

# The Avian Basal Ganglia Are a Source of Rapid Behavioral Variation That Enables Vocal Motor Exploration

Satoshi Kojima 小島哲,<sup>1</sup> Mimi H. Kao,<sup>2</sup> Allison J. Doupe,<sup>2†</sup> and Michael S. Brainard<sup>2</sup>

<sup>1</sup>Department of Structure and Function of Neural Network, Korea Brain Research Institute, Dong-gu, Daegu 41068, South Korea, and <sup>2</sup>Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, California 94143-0444

The basal ganglia (BG) participate in aspects of reinforcement learning that require evaluation and selection of motor programs associated with improved performance. However, whether the BG additionally contribute to behavioral variation (“motor exploration”) that forms the substrate for such learning remains unclear. In songbirds, a tractable system for studying BG-dependent skill learning, a role for the BG in generating exploratory variability, has been challenged by the finding that lesions of Area X, the song-specific component of the BG, have no lasting effects on several forms of vocal variability that have been studied. Here we demonstrate that lesions of Area X in adult male zebra finches (*Taeniopygia gutatta*) permanently eliminate rapid within-syllable variation in fundamental frequency (FF), which can act as motor exploration to enable reinforcement-driven song learning. In addition, we found that this within-syllable variation is elevated in juveniles and in adults singing alone, conditions that have been linked to enhanced song plasticity and elevated neural variability in Area X. Consistent with a model that variability is relayed from Area X, via its cortical target, the lateral magnocellular nucleus of the anterior nidopallium (LMAN), to influence song motor circuitry, we found that lesions of LMAN also eliminate within-syllable variability. Moreover, we found that electrical perturbation of LMAN can drive fluctuations in FF that mimic naturally occurring within-syllable variability. Together, these results demonstrate that the BG are a central source of rapid behavioral variation that can serve as motor exploration for vocal learning.

**Key words:** basal ganglia; motor exploration; reinforcement learning; social context; songbird; vocal learning

## Significance Statement

Many complex motor skills, such as speech, are not innately programmed but are learned gradually through trial and error. Learning involves generating exploratory variability in action (“motor exploration”) and evaluating subsequent performance to acquire motor programs that lead to improved performance. Although it is well established that the basal ganglia (BG) process signals relating to action evaluation and selection, whether and how the BG promote exploratory motor variability remain unclear. We investigated this question in songbirds, which learn to produce complex vocalizations through trial and error. In contrast with previous studies that did not find effects of BG lesions on vocal motor variability, we demonstrate that the BG are an essential source of rapid behavioral variation linked to vocal learning.

## Introduction

Generation of “exploratory” variability in actions is a critical component of reinforcement learning (Sutton and Barto, 1998). Although it has been proposed that the basal ganglia (BG) promote exploratory motor variability (Sridharan et al., 2006; Humphries et al., 2012; Kalva et al., 2012; Stocco, 2012), there is

little empirical evidence supporting this view (Barnes et al., 2005; Sheth et al., 2011), and neural substrates underlying motor exploration remain unclear.

Songbirds provide a tractable model system for studying the neural substrates of exploratory variability. Juvenile birds produce highly variable vocalizations akin to babbling of human infants (Marler, 1970; Tchernichovski et al., 2001; Aronov et al.,

Received Oct. 10, 2017; revised Sept. 7, 2018; accepted Sept. 8, 2018.

Author contributions: S.K., A.J.D., and M.S.B. designed research; S.K. and M.H.K. performed research; S.K. and M.H.K. analyzed data; S.K., M.H.K., A.J.D., and M.S.B. wrote the paper.

This work was supported by grants from the National Institutes of Health (Grant MH55897), the National Alliance for Research on Schizophrenia and Depression (A.J.D.), the Howard Hughes Medical Institute (M.H.K. and M.S.B.), and the Korea Brain Research Institute basic research program (Grant 18-BR-01-06 to S.K.). We thank the members of the Doupe and Brainard laboratories for discussion and comments on this manuscript, A. Arteseros for technical assistance, and R. O. Tachibana (University of Tokyo, Tokyo, Japan) for setting up the system for song-triggered microstimulation.

<sup>†</sup>Deceased, Oct. 24, 2014.

The authors declare no competing financial interests.

Correspondence should be addressed to Satoshi Kojima, Department of Structure and Function of Neural Network, Korea Brain Research Institute, 61, Cheomdan-ro, Dong-gu, Daegu 41068, South Korea. Email: satoshikojima.sk@gmail.com.

M. H. Kao's present address: Department of Biology, Tufts University, 200 Boston Avenue, Medford, MA 02155. <https://doi.org/10.1523/JNEUROSCI.2915-17.2018>

Copyright © 2018 the authors 0270-6474/18/389635-13\$15.00/0

2008), and even adult birds generate variability that enables vocal reinforcement learning (Kao et al., 2005; Tumer and Brainard, 2007; Andalman and Fee, 2009). Much of the variability in song, including cross- rendition variation in the mean fundamental frequency (FF) of individual song elements (syllables; Fig. 1A, left), is driven by a specialized BG–thalamocortical circuit, the anterior forebrain pathway (AFP; Fig. 1B, left). Although the “cortical” lateral magnocellular nucleus of the anterior nidopallium (LMAN) contributes to cross- rendition variation in mean FF (Fig. 1B, C, middle; Bottjer et al., 1984; Scharff and Nottebohm, 1991; Kao et al., 2005; Ölveczky et al., 2005, 2011; Kao and Brainard, 2006; Aronov et al., 2008; Stepanek and Doupe, 2010), the contribution of the striato- pallidal nucleus Area X to exploratory vocal variability has been controversial.

Previous studies that have lesioned Area X in zebra finches have failed to identify acute and lasting effects on variation in song tempo (Goldberg and Fee, 2011), sequence (Sohrabji et al., 1990; Scharff and Nottebohm, 1991), and acoustic features of syllables. For example, cross- rendition variability in FF was not lastingly altered by Area X lesions (Ali et al., 2013; Kojima et al., 2013; Fig. 1C, right; also see Results) and was unaffected by viral manipulations that selectively killed Area X medium spiny neurons (MSNs; Tanaka et al., 2016). Given that Area X is required for song learning in both juvenile and adult birds (Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Ali et al., 2013), these findings have motivated a hypothesis that exploratory vocal variability emerges in the thalamocortical [medial portion of the dorsolateral thalamus (DLM)–LMAN] circuit independently of Area X, whereas Area X has been hypothesized to have a role in learning which behavioral variants result in better outcomes (Fee and Goldberg, 2011; Goldberg and Fee, 2011).

Other lines of evidence, however, point to a role for Area X in actively generating song variability. Pallidal neurons in Area X (Fig. 1B, GPi) exhibit highly variable firing patterns during singing (Hessler and Doupe, 1999a; Goldberg et al., 2010; Woolley et al., 2014), and lesions of Area X mostly abolish the variable, singing- related burst firing in LMAN that has been hypothesized to drive behavioral variation (Kojima et al., 2013). Similarly, manipulations of dopamine- dependent signaling in Area X alter neural variability in LMAN and the social modulation of song variability (Leblois et al., 2010; Murugan et al., 2013), and elimination or acute manipulations of Area X neural activity can alter syllable variability at least transiently (Kojima et al., 2013; Heston et al., 2018). Together, these findings support a view that neural variability originating in Area X contributes to behavioral variability in song, contrary to results of previous lesion studies.

To test the role of Area X in generating exploratory vocal variability, we revisit the issue of whether or not lesions of Area X affect variability, focusing on rapid fluctuations in the temporal trajectory of FF within individual syllables (“within- syllable variability”; Fig. 1A, right), which has been linked to vocal learning (Charlesworth et al., 2011). We demonstrate here that Area X is required for the expression and modulation of this rapid behavioral variation, providing evidence that the BG are an essential source of exploratory variability that subserves trial- and- error learning.

## Materials and Methods

### Subjects

Subjects were juvenile and young- adult male zebra finches (*Taeniopygia guttata*), which were bred in our colony. Their care and treatment was reviewed and approved by the Institutional Animal Care and Use Com-

mittees at the University of California, San Francisco, and at the Korea Brain Research Institute.

