

Title: Discovery of a new song mode in *Drosophila* reveals hidden structure in the sensory and neural drivers of behavior

Authors: Jan Clemens^{1,3*}, Philip Coen^{1,4*}, Frederic A. Roemschied^{1*}, Talmo Pereira¹, David Mazumder^{1,2}, Diego Pacheco¹, and Mala Murthy^{1,2}

Affiliations:

¹ Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

² Department of Molecular Biology, Princeton University, Princeton, NJ, USA

³ present address: European Neuroscience Institute, Göttingen, Germany

⁴ present address: University College London, London, UK

* equal contribution

Author for correspondence: Mala Murthy (mmurthy@princeton.edu)

Keywords: *Drosophila*, courtship song, acoustic communication, behavior, neural circuits

Summary

Deciphering how brains generate behavior depends critically on an accurate description of behavior. If distinct behaviors are lumped together, separate modes of brain activity can be wrongly attributed to the same behavior. Alternatively, if a single behavior is split into two, the same neural activity can appear to produce different behaviors [1]. Here, we address this issue in the context of acoustic communication in *Drosophila*. During courtship, males utilize wing vibration to generate time-varying songs, and females evaluate songs to inform mating decisions [2-4]. *Drosophila melanogaster* song was thought for 50 years to consist of only two modes, sine and pulse, but using new unsupervised classification methods on large datasets of song recordings, we now establish the existence of at least three song modes: two distinct, evolutionary conserved pulse types, along with a single sine mode. We show how this seemingly subtle distinction profoundly affects our interpretation of the mechanisms underlying song production, perception and evolution. Specifically, we show that sensory feedback from the female influences the probability of producing each song mode and that male song mode choice affects female responses and contributes to modulating his song amplitude with distance [5]. At the neural level, we demonstrate how the activity of three separate neuron types within the fly's song pathway differentially affect the probability of producing each song mode. Our results highlight the importance of carefully segmenting behavior to accurately map the underlying sensory, neural, and genetic mechanisms.

Results

***Drosophila melanogaster* courtship song comprises at least three, not two, distinct song modes**

Manual inspection of song recordings have identified two distinct song modes in *Drosophila melanogaster*: sine song, which consists of a sustained sinusoidal oscillation, and pulse song, which comprises trains of short impulses (ref). Based on this observation, automated methods were developed to effectively detect and segment these two modes of song (Arthur et al.), and using this software, large data sets have been analyzed to link both genes and neural activity with song patterns [5-9]. However, the claim that song comprises only two modes has never been tested using statistical methods, leaving open the possibility that more than two modes exist (as has been suggested for other *Drosophila* species, [10]). We therefore took a more unbiased approach and clustered 25 millisecond waveforms of previously collected raw song recordings [9]. We first classified each sample as signal or noise based on signal amplitudes that exceeded background noise (33% of the waveforms passed this criterion); no other constraints on waveform shape were imposed (Figure 1A, see STAR Methods for details). We then aligned the signal samples to their peak energy, normalized them to correct for differences in waveform amplitude (induced by variations in the intensity of male singing or his position relative to the microphone), and adjusted their sign so that the waveform was positive immediately preceding the peak (this was done because waveform inversions are likely caused by changes in male position relative to the microphone) (Figure 1B). The resulting set of ~20,000 normalized signals from 47 wild type males of the strain NM91 contained courtship song as well as non-song noises originating from grooming, jumping or other behaviors.

To facilitate the classification and the visualization of the waveforms, we reduced the dimensionality of the data set using t-distributed stochastic neighbor embedding (tSNE) [11,12]. The tSNE method is a nonlinear dimensionality reduction method that preserves local similarity structure in a data set and is therefore particularly suited for classification. We clustered the low-dimensional representation of our signals in two steps. First, we partitioned the signal distribution along local minima using the

watershed algorithm (Figure 1C). This procedure initially yielded seven clusters which represent an over-partitioning of the waveform space, since it cuts the signal space along relatively weak local minima and thereby assigns relatively similar waveforms to different clusters. We therefore employed a second hierarchical clustering step to consolidate the watershed clusters based on the similarity of their centroids (Figure 1D) and we chose the number of modes based on the similarity of the waveforms in the watershed clusters (Figure 1E). This resulted in four distinct signal modes (Figure 1D): i) A “noise” mode lumps all signals that lack common structure across exemplars. ii) A “sine” mode joins three clusters that contain waveforms with sustained oscillations that tile a continuum of carrier frequencies between 120 and 180Hz (Supplemental Figure 1B). The three sine clusters all contain similar waveforms as indicated by the strongly overlapping and high correlation values (Figure 1E, right). This justifies them being merged into a single sine song mode by the hierarchical cluster algorithm. iii) A “pulse” mode with relatively slow (200-250Hz) and symmetrical waveforms connects two very similar clusters with either a stronger leading or lagging lobe. These two clusters form a single pulse mode, since the waveforms are highly correlated (see Figure 1E, left), and they are thus joined by the hierarchical cluster algorithm. iv) A second “pulse” mode that in contrast to the previous one is biphasic (asymmetric) with faster oscillations (250-400Hz). This pulse mode is highly dissimilar with the previous pulse mode (correlation values centered around zero, Figure 1E, right) indicating that it forms a distinct pulse mode. We term the two pulse modes “P_{slow}” and “P_{fast}”, accordingly. Note that while there is considerable variability within each song mode, the modes are clearly distinct from each other.

To facilitate analyses of the production, perception and evolution of these two pulse types within *Drosophila*, we developed a simpler and more efficient pulse type classifier. The method takes as input pulses and sines detected from an automatic song segmenter [13] – which can detect both pulse types with high speed and reliability (detection rates: 78% for all pulses, Supplemental Figure 1B, 80% for P_{slow}, 77% for P_{fast}) – and then classifies returned pulses based on their similarity with templates derived from the tSNE analysis pipeline (Figure 2A and Supplemental Figure 1C, see

STAR methods for details). This classifier reliably and efficiently reproduces the classification from the full analysis pipeline into the slow, symmetrical P_{slow} and the fast, asymmetrical P_{fast} in all individual males of eight geographically diverse wild type strains, suggesting that the existence of two pulse types is common across isolates of *Drosophila melanogaster* (Figure 2B, Supplemental Figures 1D, 2A, B).

We next ran several control analyses to ensure that the two pulse types were not data recording or analysis artifacts. First, we were able to recover both pulse types when using linear (as opposed to nonlinear) dimensionality reduction (principal component analysis instead of tSNE, Figure 2C) and also when omitting parts of the normalization procedure, demonstrating that our finding does not crucially depend on the details of our analysis pipeline (Supplemental Figure 2C). In addition, the existence of two pulse types could be due to the male changing his position relative to the directional microphones used to record song. However, when we induced singing via thermogenetic activation of P1 song pathway neurons in a tethered male walking on a spherical treadmill [5] – thereby fixing his position relative to the microphone – we still observed both pulse types (Figure 2D). Lastly, we examined the fly's pose during singing and found that the two pulse types are associated with distinct, largely unilateral, wing positions (Figure 2E-F): P_{slow} is produced with one wing fully extended (ca. 70°) while P_{fast} is produced most often with one wing only weakly extended (ca. 20°). Together, these analyses demonstrate that the P_{slow} and P_{fast} are separate song modes, likely to be produced by separate motor programs.

What is the function of the minimally extended wing? To answer that question, we recorded song from males that had one wing cut (either left or right). These males showed a reduction of singing to ~50% of intact males (Supplemental Figure 2D). While these males still produced both pulse types, the individuality of these pulses was strongly reduced (Supplemental Figure 2E, F). This suggests that the existence of two wings increases the individuality of pulse shapes produced by males and hence the function of the 2nd wing may be to make the pulse shape more idiosyncratic. However,

we observed no strong effect of cutting one wing on pulse variability (Supplemental Figure 2G).

How do the two pulse types – P_{fast} and P_{slow} – in *Drosophila melanogaster* compare to the pulse types produced in close relatives [10]? We applied our analysis pipeline to song data from two related species, *Drosophila simulans* and *Drosophila mauritiana* (Figure 2G). Since the pulses in *D. simulans* and *D. mauritiana* are about twice as fast as *D. melanogaster*, we adapted the automated song segmentation and pulse normalization procedure to reliably detect and classify pulses in these related species (see STAR Methods for details, Supplemental Figure 1B). Similar to *D. melanogaster*, pulses reproducibly clustered into two major pulse types for all strains of *D. simulans* and *D. mauritiana* examined (Figure 2H). One cluster contains slow, symmetrical pulses and the other faster, asymmetrical pulses (Figure 2H-J, Supplemental Figure 2A, B). *D. melanogaster* pulses are distinguishable from *D. simulans* and *D. mauritiana* pulses only by frequency: P_{slow} and P_{fast} are shifted from ~220 to ~470 Hz and from ~350Hz to 670 Hz with only small differences between *D. simulans* and *D. mauritiana* (Figure 2J). The existence of P_{fast} and P_{slow} in all three species supports the following model for the evolution of pulse shape in the *melanogaster* species group: *D. simulans* and *D. mauritiana* increased the speed with which two ancestral motor primitives – one for each pulse type - were executed. In the sections below, we investigate how the existence of two distinct pulse modes affects our understanding of courtship song patterning and perception in *D. melanogaster*.

Pulse type choice contributes to the modulation of song amplitude with distance

Males structure their songs into bouts composed of trains of sines and pulses [13], with the choice to sing either sine or pulse biased by sensory feedback from the female [9]. How does classifying song now into three modes affect the patterning of song? We found that males do not continually switch between P_{fast} and P_{slow} , but compose pulse trains of one type or the other more frequently than expected by chance (Figure 3A-B). By employing Generalized Linear Models [9,14] to predict pulse type from sensory and movement features (see STAR methods), we found that the distance between the male

and female was the strongest predictor of pulse type (Figure 3C). Males from eight different wild type strains consistently bias toward producing P_{fast} at larger distances from the female (Figure 3D). In addition, we observed that blind males (but not deaf or pheromone insensitive males) produce a significantly higher fraction of P_{fast} pulses (Figure 3E), indicating that vision plays a role in the choice between P_{slow} and P_{fast} . This builds on previous work that demonstrated males produce sine song when close to the female [9] – we now see that males continually bias toward modes with faster – and higher intensity —wing movements, as their distance from the female increases (Figure 3F). We tested each pulse mode for evidence of amplitude modulation with distance [5] and found that while P_{fast} amplitude increases with distance (males ratchet up P_{fast} amplitude as they get farther from the female without changing pulse shape (Figures 3H,J), P_{slow} amplitude modulation was comparatively weak (Figure 3G). Taken together, these data suggest a reinterpretation of previous work on amplitude modulation [5]: the behavior is driven by at least two processes—the choice to sing the softer or the louder pulse type (P_{slow} vs P_{fast}) coupled with analog modulation of P_{fast} amplitude.

