Performance evaluation is a critical feature of motor learning. In the vocal system, it requires the integration of auditory feedback signals with vocal motor commands. The network activity that supports such integration is unknown, but it has been proposed that vocal performance evaluation occurs offline. Recording from NIf, a sensorimotor structure in the avian song system, we show that short bouts of singing in adult male zebra finches (Taeniopygia guttata) induce persistent increases in firing activity and coherent oscillations in the fast gamma range (90–150 Hz). Single units are strongly phase locked to these oscillations, which can last up to 30 s, often outlasting vocal activity by an order of magnitude. In other systems, oscillations often are triggered by events or behavioral tasks but rarely outlast the event that triggered them by more than 1 s. The present observations are the longest reported gamma oscillations triggered by an isolated behavioral event. In mammals, gamma oscillations have been associated with memory consolidation and are hypothesized to facilitate communication between brain regions. We suggest that the timing and persistent nature of NIf’s fast gamma oscillations make them well suited to facilitate the integration of auditory and vocal motor traces associated with vocal performance evaluation.

Introduction
Synchronized network oscillations in the gamma (30–80 Hz) and fast gamma (90–150 Hz) range are observed in many different areas of the brain (Schoffelen et al., 2005; Siegel et al., 2008) and appear to be an emergent property of neural circuits (Fries, 2009). These oscillations have been observed in both vertebrate and invertebrate species (Neuenschwander et al., 1996; Engel and Singer, 2001) and are associated with a wide range of processes including sensory processing (Cardin et al., 2009), motor preparation (Schoffelen et al., 2005; Uhlhaas et al., 2009), selective attention (Fries et al., 2001; Buehmann and Deco, 2008), and memory consolidation (Fell et al., 2001; Axmacher et al., 2006). Despite the rich literature on gamma oscillations in mammals, relatively little is known regarding their role and properties in birds (Neuenschwander et al., 1996; Low et al., 2008).

In songbirds, auditory feedback is critically necessary for song learning and maintenance (Leonardo and Konishi, 1999; Lombardino and Nottebohm, 2000). To shape and maintain vocal output so that it matches a stored template of song, vocal motor patterns are evaluated through comparisons with auditory feedback. Neurons have been observed in auditory areas that are sensitive to auditory feedback perturbations during singing (Keller and Hahnloser, 2009). However, similar sensitivity to perturbations of auditory feedback has not been observed in vocal motor nuclei during singing (Konishi, 2004; Kozhevnikov and Fee, 2007; but see Sakata and Brainard, 2008). This has prompted some investigators to suggest that consolidation of premotor vocal commands with auditory feedback may occur offline (Margoliash and Schmidt, 2010). The most likely timeframes for such offline performance evaluation would be during sleep, where replay of premotor patterns is known to occur (Dave and Margoliash, 2000; Rauske et al., 2010), and shortly after the offset of vocalization, when auditory and motor traces of song would be most reliably encoded by short-term neural plasticity. We hypothesized that such offline integration of sensorimotor traces would require persistent changes in network activity following singing, and that these persistent changes may involve oscillatory activity, which has been implicated in memory consolidation and the facilitation of communication between connected brain areas (Buzsáki and Draguhn, 2004; Dan and Poo, 2004; Fries, 2005; Fries et al., 2008).

To test this hypothesis, we recorded both unit activity and local field potentials (LFPs) from the nucleus intercalaris of the nidopallium (NIf) in singing adult male zebra finches. NIf is part of the song system (Fig. 1A), a discrete network of interconnected nuclei that is necessary for the learning, maintenance, and production of song. NIf is also part of a sensorimotor loop, along with nucleus HVC (used as a proper name) and nucleus Avian (AV), that links the song system and the auditory forebrain (Akutagawa and Konishi, 2010), making it ideally situated to facilitate offline sensorimotor integration. We found that singing caused dramatic and lasting changes in firing rate and gamma oscillation power in NIf. We propose that these persistent changes in NIf’s network activity following singing could facilitate the integration of auditory and vocal motor traces associated with vocal performance evaluation.
were housed under 12 h light/dark conditions and provided with food and water ad libitum. All procedures described here were approved by an institutional animal care and use committee at the University of Pennsylvania (Philadelphia, PA).

Acute surgery
A custom-built motorized microdrive (RP Metrix) wired with 0.8–1.2 MΩ tungsten microelectrodes (FHC) was implanted in the right NIf using the following procedure. Adult male zebra finches were anesthetized using 0.02 ml of ketamine/xylazine (40 mg/kg ketamine and 8 mg/kg xylazine; Phoenix Pharmaceuticals) and 0.06–0.1 ml of diazepam (7.5 mg/kg; Abbott Laboratories). NIf was located using previously established stereotaxic coordinates and identified based on a combination of characteristic neural activity patterns associated with the bird’s own song (BOS). In addition, once the subject had recovered from surgery, premotor activity during vocalization and the state dependence of auditory response properties in NIf (Cardin and Schmidt, 2004a) were used as additional criteria for the identification of NIf. These criteria have been used successfully to identify NIf in previous studies from our laboratory (Cardin and Schmidt, 2004a; Cardin et al., 2005). Before implanting, electrodes were retracted to 200 μm above NIf to reduce gliosis. The electrodes and exposed brain were surrounded with GelFoam (Pfizer), and the microdrive was secured to the skull using dental cement. A silver wire implanted under the skull acted as ground. Subjects were allowed to recover overnight and then transferred to a modified holding cage where a tether containing a small operational amplifier (Texas Instruments) connected the implanted microdrive to a swivel connector on the roof of the cage. An elastic thread built into the tether helped to support the weight of the implant. Subjects remained tethered both during and between experiments.

