

# Behavior-Linked FoxP2 Regulation Enables Zebra Finch Vocal Learning

Jonathan B. Heston<sup>1,2</sup> and Stephanie A. White<sup>1,2</sup><sup>1</sup>Interdepartmental Program in Neuroscience, and <sup>2</sup>Department of Integrative Biology and Physiology, University of California, Los Angeles, Los Angeles, California 90095

Mutations in the FOXP2 transcription factor cause an inherited speech and language disorder, but how FoxP2 contributes to learning of these vocal communication signals remains unclear. FoxP2 is enriched in corticostriatal circuits of both human and songbird brains. Experimental knockdown of this enrichment in song control neurons of the zebra finch basal ganglia impairs tutor song imitation, indicating that adequate FoxP2 levels are necessary for normal vocal learning. In unmanipulated birds, vocal practice acutely downregulates FoxP2, leading to increased vocal variability and dynamic regulation of FoxP2 target genes. To determine whether this behavioral regulation is important for song learning, here, we used viral-driven overexpression of FoxP2 to counteract its downregulation. This manipulation disrupted the acute effects of song practice on vocal variability and caused inaccurate song imitation. Together, these findings indicate that dynamic behavior-linked regulation of FoxP2, rather than absolute levels, is critical for vocal learning.

**Key words:** basal ganglia; birdsong; forkhead; language; procedural learning; speech

## Introduction

The transcription factor FOXP2 has an unprecedented role in the formation and function of brain circuits underlying language. Individuals heterozygous for a mutant FOXP2 allele exhibit a specific language impairment characterized by deficits in the coordination and sequencing of orofacial movements required for speech (Vargha-Khadem et al., 1998; Lai et al., 2001). FOXP2 is robustly expressed in the striatum; both structural and functional imaging of individuals who harbor the mutant allele implicate this brain region, among others, in mediating the language deficits. Consistent with this notion, mice carrying this mutant allele exhibit impaired striatal synaptic plasticity and altered ultrasonic vocalizations (Groszer et al., 2008), superficially consistent with the human speech phenotype. Because these vocalizations are innate (Day and Fraley, 2013), however, their relevance to the learned vocal component of language is tenuous.

Numerous parallels between speech and birdsong make songbirds advantageous models for investigating FoxP2's role in learned vocalization (Doupe and Kuhl, 1999). Similar to human imitative learning, songbirds learn to vocalize by mimicking conspecifics. In zebra finches, song learning is sexually dimorphic such that only young males learn the songs of an adult male tutor.

This behavior relies on a set of brain nuclei collectively known as the song control circuit. In each part of the circuit, song-dedicated neurons are clustered together and identifiable; an apparently unique feature of the brains of avian vocal learners, which greatly facilitates targeted experimental interventions.

Neural expression patterns of *FoxP2* are conserved between humans and songbirds (Teramitsu et al., 2004), including robust expression within the basal ganglia. In zebra finches, FoxP2 is enriched in Area X, the song-dedicated basal ganglia nucleus necessary for vocal learning (Haesler et al., 2004; Teramitsu et al., 2004). Knockdown of FoxP2 in Area X of juvenile males leads to inaccurate song learning suggesting a postdevelopmental role for FoxP2 (Haesler et al., 2007).

Adding complexity to this role, 2 h of morning song practice results in acute decreases in FoxP2 mRNA and protein within Area X (Teramitsu and White, 2006; Miller et al., 2008; Teramitsu et al., 2010; Thompson et al., 2013). This downregulation is accompanied by acute increases in song variability (Miller et al., 2010) and the bidirectional regulation of thousands of genes, including transcriptional targets of human FOXP2 (Hilliard et al., 2012). Together, these data suggest that the dynamic regulation of FoxP2 is critical for vocal learning.

To directly test this idea, we constitutively elevated FoxP2 at the onset of the sensorimotor period for song learning by stereotaxic injection of a virus designed to express full-length FoxP2 into Area X of young male zebra finches (Fig. 1). We envisioned two potential outcomes of such a manipulation. First, if FoxP2 plays a permissive role in which adequate levels are required for song learning, then its overexpression should result in a phenotype distinct from that observed following knockdown. Alternatively, if FoxP2 plays a dynamic role, in which upregulation and downregulation are required, then its overexpression should result in a phenotype convergent with that of the knockdown.

Received Sept. 3, 2014; revised Dec. 2, 2014; accepted Dec. 27, 2014.

Author contributions: J.B.H. and S.A.W. designed research; J.B.H. performed research; J.B.H. analyzed data; J.B.H. and S.A.W. wrote the paper.

This work was supported by Grants T32MH019384, T32NS058280, UCLA Edith Hyde Fellowship (J.B.H.), R01MH070712, and Tennenbaum Center for the Biology of Creativity (S.A.W.). We thank R.L. Neve (Massachusetts Institute of Technology) for expertise on AAV, J.E. Miller (University of Arizona) for insightful comments on the manuscript, O.G. Casillas for technical assistance, and E.L. Heston for design and figure assembly.

The authors declare no competing financial interests.

Correspondence should be addressed to Stephanie A. White, University of California, Los Angeles, Box 957239, 610 Charles E. Young Drive East, Los Angeles, California 90095. E-mail: sawwhite@ucla.edu.

DOI:10.1523/JNEUROSCI.3715-14.2015

Copyright © 2015 the authors 0270-6474/15/352885-10\$15.00/0

## Materials and Methods

### Subjects

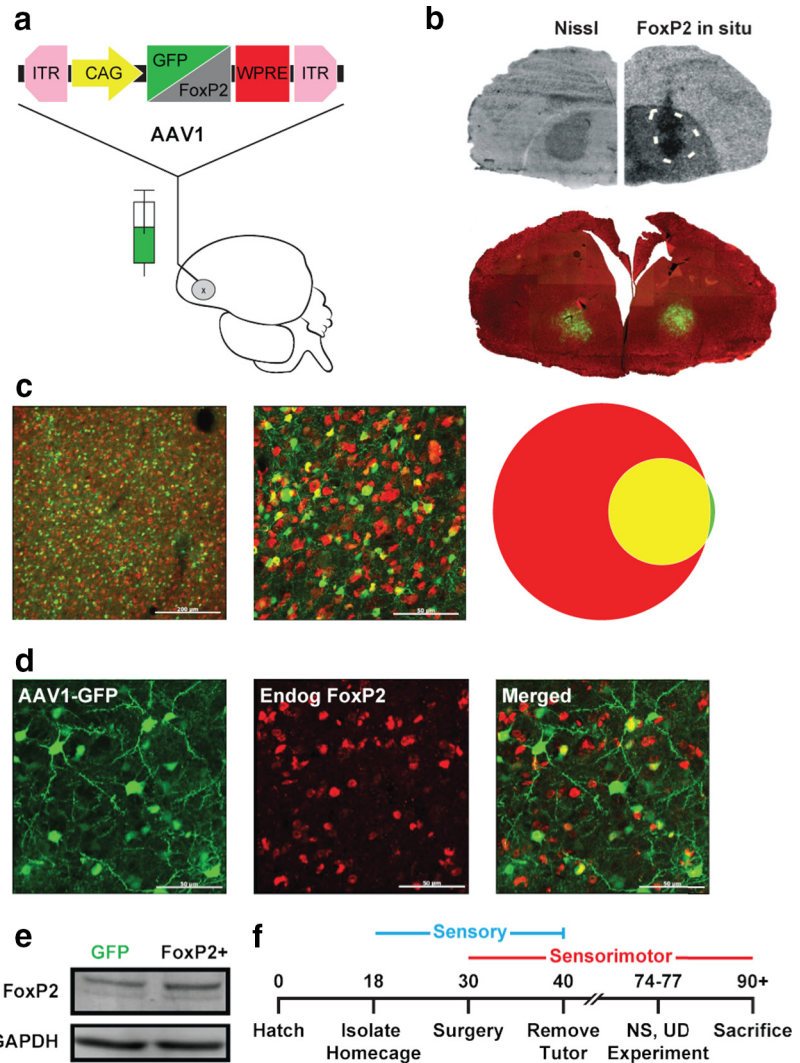
All animal use was in accordance with NIH guidelines for experiments involving vertebrate animals and approved by the University of California Los Angeles Chancellor's Institutional Animal Care and Use Committee and were consistent with the American Veterinary Medical Association guidelines. Birds were obtained from our own breeding colony, and housed in climate-controlled rooms inside cages and aviaries with a 13:11 light/dark cycle including half hours of dawn and dusk lighting conditions. Birds had unlimited access to food, grit, and water and were provided both nutritional supplements (e.g., cuttlebone, spray millet, chopped hard-boiled eggs, orange and green vegetables, Calci-boost) and environmental enrichments (e.g., a variety of perches, swings, mirrors, and water baths).