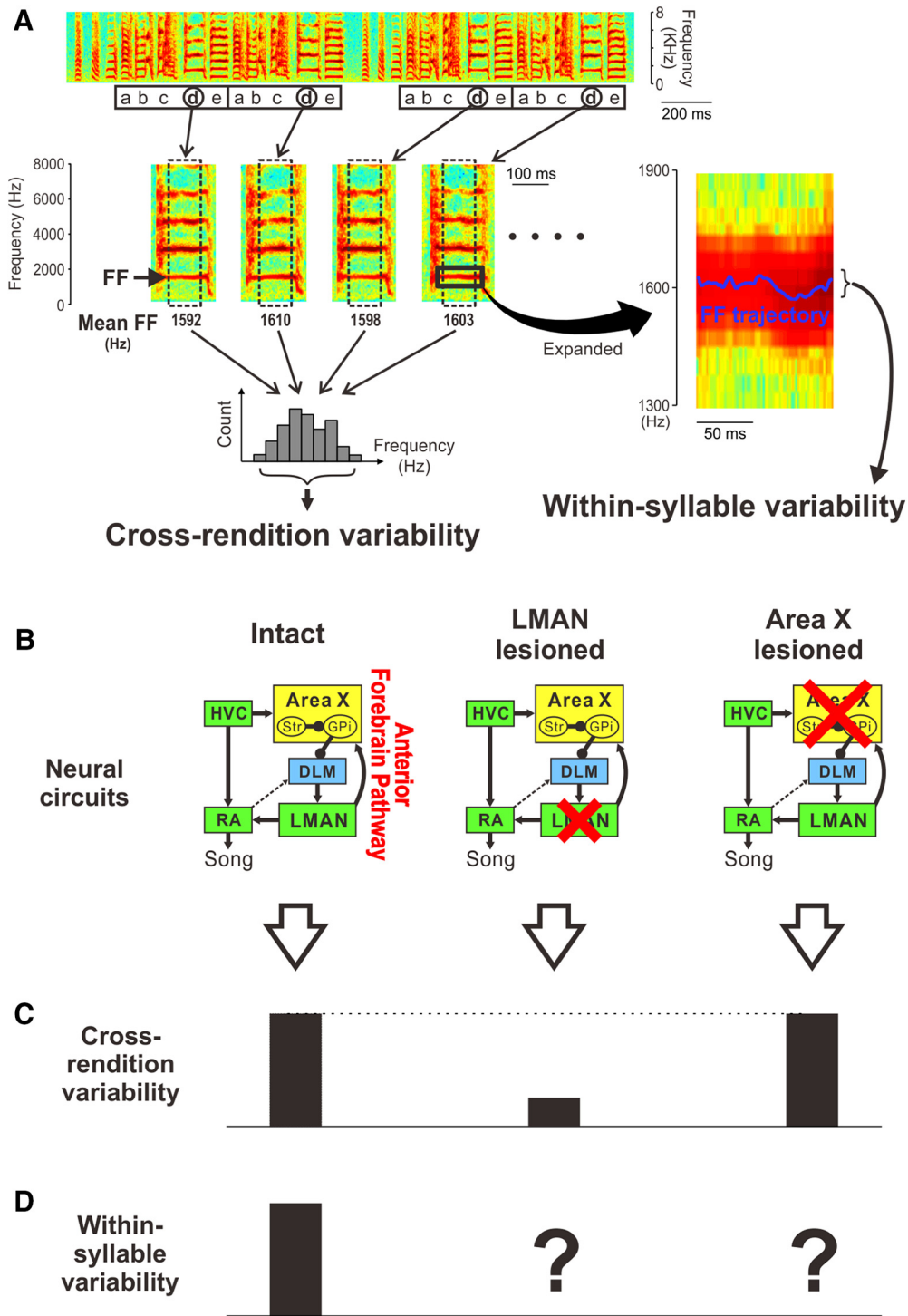
### Lesions

Under Equithesin anesthesia (Hessler and Doupe, 1999b), bilateral lesions of Area X were made in adult birds [104–118 d posthatch (dph)] or late juvenile birds (75 dph) by pressure injections of 1% ibotenic acid (0.1–0.2  $\mu$ l; PMI-100, Dagan; Kojima et al., 2013). Two to four injections were stereotaxically targeted to each side of the brain over 2 d at 1 d intervals. After all song recordings (see below), the birds were deeply anesthetized with isoflurane (Abbott Laboratories) and transcardially perfused with 0.9% saline, followed by 3.7% formaldehyde in 0.025 M phosphate buffer. Brains were postfixed, and 40  $\mu$ m sections were cut with a freezing microtome. Lesions of Area X were evaluated in brain sections labeled with an antibody to substance P (Accurate Chemical & Scientific; Carrillo and Doupe, 2004). To estimate lesion size, the remaining volume of Area X (average of right and left) was expressed as a percentage of the mean volume of intact Area X from six control birds. For all of the lesion birds, the percentage of Area X that was removed bilaterally ranged from 75 to 100%. We confirmed that LMAN was spared by measuring its volume and visually inspecting its axonal projections to the robust nucleus of the arcopallium (RA), using brain sections labeled with an antibody to calcitonin gene- related peptide (Millipore or Sigma- Aldrich; Bottjer et al., 1997). Sham lesions were made by injecting vehicle (0.2 M phosphate buffer) or a small volume of ibotenic acid into Area X, or by making small electrolytic lesions above LMAN. Analyses of the effects of Area X, LMAN, and sham lesions were performed on songs that were recorded from adult birds as part of previous studies (Kao and Brainard, 2006; Kojima et al., 2013).

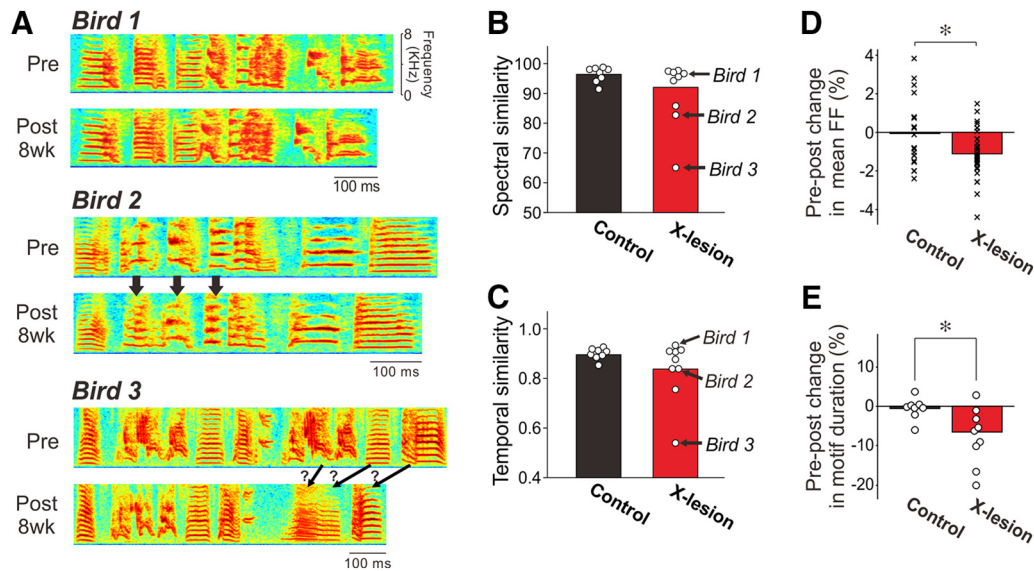
### Song analysis

To examine effects of brain lesions on song variability, we recorded songs 0–2 d before and 1–3 d and 2 weeks ( $\pm$  1 d) after Area X lesion surgery in the both female- directed (Dir) and undirected (Undir) conditions (>50 bouts of song in each condition). Undir song was defined as song produced while the male was alone in a sound- attenuating chamber. Dir song was obtained when a female zebra finch, housed in her own cage, was placed next to the male. Only songs produced when the male was facing the female were counted as Dir songs. Each female presentation lasted for  $\leq$  3 min, regardless of whether or not the male bird sang, and was repeated at least five times each day. An effort was made to record Dir and Undir song continuously throughout the day in an interleaved fashion. A small portion of syllables that substantially changed their phonology after Area X lesions (3 of 37 syllables; Fig. 2) was excluded from the analyses of song variability.

*Within- syllable vocal variability.* To examine effects of AFP lesions on song variability, we randomly chose >50 motifs each day. Within- syllable vocal variability was measured by examining the fluctuations of FF trajectories within individual syllable renditions. An FF trajectory was obtained in a segment of the sound waveform that had clear and flat harmonic structure as in a previous study (Charlesworth et al., 2011). Briefly, spectrograms were calculated using a Gaussian- windowed short- time Fourier transform ( $\sigma$  = 1 ms) sampled at 8 kHz, and a trajectory of the FF (the first harmonic frequency) was obtained by calculating the FF in individual time bins. For 45% of the syllables in Area X- lesion birds, 38% of the syllables in LMAN- lesion birds, and 50% of the syllables in control birds, the signal- to- noise ratio was better for the second harmonic than for the first harmonic, and in these cases the second harmonic frequency was used to quantify the FF trajectory; measurements made on the second harmonic were divided by two to provide an estimate of the FF. For each syllable, all renditions of FF trajectories were aligned by the onset of the syllables, based on amplitude threshold crossings, and flat portions ( $\geq$  20 ms) of FF trajectories were used for further analysis. To quantify fluctuations in FF trajectories, we removed the modulation of FF trajectories that was consistent across renditions by calculating residual FF trajectories as percent deviation from the mean trajectory across renditions, and then in each FF trajectory the within- rendition mean was subtracted (Fig. 3B). Fluctuations in FF trajectories were characterized by examining both their magnitude and frequency.



**Figure 1.** Song variability at multiple timescales and underlying neural circuits. **A**, Variation in song acoustic structure on different timescales. Top, A spectrogram illustrating four motifs of an example song consisting of syllables a–e (squares indicate motifs). In many previous studies, “cross-variation variability” in FF was used to examine song variability (left). This variability is obtained for a particular syllable (d in this example) by computing mean FF of the flat harmonic portion of individual syllable renditions (dotted squares, middle) and by measuring variations in this feature across renditions (bottom left). In contrast, the present study focuses on “within-syllable variability” (right), the rapid fluctuations of FF trajectories within individual syllable renditions (blue trace on the spectrogram that is an expanded view of the square area in the syllable on the left). **B–D**, Schematic representations of the effects of lesions in the AFP on song variability. The neural circuit underlying vocal learning and production (**B**) and magnitudes of cross-variation (**C**) and of within-syllable variability (**D**) are shown. In the circuit diagrams, red crosses indicate lesions; yellow, blue, and green boxes indicate BG, thalamic, and cortical structures, respectively. Str, Striatum; GPI, internal segment of the globus pallidus. HVC is used as a proper name, and arrowheads and filled circles indicate excitatory and inhibitory neural projections, respectively. Note that lesions of LMAN (middle) have been shown to dramatically decrease cross-variation variability compared with that in intact birds (left), whereas lesions of Area X (right) do not have sustained effects on cross-variation variability. The effects of Area X and LMAN lesions on within-syllable variability have not been previously examined (question mark in the diagram) and are the focus of this study.



**Figure 2.** Subtle effects of Area X lesions on acoustic and temporal structure in adult song motifs. **A**, Spectrograms illustrate a variable degree of changes in the motif structure of Area X-lesion birds. For each bird, a motif of pre-lesion (Pre) and 8 week (8wk) post-lesion (Post 8wk) song are shown. Arrows indicate apparent changes in syllable structure. **B, C**, Summary of spectral similarity (**B**) and temporal similarity (**C**) between pre-lesion and 8 week post-lesion song motifs (see Materials and Methods). Some of these data were published previously (Kojima et al., 2013) and are replotted here with new data. Each point corresponds to one bird; birds 1–3 shown in **A** are indicated with arrows. Note that although the majority of the birds maintained the motif structure as much as control birds, a subset of birds showed noticeable changes in both spectral and temporal structure of their song motifs. **D**, Area X lesions lowered mean FF of song syllables slightly. Percent changes in mean FF in individual syllables (averaged across renditions) from pre-lesion song to post-8 week song are shown. Each point corresponds to one syllable.  $*p < 0.02$ . **E**, Area X lesions shortened song motifs, as reported previously (Kubikova et al., 2014). Pre–post percent changes in motif duration (averaged across renditions) are shown. Each point corresponds to one bird.  $*p < 0.02$ .