Chronic recordings
During experimental sessions subjects were transferred to a recording chamber and tethered to a mercury commutator. Neural signals were collected using a BioAmp headstage and DB4 module (BioAmp headstage, DB4, and FT5 units, Tucker-Davis Technologies). Signals were amplified and filtered between 1 Hz and 7 kHz using the DB4 module and a four-channel amplifier (Model 440, Brownlee). The signal was then digitized and recorded using a Power 1401 Mark II board and Spike2 software (Cambridge Electronic Design). Neural recordings were collected while the subject vocalized in response to conspecifics housed nearby and playback of conspecific vocalizations through a hidden speaker.

Experimental conditions
The goal of this experiment was to record neural activity in NIf during natural vocal behavior. To elicit the necessary amount of vocal communication behavior in the 1–2 h that we could typically hold a multunit site in a moving subject, we adopted an adaptive experimental protocol tailored for each subject. In the simplest condition, a subject was housed alone and responded to playback of BOS, reverse BOS, and conspecific vocalizations played at ~60 dB SPL measured at the location of the subject. Playback alone was rarely enough to elicit the quantity and variety of vocalizations we required. We progressively enriched a subject’s social environment through the addition of both seen and unseen male and female zebra finches until the subject was vocalizing at a high level. The subject was able to hear and be heard by these conspecifics. For most subjects the presence of a familiar female visible in a nearby cage combined with playback was enough to induce strong vocal activity. Our most laconic subject required playback with one female visible and a combined with playback was enough to induce strong vocal activity. Our most laconic subject required playback with one female visible and a second female housed out of sight with a highly vocal male. Despite our best efforts to actively shape the experimental environment to encourage vocal behavior, we still were unable to use data from many sites. These included sites where a subject did not vocalize enough and sites where the subject vocalized too much, leaving very few periods of silence between vocalizations during which we could measure dynamic state changes. During periods of vocal inactivity, subjects were regularly inspected to ensure that they remained awake.

Materials and Methods

Animals
Adult male and female zebra finches (Taeniopygia guttata) ranging in age from 120 to 500 d were obtained from a local supplier (Canary Bird Farm, NJ). Subjects were administered antibiotics and monitored for a week before becoming eligible for experimental use. Birds

Figure 1. Connectivity and characteristic premotor activity distinguish NIf from surrounding auditory nuclei. A, Schematic of the main auditory and vocal motor pathways in the avian brain. Auditory information reaches NIf from the caudal mesopallium (CM), a secondary auditory area (Vates et al., 1996), and nucleus Avalanche (Av), which is part of the CM. HVC and the robust nucleus of the arcopallium, RA, make up the telencephalic portion of the descending motor pathway, which controls vocal output through projections onto brainstem respiratory nuclei and interhemispheric coordination are believed to involve a recurrent loop that includes HVC, the robust nucleus of the mesopallium (RA), and the pre-Boss motor system (Phillips et al., 1996). HVC and the robust nucleus of the arcopallium, RA, make up the telencephalic portion of the descending motor pathway, which controls vocal output through projections onto brainstem respiratory nuclei and the tracheo-syringeal portion of the hypoglossal nucleus (nXIIb). Song pattern generation and interhemispheric coordination are believed to involve a recurrent loop that includes HVC, RA, nucleus of the respiratory complex, and thalamic nucleus ovaformis (Uva), which also projects to NIf and Av. Av is included with the darker shaded song system nuclei to highlight its strong interconnectivity with NIf and HVC and the projection it receives from Uva. MLd, Dorsal lateral nucleus of the mesencephalon; Ov, nucleus ovaformis. B, Exemplar site demonstrating NIf's characteristic premotor activity patterns. Top, Spectrogram and amplitude waveform of a song with four motifs. Bottom, Expanded time view of multiunit activity around the first introductory note. The signal is first observed before the onset of the first introductory note (left) and the offset of the last motif (right) of song. The time lag from premotor onset to vocal onset is highlighted in green. Premotor activity ceases shortly before the offset of vocal production (highlighted in dark blue) and remains suppressed past the end of vocal output (highlighted in light blue). Red line represents baseline firing rate.
Histology
In many cases, electrode sites were confirmed by creating electrolytic lesions using an isolated pulse stimulator (model 2100; AM-Systems). Subjects were then killed using 0.2 ml of 50 mg/ml Nembutal (Abbott Laboratories). Their brains were removed and fixed in 4% formaldehyde and cryoprotected using 30% sucrose. Tissue was Nissl stained to visualize the lesion sites.

Data analysis
Data were analyzed using Spike2 and MATLAB (The MathWorks) routines written by B.C.L. unless otherwise noted. LFPs were isolated by offline filtering of neural data between 1 and 200 Hz. Multitunit activity (MUA) was isolated by highpass filtering neural data at or above 250 Hz (filtering above 250 Hz was occasionally used to remove movement artifact). Multitunit spike times were determined by setting a threshold visually to a point high enough to be clearly above background noise, but low enough to capture all distinct spike events as performed in our previous studies in NIf (Cardin and Schmidt, 2004b). Attempts to use a threshold based upon standard deviations above background noise were unreliable because of the difficulty of establishing baseline periods to measure background noise due to the high spontaneous multitunit spike rate in NIf. All LFP and MUA data used for analysis were visually inspected to ensure they were free of movement artifacts.

