METHODS

Fly strains and rearing. For all experiments, we used 3–7-day-old virgin flies. All flies were raised in density-controlled bottles seeded with 8 males and 8 females for 4 days. Bottles were kept at 25°C and 60% relative humidity. Virgin flies were then housed individually and kept in behavioural incubators (TriTech) under 12 h/12 h light/dark cycling. Male flies were 'painted' with a small spot of opaque ultraviolet-cured glue (Norland Electronic Adhesives) on the dorsal mesothorax to facilitate identification during tracking.

All wild-type strains, except for the laboratory colony Canton-S (WT2), were supplied by the Andolfatto group at Princeton University. These strains descended from single females caught in geographically diverse locations (see Extended Data Table 1). For several experiments, we used genetically modified flies. All genetic mutations were recombined into the Canton-S genetic background; their genotypes are: PIBL (pheromone-insensitive and blind): w^+/w^+; *GMR-hid/+*; $orco^-/orco^-$; oe- (oenocyte-less and PIBL): *GMR-hid/+; PromE800-Gal4, Tub-Gal80ts/UAS-hid,UAS-stingerll;* $orco^-/orco^-$; PI (pheromone-insensitive): w^+/w^+;+/+; $orco^-/orco^-$; BL (blind): w^+/w^+; *GMR-hid/GMR-hid*; Fru-A (TrpA1 expressed in all *fruitless*-positive neurons): *UAS-TrpA1/sco;* fru^{P1}*-GAL4/TM6B*; Fru-B (TrpA1 expressed in mostly brain-specific *fruitless*-positive neurons): *tsh-Gal80/UAS-TrpA1;* fru^{P1}*-GAL4/TM6B*; P1 (TrpA1 expressed in P1 song command neurons): *NP2631/UAS-frtstopfrt-TrpA1;* fru^{P1}*-FLP/TM6*B.

We acquired *GMR-hid* flies from the Bloomington Stock Center, *PromE800-Gal4, Tub-Gal80ts; UAS-hid, UAS-stingerII* flies from J. Levine, *orco*− flies from L. Vosshall, fru^{P1}-GAL4 and fru^{P1}*-FLP* from B. Dickson, and NP2631 from the Kyoto stock centre. For arista- and wing-cutting paradigms (AC and WC), to render flies deaf (AC) or mute (WC), WT1 flies were anaesthetized with CO_2, and aristae or wings were removed, with fine forceps, over 20 h before recording. *Drosophila melanogaster* sex peptide[31] or a control peptide (SIFamide[32]) were injected into PIBL virgin females 12–24 h before recording. These flies were anaesthetized on ice during injections.

Recording apparatus. The recording chamber (Fig. 1a) is octagonal and covered in a layer of copper mesh; the chamber itself was milled from white delrin by Schmit Prototypes (Menomonie, WI). The floor is sloped at an angle such that the effective chamber diameter is 25 mm (~10 fly body lengths); this is large enough to permit a range of fly motions and interactions (Supplementary Video 1). Using sloped chamber sides facilitates fly tracking[33]. The lid was made from clear plexiglass. The floor of the chamber is tiled with 9 NR23158 microphones (Knowles Electronics) to capture male song throughout the chamber. Each microphone connects to a custom-built amplifier[30]. The open areas around the microphones increase signal-to-noise ratio. The entire chamber is enclosed inside an acrylic box—with a door for loading flies—and placed on an air table to minimize microphone noise. The camera for recording fly movement, a Unibrain 540c with a Fujinon HF16HA-1B lens, was suspended above the acrylic box, and a DigiSlave L-Ring 3200 LED light was used to illuminate the behavioural chamber. Room lighting was off during all recordings.

Courtship assay. Before the first recording of each day, the chamber was prepared in two steps. First, the underside of the cover glass was rinsed and coated with Sigmacote (Sigma-Aldrich) to prevent flies from walking on the chamber ceiling. Second, to odourize the environment, >4 flies (strain WT2) were placed inside the chamber for 10 min. Flies were loaded individually into the chamber using a custom-built aspirator. To optimize for peak fly activity, we began recordings within 150 min of the behavioural incubator lights switching on. Recordings were terminated when copulation occurred or after 30 min (classified as 'no copulation'). Recordings with no song in the first 5 min (excepting the WC paradigm) were terminated early and disregarded. Data from flies that moved < 1.5 mm per minute, or sang at low amplitudes relative to other flies of the same strain, were excluded on the basis of possible poor health (this criteria applied to 25/679 males).

Dark/light transition assay. For acute transitions between dark/light conditions (see Fig. 3c), the chamber was covered completely with BK5 Blackout Fabric (Thor Labs), and flies were loaded in darkness using a Spot 90-Lumen LED headlamp (Black Diamond) on the red light setting. After 10 min of courtship in the dark, the LED L-Ring light was switched on. If flies copulated in the dark then the trial was disregarded.

Thermogenetic activation experiments. For thermogenetic activation of Fru-A, Fru-B, and P1 neurons in males (no female in the chamber), the chamber was heated with a coil (HSTAT, BriskHeat) to 25 °C (for FruA activation), 27 °C (for FruB activation), or 32 °C (for P1 activation). In the 'Fix' condition, each fly's legs were fixed in place with UV-curable glue. Songs from these males were segmented as described above, except that the criteria for bout, pulse, and sine identification were relaxed to account for small differences between heat-activated and naturally produced song (for example, increased variability in inter-pulse interval). Song from both freely moving and fixed flies were analysed with the same modified criteria.

Data processing and analysis. All data processing and analysis was conducted in MATLAB (MathWorks).

Significance testing. Unless otherwise specified in the text, significance tests were as follows:

To determine significant differences between means, we used one-way ANOVA or Kruskal–Wallis analyses, depending on whether the distributions in question were non-normal (as identified by the Jarque-Bera test for normality). For confidence intervals, both for model selection and in Fig. 4e, non-overlap was used as a conservative estimate of significant differences. For normally distributed data, variance was estimated to be similar between groups, as required for applying ANOVA. Given variability in behaviour, we sampled a large number of flies for each wild type strain (~40) before performing any statistical tests. After observing the effect size in wild type data, we then collected between 5 and 40 flies (depending on the analysis) for other genotypes and manipulations. Blinding was not necessary for these experiments because all data analyses were automated. As all statistical tests were performed between different fly strains (with the exception of Fig. 3c, for which there is an internal control), randomization of animal groups was not necessary.

Fly tracking and song segmentation. Fly positional tracking was achieved by optimizing (for flies) an algorithm originally developed for tracking bacteria[34]. Tracking errors occurred at a low rate (~1 switch in fly identity or orientation every 25 or 12 min, respectively) and were corrected manually using custom software. We performed song segmentation using automated software[30]. We combined segmentation results from the 9 microphones into a single song with false positive and negative rates similar to those reported in ref. 30. Audio (10 kHz) and video (30 Hz) sampling frequencies were synchronized using a Master 8 (A.M.P.I.), to simultaneously drive an LED and buzzer.

Quantification of movement parameters. Males can approach the female from either side during courtship. As a result, distinctions between left/right lateral movement and anticlockwise/clockwise rotations or angles are not meaningful for the analyses we performed. For this reason, only forward velocities were allowed to take negative vales and we use absolute values for lateral movements and rotations. It should be noted that the relative contributions of forward and lateral movements to song patterning could be better established with a faster video frame rate.

Quantifying the percentage of repeated bouts. Bouts were down-sampled to 50 Hz (from 10 kHz). This temporal resolution was chosen to approximate the fastest pulse rate in fly song; events on this timescale are likely to have ethological significance. Mixed bouts (containing both sine and pulse trains) were converted into strings of ones (sine) and zeros (pulses) at the 50 Hz sampling rate. For males singing more than 100 mixed bouts of song, we calculated the number of repeated bouts to be: 1 − (number of unique binary strings/total number of binary strings). Therefore, a 'percentage of repeated bouts' of zero for a fly indicates that no two mixed bouts produced by this fly had the same pulse/sine patterning, when sampled at 50 Hz.