Females slow in response to all three song modes

The correlation between the frequency of a song mode and its amplitude (Figure 3F) suggests that males may switch between song types simply as a strategy to produce song signals of different intensities. The modes, however, could also exert different effects on female behavior. Previous studies have shown that both sine and pulse song are correlated with a reduction of locomotor speed in sexually receptive females, and that this correlation depends on females being able to hear the male song [6,9]. We examined female locomotor speed relative to the amount of each song mode the male produced (per 30 seconds of courtship – previous studies showed that the strongest correlation between male song and female speed was on timescales of tens of seconds [6]). We found that all three song modes were correlated with a reduction in female speed, but in a context-dependent manner. P_{slow} and sine were correlated with reductions in female speed across a wide range of distances, whereas P_{fast} was correlated with a reduction in female speed, but only when males were far away from females (and then only weakly (Figure 3J)). All three modes were correlated with

increases in female speed when produced very close to the female. These results only emerge when examining correlations relative to male-female distance (Supplemental Figure 3A-C). That P_{slow} and P_{fast} differentially affect female behavior suggests that they have different “meanings” for the female. Further elucidation of the differential roles of P_{slow} and P_{fast} on female behavior will likely require higher resolution readouts of female postural movement changes [12,15]. The reclassification of male courtship song therefore also has implications for the study of auditory perception in the female nervous system.

Song pathway neurons modulate the choice between song modes

Above, we showed that males select between pulse types based on sensory feedback (Figure 3C, D), and that these two pulse types correspond to different wing poses (Figure 2E). This suggests that neurons within the fly’s song pathway can bias the motor output towards one of the three song modes. To test this hypothesis, we focused on three previously characterized neurons implicated in driving wing extension or generating pulse song (Figure 4A, Supplemental Figure 4B): P1, a cluster of ~20 male-specific *Fruitless*⁺ and *Doublesex*⁺ neurons per hemibrain [16,17], pIP10, a pair of male-specific *Fruitless*⁺ neurons that descend from the brain to the ventral nerve cord (VNC) [18], and ps1, a *Doublesex*⁺ wing muscle motor neuron in the ventral nerve cord [8].

Either P1 or pIP10 neural activation results in song production, even in the absence of a female [5,18]. P1 neurons are known to be involved not only in song production, but also in the integration of pheromonal and visual signals from the female [17,19-21], the production of aggressive behaviors [22], and overall mating drive [23]. We reasoned that when males are close to females, P1 activity should be elevated, since the pheromonal and visual cues provided by the female should be particularly strong. Since, males bias towards both sine song and P_{slow} production when close to the female ([9], Figure 3D), optogenetically activating P1 neurons should preferentially drive either sine or P_{slow} , but not P_{fast} , even in the absence of a female. pIP10 neurons are thought to be postsynaptic to P1 neurons [18], and therefore might exert a similar effect on song patterning.

To investigate how the activation of P1 and pIP10 affects the probability of producing the three different song modes, we used csChrimson [24] to optogenetically stimulate each neuron type in solitary males. We chose an LED stimulation protocol in which we tested varying duty cycles (DCs) between 0.1 and 0.9 during three-second long trials, thereby leaving either very long pauses (DC=0.1 corresponds to a pause of 2700ms) or very short pauses (DC=0.9 corresponds to a pause of 300ms) between subsequent stimulations – in other words, short DCs correspond to weaker stimulation and long DCs to stronger stimulation. Consistent with the idea that P1 encodes female proximity, males produced mainly sine song during P1 activation, while both P_{fast} and P_{slow} were largely suppressed (Figure 4B_{1,2}, 4C_{1,2}). We observed this effect for two different genetic drivers that target P1 neurons: R71G01 [25] and a split GAL4 driver termed P1a [22] (Supplemental Figure 4B). Interestingly, for P1a, sine song was only dominant for DCs <0.8; at higher DCs, pulse song was more dominant, suggesting sine production requires a pause between stimulations (Supplemental Figure 4A₂, Figure 4C₂). P_{fast} and P_{slow} probabilities exhibited similar dynamics for both P1 drivers (Figure 4B_{1,2}, Supplemental Figure 4A_{1,2}) suggesting that P1 mainly biases the choice between sine and pulse song, while affecting the balance between P_{fast} and P_{slow} relatively little. Moreover, P1 activation did not affect the waveform shape of either P_{fast} and P_{slow} , only the probability of producing pulses (Supplemental Figure 4C, D).

pIP10 is a descending neuron thought to be postsynaptic to P1 [18]. However, we found that pIP10 does not simply relay P1 activity, since P_{fast} , and not sine song, was the dominant mode produced during optogenetic activation (Figure 4B₃, 4C₃). Sine became the dominant song mode only after stimulation ended and only if the pauses between stimulation were sufficiently long (DCs \leq 0.9). Consistent with these results, we also found that pIP10 inactivation (using either Kir2.1 [26] or TNT [27]) during courtship with a female decreased P_{fast} production (Figure 4E). We found that the probability of producing P_{slow} was largely independent of optogenetic stimulus duration, whereas the opposite was true for P_{fast} : for short DCs (<0.6) we observed high peak probabilities, intermediate DCs (0.6-0.8) produced an onset transient with reduced steady-state

levels, and at the longest DCs, P_{fast} levels were relatively constant throughout the stimulation period and resembled those of P_{slow} (Figure 4B₃, Supplemental Figure 4A₃).

Our results with P1 and pIP10 activation highlight the necessity of carefully delineating behavioral modes and reveal complex dynamics in the neural circuits driving song production. For activation of both cell types, for example, we observed that stimulus history (and thereby the history of song produced) has a strong impact on the song types elicited by optogenetic activation (Figure 4B, Supplemental Figure 4A). While optogenetic activation of both P1 and pIP10 affected male speed (Supplemental Figure 4A), we observed no fixed relationship between male speed and song choice among the three drivers examined here, suggesting that the song dynamics induced by optogenetic activation cannot be explained as an indirect effect of changes in male speed alone.

How does the activity of P1 and pIP10 neurons ultimately lead to changes in the choice of pulse types? While much of the song pathway has not yet been mapped [28], two sets of motor neurons have been shown to be involved in either sine or pulse song [8]. Inactivation of the ps1 motor neuron reduces both pulse amplitude and carrier frequency [8]. This effect would be consistent with ps1 being involved in P_{fast} production, since P_{fast} is the louder and faster of the two pulse types. Since activation of ps1 alone does not generate song, we instead silenced ps1 using tetanus toxin (TNT), and paired males with females to determine the effect of ps1 silencing on song production. We found that ps1 inactivation does not have a strong effect on the shape of either P_{slow} or P_{fast} (Figure 4H), and only a subtle effect on the frequency of P_{slow} , but not P_{fast} , pulses (Supplemental Figure 4E, F). Instead, silencing ps1 strongly affects the switching between P_{slow} and P_{fast} with distance (Figure 4F), and thereby reduces the amount of P_{fast} pulses being produced (Figure 4G): wild type flies bias away from P_{slow} with increasing distance but ps1 inactivated flies continue to sing P_{slow} even when far from the female. ps1 inactivation also reduces the amplitude of both pulse types (Figure 4H) [5]. This implies that although ps1 is not strictly necessary for P_{fast} production – ~10% of pulses are still of the P_{fast} type and these P_{fast} pulses are indistinguishable from wild

type P_{fast} pulses (Figure 4I, Supplemental Figure 4E, F) – this motor neuron contributes instead to the choice between pulse types, and thereby the overall structure of male song.

Discussion

A central aim of systems neuroscience is to combine neural activation, silencing, and recording techniques to causally link particular neurons and circuits to behavior and to show how activity patterns within these neurons relate to behavioral activity. New computational methods (e.g. [12,29,30]) are facilitating the automated detection and classification of behaviors to drive these studies forward, particularly in genetic model systems, such as worms, flies and mice [31,32]. But solving the neural basis for behavior requires the appropriate temporal and spatial parsing of behaviors in order to generate meaningful connections with the underlying neural and muscle activity patterns that drive them. Here we use a highly quantifiable and robust behavior – the courtship song of *Drosophila melanogaster* – to investigate this fundamental issue. We show here using new unsupervised classification methods that song comprises at least three, not just two, modes, and we show how this distinction affects our interpretation of the mechanisms underlying song production (Figure 3, 4), perception (Figure 3J) and evolution (Figure 2G-J).

Rather than being a reflex-like behavior, song production relies on the integration of multiple sensory cues [5,9] and discriminating between P_{slow} and P_{fast} now reveals a novel layer of song control. Previous studies had shown that, similar to humans, *Drosophila* males increase the amplitude of their acoustic communication signals based on a visual estimate of distance to the receiver [5]. Analyzing the effect of distance on both pulse types separately, we now find that visual information mediates this amplitude modulation using two separate control mechanisms: First, a digital/binary control system biases pulse production towards the louder P_{fast} when far from the female (Figure 3D). And second, an analog control system upregulates P_{fast} – but not P_{slow} – amplitude with distance (Figure 3G). The strong interaction between the carrier frequency of each song mode and its amplitude (Figure 3F) suggests that mechanical constraints may facilitate

the production of loud signals at higher carrier frequencies. These two modes of control are likely implemented through distinct circuits, and hence discriminating the two pulse types is important for understanding the neural basis of this behavior.

Cataloguing all song modes is also important for dissecting the neural circuits driving song production. Female proximity cues activate the song command neuron P1 [17,19] and using optogenetics, we have shown that activation of P1 biases solitary males towards singing more sine song (Figure 4B_{1,2}, C_{1,2}) – which is naturally produced when the male is closest to the female [9]. However, P1 activation had no strong effect on the choice between P_{slow} and P_{fast}, suggesting that pulse choice is implemented elsewhere in the circuit (Figure 4B_{1,2}). By contrast, pIP10 is thought to be downstream of P1 [18], but its activation strongly biases males toward producing P_{fast} instead (Figure 4B₃, C₃). Thus, rather than being a relay of P1 activation, pIP10 also actively shapes the dynamics of song mode choice (Figure 4B, Supplemental Figure 4A), but not the parameters of each song mode (Supplemental Figure 4C, D). This is an important distinction – these neurons appear to bias the output of the downstream pattern generating circuits without changing their control over the features of each song mode, such as the frequency or shape of individual pulses. Moreover, activation of P1 or pIP10 in solitary males produces all 3 song modes, but with different probabilities and latencies. This suggests that a static analysis of song production based on the activation or inactivation of song command neurons is insufficient to understand how song is produced in *Drosophila*. Rather, the activity dynamics of different elements of the song pathway shape what is sung and when.

ps1 was a motor neuron previously implicated in setting the frequency and amplitude of pulses [8]. By segmenting the behavior to identify multiple pulse types, we find instead that this motor neuron does not strongly affect the parameters of each song mode (Figure 4I) (similar to the results with P1 and pIP10 activation, Supplemental Figure 4C, D), but rather the choice of which pulse type to produce relative to the distance to the female (Figure 4F). This suggests that this neuron represents one of the ends of the pathway that connects visual information with pulse choice. Further studies into the

activity of command neurons, motor neurons, and muscles [8,18,33,34] during singing will shed additional light onto the circuits choosing and shaping the three song modes.

We also found that females slow differentially to P_{slow} and P_{fast} (Figure 3J, Supplemental Figure 3), and that this slowing depends on the context: females reduce their locomotor speed to P_{fast} only when it is sung at farther distances (Figure 3J). This demonstrates that distinguishing between pulse types matters for studying the perception of song as well. Future experiments that allow precise control over the sensory cues available to the female – e.g. through sound playback – are necessary to elucidate the sensory source of the differential slowing response and ultimately its neural basis.