Measuring vocalization onsets and offsets
Microphone recordings were bandpass filtered (3–7 kHz) to eliminate ambient noise while retaining vocalization related information. The cut-off frequencies of the bandpass filter were chosen to optimize the signal-to-noise ratio to enhance the accuracy of onset and offset measurements. Vocal onsets were determined using the root mean square (RMS, 2 ms window) of the filtered microphone trace and a threshold set above background noise. To account for mild reverberation in the recording room, vocal offsets were defined as the point in the microphone RMS trace when sound amplitude decreased the most rapidly. Operationally, these inflection points were identified as minima in the average slope (10 ms window) of the RMS (1 ms window) of the filtered audio recording.

Defining the measurement windows
To determine the effect of vocal behavior on NIf’s network activity, spike rate and fast gamma power changes were measured in three nonoverlapping periods: Pre-Vocal, Post-Vocal, and Baseline.

Pre-Vocal. Defined as a period 1–3 s before vocal onset, significantly outside the range of premotor activity (see Results). To ensure that changes observed in the Pre-Vocal period were not actually Post-Vocal changes that had not fully decayed, vocal onsets had to be preceded by at least 70 s of vocal inactivity.

Post-Vocal. Defined as a window of 2–4 s after vocal offset. This window captures stable peak spike rate and oscillation values. Any vocalization that is followed by at least 4.5 s of vocal inactivity is considered for Post-Vocal analysis. Each site had to have at least seven song offsets that fit the criteria for the Post-Vocal period to be included in our final analysis.

Baseline. Defined as a period of vocal inactivity starting 90 s after vocalization and lasting until 30 s before the next vocalization.

Premotor activity
Premotor onset. Average MUA was calculated in 10 ms windows (90% overlap) aligned to the onset of at least 10 introductory notes. Counting backwards from vocal onset, the time of the first bin significantly above baseline (Student’s t test) that is preceded by two nonsignificant bins is considered premotor onset. The baseline is the average firing during Baseline periods.

Premotor offset. Defined as the measure of the lag time between the end of premotor activity and the end of vocal production. Calculated from average MUA aligned to song offsets, premotor offset for a given site is the time from the last bin significantly above baseline to the end of song.

Post-Vocal inhibition. Defined as the time it takes MUA to return to baseline levels following song offset. Using average MUA, significance with respect to baseline was determined for each bin following vocal offset (Student’s t test, p < 0.05). To account for the gradual return to baseline, we found the first 40 ms after vocal-offset where half the bins were not significantly below baseline. Post-Vocal inhibition is the time of the last significant bin within this 40 ms window.

Dynamic state changes
Early in our analyses we observed significant variability in the magnitude of changes in spike rate and synchronized spiking activity (dynamic state changes) following different vocalization types. Because song produced the most consistent changes, we restrict our analysis of dynamic state changes in NIf to song offsets unless otherwise noted. Firing rate was quantified using Z-scores for comparison both within and across sites:

\[ Z\text{-score} = \frac{x - \mu}{\sigma}, \]

where \( x \) is the experimental value, \( \mu \) is the average during Baseline periods, and \( \sigma \) is the Baseline standard deviation. Spike counts were measured in nonoverlapping 2 s bins across Baseline periods. Most sites had multiple Baseline periods separated by intervals of vocal activity. To account for any possible drift in firing rate, the nearest Baseline period was always chosen to compute the Z-score. Significance was determined using two-sample t tests comparing Pre-Vocal or Post-Vocal Z-scores to Baseline Z-scores (average Baseline Z-score will be 0, but measurements from individual bins provide variance).

For spectral analysis of spike signals, data were converted to a binary representation of spike activity with a 0.5 ms resolution. Power spectral density was measured from these binary time series using Thomson’s multitaper method (Thomson, 1982) as implemented by MATLAB with a time–bandwidth product of 6. The multitaper method computes spectral estimates using a series of orthogonal tapers derived from Slepian sequences that seek to minimize errors in the frequency domain caused by the use of finite windows. To control for the possibility that increases in spike rate might cause broadband increases in spectral power (Manning et al., 2009), power measurements across a given frequency band were normalized by the total power of the signal. The resulting “normalized power” metric represents the percentage of the total power of the signal contained within the frequency band of interest. Power spectra for LFP data were also calculated using Thomson’s multitaper method. Power in LFP spectra was unevenly distributed across frequencies in a manner resembling the 1/f power law distribution observed in mammals (Bédard and Destexhe, 2009). Due to this uneven distribution of power, low frequencies contained significantly more power than high frequencies and thus were over-represented in LFP power measurements. To control for this, we normalized LFP power spectra calculated from experimental periods (Pre-Vocal, Post-Vocal, or the final 1 s of song) by the average Baseline spectrum calculated in nonoverlapping 2 s windows across Baseline periods. The result is a modulation spectrum that shows the change in power relative to baseline for each frequency (Henrie and Shapley, 2005). Modulation spectra control for the uneven distribution of power across frequencies but do not control for broadband shifts in spectral power. Therefore, the normalized power metric described above was also used to measure power across a given frequency band in modulation spectra.

A two-sample t test was used to test for significant differences between normalized power calculated from experimental periods (Pre-Vocal, Post-Vocal, or the last 1 s of song) and normalized power calculated in nonoverlapping 2 s bins across Baseline periods (1 s bins were used for comparison with data from the last 1 s of song).