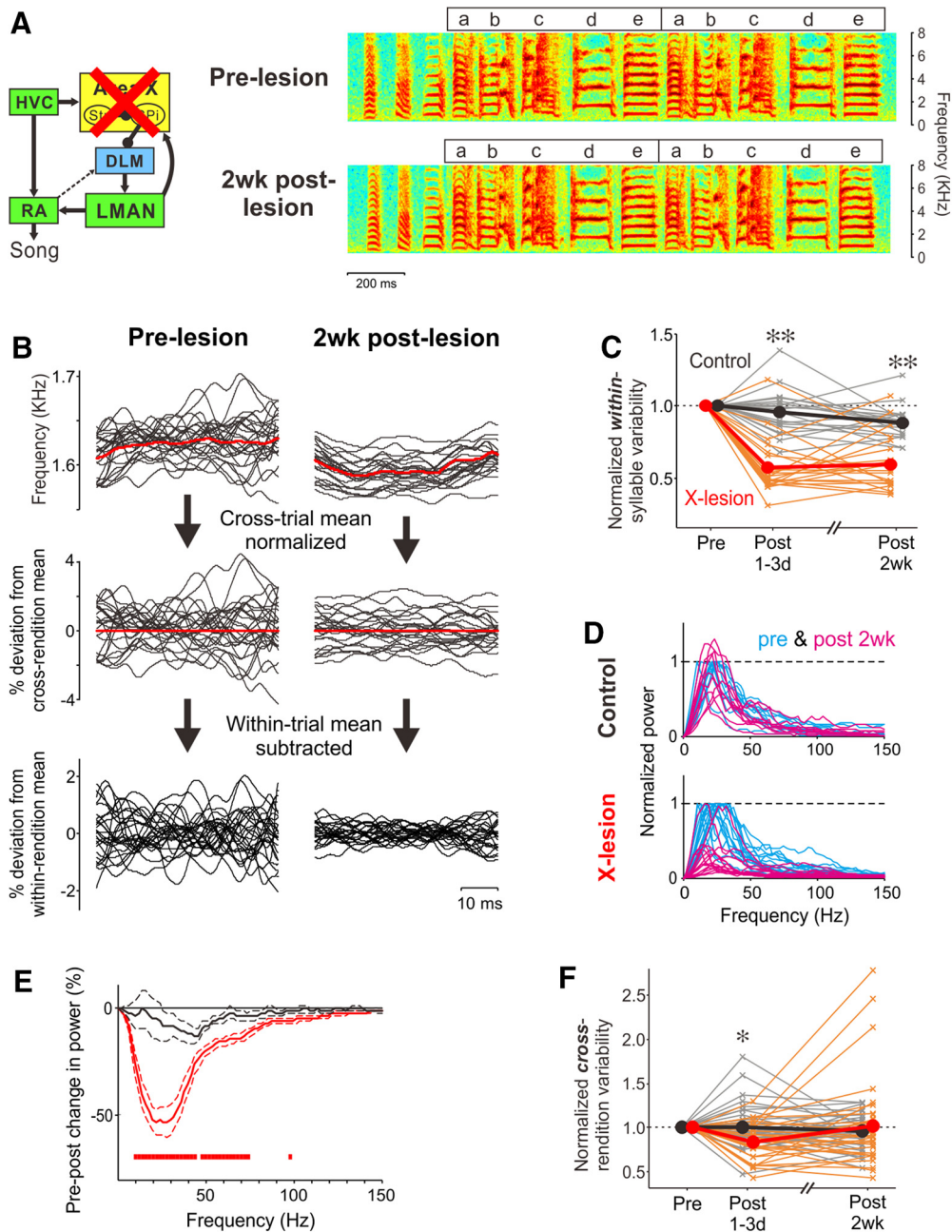
The magnitude was quantified by calculating the within-rendition SD of individual FF trajectories and averaging the SDs across renditions. For frequency analysis of FF-trajectory fluctuations, the power spectral density was computed for each FF trajectory with the FFT and complex conjugate functions in Matlab (RRID:SCR\_001622). Only FF trajectories  $\geq 30$  ms were used for this frequency analysis to ensure accurate power estimation in the range corresponding to fluctuations in FF trajectories (with peak power typically  $< 50$  Hz). To reduce the influence of outliers that could result from occasional background noise included in our song recordings, the median power spectral density of all FF trajectories for a given syllable was used for further analyses.

To examine changes in power spectral density after lesions, a pre–post change in the median power spectrum normalized by the peak height of the pre-lesion power spectrum was calculated for each syllable, and the significance of the difference between lesion and control birds was determined as follows. We first computed the magnitude of the power differences between lesion and control birds in each frequency bin (2 Hz spacing, ranging 0–150 Hz) using the  $d'_{\text{lesion-control}}$  metric (Doupe and Solis, 1997; Coleman and Mooney, 2004; Green and Swets, 2012). The  $d'_{\text{lesion-control}}$  was calculated using the following formula:  $d' = 2 * (\mu_{\text{lesion}} - \mu_{\text{control}}) / (\sigma_{\text{lesion}}^2 + \sigma_{\text{control}}^2)^{1/2}$ , where  $\mu_{\text{lesion}}$  and  $\mu_{\text{control}}$  refer to the mean of normalized power in lesion and control birds, respectively, and where  $\sigma_{\text{lesion}}^2$  and  $\sigma_{\text{control}}^2$  refer to the variance of normalized power in lesion and control birds, respectively. This measure takes into account both the mean difference in power spectra between the lesion and control groups and the variability of power spectra across different syllables to provide an indication of the discriminability between the lesion and control groups. To determine the significance of the  $d'_{\text{lesion-control}}$  values in individual frequency bins, we used a Monte Carlo randomization procedure to estimate the probability of obtaining specific  $d'_{\text{lesion-control}}$  values by chance (Sakata and Brainard, 2008). For this process, we randomly assigned, without replacement, each syllable into one of two groups representing the lesion and control groups, while conserving the sample size for each group, and calculated  $d'$  values for the shuffled dataset. This process was repeated 1000 times. When  $d'_{\text{lesion-control}}$  was greater than the 99th percentile of this distribution, we categorized this difference as significant (Fig. 3E, red bars). This statistical procedure was validated by estimating the probability that pre–post differences in power between two sets of syllables cross the threshold for

significance by chance. We randomly shuffled, without replacement, the waveform of the pre–post differences in power spectra (ranging 0–150 Hz) of each syllable in both Area X-lesion and control birds into one of two groups representing the Area X-lesion and control groups, while conserving the sample size for each group, and determined frequency bins in which the  $d'_{\text{lesion-control}}$  crossed the threshold for significance using the method described above. We repeated this process 1000 times and found that the  $d'_{\text{lesion-control}}$  crossed the threshold in 1.7% of the frequency range that we examined, on average. This confirms that our statistical procedure is equivalent to setting  $\alpha \approx 0.02$ .

**Cross-rendition vocal variability.** Cross-rendition variability of the mean FF for individual syllables was measured as described previously (Kao et al., 2005; Kao and Brainard, 2006; Stepanek and Doupe, 2010). Briefly, the auto-correlation of the harmonic sound segment was calculated, and the mean FF of the segment was defined as the reciprocal of the delay between the zero-offset peak and the highest peak in the auto-correlation function. Since many syllables show slow changes in mean FF over time during a day, we also normalized FF variability by the diurnal change as described previously (Kojima et al., 2013): for each syllable rendition, the mean FF value was divided by the mean FF averaged across all syllable renditions in a 1 h window around the target syllable (i.e.,  $\pm 30$  min), and the SD of all the normalized FF values during a given day was used as a measure of cross-rendition variability in mean FF.

**Changes in song structure.** We quantified changes in song structure after Area X lesions on the basis of song motifs. Since changes in song structure progressed gradually, we compared pre-lesion and 8 week ( $\pm 1$  week) post-lesion song motifs. Only songs recorded in Undir context were analyzed. A song motif was defined as a sequence of syllables that occurred most frequently in pre-lesion and in post-lesion songs. To assess changes in song spectral structure after Area X lesions, we randomly chose 5 pre-lesion and 10 post-lesion motifs and made pairwise comparisons between all possible pre-lesion versus post-lesion motif combinations (i.e., 50 comparisons; Kojima et al., 2013). Spectral similarity of song motifs was quantified using the “% similarity” measure of Sound Analysis Pro (SAP; version 1.04; Tchernichovski et al., 2000). Default



**Figure 3.** Area X is required for within-syllable vocal variability. **A**, Circuit diagram indicating lesions of Area X (left) and spectrograms of songs of an adult zebra finch recorded before and after 2 week (2wk) bilateral lesions of Area X (right). Conventions are as in Figure 1, *A* and *B*. **B**, Examples of FF trajectories of syllable e in pre-lesion (left) and 2 week post-lesion (right) songs, expressed as raw frequency traces (top), percent deviation from cross-rendition mean (middle), and percent deviation from within-rendition mean (bottom). Twenty consecutive trajectories are shown. Red traces indicate across-rendition means. **C**, The magnitude of within-syllable FF variability (within-rendition SD of mean-subtracted FF trajectory, averaged across renditions) normalized by that of pre-lesion song in Area X-lesion (red) and control (black) birds. Thin lines correspond to one syllable, and thick lines indicate the mean across all syllables.  $**p < 10^{-4}$ . **D**, Power spectra of FF trajectories pre-lesion (blue) and 2 weeks post-lesion (magenta) in control (top) and Area X-lesion (bottom) birds, normalized by the peak height of the spectrum in pre-lesion song; each line corresponds to one syllable. **E**, Changes in mean ( $\pm$ SEM) power spectrum of FF trajectories from pre-lesion to 2 weeks post-lesion, normalized to the peak height of the presong spectrum, in Area X-lesion (red) and control (black) birds. The red bars indicate the frequency ranges where Area X-lesion data were significantly different from control data (see Materials and Methods). **F**, For the same syllables, Area X lesions did not permanently eliminate cross-rendition variability in FF. Conventions are as in **C**.  $*p < 0.01$ .

values in SAP were used for all parameters except the time-warping tolerance, which was set to 20%.

To assess changes in song temporal structure, we compared the log-amplitude envelopes of 5 pre-lesion motifs and 10 post-lesion bouts (Kojima et al., 2013). We searched post-lesion bouts for amplitude patterns that provided the closest temporal match to the pre-lesion motif by calculating a cross-correlation function and taking the maximum value of the function. When searching for the maximum correlation, we allowed for proportional changes in the temporal pattern of songs

( $\pm 20\%$ ). A part of the data of song changes after Area X lesions was published previously (Kojima et al., 2013) and replotted with new data. For those data, mean FF of harmonic syllables and mean motif durations were also measured, and the effects of Area X lesions were examined.

