Article

Current Biology

Discovery of a New Song Mode in *Drosophila* Reveals Hidden Structure in the Sensory and Neural Drivers of Behavior

Highlights

- *D. melanogaster* males produce three song modes: one sine and two pulse modes
- Males select between pulse modes based on distance to the female
- Activity levels in song pathway neurons affect song mode choice
- The three song modes differentially affect female behavior

Authors

Jan Clemens, Philip Coen, Frederic A. Roemschied, ..., Diego E. Aldarondo, Diego A. Pacheco, Mala Murthy

Correspondence

mmurthy@princeton.edu

In Brief

For 50 years, *Drosophila melanogaster* song was thought to consist of only two modes: sine and pulse. Using unsupervised classification methods, Clemens et al. establish the existence of three distinct song modes. They show how this distinction affects the interpretation of the mechanisms underlying song production and perception.

Clemens et al., 2018, Current Biology *28*, 2400–2412
August 6, 2018
https://doi.org/10.1016/j.cub.2018.06.011

CellPress

CellPress

Current Biology
Article

Discovery of a New Song Mode in *Drosophila* Reveals Hidden Structure in the Sensory and Neural Drivers of Behavior

Jan Clemens,[1,3,5] Philip Coen,[1,4,5] Frederic A. Roemschied,[1,5] Talmo D. Pereira,[1] David Mazumder,[1,2] Diego E. Aldarondo,[1] Diego A. Pacheco,[1] and Mala Murthy[1,2,6,*]
[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
[5]These authors contributed equally
[6]Lead Contact
*Correspondence: mmurthy@princeton.edu
https://doi.org/10.1016/j.cub.2018.06.011

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. Here, we address this issue in the context of acoustic communication in *Drosophila*. During courtship, males vibrate their wings to generate time-varying songs, and females evaluate songs to inform mating decisions. For 50 years, *Drosophila melanogaster* song was thought to consist of only two modes, sine and pulse, but using unsupervised classification methods on large datasets of song recordings, we now establish the existence of at least three song modes: two distinct pulse types, along with a single sine mode. We show how this seemingly subtle distinction affects our interpretation of the mechanisms underlying song production and perception. Specifically, we show that visual feedback 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. At the neural level, we demonstrate how the activity of four separate neuron types within the fly's song pathway differentially affects the probability of producing each song mode. Our results highlight the importance of carefully segmenting behavior to map the underlying sensory, neural, and genetic mechanisms.

INTRODUCTION

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. Computational methods (e.g., [1–3]) 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 [4, 5]. But solving the neural basis for behavior requires the appropriate temporal and spatial parsing of behaviors in order to generate meaningful connections with the neural and muscle activity patterns that drive them [6]. A comprehensive classification may reveal that a seemingly variable behavior is in fact a mixture of distinct and stereotyped behavioral modes, each with its own neural basis and evolutionary trajectory. Failing to discriminate distinct behavioral modes will make the underlying brain activity appear noisy or redundant, because distinct activity patterns will appear to produce a single behavior. Here, we apply unsupervised analysis techniques to examine a behavior that has been studied for more than 50 years [7]—the courtship song produced by males of the species *Drosophila melanogaster*.

During courtship, *D. melanogaster* males vibrate their wings to produce an acoustic signal to convince the female that he is a suitable mating partner. Manual inspection of song recordings has identified two distinct song modes: sine song, which consists of a sustained sinusoidal oscillation, and pulse song, which comprises trains of short impulses [7, 8]. Based on this observation, automated methods were developed to effectively detect and segment these two modes of song [9], and using this software, large datasets have been analyzed to link both genes and neural activity with song patterns [10–14]. 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 indicated for other *Drosophila* species, including *D. simulans* and *D. mauritiana* [15]).

Whether *D. melanogaster* produces additional song modes has important consequences for the study of the neural basis of behavior. For instance, previously described neural circuits for song production might be incomplete if they only explain the production of a subset of song modes. In addition, the arrangement of song modes into bouts depends on dynamic sensory feedback from the female: males bias toward singing

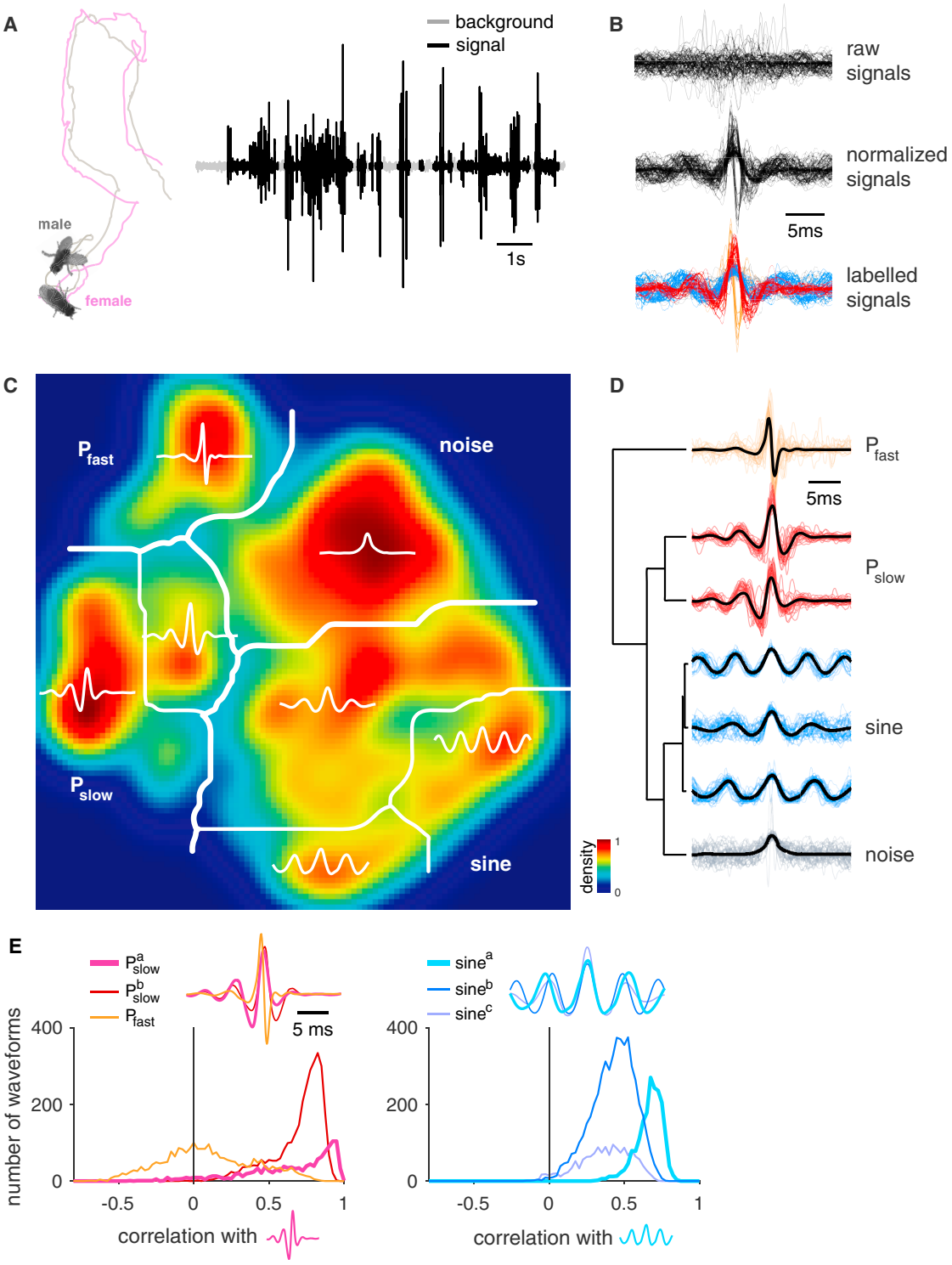

Figure 1. *Drosophila melanogaster* Courtship Song Comprises Three, Not Two, Modes

(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 (gray) and signals (black) from a wild-type *D. melanogaster* male is shown.

(B) Non-overlapping signal chunks (25 ms 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 (gray), sine (blue), P_{slow} (red), and P_{fast} (orange).

(legend continued on next page)

more pulse song when moving fast or further from the female [10] and also modulate the amplitude of their pulses with distance to the female [11]. This patterning strategy also has an impact on female behavior [12]. Elaborating the song modes of *Drosophila melanogaster* therefore has implications for studies of how sensory information is integrated into or modulates the motor circuits that generate song.

By analyzing a large dataset of song recordings with unsupervised classification methods, we here show that *D. melanogaster* produces two distinct types of pulses in addition to sine song—one pulse type with relatively low frequency and a more symmetrical shape and one pulse type with relatively faster frequency and asymmetrical shape, which we term P_{slow} and P_{fast}, respectively. We then show how the discovery of two distinct pulse modes affects our interpretation of (1) the evolution of pulse song, (2) the sensory cues controlling pulse song amplitude, and (3) the neuronal basis of pulse song production. By examining pulse structure in *D. melanogaster*'s close relatives *D. simulans*, *D. mauritiana*, and *D. sechellia*, we find that the distinction between P_{slow} and P_{fast} is common to three of these species, although *D. sechellia* only produces P_{slow}. We also show that P_{fast} is much louder than P_{slow} and produced more frequently when the male is far away from the female. The choice between these two pulse types is thus part of the male strategy to sing louder when the female is distant [11]. Additionally, we show that the activity of identified song pathway premotor and motor neurons [14, 16] affects the choice between different song modes. Lastly, we demonstrate that the two pulse types differentially affect female behavior, suggesting these communication signals have different meanings for the female. Overall, our results illustrate how correct segmentation of behavior is important for understanding the evolutionary, sensory, and neural drivers of behavior.

