X. et al., 2018. Brain-wide Organization of Neuronal Activity and Convergent Sensorimotor
594 Transformations in Larval Zebrafish. *Neuron*, 100(4), pp.876–890.e5.

595 Chiang, A.-S. et al., 2011. Three-dimensional reconstruction of brain-wide wiring networks in Drosophila
596 at single-cell resolution. *Current biology: CB*, 21(1), pp.1–11.

597 Clemens, J., Girardin, C.C., et al., 2018. Connecting Neural Codes with Behavior in the Auditory System of
598 Drosophila. *Neuron*, 97(2), p.475.

599 Clemens, J. et al., 2018. Discovery of a new song mode in Drosophila reveals hidden structure in the
600 sensory and neural drivers of behavior. *current biology*. Available at:
601 https://www.sciencedirect.com/science/article/pii/S0960982218307735?via%3Dihub.

602 Clemens, J., Deutsch, D., et al., 2018. Shared Song Object Detector Neurons in Drosophila Male and Female
603 Brains Drive Divergent, Sex-Specific Behaviors. *bioRxiv*, p.366765. Available at:
604 https://www.biorxiv.org/content/early/2018/07/10/366765.

605 Clowney, E.J. et al., 2015. Multimodal Chemosensory Circuits Controlling Male Courtship in Drosophila.
606 *Neuron*, 87(5), pp.1036–1049.

607 Coen, P. et al., 2014. Dynamic sensory cues shape song structure in Drosophila. *Nature*, 507(7491),
608 pp.233–237.

609 Coen, P. et al., 2016. Sensorimotor Transformations Underlying Variability in Song Intensity during
610 Drosophila Courtship. *Neuron*, 89(3), pp.629–644.

611 Eberl, D.F., Duyk, G.M. & Perrimon, N., 1997. A genetic screen for mutations that disrupt an auditory
612 response in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United*
613 *States of America*, 94(26), pp.14837–14842.

614 Ewing, A.W., 1978. The antenna of Drosophila as a “love song”receptor. *Physiological entomology*, 3(1),
615 pp.33–36.

616 Favre-Bulle, I.A. et al., 2018. Cellular-Resolution Imaging of Vestibular Processing across the Larval
617 Zebrafish Brain. *Current biology: CB*, 28(23), pp.3711–3722.e3.

618 Frechter, S. et al., 2018. Functional and anatomical specificity in a higher olfactory centre. *bioRxiv*.
619 Available at: https://www.biorxiv.org/content/early/2018/06/05/336982.abstract.

620 Giovannucci, A. et al., 2019. CaImAn an open source tool for scalable calcium imaging data analysis. *eLife*,
621 8. Available at: http://dx.doi.org/10.7554/elife.38173.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

622 Gong, Z. et al., 2004. Two interdependent TRPV channel subunits, inactive and Nanchung, mediate hearing
623 in Drosophila. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 24(41),
624 pp.9059–9066.

625 Göpfert, M.C. & Robert, D., 2001. Biomechanics. Turning the key on Drosophila audition. *Nature*,
626 411(6840), p.908.

627 Göpfert, M.C. & Robert, D., 2002. The mechanical basis of Drosophila audition. *The Journal of experimental*
628 *biology*, 205(Pt 9), pp.1199–1208.

629 Gouwens, N.W. & Wilson, R.I., 2009. Signal propagation in Drosophila central neurons. *The Journal of*
630 *neuroscience: the official journal of the Society for Neuroscience*, 29(19), pp.6239–6249.

631 Haesemeyer, M. et al., 2018. A Brain-wide Circuit Model of Heat-Evoked Swimming Behavior in Larval
632 Zebrafish. *Neuron*, 98(4), pp.817–831.e6.

633 Harris, D.T. et al., 2015. Representations of Taste Modality in the Drosophila Brain. *Neuron*, 86(6),
634 pp.1449–1460.

635 Hige, T. et al., 2015. Plasticity-driven individualization of olfactory coding in mushroom body output
636 neurons. *Nature*, 526(7572), pp.258–262.

637 Hsu, C.T. & Bhandawat, V., 2016. Organization of descending neurons in Drosophila melanogaster.
638 *Scientific reports*, 6, p.20259.

639 Ishikawa, Y. et al., 2017. Anatomic and Physiologic Heterogeneity of Subgroup-A Auditory Sensory
640 Neurons in Fruit Flies. *Frontiers in neural circuits*, 11, p.46.

641 Ito, K. et al., 2014. A systematic nomenclature for the insect brain. *Neuron*, 81(4), pp.755–765.

642 Jefferis, G.S.X.E. et al., 2007. Comprehensive maps of Drosophila higher olfactory centers: spatially
643 segregated fruit and pheromone representation. *Cell*, 128(6), pp.1187–1203.

644 Jenett, A. et al., 2012. A GAL4-driver line resource for Drosophila neurobiology. *Cell reports*, 2(4), pp.991–
645 1001.

646 Kallman, B.R., Kim, H. & Scott, K., 2015. Excitation and inhibition onto central courtship neurons biases
647 Drosophila mate choice. *eLife*, 4, p.e11188.

648 Kamikouchi, A. et al., 2009. The neural basis of Drosophila gravity-sensing and hearing. *Nature*, 458(7235),
649 pp.165–171.

650 Kamikouchi, A., Shimada, T. & Ito, K., 2006. Comprehensive classification of the auditory sensory
651 projections in the brain of the fruit fly Drosophila melanogaster. *The Journal of comparative*
652 *neurology*, 499(3), pp.317–356.

653 Kara, P., Reinagel, P. & Reid, R.C., 2000. Low response variability in simultaneously recorded retinal,
654 thalamic, and cortical neurons. *Neuron*, 27(3), pp.635–646.

655 Kimura, K.-I. et al., 2015. Drosophila ovipositor extension in mating behavior and egg deposition involves
656 distinct sets of brain interneurons. *PloS one*, 10(5), p.e0126445.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kvon, E.Z. et al., 2014. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature*, 512(7512), pp.91–95.

Lai, J.S.-Y. et al., 2012. Auditory circuit in the Drosophila brain. *Proceedings of the National Academy of Sciences of the United States of America*, 109(7), pp.2607–2612.

Lehnert, B.P. et al., 2013. Distinct roles of TRP channels in auditory transduction and amplification in Drosophila. *Neuron*, 77(1), pp.115–128.

Lemon, W.C. et al., 2015. Whole-central nervous system functional imaging in larval Drosophila. *Nature communications*, 6, p.7924.

Lu, R. et al., 2017. Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nature neuroscience*, 20(4), pp.620–628.

Mann, K., Gallen, C.L. & Clandinin, T.R., 2017. Whole-Brain Calcium Imaging Reveals an Intrinsic Functional Network in Drosophila. *Current biology: CB*, 27(15), pp.2389–2396.e4.

Manton, J.D. et al., 2014. Combining genome-scale Drosophila 3D neuroanatomical data by bridging template brains. *bioRxiv*. Available at: https://www.biorxiv.org/content/early/2014/06/19/006353.short.

Matsuo, E. et al., 2014. Identification of novel vibration- and deflection-sensitive neuronal subgroups in Johnston's organ of the fruit fly. *Frontiers in physiology*, 5, p.179.

Matsuo, E. et al., 2016. Organization of projection neurons and local neurons of the primary auditory center in the fruit fly Drosophila melanogaster. *The Journal of comparative neurology*, 524(6), pp.1099–1164.

Meier, M. & Borst, A., 2019. Extreme Compartmentalization in a Drosophila Amacrine Cell. *Current biology: CB*. Available at: http://dx.doi.org/10.1016/j.cub.2019.03.070.

Menda, G. et al., 2011. Classical conditioning through auditory stimuli in Drosophila: methods and models. *The Journal of experimental biology*, 214(17), pp.2864–2870.

Migault, G. et al., 2018. Whole-Brain Calcium Imaging during Physiological Vestibular Stimulation in Larval Zebrafish. *Current biology: CB*, 28(23), pp.3723–3735.e6.

Morley, E.L., Jonsson, T. & Robert, D., 2018. Auditory sensitivity, spatial dynamics, and amplitude of courtship song in Drosophila melanogaster. *The Journal of the Acoustical Society of America*, 144(2), p.734.

Murthy, M., Fiete, I. & Laurent, G., 2008. Testing odor response stereotypy in the Drosophila mushroom body. *Neuron*, 59(6), pp.1009–1023.

Musall, S. et al., 2019. Single-trial neural dynamics are dominated by richly varied movements. *bioRxiv*, p.308288. Available at: https://www.biorxiv.org/content/10.1101/308288v3.abstract [Accessed May 8, 2019].

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

691 Myronenko, A. & Song, X., 2010. Point set registration: coherent point drift. *IEEE transactions on pattern*
692 *analysis and machine intelligence*, 32(12), pp.2262–2275.

693 Namiki, S. et al., 2018. The functional organization of descending sensory-motor pathways in Drosophila.
694 *eLife*, 7. Available at: http://dx.doi.org/10.7554/eLife.34272.

695 Naumann, E.A. et al., 2016. From Whole-Brain Data to Functional Circuit Models: The Zebrafish
696 Optomotor Response. *Cell*, 167(4), pp.947–960.e20.

697 Nguyen, J.P. et al., 2016. Whole-brain calcium imaging with cellular resolution in freely
698 behavingCaenorhabditis elegans. *Proceedings of the National Academy of Sciences*, 113(8),
699 pp.E1074–E1081. Available at: http://dx.doi.org/10.1073/pnas.1507110112.

700 Otsuna, H. & Ito, K., 2006. Systematic analysis of the visual projection neurons of Drosophila
701 melanogaster. I. Lobula-specific pathways. *The Journal of comparative neurology*. Available at:
702 https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.21015.

703 Patella, P. & Wilson, R.I., 2018. Functional Maps of Mechanosensory Features in the Drosophila Brain.
704 *Current biology: CB*, 28(8), pp.1189–1203.e5.

705 Pézier, A. et al., 2014. Engrailed alters the specificity of synaptic connections of Drosophila auditory
706 neurons with the giant fiber. *The Journal of neuroscience: the official journal of the Society for*
707 *Neuroscience*, 34(35), pp.11691–11704.

708 Philipsborn, A.C. von et al., 2011. Neuronal Control of Drosophila Courtship Song. *Neuron*, 69(3), pp.509–
709 522. Available at: http://dx.doi.org/10.1016/j.neuron.2011.01.011.

710 Piazza, S. et al., 2018. Enhanced volumetric imaging in 2-photon microscopy via acoustic lens beam
711 shaping. *Journal of biophotonics*, 11(2). Available at: http://dx.doi.org/10.1002/jbio.201700050.

712 Pnevmatikakis, E.A. et al., 2016. Simultaneous Denoising, Deconvolution, and Demixing of Calcium
713 Imaging Data. *Neuron*, 89(2), pp.285–299.

714 Pnevmatikakis, E.A. & Giovannucci, A., 2017. NoRMCorre: An online algorithm for piecewise rigid motion
715 correction of calcium imaging data. *Journal of neuroscience methods*, 291, pp.83–94.

716 Riabinina, O. et al., 2011. Active process mediates species-specific tuning of Drosophila ears. *Current*
717 *biology: CB*, 21(8), pp.658–664.

718 Rideout, E.J. et al., 2010. Control of sexual differentiation and behavior by the doublesex gene in
719 Drosophila melanogaster. *Nature neuroscience*, 13(4), pp.458–466.

720 Robie, A.A. et al., 2017. Mapping the Neural Substrates of Behavior. *Cell*, 170(2), pp.393–406.e28.

721 von Schilcher, F., 1976. The role of auditory stimuli in the courtship of Drosophila melanogaster. *Animal*
722 *behaviour*, 24(1), pp.18–26.

723 Scholvinck, M.L. et al., 2015. Cortical State Determines Global Variability and Correlations in Visual Cortex.
724 *Journal of Neuroscience*, 35(1), pp.170–178. Available at: http://dx.doi.org/10.1523/jneurosci.4994-
725 13.2015.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

726 Scholz, M. et al., 2018. Predicting natural behavior from whole-brain neural dynamics. *Neuroscience*.

727 Schrödel, T. et al., 2013. Brain-wide 3D imaging of neuronal activity in Caenorhabditis elegans with
728 sculpted light. *Nature methods*, 10(10), pp.1013–1020.

729 Senthilan, P.R. et al., 2012. Drosophila auditory organ genes and genetic hearing defects. *Cell*, 150(5),
730 pp.1042–1054.

731 Shaner, N.C. et al., 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from
732 Discosoma sp. red fluorescent protein. *Nature biotechnology*, 22(12), pp.1567–1572.

733 Shorey, H.H., 1962. Nature of the Sound Produced by Drosophila melanogaster during Courtship. *Science*,
734 137(3531), pp.677–678.

735 Sofroniew, N.J. et al., 2016. A large field of view two-photon mesoscope with subcellular resolution for in
736 vivo imaging. *eLife*, 5. Available at: http://dx.doi.org/10.7554/eLife.14472.

737 Song, A. et al., Volumetric Two-photon Imaging of Neurons Using Stereoscopy (vTwINS). Available at:
738 http://dx.doi.org/10.1101/073742.

739 Stringer, C. et al., 2019. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*,
740 364(6437), p.255.

741 Suver, M.P. et al., 2019. Encoding of Wind Direction by Central Neurons in Drosophila. *Neuron*, 102(4),
742 pp.828–842.e7.

743 Talay, M. et al., 2017. Transsynaptic Mapping of Second-Order Taste Neurons in Flies by trans-Tango.
744 *Neuron*, 96(4), pp.783–795.e4.

745 Tanaka, N.K. et al., 2004. Integration of chemosensory pathways in the Drosophila second-order olfactory
746 centers. *Current biology: CB*, 14(6), pp.449–457.

747 Tirian, L. & Dickson, B., 2017. The VT GAL4, LexA, and split-GAL4 driver line collections for targeted
748 expression in the Drosophila nervous system. *bioRxiv*, p.198648. Available at:
749 https://www.biorxiv.org/content/early/2017/10/05/198648 [Accessed June 6, 2018].

750 Tootoonian, S. et al., 2012. Neural representations of courtship song in the Drosophila brain. *The Journal*
751 *of neuroscience: the official journal of the Society for Neuroscience*, 32(3), pp.787–798.

752 Varga, A.G. & Wesson, D.W., 2013. Distributed auditory sensory input within the mouse olfactory cortex.
753 *The European journal of neuroscience*, 37(4), pp.564–571.

754 Vaughan, A.G. et al., 2014. Neural pathways for the detection and discrimination of conspecific song in D.
755 melanogaster. *Current biology: CB*, 24(10), pp.1039–1049.

756 Vogt, K. et al., 2016. Direct neural pathways convey distinct visual information to Drosophila mushroom
757 bodies. *eLife*, 5. Available at: http://dx.doi.org/10.7554/eLife.14009.

758 Vogt, K. et al., 2014. Shared mushroom body circuits underlie visual and olfactory memories in Drosophila.
759 *eLife*, 3, p.e02395.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

760 Wesson, D.W. & Wilson, D.A., 2010. Smelling Sounds: Olfactory–Auditory Sensory Convergence in the
761 Olfactory Tubercle. *The Journal of neuroscience: the official journal of the Society for Neuroscience*,
762 30(8), pp.3013–3021.

763 Wilson, R.I., Turner, G.C. & Laurent, G., 2004. Transformation of olfactory representations in the
764 Drosophila antennal lobe. *Science*, 303(5656), pp.366–370.

765 Wu, M. et al., 2016. Visual projection neurons in the Drosophila lobula link feature detection to distinct
766 behavioral programs. *eLife*, 5. Available at: http://dx.doi.org/10.7554/eLife.21022.

767 Yamada, D. et al., 2018. GABAergic Local Interneurons Shape Female Fruit Fly Response to Mating Songs.
768 *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 38(18), pp.4329–
769 4347.

770 Yamamoto, D. & Koganezawa, M., 2013. Genes and circuits of courtship behaviour in Drosophila males.
771 *Nature reviews. Neuroscience*, 14(10), pp.681–692.

772 Yang, H.H. et al., 2016. Subcellular Imaging of Voltage and Calcium Signals Reveals Neural Processing In
773 Vivo. *Cell*, 166(1), pp.245–257.

774 Yoon, J. et al., 2013. Selectivity and plasticity in a sound-evoked male-male interaction in Drosophila. *PloS*
775 *one*, 8(9), p.e74289.

776 Yorozu, S. et al., 2009. Distinct sensory representations of wind and near-field sound in the Drosophila
777 brain. *Nature*, 458(7235), pp.201–205.

778 Yu, H.-H. et al., 2013. Clonal development and organization of the adult Drosophila central brain. *Current*
779 *biology: CB*, 23(8), pp.633–643.

780 Yu, J.Y. et al., 2010. Cellular Organization of the Neural Circuit that Drives Drosophila Courtship Behavior.
781 *Current biology: CB*, 20(18), pp.1602–1614.

782 Zheng, Z. et al., 2018. A Complete Electron Microscopy Volume of the Brain of Adult Drosophila
783 melanogaster. *Cell*, 174(3), pp.730–743.e22.