Dynamic state duration measurements
For each site MUA was used to calculate spike rate and normalized power in 2 s windows centered on times beginning at song offset and then moved piecewise in 50 ms steps. For each time window, the collection of values for spike rate and normalized power was tested for significance versus baseline (two-sample t test, α = 0.01). This resulted in a vector of significance vs time relative to song offset with a 50 ms resolution. The duration of spike rate and fast gamma power increases for a given site was defined as the first significant bin followed by 1 s of nonsignificant bins. This value of 1 s was determined empirically after we observed sites where brief (<1 s) dips below significance were followed by many seconds of significant values.
Single unit isolation and analysis

 Spike waveforms were identified and isolated using the WaveMark feature of Spike2. These waveforms were exported for analysis using custom built software developed in our laboratory by Dr. Chris Glaze (University of Pennsylvania, Philadelphia, PA). The software identifies single-unit clusters by modeling data from the first three principal components of spike waveforms using a mixture of t distributions (Svensen and Bishop, 2005). Each cluster was then manually inspected for separation and accuracy. Finally, Spike2’s built-in principal component analysis was used to follow the evolution of identified clusters over time. Single-unit analysis was only conducted for the period of time during which a unit formed a distinct cluster. For single unit analyses, we expanded the Post-Vocal window to 1.5–5.5 s following vocal offset to account for smaller spike sample sizes.

Phase-locking measurements

Phase times were recovered from the analytic signal of a Hilbert transform performed on signals filtered in the fast gamma range (90–150 Hz) (see Le Van Quyen et al., 2001, for a detailed description of Hilbert transforms). For each subject in which we recorded LFP signals, two electrodes were implanted. Each electrode was separated by 250 μm. Due to NIf’s small size and oblong shape, it was rare for both of these electrodes to be within the nucleus simultaneously. To avoid concerns that spike waveforms could influence the LFP signal (David et al., 2010), we used the electrode not recording the single unit to calculate phase times, hereafter referred to as Eref (reference electrode). Although LFPs lose power as they propagate through the extracellular matrix from their source, in mammals they can often be detected 1 mm away from their source (for review, see Logothetis, 2003).

As expected the LFPs that we recorded outside of NIf had lost some power as they propagated to Eref, but they still contained reliable phase information. Using fast gamma phase times measured from Eref, we determined the phase of each spike in the modified Post-Vocal window and an equal number of spikes selected randomly from Baseline. A Rayleigh test for circular uniformity was used to measure the significance of phase locking, and a Kuiper two-sample test was used to test for differences in phase locking between Baseline and Post-Vocal conditions. These tests were conducting using routines from the CircStat toolbox for MATLAB (Berens, 2009).

Spike field coherence

Spike field coherence (SFC) is a normalized measure of the degree of phase synchronization between spike activity and the LFP as a function of frequency. This method has been described in detail previously (Fries et al., 2001). Briefly, a spike–triggered waveform average (STA) is calculated by averaging the LFP around each spike. The absolute value of the STA is calculated and then normalized by the average power of the STA. The result is the SFC, a measure independent of spike rate or LFP amplitude, which takes a value between 0 and 1 for each frequency. A value of 1 indicates that all spikes occur at exactly the same phase for that frequency. A value of 0 indicates that no phase relationship exists between spike activity and oscillations at that frequency. SFC was calculated using multitaper methods (five tapers with a time–bandwidth product of 3) as implemented by the coherencectp.m function from the Chronux data analysis toolbox for MATLAB (Mitra and Bokil, 2008), available at http://chronux.org.

Auditory response testing

To test the influence of fast gamma oscillations on auditory processing, we examined auditory responses to long calls given by freely behaving female conspecifics. To determine whether a site had significant auditory responses to conspecific calls, we compared spike counts in 200 ms windows beginning at long-call onset with spike counts in the 200 ms preceding long-call onset using a paired t test. We then identified all conspecific long calls that were produced within 12 s of the experimental subject singing. Because gamma oscillations are strongest after song, we classified these conspecific calls as Strong Oscillation Calls. We compared these against conspecific calls given during the Baseline period when fast gamma power is low (Weak Oscillation Calls). Auditory responses were quantified using the auditory response index:

$$RS_{INDEX} = \frac{(X_{Response} - X_{Baseline})}{(X_{Response} + X_{Baseline})}. \quad (2)$$

The value for $X_{Response}$ was either spike count or normalized fast gamma power, both measured in 200 ms windows beginning at the onset of a conspecific call. The baseline for the Weak Oscillation Calls was the 200 ms preceding conspecific call onset. Because spike rate and oscillation power are not constant following song, we compared the measured response to a Strong Oscillation Call with the average value during the same time window following other songs. $RS_{INDEX}$ values for the two groups were compared using a two-sample t test.

Measuring premotor activity

To investigate the relationship between premotor activity and fast gamma power at each site, we compared the oscillation power measured in the Post-Vocal period after each bout of vocalizations with measures of the premotor activity that accompany these vocalizations. Vocal bouts are typically made up of some combination of song, long calls, and tet/stack calls. Male long calls have stereotyped spectrotemporal features that are learned from a tutor; these calls are typically 100–200 ms long and are used to maintain long-range vocal contact. Tet calls and stack calls are unlearned vocalizations used in short-range vocal communication; they are shorter (~40–60 ms) and quieter than long calls (Zann, 1996). For simplicity, the term “tets” is used to refer to both tet calls and stack calls in the remainder of the text.

It is not uncommon for vocalizations within a bout to be separated by 1–2 s of silence. The presence of fast gamma oscillations during these silent intervals complicates the estimation of premotor activity. The short bursts of multiunit firing associated with fast gamma oscillations have firing rates as high as those seen during premotor activity, which makes instantaneous firing rate an inaccurate tool for identifying premotor spikes. However, the increase in firing rate during premotor activity persists for tens of milliseconds, while fast gamma bursts last only ~2–4 ms. We therefore estimated premotor activity by calculating firing rate in 20 ms nonoverlapping bins starting at vocal offset and moving backwards in time. All bins with firing rates greater than two standard deviations above the mean Baseline firing rate were considered premotor activity. Spike counts derived from these “premotor” bins were used to calculate the average premotor firing rate.