While previous studies demonstrated the existence of multiple pulse types in *Drosophila* species [10,35-37] our statistical analysis of pulse shapes in *D. melanogaster*, *D. simulans* and *D. mauritiana* now shows that P_{fast} and P_{slow} shapes are conserved in these three species (Figure 2H-J, Supplemental Figure 2), implying that the existence of these two pulse types predates the species split. Interestingly, *D. yakuba* – a member of the melanogaster group outside of the branch considered here (Figure 2G) – produces two pulse types termed “thud” and “clack” that are claimed to be produced at different distances to the female {Demetriades:1999iw}, just as with P_{slow} and P_{fast} in *D. melanogaster*. It is conceivable that the thud and clack pulses in *D. yakuba* are homologous to P_{slow} and P_{fast} but further studies are required. It is also not yet clear whether distance to the female is the main driver of pulse choice in *D. simulans* and *D. mauritiana*.

In summary, we show that segmenting *Drosophila melanogaster* song into three, not two, modes affects our interpretation of the sensory, neural, and evolutionary effects on song patterning. This study highlights that a detailed characterization of behavior is required to correctly interpret data from neural activation and silencing experiments, recordings of neural activity, and comparisons of behavior across species.

Author Contributions: JC, PC, and MM designed the study. JC, PC, FR, DM, TP, and DP collected data. JC, PC, FR, and TP analyzed data. JC and MM wrote the manuscript.

Acknowledgements: We thank Asif Ghazanfar, Josh Shaevitz, and David Stern for comments on the manuscript, and Georgia Guan for technical assistance. JC was supported by a postdoctoral fellowship through the Princeton Sloan-Swartz Center, and MM was supported by an NSF BRAIN EAGER award, an NIH New Innovator Award, a McKnight Scholar Award, and a Faculty Scholar Award from the Howard Hughes Medical Institute.

References

- [1] Krakauer JW, Ghazanfar AA, Gomez-Marín A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron* 2017;93:480–90. doi:10.1016/j.neuron.2016.12.041.
- [2] Bennet-Clark HC, Ewing AW. Pulse interval as a critical parameter in the courtship song of *Drosophila melanogaster*. *Animal Behaviour* 1969;17:755–9. doi:10.1016/S0003-3472(69)80023-0.
- [3] Shorey HH. Nature of the Sound Produced by *Drosophila melanogaster* during Courtship. *Science* 1962;137:677–8. doi:10.1126/science.137.3531.677.
- [4] Bennet-Clark HC, Ewing AW. Stimuli provided by Courtship of Male *Drosophila melanogaster*. *Nature* 1967;215:669–71. doi:10.1038/215669a0.
- [5] Coen P, Xie M, Clemens J, Murthy M. Sensorimotor Transformations Underlying Variability in Song Intensity during *Drosophila* Courtship. *Neuron* 2016;89:629–44. doi:10.1016/j.neuron.2015.12.035.
- [6] Clemens J, Girardin CC, Coen P, Guan X-J, Dickson BJ, Murthy M. Connecting Neural Codes with Behavior in the Auditory System of *Drosophila*. *Neuron* 2015;87:1332–43. doi:10.1016/j.neuron.2015.08.014.
- [7] Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 2016. doi:10.1038/nature19093.
- [8] Shirangi TR, Stern DL, Truman JW. Motor control of *Drosophila* courtship song. *Cell Reports* 2013;5:678–86. doi:10.1016/j.celrep.2013.09.039.
- [9] Coen P, Clemens J, Weinstein AJ, Pacheco DA, Deng Y, Murthy M. Dynamic sensory cues shape song structure in *Drosophila*. *Nature* 2014;507:233–7. doi:10.1038/nature13131.
- [10] Riabinina O, Dai M, Duke T, Albert JT. Active process mediates species-specific tuning of *Drosophila* ears. *Current Biology : CB* 2011;21:658–64. doi:10.1016/j.cub.2011.03.001.
- [11] Maaten LVD, Hinton G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 2008;9:2579–605.
- [12] Berman GJ, Choi DM, Bialek W, Shaevitz JW. Mapping the stereotyped behaviour of freely moving fruit flies. *J R Soc Interface* 2014;11:20140672–2. doi:10.1098/rsif.2014.0672.
- [13] Arthur BJ, Sunayama-Morita T, Coen P, Murthy M, Stern DL. Multi-channel acoustic recording and automated analysis of *Drosophila* courtship songs. *BMC Biol* 2013;11:11. doi:10.1186/1741-7007-11-11.
- [14] Mineault PJ, Barthelmé S, Pack CC. Improved classification images with sparse priors in a smooth basis. *Journal of Vision* 2009;9:1–24.
- [15] Klibaite U, Berman GJ, Cande J, Stern DL, Shaevitz JW. An unsupervised method for quantifying the behavior of paired animals. *Phys Biol* 2017;14:015006. doi:10.1088/1478-3975/aa5c50.
- [16] Kimura K-I, Kimura K-I, Hachiya T, Hachiya T, Koganezawa M, Tazawa T, et al. Fruitless and doublesex coordinate to generate male-specific neurons that can initiate courtship. *Neuron* 2008;59:759–69. doi:10.1016/j.neuron.2008.06.007.
- [17] Kohatsu S, Koganezawa M, Yamamoto D. Female contact activates male-specific interneurons that trigger stereotypic courtship behavior in *Drosophila*.

- Neuron 2011;69:498–508. doi:10.1016/j.neuron.2010.12.017.
- [18] Philipsborn von AC, Liu T, Yu JY, Masser C, Bidaye SS, Dickson BJ. Neuronal control of *Drosophila* courtship song. Neuron 2011;69:509–22. doi:10.1016/j.neuron.2011.01.011.
- [19] Clowney EJ, Iguchi S, Bussell JJ, Scheer E, Ruta V. Multimodal Chemosensory Circuits Controlling Male Courtship in *Drosophila*. Neuron 2015. doi:10.1016/j.neuron.2015.07.025.
- [20] Kohatsu S, Yamamoto D. Visually induced initiation of *Drosophila* innate courtship-like following pursuit is mediated by central excitatory state. Nat Commun 2015;6:6457. doi:10.1038/ncomms7457.
- [21] Kallman BR, Kim H, Scott K, Ramaswami M. Excitation and inhibition onto central courtship neurons biases *Drosophila* mate choice. eLife 2015;4:e11188. doi:10.7554/eLife.11188.
- [22] Hoopfer ED, Jung Y, Inagaki HK, Rubin GM, Anderson DJ, Ramaswami M. P1 interneurons promote a persistent internal state that enhances inter-male aggression in *Drosophila*. eLife 2016;4:e11346. doi:10.7554/eLife.11346.
- [23] Zhang SX, Rogulja D, Crickmore MA. Dopaminergic Circuitry Underlying Mating Drive. Neuron 2016;0. doi:10.1016/j.neuron.2016.05.020.
- [24] Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, et al. Independent optical excitation of distinct neural populations. Nat Methods 2014;11:338–46. doi:10.1038/nmeth.2836.
- [25] Pan Y, Meissner GW, Baker BS. Joint control of *Drosophila* male courtship behavior by motion cues and activation of male-specific P1 neurons. Proc Natl Acad Sci U S A 2012;109:10065–70. doi:10.1073/pnas.1207107109.
- [26] Baines RA, Uhler JP, Thompson A, Sweeney ST, Bate M. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. Journal of Neuroscience 2001;21:1523–31.
- [27] Sweeney ST, Broadie K, Keane J, Niemann H, O’Kane CJ. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. Neuron 1995;14:341–51. doi:10.1016/0896-6273(95)90290-2.
- [28] Coen P, Murthy M. Singing on the fly: sensorimotor integration and acoustic communication in *Drosophila*. Current Opinion in Neurobiology 2016;38:38–45.
- [29] Wiltschko AB, Johnson MJ, Iurilli G, Peterson RE, Katon JM, Pashkovski SL, et al. Mapping Sub-Second Structure in Mouse Behavior. Neuron 2015;88:1121–35. doi:10.1016/j.neuron.2015.11.031.
- [30] Robie AA, Hirokawa J, Edwards AW, Umayam LA, Lee A, Phillips ML, et al. Mapping the Neural Substrates of Behavior. Cell 2017;170:393–406.e28. doi:10.1016/j.cell.2017.06.032.
- [31] Calhoun AJ, Murthy M. Quantifying behavior to solve sensorimotor transformations: advances from worms and flies. Current Opinion in Neurobiology 2017;46:90–8. doi:10.1016/j.conb.2017.08.006.
- [32] Anderson DJ, Perona P. Toward a Science of Computational Ethology. Neuron 2014;84:18–31. doi:10.1016/j.neuron.2014.09.005.
- [33] Shirangi TR, Wong AM, Truman JW, Stern DL. Doublesex Regulates the Connectivity of a Neural Circuit Controlling *Drosophila* Male Courtship Song.

- 461 Developmental Cell 2016;37:533–44. doi:10.1016/j.devcel.2016.05.012.
 462 [34] Lindsay T, Sustar A, Dickinson M. The Function and Organization of the Motor
 463 System Controlling Flight Maneuvers in Flies. Current Biology 2017;0.
 464 doi:10.1016/j.cub.2016.12.018.
 465 [35] Tomaru M, Yamada H. Courtship of *Drosophila*, with a special interest in
 466 courtship song 2011;69:61–85.
 467 [36] Cowling D, Burnet B. Courtship songs and genetic control of their acoustic
 468 characteristics in sibling species of the *Drosophila melanogaster* subgroup.
 469 Animal Behaviour 1981;924–35.
 470 [37] Demetriades MC, Thackeray JR, Kyriacou CP. Courtship song rhythms in
 471 *Drosophila yakuba*. Animal Behaviour 1999;57:379–86.
 472 doi:10.1006/anbe.1998.0976.
 473

CONTACT FOR REAGENTS AND RESOURCE SHARING

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

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

If not stated otherwise, virgin male and female flies were isolated within 6 hours of eclosion and aged for 3-7 days prior to experiments. Flies were raised on a 12:12 dark:light cycle, at 25°C and 60% humidity.

Figure	Genotypes
1, 2A, C, E, F, S1A, C, S2C	<i>D. melanogaster</i> NM91 (provided by Peter Andolfatto) *, **
S1B, D 2C, E, F, H-J, S2, 3A-D, F-J, S3	8 <i>D. melanogaster</i> strains CM07, CarM03, N30, NM91, TZ58, ZH23, ZW109 (these 7 strains were provided by Peter Andolfatto), and Canton S (lab stock) *, **
2D	P1-TrpA1: UAS>stop>TrpA1/+; fru ^{FLP} /NP2631 (fru ^{FLP} provided by Barry Dickson; NP2631 was obtained from the Kyoto stock center)
3E	Canton S (WT1), GMR-hid/GMR-hid;+ (blind), +;orco-/orco- (Pheromone Insensitive or PI) in a Canton S background, NM91 (WT2), arista cut 20h prior to experiment (deaf) in an NM91 background, *, **
2 H-J, S1B, S2A,B	<i>D. simulans</i> (sim194, sim195, simNS05) and <i>D. mauritiana</i> strains (mau317, mauR12, mauR61) (provided by P. Andolfatto)
4B-D, S4A-C	P1 (R71G01): w/+; GMR71G01-LexA/+; LexAop-CsChrimson/+ (GMR71G01-LexA and LexAop-CsChrimson were obtained from the Bloomington stock center) P1a: UAS-CsChrimson/+; GMR15A01-AD (attP40)/BRP>stop>-V5-2A-LexA-VP16, LexAop-myr-tdtomato; GMR71G01-DBD (attP2)/LexAop-GCaMP6s (GMR15A01-AD (attP40); GMR71G01-DBD (attP2) provided by David Anderson; UAS-CsChrimson and LexAop-GCaMP6s were obtained from the Bloomington stock center) pIP10: w/+; UAS>stop>CsChrimson/+; VT40556, Fru ^{FLP} /+ (VT40556 and Fru ^{FLP} provided by Barry Dickson; UAS>stop>CsChrimson was obtained from Vivek Jayaraman) The '+' chromosomes above come from NM91.
4E	pIP10-control: +/+; fru ^{FLP} /VT40556 control-Kir: R53G02AD attP40/UAS-Kir2.1 control-TNT: R53G02AD attP40/UAS-TNT pIP10-Kir: aus>stop>Kir2.1/+; fru ^{FLP} /VT40556** pIP10-TNT: UAS>stop>TNT/+; fru ^{FLP} /VT40556**
4F-I, S4C-D	R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4/UAS>stop>TNT (ps1>TNT R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4/UAS>stop>TNTinactive (ps1>control) (R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4 was provided by Troy Shirangi; UAS>stop>TNT and UAS>stop>TNTinactive were obtained from Bloomington)