784 Zhou, C. et al., 2014. Central brain neurons expressing doublesex regulate female receptivity in
785 Drosophila. *Neuron*, 83(1), pp.149–163.

786 Zhou, C. et al., 2015. Central neural circuitry mediating courtship song perception in male Drosophila.
787 *eLife*, 4. Available at: http://dx.doi.org/10.7554/eLife.08477.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

KEY RESOURCES TABLE:

CONTACT FOR REAGENT AND RESOURCE SHARING:

Further information and requests for resources and reagents should be directed to the Lead Contact Mala Murthy (mmurthy@princeton.edu)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Fly Stocks

Flies used to generate the IVIA atlas 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 (Coen et al. 2014). *Iav*^{1} mutants had the following genotype: *iav*^{1}/Y; GMR57C10-LexA/+; 13xLexAop-GCaMP6s, 8xLexAop-mCD8tdTomato/+. For segmentation of projections of pC1, pC2l and pC2m Dsx+ neurons w; GMR57C10-LexA/20xUAS-GCaMP6s, 20xUAS-myr-tdTomato; 8xLexAop-mCD8tdTomato/dsx^{Gal4}, for segmentation of P1 neurons and pMP7 neurons: w; NP2631-Gal4, 20XUAS>STOP>CsChrimson.mVenus/GMR57C10-LexA; 8xLexAop-mCD8tdTomato/ fruFLP. We acquired 13xLexAop-GCaMP6s (44590) from the Bloomington stock center, iav^{1} from the Kyoto stock center, GMR57C10-LexA from Gerry Rubin, 20XUAS>STOP>CsChrimson.mVenus from Vivek Jayaraman, 8xLexAop-mCD8tdTomato from Yuh Nung Jan, NP2631-Gal4 from the Kyoto stock center, fru^{FLP} from Berry Dickson, dsx^{Gal4} from Stephen Goodwin, NM91 from the Andolfatto group at Columbia University, and 20xUAS-myr-tdTomato from Glenn Turner.

METHODS

Sound delivery system

We used a sound delivery system that consists of i) a sound card (M-Audio Delta 44, 16 bits, 44100 Hz, 2-channel), ii) a 2-channel amplifier (Crown D-75A), iii) a headphone speaker (KOSS, 16 Ohm impedance; sensitivity, 112 dB SPL/1 mW), and iv) a coupling tube (length: 20 cm, diameter: 1 mm) - the sound delivery system was calibrated and all sounds were first compensated for distortions introduced by the delivery system (see Clemens, Girardin et al. 2018). During experiments, the sound tube was positioned in front of the fly at a distance of 2 mm and an angle of ~17º from the midline (elevation of ~39º). In this configuration, we observed qualitatively similar responses from the AMMC in both hemispheres.

Acoustic stimuli

Acoustic stimuli presented to flies on the two-photon imaging setup were organized in two main protocols: a) coarse tuning protocol, and b) feature tuning protocol. For both protocols the intensity of auditory stimuli presented was within the dynamic range of JO neurons 0.1-6 mm/s (Effertz et al. 2011) and within the distribution of song intensities produced during natural courtship (Morley et al. 2018) (pulse and sine song intensity 5mm/s, white noise intensity 2 mm/s).

a) Coarse tuning protocol: i) 10 seconds of synthetic pulses trains (278 pulses with pulse duration: 16 ms, pulse pause: 20 ms which represents an inter-pulse interval of 36 ms, carrier frequency: 250 Hz,

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

amplitude: 5 mm/s), ii) 10 seconds of pure tone (carrier frequency: 150 Hz, amplitude: 5 mm/s), iii) 10 seconds of white noise (frequency range: 80-1000 Hz, amplitude: 2 mm/s) iv) 10 seconds of natural song (amplitude: 1.8 mm/s), v) 10 seconds of synthetic P_{fast}-like pulse trains (278 pulses with pulse duration: 16 ms, pulse pause: 20 ms, carrier frequency: 400 Hz, amplitude: 5 mm/s). Each stimulus was preceded and followed by 10 seconds of silence, organized in blocks of three stimuli (i, ii, iii) or four stimuli (i, ii, iii, iv, or i, ii, iii, v), and blocks were repeated until reaching 6 trials per stimuli (total duration of ~9-10 mins). We used long stimuli duration (10 seconds) to boost the signal-to-noise ratio of stimuli-evoked responses in the brain.

b) Feature tuning protocol: i) pure tones at different carrier frequency (carrier frequency: 100, 150, 200, 250, 300, 500, and 800 Hz, amplitude: 5 mm/s, tone duration: 4s), different tone duration (carrier frequency: 150 Hz, amplitude: 5 mm/s, tone duration: 2, 4, and 8 s), and different intensity (carrier frequency: 150 Hz, amplitude: 0.5, 1, 2, and 5 mm/s, tone duration: 4s), ii) synthetic pulse trains at different carrier frequency (train of pulses with pulse duration: 16 ms, pulse duration: 20 ms, carrier frequency: 100, 150, 200, 250, 300, 400, 500, and 800 Hz, amplitude: 5 mm/s, train duration: 4s), pulse pause (train of pulses with pulse duration: 16 ms, pulse pause: 0, 20, 40, 60, and 80 ms, carrier frequency: 250 Hz, amplitude: 5 mm/s, train duration: 4s), and pulse duration (train of pulses with pulse duration: 4, 8, 16, 32, and 64 ms, pulse pause: 20 ms, carrier frequency: 250 Hz, amplitude: 5 mm/s, train duration: 4s), different train duration (train of pulses with pulse duration: 16 ms, pulse pause: 20 ms, carrier frequency: 250 Hz, amplitude: 5 mm/s, train duration: 2, 4, and 8s), and different intensity (train of pulses with pulse duration: 16 ms, pulse pause: 20 ms, carrier frequency: 250 Hz, amplitude: 0.5, 1, 2, and 5 mm/s, train duration: 4s), iii) natural song (stimulus iv) from Coarse tuning protocol), iv) white noise (frequency range: 80-1000 Hz, amplitude: 2 mm/s, duration: 4s), and v) pulse-to-sine morphing stimuli (pure tone with carrier frequency: 150 or 250 Hz, amplitude: 5 mm/s, duration: 4, envelope modulation at 27.8 or 30.3 Hz with different strengths (0, 0.25, 0.5, 1)). Each stimulus was preceded and followed by 4s of silence. We presented pulse (278 pulses with pulse duration: 16 ms, pulse pause: 20 ms, carrier frequency: 250 Hz, amplitude: 5 mm/s, train duration: 4s), sine (carrier frequency: 150 Hz, amplitude: 5 mm/s, tone duration: 4s) and white noise 6 times (used for detecting auditory modulated ROIs, see ROI stimuli modulation) while the other stimuli were presented 3 times, the order of each stimulus was randomized.

Fly preparation and functional Imaging

Virgin female or male flies (3-7 days old) were mounted and dissected as described previously (Murthy & Turner 2013; Tootoonian et al. 2012), with minor differences. 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 dissecting chamber floor. Following dissection of the head cuticle, we removed the air sacks and tracheas using sharp forceps (Dumont #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. We only imaged flies with intact antennae and aristae, which we tested by gently blowing on the flies.

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),

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a Chameleon Ultra II Ti:Sapphire laser (Coherent) and a water immersion objective (Olympus XLPlan 25X, 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 20mW, power above this level induced global increases in fluorescence. 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). For wild type flies, we only recorded from animals that first exhibited auditory-evoked responses in the AMMC of both hemispheres. Recordings typically lasted for 2-3 hours. Flies whose global baseline fluorescence saturated and became homogeneous were discarded.

Image preprocessing

For all experiments, we recorded either a) 13 to 16 volumetric time series consecutively along the Z axis (in a posterior-to-anterior direction) with 1.4x1.2x2 μm^3 voxel size at 1 Hz [Figures 1-5], or b) 4 to 6 volumetric time series from selected brain regions (not necessarily contiguous) with 1x1x2 μm^3, 1x1x3 μm^3 or 1.2x1.2x5 μm^3 voxel size at 2 Hz [Figure 6]. To generate the private whole brain atlas (per fly), we imaged 3-4 Z-stack volumes (in both channels) with 0.75x0.75x1 μm^3 voxel size to span the central brain in the Z dimension.

For each volumetric time series (XYZ and Time) we performed rigid motion correction on the X, Y, and Z as on the tdTomato signal using the NoRMCorre algorithm (Pnevmatikakis & Giovannucci 2017), and this transformation was then applied to the GCaMP6s signal. Images with anisotropic XY voxel size were spatially resampled to have isotropic XY voxel size (bilinear interpolation on X and Y axes), and temporally resampled to correct for different slice timing across planes of the same volume. This was also done to align timestamps of volumes relative to the start of the first stimulus (linear interpolation). Data was then up-sampled (linear interpolation) to twice the original sampling rate (2 Hz and 4 Hz, for volumes recorded at 1 Hz and 2 Hz respectively) prior to analysis.

Volumetric time series consecutively recorded along the Z axis (a) were stitched along the Z axis using NoRMCorre (similarly, we used the average tdTomato signal as the reference image and apply the transformation to the GCAMP6s signal), ending up with time series of ~307x307x252 μm^3 volumes. Motion corrected and resample volumes were then mirrored to the right hemisphere (for images recorded from the left hemisphere), flipped on the Z axis to have a standard anterior to posterior direction (from 1st to last plane imaged), and ROI segmented (see ROI segmentation below). Average tdTomato images were saved as individual NRRD (http://teem.sourceforge.net/nrrd) files (using Matlab, nrrdWriter function).

For the private whole brain atlases (per fly), z-stacks (~384x384x280 μm^3 each) were stitched on the X and Y axes using a pairwise stitching algorithm in FIJI (Preibisch et al. 2009; Schindelin et al. 2012). These whole brain images were smoothed on the X and Y axes (Gaussian kernel size of [3, 3] voxels and a standard deviation of gaussian kernel of [2, 2]), mirrored on the X axis, and flipped on the Z axis as for volumetric time series. Images were split into tdTomato and GCaMP6s channels and saved separately as NRRD files (using MatLab, nrrdWriter function).

ROI segmentation

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We used the constrained non-negative matrix factorization (CNMF) algorithm (Pnevmatikakis et al. 2016) for ROI segmentation. The algorithm was implemented and generalized to 3D-data in the MATLAB version of the CaImAn toolbox (CaImAn toolbox (Giovannucci et al. 2019)). We used the greedy initialization (Gaussian kernel size of [9, 9, 5] voxels and a standard deviation of gaussian kernel of [4, 4, 2], and 10th percentile baseline subtraction) with 600 components per sub-stack of 11 planes and walk through the entirety of the Z axis with a sub-stack overlap of 3 planes. After initialization, 1) spatial components were updated using the 'dilate' search method (the 'ellipse' method imposes a prior on the shape which is not optimal for detecting ROIs in neuropil), and default settings were used for updating temporal components, 2) spatially overlapping components were iteratively merged if the Pearson 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 sub-stacks per fly (similar to run_CNMF_patches.m), and to calculate main temporal variables per ROI (background fluorescence, Fo, delta-fluorescence DF, and DF/Fo, similar to signalExtraction.m). See https://github.com/dpacheco0921/CaImProPi.

The DF signal was detrended using a running percentile filtering (we used the 10th percentile, with a window size of 50 seconds, and a shift of 5 seconds). Fo is the modelled background over time, plus the percentile filtered out from DF. These two variables were then used to generate the DF/Fo over time.

Identification of stimulus-modulated ROIs

We linearly modelled each ROI signal as a convolution of the stimuli history and a set of three filters (one per stimulus) [Figure S1C]. Filters were estimated using ridge regression (Park & Pillow 2011), with a filter duration of 10 or 6 seconds for i) the coarse tuning protocol or ii) the feature tuning protocol, respectively. For filter estimation, we partitioned each ROI raw signal (F(t)) into all possible combinations of training (80%) and testing (20%) datasets (15 different combinations). Training and testing datasets included all presentations of pulse (with pulse duration: 16 ms, pulse pause: 20 ms, carrier frequency: 250 Hz, amplitude: 5 mm/s), sine (carrier frequency: 150 Hz, amplitude: 5 mm/s) and white noise (for both the coarse tuning and feature tuning protocol). Estimated filters (q(T)) were then convolved with the stimulus history (f(t)) to generate a set of 15 predicted signals (g(t)) for each ROI (one for each combination of training and test datasets). Prediction goodness was measured as the Pearson correlation coefficient between raw and predicted signals ρ(F, g) (using the testing dataset for each 80%/20% combination). Statistical significance of correlation coefficients (ρ(F, g)) was determined by bootstrapping. Each ROI raw signal was randomly shuffled in chunks of 10 seconds (sF(t)), and a distribution of 10,000 correlation coefficients (ρ(sF, g))) between each independent shuffle and the predicted signal was generated (using the testing dataset). p-values for ρ(F, g) significance were calculated as the fraction of ρ(sF, g) with values greater than the 30th percentile of ρ(F, g). To correct for multiple comparisons of all ROIs within a fly, p-values were corrected using the Benjamini-Hochberg (B-H) FDR correction algorithm (FDR = 0.01). This correction assumes independence or positive dependency between tests.

Construction of average *in vivo* intersex atlas (IVIA)

We built an average *in vivo* intersex brain atlas using a pool of 5 male and 8 female brains (expressing membranal tdTomato signal pan-neuronally (mCD8-tdTomato)) [Figure S2A]. 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 (Jefferis et al. 2007; Rohlfing & Maurer 2003). We used linear, followed by non-linear registration. We then generated a new average intensity and average deformation seed brain using an active deformation

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

model (implemented in CMTK as avg_adm function). We iterated this process, obtaining at the end of the 5th iteration the *in vivo* intersex atlas (IVIA).

Registration of functional data to IVIA

In vivo functional data from each fly was registered to the IVIA using the structural channel (mCD8-tdTomato signal) in a two-step fashion: i) volumes to private whole brain registration and ii) private whole brain to IVIA atlas registration [Figure 2SB and 2SD]. A 'volume' is the functionally imaged half of a hemibrain.

i) Half-hemibrain to private whole brain registration: for each fly, we registered (linear followed by nonlinear registration) the half-hemibrain tdTomato image to its own whole brain tdTomato image (Z-stack volume, see Image preprocessing). Given that the images come from the same individuals, we expected minimal nonlinear deformations. Therefore, we limited the linear transformation to only include rotation, translation and anisotropic scaling (metric: normalized mutual information). For the nonlinear transformation (metric: normalized mutual information), we used a small grid spacing (20 um, with 2 refinements), Jacobian constraint weight of 0.01, and smoothness constraint weight of 0.1.

ii) 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. In this case the 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 um, with 5 refinements), Jacobian constraint weight of 0.001, and smoothness constraint weight of 1. We noticed that in many cases the linear transformation was not great, and it seemed that this was partly due to the optic lobes. The optic lobes in our preparations had a high signal intensity (compare to the central brain regions), and motion of the optic lobes can be independent of the brain (our own observations). We therefore decided to exclude the voxels covering the optic lobes for the calculation of the linear transformation, as this provided a better initialization for the nonlinear registration calculated using all voxels. This modification improved the registration success, and the quality of registration for central brain structures (visually inspected).

These i) half-hemibrain to whole brain and ii) whole brain to IVIA atlas transformations per fly were concatenated to transform spatial components of segmented ROIs from the native half-hemibrain coordinates to *in vivo* atlas coordinates.

Construction of IVIA-IBNWB bridging registrations

From all the fix-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), limiting the accuracy of intensity-based registration algorithms. Therefore, we decided to use the rigid and non-rigid point set registration algorithm Coherent Point Drift (CPD) (Myronenko & Song 2010). We chose the surface of the antennal lobe (AL), antennal mechanosensory and motor center commissure (AMMCC), anterior optic tract (AOT), great commissure (GC), wedge commissure (WEDC), mushroom body (MB, includes mushroom body lobes, peduncle and calyx), posterior optic commissure (POC), protocerebral bridge (PB), lateral antennal lobe tract (lALT), posterior cerebro-cervical fascicle (pCCF), superior saddle commissure (sSADC), and the whole central brain as source of points to compare between

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

atlases [Figure 2SE]. These regions were distributed across the anterior-posterior and dorso-ventral axes. We converted segmented binary images from each of these brain regions to triangular surface meshes via isocontour thresholding at isovalues of 0.9, 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) [Figure 2SF].

Using the IVIA-to-IBNWB transformation, we mapped anatomical segmentation of neuropils and neurite bundles (Ito et al. 2014) to the IVIA coordinates. In this segmentation, the AMMC would also include JON projections. In addition, we merged the mushroom body accessory calyx to the mushroom body calyx, due to its small size relative to average ROI volume.