If premotor activity and fast gamma power are correlated, then the slow decay of fast gamma power beginning ~3–5 s after vocal offset (see Results, Fig. 6) suggests that the influence of premotor activity on fast gamma power decays with time. The average normalized oscillation power for all sites drops to ~70% of its maximum value 7 s after premotor activity ends. This suggests the magnitude of fast gamma oscillations is maximally influenced by premotor activity that occurred within the previous 7 s. To test whether premotor activity was correlated with fast gamma power, we counted all premotor spikes, as described above, within 7 s of the start of each Post-Vocal period. Thus, for each Post-Vocal period a normalized power value and an estimate of premotor firing rate were calculated. The correlation between these values for all vocalizations at a site was measured using Pearson’s linear correlation coefficient and tested for significance against the null hypothesis of no correlation.

Results

To investigate NIf’s role in vocal production and performance evaluation, we recorded multiunit activity and LFPs in awake behaving male zebra finches engaging in natural vocal communication behavior. Changes in network activity associated with vocalizations are reported for 18 sites across 6 birds.

NIf shows characteristic patterns of premotor activity

NIf is embedded in the auditory forebrain but is distinguishable from the surrounding auditory nuclei by its strong bursts of premotor activity during vocal production. Premotor activity during songs and long calls follows a general pattern characterized by
three distinguishing features: large increases in firing rate before and during vocalization, cessation of neural activity before the offset of the vocalization, and continued suppression of neural activity for a short time after vocal offset (Fig. 1B). Measured from the onset of the introductory notes that precede song, NIf premotor activity preceded song by an average of 45.7 ± 15.7 ms (18 sites across 6 birds). The pattern of premotor activity during song is related to the syllables that are produced and thus cannot be easily generalized across birds singing different songs. However, in all six birds recorded for this study, premotor activity ended shortly before song offset. The average time between the end of premotor activity and the offset of song was 26.5 ± 13.5 ms. Cessation of premotor activity was accompanied by a general suppression of neural activity in NIf. Firing rates remained below baseline levels for an average of 240 ± 204 ms. This general pattern of premotor activity was consistent across all sites for songs and long calls and agrees with latencies recorded in earlier studies (McCasland, 1987).

Vocal activity causes long lasting changes in the dynamic network state of NIf.

To determine whether vocal behavior caused changes in NIf’s neural activity beyond the transient increases in firing rate associated with premotor activity, we compared neural firing rates in NIf before and after periods of vocal activity with those measured during extended periods of silence. Neural activity was examined in three periods: Pre-Vocal, Post-Vocal, and a Baseline period defined by prolonged vocal inactivity (Fig. 2A; see Materials and Methods for details). To allow for comparison across sites, Z-scores were calculated for the Pre-Vocal and Post-Vocal periods based on firing rates measured during Baseline periods.

A representative example of neural activity in NIf both before and after the production of song is shown in Figure 2B. The brief suppression of activity following vocal offset transitions rapidly into a prolonged increase in firing rate. Most sites exhibited a significant increase in spiking relative to Baseline during the Post-Vocal period (15/18 sites, \( p < 0.01 \), two-sample \( t \) test; see Fig. 4).
was a strong correlation between gamma power and increased spike rate (Fig. 4A, \( r = 0.88, p < 0.001, \) Pearson linear correlation). This is in keeping with findings from previous studies where increases in gamma power are correlated with increased spike rates (Fries et al., 2001; Manning et al., 2009).

Similar to observations of gamma in mammals (Jansen et al., 2009; Hall et al., 2011), the frequency with the greatest power in the fast gamma range (peak frequency) was consistent across sites for an individual subject but could differ considerably across subjects (Fig. 4B). Across subjects, the average peak frequency of oscillation in the Post-Vocal period ranged from 110.34 Hz to 135.42 Hz; in contrast, the average frequency spread across sites within subjects was only 3.70 Hz. The difference in peak frequencies across subjects was significant (\( p < 0.0001, \) one-way ANOVA). These results suggest that each subject has an individually characteristic distribution of power across the fast gamma band.

**Local field potential recordings in NIf demonstrate selective increases in fast gamma power.**

We hypothesized that the highly coherent nature of neural firing during fast gamma oscillations should allow their detection in LFPs even though the avian telencephalon is not ideal for such measurements given its nonlaminar structure without any clear neural architecture to sum and amplify synchronous neural events. We recorded LFPs and MUA simultaneously from 15 sites in 5 birds and looked for changes in power during the Pre-Vocal and Post-Vocal periods. To control for natural variations in LFP power across frequencies, we computed “modulation” spectra (Fig. 5A; see Materials and Methods) that show the change in power for each frequency relative to the Baseline average. Most sites showed a significant increase in LFP power across the fast gamma band during the Post-Vocal period (12/15 sites, \( p < 0.01, \) t test). No sites had increased fast gamma power during the Pre-Vocal period (0/15, \( p < 0.01, \) t test). Fast gamma power measured from LFPs was positively correlated with spike rate (\( r = 0.82, p < 0.001 \) Pearson’s test for nonzero correlation). Figure 5B shows the average modulation spectrum for all sites that had significant increases in Post-Vocal fast gamma power.