Analysing song statistics. For Extended Data Fig. 2a–k, only mixed bouts (those containing both sine and pulse elements) were included in analyses. Empirical joint probability density functions (PDFs) were calculated by binning data in 2 dimensions. Independent PDFs were calculated as the outer product of the empirical marginal PDFs. For bout durations (Extended Data Fig. 2e, f), only bouts shorter than 2 s were included in PDF estimation due to the sparsity of the matrix beyond this cutoff. The kernel density estimates seen in Extended Data Fig. 2l were calculated using the ksdensity function in MATLAB.

Bout triggered averages (BTAs). Given the difference between audio/video sampling rates, we resampled relevant sections of our tracking data, keeping the sampling rate the same (30 Hz) but aligning motion samples with song bout onset. For each movement parameter, we subtracted the mean between 3 s and 2.5 s before song initiation; subtracting the mean of the entire courtship session (many minutes) instead artificially increased variability, as some flies had long periods of quiescence when not singing. Mean-subtracted trials were aligned and averaged to calculate BTAs shown in Extended Data Fig. 3.

GLM implementation. For all models, data were first split into groups hereafter termed 'fitting data' (66%) and 'testing data' (34%). We fit all models using the method described in ref. 11, which imposes a sparseness-inducing prior on basis coefficients; this method penalizes redundant features. When fitting or testing models over 1,000 iterations, data were randomly subsampled each time to equalize the frequency of each event type (a common method for dealing with uneven event frequencies). For models where we sought to determine the predictive power of a model based on male lateral velocity and male forward velocity, rather than identify the most predictive feature, model selection was not performed. If a model combined data from multiple fly strains, we first *z*-scored movement features for each strain before running the model to account for any genetic differences in behaviour.

Model selection. The fitting procedure for the binomial GLM takes as inputs features (stimulus histories for movement parameters) and binary events associated

with each feature (for example, 'no song start' versus 'song start', encoded as 0 and 1). All stimulus histories are convolved with a linear filter yielding a linear prediction. The filter represents the weight for time points before each event; positive/negative deviations in the filter indicate that an increase/decrease in the stimulus is predictive of event type 1. The linear prediction is then transformed to the predicted event probability using a logistic function. We first fit separate models for each movement feature, repeating each fit 1,000 times (the data were split evenly for fitting and cross-validation) to calculate 95% confidence intervals for the deviance reduction (for example, see Extended Data Fig. 4a, b). If the deviance reduction for a single feature was greater than zero, we then re-ran models with paired inputs (for example, male forward velocity and male lateral velocity), to determine if adding a second feature improved model performance. We continued to add features until the relative improvement from an additional feature was not significantly greater than 10%.

Model performance. We convolved each of the 1,000 filters generated during fitting iterations with the testing data; the convolved data were then passed through a nonlinear logistic function to calculate a probability score for each stimulus history predicting event 1 versus event 0 (see Fig. 2a). For each test, *P*Cor values were calculated by comparing the actual outcome (whether events were type 0 or type 1) with the predicted outcome. A predicted probability >0.5 was counted as an outcome of type 1. Receiver operating characteristic (ROC) curves where generated by varying this 0.5 threshold between 0 and 1 and running the test 1,000 times for each threshold to get estimates of the false positive and true positive rates; AUC statistics were then calculated from these ROC curves. Performance plots were generated by sorting predictions in ascending order and grouping data using 5% intervals and plotting the mean prediction value against the percentage of type 1 events for each of the 20 groups. These plots serve as a visualization tool.

Song start prediction. GLMs were trained to distinguish song starts (either pulse or sine start) from potential song starts. We considered a stimulus history of 600 ms before events, for all features. For potential song starts (times when the male is not singing), we constrained selected time points to those times when Dis <8 mm and Ang2 $<60^{\circ}$, for the entire 600 ms period. We also only included actual song starts that met these criteria ($\sim$65% of song starts). These winnowing criteria were only used for models presented in Fig. 2b–d and Extended Data Fig. 8c. As well as song start predictions, we also ran a model to classify between sine and pulse starts, given that a song start had occurred; the goal of this analysis was to eliminate noise arising from unidentified behaviours (such as attempted copulation) that prevent song starts. For this, we used the same stimulus histories described above, but only for times when song starts actually occurred and without applying any winnowing criteria (Extended Data Fig. 6b, c).

Current song mode prediction. GLMs were trained to distinguish pulse mode from sine mode. We extracted movement features during 600 ms of pulse or sine song and considered only the central 300 ms (with a 150 ms buffer either side) of this time window. This prevented contamination of the data with behaviours related to transitions, song ends, or song starts. Temporal information for this model was not relevant, so the GLM was trained and tested using the mean values of each movement feature within the central 300 ms window. Results were not significantly different for models generated using 10 separate time points rather than taking the mean (data not shown). For current song mode models applied to subsets of data, the data set was split before the separation of 'fitting' and 'testing' data. To examine model performance when females were effectively stationary, we split the data into samples when female speed was $<1\ \mathrm{mm\ s^{-1}}$ or $>1\ \mathrm{mm\ s^{-1}}$ (Fig. 3f and Extended Data Fig. 9). Similarly, to examine model performance when blind males were at large distances, or facing away from the female, we split the data into samples when Dis was <5 mm or >5 mm, or Ang2 was $<60^{\circ}$ or $>60^{\circ}$, respectively (Extended Data Fig. 8d–g).

Transition prediction. GLMs were trained to distinguish transitions (pulse-to-sine or sine-to-pulse) from continual song (pulse-to-pulse or sine-to-sine). Due to the fast nature of transitions, we considered 300 ms of stimulus history—this represents half the length considered for song starts or song ends. Song was classified as a transition if there was a switch in mode immediately following this 300 ms window.

Song end prediction. GLMs were trained to distinguish actual song bout ends from potential song bout ends. We included all bout ends with at least 600 ms of song preceding the bout end. Potential song bout ends were considered to be any 600 ms period of song that did not precede the end of a bout.

Male velocity prediction. GLMs were trained to predict male forward velocity values (at time points during male song) using the preceding 600 ms of each of the other 8 movement parameters. Owing to the analogue nature of the output data (male velocities can take a wide range of values), a nonlinear logistic function was not applied to the values after convolving with linear filters. As a result, model performance is demonstrated with the comparison of all predicted and actual male forward velocities (and corresponding r^2 value for the fit), and not with a *P*Cor (Extended Data Fig. 5).