Registration accuracy

i) IVIA within-atlas accuracy: we segmented stalks from pC1, pC2l, pC2m, and pM7 neurons, and two more tracts of Dsx+ neurons (n = 20, 10, 12, 7, 9, and 9 hemibrains for pC1, pC2l, pC2m, pM7, Dsx+ tract-a, and Dsx+ tract-b respectively), which were consistently identified across individuals in flies expressing GCaMP6s via Dsx-GAL4 or CsChrimson via NP2631-GAL4 and FruFLP combo. These stalks and tracts covered different dorso-ventral regions from the posterior half of the brain [Figure 2SC-i]. 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 [Figure 2SC-ii]. This estimate is the sum of biological variability and algorithm-associated error. Although we obtained cellular accuracy (~4.3 µm), this is slightly lower compared to values reported for fixed brain atlases (2-3 um) (Chiang et al., 2011; Jefferis et al., 2007; Peng et al., 2011; Yu et al., 2010), 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.

ii) IVIA and IBNWB between-atlas accuracy: we used segmented pC1 stalks from IVIA (described above in i)) and compared them to pC1 neurons originally collected in the FCWB atlas (imaged via Fru-GAL4 expression of GFP (Chiang et al. 2011)), which were mapped to IBNWB (error of the transformation between FCWB and IBNWB is assumed to be negligible). Similar to i), we calculated the mean pC1 trace for each atlas (µ-pC1-IVIA and µ-pC1-IBNWB) and the within-atlas accuracy (σ-pC1-IVIA and σ-pC1-IBNWB). We measured the IVIA-to-IBNWB registration accuracy as the difference between the σ-pC1-IBNWB and the standard deviation of pC1-IBNWB traces mapped to IVIA relative to µ-pC1-IVIA [Figure 2SG-i]. On the other hand, we measured the IBNWB-to-IVIA registration accuracy as the difference between the σ-pC1-IVIA and the standard deviation of pC1-IVIA traces mapped to IBNWB relative to µ-pC1-IBNWB [Figure 2SG-ii].

Deformation-Based morphometry

We measured the relative volume change between a sample brain and a reference brain (IVIA) using the Jacobian determinant. The Jacobian determinants were normalized by the mean of all voxels inside the brain (both central brain and optic lobes) per individual, and log transformed. We used a voxel-wise t-test between the groups of male (n = 22) and female (n = 23) brains (and their X-mirrored images) as described in (Cachero et al. 2010). We determined t-statistic significance by bootstrapping. The 5% and 95%

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

quantiles of the t-statistic distribution ($t_{5\%}$ = -5.19, $t_{95\%}$ = 5.17) were used as thresholds for significance of t-statistics per voxel of original female to male brain comparison.

Anatomical cataloging of central brain activity

For each fly, we combined all stimuli modulated ROIs and binarized these volumes, obtaining binary auditory maps per fly. Auditory maps were then transformed to IVIA atlas coordinates. To account for the known fly-to-atlas error, we dilated these transformed maps ([1.25, 1.25, 2] μm^3 in XYZ). We then summed these maps across flies, generating a density volume [Figure 2C, Figure S1A, Figure S4B, Figure S4F, Figure 5B, Figure 6SA-B, and Figure S7D]. 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: i) ROIs must have at least a 15% volume overlap with its assigned neuropil or neurite bundle, ii) 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) [Figure 2E and 3SA-B].

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

For all auditory modulated ROIs from intact and *iav*^{1} flies (19,389 ROIs from 21 male (4 male flies carried *iav*^{1}) and 17 female flies) presented with the coarse tuning protocol (this includes pulse, sine, white noise, natural song, and P_{fast}-like pulse, see Acoustic stimuli), we calculated the median response to each stimulus across trials, including 10 seconds before stimulus onset to 10 seconds 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's signals (pooled from all intact and *iav*^{1} 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), we then chose a distance (distance of 136) that would split both inhibitory and excitatory responses into a minimum number of distinct responses types (18 response types). Although response types could be further split, sub-clusters within each response type (obtained by reducing the distance to half its value) are still very similar to the parent response type [Figure 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 [Figure 3C-E]. Adaptation was measured as the percentage of absolute baseline subtracted signal 2 seconds before stimulus offset with respect to signal 2 seconds after stimulus onset ('adaptation index'). Therefore, adaptation indices below 100% indicates some degree of adaptation (the lower the adaptation index, the higher the adaptation), while adaptation indices above 100% indicates 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. Decay time constant of auditory responses was determined by fitting an exponential to the 20 values after stimulus offset (10 seconds), 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. Response magnitude was defined as the absolute of the percentile (80th or 20th) with the greater absolute value of baseline subtracted signal from stimulus onset

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to 2 seconds after stimulus offset (baseline was defined as the mean signal from -4 to -0.5 seconds relative to stimulus onset, dividing the response magnitude by the standard deviation during baseline gave us the SNR). Response magnitudes were normalized by the maximum value across stimuli and multiplied by 100 to have units in 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 [Figure 3F-i-iii], otherwise it was considered non-selective [Figure 3F-iv].

Clustering of tuning curves

We generated tuning curves to the 'feature tuning' stimulus set [Figure 6]. For these experiments, we sampled a fraction of each neuropil selected (24.4, 46.2, 10.5, 83, 39.8, 99.3, 63.6, and 91.8 % for the AMMC, SAD, GNG, WED, AVLP, PVLP, PLP and LH, respectively). For all auditory modulated ROIs from intact flies (10,970 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 seconds before stimulus onset to 4 seconds after stimulus offset. We measured the magnitude and the signal-to-noise ratio (SNR) of responses to auditory stimulus for each ROI. Baseline was defined as -2 to -0.25 seconds from stimuli onset, we then calculated the mean μ_b and standard deviation σ_b of the signal during this time. Response magnitude was defined as the percentile (80^{th} or 20^{th}) with the greater absolute value of μ_b subtracted signal from stimulus onset to 2 seconds after stimulus offset. The SNR was defined as the absolute response magnitude divided by σ_b. For each ROI, we generated a tuning curve by concatenating and z-scored absolute response magnitude across presented sine song at different carrier frequencies, and pulse song at different carrier frequencies, pulse pause and pulse durations (for ROIs with a SNR greater or equal than 4). We then hierarchically clustered ROI's 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 divide into three main classes, tuning to sine, tuning to pulses, and non-selective. In order to split the three main classes into the minimum number of distinct sub-clusters we decided on a distance of 60 which split all tuning curves into 7 clusters [Figure 6C].

Measuring activity diversity across brain regions

We collected ROIs belonging to the same neuropil across flies (separately for intact or *iav*^{1} 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, obtaining 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 were the values shown in Figure 3B and S5G. We then calculated the entropy for each response type distribution as a measure of activity diversity, 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 deviation of response integral ('variability index'). We measure the integral of baseline subtracted auditory responses from stimulus onset until 5 seconds after stimulus offset (baseline was defined as the mean signal from -4 to -0.5 seconds relative to stimuli onset) for each stimulus across trials. For each stimulus, response integrals were then normalized to the integral of the median signal across trials. The variability index is then defined as the standard deviation of the $\log_{10}$ transform of the median-normalized response integrals across stimuli (for all stimuli with response SNR greater than or equal to 4; see Response type kinetics and tuning for SNR calculation). We computed the variability index for each ROI and then sorted these indices by

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

neuropil [Figure 4B]. We did an all-to-all comparison of trial variability across neuropils using two-sample t-tests (we only included neuropils with at least 4 flies). We determined significance of each comparison by bootstrapping. To correct for multiple comparisons, p-values were corrected using the Benjamini-Hochberg (B-H) FDR correction algorithm (FDR = 0.01).

Measuring across-individual variability

We calculated response type distributions per neuropil (see Figure 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 use 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, restricting the distribution of similarity indices to the 0-1 range. We computed the similarity index between all possible combinations of fly pairs within a neuropil, generating a distribution of similarity indices for each neuropil. We compared these distributions across neuropils (for those with data from at least 4 flies), and significance of differences was determined using two-sample t-tests. We determined significance of each comparison by bootstrapping. To correct for multiple comparisons, p-values were corrected using the Benjamini-Hochberg (B-H) FDR correction algorithm (FDR = 0.01). This correction assumes independence of positive dependency between tests.

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. We measured the effect size using the Cohen's d [Figure 5C]. Within a neuropil, we determined the significance of female to male differences for each response type using two-sample t-tests [Figure 5C-D]. To correct for multiple comparisons within a neuropil, p-values were corrected using the Benjamini-Hochberg (B-H) FDR correction algorithm (FDR = 0.01).

References

Cachero, S. et al., 2010. Sexual dimorphism in the fly brain. *Current biology: CB*, 20(18), pp.1589–1601.

Chiang, A.-S. et al., 2011. Three-dimensional reconstruction of brain-wide wiring networks in Drosophila at single-cell resolution. *Current biology: CB*, 21(1), pp.1–11.

Coen, P. et al., 2014. Dynamic sensory cues shape song structure in Drosophila. *Nature*, 507(7491), pp.233–237.

Effertz, T., Wiek, R. & Göpfert, M.C., 2011. NompC TRP channel is essential for Drosophila sound receptor function. *Current biology: CB*, 21(7), pp.592–597.

Giovannucci, A. et al., 2019. CaImAn an open source tool for scalable calcium imaging data analysis. *eLife*, 8. Available at: http://dx.doi.org/10.7554/eLife.38173.

Ito, K. et al., 2014. A systematic nomenclature for the insect brain. *Neuron*, 81(4), pp.755–765.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Jefferis, G.S.X.E. et al., 2007. Comprehensive maps of Drosophila higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell*, 128(6), pp.1187–1203.

Morley, E.L., Jonsson, T. & Robert, D., 2018. Auditory sensitivity, spatial dynamics, and amplitude of courtship song in Drosophila melanogaster. *The Journal of the Acoustical Society of America*, 144(2), p.734.

Murthy, M. & Turner, G., 2013. Whole-cell in vivo patch-clamp recordings in the Drosophila brain. *Cold Spring Harbor protocols*, 2013(2), pp.140–148.

Myronenko, A. & Song, X., 2010. Point set registration: coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12), pp.2262–2275.

Park, M. & Pillow, J.W., 2011. Receptive field inference with localized priors. *PLoS computational biology*, 7(10), p.e1002219.

Pnevmatikakis, E.A. et al., 2016. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron*, 89(2), pp.285–299.

Pnevmatikakis, E.A. & Giovannucci, A., 2017. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of neuroscience methods*, 291, pp.83–94.

Preibisch, S., Saalfeld, S. & Tomancak, P., 2009. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* , 25(11), pp.1463–1465.

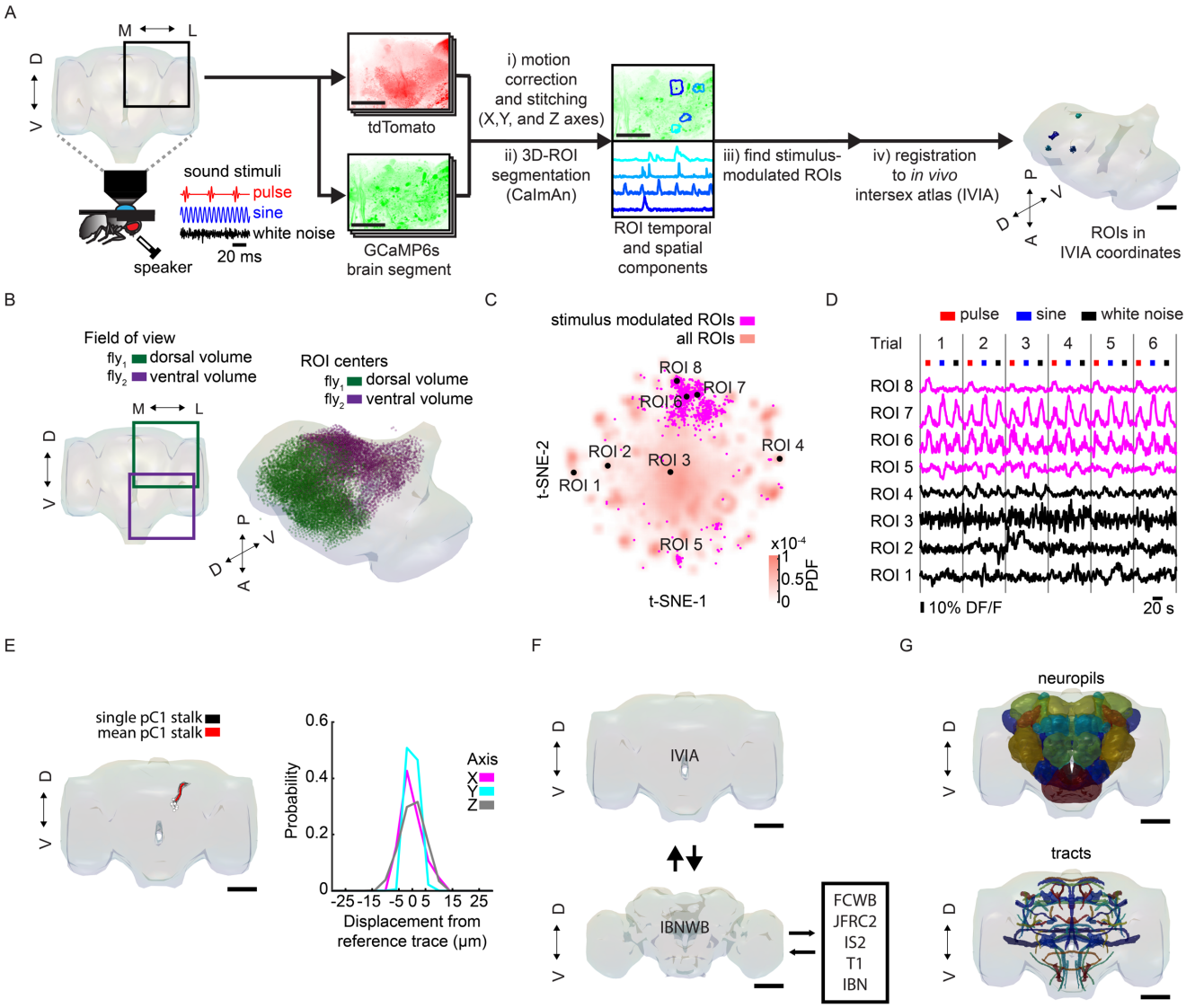
Rohlfing, T. & Maurer, C.R., Jr, 2003. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society*, 7(1), pp.16–25.

Schindelin, J. et al., 2012. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7), pp.676–682.

Tootoonian, S. et al., 2012. Neural representations of courtship song in the Drosophila brain. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 32(3), pp.787–798.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. Functional screening of auditory activity throughout the central brain

(A) Schematic of experimental setup and data-processing pipeline. Fly head is tilted 90 degrees (posterior brain is exposed for imaging), and presented with auditory stimuli (pulse, sine, and white noise; see Methods). Raw imaging data is processed as follows: tdTomato signal is used to motion-correct volumetric time-series and stitch serially imaged overlapping brain segments in XYZ axis (i). Next, the GCAMP6s signal is 3D-ROI segmented to obtain spatial and temporal components per segmented ROI via CaImAn (ii). ROIs that are modulated by auditory stimuli are detected (iii). Finally, spatial components of segmented ROIs are mapped to the IVIA space (iv).

(B) Example of segmented ROIs combined from two flies with overlapping field of views (dorsal and ventral volumes). ROI centers span the entire anterior-posterior and dorsal-ventral axis (Shown here: 11,225 ROIs from two flies).

(C) 2D t-SNE embedding of activity from all ROIs in panel (B). ROIs modulated by auditory stimuli (1,118 out of 11,225 ROIs) are in magenta.

(D) DF/F from ROIs indicated in (C). Magenta traces correspond to stimulus-modulated ROIs, while black traces are non-modulated ROIs.

(E) IVIA registration accuracy: left) 3D-rendering of traced stalks of Dsx+ neurons pC1 (black traces, n = 20 hemibrains from Dsx-GAL4/UAS-GCaMP6s flies (see Methods)) and mean pC1 stalk (red trace) across the 20 hemibrains, right) Per-axis jitter (X, Y, and Z) between traced pC1 stalk across flies relative to mean pC1 stalk.