*Microstimulation in LMAN*

For microstimulation in LMAN, custom-made bipolar tungsten electrodes (Microprobes; 100–500 kOhm, separated 200–400  $\mu$ m apart) were surgically implanted in each side of the brain. The electrodes were

targeted at LMAN using antidromic stimulation in the downstream nucleus RA. The locations of electrodes were verified histologically after each experiment. Electrical stimuli consisted of 10–50 ms trains of biphasic current pulses at 400 Hz (0.4 ms per phase, 2.5 ms between phases; Vu et al., 1994; Kao et al., 2005) and were delivered by a stimulator (A-M Systems model 2100) through wires connected to the implanted electrodes. To reproducibly inject current in LMAN at specific times in song, we detected the preceding syllables by comparing the bird's vocalizations with predefined spectral templates using custom-written software (R. O. Tachibana, University of Tokyo, Tokyo, Japan) and delivered stimulation at a fixed delay from the onset of the preceding syllables. For each experiment, the current amplitude used was the minimum intensity that elicited consistent effects on acoustic structure of the target syllables (40–400  $\mu$ A). Longer stimulations tended to require lower current amplitudes; the median intensities for each stimulation duration were 175  $\mu$ A (10 ms), 120  $\mu$ A (20 ms), and 75  $\mu$ A (50 ms), and this may account for the longer latencies observed in some of experiments with 50 ms stimulus durations (the four experiments with the longest latencies had stimulation intensities that were  $<80$   $\mu$ A). Stimulation (stim) trials were randomly interleaved with catch trials in which no stimulation was delivered (no-stim trials). In cases where stimulation caused deflection of FF trajectories without apparent disruptions of harmonic structure, at least 25 stim trials and 25 no-stim trials were collected for each syllable, and the raw FF trajectories of both stim and no-stim trials were normalized to the cross- rendition mean of no-stim trials (see Fig. 6C,H). Frequency analyses of the normalized FF trajectories were made as described above for fluctuations in FF trajectories. Latencies of LMAN stimulation-evoked FF deflections were defined for the normalized FF trajectories as the time interval between the stimulus onset and the onset of FF deflections, which was determined as the first time point where the absolute value of the cross-trial average of FF trajectories in stim trials exceeded the mean FF trajectories in no-stim trials by 2 SDs (see Fig. 6C,H, arrows).

#### Experimental design and statistical analysis

To determine whether lesions of Area X influence within-syllable FF variability (Fig. 3C) and cross- rendition FF variability (Fig. 3F), we compared the magnitude of variability between lesion and control birds using a Mann–Whitney *U* test (22 syllables in nine lesion birds and 18 syllables in seven control birds). To characterize the timescale of within-syllable variability that depends on Area X (Fig. 3E), we compared pre–post changes in the power spectral density of FF fluctuations between Area X-lesion and control birds (16 syllables in nine lesion birds and 10 syllables in five control birds), and the significance of difference was determined as described above. To determine whether within-syllable FF variability varies across social contexts (Fig. 4B), we compared the magnitude of variability across social contexts using a Wilcoxon signed-rank test (52 syllables in 18 intact birds). We also used the same test to examine whether lesions of Area X influence social context-dependent modulation of within-syllable or cross- rendition variability (Fig. 4D,E; 22 syllables in nine lesion birds and 18 syllables in seven control birds). To determine whether lesions of LMAN influence within-syllable FF variability (Fig. 5B), we compared the magnitude of variability between lesion and control birds using a Mann–Whitney *U* test (16 syllables in five lesion birds and 18 syllables in seven control birds). To determine whether within-syllable FF variability decreases with age (see Fig. 7), we used a Wilcoxon signed-rank test (12 syllables in two juvenile birds and 18 syllables in seven adult birds). All statistical analyses were performed using Matlab (RRID:SCR\_001622).

## Results

### Area X is required for within-syllable variability in song acoustic structure

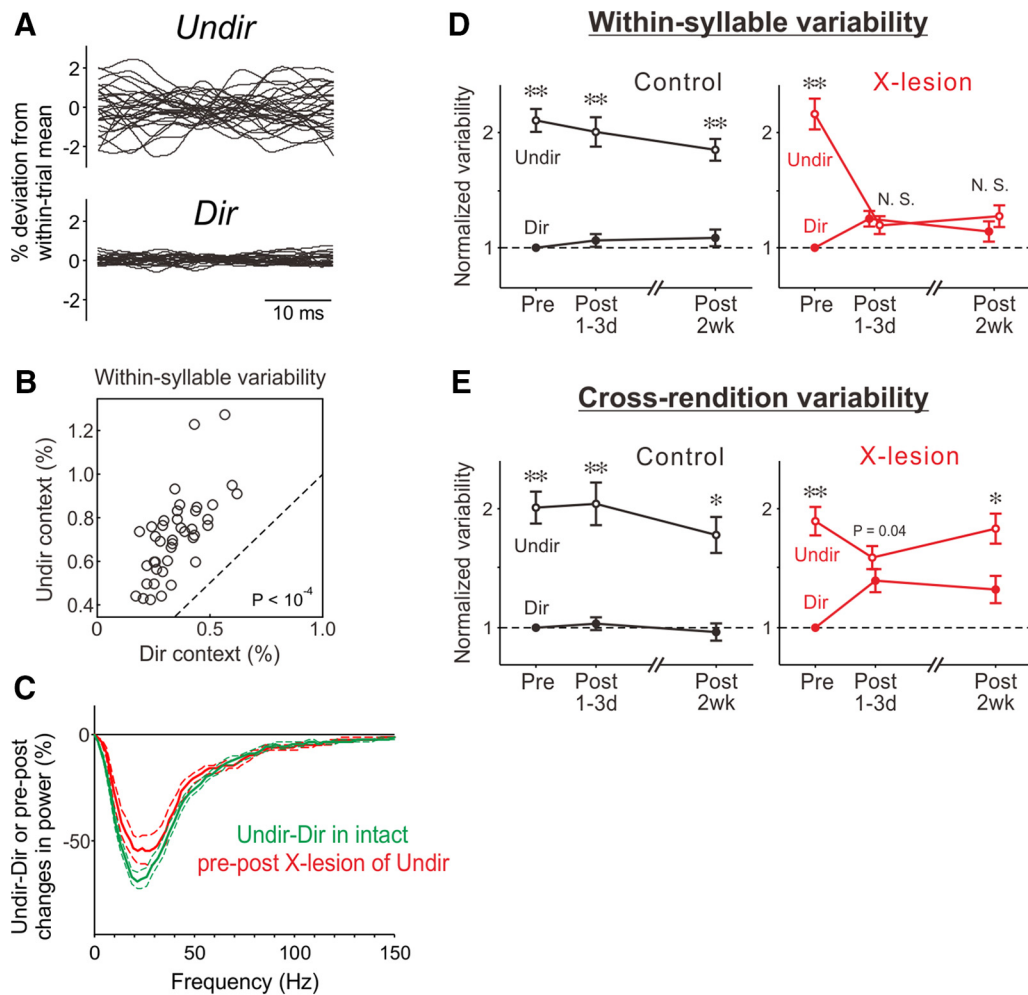
To examine the contribution of the BG nucleus Area X to vocal variability, we made bilateral lesions of Area X in adult male zebra finches at 102–118 dph. Consistent with previous studies (Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Ali et al., 2013; Kojima et al., 2013; Kubikova et al., 2014), Area X lesions

did not grossly disrupt the spectral structure of individual song syllables or the stereotyped sequence of song syllables (Figs. 2A, 3A). Songs remained easily recognizable for many weeks after lesions of Area X, although subtle changes in tempo and acoustic structure could emerge gradually over time (Fig. 2).

In contrast to prior studies that examined variability of mean FF across renditions of individual syllables (“cross- rendition” variability; Fig. 1A, left; Kao et al., 2005; Kao and Brainard, 2006; Leblois et al., 2010; Stepanek and Doupe, 2010; Ali et al., 2013; Murugan et al., 2013), we focused on vocal variability within individual syllable renditions (“within-syllable” variability; Fig. 1A, right), which has been linked to learning in response to aversive reinforcement (Charlesworth et al., 2011). For syllables with flat harmonic structure, we measured an “FF trajectory” within each syllable rendition using a short, sliding window, as described previously (Charlesworth et al., 2011; see Materials and Methods), and examined the temporal structure of fluctuations within individual FF trajectories. The black lines in Figure 3B (top left) illustrate the FF trajectories for 20 consecutive renditions of an individual syllable (Fig. 3A, e). These trajectories illustrate typical fluctuations in FF on a millisecond timescale, in contrast to the much longer timescale of cross- rendition variability (successive renditions of a given syllable are typically hundreds of milliseconds apart). Previous work has shown that this form of rapid, within-syllable fluctuation in FF can be linked to song learning in adults; the specific, idiosyncratic fluctuations on this timescale that are reinforced during learning become incorporated in lasting changes to song (Charlesworth et al., 2011).

We found that within-syllable FF variability in adult birds singing alone (“undirected” singing) dramatically decreases after Area X lesions (Fig. 3B). To measure within-syllable variability in FF, we first normalized individual FF trajectories (Fig. 3B, top) to their cross- rendition mean (Fig. 3B, middle) and then subtracted the within- rendition mean for each trajectory (Fig. 3B, bottom). Qualitatively, the magnitude of within- rendition fluctuations in FF was greater in intact birds (Fig. 3B, bottom left) than after lesions of Area X (Fig. 3B, bottom right). We quantified the magnitude of fluctuations in mean-subtracted FF trajectories as the within- rendition SD for each trajectory, averaged across all trajectories of the same syllable. This measure of within-syllable variability was significantly smaller in birds with Area X lesions ( $n = 22$  syllables in nine Area X-lesion birds) than in control birds [ $n = 18$  syllables in seven control birds (four intact and three sham-lesion birds combined)], both at 1–3 d and at 2 weeks post-lesion [Fig. 3C; Post 1–3d,  $p = 3.56 \times 10^{-6}$ ; Post 2wk,  $p = 3.38 \times 10^{-5}$ ; Mann–Whitney *U* test; there were no significant differences between sham-lesion and intact control birds at either Post 1–3d ( $p = 0.79$ ) or at Post 2wk ( $p = 0.25$ )]. These findings could not be attributed to differences between experimental and control groups in the time of day at which recordings were made (there was no significant difference in average recording times of day between pre-2 week and post-2 week data in Fig. 3C;  $p = 0.38$ , Mann–Whitney *U* test), or the FF of the analyzed syllables from control and X-lesion birds ( $p = 0.75$ , Mann–Whitney *U* test). These results demonstrate that the BG critically contribute to rapid, within-syllable variability in FF.