### Behavior

The experimental paradigm is schematized in Figure 1*f*. At 18 d young birds were moved to sound attenuation chambers (Acoustic Systems) along with both parents and any clutch-mate siblings. At 30 d, male "pupils" were stereotaxically injected with virus as described below, and then returned to their families. At 40 d, each pupil was separated from his family and placed within a sound attenuation chamber along with an adult, unrelated female (90–120 d) to enable social interactions. At 70 d, the female was removed from the chamber in preparation for an acute test of the behavioral effects of singing state on song variability, as previously described (Miller et al., 2010; Chen et al., 2013). Briefly, on the morning of 1 d between 74 and 77 d, a given male was monitored, and if he attempted to sing was distracted by the experimenter (if the bird none-the-less sang >20 complete motifs, the trial was terminated). After 2 h, he was allowed to sing and these subsequent songs were recorded. On another day within the same time window, the bird was allowed to sing undirected song for 2 h and the vocalizations during the subsequent 30 min were recorded. If on either day the bird failed to sing during the 30 min following the prior 2 h epoch, the experiment was repeated. At 77 d, the female was returned to the cage, and daily recording of song recommenced. By 100–153 d, all pupils were overdosed via isoflurane inhalation, and their brains extracted and prepared for histological analysis.

### Song recording and analysis

Vocalizations were recorded continuously from 40 to 90 d. Sounds were recorded using either a Countryman EMW omnidirectional lavalier microphone (Countryman Associates) or a Shure SM58 microphone and digitized using a PreSonus Firepod (44.1 kHz sampling rate, 24 bit depth). Recordings were acquired and song features quantified using Sound Analysis Pro (SAP) 2011 software (Tchernichovski et al., 2000). Although the investigator knew the group allocation during the experiment, this automated software was used to derive all measures of song learning and acoustic features, avoiding subjective assessment. Songs were manually hand-segmented into motifs and individual syllables by the experimenter, and then analyzed in a semiautomated manner using SAP.



**Figure 1.** Exogenous overexpression of speech-related FoxP2 during zebra finch vocal learning. *a*, Schematic depicts control (GFP) and FoxP2-expressing viral constructs delivered by stereotaxic injection into song-dedicated Area X. *b*, Top, Coronal hemisections illustrate targeting and expression in Area X, visible in the Nissl stain (left), and indicated by the dashed line (right). *In situ* hybridization signals for zebra finch FoxP2 reveal elevated mRNA at the injection site of the FoxP2-expressing virus. Bottom, Dense, restricted GFP expression at bilateral injection sites of the control virus. *c*, Mid- (left) and high- (middle) power images of the brain shown in *b* reveals overlap between viral-driven GFP-expression (green) and NeuN immunostain (red). Venn diagram (right) illustrates the quantitative overlap (yellow) between GFP and NeuN. *d*, Comparison of viral-driven GFP expression (AAV1-GFP) and immunostain signals for endogenous FoxP2 (Endog FoxP2) indicate high levels of overlap (Merged). *e*, Representative immunoblot of FoxP2 signals arising from Area X micropunches in an adult bird that was injected with AAV1-GFP in one hemisphere and AAV1-FoxP2 contralaterally. An overall 40.3% increase in FoxP2 signal was observed in hemispheres receiving the AAV-FoxP2 virus relative to the contralateral side ( $p = 0.0436$ ,  $n = 7$  pairs, paired one-tailed bootstrap). *f*, Timeline of behavioral experiments.

Motifs were identified as repeated units of song composed of multiple syllables, and excluded introductory notes. Canonical and noncanonical renditions of motifs were included in the analysis to capture the full range of singing behavior. A syllable was identified as a sound element that is separated from other syllables by silence or by local minima in the amplitude (Immelmann, 1969). Motifs, as well as the phonology and syntax of syllables, were assessed as detailed in the next section.

**Motif analysis.** We quantified how well pupils imitated their tutor's motif using similarity scores obtained in SAP from 200 asymmetric pairwise comparisons of 20 renditions of the pupil's typical motif with 10 renditions of the tutor motif. The exact same set of 10 tutor motif renditions was used for all pupils of the same tutor. Asymmetric comparisons analyze the spectrotemporal similarity of sound elements without respect to their position within a motif. This operation is well suited to the analysis of motifs because it measures large timescale resolution of acoustic similarity and makes no a priori assumptions about syllable order. We

report the upper-third quartile score from these comparisons so as not underestimate the percentage of tutor song copied. Automated analysis was supplemented by manual counting of imitated, omitted, and improvised syllables.

**Syllable level analysis of tutor copying and self-similarity.** We quantified both similarity-to-tutor and similarity-to-self using symmetric comparisons. Symmetric comparisons analyze the spectrotemporal similarity from the beginning to the end of the two sounds under investigation. This operation is well suited for the analysis of syllables where the sound elements have already been isolated and can be assumed to begin and end at corresponding time points.

For similarity-to-tutor, 20 renditions of a tutor syllable were compared with 30 renditions of that corresponding pupil syllable, generating 600 unique comparisons. Each copying metric was represented by the median of these 600 comparisons. This analysis also yielded the difference measurements, which represent the mean Euclidian distance on a feature-specific basis. Features included pitch, frequency modulation (FM), Weiner entropy, pitch goodness (PG), and amplitude modulation (AM). We also measured durational error which is operationally defined here as 100 minus sequential match. For syllable similarity-to-self measurements, the same set of 30 syllables used for the tutor comparison was compared to itself. Again, each score was represented by the median of these comparisons.

**Syllable level analysis of mean features and CV.** Each syllable is characterized by measures of acoustic features including the five listed above, as well as duration, amplitude, and mean frequency. We obtained mean and coefficient-of-variation (CV) values based on measurements of 25 renditions of a given syllable.

**Syntax analysis.** Analysis of syntactical similarity to tutor and syntax entropy was performed by an investigator aware of experimental allocation, using a string-based method as described by Miller et al., 2010. Briefly, this analysis generates strings of 300 syllables, which are annotated sequentially without respect to motif or bout terminations. The analysis has the benefit of not requiring manual selection of motifs and avoids skewing of entropy scores by the occurrence of rare or infrequent syllables. For each pupil and tutor, we manually annotated strings of 300–350 user-defined syllables, which did not include introductory notes. The range was selected to account for improvised syllables among pupils so that at least 300 tutor-copied syllables would be included in the final analysis. Based on these data, we computed a transition probability matrix. Transition probability matrices of tutors and pupils were correlated in both a punished and unpunished manner with the latter score excluding syllables that were omitted by the pupil. Because we already analyzed omissions and improvisations we report only the unpunished scores, but both analyses gave qualitatively similar results. Values for syllable syntax entropy reported are weighted entropy scores, which are adjusted for the frequency of occurrence of each syllable type when determining its contribution to overall syntactical entropy. An entropy score of 0 reflects a fixed syllable order, whereas a score of 1 indicates random syllable order (Miller et al., 2010).

**Analysis of song development.** To determine the developmental trajectory of vocal imitation we analyzed songs recorded on 50, 70, and 90 d ( $\pm 2$  d; in several cases recordings were unavailable from either 50 or 70 d due to technical issues). Twenty motifs were compared asymmetrically to a single tutor motif (this motif was chosen to be representative of the tutor's vocal repertoire and the same motif was used for comparison with each of that tutor's pupils).

**Stereotaxic neurosurgery.** At 30 d, males were anesthetized with 2% isoflurane and placed in a custom-built avian stereotax (Herb Adams Engineering). The head was held at a 45° angle relative to the vertical axis, a semicircular incision was made in the scalp to preserve vasculature which was then retracted and a small craniotomy made over the injection site (+5.15 mm anterior, +1.5–6 mm lateral to the bifurcation of the midsagittal sinus and at a depth of 3.3 mm). Virus was loaded into a glass microelectrode that had been previously broken ~8 mm from the bore to create an inner diameter of 30–50  $\mu\text{m}$ , backfilled with mineral oil, and attached to a pressure injection unit (Drummond Nanoject II, Drummond Scientific). The electrode was lowered into the brain and each hemisphere received three 27.6 nl injections over a 30 s period followed

by a 10 min wait period before the glass electrode was retracted. After completion of the injection, the scalp was replaced and the incision closed with Vetbond (Santa Cruz Animal Health).

Surgery on adults followed a nearly identical procedure with the volume of injections varying as described for each experiment. In surgeries involving both control (GFP-expressing) and experimental (FoxP2-expressing) viruses, each injected into one hemisphere, the first electrode was discarded after use on the first hemisphere and a new one was loaded for the second.

**Adeno-associated virus information.** After extensive tests, the virus that met our criteria was a custom designed AAV (serotype 1) that was cloned and produced by Virovek. Both FoxP2- and GFP-expressing viral constructs use the CMV early enhancer/chicken  $\beta$  actin (CAG) promoter to drive expression. This element, provided by Virovek, was followed by either the coding sequences for zebra finch FoxP2 or GFP (provided by Virovek) then a WPRE element. Both FoxP2- and GFP-expressing viruses had a titer of  $2.24\text{E}+13$  vg/ml justifying equal volumes of delivery.

**Histological methods.** To examine the efficacy in targeting and expression of viral injections, birds that received the GFP control virus were perfused with warm saline followed by ice-cold 4% paraformaldehyde in 0.1 M phosphate buffer, and their brains extracted for histological analysis. Characterization of viral transfection was performed using immunohistological methods described by Miller et al. (2008). To specifically assess FoxP2 mRNA and protein levels following AAV-driven FoxP2 expression, brains from FoxP2+ animals were flash frozen on liquid nitrogen. Brains were sectioned on a cryostat (Leica Microsystems) at a thickness of 30  $\mu\text{m}$  for perfused brains and 20  $\mu\text{m}$  for fresh frozen brains. Verification of targeting and overexpression of zebra finch FoxP2 mRNA following injection of the FoxP2-expressing virus (Fig. 1b) was done using *in situ* hybridization analysis as described in Teramitsu and White, 2006.

