Article

Neuron

Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch

Highlights

- A novel head-fixed singing bird preparation for two-photon imaging of network activity
- Observation of skilled motor behavior in a cortical circuit
- Use of premotor activity during singing to test a range of population models
- Song-related dynamics divorced from ongoing motor movements

Authors

Michel A. Picardo, Josh Merel, Kalman A. Katlowitz, ..., Eftychios A. Pnevmatikakis, Liam Paninski, Michael A. Long

Correspondence

mlong@med.nyu.edu

In Brief

Picardo et al. use two-photon imaging and intracellular electrophysiology in the singing zebra finch to demonstrate that premotor cortical activity does not reflect ongoing song-related kinematics but, instead, appears to form an abstract population sequence during song performance.

Picardo et al., 2016, Neuron *90*, 866–876
May 18, 2016
http://dx.doi.org/10.1016/j.neuron.2016.02.016

CellPress

CellPress

Neuron
Article

Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch

Michel A. Picardo,[1,2] Josh Merel,[3,4] Kalman A. Katlowitz,[1,2] Daniela Vallentin,[1,2] Daniel E. Okobi,[1,2] Sam E. Benezra,[1,2] Rachel C. Clary,[1,2] Eftychios A. Pnevmatikakis,[3,4,5] Liam Paninski,[3,4] and Michael A. Long[1,2,*]

[1]New York University Neuroscience Institute and Department of Otolaryngology, New York University Langone Medical Center, New York, NY 10016, USA

[2]Center for Neural Science, New York University, New York, NY 10003, USA

[3]Department of Statistics and Center for Theoretical Neuroscience, Columbia University, New York, NY 10027, USA

[4]Grossman Center for the Statistics of Mind, Columbia University, New York, NY 10027, USA

[5]Simons Center for Data Analysis, Simons Foundation, New York, NY 10010, USA

*Correspondence: mlong@med.nyu.edu

http://dx.doi.org/10.1016/j.neuron.2016.02.016

SUMMARY

The zebra finch brain features a set of clearly defined and hierarchically arranged motor nuclei that are selectively responsible for producing singing behavior. One of these regions, a critical forebrain structure called HVC, contains premotor neurons that are active at precise time points during song production. However, the neural representation of this behavior at a population level remains elusive. We used two-photon microscopy to monitor ensemble activity during singing, integrating across multiple trials by adopting a Bayesian inference approach to more precisely estimate burst timing. Additionally, we examined spiking and motor-related synaptic inputs using intracellular recordings during singing. With both experimental approaches, we find that premotor events do not occur preferentially at the onsets or offsets of song syllables or at specific subsyllabic motor landmarks. These results strongly support the notion that HVC projection neurons collectively exhibit a temporal sequence during singing that is uncoupled from ongoing movements.

INTRODUCTION

Song production in the zebra finch provides an excellent opportunity to examine the processes that shape the representation of a single complex learned behavior as it progresses from dedicated higher-order centers through downstream targets, ultimately leading to the flexion of muscles needed to produce the song (Ashmore et al., 2005; Leonardo and Fee, 2005). Although the motor pathway in the songbird is composed of well identified and spatially segregated neural circuits (Nottebohm et al., 1976), the means by which this singing behavior is represented in these regions is still a matter of debate (Troyer, 2013). One controversy centers on a single forebrain area (called HVC) that plays a central role in song production (Long and Fee, 2008; Vu et al., 1994). In one view, HVC projection neurons reflect motor-related aspects of the ongoing song (Amador et al., 2013; Boari et al., 2015), similar to the coding scheme observed in the mammalian primary motor cortex (M1) (Churchland et al., 2012; Evarts, 1968; Georgopoulos et al., 1986; Todorov, 2000). In the other view, these neurons may represent relative time within the motor act (Fee et al., 2004; Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long et al., 2010) without regard to the kinematics of the vocal apparatus. Similar abstract motor representations have been shown to exist in higher-order cortical sites such as the supplementary motor area (Matsuzaka et al., 2007; Mita et al., 2009; Nachev et al., 2008; Tanji and Shima, 1994). An intermediate hypothesis, in which HVC may represent both movement and elapsed time, may also be valid and has been suggested to exist in M1 (Carpenter et al., 1999; Lu and Ashe, 2005, 2015; Matsuzaka et al., 2007). Distinguishing between these mechanisms of motor control would be a significant step forward in our understanding of how singing behavior is encoded within HVC, potentially extending to skilled behaviors in other forebrain circuits.

Technical challenges using traditional electrophysiological approaches in singing birds have prevented a clear characterization of the HVC premotor network (Amador et al., 2013; Day et al., 2013; Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long et al., 2010; Markowitz et al., 2015). Previously, song-related neural responses were collected one at a time over the span of many weeks and then aligned to singing behavior (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long et al., 2010; Okubo et al., 2015). This process is inefficient, and coding of the behavior at an ensemble level could have shifted considerably during that time (Huber et al., 2012; Okubo et al., 2015; Peters et al., 2014), potentially leading to misinterpretations that would be avoided using a method that could allow for a "snapshot" of large neural populations in a single recording session. In the rodent, in vivo two-photon imaging has enabled the study of large populations of motor cortical neurons during the performance of

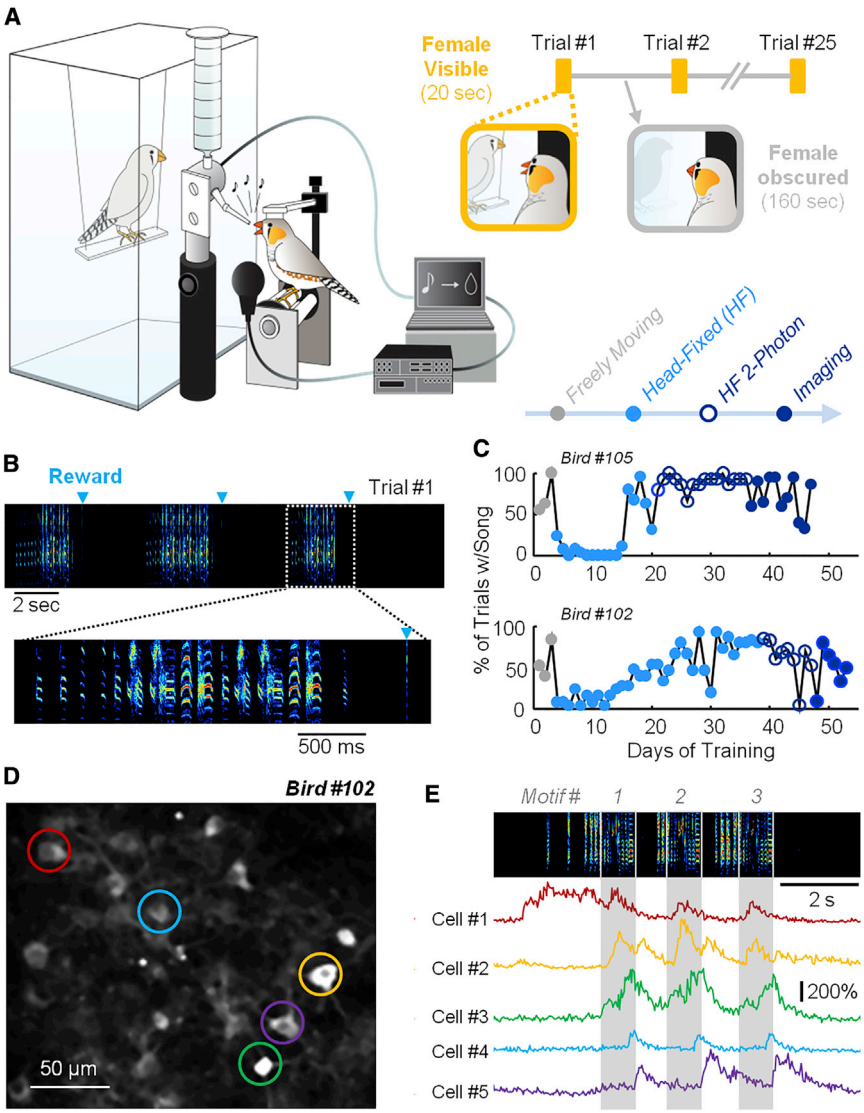

Figure 1. Imaging Neural Activity in HVC of a Head-Fixed Zebra Finch

(A) A schematic of the arena used to train head-fixed zebra finches. In our training paradigm, we used polarized glass to provide visual access to a female.