To further characterize the timescale of within-syllable variability that depends on Area X and to investigate the underlying neural mechanisms, we computed the power spectra of individual FF trajectories (mean subtracted; Fig. 3B, bottom) and compared the median spectra between pre-lesion and post-lesion song for each syllable (only FF trajectories  $\geq 30$  ms in duration were used for this analysis; see Materials and Methods). In the



**Figure 4.** Social context modulation of within-syllable variability depends on Area X. **A**, Examples of 20 consecutive FF trajectories of a single syllable produced by an intact bird in Undir and Dir contexts show context-dependent modulation of within-syllable FF variability. **B**, Scatter plots of the magnitude of within-syllable FF variability comparing Undir context with Dir context. Each point corresponds to one syllable. The dashed line indicates unity.  $p < 10^{-4}$ . **C**, Comparison of the effects of social context versus lesions of Area X on within-syllable FF variability. Green, Changes (mean  $\pm$  SEM) in the power spectrum of FF trajectories from Undir to Dir contexts in intact birds, normalized to the peak height of Undir song spectrum; red, pre–post changes in the power spectrum of Undir song in Area X-lesion birds (data from Fig. 3E replotted for comparison). Conventions are as in Figure 3E. Note the absence of significant differences between Undir–Dir data and Area X-lesion data. **D**, Within-syllable FF variability in Dir and Undir contexts in control (left) and Area X-lesion (right) birds; variability at all time points and in both contexts were normalized by that of Dir syllables in pre-lesion song. Error bars are SEM.  $*p < 10^{-3}$ ,  $**p < 10^{-4}$ . **E**, Cross- rendition FF variability in Dir and Undir contexts in control (left) and Area X-lesion (right) birds. Conventions are as in **D**.

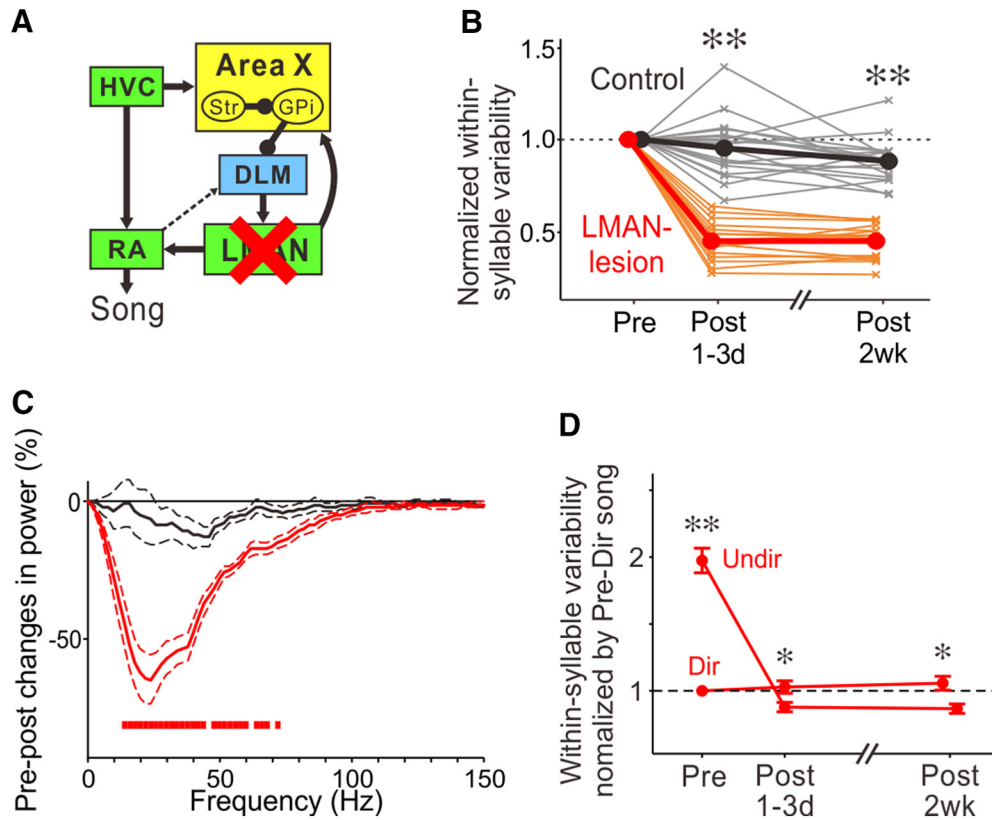
songs of intact birds, the median power spectra of FF trajectories had peaks at  $\sim 20$ – $30$  Hz for most syllables (Fig. 3D, blue traces, normalized to peak values), indicating that FF trajectories fluctuate predominantly in this frequency range (corresponding to  $\sim 30$ – $50$  ms). Area X lesions dramatically decreased FF-trajectory fluctuations in this frequency range (Fig. 3D, magenta traces, normalized to the peak of pre-lesion power spectra). The average reduction in power in the frequency range of 10–74 Hz from pre-lesion to 2 weeks post-X lesion was significantly greater than that for control birds over the same period [Fig. 3E, red bars; see Materials and Methods for significance testing;  $n = 16$  syllables in nine Area X-lesion birds and 10 syllables in five control birds (three intact birds and two sham-lesion birds)].

This marked and sustained reduction in within-syllable FF variability after lesions of Area X contrasts with a transient reduction in cross- rendition FF variability (Fig. 1A, left) after lesions of Area X [Fig. 3F; see also Ali et al. (2013) and Kojima et al. (2013)]. In the same syllables that exhibited a long-lasting reduction of within-syllable variability, the cross- rendition variability was significantly reduced at 1–3 d post-lesion ( $p = 0.003$ , Mann–Whit-

ney  $U$  test) but recovered to pre-lesion levels within 2 weeks ( $p = 0.34$ ). Thus, slower, cross- rendition variability can be generated independently of the BG, consistent with a previous report (Ali et al., 2013). These results demonstrate differential requirements of Area X for rapid within-syllable variability versus cross- rendition variability in FF.

#### Social context-dependent modulation of within-syllable variability depends on Area X

In Area X, as well as in the downstream nucleus LMAN, many neurons show singing-related patterns of activity that are modulated by social context. In particular, pallidal neurons in Area X exhibit prominent pauses in their tonic firing that are highly variable across renditions when male birds sing alone (Undir singing) but fewer and more stereotyped pauses when they sing to females (Dir singing; Hessler and Doupe, 1999a; Woolley et al., 2014). In parallel with this social context-dependent modulation of firing patterns in the BG, cross- rendition variability in FF is greater during Undir singing than during Dir singing, leading to a hypothesis that Undir singing reflects motor exploration,



**Figure 5.** LMAN is required for within-syllable variability. **A**, Circuit diagram indicating lesions of LMAN. **B**, Within-syllable FF variability in LMAN-lesion (red) and control (black) birds in Undir context; conventions are as in Figure 3C.  $**p < 10^{-5}$ . **C**, Pre–post changes in the power spectrum of FF trajectories in LMAN-lesion (red) and control (black) birds in Undir context. Conventions are as in Figure 3E. **D**, Within-syllable FF variability in Dir and Undir contexts in LMAN-lesion birds. Conventions are as in Figure 4D.  $*p < 0.02$ ;  $**p < 10^{-3}$ .

whereas Dir singing reflects motor exploitation, or the repetition of a successful behavior, by which male birds attract female birds (Jarvis et al., 1998; Kao and Brainard, 2006; Kojima and Doupe, 2011). Although this striking parallel between neural and behavioral variability supports the view that neural variability in Area X drives vocal variability, the absence of lasting changes in cross- rendition vocal variability after Area X lesions apparently contradicts this view.

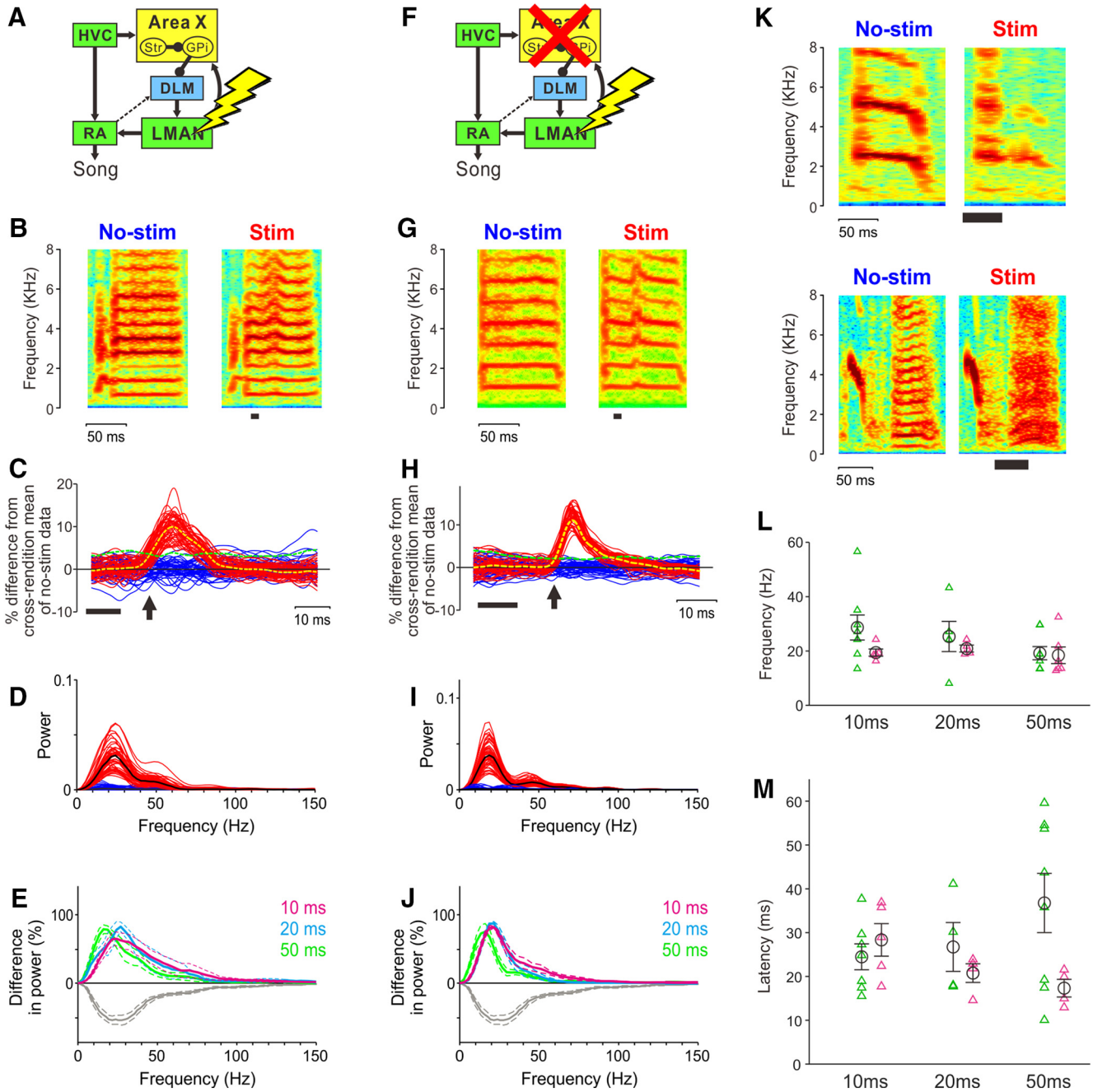
We were therefore interested in the possibility that the social modulation of neural activity in Area X might be linked to a corresponding social modulation of within-syllable variability. In intact birds, we found that the level of within-syllable variability present during Undir singing is reduced during Dir singing (Fig. 4A,B; 52 syllables in 18 birds;  $p = 4.01 \times 10^{-5}$ , Wilcoxon signed-rank test). This social context-dependent reduction in within-syllable variability was comparable with the reduction in within-syllable variability after Area X lesions: the average reduction in power between Undir and Dir songs had a frequency range similar to that of the reduction in power after Area X lesions in Undir song (~20–30 Hz), and they were not significantly different (Fig. 4C; see Materials and Methods for significance testing). Moreover, lesions of Area X eliminated the social modulation of within-syllable variability by reducing the fluctuations in FF during Undir song to the level present during Dir song (Fig. 4D; in Area X-lesion birds,  $p = 4.01 \times 10^{-5}$ , 0.29, and 0.12 for Pre, Post 1–3d, and post 2wk, respectively; in control birds,  $p = 1.96 \times 10^{-4}$  for Pre, Post 1–3d, and Post 2wk; Wilcoxon signed-rank test). These findings strongly support the view that the social modulation of neural variability in Area X pallidal neurons drives social modulation of within-syllable variability,