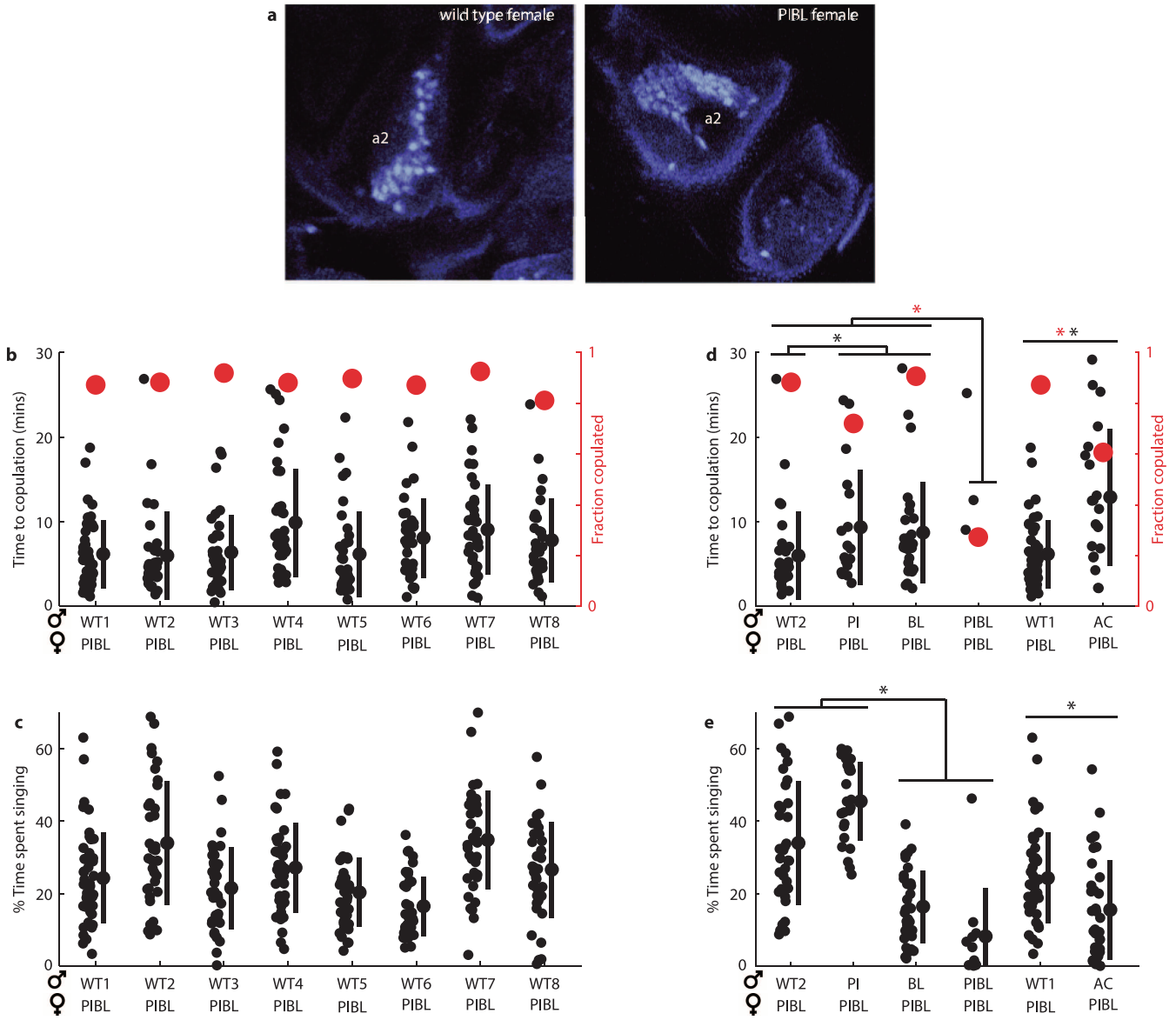
Female response to song. Female speed, along with male song data, were extracted using a sliding window of width 1 min, with 50% overlap. A range of window sizes were tested, but the results were not significantly affected unless windows became smaller than 15 s (data not shown), at which point estimates became too noisy. Time windows within 1 min of fly introduction and less than 1 min before copulation were not used, to eliminate female behaviours relating to initial introduction or copulation. A GLM was used to determine if the amount of pulse or sine song within a window predicted the mean female speed within the same window. Female speed was calculated using only time points within the window >300 ms from any song (sine or pulse), to prevent confounding effects related to male singing decisions. Only windows containing $>5\%$ (some courtship) or $<95\%$ (to allow for speed estimates) song were included.

Effect of distance on pulse/sine ratio. To eliminate the effect of velocity on song patterning (discovered in this study), we binned velocities into 1 mm $\mathrm{s^{-1}}$ windows and analysed the relationship between inter-fly distance and the pulse/sine ratio of song within each velocity bin (see Fig. 3a). Owing to the rarity of song at significant inter-fly distances, all wild-type fly strains were combined in this analysis. For each velocity window, we subtracted the mean pulse/sine ratio for all song at that velocity. Thus, negative values indicate that the pulse/sine ratio at that distance was lower than the overall ratio for that velocity. Ratios were only included if 100 song samples existed for a particular velocity/distance combination; means and standard deviations were calculated using all relative ratios that met these criteria at a particular distance. Distances below 5 mm were excluded to eliminate confounding effects of tactile interactions.

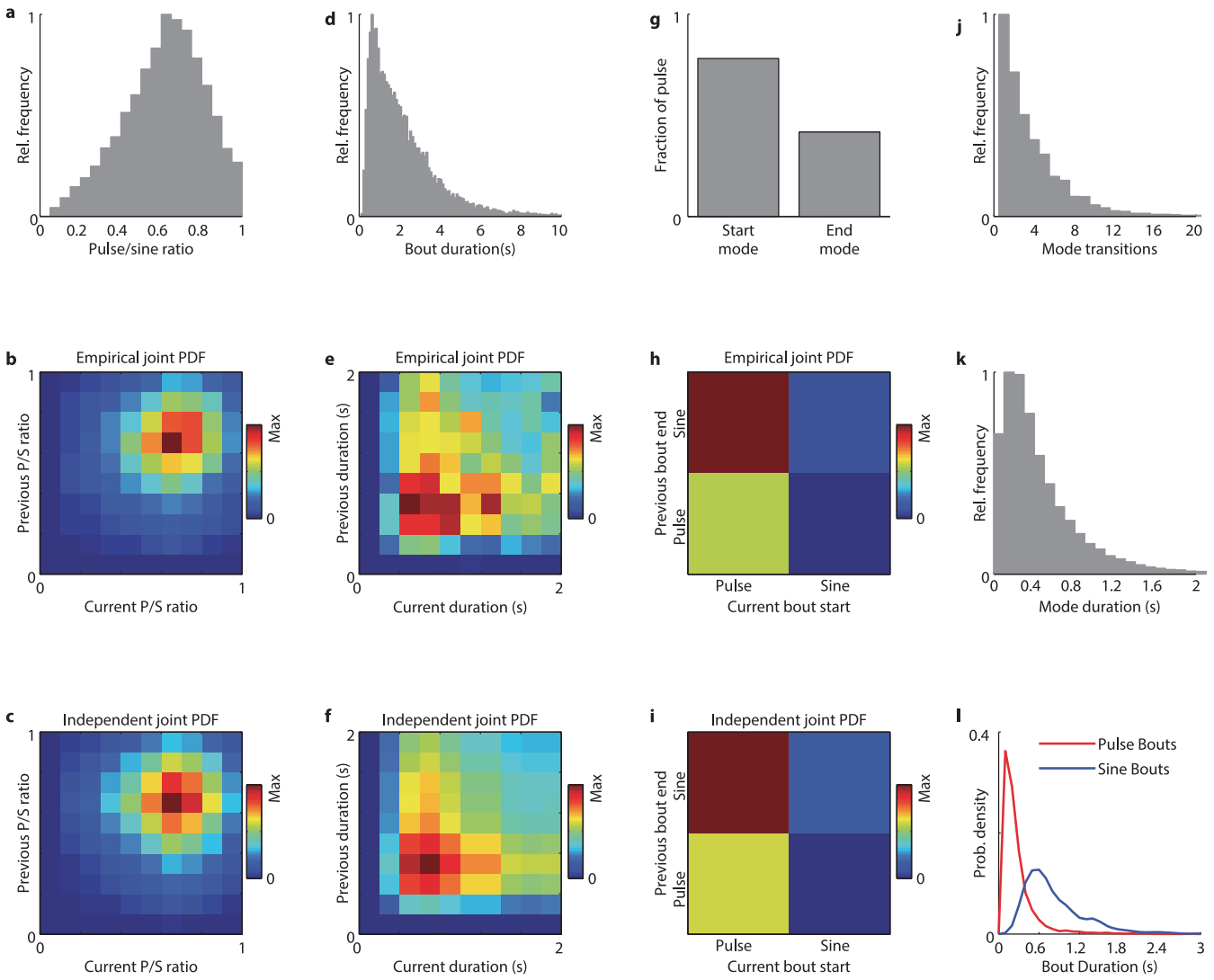
Fraction of song in pulse mode. Flies were included in this analysis (see Fig. 3b, c and j) only if they sang more than 25 bouts during the courtship period. For acute dark/light transitions (Fig. 3c), flies were required to sing at least 25 bouts during both dark and light conditions.

Female behaviour before copulation. Data from all wild type strains courting PIBL females were combined for this analysis (see Fig. 4d). Only flies that took longer than 3 min to copulate were included. The first minute of each recording was eliminated to avoid analysing behaviours associated with chamber introduction, and we examined only the 100 s preceding copulation. Each movement parameter was normalized by subtracting the mean value during 120–100 s period before copulation and dividing by the standard deviation within the final 100 s. Data were analysed similar to the BTAs (see above).