RESULTS

Drosophila melanogaster Courtship Song Comprises at Least Three, Not Two, Distinct Song Modes

To determine how many song modes *Drosophila melanogaster* males produce, we clustered 25-ms waveforms of previously collected raw song recordings [10]. 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 most 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 dataset using t-distributed stochastic neighbor embedding (tSNE) [2, 17]. The tSNE method is a nonlinear dimensionality reduction method that preserves local similarity structure in a dataset 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, because it cuts the signal space along relatively weak local minima and thereby assigns 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): (1) A "noise" mode lumps all signals that lack common structure across exemplars. (2) A "sine" mode joins three clusters that contain waveforms with sustained oscillations that tile a continuum of carrier frequencies between 120 and 180 Hz (Figure S1A). 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

(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 color bar) 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 (gray) 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 toward high values, and all three distributions overlap, indicating that sine song constitutes a single song mode.

See also Figure S1.

merged into a single sine song mode by the hierarchical cluster algorithm. (3) A "pulse" mode with relatively slow (200–250 Hz) and symmetrical waveforms connects two very similar clusters with either a stronger leading or lagging lobe. These two clusters form a single pulse mode, because the waveforms are highly correlated (see Figure 1E, left), and they are thus joined by the hierarchical cluster algorithm. (4) A second "pulse" mode that, in contrast to the previous one, is biphasic (asymmetric) with faster oscillations (250–400 Hz). This pulse mode is highly dissimilar with the previous pulse mode (correlation values centered around zero; Figure 1E, left), indicating that it forms a distinct pulse mode. We term the two pulse modes "P_{slow}" and "P_{fast}," according to their frequencies. Note that, whereas 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 detected from an automatic song segmenter [9, 10]—which can detect both pulse types with high speed and reliability (detection rates: 78% for all pulses; Figure S1B; 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 (Figures 2A and S1C; 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* (Figures 2B, S1D, S2A, and S2B).

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 hand-segmented pulses (Figure S2E), when using linear as opposed to nonlinear dimensionality reduction (principal-component analysis instead of tSNE; Figure 2C), when omitting dimensionality reduction altogether (Figure S2F, left), and also when omitting parts of the normalization procedure (Figure S2G), demonstrating that our finding does not crucially depend on the details of our analysis pipeline. 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 [11]—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 (Figures 2E and 2F): P_{slow} is produced with one wing more fully extended (ca. 60°) and 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.

For all three song modes, we observed that the secondary wing was minimally extended to angles <10degrees (Figure 2F). To determine whether the secondary wing contributes to song production, we recorded song from courting males that had one wing cut (either left or right). As expected, these males showed a reduction of singing to ~50% of intact males (Figure S3A). Whereas these males still produced both pulse types, the individuality of these pulses (how similar a given male's pulses were to the population average) was strongly reduced (Figures S3B and S3C). This suggests that the existence of two wings increases the individuality of pulse shapes produced by males, and hence, the function of the secondary wing may be to make the pulse shape more idiosyncratic. We do not know whether the female is sensitive to these small changes in pulse shape. However, we observed no strong effect of cutting one wing on pulse variability (Figure S3D).

How do the two pulse types—P_{fast} and P_{slow}—in *Drosophila melanogaster* compare to the pulse types produced in other species [15]? We applied our analysis pipeline to song data from three closely related species: *Drosophila simulans*; *Drosophila mauritiana*; and *Drosophila sechellia* (Figure 2G). Although *D. simulans* and *D. mauritiana* have been shown previously to produce two pulse types [15], little is published about the pulse shape of *D. sechellia* [19, 20]. Because 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; Figure S1B). 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 (Figures 2H–2J, S2A, S2B, and S2D). *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 ~350 Hz to 670 Hz with only small differences between *D. simulans* and *D. mauritiana* (Figures 2I and 2J). *D. sechelia* produced only a single, symmetrical pulse type that resembles P_{slow} and is of slightly lower frequency than P_{slow} in *D. melanogaster* (Figures 2H, 2I, S2A, and S2D). That our analysis does not always classify pulses into two types validates our approach. Note that our failure to find P_{fast} in *D. sechellia* is not proof of its non-existence—the species could produce P_{fast} under different experimental conditions. This result also shows how false classification can lead to false estimates of the magnitude of evolutionary changes in pulse shapes: when comparing the pulse frequencies in *D. sechellia* and *D. melanogaster*, disregarding the distinction between P_{fast} and P_{slow}, the difference appears >100% higher than when doing the correct comparison of *D. sechellia* pulses with P_{slow} in *D. melanogaster* (Figure S2H). The phylogenetic pattern of pulse shapes in the four species suggests the following hypothesis about 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. *D. sechellia* has lost P_{fast} without significant change to the frequency of P_{slow}. Further studies are necessary to test these hypotheses. 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*.

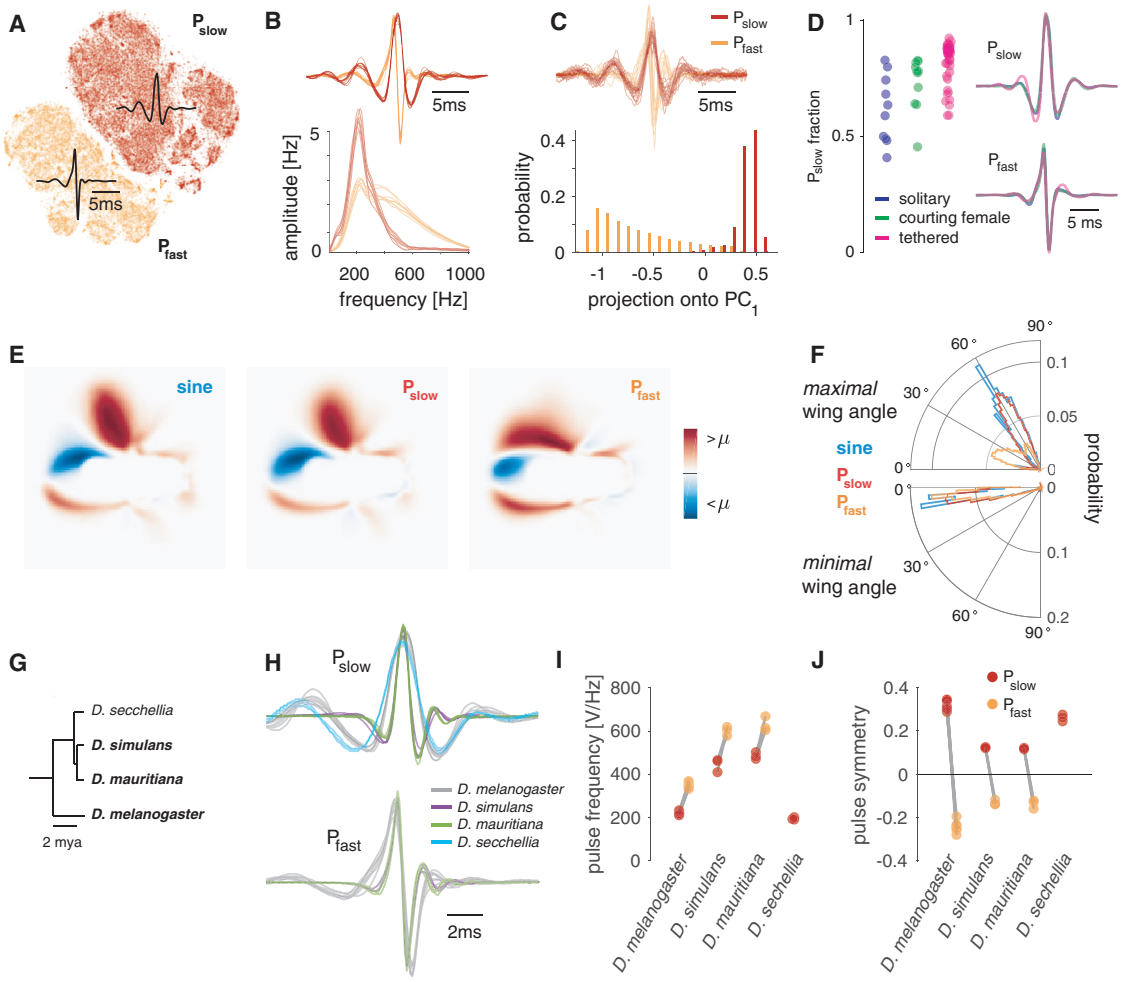

Figure 2. Song Pulses Can Be Separated into P_{fast} and P_{slow} across Strains and Species, and Each Corresponds to a Distinct Wing Pose

(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 STAR 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, which captures 55% of the variance. 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 [16] in either solitary males (purple), males courting a female (green), or males tethered and walking on a ball (pink). Flies on the ball experienced open-loop visual stimulation matching the natural statistics of female motion on the male retina during courtship [11]. 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. Number of flies (pulses) is 10 (107,710) for solitary males, 11 (138,311) for males with female, and 34 (176,526) for tethered males.

(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 (100 Hz 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 and P_{fast} (right) is produced with the wing held closer to the body. See legend for color map.

(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. Data in (E) and (F) are as follows: 99,181 ms sine, 56,386 P_{slow} pulses, and 47,241 P_{fast} pulses from 42 individuals of the *D. melanogaster* strain NM91.