thus further indicating a critical role of Area X in generating within-syllable variability. In contrast to within-syllable variability, significant social modulation of cross- rendition variability in FF was present even in the absence of Area X (Fig. 4E; in Area X-lesion birds,  $p = 5.30 \times 10^{-5}$ , 0.042, and  $7.62 \times 10^{-3}$  for Pre, Post 1–3d, and Post 2wk, respectively; in control birds,  $p = 2.33 \times 10^{-4}$ ,  $1.96 \times 10^{-4}$ , and  $1.01 \times 10^{-3}$  for Pre, Post 1–3d, and Post 2wk, respectively; Wilcoxon signed-rank test), indicating that social modulation of Area X neural activity is not required for this slower, cross- rendition acoustic variation. These results further highlight differential contributions of Area X to rapid within-syllable variability and cross- rendition variability in FF.

#### Within-syllable variability and its social modulation require LMAN

Since previous anatomical and physiological studies indicate that Area X projects to the song motor pathway (the RA) predominantly through LMAN, the output nucleus of the AFP (Okuhata and Saito, 1987; Bottjer et al., 1989; Mooney and Konishi, 1991; Gale and Perkel, 2010; Hamaguchi and Mooney, 2012), we predicted that Area X contributes to within-syllable variability and its social modulation via LMAN. If so, bilateral lesions of LMAN should also decrease the magnitude of within-syllable variability in Undir song and abolish social modulation of within-syllable variability. Indeed, we found that the magnitude of within-syllable FF variability in Undir song decreases after LMAN lesions, as much as after Area X lesions (Fig. 5A,B; 16 syllables in five birds;  $p = 7.38 \times 10^{-7}$  for both 1–3d and 2wk post-lesion, Mann–Whitney  $U$  test). The frequency range of LMAN-





**Figure 6.** Microstimulation in LMAN can elicit rapid deflections of FF trajectories in both intact birds and birds with Area X lesions. **A**, Circuit diagram showing electrical microstimulation in LMAN of an intact bird. **B**, Example of a syllable with (right; stim) and without (left; no-stim) LMAN stimulation in an intact bird. Stimulation (10 ms duration) is indicated by the black bar. **C**, Interleaved FF trajectories in stim (red) and no-stim (blue) trials for the syllable shown in **B**, normalized to the cross-rendition mean of no-stim trajectories. Dashed lines in yellow and green indicate the mean of FF trajectories in stim trials and 2 SDs of FF trajectories in no-stim trials, respectively; the thick bar indicates the timing of LMAN stimulation; and the arrow indicates the onset of FF deflections (see Materials and Methods). **D**, Power spectra of FF trajectories with and without LMAN stimulation for the syllable shown in **B** and **C**. Black lines indicate the median in each condition. **E**, Percent difference in power between median power spectra with and without stimulation (mean  $\pm$  SEM;  $n = 11$  syllables from 8 intact birds in total; colors indicate durations of LMAN stimulation (10, 20, and 50 ms). As a comparison, percent changes in the power spectrum of naturally occurring FF fluctuations after Area X lesions are also plotted (gray lines; the same data as in Fig. 3E; the direction of the difference in power is opposite between LMAN stimulation data and Area X lesion data because LMAN stimulations increase power spectra and Area X lesions decrease power spectra). Note that stimulation-driven deflections of FF trajectories have a timescale comparable with that of Area X-dependent fluctuations. **F–J**, The same experiment in birds with Area X lesions ( $n = 6$  syllables from 4 birds). Conventions are as in **A–E**. **K**, Examples of two syllables in which LMAN stimulations disrupted the harmonic structure. **L**, Peak frequencies of stimulation-dependent power spectra in individual syllables of intact (green) and lesion (magenta) birds. No significant differences in peak frequencies were observed for all the stimulation durations ( $p = 0.11, 0.24, \text{ and } 0.81$  for 10, 20, and 50 ms, respectively). **M**, Latencies of FF deflections in intact (green) and lesion (magenta) birds. No significant differences were observed for all the stimulation durations ( $p = 0.53, 0.69, \text{ and } 0.21$ , respectively).

dependent within-syllable FF variability was similar to that of Area X-dependent within-syllable FF variability (compare Figs. 5C and 3E). Moreover, social modulation of within-syllable FF variability was mostly abolished by LMAN lesions; within-

syllable variability in Undir song decreased after LMAN lesions to levels slightly below the level of variability in Dir song (Fig. 5D). These results demonstrate that much of the Area X-dependent within-syllable FF variability and its social modulation also re-

quire LMAN and suggest that Area X drives within-syllable variability by influencing the downstream thalamocortical (DLM–LMAN) circuit.

### Brief stimulation of LMAN can elicit deflections of FF trajectories on a timescale comparable with the timescale of within-syllable variability

Our results support a model in which the elevated within-syllable variability in FF present during Undir song is driven by variable burst firing in LMAN during Undir singing (Hessler and Doupe, 1999b; Kao et al., 2008), which in turn derives from variable firing patterns in Area X (Kojima et al., 2013; Woolley et al., 2014). In particular, burst firing in LMAN (1) requires Area X (Kojima et al., 2013); (2) is prominent during Undir singing (Kao et al., 2008), when within-syllable variability is higher (Fig. 4A, top, B); and (3) dramatically decreases during Dir singing (Kao et al., 2008), when within-syllable variability is low (Fig. 4A, bottom, B). Given that LMAN neurons send excitatory projections to the song motor nucleus RA (Mooney and Konishi, 1991; Stark and Perkel, 1999) and that altering LMAN activity can induce real-time changes in song structure (Kao et al., 2005; Giret et al., 2014), it is likely that burst firing in LMAN drives within-syllable variability in FF by driving rapid variation in song premotor activity in RA.

To test this idea, we examined whether brief stimulation of LMAN neurons could cause rapid fluctuations of FF trajectories similar in timescale to those normally present in Undir song. Using microelectrodes implanted in freely moving intact birds, we bilaterally stimulated LMAN at specific times in song by delivering a burst of biphasic current pulses with 10, 20, or 50 ms duration. For most syllables that we examined (11 of 14 syllables in nine birds), we observed consistent deflections of FF trajectories in response to LMAN stimulation with at least one of the three stimulus durations (10, 20, and 50 ms; Fig. 6A–C, Table 1). Upward deflections were observed in eight syllables from five birds, and downward deflections were observed in three syllables from three birds. In some cases, LMAN stimulation disrupted the harmonic structure of targeted syllables (Fig. 6K), and these syllables were excluded from further analysis. These results were qualitatively consistent with previous studies that found that LMAN stimulation (0.2–550 ms duration) could cause short latency changes to the mean FF of individual syllables, raising the possibility that LMAN activity could contribute to cross-rendition variation in FF (Kao et al., 2005; Giret et al., 2014).

Here, we were additionally interested in whether LMAN could contribute to the within-syllable fluctuations in FF that depend on Area X, and we therefore focused on the detailed temporal structure of FF deflections that were introduced by our brief burst stimulation of LMAN. We quantified the timescale of the rapid FF deflections induced by LMAN stimulation by computing the power spectra of individual FF trajectories and comparing the median spectra between stim and catch (no-stim) trials (Fig. 6D). For all the stimulus durations (10, 20, and 50 ms), the peak frequency of the stimulation-evoked deflections of FF trajectories was ~20–30 Hz (Fig. 6E, magenta, cyan, and green), similar to that of Area X-dependent within-syllable variability in FF (Fig. 6E, gray; the same data as in Fig. 3E). These results show that brief stimulation of LMAN neurons can evoke rapid FF fluctuations that are similar in timescale to those normally present in Undir song.