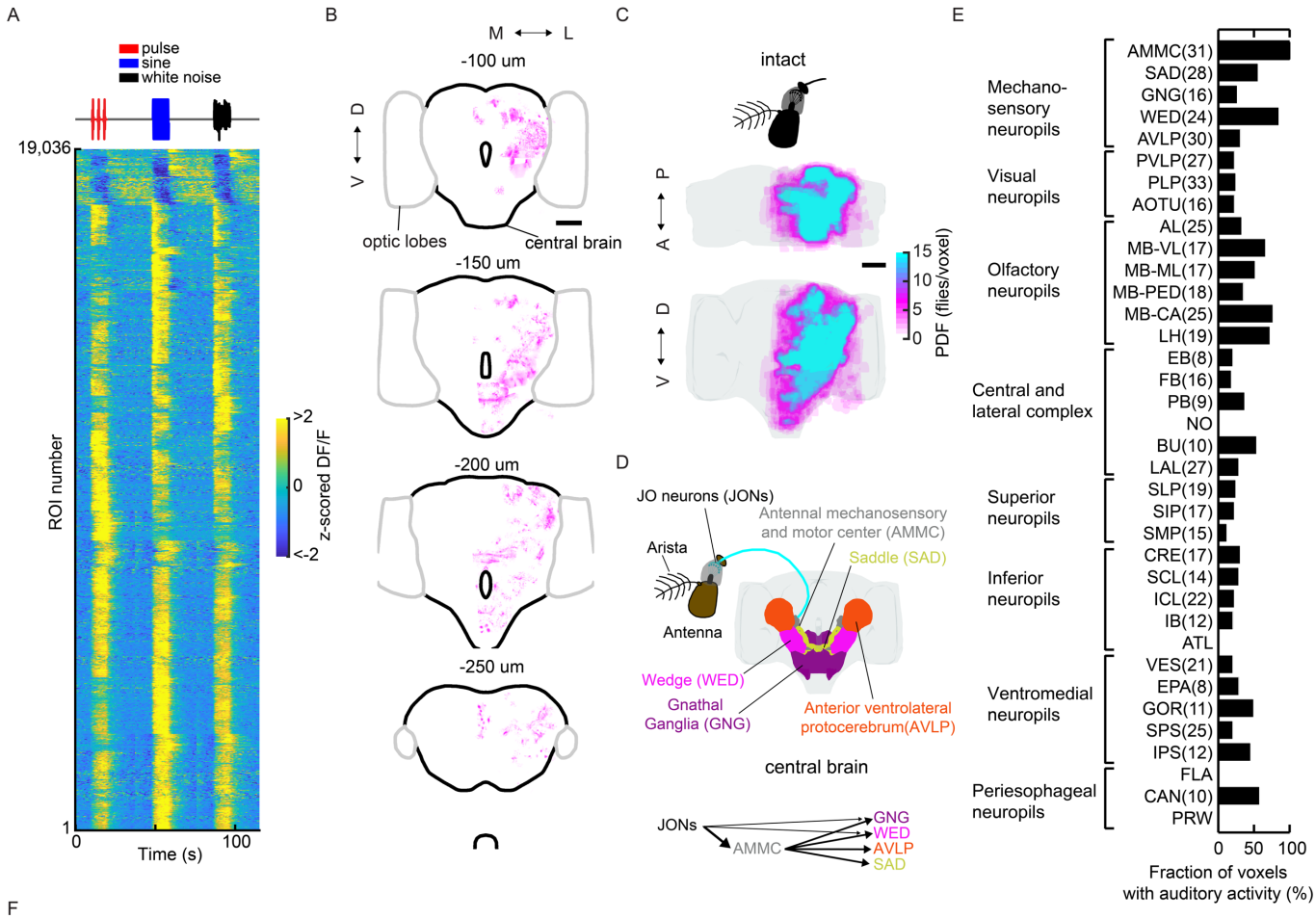
(F) Schematic of bridging registration between IVIA and IBNWB atlas (Ito et al. 2014). This bidirectional interface provides access to a network of brain atlases (IBN, FCWB, JFRC2, IS2, and T1) associated with different *Drosophila* neuroanatomy resources (Manton et al. 2014).

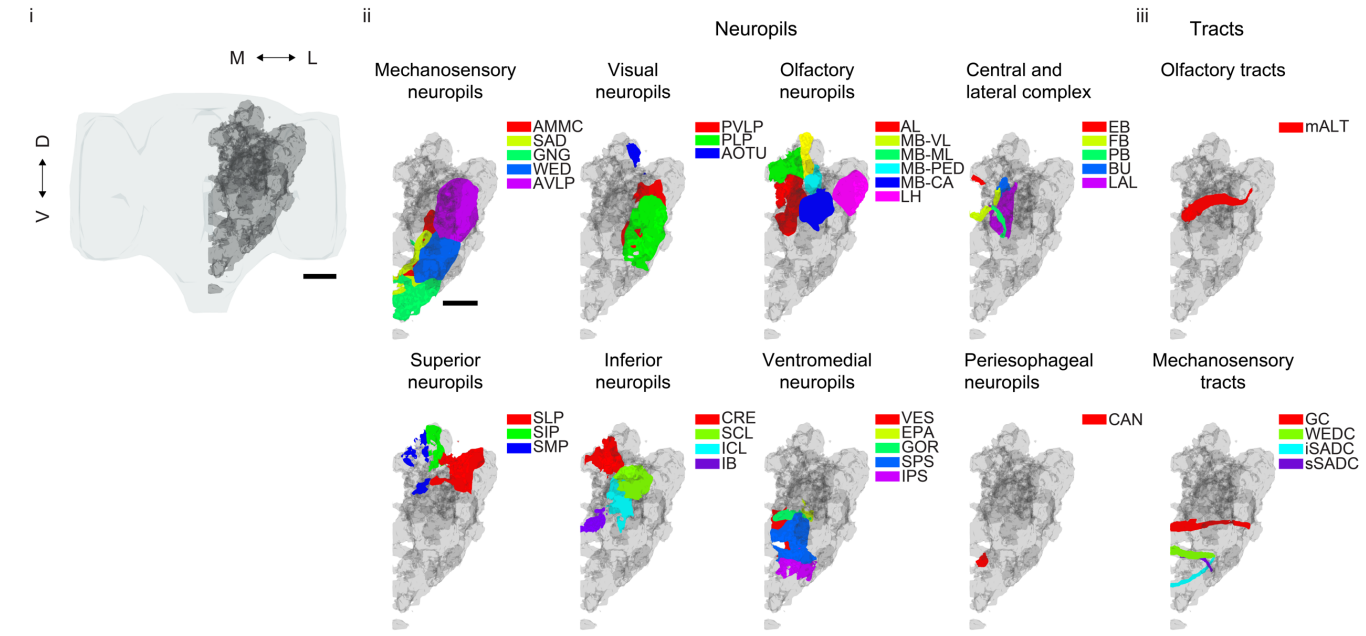
(G) Anatomical annotation of IVIA. Canonical neuropil and neurite tract segmentations from IBN mapped to IVIA (for neuropil and neurite tract names see Table S1). Scale bars in all panels, 100 µm.

See also Figures S1, S2, S3, and Movies S1-4.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2





bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. Brain-wide auditory-evoked activity requires an intact auditory receiver

(A) Responses from auditory ROIs to pulse, sine, and white noise stimuli (n = 33 flies, 19,036 ROIs). Each row is the across-trial median and z-scored DF/F response to each stimulus (6 trials per stimulus).

(B) Spatial distribution of auditory ROIs (magenta) combined from two flies (one ventral and one dorsal volume) (ROI spatial components in IVIA space). Grey contour depicts the optic lobes; black contour depicts the central brain.

(C) Maximum projection from two orthogonal views of the density of all auditory ROIs throughout the central brain (n = 33 flies, 19,036 ROIs).

(D) Schematic of the canonical mechanosensory pathway in *D. melanogaster*. JONs in the antenna project to the AMMC, GNG and WED. AMMC neurons connect with the GNG, WED, SAD, and AVLP.

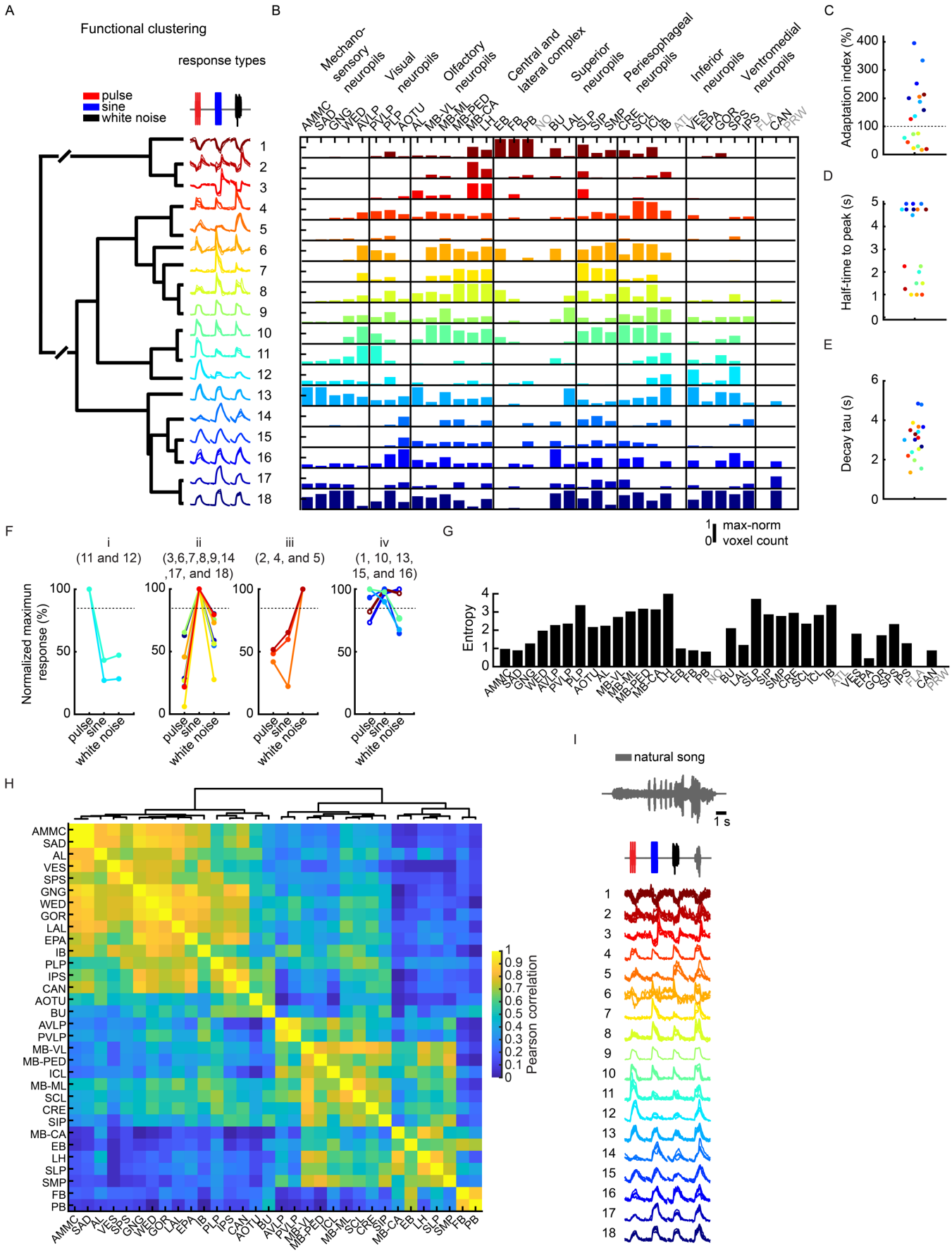
(E) Fraction of auditory voxels by central brain neuropil; percentages averaged across 33 flies (a minimum of 4 flies with auditory activity in a given neuropil was required for inclusion). The number of flies with auditory responses in each neuropil are indicated in parentheses.

(F) 3D-rendering of all auditory voxels across all flies (gray) and their distribution across brain neuropils and tracts. i) Posterior view of 3D-rendering of auditory voxels. ii) 3D-rendering of neuropils and iii) neurite tracts with auditory activity (n = 33 flies). Scale bars in all panels, 100 µm.

See also Figures S4 and S5.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3. Diverse auditory activity throughout central brain

(A) Hierarchical clustering of auditory responses into 18 distinct response types (see Methods). Thick traces are the mean response per response type to pulse, sine, and white noise. Thin traces are mean responses of subclusters within each response type (n = 33 flies, 19,036 ROIs).

(B) Distribution of response types across 36 central brain neuropils. Histogram of voxel count per response type is max-normalized for each neuropil (normalized per column). Colorcode is the same as in panel (A) (n = 33 flies). Neuropils with no auditory activity are indicated in gray font (also for panel G).

(C-E) Kinetics of auditory responses. Adaptation (C), Half-time to peak (D), and Decay time tau (E) of response types. Colorcode is the same as in panel (A).

(F) Diversity in tuning to auditory stimuli. i) pulse-preferring, ii) sine-preferring, iii) white noise preferring, and iv) non-selective response types (preference is defined by the stimulus that drives the maximum response (at least 15% greater than the second highest response). Non-selective responses are divided into sine-and-pulse preferring (solid circles) and sine-and-white-noise preferring (open circles). Colorcode is the same as in panel (A).

(G) Diversity of auditory responses across central brain neuropils. Y axis is the entropy (see Methods) of response type distributions shown in panel B.

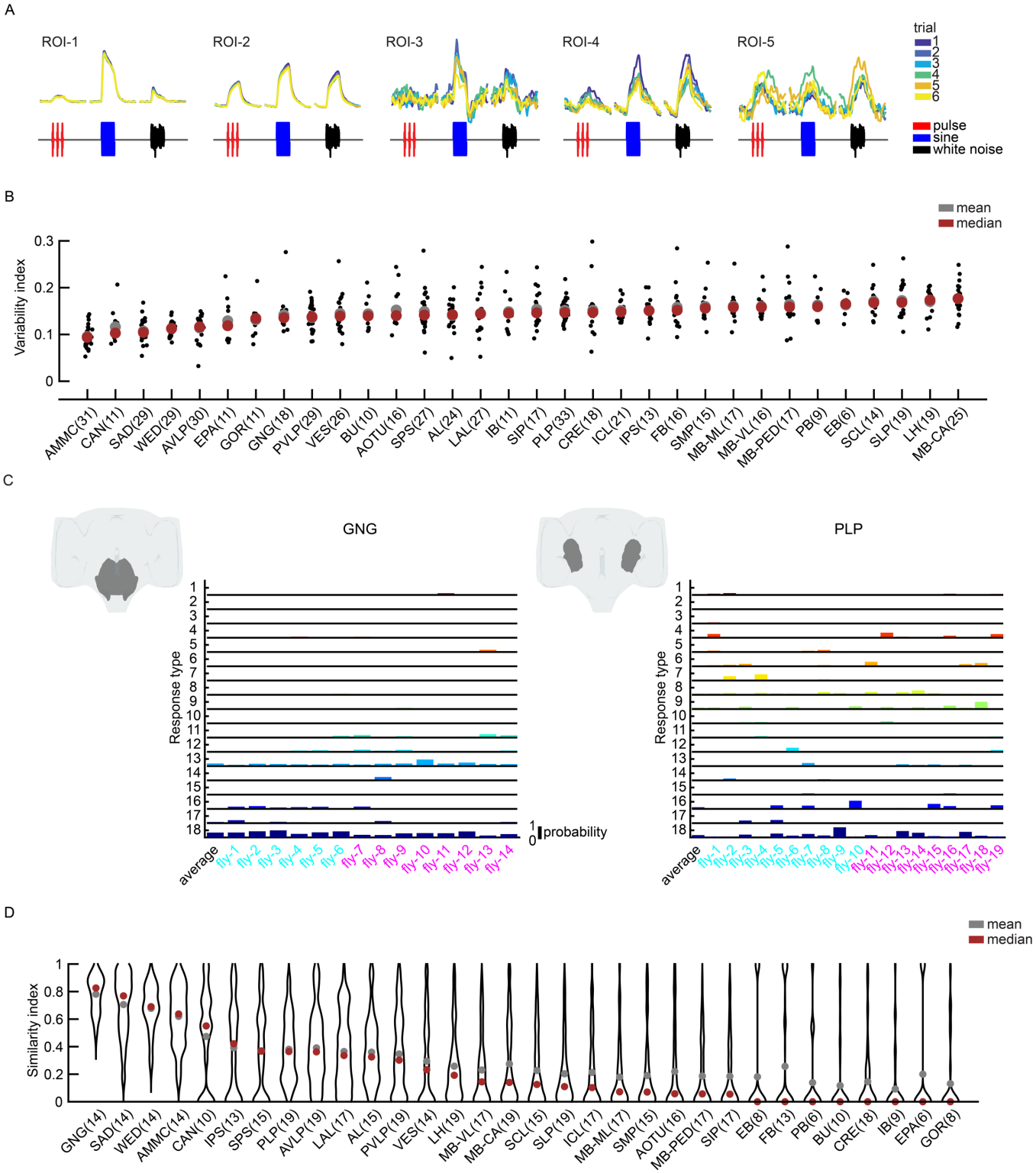
(H) Pairwise Pearson's correlation coefficients of response type distributions (B) between neuropils. Neuropils are ordered based on the hierarchical clustering of the correlation matrix. Only positive correlation values are plotted for clarity.

(I) Auditory responses to natural song; temporal responses are sorted as in panel (A) (n = 10 flies, 2,258 ROIs). Thick traces are the mean response per type and thin traces are the mean responses of each subcluster within each response type.

See also Figures S6 and S7.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4. Auditory activity is more similar across trials and individuals in early mechanosensory areas

(A) Example auditory responses to pulse, sine and white noise from five ROIs across trials.

(B) Across-trial variability index was computed by measuring the standard deviation of median-normalized responses across trials (see Methods). For each neuropil, black dots are the mean variability index per fly (across ROIs), gray and red dots are the mean and median variability index across flies, respectively. Number of flies per neuropil are shown in parenthesis.