* data previously published in [1]

**data previously published in [2]

METHOD DETAILS

Behavioral chamber

Behavioral chambers were constructed as previously described [1-3]. For optogenetic activation experiments we used a modified chamber whose floor was lined with white plastic mesh and equipped with 16 recording microphones. To prevent the LED light from interfering with the video recording and tracking, we used a short-pass filter (Thorlabs FESH0550, cut-off wavelength: 550 nm). For the analysis of fly pose during the production of all three song modes we recorded video at 100 frames per second (instead of 60 frames per second as in all other recordings) and designed a novel chamber with a domed lid and a flat floor that was lined with fine nylon mesh (Ted Pella Nylon 300 Mesh) instead of the coarse mesh used for the other recordings. This created a clean background and allowed us to resolve fine details of the fly body like the wing. When recording with a female, males were painted on the thorax with a white dot 20h prior to experiment under cold anesthesia for identification of sex during tracking. Flies were introduced gently into the chamber using an aspirator. Recordings were timed to be within 150 minutes of the behavioral incubator lights switching on to catch the morning activity peak. Recordings were stopped after 30 minutes or earlier if copulation occurred. If males did not sing in the first 5 minutes of the recording, the experiment was discarded.

Inactivation of song command and motor neurons

Tetanus toxin (TNT) was used for chronic inactivation of the motorneuron driving the direct wing muscle ps1 [4], with an inactivated TNT (TNT-in) as a negative control. Tetanus toxin (TNT) and the inward-rectifying Potassium channel $K_{ir} 2.1$ were used to inactivate the song command neuron pLP10 during courtship [5].

Optogenetic activation

Flies were kept for at least 3 days prior to the experiment on fly food supplanted with retinal (1ml all-trans retinal solution (100mM in 95% ethanol) per 100ml food). CsChrimson [6] was activated using a 627 nm LED (Luxeon Star) at an intensity of 0.46 mW/mm² (driving voltage 3.2V). Stimulus period 3sec, 9 duty cycles (DCs) ranging between 0.1 and 0.9, filling between 300 and 2700ms of the period with constant LED illumination. Each DC was presented in randomized blocks of 90s (30 trials each). Since optogenetic activation experiments were performed using solitary males, recordings were not timed to peaks in circadian activity. Sound recording and video were synchronized by positioning a red LED that blinked with a predetermined temporal pattern in the field-of-view of the camera and whose driving voltage was recorded alongside the song.

QUANTIFICATION AND STATISTICAL ANALYSIS

Song segmentation and tracking

Both tracking of flies and segmentation of song recordings were performed as previously described [1-3]. Pulses were detected with high detection rates (true positive rates 78%, 74%, 85% for *D. melanogaster*, *simulans* and *mauritiana*) and few false detections (false positive rate ~2% for *D. melanogaster* and *D. simulans*, 6% for *D. mauritiana* because some pairs produced substantial movement noise during courtship) (Supplemental Figure 1B). The song segmenter (with the parameters used in [1], not those of [7] - see [8] for a comparison of segmenter performance with these different parameters) for *D. melanogaster* detected P_{fast} and P_{slow} with similar high rates without requiring any modifications (77% and 80%, respectively). Note that we extended the song segmenter to now also classify pulses into P_{fast} and P_{slow} (available at <https://github.com/murthylab/songSegmenter>, see below). For segmenting the pulses produced by *D. simulans* and *D. mauritiana* we modified the segmenter by building strain specific pulse models from sets of manually annotated data (~500 pulses per strain).

Waveform selection for segmented pulses

For *D. melanogaster*, pulse waveforms were extracted from the recordings by taking 25.1 ms (251 samples at 10kHz) around the pulse center detected by the song segmenter from the channel on which the pulse was recorded with the highest energy. For the much shorter pulses produced by *D. simulans* and *D. mauritiana*, 25.1 ms around the pulse peak contained mostly noise. We therefore upsampled the waveforms to 20kHz and extracted the same 251 samples - now corresponding to 12.55 ms - around the pulse center.

Waveform selection for unbiased analysis of song waveforms

For the unbiased analysis of song structure, we selected waveforms from the raw recordings without using the song segmenter. As a first step, all 9 recording channels were merged by choosing the signal from the channel with the highest absolute value in 5ms windows. To avoid our analysis being drowned out by background noise (up to 90% of a song recording can contain no song), we chose only signals that exceeded the energy of the recording noise. Sound energy was calculated by estimating the signal envelope using the Hilbert transform, transforming it to a logarithmic scale and then smoothing it using a sliding window lasting 100ms. The logarithmic transformation produced an energy distribution that resembled a mixture of two Gaussians, one corresponding to baseline noise and one corresponding to song signals. We fitted a Gaussian mixture model with two components to the distribution energy values and identified as “signal” the component with the higher mean energy. From all time points classified as signal we then extracted 25ms long, non-overlapping waveforms. Note that the resulting data set contained many snippets that contained only noise if they were at the very beginning of or between song bouts.

Waveform normalization

To remove variability in the waveforms arising from the position and distance of the singing male from the microphone, we

1. **Divide** the raw waveforms, $x(t)$, by their norm: $x(t)/\sqrt{\sum x(t)^2}$.
2. **Center** to their peak energy obtained by smoothing the square pulse waveform with a rectangular window spanning 15 samples (1.50 ms for *D. melanogaster*, 0.75 ms for *D. simulans* and *D. mauritiana*).
3. **Flip** sign such that average pulse in the 10 samples (1.00 ms for *D. melanogaster*, 0.5 ms for *D. simulans* and *D. mauritiana*) preceding the pulse center is positive.

Time scales were shortened for *D. simulans* and *D. mauritiana* to enable robust alignment of the much shorter pulses produced by these species (see e.g. Figure 2H). The last step – adjusting waveform sign – does not artefactually create clustering into two pulse types but rather reduces the degeneracy of the waveform space since the same pulse can occur with different signs on different microphones, indicating that the position of the male wing relative the microphone (in front of or behind) determines waveform sign. Clustering waveforms with randomized sign yields clustering in which each pulse type appears in two clusters – one in its original shape and one with the sign inverted (Supplemental Figure 2C).

Dimensionality reduction

The dimensionality of the normalized waveforms was reduced from 251 data points per waveform to two dimensions using two different methods:

1. **T-distributed stochastic neighbor embedding (tSNE)** preserves local neighborhood structure while ignoring global similarity between data points. It is therefore ideally suited for cluster analysis, which relies on local similarity. To embed large data sets (more than 100000 pulses) we chose the computationally and memory-efficient Barnes-Hut implementation [9] (parameters: initial PCA dimensions=30, perplexity=50, accuracy=0.1). The shape of the embedded distribution was highly reproducible across runs and did not depend critically on the choice of parameter values. Since large perplexity values increased the separation between clusters - while preserving the overall shape of the embedded distribution - we chose a perplexity value of 50.
2. **Principal component analysis (PCA)** is a linear method for dimensionality reduction and works by finding a small set of orthogonal basis functions in which the global data variance is maximized. Clustering using a PCA-based representation of the waveforms yields qualitatively similar results (Figure 2C), indicating that our findings do not critically depend on the nonlinear tSNE embedding. However, the cluster separation corresponding to P_{slow} and P_{fast} was

not as strong and hence clustering was not as reproducible as with tSNE. Consequently, we chose to use tSNE embedding for all analyses.

Waveform clustering

To identify groups of similar pulses, we clustered the waveforms in two steps:

1. The 2D distribution of data was partitioned using the watershed algorithm, which cuts a distribution along local minima. To that end we estimated the distribution of the tSNE-embedded pulses using kernel density estimation (matlab function `mvksdensity`), with the bandwidth chosen automatically by the algorithm (Figure 1C). This resulted in seven clusters (thin lines in Figure 1C and cluster waveforms in Figure 1D), some of which exhibited highly similar mean waveforms (Figure 1E) indicating that this first step overpartitioned the waveform space.
2. We therefore consolidated these watershed clusters using hierarchical clustering of the cluster centroids (matlab function `clusterdata`). When clustering only the pulses detected by the song segmenter (Figure 2A), the resulting cluster trees always revealed a simple bipartite structure and we hence chose a clustering cutoff that produced two modes. For the unbiased analysis of song waveforms (Figure 1D), the cluster tree revealed four principal waveform shapes corresponding to noise, sine song and the two pulse types. Cluster consolidation
 - a. joins three sine clusters, which is supported by them containing all similar waveforms (Figure 1E, left)
 - b. joins two P_{slow} clusters but leaves the P_{fast} separate, which is again supported by the waveforms in the P_{slow} clusters being highly similar and those in the P_{fast} cluster being highly dissimilar to those of P_{slow} .

Note that there exists a myriad of approaches for clustering data. Watershed plus consolidation using hierarchical clustering on the centroids produced the most robust clustering. Direct agglomerative clustering of the tSNE embedded data – without first using the watershed cluster step – also results in four clusters (not shown), but these clusters are not as “clean” – for instance, the P_{fast} mode extends into the noise cluster. Direct clustering on the unbiased set of waveforms yields unreliable results, likely because of the large variability of waveforms in this data set (it includes noise).

The full analysis pipeline yielded near-identical results across different runs - individual pulses were almost always assigned to the same cluster with very high probability (adjusted Rand index ~ 0.97). Code illustrating all steps of the pipeline is available at <https://murthylab.github.io/pulseTypePipeline/>.

Template-based pulse type classifier

Normalized pulses from the *D. melanogaster* wild type strain NM91 were projected onto the centroids of the P_{slow} and P_{fast} clusters and a quadratic boundary separating both pulse types was obtained using quadratic discriminant analysis cross-validated using an 80:20 partition into training and test data (Supplemental Figure 1C). The classifier performs well on the test data and generalizes well to different *D. melanogaster* strains (Supplemental Figure 1D). It is therefore a fast and robust method for classifying pulses that does not require the computationally extensive pipeline outlined above. Code for building and evaluating the classifier is available at: <https://github.com/murthylab/pulseTypeClassifier>. The classifier has also been integrated into the song segmenter software (available at <https://github.com/murthylab/songSegmenter>).