We observed that the latencies of LMAN stimulation effects on FF trajectories were generally short (median, 20 ms; range, 10–60 ms) and likely reflected an upward bound, since in each

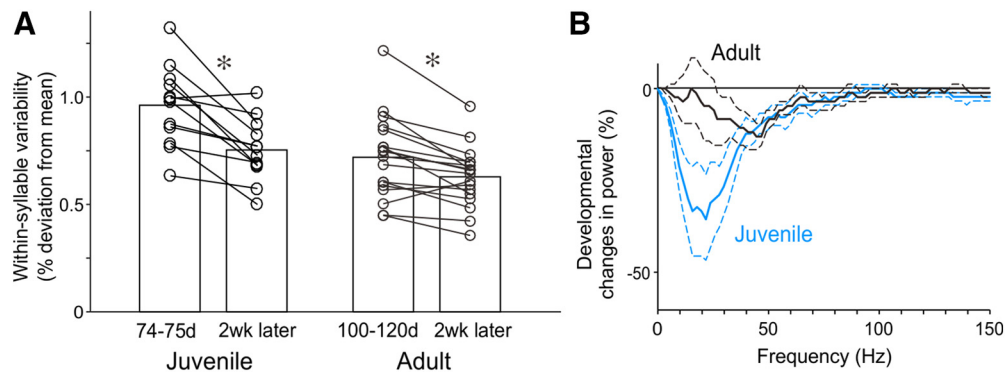
**Table 1. Responses of FF trajectories to LMAN stimulation with different durations in intact birds and birds with Area X lesions**

	Bird ID	Syllable ID number	Stimulus duration		
			10 ms	20 ms	50 ms
Intact bird	blk63mgt24	1	–	–	D
	blk93red55	1	U	–	N
	ppl56red81	1	U	–	N
	red56ppl9	1	–	–	D
	ylw62mgt75	1	–	–	U
	org26org79	1	N	N	N
	org92mgt87	1	D	D	D
		2	N	N	N
		3	U	U	U
	ylw20org26	1	–	–	N
		2	U	U	U
	ylw86wht87	1	U	U	U
		2	U	–	N
		3	U	U	U
Lesion bird	ylw25red65	1	N	N	N
		2	–	–	U
	ylw28mgt76	1	U	U	U
		2	U	U	U
	blk92wht33	1	U	U	U
	mgt88wht94	1	U	U	U
		2	U	–	U

D, Downward deflection; U, upward deflection; N, no FF trajectory was measured because LMAN stimulation disrupted the harmonic structure.

experiment we used the minimal current level that was sufficient to elicit effects on syllable structure (see Materials and Methods). These latencies were comparable with those reported in previous studies of the effects of LMAN stimulation [20–42 ms (Giret et al., 2014) and an average of 50 ms (Kao et al., 2005)] and are consistent with the possibility that the effects of LMAN burst stimulation reflect direct influences of activated LMAN neurons on downstream neurons in RA. However, in intact birds, LMAN provides recurrent input to Area X, which could act to introduce variability to song via indirect pathways from Area X to the song motor pathway (Gale et al., 2008; Hamaguchi and Mooney, 2012). To rule out this possible interpretation, we therefore performed an additional set of LMAN stimulation experiments in which we first lesioned Area X bilaterally.

The results of LMAN stimulation experiments in Area X lesion birds were substantially the same as those in intact birds. For six of seven syllables in four lesion birds, consistent deflections of FF trajectories were observed in response to LMAN stimulation with at least one of the three stimulus durations (10, 20, and 50 ms; Fig. 6F–J, Table 1). Upward deflections were observed in all 6 syllables, and no downward deflections were observed. The peak frequency of the stimulation-evoked deflections of FF trajectories was ~20–30 Hz for all the stimulus durations (10, 20, and 50 ms; Fig. 6J, magenta, cyan, and green), similar to that of Area X-dependent within-syllable variability in FF (Fig. 6J, gray). Also, peak frequencies of stimulation-dependent power spectra in individual syllables were not significantly different between intact and lesion birds (Fig. 6L;  $p = 0.11, 0.24, \text{ and } 0.81$  for 10, 20, and 50 ms, respectively). Moreover, there were no significant differences in latencies of the evoked deflections between intact and lesion birds (Fig. 6M;  $p = 0.53, 0.69, \text{ and } 0.21$ , respectively). Thus, even in the absence of Area X, brief stimulation of LMAN neurons can evoke rapid FF fluctuations and thus provide further support for the link between LMAN burst firing and within-syllable variability in FF trajectories.



**Figure 7.** Within-syllable FF variability decreases with age. **A**, Changes in the magnitude of within-syllable FF variability over 2 weeks in juvenile (left) and adult (right) birds. Each line corresponds to one syllable.  $*p < 0.002$ . **B**, Percent changes (mean  $\pm$  SEM) in the power spectrum of FF trajectories in juvenile (blue) and adult (black) birds over the 2 week period shown in **A**. Note the greater reduction of power in juvenile birds compared with that in adult birds.

### Age-dependent reduction in within-syllable variability

Previous studies have shown that both cross-rendition variation in FF and song plasticity decline with age, consistent with the notion that variability in syllable structure contributes to motor exploration and plasticity (Lombardino and Nottebohm, 2000; Brainard and Doupe, 2001; Kao and Brainard, 2006; Aronov et al., 2008). To examine whether within-syllable variability in song also declines with age, we recorded songs of juvenile zebra finches over time (Fig. 7). We found that within-syllable variability in FF decreases significantly with age. There was a large reduction in within-syllable variability between  $\sim 75$  dph and 2 weeks later (Fig. 7,  $n = 12$  syllables in two intact birds;  $p = 1.3 \times 10^{-4}$ , Wilcoxon signed-rank test). Similarly, there was a significant, but more modest, decrease in within-syllable variability between  $\sim 110$  dph and 2 weeks later (Fig. 7,  $n = 18$  syllables in seven intact birds;  $p = 0.002$ ). The observed age-dependent decline of within-syllable variability parallels an age-dependent reduction in song plasticity (Lombardino and Nottebohm, 2000; Brainard and Doupe, 2001), consistent with the notion that BG-driven within-syllable variation may be a critical substrate for song learning and plasticity.

### Discussion

The BG receive signals relating to reward evaluation and action selection, but whether they are also a source of behavioral variation that enables trial-and-error learning has remained unclear. Previous studies in songbirds that examined the effects of lesions of the BG nucleus Area X, which is specialized for singing, found no lasting changes to several measures of song variability. These studies all focused on variations in song structure across successive renditions of syllables (a timescale of hundreds of milliseconds; Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Goldberg and Fee, 2011; Ali et al., 2013), raising the hypothesis that exploratory vocal variability does not depend on Area X but instead likely emerges in the downstream thalamocortical (DLM–LMAN) circuit independently of Area X (Fee and Goldberg, 2011; Goldberg and Fee, 2011). In contrast, here we show that lesions of Area X dramatically decrease fluctuations of FF trajectories within individual syllables (a timescale of tens of milliseconds). Importantly, this form of rapid, within-syllable variation has been linked to adult song learning; the specific, idiosyncratic fluctuations on this timescale that are reinforced during learning become incorporated in lasting changes to song (Charlesworth et al., 2011). Moreover, FF differences across song syllables are principal characteristics that give each

learned song its unique identity. Hence, our results implicate the BG in generating a form of vocal variability that is crucial for learning and optimizing a key feature of song.

Our results also can explain discrepancies in previous studies between neural variability in the BG–thalamocortical circuit and song variability. In both Area X and downstream nucleus LMAN, neurons exhibit highly variable firing patterns during Undir singing (Hessler and Doupe, 1999a,b; Ólveczky et al., 2005; Kao et al., 2005, 2008; Goldberg et al., 2010; Woolley et al., 2014). It has been unclear whether this variable neural activity contributes to song variability because lesions of Area X that eliminate variable burst firing in LMAN do not cause lasting changes to any forms of song variation that were examined previously (Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Goldberg and Fee, 2011; Ali et al., 2013). Our results reveal that it is the rapid within-syllable variability in FF that is permanently eliminated after Area X lesions.

Our results support a model in which variability in Area X pallidal neurons drives downstream burst firing in LMAN that in turn drives rapid fluctuations in FF. This model is supported by the following experimental findings in addition to our current results: (1) variable firing in Area X pallidal neurons does not depend on recurrent inputs from LMAN and appears to emerge in Area X (Woolley et al., 2014; Budzillo et al., 2017); (2) lesions of Area X abolish variable burst firing in LMAN (Kojima et al., 2013), indicating that neural variability is propagated from Area X to LMAN; (3) viral manipulations that alter Area X medium spiny neuron activity can cause changes in neural variability measured in LMAN and in song (Tanaka et al., 2016; Heston et al., 2018); (4) LMAN exhibits highly variable burst firing only during Undir song, when within-rendition behavioral variability is high, and such bursting is largely absent during Dir song, when within-rendition variability is low (Hessler and Doupe, 1999a; Kao et al., 2005, 2008); (5) mimicking LMAN bursts by electrical stimulation in LMAN, even in the absence of Area X, can drive rapid deflections of FF trajectories with a similar timescale to within-syllable variation in FF normally present in Undir song (Fig. 6); and (6) both lesions of LMAN and lesions of Area X reduce within-syllable song variation and its social modulation (Figs. 3–5).

This model provides some insight into results from recent studies that manipulated, rather than eliminated, Area X activity. Heston et al. (2018) found that transient manipulations of activity in a subpopulation of Area X MSNs could alter variability in FF both across renditions and within syllables; based on their

observations, they hypothesized that Area X may act as a variability “repressor” that functions reciprocally to LMAN, with increased Area X activity attenuating variability present in LMAN and decreased activity enabling the expression of variability that arises independently of Area X. However, these studies did not determine how the manipulations of MSNs altered activity of the pallidal projection neurons that contribute to downstream activity in LMAN. Our results provide a direct demonstration that elimination of pallidal projection neuron activity (via lesions of Area X) permanently decreases, rather than increases, within-syllable variability and thereby indicate a necessary role of Area X in the generation of this form of exploratory variability in song.

While our results indicate that within-syllable fluctuations in FF arise from Area X, they do not identify the source of slower cross- rendition variation in FF. Such cross- rendition variation is transiently decreased by Area X lesions, but recovers over days (Figs. 3*F*, 4*E*; Ali et al., 2013). In contrast, cross- rendition variation is permanently eliminated by LMAN lesions (Kao et al., 2005; Kao and Brainard, 2006). Therefore, cross- rendition variation might normally arise in LMAN independently of Area X (Fee and Goldberg, 2011; Goldberg and Fee, 2011). However, the transient reduction in cross- rendition variability by Area X lesions suggests that, in the intact brain, Area X normally contributes to this form of variability. In this case, the recovery of cross- rendition variability might reflect compensatory changes in LMAN or other circuitry in response to the removal of normal activity originating from Area X. The possibility that Area X contributes to cross- rendition variability is also supported by previous studies showing that manipulations of dopamine- dependent signaling in Area X can alter social modulation of cross- rendition variability (Leblois et al., 2010; Murugan et al., 2013). Regardless of the source, the slower cross- rendition variation that persists after Area X lesions is not sufficient to enable adaptive changes in FF in response to contingent reinforcement (Ali et al., 2013). Whether such cross- rendition variation in intact birds contributes to learning (independent of the contributions of within- rendition variation) remains to be determined.