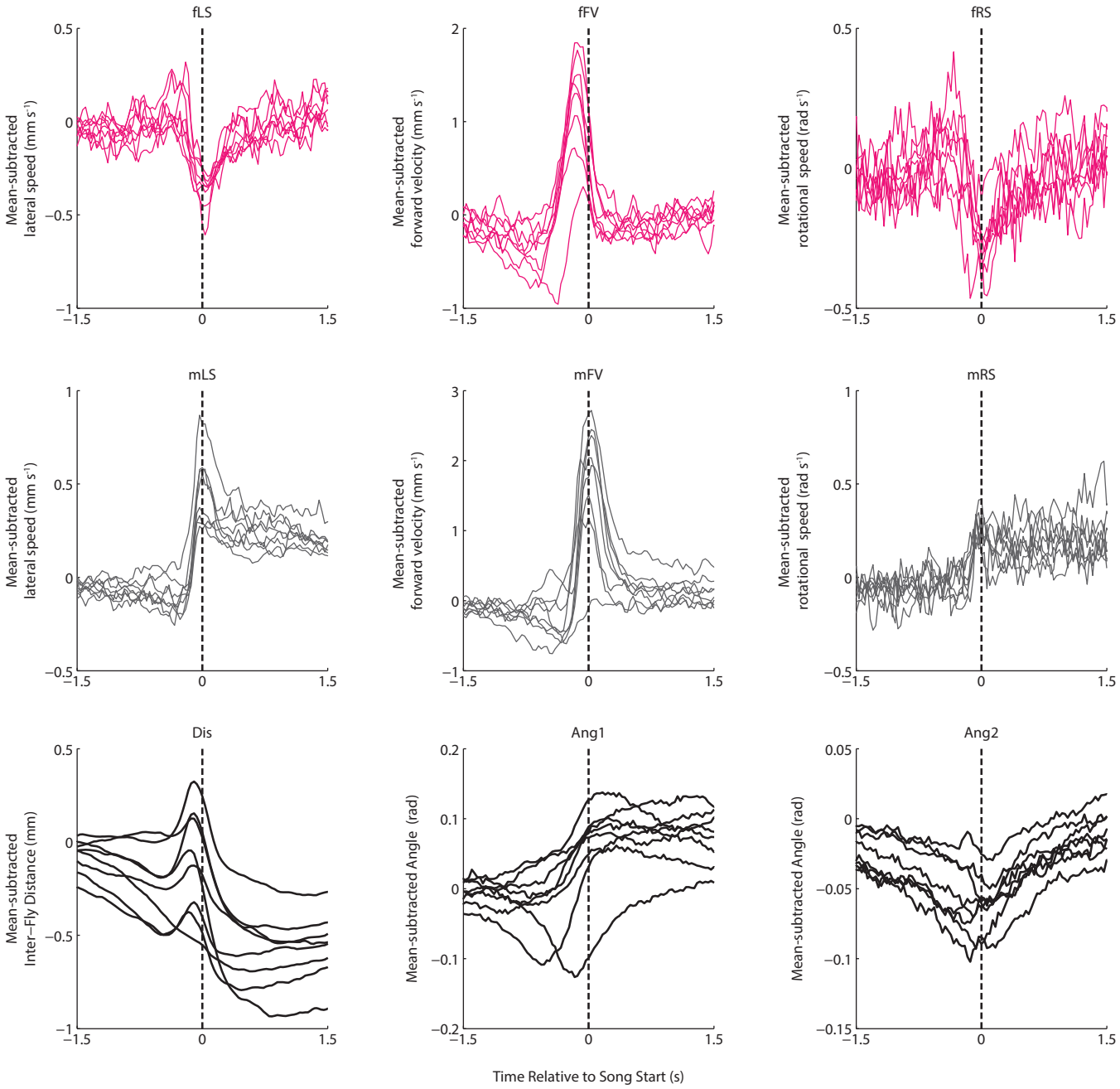
31. Yapici, N., Kim, Y. J., Ribeiro, C. & Dickson, B. J. A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour. *Nature* **451,** 33–37 (2008).
32. Verleyen, P. *et al.* SIFamide is a highly conserved neuropeptide: a comparative study in different insect species. *Biochem Biophys Res Commun.* **320,** 334–341 (2004).
33. Simon, J. C. & Dickinson, M. H. A new chamber for studying the behavior of *Drosophila*. *PLoS ONE* **5,** e8793 (2010).
34. Deng, Y., Coen, P., Sun, M. & Shaevitz, J. W. Efficient multiple object tracking using mutually repulsive active membranes. *PLoS ONE* **8,** e65769 (2013).
35. Senthilan, P. R. *et al. Drosophila* auditory organ genes and genetic hearing defects. *Cell* **150,** 1042–1054 (2012).



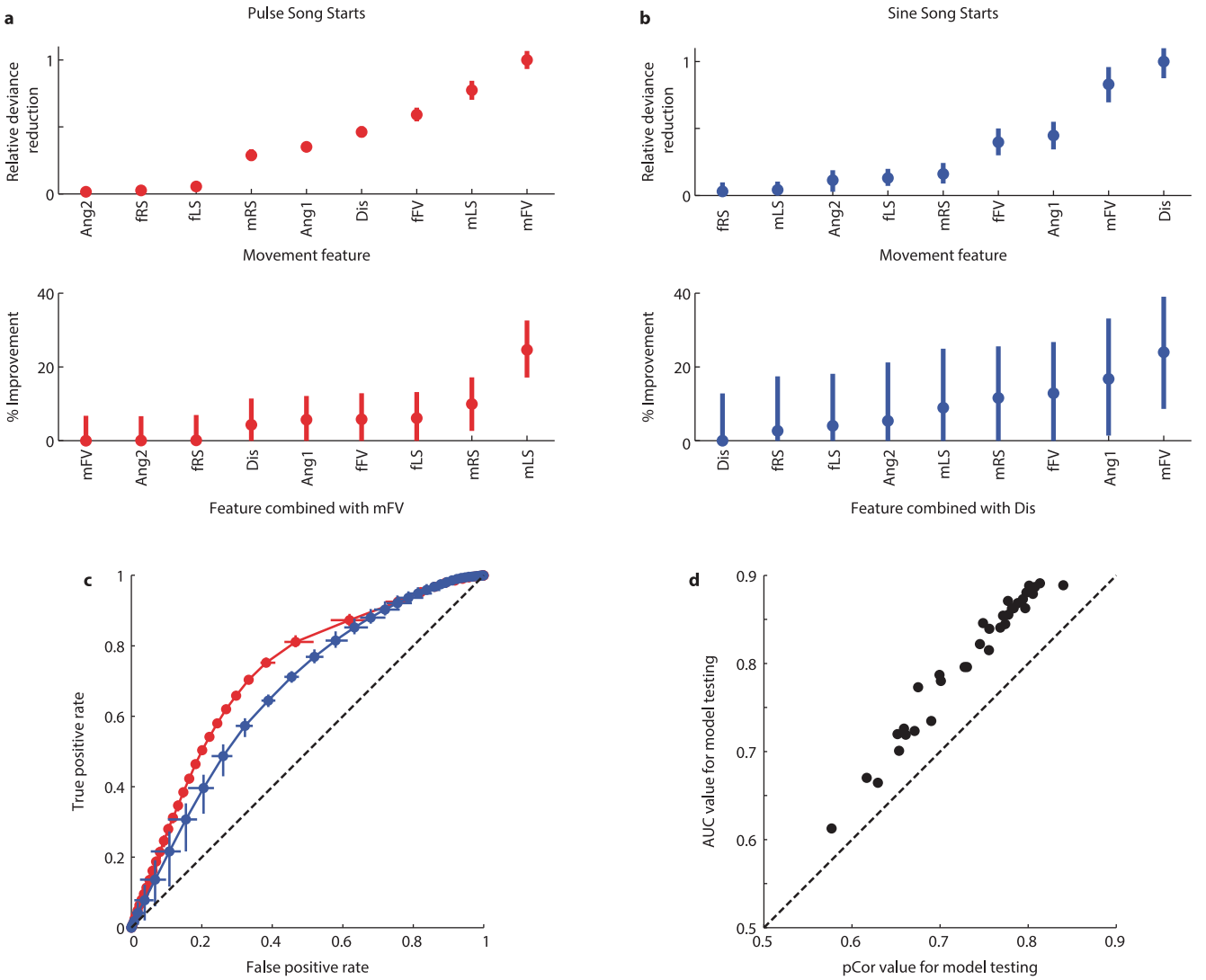
Extended Data Figure 1 | Courtship behaviour with PIBL females. **a**, Recently, genes involved in photoreceptor development have been implicated in JO neuron function[35], so we confirmed that the *GMR*-hid mutation (which induces photoreceptor apoptosis) did not affect JO neuron development. Here we show a single *z* plane image of the antenna of a wild-type (left) or PIBL (right) female fly, labelled with anti-elav (blue) to mark the nuclei of JO neurons. **b,** Time to copulation (black) and fraction of copulating pairs (red) are similar for all 8 wild-type strains. $n = 34–48$ males for each strain. **c,** The percentage of time males spent singing for all 8 wild-type strains. $n = 34–48$ males for each strain. **d,** As in **b**, but for arista cut, pheromone-insensitive, blind and PIBL males compared with wild-type strains of matched genetic background (WT2 for pheromone-insensitive, blind, and PIBL and WT1 for arista cut). $^*P < 0.05$, $n = 11–48$ males for each genotype. **e,** As in **c**, but for arista-cut, pheromone-insensitive, blind, and PIBL males compared with wild-type strains of matched genetic background. $^*P < 0.01$, $n = 11–48$ males for each genotype. **b–e,** Individual points, mean and s.d. are given for each strain/genotype.