To validate overexpression of FoxP2 protein, adult male zebra finches were injected with FoxP2-expressing virus in one hemisphere and GFP-expressing virus in the contralateral hemisphere (502 nl per injection site). This approach allowed us to control for any difference in FoxP2 levels that are a result of dynamic behavioral regulation or interbird differences. Three to 4 weeks later, the birds were killed immediately after singing 2 h of undirected song. Area X tissue punches were obtained using methods previously described by Miller et al. (2008). Briefly, sections of 20  $\mu\text{m}$  thickness were cut before visualization of Area X, then bilateral tissue punches of Area X were obtained at a depth of 1 mm using a 20 gauge Luer adaptor (Beckton Dickinson) attached to a 1 ml syringe. Unilateral tissue punches were homogenized in 30  $\mu\text{l}$  of ice-cold modified RIPA lysis buffer with protease inhibitors using a hand-held homogenizer, mixed with an equal volume of 2 $\times$  Laemmli loading buffer (Bio-Rad) containing 0.1%  $\beta$ -mercaptoethanol and stored at  $-80^\circ\text{C}$  until use.

Samples were boiled for 3–5 min and loaded on a 10% acrylamide SDS-PAGE gel along with Prestained Precision Plus ladders (Bio-Rad, Pierce) as a molecular mass marker. Samples were then subjected to electrophoresis, electroblotted onto PVDF membranes (Millipore) for 4 h at 400 mA, and analyzed with rabbit antibody against FoxP2 (1:500, Millipore, ABE73; Miller et al., 2008), mouse antibody to GAPDH (1:30,000, Millipore, MAB374; Miller et al., 2008), and a rabbit antibody to DARPP-32 (1:5000, Abcam, ab1855; Murugan et al., 2013). Finally, blots were probed with horseradish peroxidase-conjugated anti-rabbit IgG (1:2000 dilution) and anti-mouse IgG (1:10,000 dilution; GE Healthcare Pharmacia Biotech). As previously reported, we detected the presence of two bands (~69, ~66 kDa; see Fig. 1e) with the lighter band of lower molecular mass potentially representing another isoform of FoxP2. We quantified only the expression of the higher molecular weight band as this represents the full-length isoform that is being overexpressed. Expression levels of FoxP2 in Figure 1e are presented as percentage change in the FoxP2+ hemisphere relative to the GFP hemisphere.

**Statistics.** Resampling statistics were used throughout our analysis, including either paired or unpaired resample tests. Comparisons between groups were done primarily using an unpaired resampling test (the one exception being the feature-specific errors, which used a paired test as described below). The unpaired test begins by calculating the difference in group means. This value represents the test statistic,  $M$ . We then

created pseudo datasets with same  $n$  as the actual group sizes and randomly drawn with replacement from a combined set of actual data points. This process was repeated 10,000 times, keeping track of pseudo- $M$  values. These values formed the distribution of  $M$  under the null hypothesis, reflecting the values of  $M$  we could have expected if the direction if the distribution of data points was random, and was not an effect of the experimental paradigm. Finally, the number of pseudo- $M$  values that were as large or larger than the actual  $M$  was determined and this number divided by 10,000. This value reflects the reported  $p$  value.

The paired test followed a similar procedure and is described in Miller et al. (2010). Briefly, this test begins by first calculating the group mean of the individual samples' conditional differences. This mean was our test statistic,  $M$ . Then, we randomly sampled  $n$  times from a vector containing 1 and  $-1$ , where  $n$  was the number of samples. The  $n$  element long vector of 1's and  $-1$ 's was multiplied by the vector containing the actual differences, effectively randomizing the direction of the conditional differences. Then, we took the mean of this randomized data and repeated the randomization process 10,000 times, keeping track of the mean each time. These means formed the distribution of  $M$  under the null hypothesis, reflecting the values of  $M$  we could have expected if the direction of the individual conditional differences was random, and was not an effect of the experimental paradigm.

All hypotheses related to tutor imitation were tested using a one-tailed test because young zebra finches characteristically imitate their tutors' songs. All deviations from the tutor model will result in lower imitation scores and larger error scores, thus measures were bounded on one side. A one-tailed test was also used to determine whether or not injection of the FoxP2+ AAV resulted in higher FoxP2 protein levels because there was no prediction for it to decrease FoxP2. All hypotheses related to vocal variability were tested using a two-tailed test because both increases and decreases in variability were possible consequences of viral overexpression and/or behavioral context.

Comparisons within a bird were performed using paired tests. Comparisons between groups of birds were done using unpaired tests. The one exception was the analysis of syllable error magnitude, which compared feature-specific errors on a paired syllable-by-syllable manner.

## Results

To date, no virus has been successfully used to drive overexpression of a gene that is normally present in the zebra finch song circuit. After unsuccessful tests of several virus types and serotypes, we assessed the ability of adeno-associated virus serotype 1 (AAV1) to drive overexpression of either zebra finch FoxP2 (FoxP2+) or GFP off of the CAG promoter (Fig. 1a). Both the FoxP2+ and GFP-expressing AAV1 viruses transduced a significant portion of Area X without spillover into adjacent regions of the brain as measured by both *in situ* hybridization signals for zebra finch FoxP2 mRNA, as well as GFP reporter signals (Fig. 1b) and FoxP2 protein expression (Fig. 1e). Importantly, no GFP fluorescence was detected in LMAN or HVC (data not shown), which both project to Area X, indicating that the virus does not retrogradely transfect its afferent inputs.

At the epicenter of viral transfection in Area X,  $24.0 \pm 5.5\%$  of the NeuN-positive cells also expressed GFP and, of the total number of transduced cells,  $96.7 \pm 1.7\%$  were NeuN-positive ( $n = 4$  birds; Fig. 1c), indicating a robust level of neuron-specific expression. Additionally, we found a high degree of overlap between GFP and endogenous FoxP2 (Fig. 1d), consistent with our goal of overexpressing FoxP2 within the subset of striatal neurons that normally express it. There was little overlap between GFP and Lant6, indicating that the virus tends not to transduce pallidum-like projection neurons, consistent with AAV1's low efficiency in mammalian pallidum (Burger et al., 2004). Thus, the AAV1-CAG construct meets the minimum requirements of robustly and specifically transducing FoxP2-expressing neurons in Area X without damaging this nucleus nor infecting other regions of the

brain which either do not express high levels of FoxP2 or are not part of the song circuit.

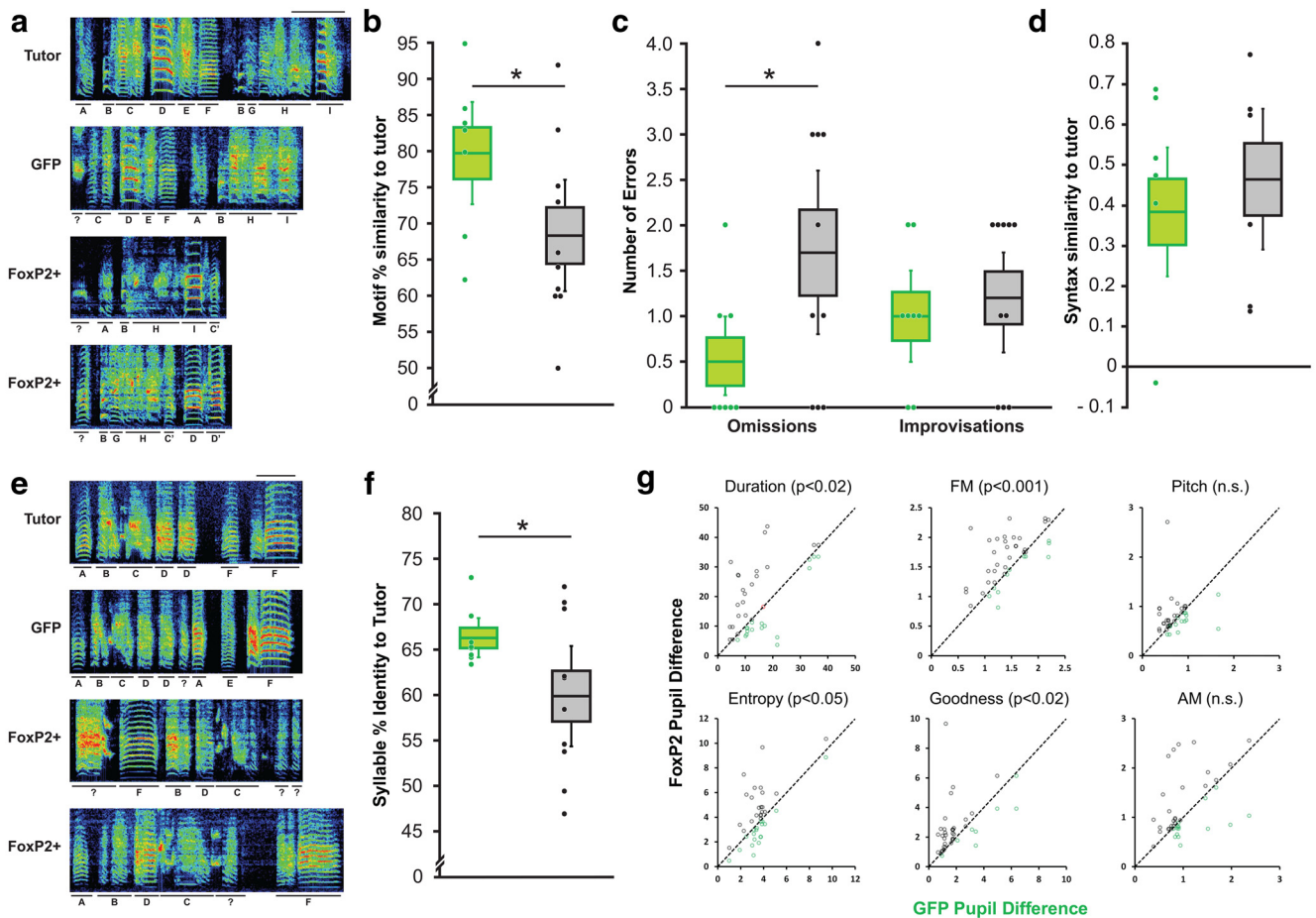
Next, we verified that overexpression leads to detectable increases in FoxP2 protein levels *in vivo*. Individual birds were injected with AAV1-GFP in one hemisphere and AAV1-FoxP2 in the other. Because FoxP2 protein levels vary as a function of behavior (Miller et al., 2008; Thompson et al., 2013), all comparisons were made within the same bird, using the GFP injected hemisphere as a control. Birds were allowed to sing for 2 h before being killed. Western blot analysis of protein from Area X micro punches revealed that injection of AAV1-FoxP2 was effective in increasing FoxP2 protein levels relative to levels in the GFP-injected control hemisphere (Fig. 1e).