Characterization of pulse shapes

Song pulses are transient signals and their raw magnitude spectra are relatively broad (Figure 2B, Supplemental Figure 4E). Accordingly, the *peak* frequency of the spectrum is an unreliable measure of pulse carrier frequency. We hence used the center of mass of the magnitude spectra thresholded at $e^{-1}=0.37$ as a measure of frequency. The values obtained tended to closely resemble 1) the spectral peak frequencies in cases of sufficiently peaked spectra (e.g. for most instances of P_{slow}) and 2) the carrier frequency values extracted from fits of Gabor functions to the pulse waveforms (not shown), demonstrating that the center of mass faithfully extracts the pulse carrier frequency. A pulse symmetry index was calculated as the dot product of the first half of the pulse and the flipped second half of the pulse. Positive indices correspond to even symmetry and negative indices to odd symmetry. Pulse amplitudes were normalized for the gain of individual microphones as described in [2].

Identification of sensory cues driving pulse choice

The GLM analysis for identifying the song features driving the choice between P_{slow} and P_{fast} was performed as in [2].

Amplitude modulation plots

We evaluated the relation between pulse amplitude and distance to the pulse at a delay of 470ms since the distance at this delay is most predictive of pulse amplitude [2]. Distance was binned between 1.75 and 10.25 mm in 0.5mm steps and the mean pulse amplitude (and its standard error) were calculated for each. The first three pulses of each train were excluded from the analysis since flies do not modulate pulse amplitude at pulse train start [2]. Distance bins with <50 pulses across all flies were excluded before calculating r^2 values.

Analysis of optogenetic activation data

The song probability traces (Figure 4B, Supplemental Figure 4A) were constructed by calculating the fraction of trials (for each time bin, pooled across all flies) during which a male produced sine, P_{slow} or P_{fast} . Traces do not add up to 1.0 since there are trials during which males did not sing any song. Song probability traces and male speed were down sampled to 400Hz and smoothed with a sliding Gaussian kernel lasting 150ms with standard deviation $\sigma=64\text{ms}$.

DATA AND SOFTWARE AVAILABILITY

Song segmenter with pulse classifier: <https://github.com/murthylab/songSegmenter>.

Stand-alone pulse classifier: <https://github.com/murthylab/pulseTypeClassifier>.

Script illustrating the analysis pipeline: <https://murthylab.github.io/pulseTypePipeline/>.

Raw data is available upon reasonable request from the Lead Contact, Mala Murthy (mmurthy@princeton.edu)

- [1] Coen P, Clemens J, Weinstein AJ, Pacheco DA, Deng Y, Murthy M. Dynamic sensory cues shape song structure in *Drosophila*. *Nature* 2014;507:233–7. doi:10.1038/nature13131.
- [2] Coen P, Xie M, Clemens J, Murthy M. Sensorimotor Transformations Underlying Variability in Song Intensity during *Drosophila* Courtship. *Neuron* 2016;89:629–44. doi:10.1016/j.neuron.2015.12.035.
- [3] Deng Y, Coen P, Sun M, Shaevez JW. Efficient Multiple Object Tracking Using Mutually Repulsive Active Membranes. *PLoS ONE* 2013;8:e65769. doi:10.1371/journal.pone.0065769.s004.
- [4] Shirangi TR, Stern DL, Truman JW. Motor control of *Drosophila* courtship song. *Cell Reports* 2013;5:678–86. doi:10.1016/j.celrep.2013.09.039.
- [5] Philipsborn von AC, Liu T, Yu JY, Masser C, Bidaye SS, Dickson BJ. Neuronal control of *Drosophila* courtship song. *Neuron* 2011;69:509–22. doi:10.1016/j.neuron.2011.01.011.
- [6] Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, et al. Independent optical excitation of distinct neural populations. *Nat Methods* 2014;11:338–46. doi:10.1038/nmeth.2836.
- [7] Arthur BJ, Sunayama-Morita T, Coen P, Murthy M, Stern DL. Multi-channel acoustic recording and automated analysis of *Drosophila* courtship songs. *BMC Biol* 2013;11:11. doi:10.1186/1741-7007-11-11.
- [8] Stern DL, Clemens J, Coen P, Calhoun AJ, Hogenesch JB, Arthur BJ, et al. Experimental and statistical reevaluation provides no evidence for *Drosophila* courtship song rhythms. *Proceedings of the National Academy of Sciences of the United States of America* 2017;89:201707471–9983.

doi:10.1073/pnas.1707471114.

- [9] van der Maaten L. Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research 2014;15:3221–45.

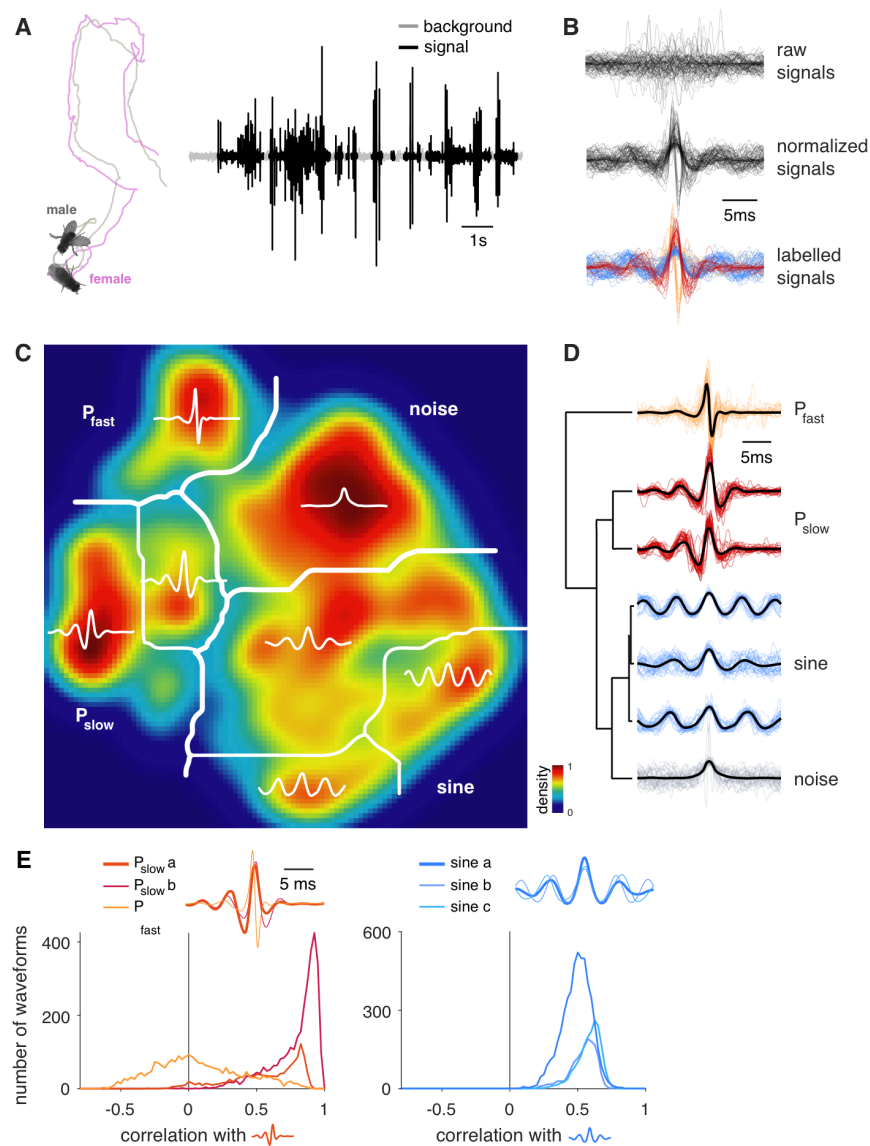


Figure 1: *Drosophila melanogaster* courtship song comprises three, not two, modes

A During courtship, the male chases the female and produces a courtship song through unilateral wing vibrations. Shown are male (gray) and female (magenta) traces during courtship. Song recording (right) with background noise (grey) and signals (black) from a wild type *D. melanogaster* male.

B Non-overlapping signal chunks (25ms duration) were extracted (top) and then normalized (middle) to scale all signals to the same peak amplitude and to flip the sign of signal waveforms such that the largest peak prior to the pulse center is always positive. Variation in signal amplitude or sign could be due to the male's position relative to the microphone. Here we show a random subset of 86 out of all 21,104 chunks. After clustering, signals produced by males during courtship fall into four classes: non-song noises (grey), sine (blue), P_{slow} (red) and P_{fast} (orange).

C We reduced the dimensionality of 21,104 waveforms from 47 individual males of the *D. melanogaster* strain NM91 using t-distributed stochastic neighbor embedding (tSNE) and the resulting two-dimensional distribution (density color coded, see colorbar) of signals was partitioned using the watershed algorithm. This yielded seven clusters, which upon inspection of the waveforms, correspond to noise, sine song (three clusters) and two distinct pulse song modes – P_{slow} (two clusters) and P_{fast} (see text labels). Thick white lines mark the main mode boundaries after cluster consolidation (see D) and thin white lines mark the submodes within each main mode.

D Hierarchical clustering of the average waveform for each watershed cluster (centroids, thick black lines) supports the grouping of clusters into 4 main modes. Horizontal distance in the tree (left) corresponds to the dissimilarity between cluster centroids. Thin colored lines (right) show individual waveforms for each cluster. The waveforms corresponding to non-song noises (grey) are highly heterogeneous and have in common only the peak in energy to which they were aligned. Three sine song clusters (blue) mainly differ in frequency and constitute a continuum of waveforms. Note the second sine cluster is heterogeneous. Two P_{slow} clusters (red) differ only in their asymmetry, with more weight on either the negative lobe leading (bottom) or lagging (top) relative to the main positive peak. P_{fast} (orange) joins all fast and biphasic pulses.

E Similarity between all pulse waveforms (left) or all sine waveforms (right) obtained from the watershed clustering (Figure 1C). Similarity was computed by correlating all pulse or sine waveforms with the centroids for one of the clusters (the waveform used for correlation is thicker in the inset). Distributions do not strongly depend on which of the clusters was chosen as a template for correlation. The distribution for the pulse clusters is bimodal: the two P_{slow} clusters strongly overlap and have high correlation values with the centroid, indicating that they belong to a single song mode. The P_{fast} pulses exhibit projection values onto the P_{slow} centroid symmetrically distributed around 0, indicating very low similarity with the P_{slow} pulses and supporting its classification into a distinct pulse song mode. By contrast, the distribution for all three sine song clusters (right) is skewed towards high values and all three distributions overlap strongly, indicating that sine song constitutes a single song mode.