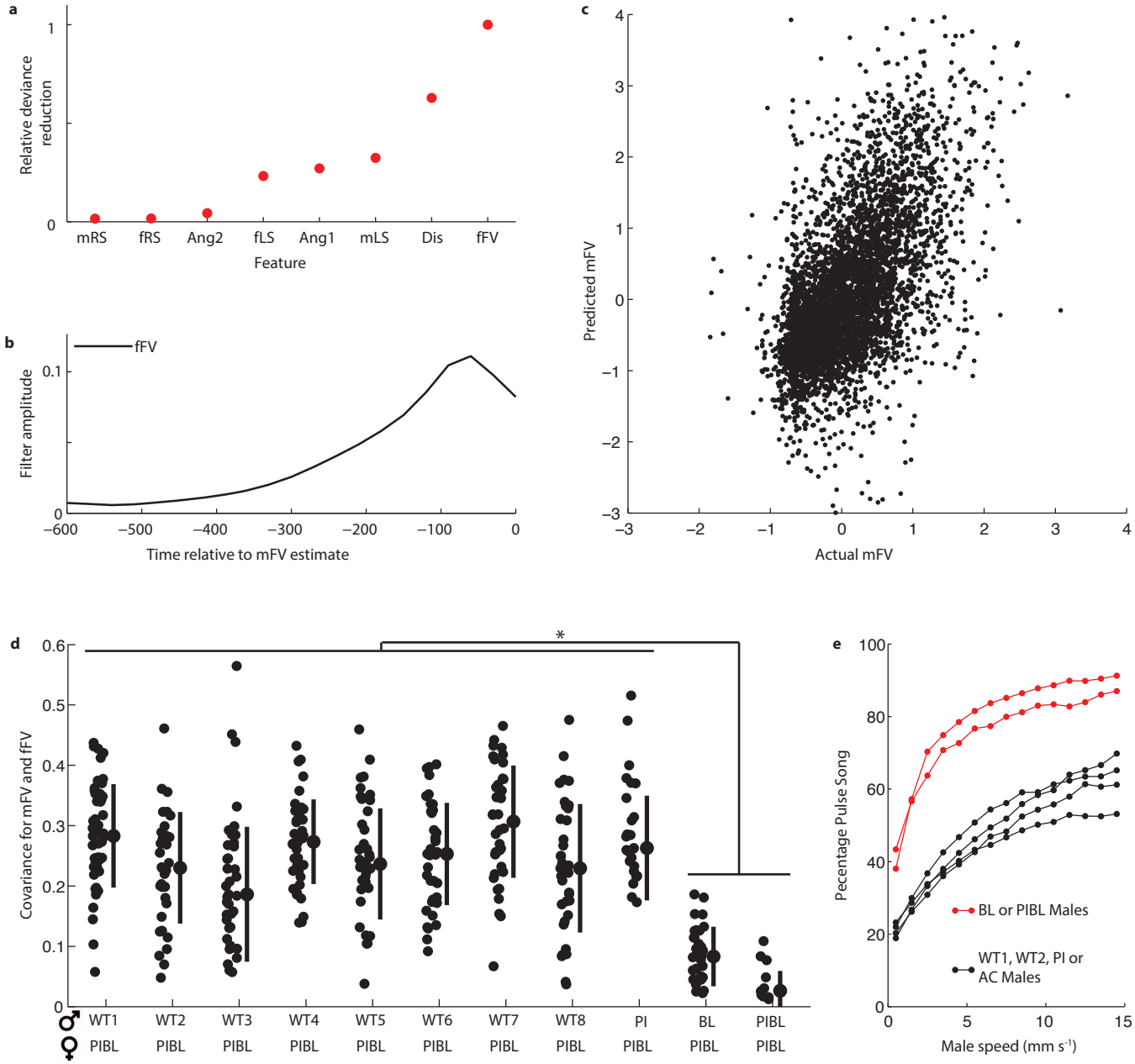
Extended Data Figure 2 | Song bout statistics for wild-type strains courting PIBL females. For all panels, data come from the 116 males singing more than 100 song bouts. **a,** Relative frequency of the pulse/sine ratio for mixed bouts (song bouts containing both sine and pulse elements). $n = 15{,}489$ bouts. **b,** The empirical joint probability density function (PDF) of pulse/sine ratios for consecutive pairs of mixed bouts (see Methods). $n = 10{,}805$ bouts. **c,** As in **b**, but the independent, rather than empirical, joint PDF. The independent joint PDF is given by multiplying the individual 1D distributions of current and previous bouts for each bin within the 2D space. The distributions in **b** and **c** are not significantly different ($P = 0.99$, two-sample Kolmogorov–Smirnov test). **d,** Relative frequency of bout durations for mixed bouts. $n = 15{,}489$ bouts. **e,** The empirical joint PDF of bout durations for consecutive pairs of mixed bouts lasting less than 2 s. $n = 3{,}535$ bouts. **f,** As in **e**, but the independent, rather than empirical, joint PDF. The distributions in **e** and **f** are not significantly different ($P = 0.19$, two-sample Kolmogorov–Smirnov test). **g,** The fraction of mixed bouts starting and ending in pulse mode. $n = 15{,}489$ bouts. **h,** The empirical joint PDF of the ending and starting modes for consecutive pairs of mixed bouts. $n = 10{,}805$ bouts. **i,** As in **h**, but the independent, rather than empirical, joint PDF. The distributions in **h** and **i** are not significantly different ($P = 0.99$, two-sample Kolmogorov–Smirnov test). **j,** Relative frequency of the number of mode (sine or pulse) transitions within mixed bouts. $n = 15{,}489$ bouts. **k,** Relative frequency of the durations of each song mode (sine or pulse) within mixed bouts. $n = 76{,}979$ song modes. **l,** Relative frequency of durations of non-mixed song bouts, comprising a single song mode. $n = 8{,}624$ or $n = 772$ for pulse only or sine only, respectively.