(G) Phylogenetic tree for the four *melanogaster* species used (reproduced from [18]). The species that produce P_{fast} 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) pulse type, and all species but *D. sechellia* produce a faster asymmetrical pulse (P_{fast}, bottom) type.

(I and 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*. *D. sechellia* produces slower P_{slow} pulses than *D. melanogaster* and no P_{fast}. For N flies and pulses, see Figure S2A.

See also Figures S2 and S3.

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 [9], with the choice to sing either sine or pulse biased by sensory feedback from the female [10]. How does classifying song into three modes affect our understanding of how song is patterned by feedback cues? 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 (Figures 3A and 3B). By employing generalized linear models [10, 21] 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 (Figures 3C and S3E). 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 [10]; 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 [11] and found that, although P_{fast} amplitude increases with distance (males ratchet up P_{fast} amplitude as they get farther from the female without changing pulse shape; Figures S3F and S3G), P_{slow} amplitude modulation was comparatively weak (Figure 3G). Taken together, these data suggest a reinterpretation of previous work on amplitude modulation [11]: the behavior is driven by at least two processes—the binary choice to sing the softer or the louder pulse type (P_{slow} versus P_{fast}) coupled with continuous modulation of P_{fast} amplitude.

Individual Cell Types within the Song Pathway Modulate the Choice between Song Modes

Above, we showed that males select between pulse types based on sensory feedback (Figures 3C and 3D) 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 toward one of the three song modes. To test this hypothesis, we focused on four previously characterized *Fruitless*+ neurons implicated in driving wing extension and/or generating pulse song (Figures 4A and S4B): P1, a cluster of ~20 central brain neurons per hemibrain [24, 25]; pIP10, a pair of neurons that descend from the brain to the ventral nerve cord (VNC) [16]; vPR6, a set of ~5 thoracic interneurons per side [16]; and ps1, a wing muscle motor neuron in the VNC [14]. Previous work indicates that P1, pIP10, and vPR6 constitute a connected pathway from the brain to the wing motor neuropil of the VNC [16].

P1, pIP10, or vPR6 neural activation results in song production, even in the absence of a female [11, 16]. 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 [25–28], the production of aggressive behaviors [23], and overall mating drive [29]. We reasoned that, when males are close to females, the pheromonal and visual cues provided by the female would be particularly strong and hence P1 activity should be elevated. Because males bias toward both sine song and P_{slow} production when close to the female [10] (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 and vPR6 neurons, via connections to P1 [16], might therefore exert similar effects 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 [30] to optogenetically stimulate each neuron type in solitary males. We chose a light-emitting diode (LED) stimulation protocol that allowed us to vary the duty cycle (DC) 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 2,700 ms) or very short pauses (DC = 0.9 corresponds to a pause of 300 ms) 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, and both P_{fast} and P_{slow} were largely suppressed (Figures 4B1, 4B2, 4C1, and 4C2). We observed this effect for two genetic drivers that target P1 neurons: R71G01 [22] and a split GAL4 driver termed P1a that is the intersection of R71G01 with a second driver [23] (Figure S4B). Interestingly, for P1a, sine song was only dominant for DCs < 0.8; at higher DCs, pulse song was more dominant, suggesting that sine production requires a pause in P1a activity (Figures S4A2 and 4C2). P_{fast} and P_{slow} probabilities exhibited similar dynamics for both P1 drivers (Figures 4B1, 4B2, S4A1, and S4A2), suggesting that P1 mainly biases the choice between sine and pulse song, affecting the balance between P_{fast} and P_{slow} relatively little. Moreover, P1 activation did not affect the waveform shape of either P_{fast} or P_{slow}, only the probability of producing pulses (Figures S4C and S4D).

pIP10 is a descending neuron thought to be postsynaptic to P1 [16]. However, we found that pIP10 does not simply relay P1 activity, because P_{fast}, and not sine song, was the dominant mode produced during optogenetic activation (Figures 4B3 and 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 [31] or TNT [32]) 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} (Figures 4B3 and S4A3).

vPR6 is a thoracic local neuron thought to be postsynaptic to pIP10—constitutive thermogenetic activation of vPR6 is known to modulate the inter-pulse interval of pulse trains and so this neuron was proposed to be part of the central pattern generator for pulse song (Figure 4B4) [16]. We found that optogenetic activation of vPR6 drives only pulses (and no sine) and that P_{fast} and P_{slow} production dominate early and late during stimulation, respectively (Figure 4C4). Thus, this neuron affects the choice

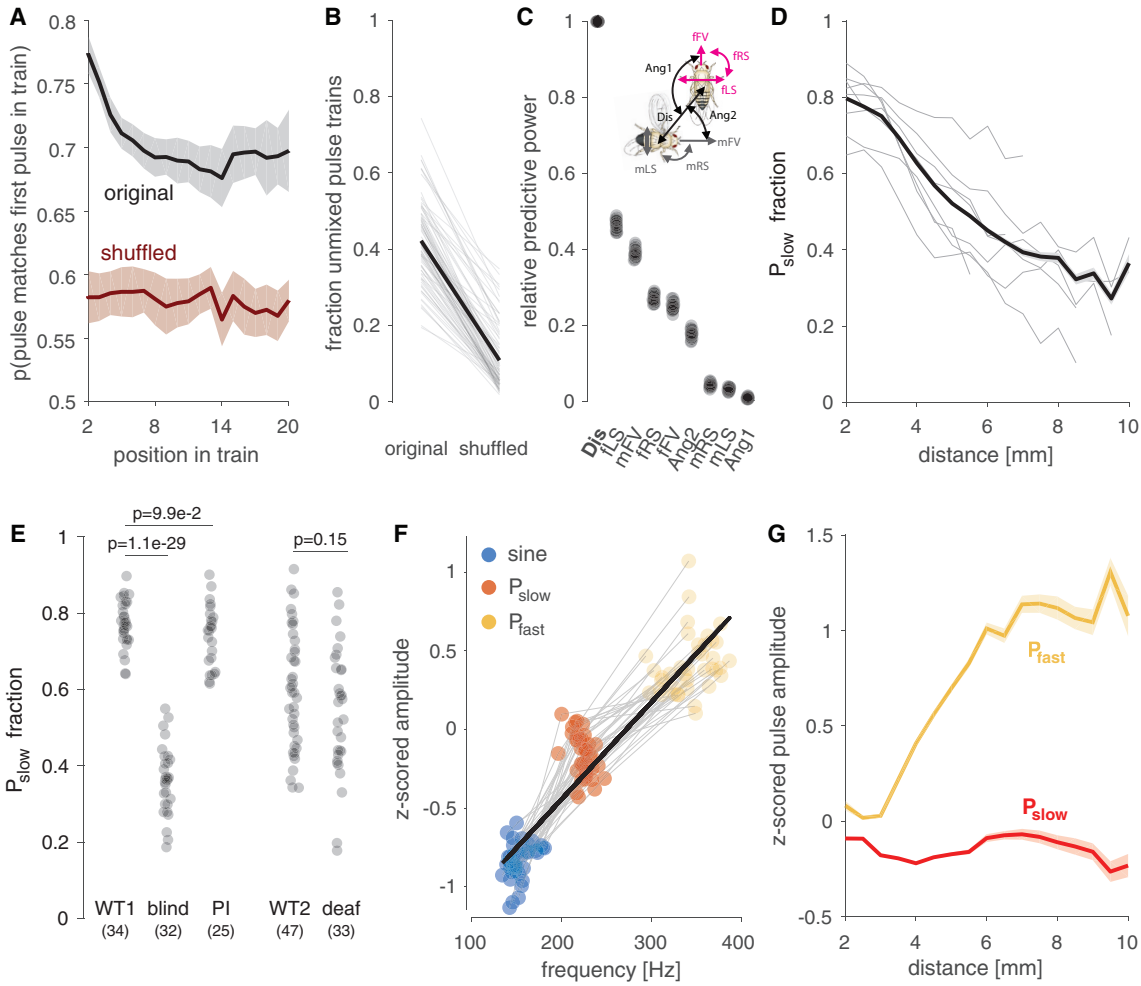

Figure 3. The Choice between P_{slow} and P_{fast} Is Driven by Sensory Feedback

(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 ± SEM across 8 *D. melanogaster* strains is shown.

(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). 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 power of different sensory cues for predicting the choice between pulse types (P_{slow} versus P_{fast}) from different sensory features (see STAR Methods for details). Distance (Dis) is most predictive of pulse choice. Each dot corresponds to the performance obtained for a fit to a subset (80%) of the data. Abbreviations for the sensory cues tested are defined in the pictogram: Ang1, angle between female forward velocity and male position; Ang2, angle between male forward velocity and female position; Dis, distance; fFV, female forward velocity; fLS, female lateral speed; fRS, female rotational speed; mFV, male forward velocity; mLS, male lateral speed; mRS, male rotational velocity.

(D) Fraction of P_{slow} out of all pulses produced as a function of distance. Flies bias toward 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 ± 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. Numbers of flies (pulses) are 34 (58,099) for WT1, 32 (56,176) for blind flies, 25 (99,718) for PI flies, 47 (71,029) for WT2, and 33 (47,059) for deaf flies.

(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 (71,029 pulses) 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 ± SEM across 8 *D. melanogaster* strains).

Analyses in (A)–(D) and (G) based on 315 flies (330,759 pulses) from 8 wild-type strains (see Figure S2A for details). See also Figure S3.

between P_{fast} and P_{slow}—weak stimulation drives P_{fast}, but stronger stimulation can drive P_{slow}.