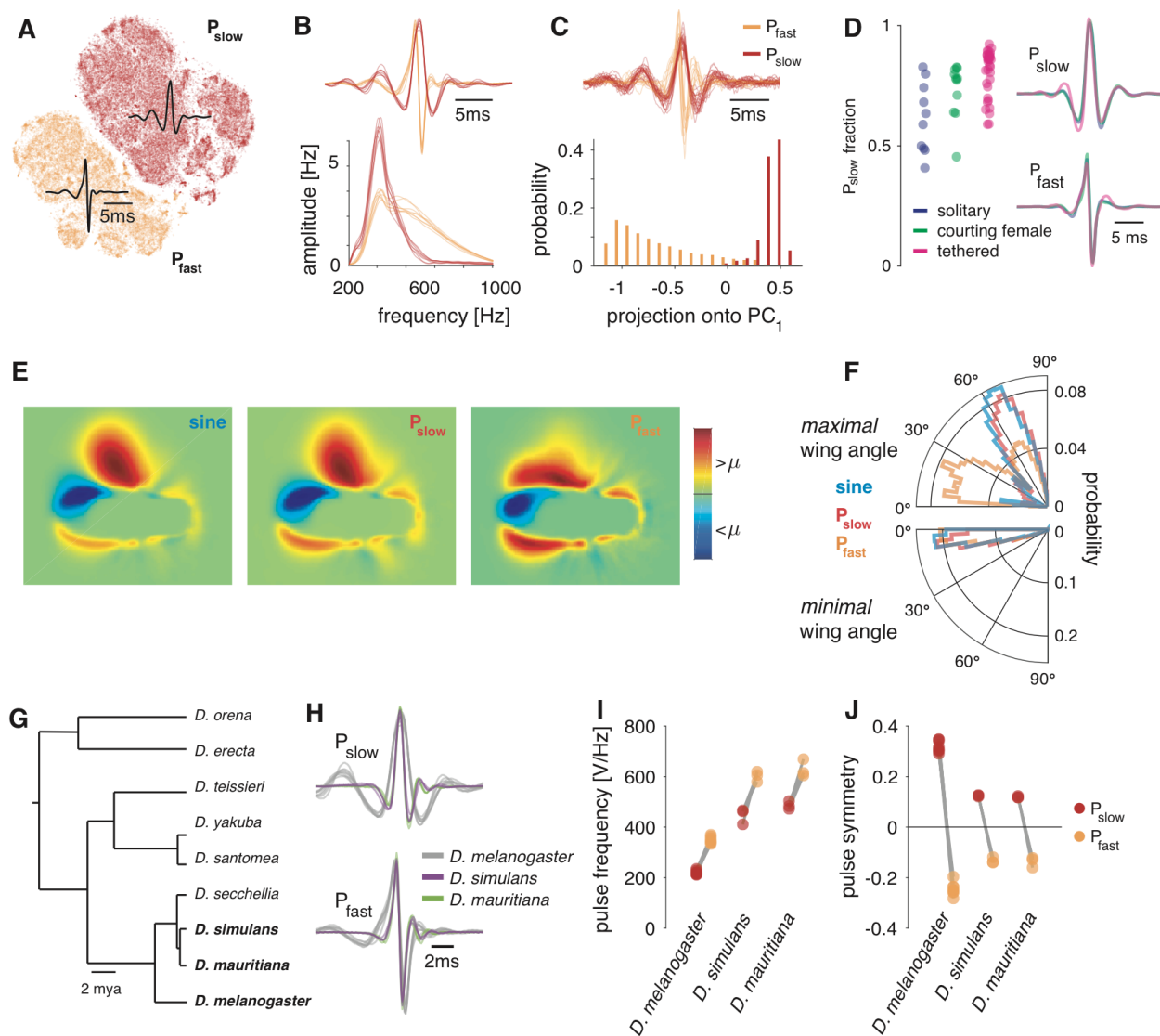


Figure 2: Song pulses can be separated into P_{fast} and P_{slow} across strains and species, and each corresponds to a distinct wing pose

A Two-dimensional tSNE of all pulses detected by the automated segmenter reveals a clear separation of the detected pulses into P_{slow} (red) and P_{fast} (orange). Shown are 71,029 pulses from 47 males of the *D. melanogaster* strain NM91. Black waveforms correspond to the average waveform of all pulses classified as P_{slow} (top) or P_{fast} (bottom), respectively.

B Average waveforms (top) and spectra (bottom) of P_{fast} and P_{slow} for all eight *D. melanogaster* strains investigated (see Methods for strain names).

C The P_{fast} versus P_{slow} distinction is evident upon inspection of individual pulses (top). Shown are 32 randomly selected, normalized pulses from NM91 males. Both pulse types are also separable using a linear dimensionality reduction technique – PCA (bottom). Plotted is the distribution of projection values of all 71,029 NM91 pulses onto the first principal component. Pulses (top) and projection values (bottom) are colored by the pulse type they were classified as (see legend).

D Pulses driven by thermogenetic activation of P1 neurons [1] in either solitary males (purple), males courting a female (green), or males tethered and walking on a ball (pink). Flies experienced open-loop visual stimulation matching the natural statistics of female motion on the male retina during courtship [2]. In all three conditions, both pulse types occur though with varying probability (left). Pulse shapes for P_{slow} (top right) and P_{fast} (bottom right) are nearly indistinguishable for all three conditions.

E Pose of flies during the production of sine, P_{slow} and P_{fast} . Shown is the average, mean subtracted frame at the time of the peak amplitude of sine and pulses, respectively (100Hz frame rate, 25 px/mm). Frames were oriented such that the male faces rightward, and flipped such that maximally extended wing is up. Both sine (left) and P_{slow} (middle) are produced with a strongly extended wing while P_{fast} (right) is produced with the wing held closer to the body.

F Polar histograms of the angle of the maximally (top) and minimally (bottom) extended wing during sine (blue), P_{slow} (red) and P_{fast} (orange) production. Distribution of the maximal wing angles for sine and P_{slow} are similar (50-70°). The distribution of P_{fast} angles has a main mode at 5-30° and a smaller second mode overlapping with the angles observed for sine and P_{slow} . Minimal wing angles (bottom) overlap for all three song modes. *Drosophila melanogaster* song is thought to be produced by unilateral wing vibration – the minimally extended wing is therefore likely silent. Data in E, F: A sine, B P_{slow} and C P_{fast} events for from 34 individuals of the *D. melanogaster* strain NM91.

G Phylogenetic tree of the *melanogaster* species subgroup (reproduced from [3]). The species investigated here are highlighted in bold font.

H Average pulse waveforms for all strains analyzed, colored by species (see legend). Each strain of each species produces a slower symmetrical (P_{slow} , top) and a faster asymmetrical pulse (P_{fast} , bottom) type.

I, J Pulse frequency (I) and symmetry (J) for the pulse waveforms in H by pulse type. *D. simulans* and *D. mauritiana* pulses are faster than those of *D. melanogaster*. For N flies and pulses see Supplemental Figure 2.

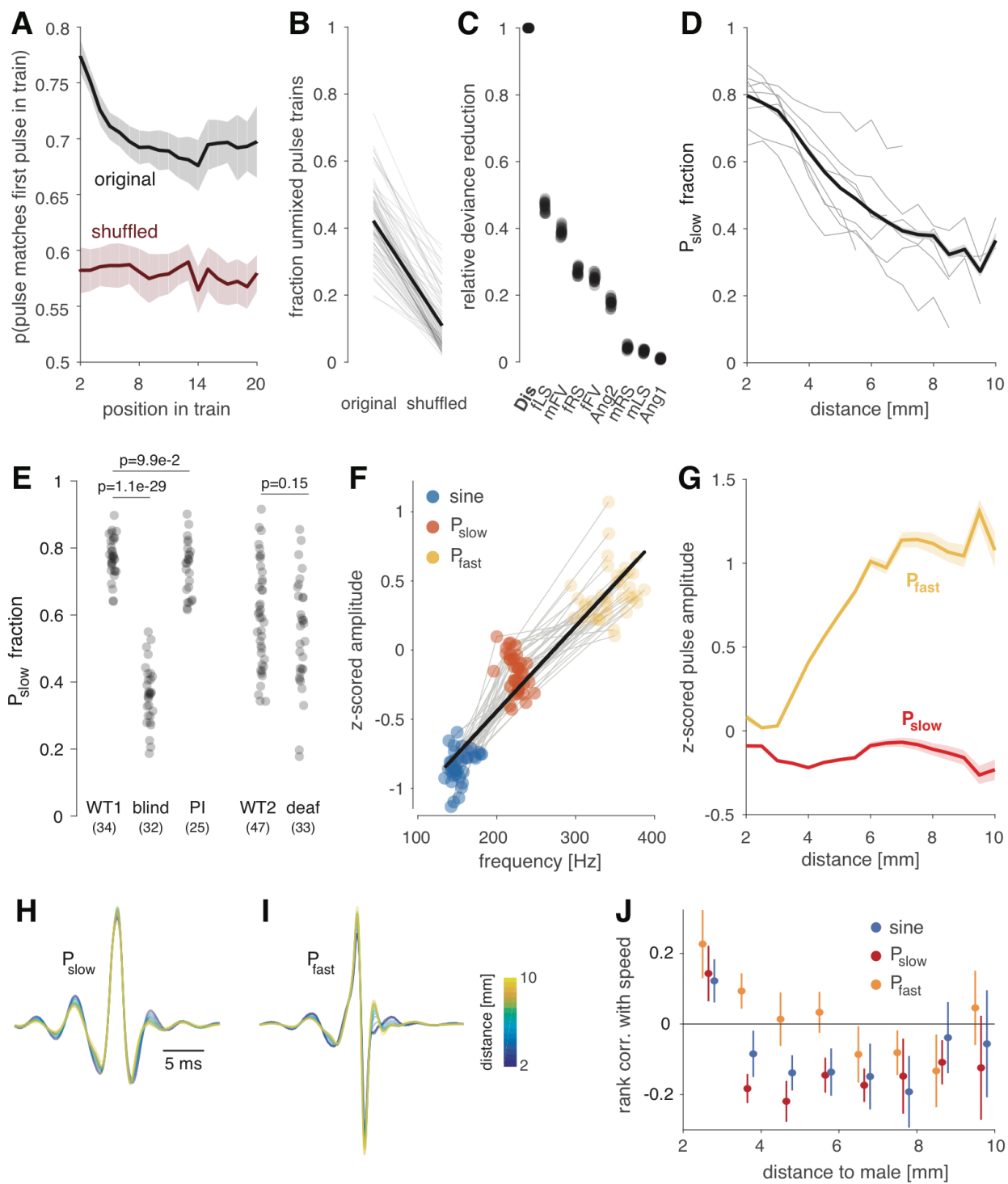


Figure 3: The choice between P_{slow} and P_{fast} is driven by sensory feedback and this choice impacts female responses

A When starting a pulse train with one pulse type, males continue to sing that type more than expected by chance. To estimate chance level, the sequence of pulse types was randomly permuted (shuffled) for each fly. Mean \pm SEM across 8 *D. melanogaster* strains. N=315 flies.

B The persistence observed in (A) leads to more pulse trains containing only one pulse type than would be expected by chance (original: 0.42 ± 0.11 , shuffled: 0.10 ± 0.07 , $p = 1.7 \times 10^{-29}$, rank sum test, N=315 flies). Thin lines show the probability, for several individuals, of observing unmixed pulse trains with at least 3 pulses using original (left) and shuffled (right) pulse labels. The thick black line corresponds to the average over all flies.

C Relative deviance reduction in a generalized linear model that predicts the choice between pulse types (P_{slow} versus P_{fast}) from different sensory features (see Methods for details). Distance is most predictive of pulse choice. Each dot corresponds to the performance obtained for a fit to subset (80%) of the data.

D Fraction of P_{slow} out of all pulses produced as a function of distance. Flies bias towards P_{slow} when close to the female ($r^2 = 0.92$, $p = 1.0 \times 10^{-9}$). Thin lines correspond to individual *D. melanogaster* strains, and the thick shaded line depicts mean \pm SEM.

E Fraction of P_{slow} in flies with different sensory manipulations. Blind flies but not pheromone insensitive (PI) or deaf flies produce more P_{fast} , indicating that distance estimation requires visual cues. N flies indicated in parentheses and p-values are from a two-tailed t-test.

F A song mode's carrier frequency strongly correlates with its amplitude ($r^2 = 0.86$, $p = 2 \times 10^{-51}$). Dots correspond to the mean carrier frequency and amplitude (peak to peak) of sine (blue), P_{slow} (red) or P_{fast} (orange) from 47 males of the *D. melanogaster* strain NM91. Grey lines connect song modes of individuals, thick black line is the result of linear regression.