(B) An example spectrogram (frequency, 0.5–8 kHz) showing singing behavior elicited during a trial.

(C) A timeline displaying the progression of training conditions. The percentage of trials with song for each day is plotted for two birds.

(D) GCaMP6s-expressing neurons within HVC.

(E) Examples of fluorescence measurements in the neurons circled in (D). Individual song motifs are highlighted in gray.

RESULTS

Two-Photon Calcium Imaging in the Head-Fixed Singing Bird

We wanted to test the proposed hypotheses of song representation by examining the neuronal activity at a population level with two-photon microscopy (Denk et al., 1990). Recent experiments have used two-photon imaging of HVC activity in the anesthetized zebra finch to measure sensory responses within that structure (Graber et al., 2013; Peh et al., 2015). To adapt this approach to address motor dynamics, we developed a head-fixed singing bird preparation. Because we were not able to elicit head-fixed song in initial attempts (n = 6 birds), we used operant conditioning to encourage this behavior. Male zebra finches were separated from females by a glass partition whose transparency was under experimental control, and they were given visual access to females at regular intervals (Figure 1A). Males were rewarded for singing during those epochs (Figure 1B; Figures S1A–S1D) and then transitioned to a head-fixed context when song could be evoked reliably (Figure 1C; Figures S1E, S1F, and S1I). The vast majority of these birds (71 of 78) produced at least one head-fixed song motif (Table 1). Although head fixation can alter behavior and the associated neural activity for certain tasks (Ravassard et al., 2013), we found that the quality ($p = 0.27$, $F = 1.47$, one-way ANOVA) and timing ($p = 0.95$, paired t test) of singing behavior for five individual birds were not significantly different compared with song produced during free movement (Figures S1J–S1L).

A subset of the head-fixed singing birds was acclimated to the two-photon microscope (Figures S1G and S1H; Movie S1). In those individuals, we virally expressed a genetically encoded calcium indicator (GCaMP6s or GCaMP6f) (Chen et al., 2013) to report the activity of HVC neurons during singing. To calibrate the calcium indicator in vivo, we performed simultaneous

tractable behaviors (Huber et al., 2012; Li et al., 2015; Peters et al., 2014).

For these experiments, we developed a head-fixed preparation that enabled us to image populations of HVC projection neurons during singing. We then applied a new computational algorithm that enabled high-precision estimates for event timing by integrating across song trials. Additionally, we analyzed a series of intracellular recordings in which spikes and excitatory postsynaptic events were used to reflect motor-related network activity within HVC. First, we demonstrate that the rate of HVC premotor bursts during silent gaps in the song does not differ relative to epochs of active singing. We then found that HVC activity occurs continuously within the context of a syllable rather than concurrent with identified motor components of the song. Our results show that behaviorally relevant temporal sequences within HVC of the zebra finch are uncoupled from the properties of each constitutive movement, akin to higher-order cortical areas in primates (e.g., Tanji and Shima, 1994).

Table 1. Zebra Finch Training

	Freely Moving	Head-Fixed (No 2P)	Head-Fixed 2P	Head-Fixed 2P and Imaging
Number of birds trained	109	78	17	6
Number of sessions	540	2,474	187	85
Number of trials	18,500	61,750	2,834	1,405
Trials with song	7,463	11,811	1,509	756
Number of song motifs	37,112	66,025	6,800	2,868
Motifs per trial	2.0	1.1	2.4	2.0

Shown is the performance of zebra finches in each phase of the training program. Under the last two conditions, only a subpopulation of head-fixed birds was used for this study, and further training was carried out under the two-photon microscope (2P).

juxtacellular recordings and calcium imaging (Figure S2A) of bursting activity in an anesthetized and pharmacologically disinhibited zebra finch (Figures S2B–S2E). The latency from the first spike of the burst to the onset of the calcium transient was 5.9 ± 3.7 ms (n = 3 neurons from 2 birds). The standard deviation of onset timing within individual neurons was 3.5 ± 1.3 ms (e.g., Figure S2E), which supported the notion that GCaMP6 could reliably report neural activity with high temporal precision. During singing, the majority of HVC neurons exhibited one (cells 1, 4, and 5) or a few (cells 2 and 3) distinct calcium transients (Figures 1D and 1E), consistent with previous measurements taken with a head-mounted CMOS camera (Markowitz et al., 2015). For this study, we restricted our analysis to these sparsely active neurons that are likely to represent HVC premotor [$HVC_{(RA)}$] or basal ganglia-projecting [$HVC_{(X)}$] cells (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Long et al., 2010), as opposed to other observed neurons that showed sustained fluorescence during singing and probably represent inhibitory interneurons (Hahnloser et al., 2002; Kosche et al., 2015).

Increasing Temporal Resolution by Integrating Data across Behavioral Renditions

We next asked whether our approach could provide sufficient temporal resolution to test hypotheses concerning the representation of singing behavior within the HVC circuit. High temporal resolution is critical for detecting the precise onset of (~10 ms) bursting responses within HVC (Hahnloser et al., 2002; Long et al., 2010). However, the onset time of putative burst events was difficult to estimate precisely in individual trials (e.g., Figures 2A, 2C, and 2E). To address this problem, we took advantage of the stereotyped timing inherent in both the zebra finch song as well as the underlying HVC bursting activity (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007). For a given focal plane, we imaged during several renditions of the song (Figure 2B) and aligned calcium traces from multiple trials to the singing behavior (Figure 2D). Using a Bayesian method (Figures S2D and S2F–S2I), we found that combining fluorescence measurements across trials could greatly sharpen the temporal resolution of our estimates for burst onsets (Figures 2C–2G; see Experimental Procedures for details). For instance, in an example cell that had been imaged for 23 song motifs, the burst events could be detected with significantly more precision when considering all trials (mean σ_{onset} = 3.0 ms; Figure 2F) compared with estimates taken from single trials (mean σ_{onset} = 16.3 ms; Figure 2E). Across our dataset, the uncertainty of our estimate for burst onsets decreased as a function of the number of motifs analyzed (Figure 2G) and as a function of the signal-to-noise ratio of measurements from individual neurons (Figure 2H).

Using Calcium Imaging to Investigate Motor Representation at a Population Level

When the timing of HVC bursts was determined, we examined the relationship between these events and singing behavior. We asked whether bursts occurred preferentially during periods of active singing. Zebra finches produce a repeated series of vocal elements, known as "syllables," that are approximately 75–300 ms in length and are separated by shorter (15- to 100-ms) silent "gaps" during inspiration (Hartley and Suthers, 1989) (Figures 3A and 3B). Although syringeal muscles can be activated during both syllables and gaps, electromyogram recordings during song production show dynamic patterns of muscle activation that correlate with the rapidly changing acoustic features of song (Goller and Cooper, 2004; Riede and Goller, 2010; Suthers et al., 1994; Vicario, 1991a). As a result, the kinematic hypothesis would predict that HVC projection neurons would be significantly more active during syllables than during silent gaps, whereas the temporal sequence hypothesis would not anticipate a difference between these conditions. In total, we considered 294 burst events (250 cells) from five birds producing a total of 45 syllables (Figure 3; Figure S3). We found no relationship between the relative frequency of HVC projection neuron bursts and the presence of a syllable (74.0 bursts/s) or a gap (76.7 bursts/s) ($p > 0.05$, paired t test), supporting the notion that bursts are forming a continuous sequential representation throughout the song motif. Additionally, there was no temporal alignment of burst events to the onsets (Figure 3C) or offsets (Figure 3D) of syllables.