Our results with P1, pIP10, and vPR6 activation reveal complex dynamics in the neural circuits driving song production. For activation of all three cell types, we observed that stimulus history (and thereby the history of song produced) has a strong impact on the song types elicited by optogenetic activation (Figures 4B and S4A). Although optogenetic activation of all three

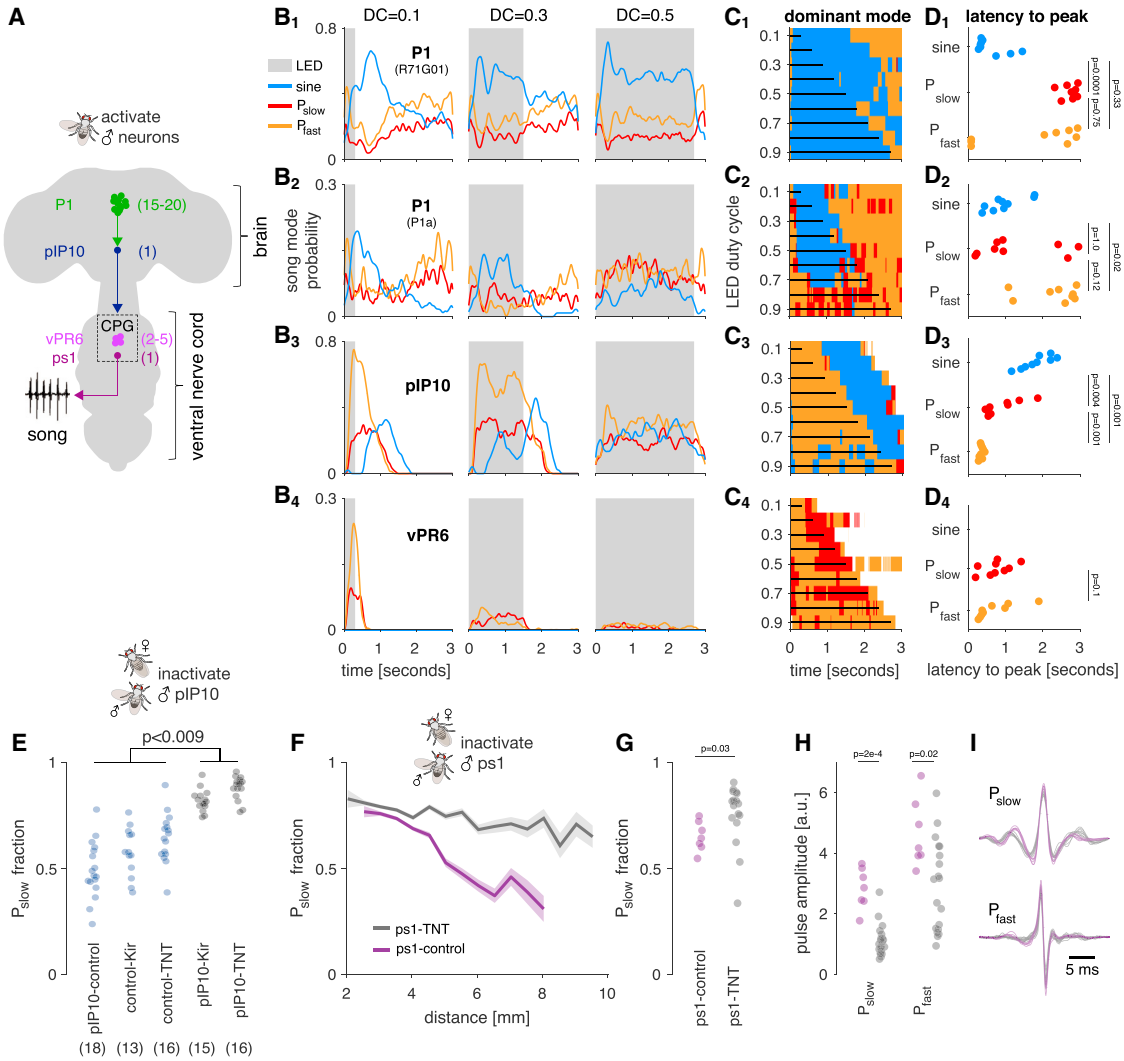

Figure 4. The Activity of Song Pathway Neurons Modulates the Choice between Song Modes

(A) Schematic of four elements of the putative song pathway [14, 16] 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 [22] and P1a [23]—see pictogram in Figure S4B for more information on which neurons are labeled by these drivers). pIP10 is a descending neuron. vPR6 is a VNC local neuron [16]. ps1 is a motor neuron that innervates the ps1 wing muscle [14].

(B_{1-4}) Population average song probability for production of sine song (blue), P_{slow} (red), or P_{fast} (orange) upon optogenetic activation of P1 (R71G01; B_1), P1 (P1a; B_2), pIP10 (B_3), and vPR6 (B_4). The trials lasted 3 s with the LED being activated (shaded gray area) for 300 ms (duty cycle [DC] 0.1) and 2,700 ms (DC 0.9). Shown are data for three out of nine DCs tested (see Figure S4A for the full dataset).

(C_{1-3}) 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; C_1) and P1 (P1a; C_2) stimulation strongly drives sine song production, and pIP10 (C_3) drives mainly P_{fast} (orange). Sine song is produced after LED stimulation ends. vPR6 activation only drives pulse song, with P_{fast} and P_{slow} being dominant during early and late phases, respectively.

(D_{1-4}) Latency to peak probability for sine (blue), 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} (D_3). There is no such consistent difference in latency between pulse types for P1 (R71G01; D_1) and P1a (D_2) activation. There are no consistent differences in latencies for vPR6 (D_4) activation. Analyses in (B)–(D) are based on 49,955 pulses from 5 flies for P1 (R71G01), 13,364 pulses from 4 flies for P1 (P1a), 30,270 pulses from 5 flies for pIP10, and 4,516 pulses from 8 flies for vPR6, 30 trials per fly and DC.

(E) P_{slow} fraction produced by males courting a female when pIP10 is inactivated (gray, using TNT or Kir2.1) 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 is given in parentheses. Number of flies (pulses) are 18 (147,261) for pIP10-control, 13 (270,997) for control-Kir, 16 (309,723) for control-TNT, 15 (339,398) for pIP10-Kir, and 16 (386,819) for pIP10-TNT.

(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 motor neuron) and ps1-inactivated flies (gray, functional TNT expressed in ps1). Shaded lines indicate mean ± 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: $r^2 = 0.72$, $p = 7 \times 10^{-7}$).

(legend continued on next page)

cell types affected male speed (Figure S4A), we observed no fixed relationship between male speed and song choice among the three cell types examined here, suggesting that the song dynamics induced by optogenetic activation cannot be explained as an indirect effect of changes in male speed alone.

Finally, we investigated the activity of a song pathway motor neuron implicated in regulating both the amplitude and carrier frequency of song pulses [14]. This effect is consistent with ps1 being involved in P_{fast} production, because P_{fast} is the louder and faster of the two pulse types. Because 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 males with ps1 neurons inactivated still produce both pulse types, with no deficit in the shape of either P_{slow} or P_{fast} (Figure 4I) and only a subtle effect on the frequency of P_{slow} pulses (Figures S4E and S4F). 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, demonstrating that it has additional functions beyond pulse choice (Figure 4H). This implies that, although ps1 is not strictly necessary for P_{fast} production—~10% of pulses are still of the P_{fast} type and indistinguishable from wild-type P_{fast} pulses (Figures 4I, S4E, and S4F)—this motor neuron contributes to the choice between pulse types and thereby the overall structure of male song.

Female Song Responses Depend on Pulse Type Choice in Males

Up to now, we have shown that the distinction between P_{fast} and P_{slow} is functionally significant for the male, for instance, as a strategy to efficiently produce pulse song at different amplitudes (Figures 3F and 3G). That is, it may be energetically more efficient for the male to produce P_{fast} at high amplitudes. Under this scenario, two pulse types could evolve without having different effects on the female, and experimentally discriminating the two pulse types would still be important for our understanding of song evolution (Figures 2G–2J) and of the role of particular neuron types in patterning song (Figures 3 and 4). However, independent of their function for the male, the two pulse types could also serve as distinct signals to the female. We hence examined whether females respond differentially to P_{fast} and P_{slow} and whether regular male song mode choice—e.g., producing more P_{slow} when close to the female (Figure 3D)—influences mating success.

Previous studies have shown that 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 [10, 12]. We examined female locomotor speed relative to the amount of each song mode the male produced. We used time windows of 30 s to investigate this correlation, because previous work demonstrated that females integrate across song bouts [12]. When examined across all distances, P_{slow} and sine are negatively correlated with female speed (Figures 5A and 5B), and there is no correlation between P_{fast} and female speed (Figure 5C). However, males choose between pulse types in a context-dependent manner (Figure 3D)—distance-dependent female responses would indicate that females discriminate between pulse types and between the context in which they are produced. Indeed, when examining correlations relative to male-female distance, we found that all three song modes have the potential to slow females (Figure 5D). 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 only when males were far away from females (and then only weakly). All three modes were correlated with increases in female speed when produced very close to the female.