G P_{fast} (orange) is louder than P_{slow} (red) at all distances and is the only pulse type that is amplitude modulated (P_{fast} $r^2 = 0.86$, $p = 1.1 \times 10^{-7}$, P_{slow} $r^2 = 0.03$, $p = 0.5$) (mean \pm SEM across 8 *D. melanogaster* strains).

H, I Average P_{slow} (H) and P_{fast} (I) shape produced at different distances from the female (see colorbar). N= 330759 pulses from 315 individuals and 8 different *D. melanogaster* wild types strains. The shape of both pulse types changes only minimally with distance.

J Rank correlation between female speed and the amount of sine song (blue) or the number of P_{slow} (red) or P_{fast} (orange) pulses per 30 second time window as a function of distance (mean \pm SEM across 8 strains). Sine and P_{slow} are negatively correlated with female speed for almost all distances. P_{fast} only slows females down for distances between 6 and 9 mm. Note the positive correlation when the male is very close to female for all song modes.

Figure 4: The activity of song command neurons and wing muscles affects the choice between song modes.

A Schematic of three elements of the putative song pathway [1,4] manipulated in this study. The total number of neurons per hemisphere is indicated in parentheses. P1 neurons are activated by female cues and are local to the brain. We used two separate drivers to label P1 neurons (R71G01 [5] and P1a [6] - see pictogram in Supplementary Figure 4B for more information on which neurons are labeled by these drivers). pIP10 is a descending neuron. ps1 is a motor neuron that innervates the ps1 wing muscle [4].

B₁₋₃ Population average song probability for producing sine song (blue), P_{slow} (red) or P_{fast} (orange) upon optogenetic activation of P1 (R71G01) (B₁), P1 (P1a) (B₂), and pIP10 (B₃). The trials lasted 3 seconds with the LED being activated (shaded gray area) for 300ms (duty cycle (DC) 0.1) and 2700ms (DC 0.9). Shown are data for three out of nine DCs tested (see Supplemental Figure 4A for the full data set). B-D: Average over 6 P1 (R71G01), 6 P1 (P1a), 3 pIP10 flies, 30 trials per fly and DC.

C₁₋₃ The dominant song mode (color coded, see legend in A) as a function of time for all nine LED DCs tested. Vertical black lines indicate the duration of the stimulation. When the LED is on, P1 (R71G01) (C₁) and P1 (P1a) (C₂) stimulation strongly drives sine song production, while pIP10 (C₃) drives mainly P_{fast} (orange). Sine song is produced after LED stimulation in pIP10.

D₁₋₃ Latency to peak probability for P_{slow} (red) and P_{fast} (orange) for all nine DCs tested (p-values in the plot are the result of two-tailed Wilcoxon rank sum tests between all pairs, corrected for multiple comparisons using the Bonferroni method). pIP10 activation consistently drives P_{fast} with shorter latency than P_{slow} (D₃). There is no such consistent difference in latency between pulse types for P1 (R71G01) (D₁) and P1a (D₂).

E P_{slow} fraction produced by males courting a female when pIP10 is inactivated (gray, using TNT or Kir2.1, respectively) and in control flies (blue). pIP10 inactivation during courtship leads to significantly more P_{slow} (ANOVA ($p=6 \times 10^{-17}$) followed by a Tukey-Kramer post hoc test, p-value is corrected for multiple comparisons). Number of flies for each genotype given in parentheses.

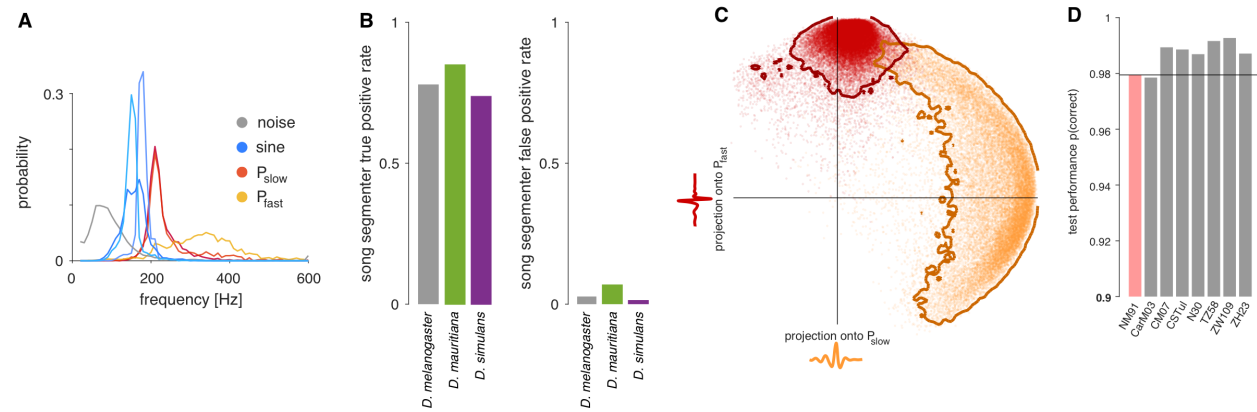
F Fraction of P_{slow} pulses out of all pulses produced as a function of distance to the female in ps1-control flies (purple, inactive TNT expressed in the ps1 motoneuron) and ps1-inactivated flies (gray, functional TNT expressed in ps1). Shaded lines indicate mean \pm SEM across flies). The magnitude of the modulation of pulse choice with distance is strongly reduced upon ps1 inactivation (ps1-control: $r^2=0.92$, $p=3 \times 10^{-5}$; ps1-TNT: $r^2=0.72$, $p=7 \times 10^{-7}$).

G P_{slow} fraction in ps1-TNT flies (gray) and ps1-control flies (purple). Ps1 inactivation reduces the fraction of P_{slow} pulses produced ($p=0.03$, two-tailed Wilcoxon rank sum test).

H Raw microphone amplitude of P_{slow} (left) and P_{fast} (right) pulses in ps1-TNT (gray) and ps1-control (purple) flies. Ps1 inactivation reduces the amplitude of both pulse types ($p=2 \times 10^{-4}$ for P_{slow} and 0.02 for P_{fast} , two-tailed Wilcoxon rank sum test).

I Shapes of the P_{slow} (top) and P_{fast} (bottom) pulses in ps1-ctrl (purple) and ps1-TNT flies (gray). Lines correspond to averages over pulses of each type for each fly. P_{slow} slows and widens weakly with ps1 activation. F-I: N=7 control flies (23636 pulses) and 18 experimental flies (14417 pulses).

SUPPLEMENTAL FIGURES:



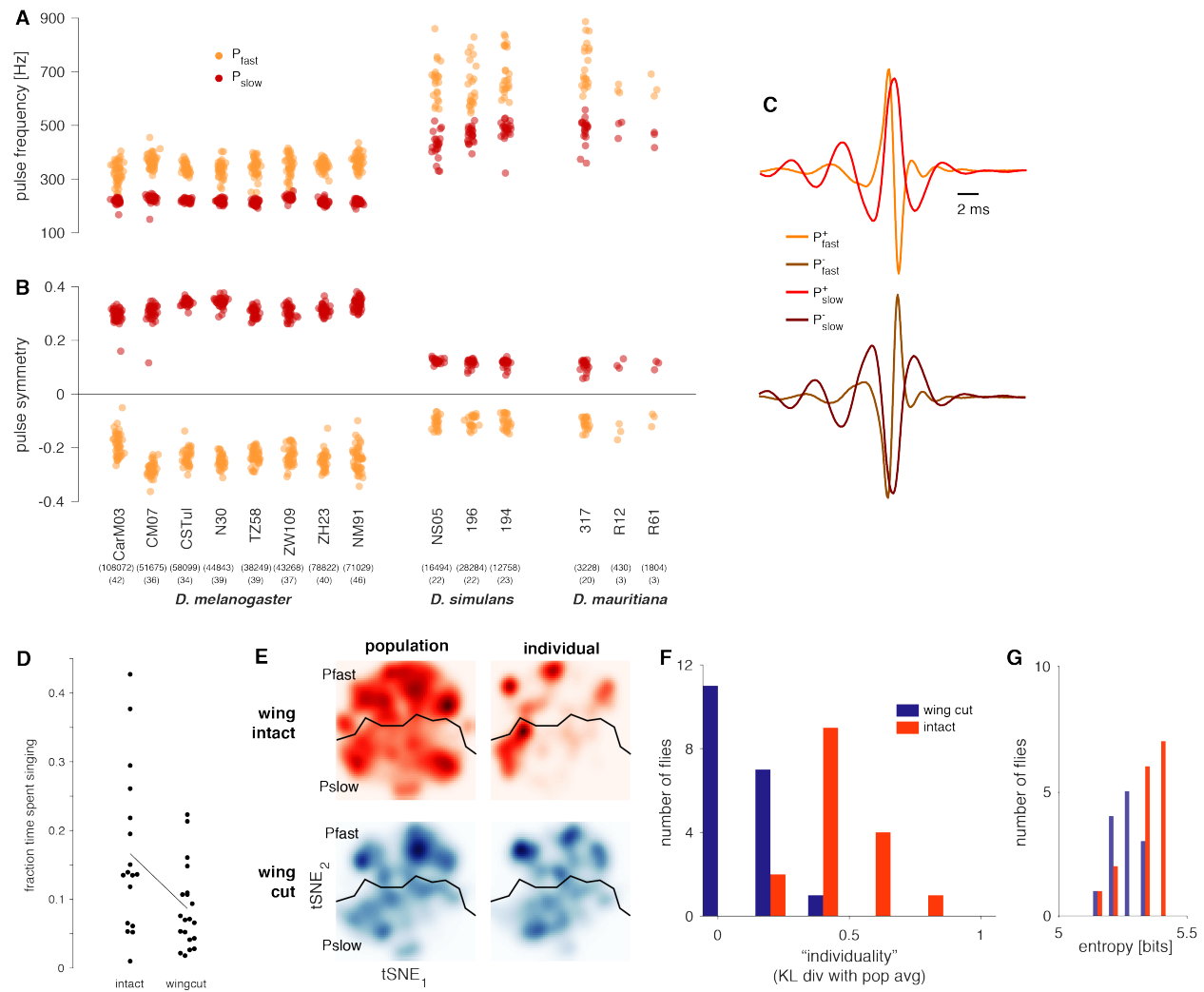
Supplemental Figure 1: P_{fast} and P_{slow} are distinct song modes

A Distribution of carrier frequencies for all 21,104 waveform shapes in Fig. 1C, separated by watershed clusters from Fig. 1E (see legend for color code).

B True (left) and false (right) positive rates for the automatic segmentation of pulses from *D. melanogaster* (grey), *D. mauritiana* (green), and *D. simulans* (purple). True positive rates (left) are high for all species (>74%) with only few false positives (right, <7%). *D. mauritiana* (green, left) exhibits slightly elevated false positive rates because this species tends to produce a lot of movement noise during courtship.

C A template-based pulse type classifier was trained by projecting each pulse onto the P_{slow} and P_{fast} centroids obtained from the cluster analysis. Each dot represents on pulse colored by the pulse label assigned in the cluster analysis. Red and orange lines correspond to the contour enveloping 95% of the pulses for each pulse type. A quadratic decision boundary is then determined using Quadratic Discriminant Analysis (QDA) to classify novel pulses based on their projection values.