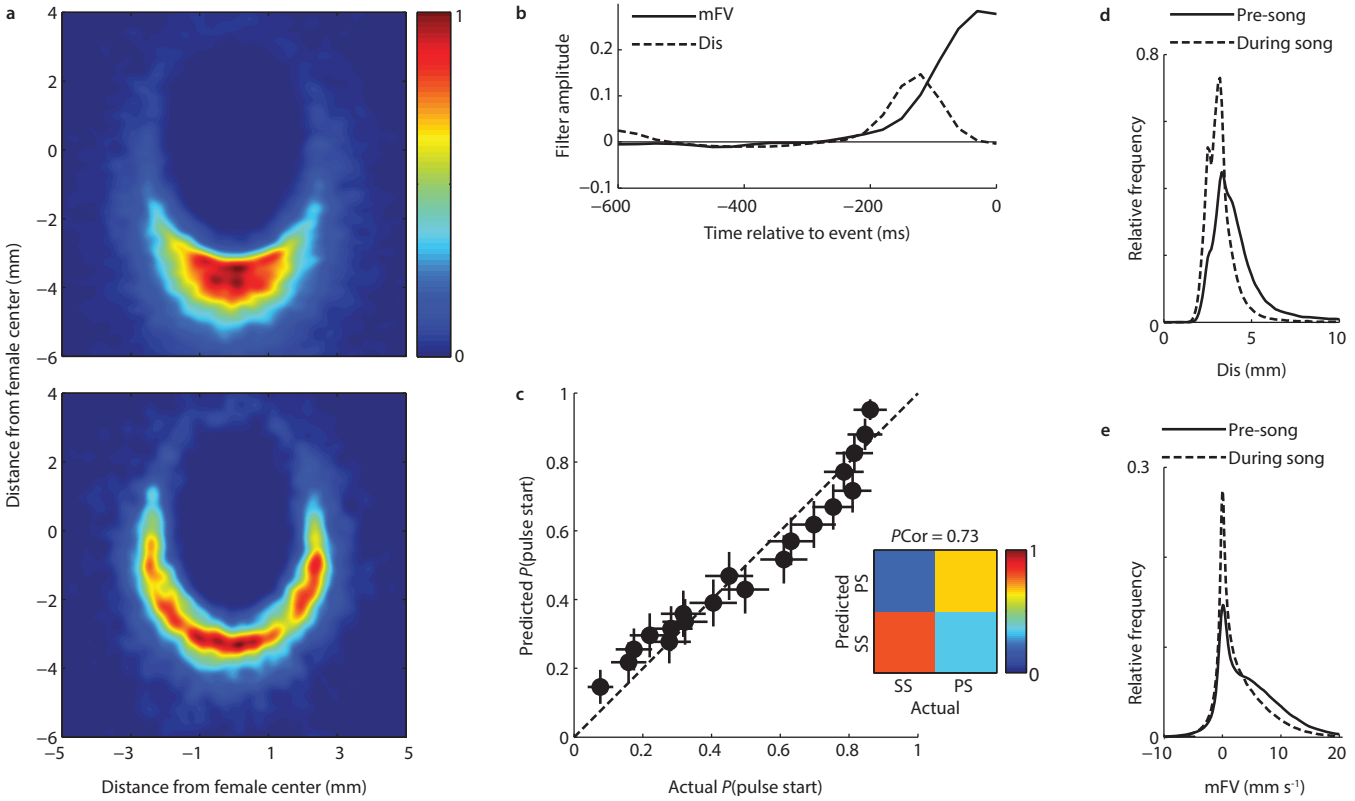
Extended Data Figure 3 | Bout triggered averages (BTAs) for all nine movement parameters for song starts. BTAs are formed similar to spike-triggered averages (STAs) for neural data. Movement parameters for each of the 8 wild-type strains were aligned to the start of song (n = 2,427–7,586 bouts from 34–48 males). All males were paired with PIBL females. Female and male parameters are coloured magenta and grey, respectively. For each trial, movement parameters were mean-subtracted before averaging (see Methods).



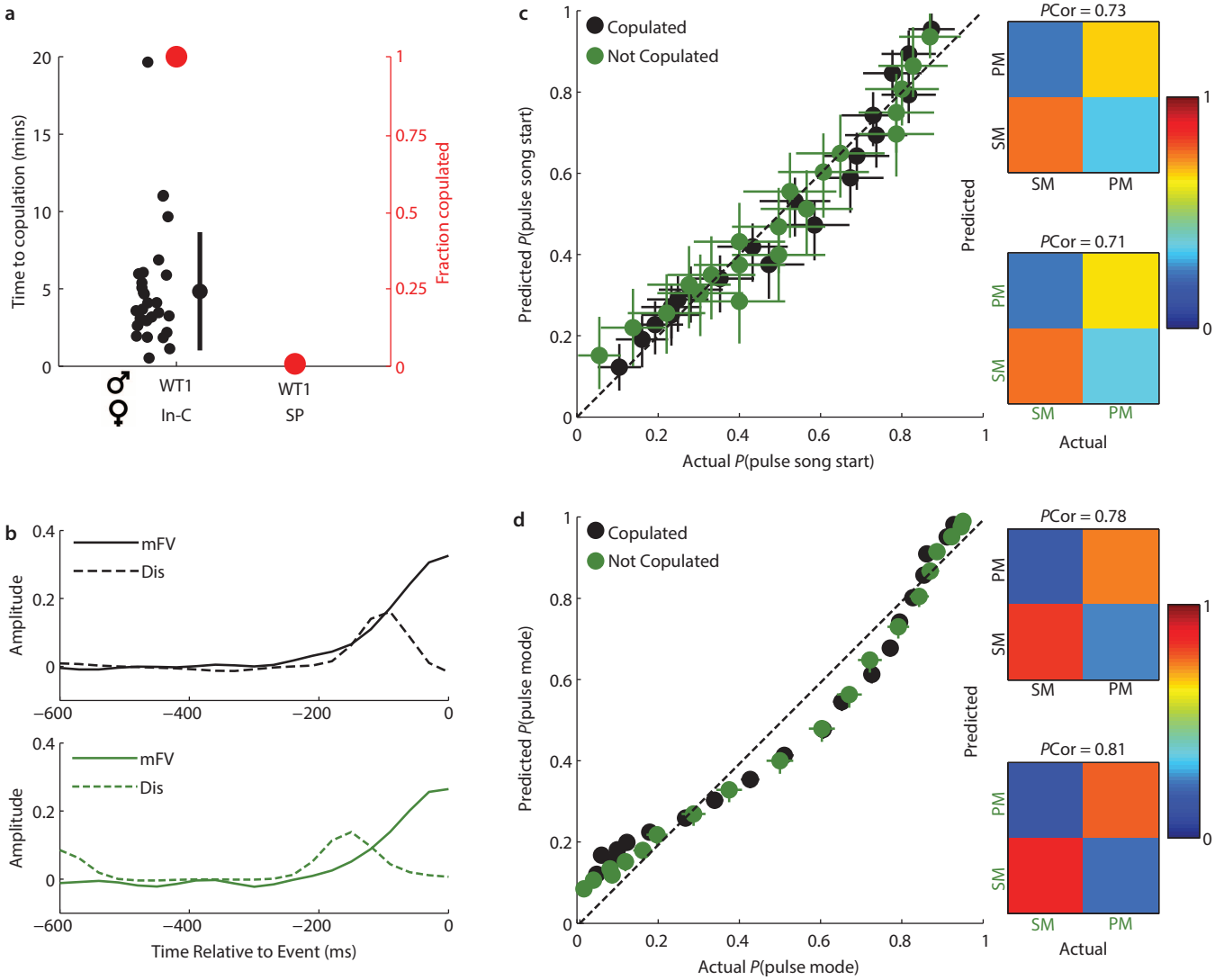
Extended Data Figure 4 | Model selection criteria examples and comparison of model performance statistics. **a,** Top, first, we train nine separate GLMs, each based on a single feature, followed by cross-validation on two-thirds of the data, with 1,000 repetitions. The single feature which gives the greatest reduction in deviance is chosen—here male forward velocity for the detection of song bouts that start in pulse mode. Bottom, a second feature is included in the model if the additional reduction in deviance improves the model by a minimum of 10%—here male lateral speed. **b,** Top, as in **a** but for song bouts that start in sine mode. Dis is selected as the most predictive feature. Bottom, as in **a**, but no second feature results in a significant model improvement, so only the one feature model is used. **c,** Receiver operating characteristic curves for GLM models designed to identify pulse (red) and sine (blue) song starts. Integrating the area under the curve (AUC) shows that both models perform significantly better than chance, for which AUC would be 0.5. AUC = 0.72 (for the pulse starts model) and 0.62 (for the sine starts model). **d,** Comparison between the *P*Cor and AUC values for every model presented in this study, showing a high correlation between the two measures: $r^2 = 0.98$. For every model tested, the *P*Cor value is a more conservative measure of performance. Error bars indicate 95% confidence intervals, although some are too small to visualize (**a**–**c**).