If females are sensitive to the "correct" song mode choice, then experimental disruption of the male song pattern during courtship should result in a less attractive song and consequently reduced copulation. Males with inhibited pIP10 or ps1 still sing abundantly but produce either less P_{fast} or do not effectively switch between pulse types based on distance (Figures 4E–4G; see figure legends for number of pulses produced); we found that these males also copulated less with wild-type females (Figures 5E and 5F). Likewise, optogenetic activation of pIP10 in males during courtship abrogates distance-dependent pulse choice (Figure S5A) and disrupts song mode choice in general (Figure S5B). The resulting song was unattractive to females: relative to controls for this genotype, both pulse types are now correlated with female acceleration (versus slowing) when pIP10 neurons are optogenetically activated in males (Figure 5G). In addition, pIP10-activated males copulated less (Figure 5H). Overall, this suggests that females discriminate between the two pulse types and that the context in which each pulse type is produced is likely to be an important determinant of a song's attractiveness. Further elucidation of the differential roles of P_{slow} and P_{fast} on female behavior will most likely require higher resolution readouts of female postural movement changes [2, 33, 34]. The reclassification of male courtship song therefore also has implications for the study of auditory perception in the female nervous system.

DISCUSSION

Here, we show that *Drosophila melanogaster* courtship song comprises at least three distinct modes, and we demonstrate

(G) P_{slow} fraction in ps1-TNT flies (gray) and ps1-control flies (purple). ps1 inactivation increases 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 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.
(F)–(I): 23,636 pulses from 7 control flies and 14,417 pulses from 18 experimental flies are shown.
See also Figure S4.

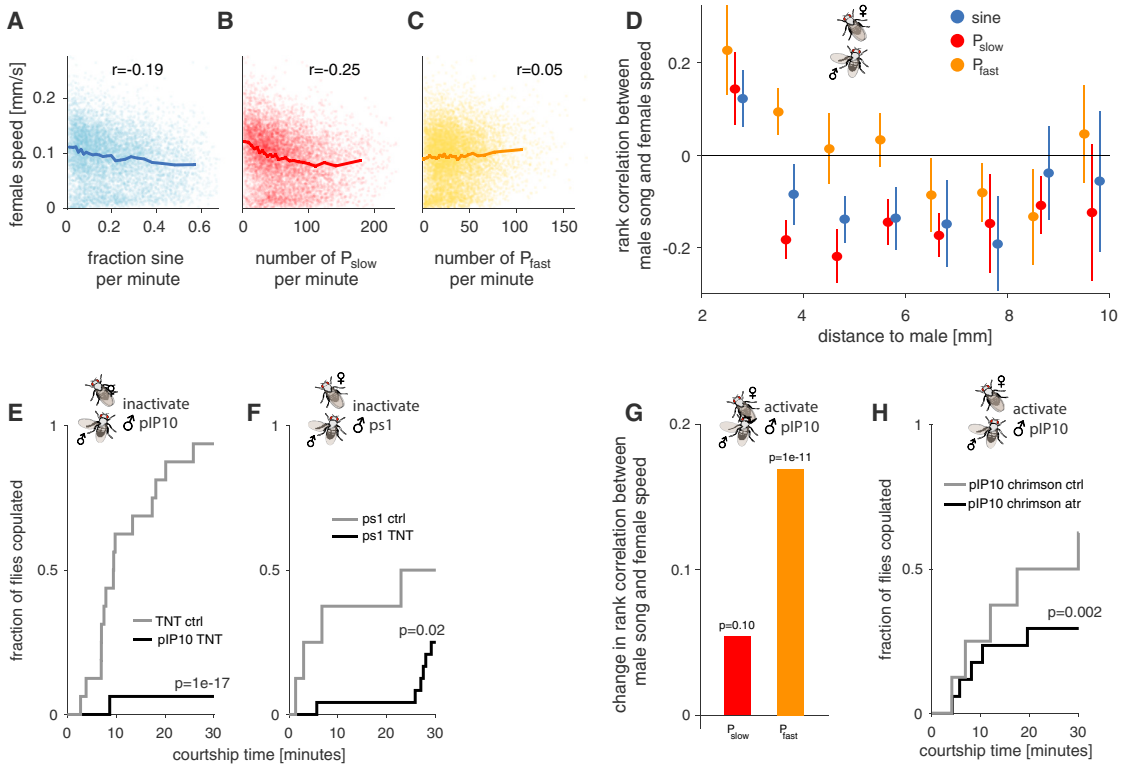

Figure 5. Females Discriminate between Pulse Types and Are Sensitive to Song Context

(A–C) Rank correlation between female speed and amount of sine (A) or number of P_{slow} (B) or P_{fast} (C) pulses per 30 s time window of courtship (8,826 windows from 315 *D. melanogaster* pairs used). Receptive females slow in response to song [10]. Individual points correspond to individual time windows; shaded lines are the binned averages (mean ± 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} $r = 0.05$, $p = 2 \times 10^{-5}$). Individual points correspond to individual time windows; shaded lines are the binned averages (mean ± SEM).

(D) Rank correlation between female speed and the amount of sine song (blue) or the number of P_{slow} (red) or P_{fast} (orange) pulses as a function of distance (mean ± SEM; same data as in A–C). Sine and P_{slow} are negatively correlated with female speed for almost all distances. P_{fast} only slows females for distances between 6 and 9 mm. Note the positive correlation when the male is very close to female for all song modes.

(E and F) Copulation success of males in which song mode choice is disrupted by inactivating the premotor neuron pIP10 (E) or the motor neuron ps1 (F) using TNT (black). Control flies (gray; same as in Figures 4E and 4G) are more likely to copulate within 30 min ($p = 1 \times 10^{-17}$ for pIP10 and 0.02 for ps1; two-sided binomial test). Number of flies is as in Figures 4E and 4G.

(G) Relative rank correlation between female speed and the number of P_{slow} (red) or P_{fast} (orange) pulses per 30-s time window of courtship (rank correlations from controls with males not fed all-*trans* retinal [286 time windows from 8 flies] are subtracted from rank correlations from experiments with males fed all-*trans* retinal [656 time windows from 17 flies]—in both conditions, there is red light stimulation). Song mode choice was disrupted by optogenetic activation of pIP10 in experimental males (see Figure S5B) that courted wild-type females. All males expressed the red-shifted channelrhodopsin csChrimson [30] in pIP10—all-*trans* retinal is a cofactor required for functional channelrhodopsin in flies. Control flies were raised on regular food.

(H) Copulation success for experimental (black, 17 flies) and control (gray, 8 flies) males from (G). Disrupting natural song mode choice via optogenetic activation of pIP10 reduces mating success ($p = 1.6 \times 10^{-3}$; two-sided binomial test).

See also Figure S5.

how this distinction affects our interpretation of the mechanisms underlying evolution (Figure 2), song patterning (Figures 3 and 4), and song perception (Figure 5). Song patterning relies on the integration of multiple sensory cues, 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 relative to a visual estimate of distance to the receiver [11]. These studies implied a single control system that upregulates pulse amplitude with perceived distance. 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 control system affects the binary pulse choice and biases pulse production toward the louder P_{fast} when far from the female (Figure 3D), and second, an analog control system upregulates a continuous variable—the amplitude of P_{fast}, but not of P_{slow}—with distance (Figure 3G). These two modes of control are most likely implemented through distinct circuits.

Female proximity cues activate the song premotor neuron P1 [25, 26], and using optogenetics, we have shown that activation of P1 biases solitary males toward singing more sine song (Figures 4B1, 4B2, 4C1, and 4C2)—which is naturally produced when the male is closest to the female [10]. 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 (Figures 4B1 and 4B2). By contrast, pIP10 is thought to be downstream of P1 [16], but its activation strongly biases males toward producing P_{fast} instead (Figures 4B3 and 4C3). Thus, rather than being a relay of P1 activation, pIP10 also actively shapes the dynamics of song mode choice (Figures 4B and S4A). Parallel descending pathways (e.g., pMP2 [25, 35]) most likely bias motor circuits toward the production of sine song or P_{slow}, but these pathways have not yet been functionally characterized due to a lack of sparse and specific genetic drivers to target these descending neurons. The activation of P1 or pIP10 in solitary males produces all three song modes but with different probabilities and latencies. This suggests that a static analysis of song production is insufficient for understanding how song is generated in *Drosophila*. Rather, the activity dynamics of different elements of the song pathway shape what is sung and when. These activity dynamics will be a product of connections between song pathway neurons, the history of activation of each neuron, and the sensory pathways that bias the output toward each of the three song modes.

ps1 is a motor neuron previously implicated in setting the frequency and amplitude of pulses [14]. By segmenting the behavior to identify multiple pulse types, we find instead that this motor neuron does not strongly affect pulse shape (Figure 4I) 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. The dynamical preference of the premotor neurons P1 and pIP10 for driving the production of particular pulse types (Figures 4B and 4C) could be mediated by differential recruitment of the ps1 muscle, which in turn could push the song production system into the P_{fast} mode by increasing thoracic rigidity [14]. Further studies into the activity of premotor neurons, motor neurons, and muscles [14, 16, 35, 36] during singing will shed additional light on the circuits that drive the three song modes.

We also found that females slow differentially to P_{slow} and P_{fast} (Figures 5A–5D) 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 5D)—females may accelerate in response to P_{fast} when it is sung at close distances because they prefer quieter songs. Further, manipulations that disrupted the natural song mode choice strongly reduced male mating success (Figures 5E–5H), suggesting that context is an important determinant of the female mating decision. These results highlight 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 [37]—will be necessary to elucidate the neural basis of the differential slowing response.