D Test performance of the QDA classifier for the strain on which the classifier was trained on (NM91, red) - using a set of pulses not used for training - and for 7 completely new *D. melanogaster* wild type strains (grey). Note the y-axis starts at 0.9 to highlight differences between strains. The classifier generalizes excellently across wild type strains of *D. melanogaster*.



Supplemental Figure 2: P_{fast} and P_{slow} are produced by strains of several *Drosophila* species

A, B Pulse carrier frequency (A) and pulse symmetry (B) for P_{fast} (orange) and P_{slow} (red) centroids of all individuals for the 8 *D. melanogaster*, 3 *D. simulans*, and 3 *D. mauritiana* strains. P_{slow} tends to be the slow and symmetrical, and P_{fast} tends to be the fast and asymmetrical pulse type in all individuals. Numbers in parentheses indicate the number of pulses (top) and flies (bottom) tested for each strain and species.

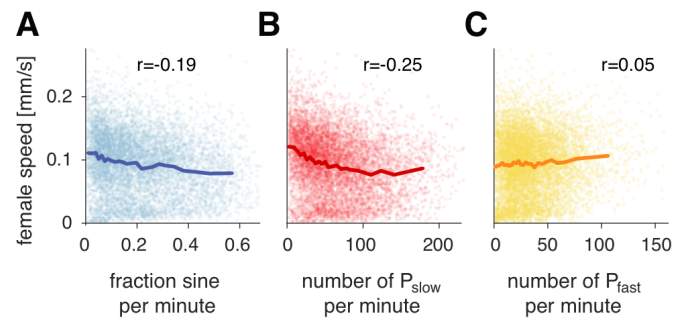
C Clustering pulses with randomized sign produces clusters in which each pulse type appears in a version with “correct” (top, bright colors) and inverted (bottom, darker colors) sign.

D Wing cut males (N=20 pairs, a single wing of the male (either left or right) cut 24h before the experiment) produce less song than intact controls (N=17 pairs) ($p=0.01$, two-tailed Wilcoxon rank sum test).

E All pulses (wing intact – red, top; single wing cut – bottom, blue) were embedded into the same tSNE space and clustered to identify P_{slow} and P_{fast} pulses (boundary between pulse types is marked by a thick black line). We then compared the distribution of pulses in tSNE space between the population (left) and individuals (right) (distributions were estimated as for the pulse clustering, see STAR Methods for details). Wing cut males (bottom) produce pulses with a distribution that tends to resemble that of the population. By contrast, the pulses of intact males seem to be more “idiosyncratic” than those of wing cut males – meaning the pulse distribution in intact males is very dissimilar to that of the population.

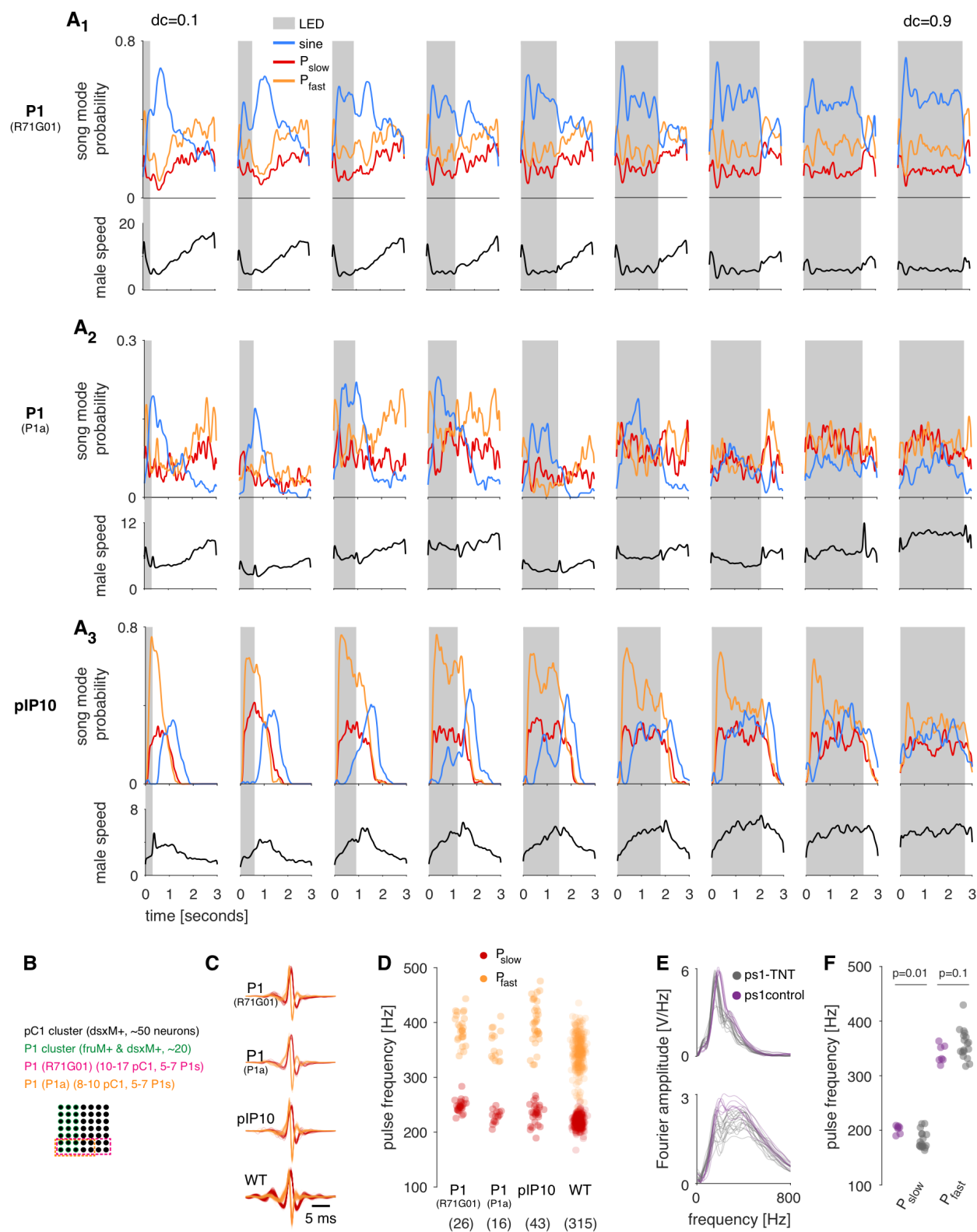
F We quantified the similarity of the pulse distributions of individuals and the population in tSNE space using the Kullback-Leibler divergence (KL) for both pulse types: $D_{\text{KL}}(p_{\text{pop}}, p_{\text{ind}}) = \sum p_{\text{pop}} \log_2(p_{\text{pop}}/p_{\text{ind}})$ where p_{pop} and p_{ind} are the distributions of pulses in tSNE space for the population and for individuals, respectively, as shown in E. Wing cut flies (blue bars) have smaller KL values (meaning higher similarity to the distribution of the population) than intact flies (red bars) ($p=2 \times 10^{-5}$, two-tailed Wilcoxon rank sum test).

G The variability of pulse shapes – given by the entropy of the distribution of pulse shapes in tSNE space – changes only little, indicating that the lack of a wing does not strongly affect the diversity of pulses produced ($p=6 \times 10^{-6}$, two-tailed Wilcoxon rank sum test).



Supplemental Figure 3: Correlation between song modes and female speed

A, B, C Rank correlation between female speed and amount of sine (A) or number of P_{slow} (B) or P_{fast} (C) pulses per 30 second time window of courtship (8826 windows from 315 *D. melanogaster* pairs used). If receptive, females slow down in response to song {Coen et al. 2014}. Individual points correspond to individual time windows, shaded lines are the binned averages (mean \pm SEM). Bin width was chosen adaptively such that each bin contains the same number of data points. Only the amount of sine song and P_{slow} are strongly correlated with a reduction in female speed (rank correlation values indicated in each panel, sine $r=-0.19$, $p=2 \times 10^{-74}$; P_{slow} $r=-0.25$, $p=2 \times 10^{-129}$; P_{fast} $r=0.05$, $p=2 \times 10^{-5}$).



Supplemental Figure 4: Three different song pathway neurons and their role in song patterning

A Population average song probability for producing sine song (blue), P_{slow} (red) or P_{fast} (orange) upon optogenetic activation of P1 (A_1), P1a (A_2), and pIP10 (A_3). Same data as in Fig. 1B in the main manuscript but with all nine DCs used. Black lines show the trial and population averaged male speed.

B Schematic of the labelling of P1 (R71G01) and P1 (P1a). pC1 (black) is an anatomical cluster of neurons that are *DsxM+*. P1 (green) is a subset of pC1 that is also *FruM+*. R71G01 [5] (labelled “P1 (R71G01)” in our figures) labels 5-7 P1 neurons, in addition to 10-17 pC1 neurons and many other neurons in the fly. P1a [6] labels the same set of P1 neurons as R71G01 but only 8-10 *FruM-* pC1 neuron and a few other neurons.

C, D P_{slow} (red) and P_{fast} (orange) shapes (C) and frequencies (D) for optogenetic activation of P1 (R71G01), P1a, and pIP10. *D. melanogaster* wild type data shown for comparison. Number of flies for panels C and D are given in parentheses in panel D.

E Spectra of P_{slow} (top) and P_{fast} (bottom) in ps1-ctrl (purple) and ps1-TNT flies (grey). Note the small peak shift for P_{slow} for ps1 inactivation.

F Frequency of P_{slow} (right) and P_{fast} (left) pulses for ps1 control (purple) and ps1 experimental (gray) flies. Each dot corresponds to the frequency of the average pulse for one individual (N=24 experimental flies (36358 pulses) and 8 control flies (25850 pulses)). Inactivation of ps1 decreases P_{slow} but not P_{fast} frequency (P_{fast} , $p=0.1$; P_{slow} $p=0.01$, two-tailed Wilcoxon rank sum test). E-F: N=7 experimental flies (36358 pulses) and 24 control flies (25850 pulses).

References

- [1] Philipsborn von AC, Liu T, Yu JY, Masser C, Bidaye SS, Dickson BJ. Neuronal control of *Drosophila* courtship song. *Neuron* 2011;69:509–22. doi:10.1016/j.neuron.2011.01.011.
- [2] Coen P, Xie M, Clemens J, Murthy M. Sensorimotor Transformations Underlying Variability in Song Intensity during *Drosophila* Courtship. *Neuron* 2016;89:629–44. doi:10.1016/j.neuron.2015.12.035.
- [3] David JR, Lemeunier F, Tsacas L, Yassin A. The Historical Discovery of the Nine Species in the *Drosophila melanogaster* Species Subgroup. *Genetics* 2007;177:1969–73. doi:10.1534/genetics.104.84756.
- [4] Shirangi TR, Stern DL, Truman JW. Motor control of *Drosophila* courtship song. *Cell Reports* 2013;5:678–86. doi:10.1016/j.celrep.2013.09.039.
- [5] Pan Y, Meissner GW, Baker BS. Joint control of *Drosophila* male courtship behavior by motion cues and activation of male-specific P1 neurons. *Proc Natl Acad Sci U S A* 2012;109:10065–70. doi:10.1073/pnas.1207107109.
- [6] Hoopfer ED, Jung Y, Inagaki HK, Rubin GM, Anderson DJ, Ramaswami M. P1 interneurons promote a persistent internal state that enhances inter-male aggression in *Drosophila*. *eLife* 2016;4:e11346. doi:10.7554/eLife.11346.