We next wanted to examine the fine structure of syllables to test whether HVC projection neuron activity is best explained by the presence of specific motor movements or by a continuous temporal sequence. Because a direct assessment of muscular activity was not feasible in our preparation, we turned to a previously established tool for estimating vocal muscle gestures from recorded songs (Amador et al., 2013; Boari et al., 2015; Figure S4). These events, called gesture-trajectory extrema (GTEs), are associated with changes in syringeal tension or subsyringeal pressure (Amador et al., 2013) and can be estimated by analyzing song structure (Boari et al., 2015). We first examined the relationship between GTEs and burst onsets (e.g., Figure 3B) and found no significant correlation between these events at any time lag in both individual birds and across the population (Figure 3E). To validate our method, we performed a simulation where burst events were placed at syllabic and subsyllabic time points and adjusted for associated experimental uncertainties as well as the inherent temporal variance suggested previously to exist between these events (Amador et al., 2013). In

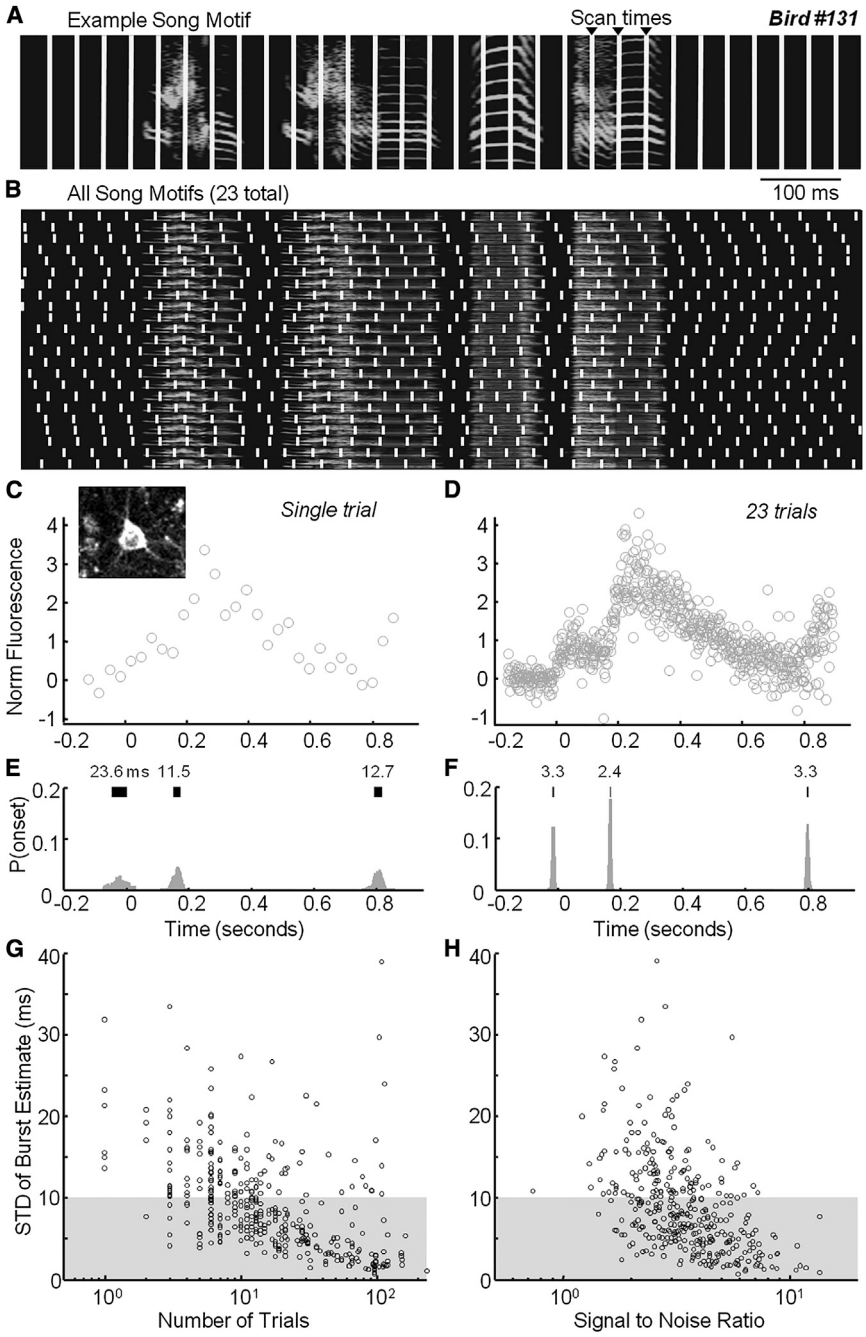

Figure 2. Integrating Imaging Data across Trials

(A and B) Spectrograms with vertical lines representing the scan times for a single motif (A) and across 23 motifs (B).

(C and D) Fluorescence transients for one neuron (inset). We show the results from a single trial (29 points) (C) as well as measurements taken across all trials (667 points) (D).

(E and F) Histograms represent an estimate of the posterior probability of burst onsets in the above panels. SDs (in milliseconds) are provided for each peak (horizontal bar, mean ± 1 SD).

(G and H) The SDs of onset estimates varied as a function of the number of trials (n = 376 bursts) (G) and the signal-to-noise ratio of the fluorescence measurements (n = 361 bursts) (H) for single neurons. The region shaded in gray indicates values less than 10 ms.

these simulated positive controls, we were able to detect a significant correlation between motor events and bursting activity (Figure 3F).

Hypothesis Testing of Network Representation

To formally test hypotheses concerning motor representations within HVC, we defined the predictions of two models: a movement model based on GTE events and a uniform distribution model in which bursts are spread evenly throughout the song (Figure 3G). When we smoothly interpolated between these two models with a likelihood-ratio test to consider a range of mixed coding schemes (see Experimental Procedures for details), we found that our data strongly favored a uniform distribution (Figure 3H), supporting the notion of a song-related temporal sequence within HVC. As before, we confirmed the efficacy of this method by finding a strong preference for the GTE model in our simulated positive control (Figure 3I). One limitation of the likelihood-ratio approach is that it tests the GTE model with no temporal offset, which was a choice that was originally motivated by the previous claim that GTE and bursts co-occur (Amador et al., 2013). Although our decision to focus on a zero latency relationship is consistent with our observation of a lack of strong correlations at a variety of time offsets (e.g., Figure 3E), we repeated the log likelihood analysis at a range of possible offsets (−30 to 30 ms). The repeated testing at various negative time offsets can help to counteract any systematic inaccuracy in our estimate of the lag between the true burst time and the calcium onset resulting from our calibration conditions (Figure 4A). The positive time offsets enable us to consider a "causal" scenario in which the accepted premotor delays (~15–20 ms) existing between HVC activity and song (Kozhevnikov and Fee, 2007) are incorporated into the analysis (Figure 4B). Across these various tests, we continued to find strong support for a uniform distribution compared with a subsyllabic kinematic model.