Whereas previous studies demonstrated the existence of multiple pulse types in *Drosophila* species [15, 38–40], our statistical analysis of pulse shapes in four species of the *melanogaster* group now shows that pulse shapes are conserved in these species (Figures 2H–2J and S2): *D. melanogaster*, *D. simulans*, and *D. mauritiana* all produce P_{fast} and P_{slow}, and *D. sechellia* only produces P_{slow}. The most parsiminous explanation for this pattern is that the existence of these two pulse types predates the species split and that *D. sechellia* has lost the ancestral P_{fast} type. 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 produced at different distances to the female [40, 41], just as with P_{slow} and P_{fast} in *D. melanogaster*. Notably, pIP10 activation in *D. yakuba* [41] and in *D. melanogaster* (Figure 4B3) biases the song toward the louder and higher frequency pulse type. This common neural control suggests that *D. melanogaster* P_{fast} and *D. yakuba* clack are functionally homologous, 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, by now segmenting *Drosophila melanogaster* song into three (not two, as previously thought) modes, we have revealed two major pitfalls when lumping distinct behavioral modes. First, the changes in the prevalence or the complete loss of one mode upon experimental manipulation will appear as changes in the shape of the lumped mode. For instance, the motor neuron ps1 was reported to change the frequency of pulses [14]. By discriminating two pulse types, we show that pulse frequency is relatively stable (Figure S4E) and that what changes is the prevalence of each pulse type with distance to the female (Figure 4G). The same danger looms for studies on the genetics and evolution of behavior—for instance, the loss of one pulse type can appear as a large change in overall pulse shape, even when the homologous pulse types change only little (Figure S2H). Second, lumping distinct behavioral modes may lead one to completely miss levels of neural control. For instance, previous studies showed that males produce louder pulses with increasing distance to the female [11]. Assuming the existence of a single pulse type, this implied a single, analog control mode. The discovery of two pulse modes now implies two control circuits for producing louder pulses: one for biasing toward the louder P_{fast} pulse type (Figure 3D) and one for upregulating the amplitude P_{fast}, but not of P_{slow} (Figure 3G). The existence of two pulse types also implies two functional modes of the song-pattern-generating network, potentially occupying non-overlapping sets of neurons. Premotor and motor neurons bias not only between sine and pulse but also between the two pulse types (Figure 4), suggesting that they are able to engage these different network modes as neural substrates for the distance-dependent pulse choice. The multiple hypothesized—and previously seemingly redundant—song-triggering descending neurons pIP10 [16] and pMP2 [24, 42] most likely play specific roles in biasing the downstream circuits toward producing one of the three song modes based on sensory cues. Likewise, although females respond to both pulse types, they do so in a context-dependent manner, and the effectiveness of courtship song depends on the male singing the right song mode in the right context (Figure 5). The female auditory system must therefore be able to discriminate between pulse types and to integrate acoustic information with context information to inform her behavior. All of the above examples hint at the existence of novel but still unknown circuits. This highlights the benefits of a detailed characterization of behavior for interpreting data from neural activation and

silencing experiments, recordings of neural activity, and comparisons of behavior across species.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
- METHOD DETAILS
  - Behavioral chamber
  - Optogenetic activation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Song segmentation and tracking
  - Waveform selection for segmented pulses
  - Waveform selection for unsupervised analysis of song waveforms
  - Waveform normalization
  - Dimensionality reduction
  - Waveform clustering
  - Template-based pulse type classifier
  - Gaussian Mixture Modeling
  - Female song responses
  - Effect of disrupted pulse choice on mating success
  - Characterization of pulse shapes
  - Identification of sensory cues driving pulse choice
  - Amplitude modulation plots
  - Analysis of optogenetic activation data
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and can be found with this article online at https://doi.org/10.1016/j.cub.2018.06.011.

ACKNOWLEDGMENTS

We thank Asif Ghazanfar, Joshua Shaevitz, and David Stern for comments on the manuscript and Georgia Guan for technical assistance. J.C. was supported by a postdoctoral fellowship through the Princeton Sloan-Swartz Center, F.A.R. was supported by the German Research Foundation (DFG Forschungsstipendium RO 5787/1-1), and M.M. was supported by an NSF BRAIN EAGER award, an NIH New Innovator Award (DP2 NS092378), a McKnight Scholar Award, and a Faculty Scholar Award from the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

J.C., P.C., and M.M. designed the study. J.C., P.C., F.A.R., T.D.P., D.M., D.E.A., and D.A.P. collected data. J.C., P.C., F.A.R., T.D.P., and D.E.A. analyzed data. J.C. and M.M. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 16, 2017
Revised: March 10, 2018
Accepted: June 7, 2018
Published: July 26, 2018

REFERENCES

1. Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abraira, V.E., Adams, R.P., and Datta, S.R. (2015). Mapping sub-second structure in mouse behavior. Neuron *88*, 1121–1135.
2. Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. J. R. Soc. Interface *11*, 20140672.
3. Robie, A.A., Hirokawa, J., Edwards, A.W., Umayam, L.A., Lee, A., Phillips, M.L., Card, G.M., Korff, W., Rubin, G.M., Simpson, J.H., et al. (2017). Mapping the neural substrates of behavior. Cell *170*, 393–406.e28.
4. Calhoun, A.J., and Murthy, M. (2017). Quantifying behavior to solve sensorimotor transformations: advances from worms and flies. Curr. Opin. Neurobiol. *46*, 90–98.
5. Anderson, D.J., and Perona, P. (2014). Toward a science of computational ethology. Neuron *84*, 18–31.
6. Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. Neuron *93*, 480–490.
7. Bennet-Clark, H.C., and Ewing, A.W. (1967). Stimuli provided by courtship of male *Drosophila melanogaster*. Nature *215*, 669–671.
8. Bennet-Clark, H.C., Dow, M., Manning, A., and von Schilcher, F. (1976). Courtship stimuli in Drosophila melanogaster. Behav. Genet. *6*, 93, 95.
9. Arthur, B.J., Sunayama-Morita, T., Coen, P., Murthy, M., and Stern, D.L. (2013). Multi-channel acoustic recording and automated analysis of Drosophila courtship songs. BMC Biol. *11*, 11.
10. Coen, P., Clemens, J., Weinstein, A.J., Pacheco, D.A., Deng, Y., and Murthy, M. (2014). Dynamic sensory cues shape song structure in Drosophila. Nature *507*, 233–237.
11. Coen, P., Xie, M., Clemens, J., and Murthy, M. (2016). Sensorimotor transformations underlying variability in song intensity during Drosophila courtship. Neuron *89*, 629–644.
12. Clemens, J., Girardin, C.C., Coen, P., Guan, X.J., Dickson, B.J., and Murthy, M. (2015). Connecting neural codes with behavior in the auditory system of Drosophila. Neuron *87*, 1332–1343.
13. Ding, Y., Berrocal, A., Morita, T., Longden, K.D., and Stern, D.L. (2016). Natural courtship song variation caused by an intronic retroelement in an ion channel gene. Nature *536*, 329–332.
14. Shirangi, T.R., Stern, D.L., and Truman, J.W. (2013). Motor control of Drosophila courtship song. Cell Rep *5*, 678–686.
15. Riabinina, O., Dai, M., Duke, T., and Albert, J.T. (2011). Active process mediates species-specific tuning of Drosophila ears. Curr. Biol. *21*, 658–664.
16. von Philipsborn, A.C., Liu, T., Yu, J.Y., Masser, C., Bidaye, S.S., and Dickson, B.J. (2011). Neuronal control of Drosophila courtship song. Neuron *69*, 509–522.
17. van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*, 2579–2605.
18. David, J.R., Lemeunier, F., Tsacas, L., and Yassin, A. (2007). The historical discovery of the nine species in the *Drosophila melanogaster* species subgroup. Genetics *177*, 1969–1973.
19. Cobb, M., Burnet, B., Blizard, R., and Jallon, J.-M. (1989). Courtship in *Drosophila sechellia*: its structure, functional aspects, and relationship to those of other members of the *Drosophila melanogaster* species subgroup. J. Insect Behav. *2*, 63–89.
20. Tomaru, M., and Oguma, Y. (2000). Mate choice in *Drosophila melanogaster* and *D. sechellia*: criteria and their variation depending on courtship song. Anim. Behav. *60*, 797–804.
21. Mineault, P.J., Barthelmé, S., and Pack, C.C. (2009). Improved classification images with sparse priors in a smooth basis. J. Vis. *9*, 1–24.
22. Pan, Y., Meissner, G.W., and Baker, B.S. (2012). Joint control of Drosophila male courtship behavior by motion cues and activation of male-specific P1 neurons. Proc. Natl. Acad. Sci. USA *109*, 10065–10070.