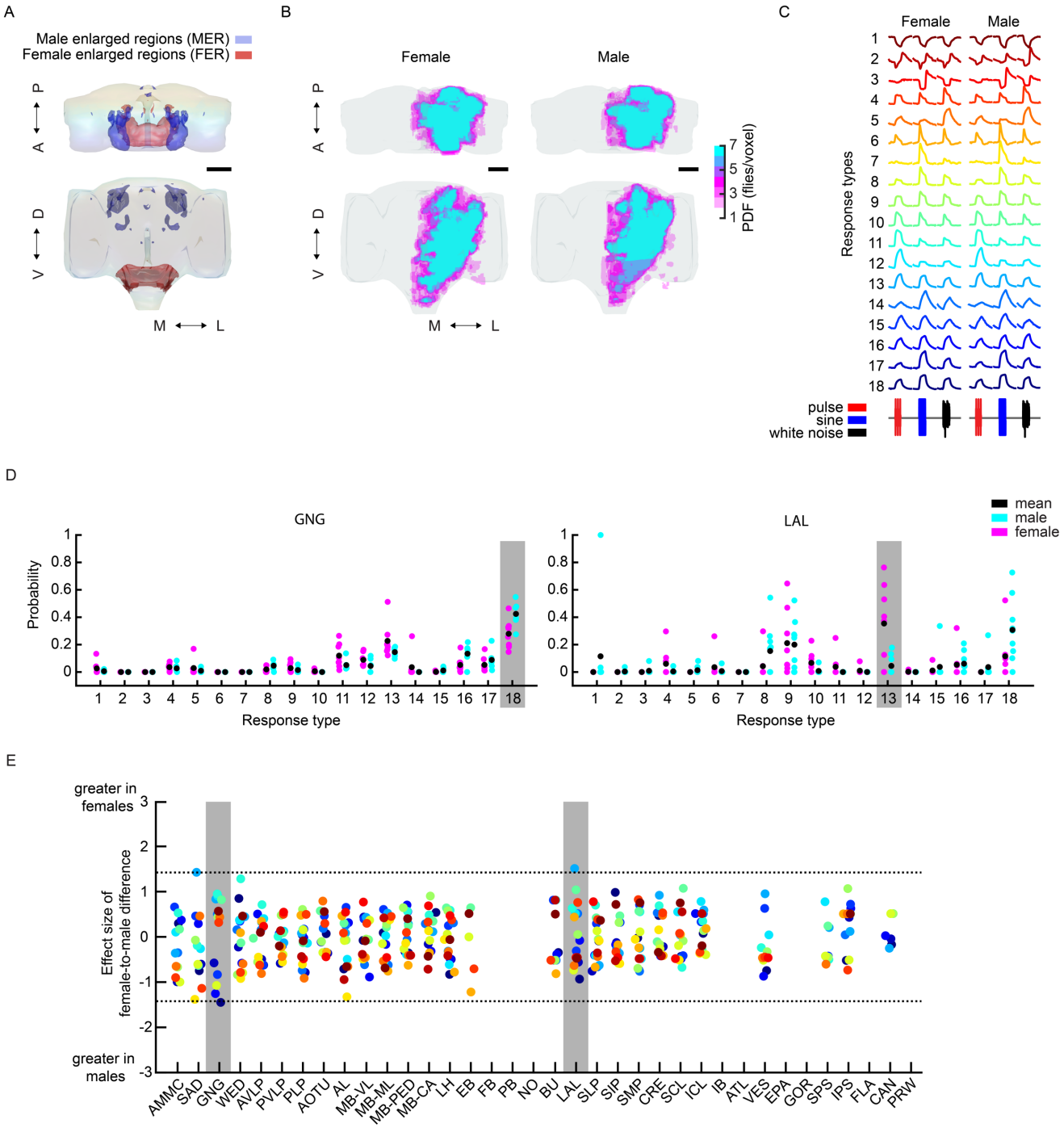
(C) Stereotypy of auditory response types across individuals for the GNG and PLP. Color code of bars per response type are the same as in 3B.

(D) Across-individual similarity index was computed by measuring the dot product between response type distributions (as in 3B), per neuropil and individual (see Methods). Gray and red dots are the mean and median across-individual similarity index, respectively. Number of flies per neuropil are shown in parenthesis.

See also Figures S8 and S9.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5. Auditory responses are not sexually dimorphic

(A) Structural dimorphism in IVIA atlas. Images are the projection from two orthogonal views of male and female enlarged regions (MER and FER, depicted in blue and red respectively). MER are mostly confined to AVLP, AOTU, MB, SLP and ICL, while FER is confined to the GNG (see figure S11A for more details).

(B) Spatial distribution of auditory activity across sexes. Images are the maximum projection from two orthogonal views of the density of auditory responses in females (n = 17) or male flies (n = 16).

(C) Auditory activity per response type across sexes. Auditory responses from female and male flies are sorted as in Figure 3A, and mean response across ROIs belonging to a response type is shown. Auditory responses are highly similar across sexes for all response types.

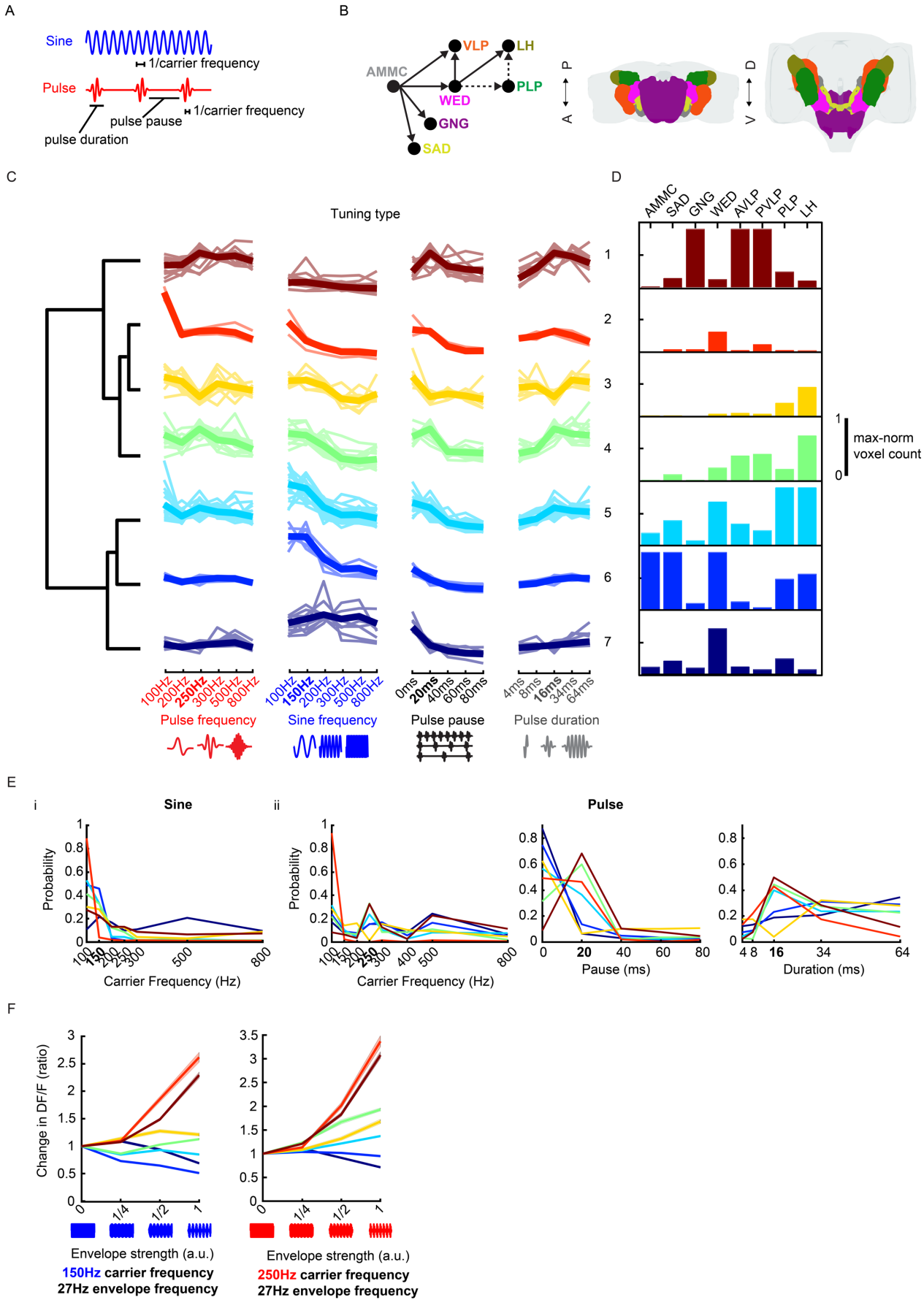
(D) Activity profiles across flies sorted by sex (males in cyan, and females in magenta) are shown for two example neuropils, the GNG and LAL. Probability of response types is variable across individuals, as exemplified in the LAL, but slight sex differences are nonetheless observed (response type 18 in the GNG, or response type 13 in the LAL, both shaded in gray).

(E) Sex-related differences in auditory activity profiles across neuropils. Each dot is effect size (see Methods) of the difference in probability for each response type across sexes, color denotes response type ID as in 5C. Neuropils with the greatest effect size are the GNG (response type 18) and LAL (response type 13), highlighted in gray shading. However these differences are not significant ($p > 0.05$). Scale bars in all panels, 100 µm.

See also Figure S10.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6. Auditory tuning to song features across brain neuropils

(A) Spectral and temporal features of song. Pulse song is characterized by its carrier frequency, pulse pause and duration; while sine song is mainly characterized by its carrier frequency.

(B) Schematic of putative pathway starting in the AMMC mechanosensory and selected downstream neuropils imaged, and their connectivity (Lai et al. 2012; Matsuo et al. 2016; Kamikouchi, Shimada, and Ito 2006; Frechter et al. 2018; Yu et al. 2013; M. Ito et al. 2013; Suver et al. 2019). Dotted lines indicate proposed connections based on anatomy and herein functional results.

(C) Hierarchical clustering of auditory tuning curves into 7 distinct tuning types. Thick traces are the mean z-scored response magnitudes (80th percentile of activity during stimuli plus 2 seconds after) to the stimuli indicated on the x axis (across ROIs and within each tuning type). Thin traces are mean z-scored response magnitudes of subclusters within each tuning type (n =21 flies, 10,970 ROIs).

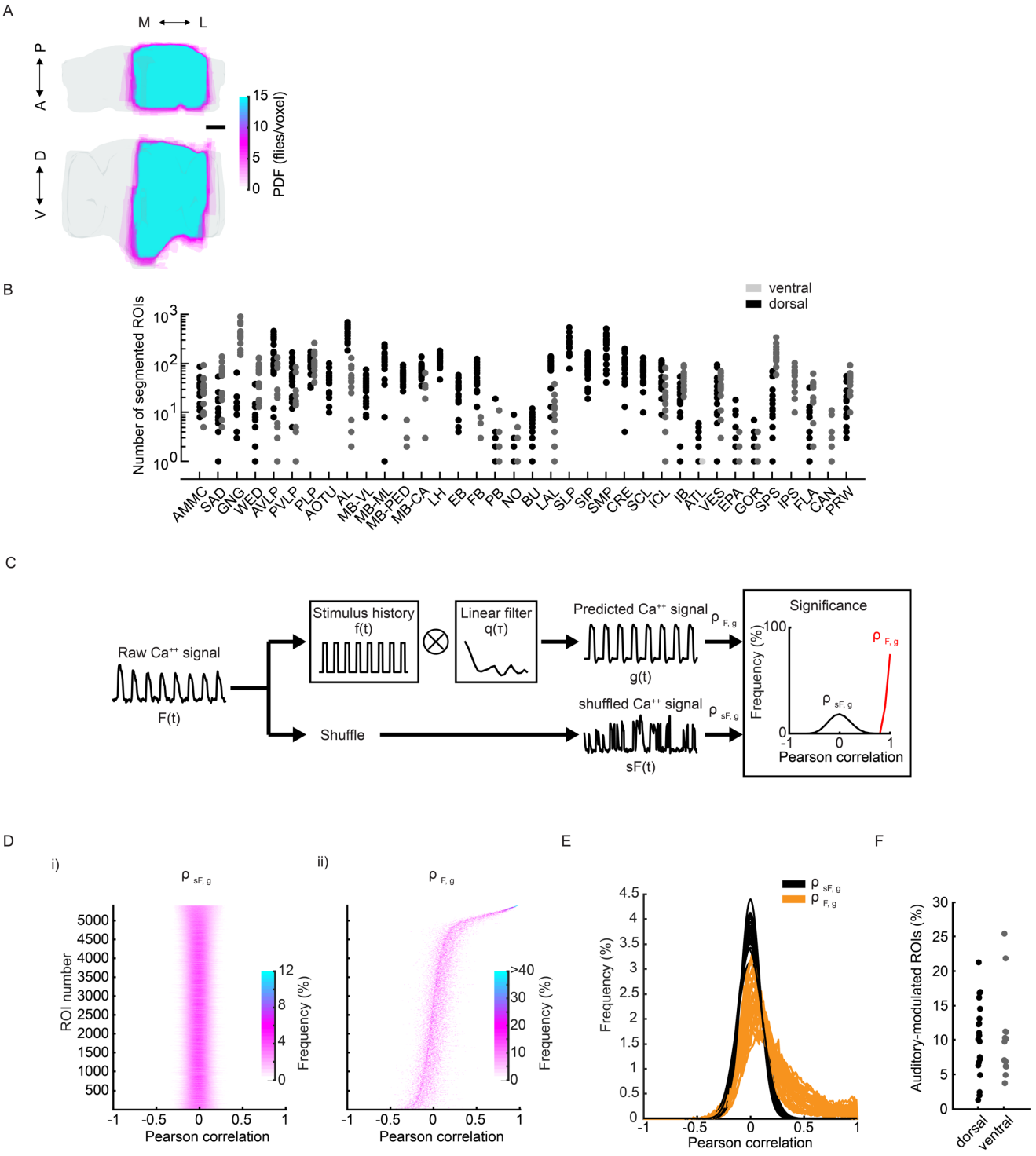
(D) Distribution of tuning types across selected brain neuropils. Histogram of voxel count per response type is max-normalized for each neuropil (normalized per column). Colorcode is the same as in panel C (n = 21 flies).

(Ei-ii) Probability distributions of responses to sine (i) and pulse (ii) stimuli, separated by tuning type (preference within each song feature is defined as the stimulus with the maximum response magnitude). Colorcode is the same as in panel C.

(F) Effect of amplitude envelope on responses to sine tones. Amplitude envelopes (~27Hz envelope frequency) of different strengths 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 of strength 0 (i.e., a sine tone of 150 or 250 Hz, with no envelope modulation) and sorted by tuning type. Thick traces are the mean normalized response magnitudes per tuning type, and shading is the s.e.m. Colorcode is the same as in panel C (n = 21 flies).

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S1



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S1. Spatial coverage of imaged activity and measurement of auditory-evoked activity

(A) Maximum projection from two orthogonal views of the density of segmented ROIs from all imaged volumes (n = 33 flies, 185,395 ROIs) - segmented ROIs (all projected onto ½ of the brain) cover the entirety of the D-V and A-P axes.

(B) Number of segmented ROIs across neuropils (see Figure 1F-G) sampled by ventrally versus dorsally imaged volumes (n = 33 flies, 185,395 ROIs).

(C) Schematic of the method to identify stimulus-modulated ROIs. Raw Ca^{++} signal (F(t)) is modeled as a convolution of stimulus history (f(t)) and a set of filters per stimulus type (q(τ)). Auditory modulation is measured by the cross-validated correlation scores ($\rho_{F,g}$) between raw and predicted Ca^{++} signal. Correlation of shuffled 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 of $\rho_{F,g}$.