### FoxP2 overexpression impairs vocal imitation

With a suitable virus in hand, we tested the effect of constitutive FoxP2 overexpression (FoxP2+) in Area X throughout the sensorimotor learning period on multiple facets of song behavior. The experimental timeline is shown in Figure 1f. We first examined overall tutor song imitation. Zebra finches learn a stereotyped and repeated unit of song known as a motif which is composed of a sequence of spectrally distinct units, or syllables (Immelmann, 1969). Song imitation was assessed at both the motif and syllable levels. Strikingly, the motifs of FoxP2+ birds were truncated and contained less of the tutor's source motif than did motifs of GFP control siblings (Fig. 2a). A motif similarity score was calculated using SAP (Tchernichovski et al., 2000) as one metric for quantifying this observation, because it provides unbiased information about the percentage of sound from the tutor's motif that was included in the pupil's motif. We found that FoxP2 overexpression resulted in a profound decrease in the motif similarity (Fig. 2b). To elaborate on these findings, the number of imitated syllables was counted manually, as well as the number of syllables from the pupil's vocal repertoire that were improvised. We found that FoxP2+ birds omitted more syllables but were no more likely to create an improvised syllable than were control birds (Fig. 2c). No difference in syntax accuracy was detected (Fig. 2d), indicating that both sets of birds tended to arrange their syllables in the same order as their tutor.

Next, for those syllables that were copied from the tutor, SAP was used to test for any differences in the quality of those copies (Fig. 2e). Across our dataset, syllable identity scores were much lower in FoxP2+ birds compared with GFP controls (Fig. 2f). To determine whether feature-specific errors could account for the poor copying of these syllables, feature difference measures were examined. These measurements represent feature-specific Euclidean distances such that larger differences correspond to larger errors. Because these measurements tend to be affected by syllable type, we compared individual syllables from FoxP2+ birds with their corresponding syllable performed by their GFP control brother. This syllable-by-syllable analysis revealed that FoxP2+ birds exhibited larger differences in FM, entropy, PG, and duration error than the corresponding syllables imitated by their GFP control brother (Fig. 2g). No differences were found for pitch or AM distances.

To assess the developmental trajectory of these impairments, the similarity of a given pupil's songs to its tutor's motif was examined across the critical period for sensorimotor learning, at 50, 70, and 90 d (Fig. 3a). At all three ages, the FoxP2+ birds had lower motif similarity scores compared with GFP pupils. From 50–70 d, the songs of both sets of birds became more similar to those of their tutors as evidenced by motif similarity scores, which stabilized between 70 and 90 d (Fig. 3b). This suggests that



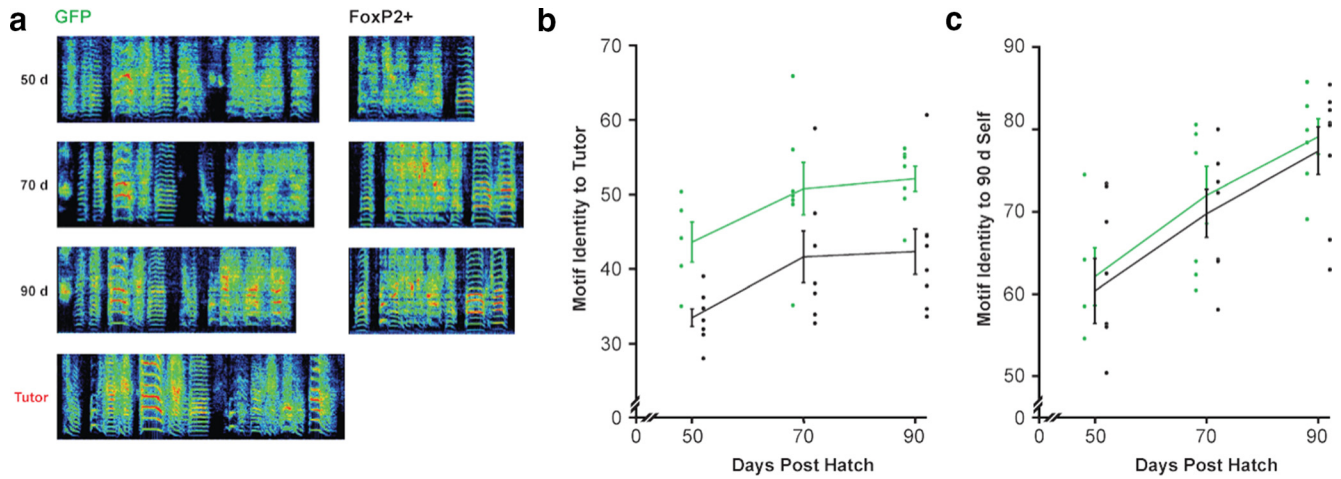
**Figure 2.** FoxP2 overexpression during the sensorimotor critical period disrupts vocal learning. *a*, Spectrograms (frequency range of 0–11 kHz) depict motifs from a tutor and his three pupils which each received a stereotaxic injection of AAV1 driving either GFP expression (GFP) or FoxP2 (FoxP2+). Scale bars, 200 ms. Syllables that correspond across motifs are underlined with black bars and identified by letters (question marks indicate unidentifiable syllables). *b*, Quantification of the similarity of each pupil’s motif to its tutor reveals that FoxP2+ birds (gray bars) have lower scores than those of GFP birds (green bars;  $p = 0.0269$ ,  $n = 8\text{GFP}/10\text{FoxP2+}$ , unpaired one-tailed bootstrap). Midline represents mean, upper and lower bounds of the box represent SE, upper and lower whiskers represent 95% confidence intervals, and points represent individual birds. *c*, Manual counting of syllables revealed that FoxP2 overexpression leads to an increase in the number of tutor syllables that were omitted by the pupil ( $p = 0.0247$ ,  $n = 8\text{GFP}/10\text{FoxP2+}$ , unpaired one-tailed bootstrap). In contrast, GFP and FoxP2+ pupils exhibit similar levels of improvised syllables ( $p = 0.3040$ ,  $n = 8\text{GFP}/10\text{FoxP2+}$ , unpaired one-tailed bootstrap). *d*, The motifs of GFP and FoxP2+ pupils exhibit similar levels of syntax similarity to their tutor’s motif ( $p = 0.2276$ ,  $n = 7\text{GFP}/8\text{FoxP2+}$ , unpaired one-tailed bootstrap). *e*, Exemplar spectrograms of a different tutor and his three pupils (1 GFP, 2 FoxP2+) highlight the low fidelity imitation of tutor syllables by FoxP2+ pupils. *f*, Poor syllable imitation by FoxP2+ relative to GFP pupils is reflected in lower-syllable identity scores ( $p = 0.0067$ ,  $n = 8\text{GFP}/10\text{FoxP2+}$ , unpaired one-tailed bootstrap). *g*, Syllable-by-syllable comparison of the feature-specific errors made by FoxP2+ pupils versus their GFP sibling. Black points above unity represent syllables for which the FoxP2+ sibling made larger errors, whereas green points below unity represent syllables for which the GFP sibling made larger errors. FoxP2+ pupils made larger errors for duration (100-sequential match), FM entropy, and goodness, but not pitch nor AM ( $p = 0.0115$ ,  $0.0004$ ,  $0.0239$ ,  $0.0108$ ,  $0.1196$ ,  $0.0761$ , respectively;  $n = 41$  syllables, paired one-tailed bootstrap).