We considered a number of potential sources of error that may have biased our findings. First, our ability to precisely estimate burst onset times for individual neurons was variable (Figures

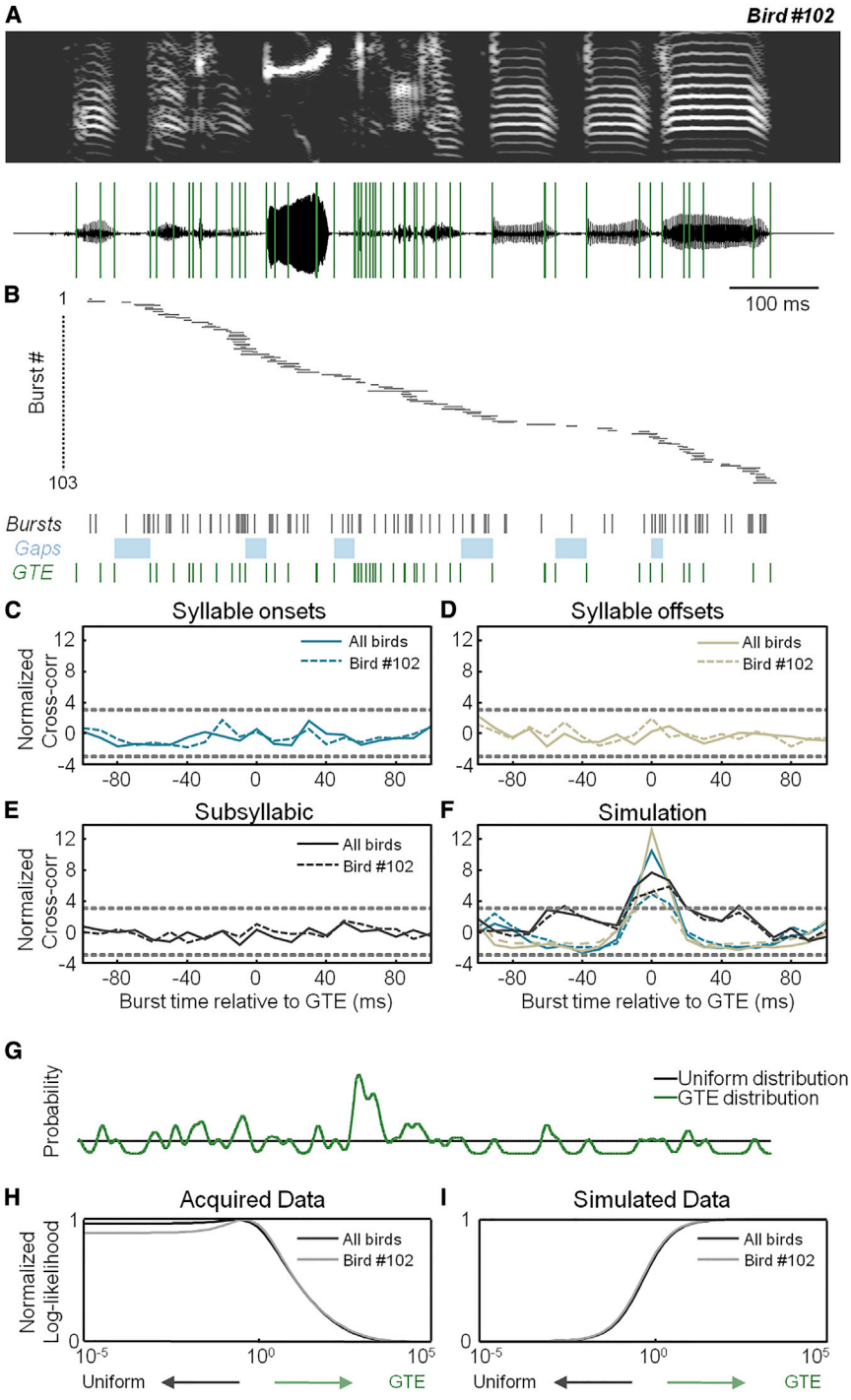

Figure 3. Timing of HVC Network Activity during Singing

(A) The spectrogram (top) and waveform (bottom) of a song motif with 58 GTE overlaid (green lines).

(B) Inferred onset times of 103 burst events from 90 HVC projection cells aligned to the song motif. Horizontal bars represent burst onset times (±1 SD). Below, we show burst onsets (black), silent gaps (light blue), and GTE times (green).

(C–E) Normalized cross-correlation (Experimental Procedures) between burst onset times and syllable onsets (C), syllable offsets (D), and subsyllabic time points (E). Negative offsets mean that the burst precedes the motor event. Dashed lines, ± 3 SD from the null model (Experimental Procedures).

(F) The normalized cross-correlation for simulated data based on syllable onsets, offsets, and subsyllabic time points.

(G) Probabilities of burst times in two alternative models (uniform, black; GTE, green).

(H) The normalized log-likelihood ratio of the predictive model as a function of the relative contributions from the uniform and GTE models, with positive values for R favoring the uniform model. The R value for bird 102 is 23.6, and that for the population of five birds is 84.2, which provides strong support for the uniform model over the GTE model. Note that the value for "c" is plotted on the x axis (Experimental Procedures).

(I) The normalized log-likelihood for simulated data. The large negative values for R (R_{102}, −45.6; R_{All}, −94.7) demonstrate that our positive control simulation strongly supports the GTE model.

2G and 2H). Although our analyses take this variability into account (see Experimental Procedures), it is possible that neurons with the highest amount of uncertainty were obscuring an existing relationship between bursts and GTEs. To address this concern, we restricted our analysis to only bursts with an onset uncertainty of less than 10 ms (186 of 294 events) and again saw no significant relationship between GTE and HVC burst timing (Figures S5A–S5D). Second, zebra finches may generate some muscle movements during silent gaps, such as "mini-breaths" (Hartley and Suthers, 1989), whose relationship with HVC activity is unclear (Andalman et al., 2011). Because we are unable to reconstruct gestures outside of phonation, we repeated the analysis after removing bursts that occurred during gaps and found no significant change in the results (Figures S5E and S5F).

Using Intracellular Recording to Investigate Motor Representation at a Population Level

To provide an additional test of our hypotheses, we next used intracellular recordings to examine the activity of song-related populations of cells within HVC. We first considered the onset of burst responses of individual projection neurons (bird 33, n = 12 cells, 17 bursts; Figures 5A and 5B) to provide an additional validation of our analytical approach in a condition with traditionally less temporal uncertainty than calcium imaging. In our analysis of 28 bursting events in 21 cells from 3 birds, we did not find any evidence of a relationship between burst times and GTE (Figure 5C), despite simulations indicating that we would be able to find a relationship if one existed (Figure 5D).

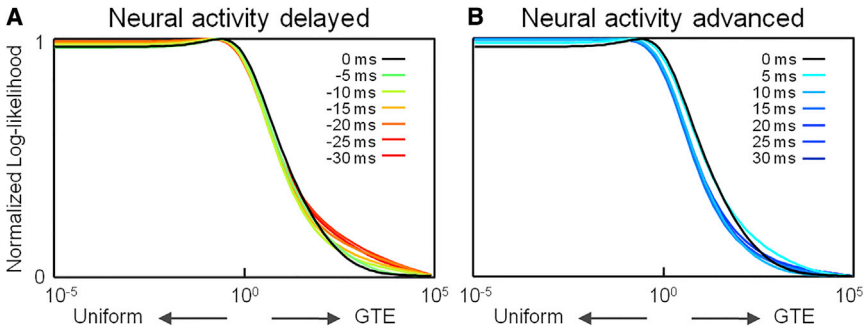

Figure 4. Testing the Temporal Relationship between Neural Activity and Movement

(A and B) Normalized log-likelihood of data acquired across all five imaged birds as a function of the relative strength of the GTE model presented with a series of additional time offsets in which burst times are either delayed (A) or advanced (B) relative to gestural time points. The values for R_{All} across all delays range from 77.4 to 118.1.