The ability to detect fast gamma oscillations in the spike output of NIf neurons (Fig. 3) is strong evidence that either fast gamma oscillations are generated locally by the neural network in NIf, or else a second structure connected to NIf oscillates in the fast gamma range and entrains NIf’s network through strong synaptic connections. NIf is not connected to the auditory structures that surround it, and therefore it is unlikely that these structures are the source of the fast gamma oscillations that shape neural activity within NIf. Further evidence that the increase in LFP fast gamma power originates from NIf comes from a comparison of LFPs recorded from electrodes located within NIf with reference electrodes.
located a short distance (~250 μm) outside of NIf. At all sites, fast gamma power was stronger at the LFP recorded within NIf than at the LFP recorded from outside of NIf; at all sites but one, this difference reached significance (10/11, \( p < 0.01 \), paired \( t \) test).

To further confirm that fast gamma power observed in the LFP recordings originates from NIf, we compared activity across five recording sites from a single electrode penetration through NIf. We hypothesized that the power of fast gamma oscillations should increase with proximity to the center of NIf, given that NIf has no clear neural architecture that would support segregated layers of sinks and sources (Fortune and Margoliash, 1992). As shown in Figure 5C, as the recording electrode moved through NIf there was first an increase and then a decrease in the magnitude of premotor activity, a characteristic feature of neural activity in NIf (see Fig. 1). These changes in premotor activity across sites were strongly correlated with changes in the magnitude of normalized fast gamma measured from the LFPs recorded at each site (\( r = 0.8905, p = 0.043 \), Pearson’s test for nonzero correlation).

Together, these findings all support the conclusion that NIf is the source of the increases in fast gamma power observed in our local field potential recordings.

Increases in fast gamma power and neural firing rate often persist for tens of seconds after the offset of song

Increases in gamma band activity triggered by a stimulus or event typically last for a second or less (Laurent and Naraghi, 1994; Siegel and Konig, 2003; Brugge et al., 2009; Fukuda et al., 2010). In stark contrast, NIf’s vocalization-induced increases in fast gamma power often persisted for 10 or more seconds, with some
The duration of the changes in NIf’s network activity varied considerably from site to site (Fig. 6C). The duration of elevated fast gamma power following the offset of song ranged between 9.4 and 34.7 s, with an average of 21.62 ± 8.76 s (17 sites in 6 birds). Spike rate in NIf remained elevated following the offset of song for periods ranging from 10.2 s to 51.1 s, with an average of 32.09 ± 13.96 s (13 sites in 5 birds). Only sites that demonstrated significant increases in fast gamma power or spike rate were included in this analysis.

Fast gamma oscillations in NIf during song production

The presence of fast gamma oscillations during song could indicate that fast gamma is involved in providing a temporal framework for motor production. This would fit with previous results showing that network dynamics within the robust nucleus of the arcopallium (RA) evolve on a timescale of ~125 Hz (Leonardo and Fee, 2005). Measurements of fast gamma power during song production must be assessed using LFPs because the power spectrum derived from multunit spike times is dominated by the premotor bursts associated with the production of song syllables. To determine whether fast gamma power increased before the offset of song, we measured normalized LFP power during the final 1 s of each song that directly preceded a Post-Vocal period. In both the recording (E_rec) and reference (E_ref) electrodes, we observed either no change or a significant decrease (E_rec, 10/15 sites; E_ref, 5/10 sites; p < 0.01, t-test) in normalized LFP fast gamma power during singing as compared to Baseline.

Isolated single units are phase locked to fast gamma oscillations

We were able to isolate 14 single units from our multunit recording sites in NIf (see Materials and Methods for details). To test for phase locking, we determined the phase of fast gamma oscillations corresponding to single-unit spike times from the Post-Vocal and Baseline periods. To control for any possible influence of spike waveforms on the simultaneously recorded LFP, we measured phase locking of single units against LFPs recorded from a second electrode (E_ref) located outside of NIf for all but two units, and neither of these two units showed significant phase locking.

Of 14 single units, 8 were significantly phase locked to the LFP during the Post-Vocal period (8/14, p < 0.05, Rayleigh test). Figure 7A shows the phase distribution and spike-triggered waveform average of an example unit during Post-Vocal and Baseline periods. The clearly visible fast gamma waveform in the STA demonstrates the tight coupling between single-unit firing and oscillations during the Post-Vocal period. In this example, the STA calculated from Baseline spike times shows a small but significant amount of phase locking during nonvocal periods. Four units showed weak phase locking during the baseline period (4/14, p < 0.05, Rayleigh test). Figure 7B shows the phase preference for each of the phase-locked neurons. Most phase-locked neurons fired preferentially during the trough or rising phase of fast gamma.

To quantify the phase-locking properties of identified single units across frequencies, we calculated their spike-field coherence. In this analysis a coherence of 1 for a given frequency indicates that spikes always occur at the same phase, while a value of 0 indicates no correlation between spike times and the phase of oscillation. Figure 7C shows the SFC for an exemplar unit. For this unit the increase in phase locking during the Post-Vocal period is restricted to a frequency band between 95 and 120 Hz.
The SFC during the Post-Vocal period for each of the significantly phase-locked units is shown in Figure 7D.