FoxP2+ birds follow similar sensorimotor learning trajectories as those of normal birds, despite ultimate differences in tutor similarity. In keeping with this interpretation, when using the bird’s own song as an endpoint comparison, the two sets of birds made similar progress in attaining the adult version of their songs (Fig. 3c). Thus, FoxP2 overexpression appears to specifically interfere with vocal mimicry but does not nonspecifically interfere with the bird’s ability to modify its song over the course of sensorimotor learning.

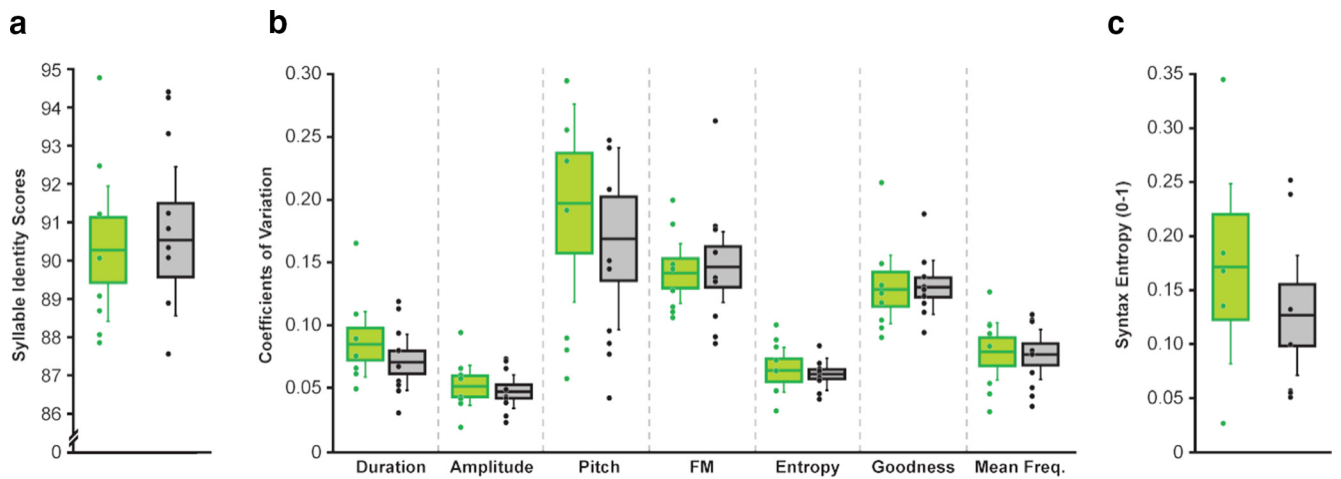
In sum, constitutive FoxP2 overexpression in Area X during sensorimotor learning led to incomplete copying of the tutor motif and poor copying of the tutor syllables with errors that spanned multiple song features. These learning deficits emerged early and persisted into adulthood because FoxP2+ birds failed to adaptively modify their songs to produce a copy of their tutors’ songs. Interestingly, these results are not directly opposite to those found following knockdown of FoxP2 (Haesler et al., 2007;

Murugan et al., 2013), providing support for the importance of behavior-driven cycling in FoxP2 expression in vocal learning.

We next examined the mature songs of FoxP2+ and GFP birds by assessing rendition-to-rendition variability in adulthood. Because both artificially and naturally low levels of FoxP2 in Area X are associated with increased variability (Haesler et al., 2007; Miller et al., 2010), one simple prediction was that, conversely, FoxP2 overexpression would decrease variability; a feature that is typically indicated by comparing multiple renditions of the same syllable to itself. Decreased variability would then be reflected in high self-similarity scores. Alternatively, FoxP2 overexpression may lead to decreased self-similarity much like that observed following knockdown. Contrary to either prediction, there were no differences in self-similarity for any of the measures that had been assessed in the pupil-to-tutor comparisons described above (Fig. 4a). Another way of analyzing vocal variability is to examine the coefficient of variation for specific features.



**Figure 3.** FoxP2 overexpression leads to imitation deficits that emerge early despite similar developmental trajectories. **a**, Spectrograms depict representative motifs of two pupils (GFP control and FoxP2+) at 3 stages of sensorimotor learning (50, 70, and 90 d) and that of their shared tutor. **b**, Motif identity scores indicate the similarity of a pupil’s motif to that of its tutor. Scores are plotted for three ages of each control GFP (green) and FoxP2+ (black) pupil. The latter group had lower scores at 50 d ( $p = 0.0028$ ,  $n = 5$  GFP/8FoxP2+, unpaired one-tailed bootstrap), 70 d ( $p = 0.0398$ ,  $n = 7$  GFP/7FoxP2+, unpaired one-tailed bootstrap), and 90 d ( $p = 0.0082$ ,  $n = 8$  GFP/8FoxP2+, unpaired one-tailed bootstrap). The motifs of both groups of pupils became increasingly similar to that of the tutor between 50 and 70 d (GFP:  $p = 0.0471$ ,  $n = 5$ ; FoxP2+:  $p = 0.0166$ ,  $n = 7$ ; paired one-tailed bootstrap), but not between 70 and 90 d (GFP,  $p = 0.2772$ ,  $n = 7$ ; FoxP2+:  $p = 0.1786$ ,  $n = 7$ ; paired one-tailed bootstrap). **c**, Motif identity scores indicate the similarity of a pupil’s motif to its own adult version. Scores are plotted for three ages of each control GFP (green) and FoxP2+ (black) pupil. Both groups followed similar developmental trajectories manifested by increases in similarity to adult song between 50 and 70 d (GFP:  $p = 0.0204$ ,  $n = 5$ ; FoxP2+:  $p = 0.0214$ ,  $n = 8$ ; paired one-tailed bootstrap), and between 70 and 90 d (GFP:  $p = 0.0004$ ,  $n = 7$ ; FoxP2+:  $p = 0.0214$ ,  $n = 7$ ; paired one-tailed bootstrap) and no difference between groups at the 50, 70, or 90 d time points ( $p = 0.7517, 0.6381, 0.6366$ , respectively; unpaired one-tailed bootstrap).



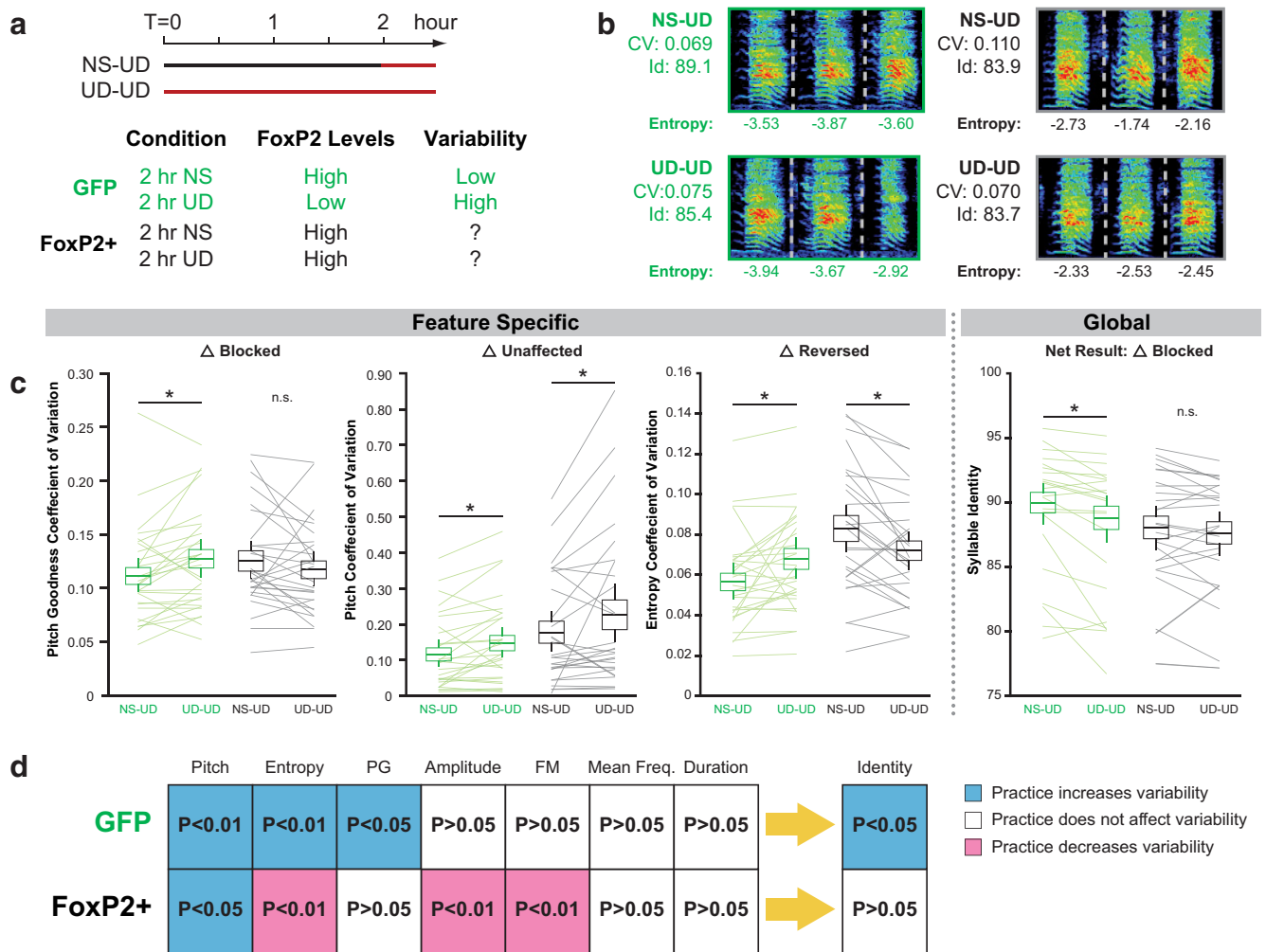
**Figure 4.** FoxP2 overexpression does not affect variability at adulthood. Green and gray bars represent GFP controls and FOXP2+ pupils, respectively. **a**, Overall rendition-to-rendition variability of syllables, as measured by syllable identity, was not significantly different between groups ( $p = 0.8489$ ,  $n = 8$  GFP/10FoxP2+, unpaired two-tailed bootstrap). **b**, FoxP2 overexpression did not affect feature-specific variability, as measured by the coefficient of variability for duration, amplitude, pitch, FM, entropy, PG, mean frequency ( $p = 0.2685, 0.5548, 0.5703, 0.7217, 0.8237, 0.928, 0.8371$ , respectively;  $n = 8$  GFP/10FoxP2+, unpaired two-tailed bootstrap). **c**, FoxP2 overexpression did not affect syntax variability ( $p = 0.8489$ ,  $n = 7$  GFP/8FoxP2+, unpaired two-tailed bootstrap).