23. Hoopfer, E.D., Jung, Y., Inagaki, H.K., Rubin, G.M., and Anderson, D.J. (2015). P1 interneurons promote a persistent internal state that enhances inter-male aggression in Drosophila. eLife *4*, e11346.
24. Kimura, K., Hachiya, T., Koganezawa, M., Tazawa, T., and Yamamoto, D. (2008). Fruitless and doublesex coordinate to generate male-specific neurons that can initiate courtship. Neuron *59*, 759–769.
25. Kohatsu, S., Koganezawa, M., and Yamamoto, D. (2011). Female contact activates male-specific interneurons that trigger stereotypic courtship behavior in Drosophila. Neuron *69*, 498–508.
26. Clowney, E.J., Iguchi, S., Bussell, J.J., Scheer, E., and Ruta, V. (2015). Multimodal chemosensory circuits controlling male courtship in Drosophila. Neuron *87*, 1036–1049.
27. Kohatsu, S., and Yamamoto, D. (2015). Visually induced initiation of Drosophila innate courtship-like following pursuit is mediated by central excitatory state. Nat. Commun. *6*, 6457.
28. Kallman, B.R., Kim, H., and Scott, K. (2015). Excitation and inhibition onto central courtship neurons biases Drosophila mate choice. eLife *4*, e11188.
29. Zhang, S.X., Rogulja, D., and Crickmore, M.A. (2016). Dopaminergic circuitry underlying mating drive. Neuron *91*, 168–181.
30. Klapoetke, N.C., Murata, Y., Kim, S.S., Pulver, S.R., Birdsey-Benson, A., Cho, Y.K., Morimoto, T.K., Chuong, A.S., Carpenter, E.J., Tian, Z., et al. (2014). Independent optical excitation of distinct neural populations. Nat. Methods *11*, 338–346.
31. Baines, R.A., Uhler, J.P., Thompson, A., Sweeney, S.T., and Bate, M. (2001). Altered electrical properties in Drosophila neurons developing without synaptic transmission. J. Neurosci. *21*, 1523–1531.
32. Sweeney, S.T., Broadie, K., Keane, J., Niemann, H., and O'Kane, C.J. (1995). Targeted expression of tetanus toxin light chain in Drosophila specifically eliminates synaptic transmission and causes behavioral defects. Neuron *14*, 341–351.
33. Klibaite, U., Berman, G.J., Cande, J., Stern, D.L., and Shaevitz, J.W. (2017). An unsupervised method for quantifying the behavior of paired animals. Phys. Biol. *14*, 015006.
34. Pereira, T., Aldarondo, D., Willmore, L., Kislin, M., Wang, S.S., Murthy, M., and Shaevitz, J.W. (2018). Fast animal pose estimation using deep neural networks. bioRxiv. https://doi.org/10.1101/331181.
35. Shirangi, T.R., Wong, A.M., Truman, J.W., and Stern, D.L. (2016). Doublesex regulates the connectivity of a neural circuit controlling Drosophila male courtship song. Dev. Cell *37*, 533–544.
36. Lindsay, T., Sustar, A., and Dickinson, M. (2017). The function and organization of the motor system controlling flight maneuvers in flies. Curr. Biol. *27*, 345–358.
37. Clemens, J., Deutsch, D., Thiberge, S.Y., and Murthy, M. (2018). Shared song object detector neurons in Drosophila male and female brains drive divergent, sex-specific behaviors. bioRxiv. https://doi.org/10.1101/366765.
38. Tomaru, M., and Yamada, H. (2011). Courtship of *Drosophila*, with a special interest in courtship songs. Low Temperature Science *69*, 61–85.
39. Cowling, D.E., and Burnet, B. (1981). Courtship songs and genetic control of their acoustic characteristics in sibling species of the *Drosophila melanogaster* subgroup. Anim. Behav. *29*, 924–935.
40. Demetriades, M.C., Thackeray, J.R., and Kyriacou, C.P. (1999). Courtship song rhythms in *Drosophila yakuba*. Anim. Behav. *57*, 379–386.
41. Ding, Y., Lillvis, J.L., Cande, J., Berman, G.J., Arthur, B.J., Xu, M., Dickson, B.J., and Stern, D.L. (2017). Neural changes underlying rapid fly song evolution. bioRxiv. https://doi.org/10.1101/238147.
42. Yu, J.Y., Kanai, M.I., Demir, E., Jefferis, G.S.X.E., and Dickson, B.J. (2010). Cellular organization of the neural circuit that drives Drosophila courtship behavior. Curr. Biol. *20*, 1602–1614.
43. Kyriacou, C.P., Green, E.W., Piffer, A., and Dowse, H.B. (2017). Failure to reproduce period-dependent song cycles in Drosophila is due to poor automated pulse-detection and low-intensity courtship. Proc. Natl. Acad. Sci. USA *114*, 1970–1975.
44. van der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. J. Mach. Learn. Res. *15*, 3221–3245.
45. Larsson, M.C., Domingos, A.I., Jones, W.D., Chiappe, M.E., Amrein, H., and Vosshall, L.B. (2004). Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. Neuron *43*, 703–714.
46. Grether, M.E., Abrams, J.M., Agapite, J., White, K., and Steller, H. (1995). The head involution defective gene of Drosophila melanogaster functions in programmed cell death. Genes Dev. *9*, 1694–1708.
47. Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295–300.
48. Deng, Y., Coen, P., Sun, M., and Shaevitz, J.W. (2013). Efficient multiple object tracking using mutually repulsive active membranes. PLoS ONE *8*, e65769.
49. Stern, D.L., Clemens, J., Coen, P., Calhoun, A.J., Hogenesch, J.B., Arthur, B.J., and Murthy, M. (2017). Experimental and statistical reevaluation provides no evidence for *Drosophila* courtship song rhythms. Proc. Natl. Acad. Sci. USA *114*, 9978–9983.
50. Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Automat. Contr. *19*, 716–723.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
all-trans retinal	Sigma-Aldrich	R2500
Sigmacote	Sigma-Aldrich	SL2
Deposited Data		
Manually annotated *D. melanogaster* pulses	[43]	http://www.pnas.org/content/114/8/1970/tab-figures-data#fig-data-additional-files
Experimental Models: Organisms/Strains		
All fly strains including their complete genotypes, sources, and the figures they appear in are listed in a table in Experimental Model and Subject Details.	N/A	N/A
Software and Algorithms		
MATLAB R2017a	MathWorks	https://mathworks.com/products/matlab.html
tSNE code	[17, 44]	https://github.com/lvdmaaten/bhtsne
Fly song segmenter with pulse classifier	[9], this paper	https://github.com/murthylab/songSegmenter
Stand-alone pulse classifier	This paper	https://github.com/murthylab/pulseTypeClassifier
Pulse classification pipeline	This paper	https://murthylab.github.io/pulseTypePipeline/

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 hr of eclosion and aged for 3-7 days prior to experiments. Flies were raised at low density on a 12:12 dark:light cycle, at 25°C and 60% humidity.

Figure	Genotypes
1, 2A, 2B S1A, S1C, S2E, S3A–S3D, and 5A–5C	*D. melanogaster* NM91 males (provided by Peter Andolfatto) courting pheromone insensitive and blind (PIBL) females (GMR-hid/GMR-hid; orco-/orco- [45, 46]). *, **
2E, 2F, and S3A–S3D	*D. melanogaster* NM91 males (provided by Peter Andolfatto) courting NM91 wild type females.
S1B, S1D, 2C, 2H–2J, S2, 3A–3D, 3F, 3G, S3E–S3G, and 5A–5C	8 *D. melanogaster* strains CM07, CarM03, N30, NM91, TZ58, ZH23, ZW109 (provided by Peter Andolfatto or the Drosophila species stock center), and Canton S (lab stock). Males courting PIBL females *,**.
2D	P1-TrpA1: UAS>stop>TrpA1/+; fru^{FLP}/NP2631 (fru^{FLP} [42] provided by Barry Dickson; UAS>stop>TrpA1 was generated in [16]; NP2631 [16, 42] was obtained from the Kyoto stock center)
3E	Canton S (WT1), GMR-hid/GMR-hid;+ (blind [46]), +; orco-/orco- (Pheromone Insensitive or PI [45]) in a Canton S background, NM91 (WT2), arista cut 20 hr prior to experiment (deaf) in an NM91 background, *,**. Males court PIBL females.
2H–2J, S1B, S2A–S2D, and S2H	*D. simulans* (sim194, sim195, simNS05), *D. mauritiana* strains (mau317, mauR12, mauR61), and *D. sechellia* (sec137, sec138, secAQTNF25) (provided by P. Andolfatto). Males court females of the same strain.

(*Continued on next page*)

Continued

Figure	Genotypes
4B–4D and S4A–S4D	P1 (R71G01): w/+; GMR71G01-LexA/+; LexAop-CsChrimson/+ (GMR71G01-LexA [22] and LexAop-CsChrimson [30] were obtained from the Bloomington stock center, GMR71G01-LexA [22] was contributed to the Bloomington stock center by Gerry Rubin, and LexAop-csChrimson [30] by Vivek Jayaraman). 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) [23] provided by David Anderson; UAS-CsChrimson [30] and LexAop-GCaMP6s [47] were obtained from the Bloomington stock center (UAS-CsChrimson [30] was contributed to the stock center by Vivek Jayaraman and LexAop-GCaMP6s [47] by Douglas Kim)). pIP10: w/+; UAS>stop>CsChrimson/+; VT40556, Fru^{FLP}/+ (VT40556 [16] and Fru^{FLP} [42] were provided by Barry Dickson; UAS>stop>CsChrimson [30] was provided by Vivek Jayaraman). vPR6: w/+; UAS>stop>CsChrimson/+; VT57239, Fru^{FLP}/+ (VT57239 [16, 42] was provided by Barry Dickson).
4E and 5E	pIP10-control: +/+; fru^{FLP}/VT40556. control-Kir: R53G02AD attP40/UAS-Kir2.1. control-TNT: R53G02AD attP40/UAS-TNT. pIP10-Kir: UAS>stop>Kir2.1/+; fru^{FLP}/VT40556**. pIP10-TNT: UAS>stop>TNT/+; fru^{FLP}/VT40556**. Males court PIBL females. (R53G02AD attP40 was obtained from the Bloomington stock center and contributed by Gerry Rubin; UAS>stop>Kir2.1 was provided by Troy Shirangi; Kir2.1 and TNT where described in [31] and [32], respectively).
4F–4I, S4E, S4F, and 5F	ps1-TNT: R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4/UAS>stop>TNT [14]. ps1-control: R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4/UAS>stop>TNTinactive (R48F07-LexA, p65/LexAop2-FlpL; R73C03-GAL4 [14] was provided by Troy Shirangi; UAS>stop>TNT and UAS>stop>TNTinactive [16] were obtained from the Bloomington stock center and contributed by Barry Dickson). Males court NM91 wild type females.
5G, 5H, S5A, and S5B	pIP10: w/+; UAS>stop>CsChrimson/+; VT40556, Fru^{FLP}/+. Males court NM91 wild type females.