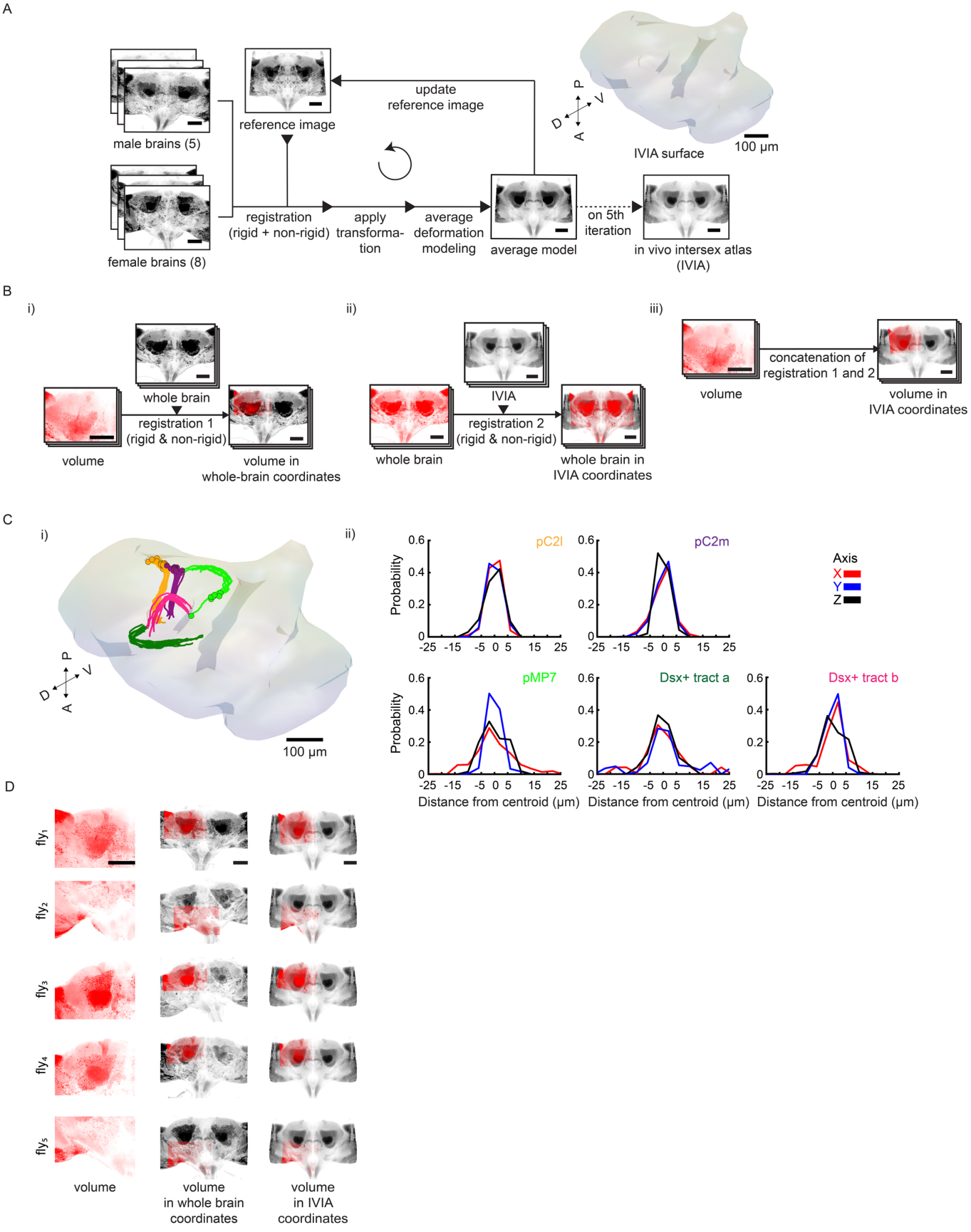
(D) Distribution of Pearson correlation coefficients of i) shuffled-vs-predicted ($\rho_{sF,g}$) and ii) raw-vs-predicted ($\rho_{F,g}$) signals per ROI from a single fly. ROIs were sorted by the median $\rho_{F,g}$ (n = 1 fly, 5,410 ROIs). Only a fraction of ROIs have $\rho_{F,g}$ beyond the null distribution ($\rho_{sF,g}$).

(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.

(F) Frequency of detected auditory ROIs across flies using method described in panel C (n = 33 flies).

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S2



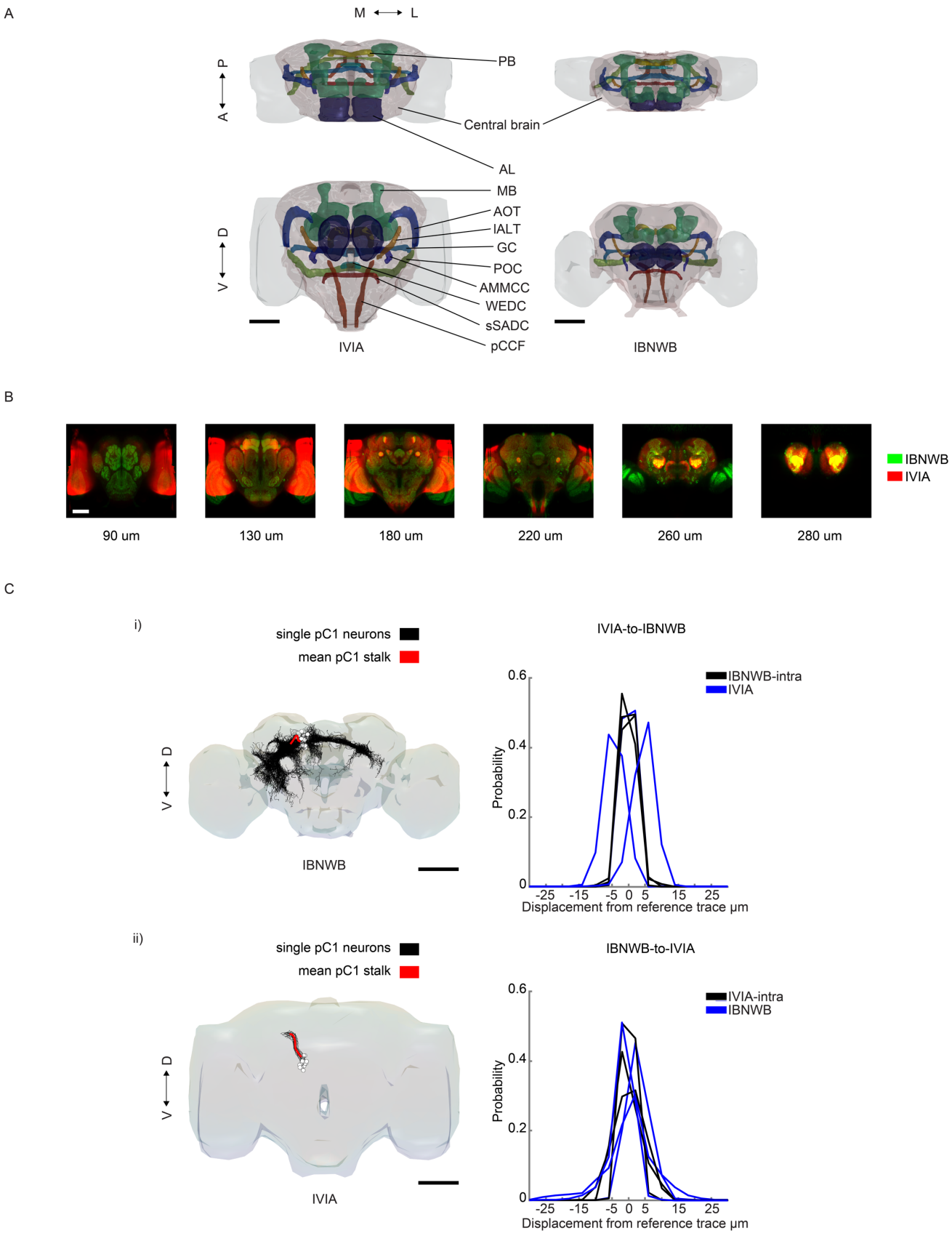
bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S2. Building the *in vivo* intersex atlas (IVIA)

A) Schematic of IVIA generation: images of male (n = 5) and female (n = 8) brains expressing membranal 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) Schematic of registration pipeline of functionally-imaged volumes to IVIA. i) volumes per fly are registered to their own private whole brain atlas, ii) private whole brain atlases (one per fly) are registered to the IVIA, iii) registrations from i) and ii) are concatenated to map volumes to IVIA space.
(C) IVIA registration accuracy: i) 3D-rendering of traced tracts of Dsx+ neurons (pC2m, pC2l), Fru+ neurons pMP7, and extra Dsx+ neuronal tracts (a and b), ii) Per-axis jitter (X, Y, and Z) between matched traced tracts of pC2m, pC2l, pMP7, and Dsx+ tract-a/b across flies (n = 10, 12, 7, 9, and 9 hemibrains respectively).
(D) Example imaged tdTomato volumes 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 panel Bii). Scale bars in all panels, 100 µm.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S3



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S3. Registration between *in vivo* intersex atlas (IVIA) and fixed-brain atlas (IBNWB)

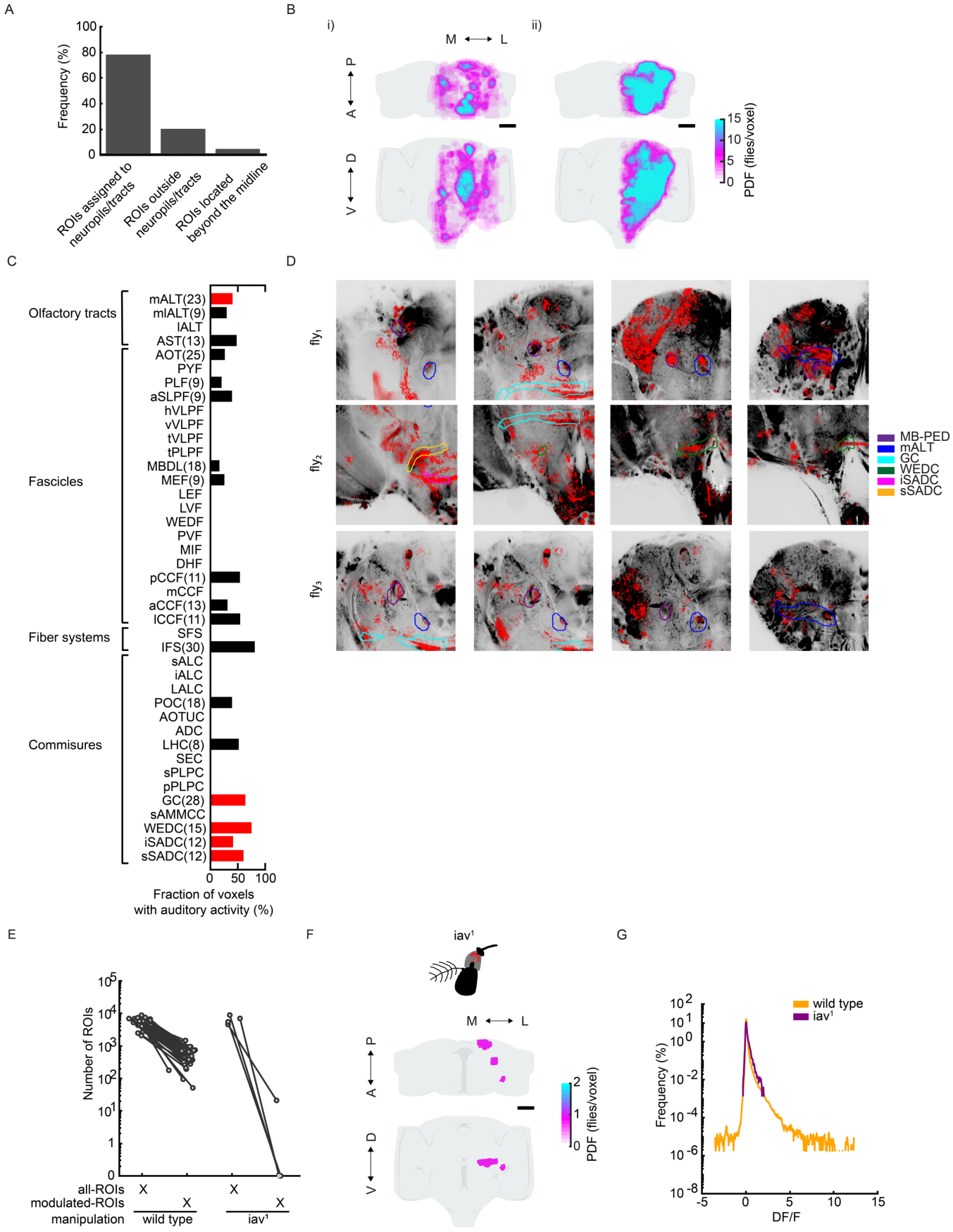
(A) Central brain neuropils and fiber bundles used for point set registration. Points from meshes of segmented antennal lobe (AL), mushroom body MB (includes mushroom body lobes, peduncle and calyx), protocerebral bridge (PB), antennal mechanosensory and motor center commissure (AMMCC), anterior optic tract (AOT), great commissure (GC), wedge commissure (WEDC), posterior optic commissure (POC), lateral antennal lobe tract (lALT), 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.

(B) Overlay of IVIA (red) and registered IBNWB (in IVIA space, green) at different depths (90, 130, 180, 220, 260, and 280 um). 0 um is the most anterior section of the brain and 300 um the most posterior.

(C) Atlas-to-atlas registration accuracy measured using pC1 stalks from IBNWB and IVIA. i) pC1 traces from IBNWB; black traces are single pC1 neurons (IBNWB) and red trace is the mean reference pC1 stalk. IVIA-to-IBNWB transformation increases the jitter across all axes from the reference mean pC1 trace by ~2.24 mm. ii) pC1 traces from IVIA; black traces are single pC1 stalks (IVIA) and red trace is the mean reference pC1 stalk, IBNWB-to-IVIA transformation increases the jitter by ~2.8 µm. Scale bars in all panels, 100 µm.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S4



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S4. Auditory activity in neurite tracts (wild type flies) and in *iav*^{1} flies

(A) Distribution of auditory ROIs within and outside neuropils and tracts. 77.17% of auditory ROIs are assigned to neuropils/neurite tracts, 22.83% of auditory ROIs fall outside neuropils/neurite tracts (but lie within the same hemibrain), and 0.3% of auditory ROIs are beyond the midline.

(B) Maximum projection from two orthogonal views of the density of i) auditory ROIs outside neuropils/neurite tracts (n = 33 flies, 4,346 ROIs), and ii) auditory ROIs within neuropils/neurite tracts (n = 33 flies, 14,658 ROIs).

(C) Fraction of auditory voxels by central brain neurite tracts, data averaged across 33 flies. A neurite tract is considered to have auditory activity if auditory activity is found in at least 4 flies (Fraction for neurite tract with auditory activity in less than 4 flies are not shown). Red represents tracts that were clearly distinguishable from neuropil by visual inspection.

(D) Auditory ROIs from three flies - clustering of ROIs in several tracts is indicated.

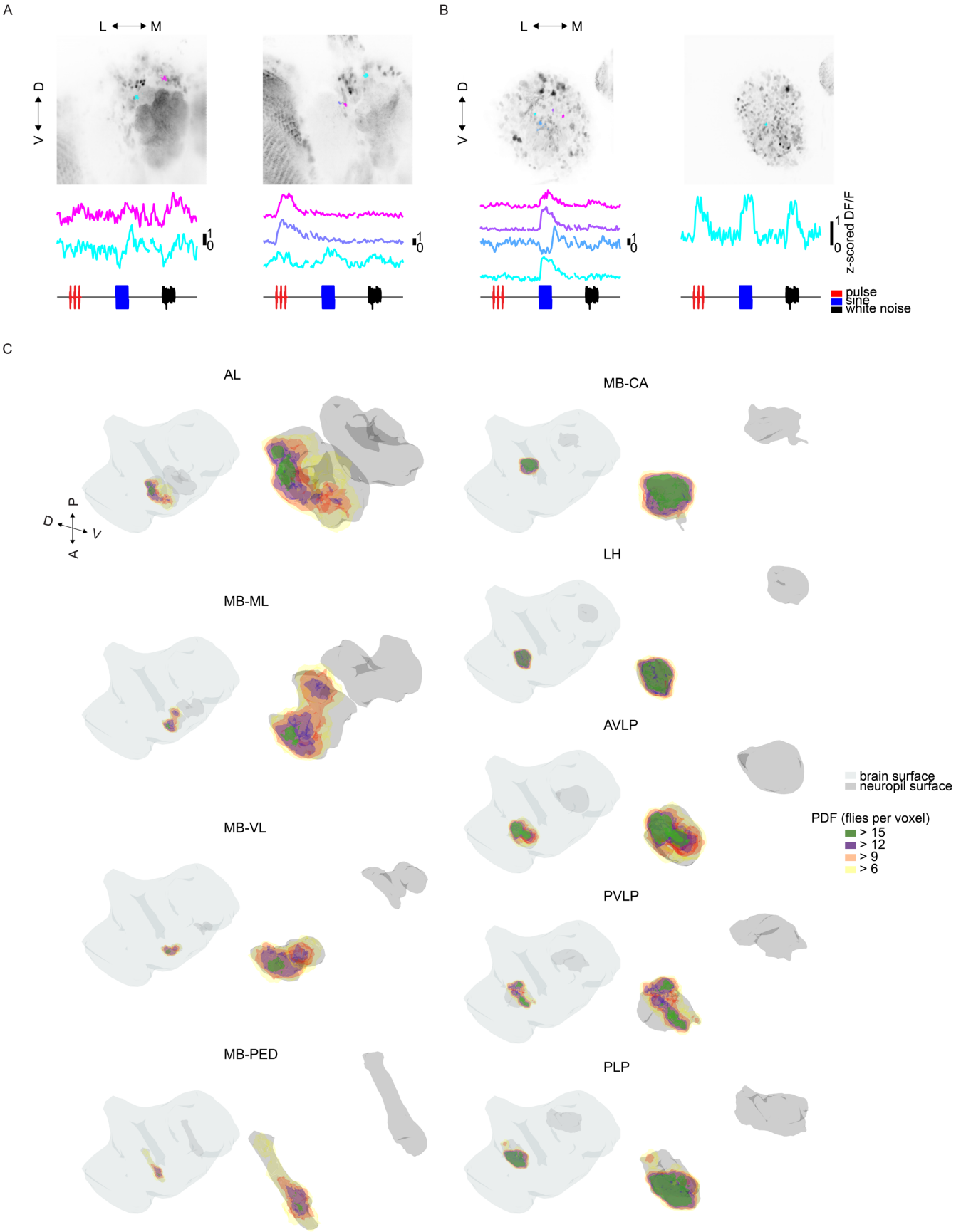
(E) Detected auditory ROIs in wild type (n = 33 flies, 19,036 ROIs) and *iav*^{1} flies (n = 4 flies, 21 ROIs).