Of the 14 single units, 5 had significantly increased spike rates in the Post-Vocal period relative to Baseline (5/14, \( p < 0.05 \), two-sample \( t \) test), 8 had significantly decreased firing rates (8/14, \( p < 0.05 \), two-sample \( t \) test), and 1 showed no significant change in firing rate. All of the single units with increased Post-Vocal firing rates were also significantly phase-locked during the Post-Vocal period. Of the neurons that did not increase their firing rate in the Post-Vocal period, only three were significantly phase-locked (3/9, \( p < 0.05 \) Raleigh test). The fact that most of our single units had decreased firing rates in the Post-Vocal period, while nearly all of our multiunit recordings increased firing during this time, may be a reflection of the fact that larger neurons are easier to isolate from multiunit recordings (Logothetis, 2003). NIf is known to contain populations of neurons with different soma sizes (Fortune and Margoliash, 1992), and the larger of these neural subtypes will be naturally over-represented in single unit isolations from extracellular recordings.

Auditory responses during fast gamma oscillations in NIf

NIf shows a high degree of variability in both the strength and selectivity of auditory responses in awake birds (Cardin and Schmidt, 2004a). It is possible that fast gamma oscillations in NIf represent a specialization for processing conspecific responses during vocal communication. We compared auditory responses to conspecific long calls given during periods of strong oscillations (following song production) and weak oscillations (during the Baseline period). Because these calls were produced during natural vocal interactions, only six sites had enough activity to be analyzed. All six of these sites had significant auditory responses to conspecific calls, measured as an increase in multiunit firing (6/6 sites, \( p < 0.01 \), \( t \) test). However, none of the sites showed any significant difference in auditory responses during periods of strong vs weak oscillations (6/6 sites, \( p < 0.05 \), \( t \) test). In addition, no sites demonstrated significant changes in normalized fast gamma power in response to conspecific calls (6/6 sites, \( p < 0.05 \), \( t \) test). Because of our small sample size we cannot conclude with certainty that fast gamma oscillations have no influence on auditory responses in NIf. It is worth noting, however, that antiphonal calling (the exchange of calls between two or more conspecifics) is the primary form of reciprocal vocal communication in zebra finches (Zann, 1996). Antiphonal call responses are typically produced within a few hundred milliseconds after hearing a conspecific long call. Thus, an antiphonal response call will often arrive during the period of post-vocal inhibition in NIf when both spike rate and oscillatory activity are largely suppressed (see Figs. 1, 2).

The relationship between fast gamma oscillations and premotor activity in NIf

To clarify the relationship between vocal activity and fast gamma oscillations, we expanded our analysis to include both tets (short
against normalized fast gamma power in the Post-Vocal period. The vocalization type preceding vocalization tended to asymptote before maximal levels of premotor activity were reached (Fig. 8B). As a result of this, oscillations induced by singing would often approach the maximal power for fast gamma at a given site. Thus, while fast gamma power increased following tets and long calls, most strong oscillations occurred after song.

Discussion

In this study, we show that vocal activity causes dramatic and prolonged changes in network dynamics in NIf that are characterized by an increase in spike rate and neural oscillations in the fast gamma range (90–150 Hz). The observation of fast gamma oscillations in LFPs recorded outside of NIf suggests an unusually high degree of coherency in these oscillations, given that NIf and most structures in the avian brain (Wang et al., 2010) lack a laminated architecture and any clear concentration of sinks and sources that would amplify synchronous synaptic activity (Fortune and Margoliash, 1992; Lethologita, 2003). NIf’s fast gamma oscillations could play an important role in integrating vocalization-related sensorimotor information important for the encoding and maintenance of learned vocalizations.

A remarkable characteristic of our observations is that relatively short bouts of vocal activity (2–3 s) can induce oscillations that persist for an order of magnitude longer than the stimulus that initiated them. We are unaware of any examples where an isolated stimulus or behavioral event causes oscillations that last more than a few seconds (Laurent and Naraghi, 1994; Siegel and Konig, 2003; Brugge et al., 2009; Fukuda et al., 2010). Most models of gamma-producing networks require tonic input to reproduce oscillatory activity (Bartos et al., 2007). While we cannot rule out that NIf receives such input from its auditory afferents, it is unlikely to come from within the song system, because activity in the premotor thalamic nucleus Uva, which sends a strong projection to NIf, remains elevated for only 1–2 s following song offset (Williams and Vicario, 1993). Another possibility is that neuromodulatory inputs (Ryan and Arnold, 1981; Cardin and Schmidt, 2004b) are active during and/or following vocalization in NIf and cause the network to express oscillatory activity (Fisahn et al., 1998).

Fast oscillations and the song system

Although no previous studies have reported fast gamma oscillations in the song system, results from studies in RA hint at the involvement of fast gamma in the encoding and production of song. In RA, each moment of song is encoded by a stereotyped ensemble of neurons (Yu and Margoliash, 1996; Leonardo and Fee, 2005). The motor code for song in RA consists of a succession of unique neural ensembles, each of which is active for ~8 ms (125 Hz), suggesting that, at the level of RA, the structure of song may be encoded on a fixed time scale in the fast gamma range. Activity patterns in RA also suggest that fast gamma may be important for the offline encoding of tutor song during song learning in juvenile birds. During the first night following exposure to the tutor song, RA neurons transition from being tonically active to becoming more phasic in nature, with a dramatic increase in interspike intervals in the fast gamma range (Shank and Margoliash, 2009). Finally, RA and HVC both contain a population of fast-spiking, parvalbumin-expressing interneurons (Wild et al., 2001, 2005; Rauske et al., 2003) similar to those that underlie gamma oscillations in mammals, and experiments in RA have
shown that activation of this interneuronal network can synchronize firing in projection neurons (Spiro et al., 1999).