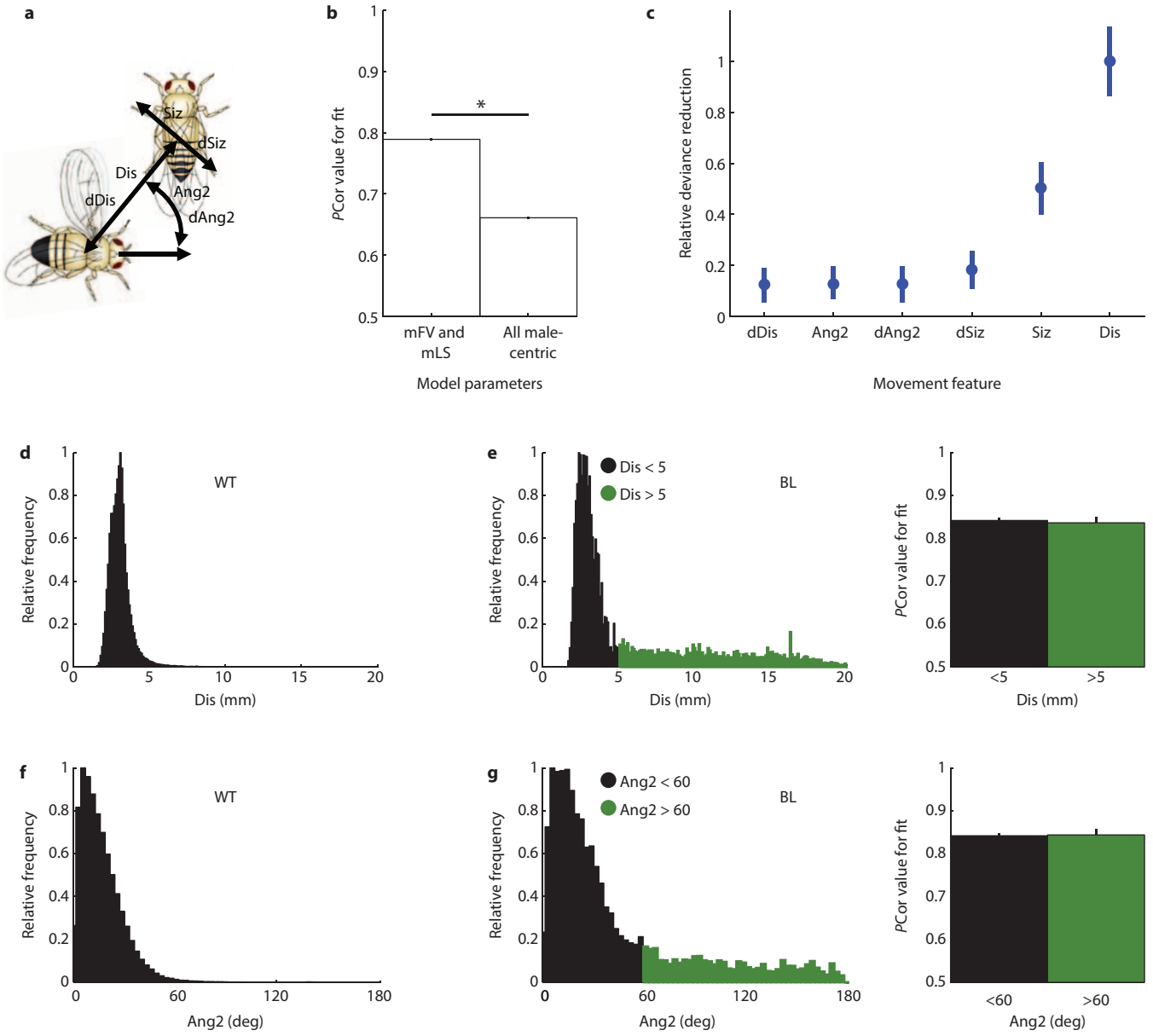
Extended Data Figure 5 | Female forward velocity changes predict male forward velocity changes in wild-type and pheromone-insensitive males, but not blind or PIBL males. **a,** Relative deviance reduction for GLMs, one for each movement feature, to predict male forward velocity at time points during song. Female forward velocity is the optimal predictor. Error bars indicate 95% confidence intervals, although some are too small to visualize. **b,** The female forward velocity linear filter is most predictive of male forward velocity values at a lag of ~60 ms. **c,** GLM performance for predicting male forward velocity based only on female forward velocity ($n = 58{,}1814$ test events from 315 pairs, $r^2 = 0.39$). Values of male forward velocity and female forward velocity were normalized such that $\mu = 0$ and s.d. $= 1$ (see Methods). A total of 1% of the data (randomly selected) is plotted here for illustrative purposes. **d,** As an estimate of the time males spent following females, we measured the maximum cross-covariance (normalized by the auto-covariance) between male and female forward velocities, $n = 11$–48 males for each strain. Perfect following behaviour, over the entire trial, would produce a value of 1. We tested all following delays between 0 and 300 ms. BL and PIBL, but not PI, males show significantly reduced following compared with all other WT strains, $^*P < 0.05$. Individual points, mean, and s.d. are given for each strain/genotype. **e,** The two blind male genotypes (blind and PIBL, red) sing a higher percentage of pulse song at all male speeds (binned to nearest mm s^{-1}) compared with wild-type males or males with other sensory manipulations (WT1, WT2, pheromone-insensitive, and arista-cut, black). In all cases, females were PIBL. Speeds > 15 mm s^{-1} were excluded owing to insufficient data. For each point, $n = 1{,}208$–15,736 samples.



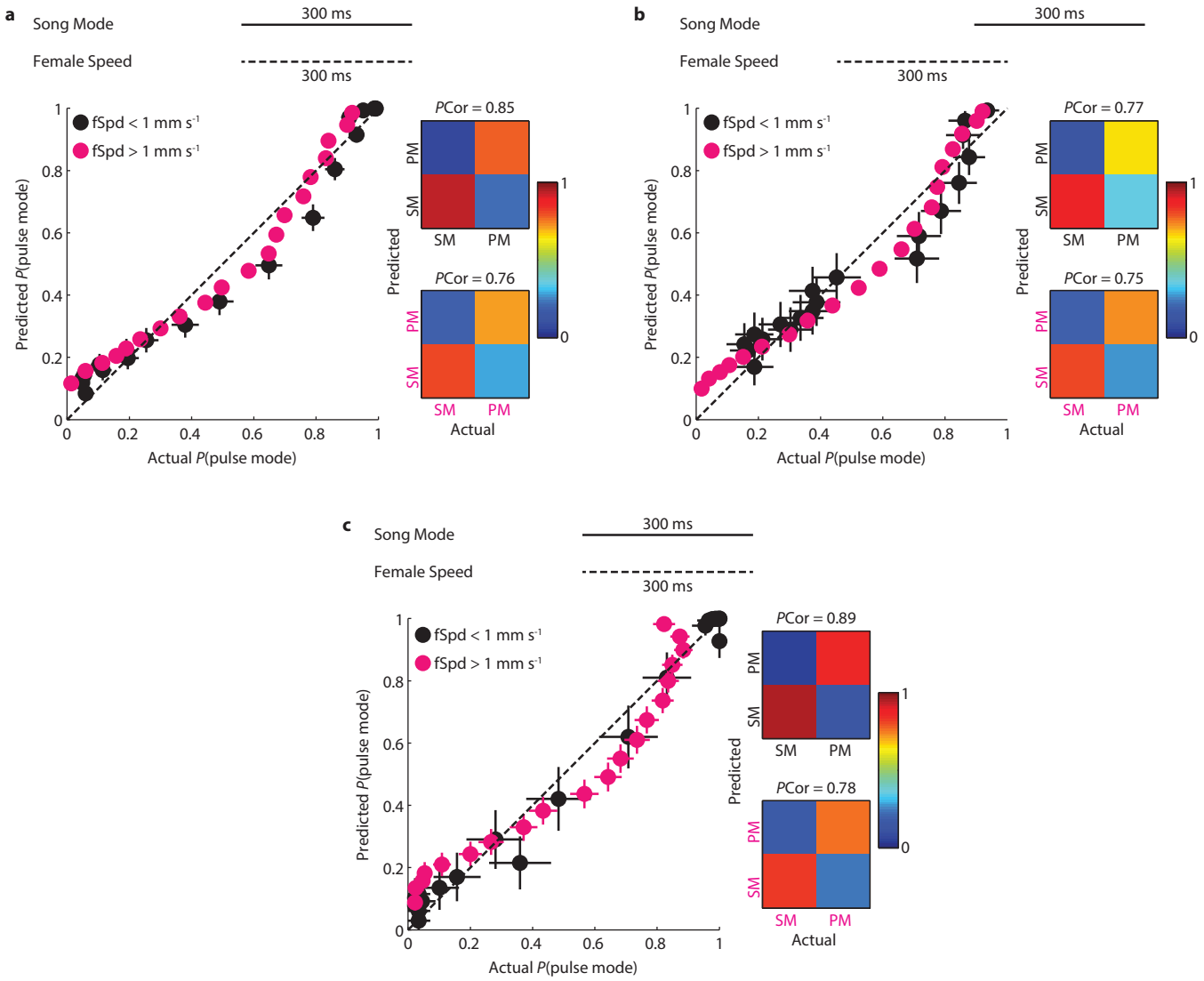
Extended Data Figure 6 | Relationships between male–female distance, male velocity and song bout starts. **a,** A two-dimensional normalized kernel density estimate of the male centre relative to the female centre (0,0) at the time of song bout initiation using combined data from all wild-type males. Males are positioned further from the female when they start a song bout in pulse mode (top, $n = 27{,}820$ bouts from 315 males) versus sine mode (bottom, $n = 5{,}749$ bouts from 315 males). **b,** Linear filters for male forward velocity and Dis, the most predictive features for the song start mode classification GLM (predicting sine song starts (SS) versus pulse song starts (PS)). **c,** GLM performance for classifying song start mode with male forward velocity and Dis filters ($PCor = 0.73$, $n = 3{,}904$ test events from 315 males). Error bars indicate 95% confidence intervals. **d,** Relative frequency distribution of Dis for periods 150 ms before the start of song bouts (solid) and during song (dashed), $n \geq 20{,}1414$ time points from 315 males. The variance in Dis is larger, by 229%, for time points before song. **e,** As in **d**, but for male forward velocity. The variance increase in male forward velocity for time points before song is 58%, much smaller than the increase observed with Dis.