(F) Maximum projection from two orthogonal views of the density of auditory ROIs in *iav*^{1} flies (all 21 ROIs come from 1 out of 4 flies imaged). These ROIs were located in the mushroom body calyx (MB-CA), posterior VLP (PVLP), and the posterior lateral protocerebrum (PLP).

(G) Distribution of DF/F during stimulus (from stimulus start until 5 seconds after stimulus offset) for auditory ROIs in wild type (n = 33 flies, 19,036 ROIs) and *iav*^{1} flies (n = 4 flies, 21 ROIs).

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S5



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S5. Auditory activity in olfactory and visual brain regions

(A) Auditory responses from individual antennal lobe projection neurons. Gray image is the average GCaMP6s image (averaged over time) and individual ROIs are indicated in color.

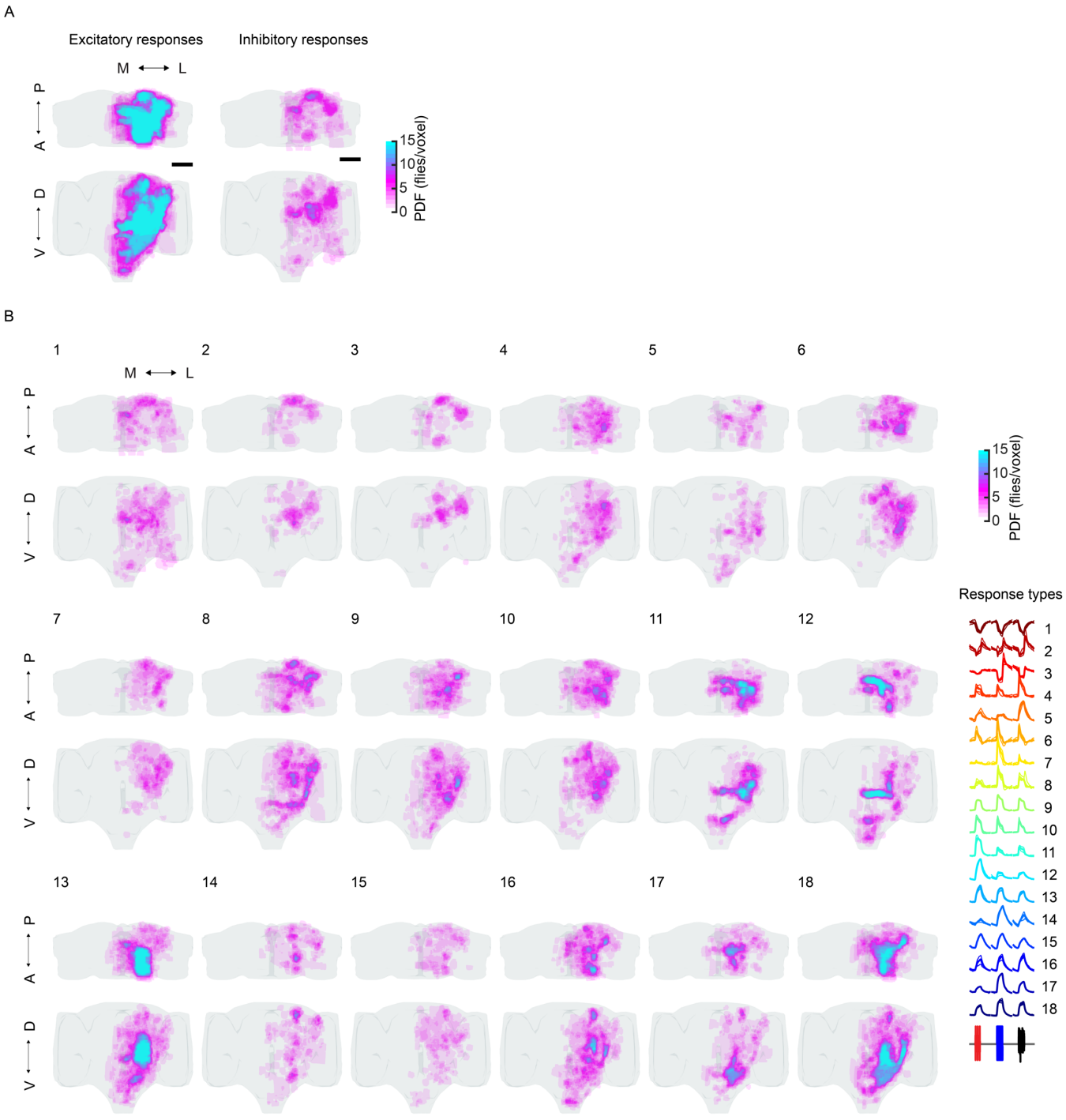
(B) Auditory responses from individual Kenyon cells. Gray image is the average GCaMP6s image (averaged over time) and individual ROIs are indicated in color.

Bottom traces for panels A and B are the median z-scored DF/F responses (across 6 trials) to pulse, sine, and white noise stimuli.

(C) Spatial distribution of auditory responses within selected olfactory (AL, MB-ML, MB-VL, MB-PED, MB-CA, and LH), visual (PVLP, and PLP) and mechanosensory neuropils (AVLP). Auditory activity is spatially restricted within each neuropil.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S6



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

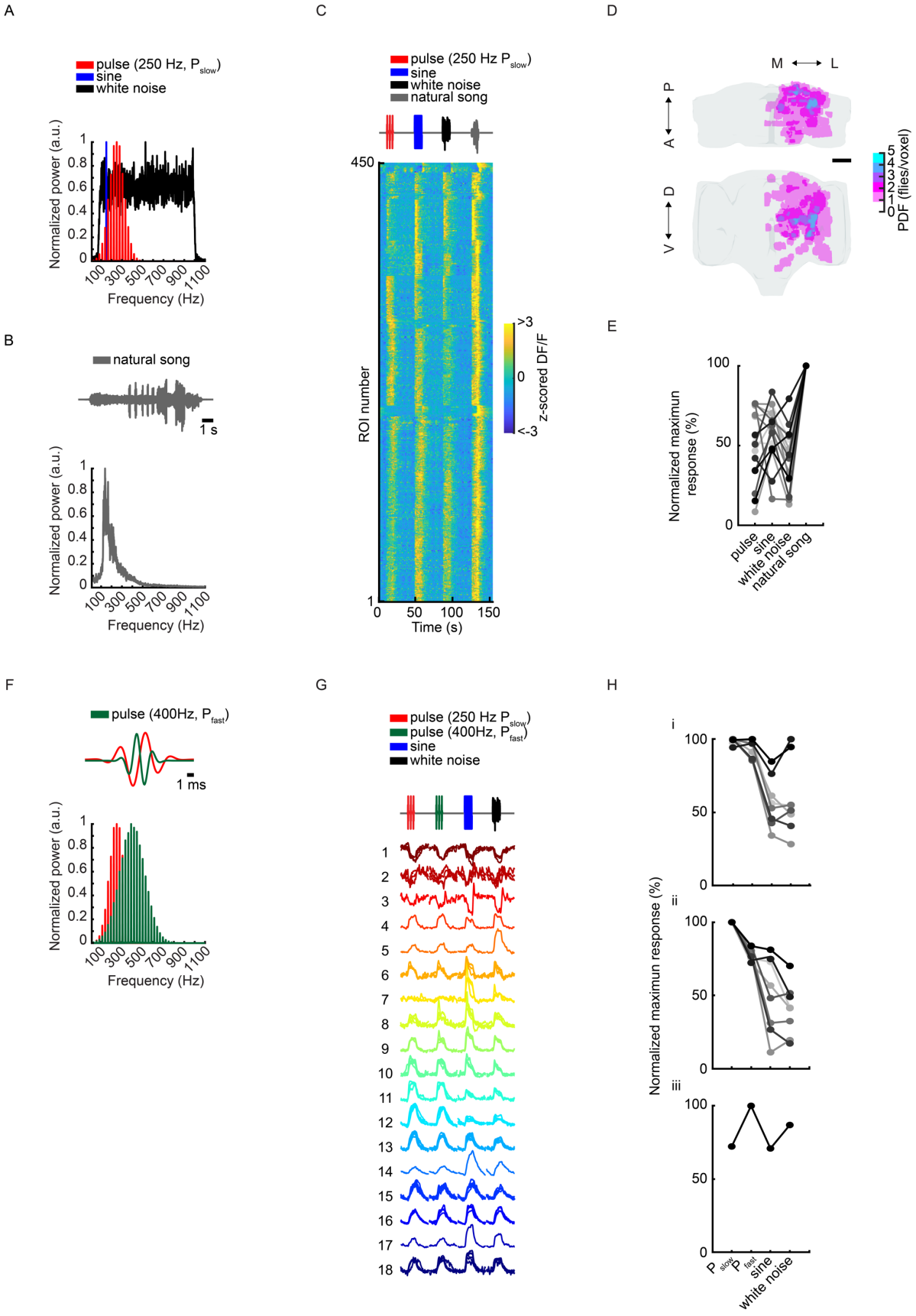
Figure S6. Spatial distributions of auditory response types

(A) Spatial distribution of excitatory vs inhibitory auditory responses across the central brain. Images are the maximum projection from two orthogonal views of the density of excitatory and inhibitory response types across flies (n = 33).

(B) Spatial distribution of auditory activity belonging to each response type across the central brain. Images are the maximum projection from two orthogonal views of the density of each response type across flies (n = 33). Scale bars in all panels, 100 µm.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S7



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S7. Auditory responses to natural song and fast pulses

(A) Spectral profile of auditory stimuli - pulse (P_{slow}), sine, and white-noise - used to classify response types in Figure 3A, and their spectral features.

(B) Spectral profile of natural song used in Figure 3G.

(C) Responses from auditory ROIs (to pulse, sine, white noise, and natural song) with larger response to natural song (includes all subclusters from Figure 3I, n = 10 flies, and 450 ROIs with larger response to natural song stimulus). Each row is the median z-scored DF/F response to each stimulus across trials.

(D) 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 = 10, 2,258 ROIs).

(E) Subclusters from Figure 3G that preferred natural song (~21% of subclusters from panel (G)).

(F) Spectral profile of P_{fast} and P_{slow}.

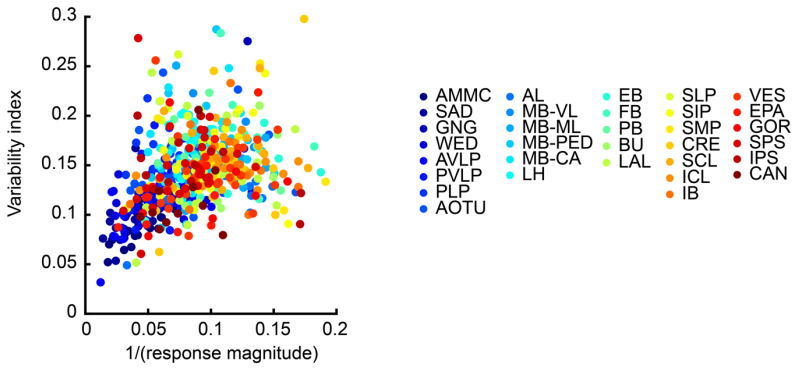
(G) Auditory responses to second pulse mode P_{fast}; auditory activity is sorted by response type as in Figure 3A (n = 2 flies, 2,193 ROIs).

(H) Subclusters from panel (G) with i) broad preference for pulse, ii) P_{slow} preference, or iii) P_{fast} preference.

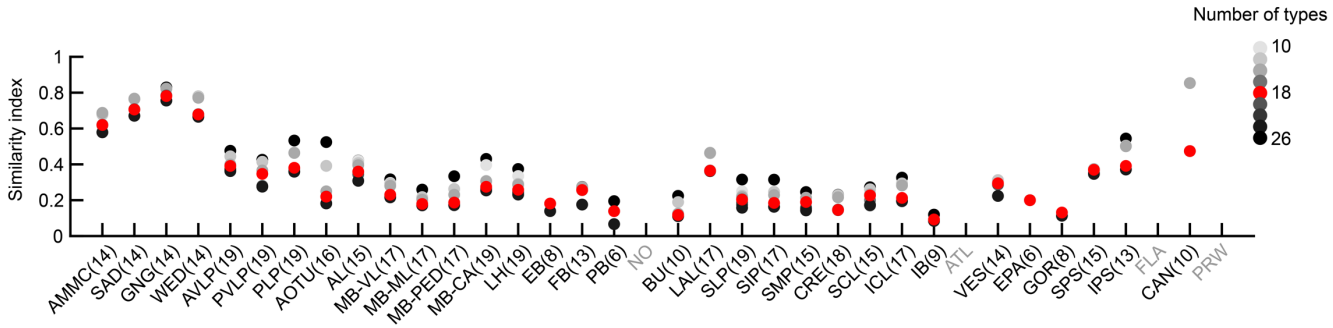
bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S8

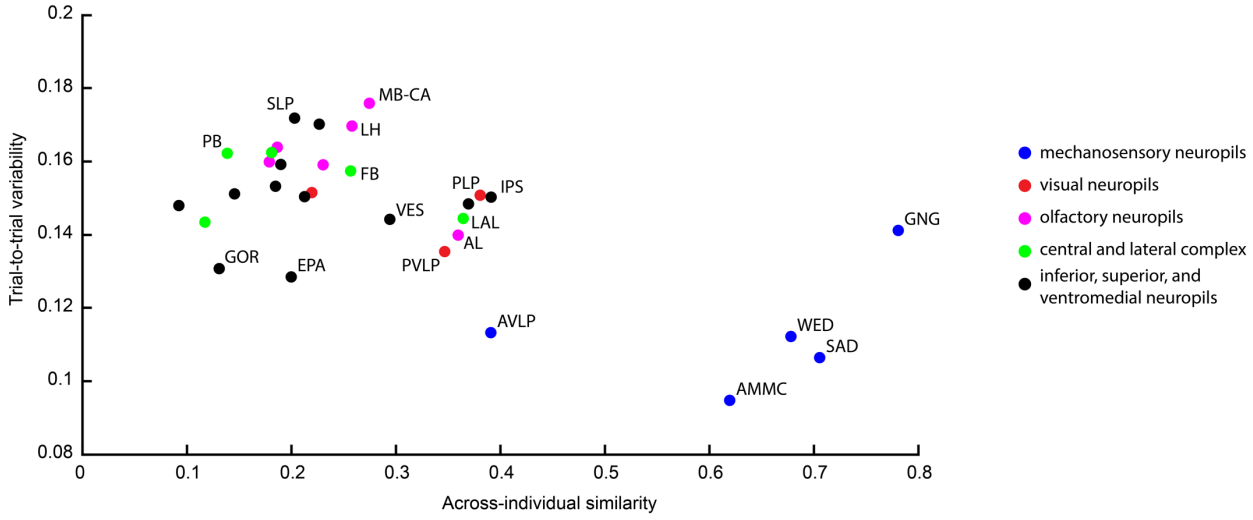
A



B



C



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S8. Across-trial and across-individual comparisons of auditory activity by neuropil

(A) Comparison of across-trial variability index and response magnitude. Each dot is the mean variability index and response magnitude (80th percentile of z-scored DF/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 to response magnitude.