Consistent with the self-similarity results, we were unable to detect any difference in CV values for any feature (Fig. 4b). Finally, the adult songs of FoxP2+ and GFP control birds did not differ in syntax entropy (Fig. 4c), a measure of vocal sequence variability. Together, these data indicate that FoxP2 overexpression throughout sensorimotor learning does not result in altered vocal variability in adulthood when vocal learning is complete.

Going further, we examined dynamic behavior-driven changes in vocal variability in ~75 d birds by comparing songs after the bird had spent the first 2 h of the day singing by itself [i.e., undirected singing (UD)] and after the bird had spent the first 2 h of another day not singing (NS), on 2 adjacent days (Fig. 5a). We have previously shown that this behavioral manipulation both decreases Area X FoxP2 mRNA and protein levels and increases vocal vari-

ability in the UD condition (Teramitsu et al., 2006; Miller et al., 2008; Miller et al., 2010; Hilliard et al., 2012). If these two phenomena are causally related, then preventing FoxP2 downregulation should prevent the acute increase in vocal variability. In support of our prior studies, 2 h of UD singing decreased the syllable identity scores of GFP control birds. Interestingly, this effect appeared blocked in FoxP2+ birds (Fig. 5a,b).

To gain insight into this observation, feature-specific changes in variability were again examined by evaluating the coefficient of variation for the features examined above (Fig. 5c,d). In GFP control birds, after UD singing, CVs were higher for pitch, PG, and entropy. No effect of condition was found for amplitude or frequency modulation, mean frequency, or duration. In sum, vocal practice in control birds leads to a semicoordinated increase



**Figure 5.** FoxP2 overexpression disrupts behavior-dependent transitions between low and high variability during learning. **a**, The approach and hypotheses are schematized. On adjacent days, birds were prevented from singing for 2 h or were allowed to sing undirected song for 2 h. The vocal variability immediately following these two epochs was measured. We predicted transitions between variability states in the GFP birds but not FoxP2+ birds. **b**, Exemplar syllables are shown here with their individual measurements and entropy CV and self-identity measurements based on 20 renditions of the syllable. **c**, We found divergent effects of FoxP2 overexpression on feature-specific variability, exemplified here by PG, pitch, and entropy CV. In each example, GFP birds showed significantly elevated variability following vocal practice (UD-UD). In FoxP2+ birds, however, the effect of vocal practice depended on the feature being measured: there was no effect on PG CVs, an increase in pitch CV in UD-UD condition, and a decrease in entropy CV in the UD-UD condition. The net result of these changes is a global practice induced increase in variability in GFP birds, which is blocked in FoxP2+ birds. **d**, A summary diagram of all the feature-specific and global changes observed. Notably, in GFP birds the semicoordinated practice-induced change in variability across multiple features gives rise to a global increase in variability. In contrast, in FoxP2+ birds feature-specific changes are not coordinated and do not give rise to a global change in variability. The means, confidence intervals, and *p* values represented here are shown in Table 1.

in variability across multiple features of song which likely accounts for higher levels of global variability, consistent with our prior studies in uninjected birds (Miller et al., 2010; Hilliard et al., 2012).

By contrast, in FoxP2+ birds, the feature level results were mixed (Fig. 5*b,c*). In line with our hypothesis and with the global similarity results, the increased CV for pitch goodness observed here in GFP controls (Miller et al., 2010; Hilliard et al., 2012; and previously in uninjected birds) was blocked by overexpression of FoxP2. Pitch variability, on the other hand, was unaffected by FoxP2 overexpression with both sets of birds showing increased CVs following vocal practice. Most surprisingly, we found a reversal of the effect for entropy, amplitude, and frequency modulation. These latter features were less variable after 2 h of UD singing in FoxP2+ birds. In short, vocal practice under conditions of constitutive FoxP2 overexpression results in a mixture of increases, decreases, and no effect on feature-specific variability. When considered together, these uncoordinated

effects do not translate into an overall change in vocal variability (Fig. 5*d*).

In addition to differences in dynamic regulation of variability within a given ~75 d bird across singing conditions (NS vs UD), we also observed intergroup (GFP vs FoxP2+) differences in overall variability. To control for dynamic changes in gene expression and vocal variability, we compared GFP control and FoxP2+ birds in the same singing conditions. For example, we compared the entropy CV of GFP birds in the NS condition with values for FoxP2+ birds in the NS condition. This was done for different measures of variability as well as by comparing these features in the UD condition. The analysis revealed global increases in variability as measured by syllable identity (in both the NS and UD conditions), as well as feature-specific increased CVs for entropy (in both the NS and UD conditions), pitch (in the UD condition), mean frequency (in the UD conditions), and amplitude (in the NS condition). Thus, in addition to disrupting practice-induced changes in variability at ~75 d, FoxP2 overex-

**Table 1.** Mean or CV values with 95% confidence intervals and exact *p* values from paired two-tailed bootstrapped test comparing variability in the NS-UD and UD-UD conditions within each group

		GFP			FoxP2+				
Feature		NS mean (95% CI)	UD mean (95% CI)	<i>p</i>	Direction	NS mean (95% CI)	UD mean (95% CI)	<i>p</i>	Direction
Mean	Duration	0.054 (0.043–0.065)	0.061 (0.048–0.075)	0.136	—	0.068 (0.055–0.083)	0.065 (0.055–0.074)	0.677	—
Mean	Amplitude	0.038 (0.032–0.045)	0.039 (0.034–0.044)	0.806	—	0.050 (0.044–0.055)	0.041 (0.036–0.047)	<b>0.006</b>	NS > UD
Mean	Pitch	<b>0.116 (0.082–0.154)</b>	<b>0.147 (0.109 to –0.189)</b>	<b>2 × 10<sup>4</sup></b>	UD > NS	<b>0.178 (0.125–0.235)</b>	<b>0.227 (0.152–0.312)</b>	<b>0.042</b>	UD > NS
Mean	FM	0.130 (0.108–0.157)	0.132 (0.108–0.158)	0.785	—	0.134 (0.11 to –0.160)	0.119 (0.100–0.140)	<b>0.008</b>	NS > UD
Mean	Entropy	<b>0.056 (0.048–0.065)</b>	<b>0.068 (0.058–0.078)</b>	<b>3 × 10<sup>4</sup></b>	UD > NS	<b>0.083 (0.071–0.094)</b>	<b>0.072 (0.063–0.081)</b>	<b>0.002</b>	NS > UD
Mean	PG	<b>0.111 (0.097–0.128)</b>	<b>0.128 (0.111–0.145)</b>	<b>0.002</b>	UD > NS	0.126 (0.110–0.143)	<b>0.117 (0.102–0.134)</b>	0.215	—
Mean	Mean frequency	0.074 (0.062–0.088)	0.081 (0.066–0.095)	0.263	—	0.099 (0.084–0.115)	0.113 (0.088–0.14)	0.209	—
Variance	Pitch	<b>0.412 (0.321–0.520)</b>	<b>0.492 (0.372–0.621)</b>	<b>0.004</b>	UD > NS	0.427 (0.326 to –0.546)	0.408 (0.331–0.488)	0.555	—
Variance	FM	0.170 (0.144–0.198)	0.175 (0.147–0.206)	0.554	—	<b>0.168 (0.142–0.197)</b>	<b>0.143 (0.121–0.167)</b>	<b>0.014</b>	NS > UD
Variance	Entropy	0.343 (0.290–0.405)	0.349 (0.297–0.403)	0.780	—	0.422 (0.358–0.497)	0.372 (0.326–0.419)	0.051	NS > UD
Variance	PG	0.431 (0.367–0.516)	0.435 (0.382–0.492)	0.917	—	0.454 (0.374–0.544)	0.400 (0.331–0.476)	<b>0.068</b>	NS > UD
Variance	Mean frequency	0.559 (0.473–0.655)	0.625 (0.511–0.751)	0.276	—	0.506 (0.407–0.610)	0.583 (0.437–0.738)	0.146	—
Global	Similarity	<b>96.7 (95.7–97.5)</b>	<b>95.8 (94.5–97.0)</b>	<b>0.001</b>	UD > NS	96.1 (95.1–97.0)	91.5 (90.6–92.3)	0.136	—
Global	Accuracy	<b>92.9 (92.0 to –93.7)</b>	<b>92.3 (91.4–93.2)</b>	<b>0.002</b>	UD > NS	91.7 (90.7–92.6)	91.5 (90.6–92.3)	0.339	—
Global	Sequence match	93.5 (92.6–94.4)	93.3 (92.4–94.2)	0.639	—	93.3 (92.2–94.4)	93.4 (92.6–94.3)	0.825	—
Global	Pitch difference	<b>1.159 (0.877 to –1.496)</b>	<b>1.284 (0.928–1.691)</b>	<b>0.032</b>	UD > NS	1.983 (1.597–2.396)	2.030 (1.703–2.362)	0.669	—
Global	FM difference	1.323 (1.188 to –1.458)	1.346 (1.206–1.488)	0.392	—	<b>1.538 (1.388–1.688)</b>	<b>1.474 (1.10315–1.683)</b>	<b>0.006</b>	NS > UD
Global	Entropy difference	3.680 (3.042 to –4.371)	3.945 (3.282–4.379)	0.051	UD > NS	5.235 (4.186–6.343)	4.738 (3.657–5.985)	0.055	NS > UD
Global	Goodness difference	1.839 (1.357–2.417)	1.946 (1.530 to –2.413)	0.386	—	1.701 (1.432–1.993)	1.749 (1.426–2.098)	0.500	—
Global	AM difference	0.896 (0.789–1.010)	0.927 (0.818–1.050)	0.157	—	1.058 (0.940–1.188)	1.052 (0.921–1.199)	0.790	—
Global	Identity	<b>90.0 (88.3–91.4)</b>	<b>88.8 (87.0–90.4)</b>	<b>0.001</b>	UD > NS	88.1 (86.3–89.7)	87.6 (85.9–89.2)	0.212	—
Global	Global match	<b>84.1 (82.2–85.8)</b>	<b>83.1 (81.1–84.9)</b>	<b>0.034</b>	UD > NS	81.5 (79.3–83.7)	81.2 (79.2–83.1)	0.556	—