Extended Data Figure 7 | Failed copulations do not result from differences in song patterning decisions. a, Time to copulation (black) and fraction of copulating pairs (red) for sex-peptide-injected (SP) or control-peptide-injected (In-C) females paired with WT1 males ($n = 28$ or 30 males). Individual data points, mean and s.d. are given for In-C females (no SP females copulated). **b,** Male forward velocity (solid) and Dis (dashed) linear filters for song start classification (predicting songs that start in pulse mode versus sine mode). The GLM is based on data from wild-type flies that copulated (top, black) or did not copulate (bottom, green). **c,** GLM performance for classifying song starts with male forward velocity and Dis filters for wild type flies that copulated (black, PCor = 0.72, $n = 2{,}490$ test events from 278 males) or did not copulate (green, PCor = 0.71, $n = 1{,}458$ test events from 37 males). **d,** GLM performance for classifying current song mode (based on mean male forward velocity and male lateral speed) for wild-type flies that copulated (black, PCor = 0.78, $n = 36{,}094$ test events from 278 males) or did not copulate (green, PCor = 0.81, $n = 17{,}666$ test events from 37 males). Error bars indicate 95% confidence intervals, although some are too small to visualize (**c, d**).



Extended Data Figure 8 | Male velocity consistently predicts the current mode of song. **a,** Male-centric features used to examine model performance. Dis and Ang2 are the same as described in Fig. 1a. Siz represents a projection of the female's body axis onto a plane perpendicular to a line joining the two fly centres (that is, the absolute value of the sine of the angle between female body axis and Dis). Thus, when Siz = 0 or 1, the female occupies a minimal or maximal region of the male's visual space respectively. dDis, dAng2 and dSiz represent the rate of change of Dis, Ang2 and Siz. **b,** Comparison of models to classify current song mode based on only male forward velocity and male lateral speed versus all 6 male-centric features combined (*$P < 0.001$). Models were tested using the same data set ($n = 55{,}464$ test events from 315 males). **c,** Deviance reduction statistics for models predicting song bouts starting in sine mode, using only male-centric features as inputs. Dis remains the most important feature (compare with Fig. 2). **d,** The distribution of Dis during song for wild-type males ($n = 338{,}238$ time points from 315 males). **e,** Left, the distribution of Dis during song for blind males ($n = 10{,}802$ time points from 33 males) is broader than for wild type. However, model performance (right) for GLMs using male forward velocity and male lateral speed to classify current song mode remains high for song samples produced at < 5 mm (black, $n = 2{,}074$ test events) or > 5 mm (green, $n = 650$ test events) from 33 males. **f,** As in **d**, but for Ang2 rather than Dis. **g,** As in **e**, but for Ang2 rather than Dis and splitting the data for Ang2 $< 60^\circ$ (black, $n = 2{,}258$ test events) or $> 60^\circ$ (green, $n = 534$ test events). Error bars indicate 95% confidence intervals, although some are too small to visualize (**b, c**, **e**, **g**).



Extended Data Figure 9 | Models to predict current song mode during times when the female is stationary. **a,** GLM performance for classifying current song mode (based on mean male forward velocity and male lateral speed) for wild-type flies. The data set was divided into samples for which the female was effectively stationary during the song sample (black, *P*Cor = 0.85, $n = 9{,}454$ test events from 315 males), and those where she was moving (magenta, *P*Cor = 0.76, $n = 46{,}204$ test events from 315 males). Model performance remains high even without any motion cues from the female. **b,** As in **a**, but the data set is divided according to a shifted estimate of female speed, 240 ms before the song sample. This matches the most predictive region of the female BTA (Extended Data Fig. 3). Model performance remains high whether the female is stationary (black, *P*Cor = 0.77, $n = 3{,}110$ test events from 315 males) or moving (magenta, *P*Cor = 0.75, $n = 44{,}314$ test events from 315 males) 240 ms before the song sample. **c,** As in **a**, but using data from pheromone-insensitive males. Male velocity remains a successful predictor even when males cannot detect pheromones and when the female is stationary (black, *P*Cor = 0.89, $n = 1{,}788$ test events from 25 males) or moving (magenta, *P*Cor = 0.78, $n = 9{,}018$ test events from 25 males) during the song sample. Error bars in all plots indicate 95% confidence intervals, although some are too small to visualize.

Extended Data Table 1 | Descriptions and acronyms for all fly strains/genotypes

Acronym	Description
WT1	Wild type from Nairobi, Kenya; collected by Andolfatto & Bachtrog (2006)
WT2	Canton-S laboratory strain
WT3	Wild type from San Diego, California; collected by Andolfatto (2006)
WT4	Wild type from Cartagena, Colombia; collected by Andolfatto (2009)
WT5	Wild type from the Netherlands; collected by Davis (2000)
WT6	Wild type from Zanzibar, Tanzania; collected by Andolfatto & Bachtrog (2006)
WT7	Wild type from Harare, Zimbabwe; collected by Begun (1993)
WT8	Wild type from Victoria Falls, Zimbabwe; collected by Ballard (2002)
PI	Genetic manipulation to remove all volatile pheromone receptors and many general olfactory receptors
BL	Genetic manipulation to remove photoreceptors
PIBL	Genetic manipulation to remove photoreceptors and volatile pheromone receptors
AC	PIBL (female) or WT1 (male) flies had aristae removed > 20 hours before recording
WC	Male flies had wings removed > 20 hours before recording
SP	PIBL females injected with *Drosophila melanogaster* sex peptide 12-20 hours before recording
In-C	PIBL females injected with *Drosophila melanogaster* SIFamide peptide 12-20 hours before recording
oe^{-}	Genetic manipulation to remove pheromone producing cells (in PIBL females)
sim	*Drosophila simulans* strain from UCSD species stock center
Fru-A	Males expressing thermosensitive TrpA1 channels in all *fruitless* positive neurons.
Fru-B	Males expressing thermosensitive TrpA1 channels in all *fruitless* positive neurons, but with suppressed expression in the ventral nerve cord via the GAL80 system.
P1	Males expressing thermosensitive TrpA1 channels sparsely in putative song command (P1) neurons