We next turned our attention to the subthreshold membrane potential of these HVC neurons. We have recently demonstrated that the series of stereotyped synaptic events onto identified HVC projection neurons is driven by motor-related inputs (Vallentin and Long, 2015). Previous paired recordings in slices of HVC have shown extensive interconnectivity within that nucleus, including links between excitatory neurons (Kosche et al., 2015; Mooney and Prather, 2005). As a result, we reasoned that identified excitatory events onto individual HVC projection neurons (Figure 6B, see example traces) could be used to simultaneously monitor the activity of multiple premotor neurons. We estimated onset times for the most consistently detected postsynaptic potentials (Figures 6A–6C; Experimental Procedures). We then repeated this analysis for all 15 HVC projection neurons within one bird (n = 187 events; Figure 6D) and across a group of three birds (n = 333 events from 26 cells; Figures S6A–S6C). Additionally, there was no connection between the incidence of synaptic onset times and the presence of either a syllable (165.7 events/s) or a gap (164.1 events/s) ($p > 0.05$, paired t test). When considering syllabic and subsyllabic structure, we also found no significant relationship between synaptic events and behavior (Figures 6E–6K; Figures S6D and S6E), including separate analyses where gaps were not included (Figures S6F and S6G) and a range of delays were introduced (Figure S6H). The possibility existed that GTE-related events may be observed in one cell class but not the other (n = 19 $HVC_{(RA)}$ and 7 $HVC_{(X)}$; Figure 6D), but we found no additional relationship when considering each cell type separately (Figure S7). Because of potential errors in our ability to detect synaptic events, we also analyzed the average membrane potential at syllable onsets and offsets as well as GTE times, and we did not find any significant deviation from zero across all neurons tested (n = 33 cells).

DISCUSSION

In our investigation, we focused on the forebrain nucleus HVC of the zebra finch because these neurons represent a well-defined skilled behavior using a sparse and reliable code. We developed a method for investigating the neural mechanisms underlying song production in a head-fixed context, which represents a rare example of a restrained social behavior (Lenschow and Brecht, 2015; Oomura et al., 1983). Using this approach, we provided population-level support for the notion that HVC projection neurons exhibit an abstract representation of elapsed time, suggesting a scheme for encoding behavior that is divorced from the kinematics of ongoing movements (Lu and Ashe, 2015; Matsuzaka et al., 2007; Tanji and Shima, 1994).

Our first step was to determine whether the population data we collected from the imaging (n = 5 birds) and electrophysiology (n = 3 birds) experiments were modulated temporally. Despite observing a large number of motor time points in each bird, obvious gaps in activity were still evident (e.g., Figures S3A and S6C). These nonuniformities may arise from a number of sources. For instance, because our recordings were spatially restricted within HVC, a sampling bias may exist when burst events from nearby cells exhibit similar timing (Graber et al., 2013; Markowitz et al., 2015; Peh et al., 2015). We next tested whether any nonuniformities could be explained by the coincident activity of neurons with motor movements. We greatly expanded on previous efforts to address this issue (Amador et al., 2013; Kozhevnikov and Fee, 2007) by quantitatively examining a number of possible relationships between HVC neuronal activity and singing. First, we find that HVC projection neurons are not preferentially active during syllables compared with silent gaps despite the clear increase in the number and complexity of motor commands during these times (Goller and Cooper, 2004; Riede and Goller, 2010; Suthers et al., 1994; Vicario, 1991a). Second, we added to this analysis by further showing that no structured neural activity existed when considering syllable onsets or offsets separately. Third, we tested whether the timing of HVC projection neuron activity is modulated by subsyllabic motor gestures, as measured by GTE (Amador et al., 2013; Boari et al., 2015).

To analyze the relationship between GTE and neural activity, we used a quantitative approach to demonstrate that the data are better described by a simple uniform model than by a gestural hypothesis. Moreover, this analysis allowed us to test for the possibility of a mixed coding scheme in which a population of neurons represents both kinematics and some other feature independent of the ongoing movement. By examining these intermediate models in our likelihood analysis, we did not see support for the idea of partial coding of GTEs by the population. It should be stated that the poor fit of our data with the GTE model does not necessarily mean that the underlying event times are distributed uniformly but, rather, that any nonuniformities cannot be explained by aspects of motor behavior tested as part of this study. The method for identifying these GTE time points was derived from models of subsyringeal pressure and syringeal muscle activity (Goller and Cooper, 2004; Méndez et al., 2010; Perl et al., 2011). However, future work should directly measure these factors in conjunction with population activity to further confirm our findings.

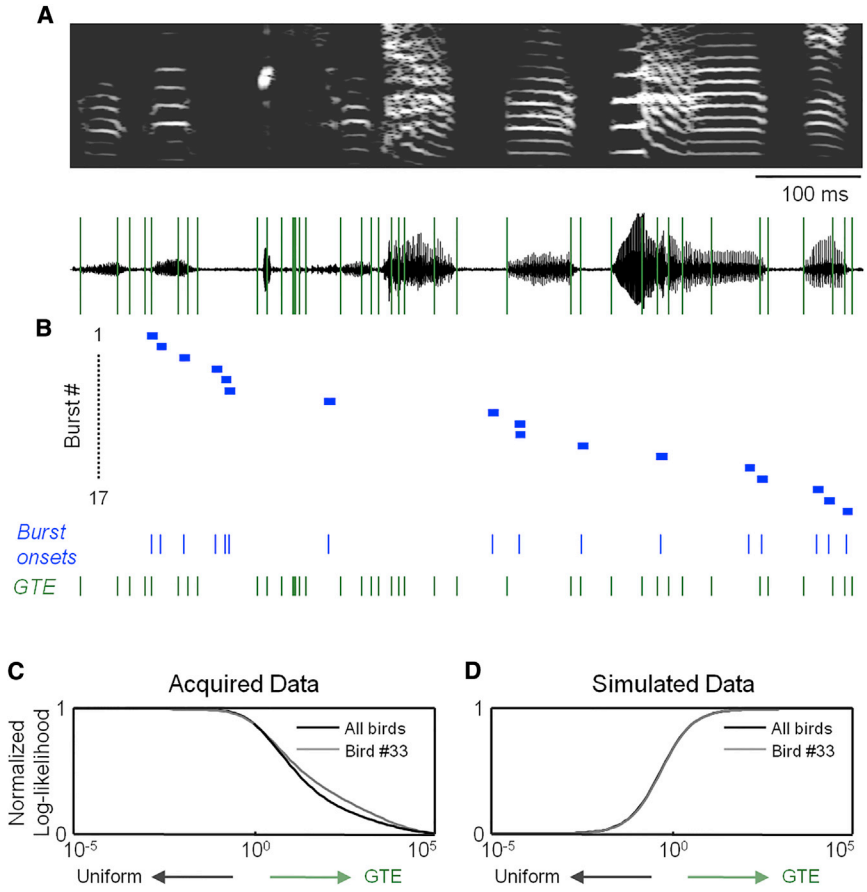

Figure 5. Timing of HVC Bursts during Singing

(A) The spectrogram (top) and waveform (bottom) of a song motif with 45 GTE overlaid (green lines).

(B) Onset times of 17 burst events measured electrophysiologically from 12 HVC projection cells aligned to the song motif. Horizontal bars, depicted with a standard width of 10 ms, are centered on burst onset times. Below, we show burst onsets (blue) and GTE times (green).

(C) The normalized log-likelihood of the acquired data as a function of the relative strength of the GTE model (R_{33}, 32; R_{All}, 49.7).

(D) The normalized log-likelihood for simulated data (R_{33}, −9.0; R_{All}, −17.7).

An important future direction is to understand the processes that enable a transformation between the abstract temporal sequence within HVC and the production of singing behavior. A candidate site that is likely to be central to this conversion is the robust nucleus of the archopallium (RA), which sits between HVC and the motor neurons that coordinate syringeal and respiratory activity (Vicario, 1991b). Each RA neuron receives synaptic input from multiple HVC neurons (Fee et al., 2004; Garst-Orozco et al., 2014), and spiking within this structure covaries with certain song features (Sober et al., 2008). Despite a wealth of single-unit recordings in this region during singing (Leonardo and Fee, 2005; Yu and Margoliash, 1996), an analysis linking RA spiking activity to GTE has not yet been performed.