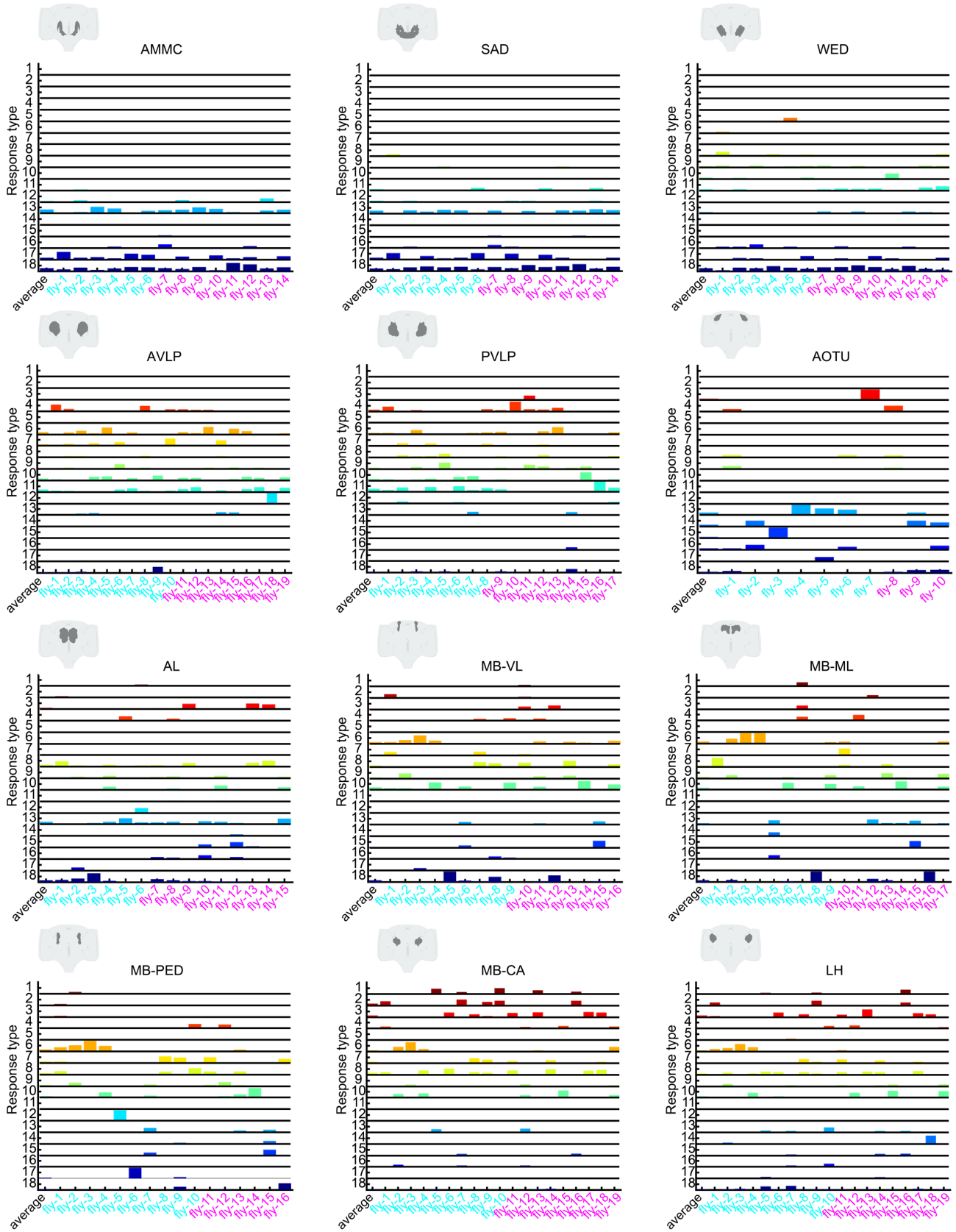
(B) Robustness of differences in similarity index across neuropils to the number of response types (see Figure 3A). Similarity index per neuropil (as in Figure 4D) is calculated for different numbers of response types (from 10 to 26 - in Figure 3 we used 18 types).

(C) Comparison of across-trial variability and across-individual similarity. Each dot is the mean trial variability index and across-individual similarity index per neuropil (across flies) as in Figures 4B and 4D. Types of neuropils are color coded according to the legend. Early mechanosensory neuropils have low across-trial variability and high across-individual similarity.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

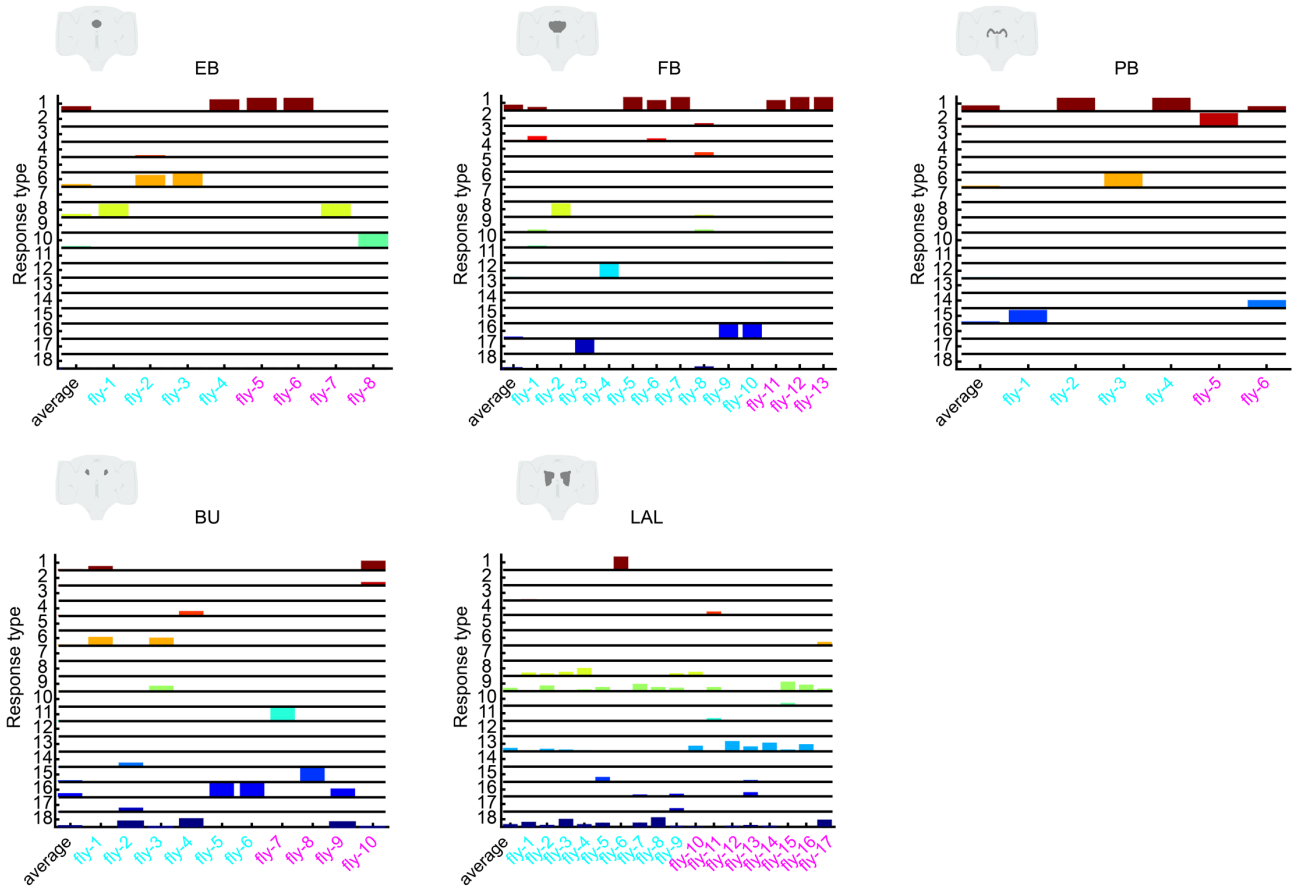
Figure S9

A



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S9. Auditory response types separated by individual

(A) Temporal response profiles, similar to Figure 4C, for all mechanosensory, visual, olfactory, central and lateral complex neuropils. For each neuropil, male and female flies are depicted in cyan and magenta respectively.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

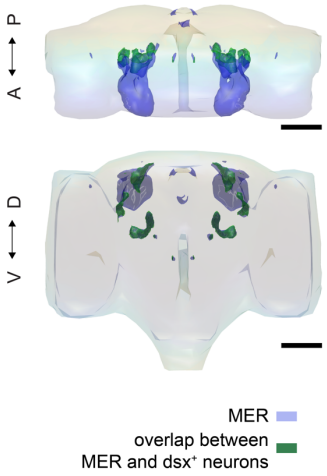
Figure S10

A

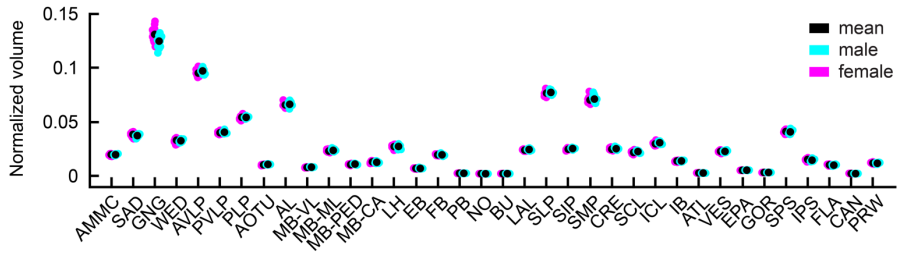
Overlap with neuropils (%)

	IVIA		IS2	
Neuropil	MER	FER	MER	FER
AMMC	0	0	12.3674	0
SAD	0	1.9599	0	0.010439
GNG	0	23.0398	0.32129	0.064546
WED	0	0.002494	0.066076	0.16347
AVLP	4.2354	0	3.6569	0.42645
PVLP	0.60815	0	3.0846	0.01191
PLP	0	0.0038051	0	8.9941
AOTU	34.5342	0	38.8137	5.9317
AL	1.211	0	0	8.9314
MB-VL	16.3986	0	25.83	0.51782
MB-ML	1.0809	0	4.4376	2.8325
MB-PED	0.88435	0	0	0.022422
MB-CA	0	0	0.15293	0
LH	0	0	0.17996	6.6362
EB	0	0	3.2346	8.4211
FB	0	0	0	0.52955
PB	0	0	0	0.4122
NO	0	0.22349	1.8496	0
BU	0	0	6.5153	2.0731
LAL	2.8113	0	6.8849	7.4393
SLP	0.60691	0	35.2026	8.672
SIP	15.159	0	0	0.59029
SMP	0.7098	0	6.0956	0.24041
CRE	0	0	14.9103	3.4219
SCL	0.83079	0	23.9035	0.35677
ICL	6.5181	0	0	1.6302
IB	0	0	1.1539	0
ATL	0	0	0	0
VES	0.60022	0	9.8614	0
EPA	2.8483	0	54.4656	0
GOR	0.81936	0	66.8865	0
SPS	0	0	0.16545	0
IPS	0	0	0	0
FLA	0	0	0.39845	0
CAN	0	0	0	0
PRW	0	0	0.87759	0

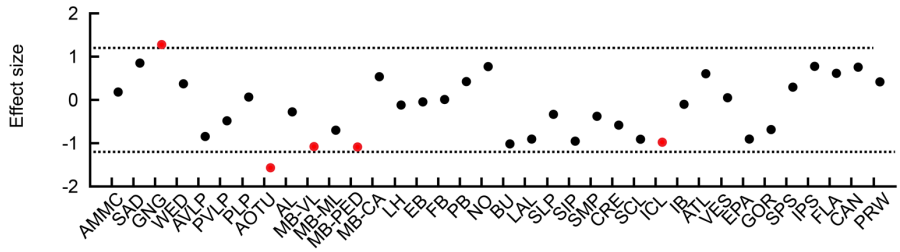
B



C



D



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S10. Structural dimorphisms in IVIA

(A) Fraction of MER or FER voxels by central brain neuropil in *in vivo* (IVIA) vs fixed brain (IS2) atlases. IVIA data is from 22 males and 23 females, IS2 data comes from (Cachero et al. 2010). Neuropils with an overlap of more than 3% of voxels are highlighted in red.

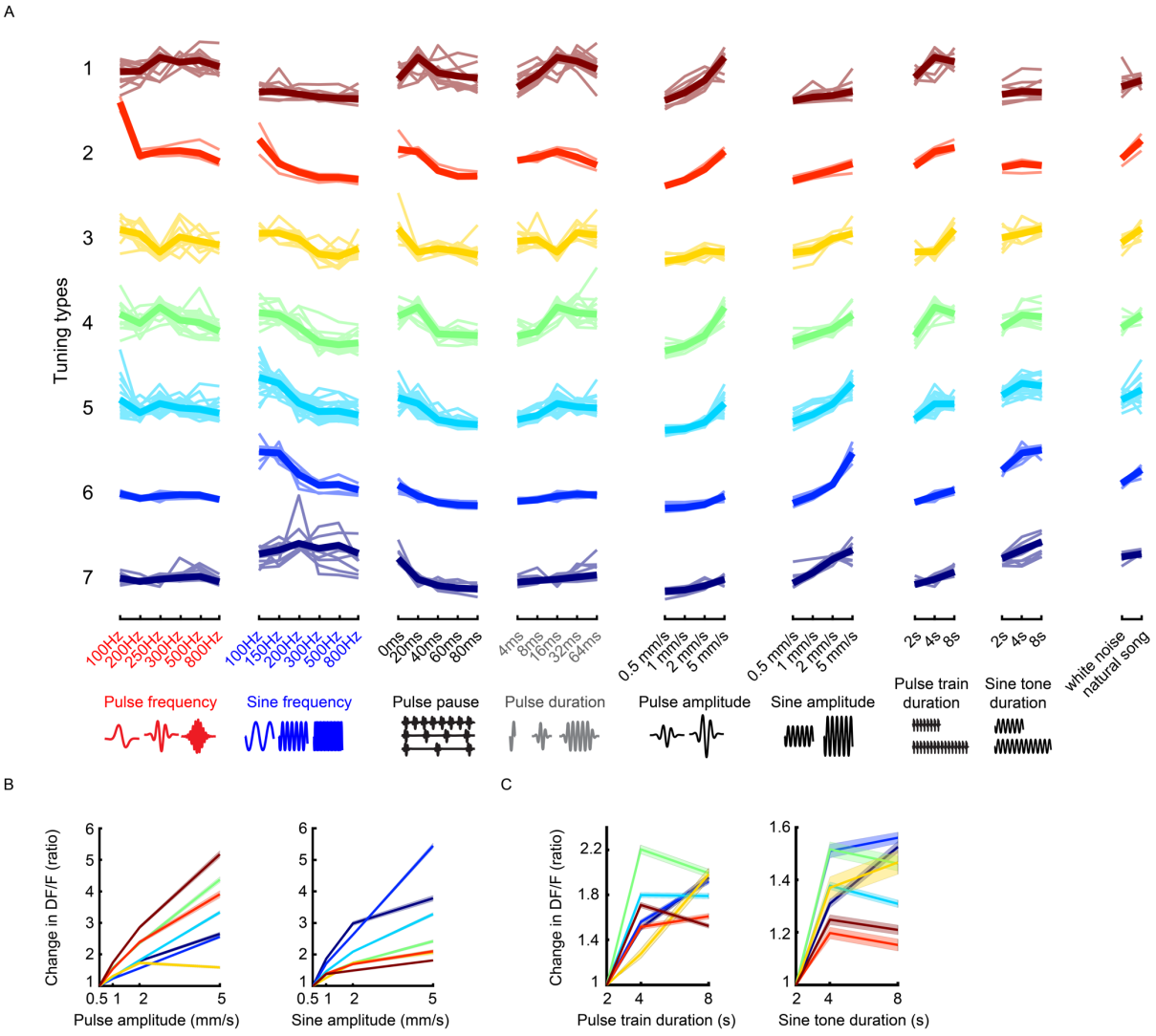
(B) Overlap between *in vivo* MER and the lateral protocerebral complex. Using the dsx^{Gal4} we generated an average male Dsx image that we used to determine the overlap between MER and the lateral protocerebral complex. MER voxels overlapping with the lateral protocerebral complex are depicted in green.

(C) Normalized brain volume of central brain neuropils across flies and sorted by sex (males in cyan, and females in magenta). Volume of each neuropil is normalized by the full central brain volume per fly.

(D) Effect size (Cohen's d) of the normalized brain volume difference between sexes and across neuropils. Red dots indicate significant differences ($p < 0.05$). The GNG is enlarged in females, while the AOTU, MB-VL, MB-PED, and ICL are enlarged in males.

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S11



bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S11. Related to Figure 6

(A) Response of tuning types to song features as in Figure 6A, but including responses to pulses and sines of different amplitudes, different train duration, and also to white noise and natural song. Thick traces are the mean z-scored response magnitudes (80th percentile of activity during stimuli plus 2 seconds) to pulse and sine features per tuning type (across ROIs within each type). Thin traces are mean z-scored response magnitudes of subclusters within each tuning type (n = 21 flies, 10,970 ROIs). Colorcode is the same as in figure 6C (n = 21 flies).

(B) Auditory responses to different sine and pulse intensities. Response magnitude (80th percentile of the activity during stimuli) per ROI is normalized to the response at 0.5 mm/s, and sorted per tuning type. Thick traces are the mean normalized response magnitude per tuning type, and shading is the s.e.m. Colorcode is the same as in figure 6C (n = 21 flies).

(C) Auditory responses to different sine and pulse train durations. Response magnitude (80th percentile of the activity during stimuli) per ROI is normalized to the response to the shortest stimulus (2s), and sorted per tuning type. Thick traces are the mean normalized response magnitude per tuning type, and shade is the s.e.m. Colorcode is the same as in figure 6C (n = 21 flies).

bioRxiv preprint doi: https://doi.org/10.1101/709519; this version posted July 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Movie legends

Movies S1-S4

DF/Fo time-series movies (maximum intensity projected over entire z axis of the Drosophila brain) of male (S1-2) and female (S3-4) brain dorsal (S1 and S3) and ventral (S2 and S4) quadrants. Responses to pulse (250Hz), sine (150Hz) and white noise are shown. Pseudocolor scale is from 0 to 200 %. Movie speed is sped up 5X.

Movie S5

Z-stack of the probability density of auditory ROIs (number of flies containing an auditory ROI at each voxel) throughout the central brain (n = 33 flies, 19,036 ROIs). Mechanosensory neuropils have grey contours, while the rest of the central brain neuropils imaged have red contours. Pseudocolor scale is from 0 to 15 flies/voxel.