All '+' chromosomes come from NM91. * indicates data previously published in [10]. ** indicates data previously published in [11].

METHOD DETAILS

Behavioral chamber

Behavioral chambers were constructed as previously described [10, 11]. 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). When recording with a female, males were painted on the thorax with a white dot 20 hr 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 min of the behavioral incubator lights switching on to catch the morning activity peak. Recordings were stopped after 30 min or earlier if copulation occurred. If males did not sing in the first 5 min of the recording, the experiment was discarded. For the analysis of wing angles 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 used a chamber 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 cleaner background for video recordings and facilitated resolving the wings.

Optogenetic activation

Flies were kept for at least 3 days prior to the experiment on fly food supplanted with retinal (1 mL all-trans retinal solution (100mM in 95% ethanol) per 100 mL food).

CsChrimson [30] was activated using a 627 nm LED (Luxeon Star) at an intensity of 0.46 mW/mm^2 (driving voltage 3.2V). Stimulus period 3 s, 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 90 s (30 trials each). For the experiments of males paired with females (Figures 5G and 5H), we only used DCs 0.3 and 0.9. When optogenetic activation experiments were performed using solitary males (Figure 4), recordings were not timed to peaks in circadian activity. When paired with wild-type females (Figure 5), recordings were timed to be within 150 min of the behavioral incubator lights switching on to catch the morning activity peak. For these experiments, we used control flies that had the same genotype but were kept on regular food without added retinal. Sound recording and video were synchronized by positioning a green 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 [10, 11, 48]. 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*, 7% for *D. mauritiana* because some pairs produced substantial movement noise during courtship) (Figure S1B). Detections of the same pulse on different microphones of the chamber were consolidated and the pulse waveform was taken from the loudest channel. The song segmenter (with the parameters used in [10], not those of [9] - see [49] for a comparison of segmenter performance with these different parameters) for *D. melanogaster* detected both P_{fast} and P_{slow} at high rates without requiring any modifications (77% and 80%, respectively). We now extended the song segmenter to 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*, *D. mauritiana* we modified the segmenter by building strain specific pulse models from sets of manually annotated data (~500 pulses per strain).

The song segmenter performed poorly for recordings from *D. sechellia* with a low true positive rate of only 53% and a high false positive rate 22%. Classification of these automatically segmented pulses yielded two clusters: One that contained virtually all the true positive pulses and resembled P_{slow}, and one that was dominated by the false positives. The "pulses" from this second cluster did not occur in regular trains of fixed IPI. By contrast, the waveforms in the P_{slow} cluster almost always occurred in regular trains with clearly peaked IPI distributions. For all other species, pulses of both types occurred in regular trains with near identical IPI distributions (Figure S2C) and with similar and small false positive rates. We therefore used hand-segmented waveforms for classification of *D. sechellia* pulses. Both pulse types were detectable in hand-segmented data for all other strains used (e.g., Figure S2E).

Analysis of wing angles during song production

To compute wing angles during bouts of song, we estimated the positions of wing tips and the center line of the thorax and abdomen in each song frame. Each video frame was cropped and oriented such that the fly was always facing right, irrespective of its position in the behavioral chamber. We manually annotated single frames and then trained a convolutional neural network to automate the annotation of wings and the center line of the thorax and abdomen [34]. Wing angles were normalized to fall in the range [-180°, 180°] with 0° denoting the wing retracted state and >0° denoting wing positions away from the center line. To plot the fly body position density (Figure 2E) for each song mode, we flipped the foreground mask of each fly in each video frame such that the maximally extended wing was always on the left side of the body. For these experiments, song was segmented using custom software optimized for low signal to noise scenarios.

Waveform selection for segmented pulses

For *D. melanogaster* and *D. sechellia* pulse waveforms were extracted from the recordings by taking 25.1 ms (251 samples at 10 kHz) 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 20 kHz and extracted the same 251 samples - now corresponding to 12.55 ms - around the pulse center.

Waveform selection for unsupervised analysis of song waveforms

For the unsupervised analysis of song structure (Figure 1), 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 5 ms 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 25 ms 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* and *D. sechellia*, 0.75 ms for *D. simulans* and *D. mauritiana*).
3. **Flip** sign such that the average of the 10 samples (1.00 ms for *D. melanogaster* and *D. sechellia*, 0.5 ms for *D. simulans* and *D. mauritiana*) preceding the pulse center is positive.

Timescales 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 artifactually 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 (Figure S2G).

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 datasets (more than 100000 pulses) we chose the computationally and memory-efficient Barnes-Hut implementation [44] (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). For the unsupervised 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, right), or (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}.

When clustering only the pulses detected by the song segmenter (Figure 2A), we chose a clustering cutoff that produced two modes.

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 all waveforms yields unreliable results, likely because of the large variability of waveforms in this dataset (it includes noise). However, Gaussian Mixture Modeling of the full waveforms without dimensionality reduction also reliably clusters pulses into P_{fast} and P_{slow} (for details see below and Figure S2F).

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 (Figure S1C). The classifier performs well on the test data and generalizes well to different *D. melanogaster* strains (Figure S1D). 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).

Gaussian Mixture Modeling

As an alternative to the above methods, we clustered pulses using a Gaussian Mixture Model (GMM). GMMs fit the data as a sum of Gaussian distributions with one Gaussian per cluster in the data. For fitting, we used MATLAB's ''fitgmdist'' function, with regularization set to 0.01 to ensure convergence of the algorithm. To speed up fitting, we used random subsets of 3551 out of the 71029 pulses ($1/20^{th}$) from the wild-type strain NM91. To determine the optimal number of clusters, we fit GMMs to 20 different random subsets each with 1-6 clusters. The model likelihood is prone to overfitting, since adding more parameters to a model will almost always improve the model fit for instance if the true data distribution is not perfectly Gaussian. We therefore used Akaike's Information Criterion (AIC) [50] as a measure of error adjusted for the number of parameters (Figure S2F, right panel). When fitted using the full waveforms (all 251 points per pulse), a fit with 2 clusters reproduced the partition into P_{fast} and P_{slow} obtained with our standard procedure with >97% accuracy for all subsets (centroids shown in Figure S2F, left panel). Model likelihood and AIC were low for the GMMs fit to the full waveforms since it included many data points at the beginning and end of the pulse waveforms that only contributed noise to the data. We therefore used only time points with sufficient signal by selecting those whose variance across the full dataset exceeded 2^{-8} roughly corresponding to the 40 central time points of each pulse. Note that the results obtained were robust to the specific threshold chosen for inclusion of time points. Using this method, the AIC exhibited a minimum at 2 components for all 20 random subsets of the data indicating that the data is most parsimoniously represented using two pulse types.

Female song responses

For each pair of flies, we split the annotated song and the tracking data into 30 s time windows with 15 s overlap and computed the number of P_{fast} and P_{slow} pulses, the fraction of the time window occupied by sine song, the average female speed and the average male-female distance. Time windows that were occupied by song for <5% or >95% of the time were excluded from the analysis. Song features and speed values were z-scored for each pair to correct for individual differences in overall speed or singing. We then computed the rank correlation between each song feature and female speed either across all distances or binned by distance (2-10 mm, bin width 1 mm).

Effect of disrupted pulse choice on mating success

To test the effect of disrupting song mode choice on the male's mating success, we paired genetically manipulated males with wild-type females of the strain NM91 and scored the times to copulation up to 30 min. UAS-Kir controls did not copulate and we hence did not use this effector to assess male copulation success. For statistics, we performed a two-sided binomial test on the fraction of pairs that had copulated at the end of the scoring window.

Characterization of pulse shapes

Song pulses are transient signals and their raw magnitude spectra are relatively broad (Figures 2B and S4E). 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 [11].

Identification of sensory cues driving pulse choice

The Generalized Linear Model analysis for identifying the song features driving the choice between P_{slow} and P_{fast} was performed as in [11].

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 [11].

Pulse amplitude was taken as the range (peak to trough) of the recorded waveform, corrected for position in the chamber since actual and recorded pulse amplitude can differ due to differences in the gain between the microphones or due to the male being on or off the microphone. For correction, a position-dependent gain map of the recording chamber was estimated from all recorded pulses based on the assumption that the average amplitude of the pulse produced by the male should be the same throughout the chamber (see [11] for details). 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 bin. The first three pulses of each train were excluded from the analysis since flies do not modulate pulse amplitude at pulse train start [11]. Distance bins with <50 pulses across all flies were excluded before calculating r^2 values.

Analysis of optogenetic activation data

The song probability traces (Figures 4B and S4A) 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 σ = 64ms.

DATA AND SOFTWARE AVAILABILITY

The song segmenter with pulse classifier is available at https://github.com/murthylab/songSegmenter. The stand-alone pulse classifier is available at https://github.com/murthylab/pulseTypeClassifier. The script illustrating the analysis pipeline is available at https://murthylab.github.io/pulseTypePipeline/. Raw data are available upon reasonable request from the Lead Contact, Mala Murthy (mmurthy@princeton.edu). Pictograms of flies were modified from Benjamin de Bivort's lab web page (http://debivort.org/docs/flyclipart.pdf)