Sequences similar to those described in this study have been shown to exist across a variety of brain regions, such as the hippocampus (Pastalkova et al., 2008), parietal cortex (Harvey et al., 2012), and basal ganglia (Mello et al., 2015). Within HVC, computational models (Fiete et al., 2010; Jun and Jin, 2007) and recent experimental findings (Okubo et al., 2015) suggest that this kind of population sequence may arise as part of a developmental process. Previous electrophysiological results have shown that the circuitry within HVC (Kosche et al., 2015; Mooney and Prather, 2005; Solis and Perkel, 2005) is likely to play an important role in generating these sequences (Long and Fee, 2008; Vu et al., 1994). Specifically, strong excitatory collaterals within HVC appear to underlie the propagation of network activity through a feedforward synaptic chain (Abeles, 1991; Long et al., 2010). A combination of our high temporal resolution imaging technique with emerging anatomical approaches for circuit reconstruction (Briggman et al., 2011) or electrophysiological methods (Ko et al., 2011) could help to identify mechanisms underlying sequence generation within HVC.

EXPERIMENTAL PROCEDURES

Animals

We used adult (>90 days post-hatching) male and female zebra finches that were obtained from an outside breeder and maintained in a temperature- and humidity-controlled environment with a 12 hr:12 hr light:dark schedule. All animal maintenance and experimental procedures were performed according to the guidelines established by the Institutional Animal Care and Use Committee at the New York University Langone Medical Center.

Song Detection and Reward

We recorded singing behavior with an omni-directional lavalier condenser microphone (AT803, Audio-Technica) and amplified the signal with a solid-state preamplifier (Ultragain Pro MIC2200, Behringer). Song was detected using a digital signal processor (RX8, Tucker-Davis Technologies), and custom software was written using the RPvdsEx interface (Tucker-Davis Technologies) and MATLAB. Because the zebra finch produces a song containing short gaps between syllables and motifs, singing behavior was defined as time periods in which the ratio of high-frequency power to low-frequency power (0–1 kHz and 1–7 kHz, respectively) was greater than 3 for more than 50% of a 1-s sliding window. Singing behavior was only evaluated during 20-s trial periods when the female zebra finch was visible (e.g., Figure 1A). Liquid rewards were administered using a gravity-fed line though a solenoid (NResearch) that was opened for 200 ms to dispense approximately 20 μl of water. To evaluate the detection algorithm, we manually noted false negatives and false positives, enabling us to calculate the sensitivity [true positives/(true positives + false negatives)] and positive predictive value [true positives/(true positives + false positives)].

Surgical Procedures

All surgical procedures were performed under isoflurane anesthesia (1%–3% in oxygen) following established guidelines. Prior to the onset of training, a small (3.9 × 4.55 × 1.3 mm) stainless steel head plate with two threaded holes was affixed with dental acrylic over the inner leaflet of the skull, with the

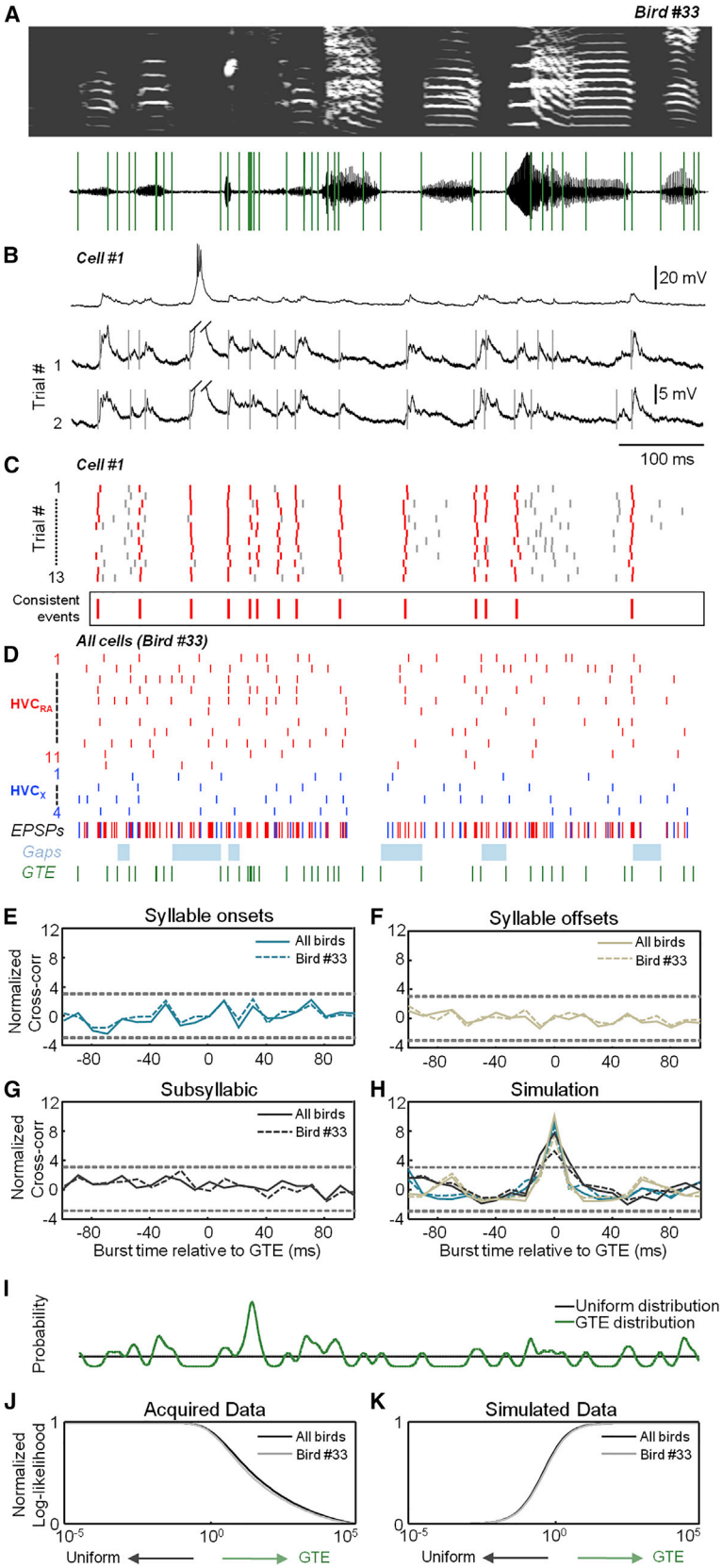

Figure 6. Timing of HVC Synaptic Inputs during Singing

(A and B) A sonogram and GTE time points (A) for the bird shown in Figure 5 and song-related intracellular recordings from an $HVC_{(RA)}$ neuron aligned to the song motif (B). The traces below have been truncated (oblique lines). Vertical gray lines indicate identified synaptic events.

(C) All excitatory postsynaptic potentials (EPSPs) detected across 13 song motifs. Consistently detected events (Experimental Procedures) are indicated with red lines (mean values below).

(D) The consensus synaptic event times for 11 $HVC_{(RA)}$ and 4 $HVC_{(X)}$ neurons (red and blue, respectively) recorded in the same bird. Below, we show EPSP onsets (red or blue, as above), silent gaps (light blue), and GTE times (green).

(E–G) Normalized cross-correlation (Experimental Procedures) between burst onset times and syllable onsets (E), syllable offsets (F), and subsyllabic time points (G). Dashed lines, ±3 SD from the null model.

(H) Normalized cross-correlation for simulated data based on syllable onsets, offsets, and subsyllabic time points.

(I) Probabilities of burst times in two alternative models (uniform, black; GTE, green).