Significant *p* values ( $p < 0.05$ ) are shown in **bold** whereas trend like *p* values ( $0.05 < p < 0.10$ ) are shown in *italics*. Where a significant or trend-like conditional effect was observed the direction column displays which condition showed higher levels of variability. Whether this corresponds to a larger or smaller value depends on the measure.

pression tends to increase variability for multiple song features and in both NS and UD conditions. Unlike the increased variability observed following FoxP2 knockdown (Haesler et al., 2007; Murugan et al., 2013); however, this increase does not persist in adulthood (Fig. 4).

## Discussion

Prior studies in the zebra finch species of songbird suggest that behaviorally linked downregulation of the speech-related gene FoxP2 plays an important role in vocal learning (Miller et al., 2008; Teramitsu et al., 2010; Shi et al., 2013). To test this idea, we used AAV-mediated gene expression to constitutively elevate FoxP2 in song-dedicated Area X of juvenile birds undergoing sensorimotor learning, which we confirmed by observing increased levels of FoxP2 mRNA, as well as GFP reporter (Fig. 1*b*) and FoxP2 protein (Fig. 1*e*) in Area X. Consistent with our expectation, these constitutively elevated levels impaired song learning and disrupted behaviorally induced changes in vocal variability. Systematic examination of the nature and timing of the vocal learning deficits allowed us to discern between a permissive versus dynamic model of FoxP2 function. Our data support a model in which dynamic regulation plays a necessary role in vocal learning.

A core finding was that FoxP2 overexpression leads to poor copying of tutor song. Errors occurred at the level of the motif and of individual syllables, indicating deficits in both the selection and execution, respectively, of correct vocal motor patterns. We argue that these behavioral deficits were the result of disrupted FoxP2 regulation within an otherwise intact circuit. Indeed, FoxP2 overexpression led to a behavioral phenotype distinct from that previously observed following electrolytic lesions of juvenile Area X (Scharff and Nottebohm, 1991). Such lesions lead to high sequence entropy and unusually long syllables, neither of which were observed here in FoxP2+ birds.

Several observations suggest that we disrupted what is normally a direct relationship between singing-related neural activity

and Area X FoxP2 levels. First, the relationship is robust: that FoxP2 levels are lower in Area X of singing, compared with nonsinging, zebra finches has been replicated at both the mRNA and protein level by other laboratories (Shi et al., 2013; Thompson et al., 2013) and ourselves (Teramitsu and White, 2006; Miller et al., 2010; Hilliard et al., 2012), and in another songbird species (Chen et al., 2013). None-the-less, neuromodulators could affect both singing behavior and FoxP2 levels in parallel. The impact of stress as such a factor appears unlikely based on two observations. First, in our experience, distracting a bird from singing (to obtain sufficient nonsingers in behavioral studies) did not lead to detectable changes in serum cortisol (Miller et al., 2008). Second, and in line with this, our microarray-based study (Hilliard et al., 2012) revealed that gene expression patterns, including FoxP2 levels, are similar between birds who were distracted from singing and those who did not sing by their own volition. This renders the idea that pre-existing FoxP2 levels drive singing levels unparsimonious. Further, it seems unlikely that motivational factors were similar in the two sets of nonsinging birds. Rather, the shared gene expression pattern of nonsingers, which were distinct from those in singers, more likely reflects the shared feature of not singing. Finally, singing upregulates micro-RNAs which directly target FoxP2 and repress its levels (Shi et al., 2013), providing a mechanism whereby singing can downregulate FoxP2. Further work will be necessary to determine the intervening steps in the pathway between song-related neural activity and FoxP2 downregulation. In any case, our results suggest that FoxP2 behavior-linked changes in Area X are critical for vocal learning.

In many respects, the learning-related behavioral phenotypes observed here match those observed following knockdown of FoxP2 (Haesler et al., 2007; Murugan et al., 2013). Specifically, both manipulations resulted in incomplete motif copying and poor syllable copying, including feature-specific-errors such as duration and entropy. In neither manipulation was there an ef-



fect on the developmental progress toward the bird's own final song. Although it could be argued that both outcomes are a non-specific consequence of altering the activity of any major transcription factor, there is precedence for specific but nonopposing effects following similar manipulations of the cAMP response element binding protein (CREB) in zebra finches: Overexpression of a dominant-negative isoform of CREB impaired song learning whereas overexpression of activated CREB had no effect (Abe K and Watanabe D, *SfN* 2013, JJJ16). A more parsimonious interpretation of our findings and those of Haesler et al. (2007) is that the convergent behavioral deficits point to a commonality of both interventions: disruption of behavior-driven cycling of FoxP2 in the basal ganglia, highlighting the importance of its dynamic regulation as a key determinant of normal vocal learning.

Given the increasing evidence that motor circuits play critical roles in encoding sensory representations (Iacononi et al., 1999; Roberts et al., 2012) one might argue that the observed deficits reflect impaired sensory, rather than sensorimotor, learning. This seems unlikely for several reasons. First, by isolating young pupils from their tutors at 45 d, before the peak of AAV-driven gene expression, we minimized the overlap between the presence of the tutor and the expression of the virus. Thus, the virus had little opportunity to affect acquisition of the sensory template, which can be complete within 2 weeks of exposure to the tutor (Böhner, 1990) and can be formed in as little as 2 h under operant conditions (Deshpande et al., 2014). Alternatively, viral-driven overexpression of FoxP2 could have interfered with retention of the sensory template. Again, this is unlikely given the wealth of evidence that the template is stored in primary and secondary auditory regions which then feed into the afferents of Area X (London and Clayton, 2008; Gobes et al., 2010). Last, FoxP2 downregulation occurs even in deafened birds (Teramitsu et al., 2010) and therefore is unlikely to be important for learning of a purely auditory memory. The observed impairments in FoxP2+ birds thus appear to be of sensorimotor origin.

In addition to being important for learning (Bottjer et al., 1984; Scharff and Nottebohm, 1991; Andalman and Fee, 2009), the corticobasal ganglia song control circuit is important for generating vocal variability (Kao et al., 2005; Aronov et al., 2008). Accordingly, we investigated the effect of FoxP2 overexpression on this latter role. Given the observation that viral knockdown and behavior-driven decreases in FoxP2 levels lead to increased vocal variability, one prediction was that FoxP2 overexpression would decrease variability. Contrary to this hypothesis, we found no effect of FoxP2 overexpression on song variability at adulthood, presenting a major difference between the knockdown and overexpression phenotypes. One explanation for these asymmetric results may be a saturation-effect of FoxP2 overexpression on dopamine signaling. Viral-driven FoxP2 knockdown in Area X of adult zebra finches was previously linked to decreased expression of dopaminergic signaling molecules, including DARPP-32, and increased vocal variability (Murugan et al., 2013). Here, in adult birds with chronically elevated levels of FoxP2, we found no effect on Area X DARPP-32 levels (data not shown), consistent with the lack of effect on variability at adulthood. Together, these results suggest that high FoxP2 levels are sufficient for a grossly normal basal ganglia circuit that generates normal levels of variability.