**Gamma oscillations in NIF are unlikely to play a direct role in motor production or sensory processing**

Despite its song-related auditory and motor activity, it remains unclear what role NIF plays in vocal production and/or auditory processing. Anatomically, NIf is located at the interface of the auditory and vocal motor systems. It receives motor projections from the thalamic nucleus Uva and auditory input from the secondary auditory area CM (Vates et al., 1996), including a strong projection from nucleus Avalanché, a subdivision of CM that is connected reciprocally with both NIF and HVC (Akutagawa and Konishi, 2010). In addition to auditory information, NIF provides strong excitatory drive to nucleus HVC (Coleman and Mooney, 2004; Hahnloser and Fee, 2007), which along with RA makes up the telencephalic portion of the direct motor pathway of the song system. Unlike HVC and RA, however, lesions of NIF do not cause immediate degradation of the spectrotemporal features of song (Simpson and Vicario, 1990; Cardin et al., 2005; Roy and Mooney, 2009). This lack of short-term spectral degradation of song following NIF lesions, combined with our failure to find significant increases in normalized fast gamma power before or during song, make it unlikely that NIF’s oscillations are directly involved in song production.

Auditory responses in NIF range from variable and nonselective in awake birds to strong and highly selective for BOS during sleep (Janata and Margoliash, 1999; Cardin and Schmidt, 2004a). In mammals, gamma oscillations have been linked to attentional modulation of sensory responses (Fries et al., 2001; Buehlmann and Deco, 2008); similar modulation in NIF could help explain the variability of auditory responses in the awake bird (Cardin and Schmidt, 2004a). To test this possibility, we measured auditory responses to calls produced by freely behaving conspecifics. We failed, however, to observe any differences in auditory response strength between calls heard under conditions of high or low gamma power.

**NIF and the auditory feedback-dependent maintenance of song**

Bilateral lesions of NIF cause a significant decrease in spontaneous activity and auditory responses in HVC, which is the source of auditory information to both the descending motor and anterior forebrain pathways (Coleman and Mooney, 2004; Cardin et al., 2005; Roy and Mooney, 2009). Given the necessity of auditory feedback for the maintenance of crystallized song, lesions of NIF in zebra finches might be expected to cause spectrotemporal degradation of song similar to that seen following deafening (Nordeen and Nordeen, 1992; Lombardino and Nottebohm, 2000). In Bengalese finches, both deafening (Woolley and Rubel, 1997) and NIF lesions (Hosino and Okanoya, 2000) cause, within a relatively short time period (days to weeks), syntactically complex songs to become less variable and more deterministic. In contrast, zebra finch song retains its spectrotemporal structure following bilateral NIF lesions for at least 5 weeks (Cardin et al., 2005; Roy and Mooney, 2009), which is longer than song remains intact in deafened birds of the same age (Lombardino and Nottebohm, 2000). However, the songs of older zebra finches can maintain their spectrotemporal structure for more than a year following deafening (Lombardino and Nottebohm, 2000), suggesting that the motor code of crystallized zebra finch song is resilient to disruptions of auditory feedback. It is perhaps not surprising then that lesions of NIF do not have an immediate or near immediate effect on the spectrotemporal structure of adult song (Cardin et al., 2005; Roy and Mooney, 2009), especially since HVC also receives auditory inputs from Av (Akutagawa and Konishi, 2010). Whether NIF lesions would eventually cause degradations in song similar to those observed following deafening is unknown, and the role of NIF in auditory feedback-based song learning and maintenance remains unclear.

**Fast gamma oscillations as a possible mechanism for offline vocal-motor performance evaluation**

In other systems, the presence of sustained oscillations is an indication of sustained neural processing and/or communication between brain regions (Hoogenboom et al., 2006; Fries et al., 2008; Koch et al., 2009). In NIF, persistent oscillatory activity is specifically induced by the production of vocalizations, particularly song. It is reasonable to hypothesize, therefore, that fast gamma oscillations in NIF represent sustained processing of information associated with vocal production. The question then is: what sort of information a song system nucleus would be processing in the period immediately following song production?

Given the lack of immediate effects of feedback perturbation on song motor activity (Konishi, 2004; Kozhevnikov and Fee, 2007; but see Sakata and Brainard, 2008), offline processing has been proposed as a possible mechanism for comparing motor performance (auditory feedback) with motor instruction (premotor activity) (Margoliash and Schmidt, 2010). Fast gamma oscillations in NIF are tightly linked to premotor activity and may therefore be well suited to facilitate offline integration of auditory and premotor information following singing bouts. The anatomical connectivity between Av, NIF, and HVC suggests a sensorimotor loop that is well situated to play a role in offline vocal processing, especially given that Av forms part of the auditory forebrain known to contain neurons sensitive to perturbations in vocalization-related auditory feedback (Keller and Hahnloser, 2009). If Av tracks auditory feedback during vocal production, fast gamma oscillations could simultaneously enhance communication between NIF and Av while providing a temporal framework for the integration of vocal motor and auditory feedback traces. If such offline sensorimotor integration were occurring, then we would predict that field oscillations and/or single unit activity in Av should phase lock to fast gamma oscillations in NIF. Furthermore, if NIF is involved in the process of evaluating vocal performance to correct errors in vocal output, then lesions of NIF should affect the recovery of normal song following decrystallization caused by distorted auditory feedback (Leonardo and Konishi, 1999).

Comparisons of gamma oscillations in the songbird brain with those in mammalian systems will expand our understanding of the universality and versatility of gamma oscillations (Fries, 2009). Furthermore, as a discrete nucleus with a limited set of inputs, NIF provides a rich model system for the study of gamma oscillations, particularly given the prolonged duration of the oscillations and their established relationship with an easily quantifiable vocal behavior.

**References**