(J) The normalized log-likelihood of the acquired data as a function of the relative strength of the GTE model (R_{33}, 240.9; R_{All}, 493.9).

(K) The normalized log-likelihood for simulated data (R_{33}, −97.8; R_{All}, −157.7).

posterior edge approximately 5 mm rostral to the bifurcation of the sagittal sinus. Cranial window surgery and viral injection were performed when the zebra finch demonstrated the ability to sing reliably under the head-fixed condition (Figure 1; Figure S1). HVC was identified electrophysiologically during surgery using a bipolar stimulating electrode placed into RA (Long et al., 2010). In most birds, we injected AAV9.Syn.GCaMP6s.WPRE.SV40 (Penn Vector Core) into HVC using a beveled pipette (opening diameter, 30 μm; length of bevel, 100 μm). In two birds, we injected a 1:1 mix of AAV9.CamKII0.4.Cre.SV40 and either AAV9.CAG.Flex.GCaMP6f.WPRE.SV40 (bird 192) or AAV9.CAG.Flex.GCaMP6s.WPRE.SV40 (bird 193). We performed three to six injections (30–100 nl/site) separated by approximately 300 μm using an oil-based pressure injection system (Nanoject II, Drummond Scientific). At the end of the injection procedure, we sealed a 3-mm-diameter circular coverglass (#1 thickness, Warner Instruments) with Kwik-Sil adhesive (WPI) and fixed the edge of the glass with cyanoacrylate. We also cemented a black plastic ring (inner diameter, 5 mm; outer diameter, 7.5 mm) around the cranial window to prevent light contamination during imaging.

Two-Photon Imaging

We used a customized movable objective microscope (Sutter Instrument) to scan our field of view (frame rate, 28.8 Hz unless noted otherwise) with a resonant system (Thorlabs) and ScanImage 4.2 software (Pologruto et al., 2003). All imaging was done using a 16× water immersion objective (Nikon) with a numerical aperture of 0.8 and a working distance of 3 mm. To protect the microscope from external light contamination, we wrapped it with a light-attenuating material. Additionally, we fit a black balloon to the tip of the objective on one end and the black ring surrounding the optical window on the other (Dombeck et al., 2010). The excitation source was a mode-locked Ti:sapphire laser (Chameleon, Coherent) tuned at 920 nm and controlled by a Pockels cell (Conoptics 302RM). Fluorescent light was detected using a GaAsP photomultiplier tube (H10770PA-40 PMT module, Hamamatsu) and a wide detection path (2-in collection lens).

Two-Photon Targeted Electrophysiological Recordings

In vivo electrophysiological recordings of HVC neurons expressing GCaMP6s were performed in isoflurane-anesthetized zebra finches. For these birds, we did not affix the cranial window to the brain with Kwik-Sil. Rather, the glass coverslip was sealed at the edges with cyanoacrylate, and a small opening (~300 μm in diameter) was produced near the center of the glass using a carbide burr. Borosilicate glass pipettes were fabricated on a horizontal puller (Sutter Instrument) with impedance values in the range of 4–5 MΩ and filled with an solution of K-gluconate (150 mM) and 5 μM of Red Alexa 594 (Molecular Probes, Invitrogen) to visualize the pipette under the microscope. Because HVC projection neurons produce only infrequent bursts outside of the context of singing (Hahnloser et al., 2002; Long et al., 2010), we applied 0.1 mM of the $GABA_A$ receptor antagonist gabazine (Sigma) to the surface of the craniotomy to increase the rate of spontaneous bursting activity (Mooney and Prather, 2005). Signals were recorded using a Neurodata IR183A single-channel amplifier (Cygnus Technology) and custom MATLAB acquisition software. Data were low pass-filtered at 5 kHz and digitized with a National Instruments digital-to-analog converter (acquisition rate, 40 kHz). For these experiments, the soma was scanned at 52 Hz, and these data were aligned with spiking activity.

Intracellular Recordings

Intracellular recordings during singing were carried out as described previously (Vallentin and Long, 2015). Briefly, a motorized intracellular microdrive was installed on the head of the zebra finch. For antidromic identification of projection neurons, we implanted a bipolar stimulating electrode into the RA and/or area X. Sharp electrodes with an impedance of 70–130 MΩ were backfilled with 3 M potassium acetate and inserted into HVC. Acceptable recordings were defined as having a spike height of more than 40 mV, a resting membrane potential more hyperpolarized than −50 mV, and a total recording duration of more than 3 min. When stable recordings were achieved, a female bird zebra finch was presented to elicit directed singing. Neurons with fewer than two song motifs were excluded from our analysis.

Additional experimental procedures related to analytical methods can be found in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.02.016.

AUTHOR CONTRIBUTIONS

M.P., D.O., and M.L. designed the experiments. M.L., M.P., J.M., L.P., and K.K. wrote the paper. M.P., D.V., S.B., R.C., and K.K. collected the data. M.P., E.P., J.M., and K.K. analyzed the data. J.M., E.P., L.P., and K.K. created relevant computational tools for analysis.

ACKNOWLEDGMENTS

This research was supported by the NIH (R01NS075044), the New York Stem Cell Foundation, the Rita Allen Foundation, the Esther A. and Joseph Klingenstein Foundation, the Simons Foundation (Global Brain Initiative), DARPA N66001-15-C-4032 (SIMPLEX), and EMBO (ALTF 1608-2013). We thank Aimee Chow, Celine Cammarata, and Brandon Robinson for technical assistance and help with further development of our training protocol. We thank Florin Albeanu and Simon Peron for assistance with two-photon imaging and Jeff Gauthier and Arnaud Malvache for help with analysis. We thank Florin Albeanu, Dmitriy Aronov, Brenton Cooper, Kishore Kuchibhotla, and John Long for comments on earlier versions of this manuscript. We thank Gabriel Mindlin and Marcos Trevisan for assistance with identifying GTE time points. We acknowledge the GENIE Program and the Janelia Farm Research Campus; specifically, Vivek Jayaraman, Ph.D.; Rex A. Kerr, Ph.D.; Douglas S. Kim, Ph.D.; Loren L. Looger, Ph.D.; and Karel Svoboda, Ph.D. from the GENIE Project, Janelia Farm Research Campus, Howard Hughes Medical Institute.

Received: December 5, 2015
Revised: January 14, 2016
Accepted: February 4, 2016
Published: May 18, 2016

REFERENCES

Abeles, M. (1991). Corticonics: Neural Circuits of the Cerebral Cortex (Cambridge University Press).

Amador, A., Perl, Y.S., Mindlin, G.B., and Margoliash, D. (2013). Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature *495*, 59–64.

Andalman, A.S., Foerster, J.N., and Fee, M.S. (2011). Control of vocal and respiratory patterns in birdsong: dissection of forebrain and brainstem mechanisms using temperature. PLoS ONE *6*, e25461.

Ashmore, R.C., Wild, J.M., and Schmidt, M.F. (2005). Brainstem and forebrain contributions to the generation of learned motor behaviors for song. J. Neurosci. *25*, 8543–8554.

Boari, S., Sanz Perl, Y., Amador, A., Margoliash, D., and Mindlin, G.B. (2015). Automatic reconstruction of physiological gestures used in a model of birdsong production. J. Neurophysiol. *114*, 2912–2922.

Briggman, K.L., Helmstaedter, M., and Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. Nature *471*, 183–188.

Carpenter, A.F., Georgopoulos, A.P., and Pellizzer, G. (1999). Motor cortical encoding of serial order in a context-recall task. Science *283*, 1752–1757.

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.

Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature *487*, 51–56.

Day, N.F., Terleski, K.L., Nykamp, D.Q., and Nick, T.A. (2013). Directed functional connectivity matures with motor learning in a cortical pattern generator. J. Neurophysiol. *109*, 913–923.