We were surprised then to find that, rather than decreasing vocal variability, FoxP2 overexpression actually increased it at ~75 d. This effect on variability is distinct from FoxP2 knockdown described above which causes an increase in variability that persists into adulthood (Haesler et al., 2007; Murugan et al., 2013). We suggest that the developmentally increased variability

observed here is an effect of retarded vocal imitation rather than an intrinsic effect of FoxP2 overexpression *per se*. Ravbar et al. (2012) have shown that variability decreases as syllables hone in on their target syllable in the model song. A corollary observed here is that variability was maintained in syllables that remained distant to their target. This may appear at odds with the observation that intrinsic vocal development appears intact. Our results, however, indicate that developmental timelines for intrinsic vocal repertoire, vocal imitation, and vocal variability can follow distinct trajectories. To recap, we found the following: no deficits in the trajectory toward the birds' final song, early deficits in the capacity to imitate the tutor song, and increased variability early in development that is corrected by adulthood.

More intriguingly, FoxP2 overexpression disrupted the normal practice-induced increase in vocal variability observed previously and in control birds. In unmanipulated birds (Miller et al., 2010; Hilliard et al., 2012) and the GFP birds in this study, multiple feature-specific increases in variability act in a semicoordinated manner to increase global variability. In FoxP2+ birds, these feature-specific effects are altered in various ways with the net outcome of blocking the global increase in variability. Based on these observations, we suggest that FoxP2 downregulation plays a critical role in coordinating behavior- or use-dependent transitions between brain states. These different states could represent plasticity and consolidation, exploration and exploitation, or some nonmutually exclusive combination. For example, viral knockdown of FoxP2 accelerates the propagation of neural activity through the AFP and, in principal, could increase vocal variability by affecting spike timing, jitter, and reliability (Murugan et al., 2013). This same mechanism could be harnessed by natural FoxP2 downregulation to not only increase variability but to also regulate Hebbian spike timing plasticity which is hypothesized to underlie zebra finch vocal learning (Troyer and Doupe, 2000a,b; Fiete et al., 2010).

The idea that a single molecule could be involved in both plasticity and consolidation depending on its expression level is supported by a recent study examining the role of circadian glucocorticoid fluctuations in motor learning (Liston et al., 2013). The authors found that high glucocorticoid levels were important for learning and dendritic spine formation, whereas low levels were important for consolidation and the stabilization of spines. Pharmacologically interfering with either of these two states led to a common deficit in motor learning. Similarly, both overexpression and knockdown of the molecule *gadd45 $\alpha$*  results in comparable decreases in dendrite complexity *in vitro* (Sarkisian and Siebzehnrbubl, 2012).

Together with our work, these studies support the dichotomy between permissive molecules, for which constitutively high (or low) levels are able to support learning versus gating molecules, which require dynamic transitions between high and low expression levels to switch between states of plasticity and consolidation, and suggest that motor learning requires a complex integration of both types of molecules. Our results lend insight into the treatment of both genetic and nongenetic speech and language disorders. In the case of genetically based disorders, simple gene replacement may be insufficient, as this would not address the importance of behaviorally linked on-line gene regulation. On the other hand, analogous speech-dependent gene cascades in humans could be taken advantage of to optimize behavioral speech therapy by aligning therapy sessions with points of maximum vocal plasticity.

## References

- 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](#)
- Böhner J (1990) Early acquisition of song in the zebra finch, *Taeniopygia guttata*. *Anim Behav* 39:369–374. [CrossRef](#)
- 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](#)
- Burger C, Gorbatyuk OS, Velardo MJ, Peden CS, Williams P, Zolotukhin S, Reier PJ, Mandel RJ, Muzyczka N (2004) Recombinant AAV viral vectors pseudotyped with viral capsids from serotypes 1, 2, and 5 display differential efficiency and cell tropism after delivery to different regions of the central nervous system. *Mol Ther* 10:302–317. [CrossRef Medline](#)
- Chen Q, Heston JB, Burkett ZD, White SA (2013) Expression analysis of the speech-related genes FoxP1 and FoxP2 and their relation to singing behavior in two songbird species. *J Exp Biol* 216:3682–3692. [CrossRef Medline](#)
- Day NF, Fraley ER (2013) Insights from a nonvocal learner on social communication. *J Neurosci* 33:12553–12554. [CrossRef Medline](#)
- Deshpande M, Pirlepsov F, Lints T (2014) Rapid encoding of an internal model for imitative learning. *Proc Biol Sci* 281:20132630. [CrossRef Medline](#)
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22:567–631. [CrossRef Medline](#)
- Fiete IR, Senn W, Wang CZ, Hahnloser RH (2010) Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron* 65:563–576. [CrossRef Medline](#)
- Gobes SM, Zandbergen MA, Bolhuis JJ (2010) Memory in the making: localized brain activation related to song learning in young songbirds. *Proc Biol Sci* 277:3343–3351. [CrossRef Medline](#)
- Groszer M, Keays DA, Deacon RM, de Bono JP, Prasad-Mulcare S, Gaub S, Baum MG, French CA, Nicod J, Coventry JA, Enard W, Fray M, Brown SD, Nolan PM, Pääbo S, Channon KM, Costa RM, Eilers J, Ehret G, Rawlins JN, et al. (2008) Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol* 18:354–362. [CrossRef Medline](#)
- Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C (2004) FoxP2 expression in avian vocal learners and non-learners. *J Neurosci* 24:3164–3175. [CrossRef Medline](#)
- Haesler S, Rochefort C, Georgi B, Licznarski P, Osten P, Scharff C (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. *Plos Biol* 5:e321. [CrossRef Medline](#)
- Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA (2012) Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. *Neuron* 73:537–552. [CrossRef Medline](#)
- Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. *Science* 286:2526–2528. [CrossRef Medline](#)
- Immelmann K (1969) Song development in the zebra finch and other estrildid finches. In: *Bird vocalizations* (Hinde RA, ed), pp 64–74. Cambridge: Cambridge UP.
- 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](#)
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–523. [CrossRef Medline](#)
- Liston C, Cichon JM, Jenneteau F, Jia Z, Chao MV, Gan WB (2013) Circadian glucocorticoid oscillations promote learning-dependent synapse formation and maintenance. *Nat Neurosci* 16:698–705. [CrossRef Medline](#)
- London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. *Nat Neurosci* 11:579–586. [CrossRef Medline](#)
- Miller JE, Spiteri E, Condro MC, Dosumu-Johnson RT, Geschwind DH, White SA (2008) Birdsong decreases protein levels of FoxP2, a molecule required for human speech. *J Neurophysiol* 100:2015–2025. [CrossRef Medline](#)
- Miller JE, Hilliard AT, White SA (2010) Song practice promotes acute vocal variability at a key stage of sensorimotor learning. *Plos One* 5:e8592. [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](#)
- Ravbar P, Lipkind D, Parra LC, Tchernichovski O (2012) Vocal exploration is locally regulated during song learning. *J Neurosci* 32:3422–3432. [CrossRef Medline](#)
- Roberts TF, Gobes SM, Murugan M, Ölveczky BP, Mooney R (2012) Motor circuits are required to encode a sensory model for imitative learning. *Nat Neurosci* 15:1454–1459. [CrossRef Medline](#)
- Sarkisian MR, Siebzehrubl D (2012) Abnormal levels of Gadd45alpha in developing neocortex impair neurite outgrowth. *Plos One* 7:e44207. [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. [Medline](#)
- Shi Z, Luo G, Fu L, Fang Z, Wang X, Li X (2013) miR-9 and miR-140-5p target FoxP2 and are regulated as a function of the social context of singing behavior in zebra finches. *J Neurosci* 33:16510–16521. [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](#)
- Teramitsu I, White SA (2006) FoxP2 regulation during undirected singing in adult songbirds. *J Neurosci* 26:7390–7394. [CrossRef Medline](#)
- Teramitsu I, Kudo LC, London SE, Geschwind DH, White SA (2004) Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *J Neurosci* 24:3152–3163. [CrossRef Medline](#)
- Teramitsu I, Poopatanapong A, Torrissi S, White SA (2010) Striatal FoxP2 is actively regulated during songbird sensorimotor learning. *Plos One* 5:e8548. [CrossRef Medline](#)
- Thompson CK, Schwabe F, Schoof A, Mendoza E, Gampe J, Rochefort C, Scharff C (2013) Young and intense: FoxP2 immunoreactivity in Area X varies with age, song stereotypy, and singing in male zebra finches. *Front Neural Circuits* 7:24. [CrossRef Medline](#)
- Troyer TW, Doupe AJ (2000a) An associational model of birdsong sensorimotor learning I: efference copy and the learning of song syllables. *J Neurophysiol* 84:1204–1223. [Medline](#)
- Troyer TW, Doupe AJ (2000b) An associational model of birdsong sensorimotor learning II: temporal hierarchies and the learning of song sequence. *J Neurophysiol* 84:1224–1239. [Medline](#)
- Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J, Alcock KJ, Connelly A, Frackowiak RS, Friston KJ, Pembrey ME, Mishkin M, Gadian DG, Passingham RE (1998) Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci U S A* 95:12695–12700. [CrossRef Medline](#)