The rapid, BG- dependent fluctuations in FF that we have described are well suited to serve a function in learning. First, when birds are trained via aversive reinforcement that is contingent on FF at a precise time point within a targeted syllable, they learn to produce the average of the FF trajectories that escape from the aversive reinforcement (i.e., those that are associated with successful outcomes; Charlesworth et al., 2011). This indicates that the nervous system can keep track of variation on this rapid time- scale and its association with better versus worse outcomes, which are critical prerequisites for behavioral variation that can serve as motor exploration for learning. Since this learning to modify FF trajectories can lead to global and complex changes in syllable structure when aversive reinforcement is contingent on FF at multiple time points in FF trajectories (Charlesworth et al., 2011), rapid fluctuations in acoustic trajectories may also be involved in naturally occurring song learning and maintenance. Second, this component of variation is strongly modulated by social context: within- syllable fluctuations in FF are large when birds sing alone (Undir song) relative to when birds sing courtship song to females (Dir song; Fig. 4*A,B*). This supports the hypothesis that Undir song reflects a state of vocal practice in which birds actively generate variation in performance (motor exploration) that enables fine- tuning of premotor control to produce and maintain a good match to a song template, whereas Dir song is a performance state in which the current best version of learned song is produced with minimal variation (motor exploi-

tion; Jarvis et al., 1998; Kao et al., 2005; Kojima and Doupe, 2011). Third, there is an age- dependent decline in this form of vocal variability (Fig. 7) that is appropriate if greater variation is required to refine initially coarse vocalizations during developmental song learning whereas less variation is needed to maintain already well learned song.

In summary, we have demonstrated a role for BG in generating variation in the kinematics of individual actions in birdsong, rapid fluctuations in FF within renditions of individual syllables. Although studies of motor skill learning and performance in mammals have directed attention particularly to the role of BG in selecting between distinct alternative actions, recent findings implicate the BG in modulating the kinematics of individual actions (Yttri and Dudman, 2016). Our finding that avian BG actively generate and regulate a form of rapid behavioral variation that is important for optimization of individual actions also indicates that this may be a general function of BG circuitry.

## References

- Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP (2013) The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80:494–506. [CrossRef Medline](#)
- Andalman AS, Fee MS (2009) A basal ganglia- forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U S A* 106:12518–12523. [CrossRef Medline](#)
- Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* 320:630–634. [CrossRef Medline](#)
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM (2005) Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437:1158–1161. [CrossRef Medline](#)
- Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224:901–903. [CrossRef Medline](#)
- Bottjer SW, Halsema KA, Brown SA, Miesner EA (1989) Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *J Comp Neurol* 279:312–326. [CrossRef Medline](#)
- Bottjer SW, Roselinsky H, Tran NB (1997) Sex differences in neuropeptide B staining of song- control nuclei in zebra finch brains. *Brain Behav Evol* 50:284–303. [CrossRef Medline](#)
- Brainard MS, Doupe AJ (2001) Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J Neurosci* 21:2501–2517. [CrossRef Medline](#)
- Budzillo A, Duffy A, Miller KE, Fairhall AL, Perkel DJ (2017) Dopaminergic modulation of basal ganglia output through coupled excitation–inhibition. *Proc Natl Acad Sci U S A* 114:5713–5718. [CrossRef Medline](#)
- Carrillo GD, Doupe AJ (2004) Is the songbird area X striatal, pallidal, or both? An anatomical study. *J Comp Neurol* 473:415–437. [CrossRef Medline](#)
- Charlesworth JD, Tumer EC, Warren TL, Brainard MS (2011) Learning the microstructure of successful behavior. *Nat Neurosci* 14:373–380. [CrossRef Medline](#)
- Coleman MJ, Mooney R (2004) Synaptic transformations underlying highly selective auditory representations of learned birdsong. *J Neurosci* 24:7251–7265. [CrossRef Medline](#)
- Doupe AJ, Solis MM (1997) Song- and order- selective neurons develop in the songbird anterior forebrain during vocal learning. *J Neurobiol* 33:694–709. [CrossRef Medline](#)
- Fee MS, Goldberg JH (2011) A hypothesis for basal ganglia- dependent reinforcement learning in the songbird. *Neuroscience* 198:152–170. [CrossRef Medline](#)
- Gale SD, Perkel DJ (2010) Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. *J Chem Neuroanat* 39:124–131. [CrossRef Medline](#)
- Gale SD, Person AL, Perkel DJ (2008) A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. *J Comp Neurol* 508:824–839. [CrossRef Medline](#)
- Giret N, Kornfeld J, Ganguli S, Hahnloser RHR (2014) Evidence for a causal inverse model in an avian cortico- basal ganglia circuit. *Proc Natl Acad Sci U S A* 111:6063–6068. [CrossRef Medline](#)

- Goldberg JH, Fee MS (2011) Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. *J Neurophysiol* 105:2729–2739. [CrossRef Medline](#)
- Goldberg JH, Adler A, Bergman H, Fee MS (2010) Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. *J Neurosci* 30:7088–7098. [CrossRef Medline](#)
- Green D, Swets J (2012) Signal detection theory and psychophysics. Tallahassee, FL: Peninsula Publishing.
- Hamaguchi K, Mooney R (2012) Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci* 32:11671–11687. [CrossRef Medline](#)
- Hessler NA, Doupe AJ (1999a) Social context modulates singing-related neural activity in the songbird forebrain. *Nat Neurosci* 2:209–211. [CrossRef Medline](#)
- Hessler NA, Doupe AJ (1999b) Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. *J Neurosci* 19:10461–10481. [CrossRef Medline](#)
- Heston JB, Simon J 4th, Day NF, Coleman MJ, White SA (2018) Bidirectional scaling of vocal variability by an avian cortico-basal ganglia circuit. *Physiol Rep* 6:e13638. [CrossRef Medline](#)
- Humphries MD, Khamassi M, Gurney K (2012) Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Front Neurosci* 6:9. [Medline](#)
- Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F (1998) For whom the bird sings: context-dependent gene expression. *Neuron* 21:775–788. [CrossRef Medline](#)
- Kalva SK, Rengaswamy M, Chakravarthy VS, Gupte N (2012) On the neural substrates for exploratory dynamics in basal ganglia: a model. *Neural Netw* 32:65–73. [CrossRef Medline](#)
- Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96:1441–1455. [CrossRef Medline](#)
- Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433:638–643. [CrossRef Medline](#)
- Kao MH, Wright BD, Doupe AJ (2008) Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. *J Neurosci* 28:13232–13247. [CrossRef Medline](#)
- Kojima S, Doupe AJ (2011) Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci U S A* 108:1687–1692. [CrossRef Medline](#)
- Kojima S, Kao MH, Doupe AJ (2013) Task-related “cortical” bursting depends critically on basal ganglia input and is linked to vocal plasticity. *Proc Natl Acad Sci U S A* 110:4756–4761. [CrossRef Medline](#)
- Kubikova L, Bosikova E, Cvikova M, Lukacova K, Scharff C, Jarvis ED (2014) Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. *Sci Rep* 4:6590. [Medline](#)
- Leblois A, Wendel BJ, Perkel DJ (2010) Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *J Neurosci* 30:5730–5743. [CrossRef Medline](#)
- Lombardino AJ, Nottebohm F (2000) Age at deafening affects the stability of learned song in adult male zebra finches. *J Neurosci* 20:5054–5064. [CrossRef Medline](#)
- Marler P (1970) Birdsong and speech development: could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man. *Am Sci* 58:669–673. [Medline](#)
- Mooney R, Konishi M (1991) Two distinct inputs to an avian song nucleus activate different glutamate receptor subtypes on individual neurons. *Proc Natl Acad Sci U S A* 88:4075–4079. [CrossRef Medline](#)
- Murugan M, Harward S, Scharff C, Mooney R (2013) Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron* 80:1464–1476. [CrossRef Medline](#)
- Okuhata S, Saito N (1987) Synaptic connections of thalamo-cerebral vocal nuclei of the canary. *Brain Res Bull* 18:35–44. [CrossRef Medline](#)
- Ölveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3:e153. [CrossRef Medline](#)
- Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS (2011) Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol* 106:386–397. [CrossRef Medline](#)
- Sakata JT, Brainard MS (2008) Online contributions of auditory feedback to neural activity in avian song control circuitry. *J Neurosci* 28:11378–11390. [CrossRef Medline](#)
- Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11:2896–2913. [CrossRef Medline](#)
- Sheth SA, Abuelem T, Gale JT, Eskandar EN (2011) Basal ganglia neurons dynamically facilitate exploration during associative learning. *J Neurosci* 31:4878–4885. [CrossRef Medline](#)
- Sohrabji F, Nordeen EJ, Nordeen KW (1990) Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav Neural Biol* 53:51–63. [CrossRef Medline](#)
- Sridharan D, Prashanth PS, Chakravarthy VS (2006) The role of the basal ganglia in exploration in a neural model based on reinforcement learning. *Int J Neural Syst* 16:111–124. [CrossRef Medline](#)
- Stark LL, Perkel DJ (1999) Two-stage, input-specific synaptic maturation in a nucleus essential for vocal production in the zebra finch. *J Neurosci* 19:9107–9116. [CrossRef Medline](#)
- Stepanek L, Doupe AJ (2010) Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J Neurophysiol* 104:2474–2486. [CrossRef Medline](#)
- Stocco A (2012) Acetylcholine-based entropy in response selection: a model of how striatal interneurons modulate exploration, exploitation, and response variability in decision-making. *Front Neurosci* 6:18. [Medline](#)
- Sutton RS, Barto AG (1998) Reinforcement learning. Cambridge, MA: MIT.
- Tanaka M, Singh Alvarado J, Murugan M, Mooney R (2016) Focal expression of mutant huntingtin in the songbird basal ganglia disrupts cortico-basal ganglia networks and vocal sequences. *Proc Natl Acad Sci U S A* 113:E1720–E1727. [CrossRef Medline](#)
- Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. *Anim Behav* 59:1167–1176. [CrossRef Medline](#)
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291:2564–2569. [CrossRef Medline](#)
- Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature* 450:1240–1244. [CrossRef Medline](#)
- Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14:6924–6934. [CrossRef Medline](#)
- Woolley SC, Rajan R, Joshua M, Doupe AJ (2014) Emergence of context-dependent variability across a basal ganglia network. *Neuron* 82:208–223. [CrossRef Medline](#)
- Yttri EA, Dudman JT (2016) Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* 533:402–406. [CrossRef Medline](#)