Denk, W., Strickler, J.H., and Webb, W.W. (1990). Two-photon laser scanning fluorescence microscopy. Science *248*, 73–76.

Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., and Tank, D.W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nat. Neurosci. *13*, 1433–1440.

Evarts, E.V. (1968). Relation of pyramidal tract activity to force exerted during voluntary movement. J. Neurophysiol. *31*, 14–27.

Fee, M.S., Kozhevnikov, A.A., and Hahnloser, R.H. (2004). Neural mechanisms of vocal sequence generation in the songbird. Ann. N Y Acad. Sci. *1016*, 153–170.

Fiete, I.R., Senn, W., Wang, C.Z., and Hahnloser, R.H. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. Neuron *65*, 563–576.

Garst-Orozco, J., Babadi, B., and Ölveczky, B.P. (2014). A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. eLife *3*, e03697.

Georgopoulos, A.P., Schwartz, A.B., and Kettner, R.E. (1986). Neuronal population coding of movement direction. Science *233*, 1416–1419.

Goller, F., and Cooper, B.G. (2004). Peripheral motor dynamics of song production in the zebra finch. Ann. N Y Acad. Sci. *1016*, 130–152.

Graber, M.H., Helmchen, F., and Hahnloser, R.H. (2013). Activity in a premotor cortical nucleus of zebra finches is locally organized and exhibits auditory selectivity in neurons but not in glia. PLoS ONE *8*, e81177.

Hahnloser, R.H., Kozhevnikov, A.A., and Fee, M.S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature *419*, 65–70.

Hartley, R.S., and Suthers, R.A. (1989). Airflow and pressure during canary song: direct evidence for mini-breaths. J. Comp. Physiol. A *165*, 15–26.

Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature *484*, 62–68.

Huber, D., Gutnisky, D.A., Peron, S., O'Connor, D.H., Wiegert, J.S., Tian, L., Oertner, T.G., Looger, L.L., and Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. Nature *484*, 473–478.

Jun, J.K., and Jin, D.Z. (2007). Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. PLoS ONE *2*, e723.

Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjöström, P.J., and Mrsic-Flogel, T.D. (2011). Functional specificity of local synaptic connections in neocortical networks. Nature *473*, 87–91.

Kosche, G., Vallentin, D., and Long, M.A. (2015). Interplay of inhibition and excitation shapes a premotor neural sequence. J. Neurosci. *35*, 1217–1227.

Kozhevnikov, A.A., and Fee, M.S. (2007). Singing-related activity of identified HVC neurons in the zebra finch. J. Neurophysiol. *97*, 4271–4283.

Lenschow, C., and Brecht, M. (2015). Barrel cortex membrane potential dynamics in social touch. Neuron *85*, 718–725.

Leonardo, A., and Fee, M.S. (2005). Ensemble coding of vocal control in birdsong. J. Neurosci. *25*, 652–661.

Li, N., Chen, T.W., Guo, Z.V., Gerfen, C.R., and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. Nature *519*, 51–56.

Long, M.A., and Fee, M.S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature *456*, 189–194.

Long, M.A., Jin, D.Z., and Fee, M.S. (2010). Support for a synaptic chain model of neuronal sequence generation. Nature *468*, 394–399.

Lu, X., and Ashe, J. (2005). Anticipatory activity in primary motor cortex codes memorized movement sequences. Neuron *45*, 967–973.

Lu, X., and Ashe, J. (2015). Dynamic reorganization of neural activity in motor cortex during new sequence production. Eur. J. Neurosci. *42*, 2172–2178.

Markowitz, J.E., Liberti, W.A., 3rd, Guitchounts, G., Velho, T., Lois, C., and Gardner, T.J. (2015). Mesoscopic patterns of neural activity support songbird cortical sequences. PLoS Biol. *13*, e1002158.

Matsuzaka, Y., Picard, N., and Strick, P.L. (2007). Skill representation in the primary motor cortex after long-term practice. J. Neurophysiol. *97*, 1819–1832.

Mello, G.B., Soares, S., and Paton, J.J. (2015). A scalable population code for time in the striatum. Curr. Biol. *25*, 1113–1122.

Méndez, J.M., Dall'Asén, A.G., Cooper, B.G., and Goller, F. (2010). Acquisition of an acoustic template leads to refinement of song motor gestures. J. Neurophysiol. *104*, 984–993.

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. Nat. Neurosci. *12*, 502–507.

Mooney, R., and Prather, J.F. (2005). The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. J. Neurosci. *25*, 1952–1964.

Nachev, P., Coulthard, E., Jäger, H.R., Kennard, C., and Husain, M. (2008). Enantiomorphic normalization of focally lesioned brains. Neuroimage *39*, 1215–1226.

Nottebohm, F., Stokes, T.M., and Leonard, C.M. (1976). Central control of song in the canary, Serinus canarius. J. Comp. Neurol. *165*, 457–486.

Okubo, T.S., Mackevicius, E.L., Payne, H.L., Lynch, G.F., and Fee, M.S. (2015). Growth and splitting of neural sequences in songbird vocal development. Nature *528*, 352–357.

Oomura, Y., Yoshimatsu, H., and Aou, S. (1983). Medial preoptic and hypothalamic neuronal activity during sexual behavior of the male monkey. Brain Res. *266*, 340–343.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. Science *321*, 1322–1327.

Peh, W.Y., Roberts, T.F., and Mooney, R. (2015). Imaging auditory representations of song and syllables in populations of sensorimotor neurons essential to vocal communication. J. Neurosci. *35*, 5589–5605.

Perl, Y.S., Arneodo, E.M., Amador, A., Goller, F., and Mindlin, G.B. (2011). Reconstruction of physiological instructions from Zebra finch song. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *84*, 051909.

Peters, A.J., Chen, S.X., and Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. Nature *510*, 263–267.

Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. Biomed. Eng. Online *2*, 13.

Ravassard, P., Kees, A., Willers, B., Ho, D., Aharoni, D., Cushman, J., Aghajan, Z.M., and Mehta, M.R. (2013). Multisensory control of hippocampal spatiotemporal selectivity. Science *340*, 1342–1346.

Riede, T., and Goller, F. (2010). Functional morphology of the sound-generating labia in the syrinx of two songbird species. J. Anat. *216*, 23–36.

Sober, S.J., Wohlgemuth, M.J., and Brainard, M.S. (2008). Central contributions to acoustic variation in birdsong. J. Neurosci. *28*, 10370–10379.

Solis, M.M., and Perkel, D.J. (2005). Rhythmic activity in a forebrain vocal control nucleus in vitro. J. Neurosci. *25*, 2811–2822.

Suthers, R.A., Goller, F., and Hartley, R.S. (1994). Motor dynamics of song production by mimic thrushes. J. Neurobiol. *25*, 917–936.

Tanji, J., and Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. Nature *371*, 413–416.

Todorov, E. (2000). Direct cortical control of muscle activation in voluntary arm movements: a model. Nat. Neurosci. *3*, 391–398.

Troyer, T.W. (2013). Neuroscience: The units of a song. Nature *495*, 56–57.

Vallentin, D., and Long, M.A. (2015). Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. J. Neurosci. *35*, 299–307.

Vicario, D.S. (1991a). Contributions of syringeal muscles to respiration and vocalization in the zebra finch. J. Neurobiol. *22*, 63–73.

Vicario, D.S. (1991b). Organization of the zebra finch song control system: II. Functional organization of outputs from nucleus Robustus archistriatalis. J. Comp. Neurol. *309*, 486–494.

Vu, E.T., Mazurek, M.E., and Kuo, Y.C. (1994). Identification of a forebrain motor programming network for the learned song of zebra finches. J. Neurosci. *14*, 6924–6934.

Yu, A.C., and Margoliash, D. (1996). Temporal hierarchical control of singing in birds. Science *273*, 1871–1875.