eNeuro

Research Article: Confirmation | Cognition and Behavior

Beyond critical period learning: Striatal FoxP2 affects the active maintenance of learned vocalizations in adulthood

Nancy F. Day[1], Taylor G. Hobbs[1], Jonathan B. Heston[1] and Stephanie A. White[1]

[1]*Department of Integrative Biology and Physiology, University of California Los Angeles, Los Angeles, California 90095*

https://doi.org/10.1523/ENEURO.0071-19.2019

Received: 26 February 2019

Revised: 5 March 2019

Accepted: 6 March 2019

Published: 22 March 2019

N.F.D., J.B.H., and S.A.W. designed research; N.F.D., T.H., and J.B.H. performed research; N.F.D. and T.H. analyzed data; N.F.D. and S.A.W. wrote the paper.

Funding: HHS | NIH | National Institute of Mental Health (NIMH)
MH070712
;

Funding: HHS | National Institutes of Health (NIH)
HD07228
.

Conflict of Interest: The authors declare no conflicts of interest

NIH T32 HD07228 (NFD), NIH R01 MH070712 (SAW).

Correspondence should be addressed to Nancy F. Day at daynf@ucla.edu

Cite as: eNeuro 2019; 10.1523/ENEURO.0071-19.2019

Beyond critical period learning: Striatal FoxP2 affects the active maintenance of learned vocalizations in adulthood

Abbreviated title: FoxP2 is necessary for song maintenance

Nancy F. Day, Taylor G. Hobbs, Jonathan B. Heston*, and Stephanie A. White

Department of Integrative Biology and Physiology, University of California Los Angeles, Los Angeles, California 90095

*current address:
Department of Neurosciences, University of California, San Diego, CA 92093, USA

Author Contributions: NFD and SAW designed research; NFD, TGH, and JBH Performed Research; NFD and TGH analyzed data; NFD and SAW Wrote the paper

Corresponding author:
Nancy F. Day, Ph.D.
University of California Los Angeles
Department of Integrative Biology and Physiology
610 Charles E Young Dr. East
Los Angeles, CA 90095
Email: daynf@ucla.edu

Number of figures: 5
Number of Tables: 0
Number of Multimedia: 0

Number of words for:
Abstract – 251/250
Significance Statement – 120/120
Introduction – 635/750
Discussion – 1611/3000

Acknowledgements: We thank Chae Y. Kim, Petra Grutzik, and Aneesa Yousefi for pilot experiments and behavioral analysis; Todd H. Kimball, Sara N. Freda, Lauren Eisenman, and Mara Burns for RT-qPCR assistance; Samuel J. Sober and Lukas Hofmann for sharing code; and Caitlin M. Aamodt and Melissa J. Coleman for reviewing a preliminary draft of the manuscript.

Conflict of Interest: The authors declare no conflicts of interest

Funding: NIH T32 HD07228 (NFD), NIH R01 MH070712 (SAW).

Abstract

In humans, mutations in the transcription factor FOXP2 result in language disorders associated with altered striatal structure. Like speech, birdsong is learned through social interactions during maturational critical periods, and relies on auditory feedback during initial learning and on-going maintenance. Hearing loss causes learned vocalizations to deteriorate in adult humans and songbirds. In the adult songbird brain, most FoxP2-enriched regions (e.g. cortex, thalamus) show a static expression level, but in the striatal song control nucleus, Area X, FoxP2 is regulated by singing and social context: When juveniles and adults sing alone, its levels drop, and songs are more variable. When males sing to females, FoxP2 levels remain high and songs are relatively stable: This 'on-line' regulation implicates FoxP2 in ongoing vocal processes, but its role in the auditory-based maintenance of learned vocalization has not been examined. To test this, we overexpressed FoxP2 in both hearing and deafened adult zebra finches and assessed effects on song sung alone versus songs directed to females. In intact birds singing alone, no changes were detected between songs of males expressing FoxP2 or a GFP construct in Area X, consistent with the marked stability of mature song in this species. In contrast, songs of males overexpressing FoxP2 became more variable and were less preferable to females, unlike responses to songs of GFP-expressing control males. In deafened birds, song deteriorated more rapidly following FoxP2 overexpression relative to GFP controls. Together, these experiments suggest that behavior-driven FoxP2 expression and auditory feedback interact to precisely maintain learned vocalizations.

Significance Statement

Mutations within the FOXP2 gene impair speech and language. In zebra finch songbirds, the predominant model for investigating the neural and genetic mechanisms underlying human speech, FoxP2 is critical for song learning. Striatal FoxP2 expression levels correlate with song variability. We overexpressed FoxP2 in the striatopallidum of adult male zebra finches to assess its contribution to the maintenance of adult vocalizations independent of developmental perturbations. We tested the hypothesis that high FoxP2 expression promotes song stability by longitudinally assessing song in the presence and absence of auditory feedback and in two social contexts. We found that dysregulated FoxP2 interferes with hearing-dependent song maintenance. These results suggest that auditory-based regulation of FoxP2 is critical for the ongoing maintenance of adult vocalizations.

Introduction

A foundation for humans' ability to acquire language is speech, a learned vocal behavior that relies on sensorimotor experience. The discovery of a point mutation in the DNA binding domain of the forkhead box P2 (FOXP2) transcription factor in a British family with an inherited language impairment provided the first definitive link between this gene and speech and language (Lai et al., 2001). Individuals who inherit this mutation have speech deficits and structural abnormalities in the striatum, among other brain areas (Watkins et al., 2002).

The zebra finch songbird (*Taeniopygia guttata*), a species in which only males sing, is an essential animal model for studying learned vocal communication (Brainard and Doupe, 2013). Zebra finch song and human speech exhibit many parallels (Doupe and Kuhl, 1999), including:

1) acquisition of species-specific acoustic signals (e.g. native language/tutor song) during a sensory critical period; and 2) refinement of immature vocal signals (e.g. babbling/subsong) into precisely-controlled, mature vocalizations (e.g. words/crystallized song) using auditory-guided learning during a sensorimotor critical period (Bolhuis et al., 2010; Brainard and Doupe, 2013). Vocal plasticity persists into adulthood such that both groups are able to continually modify their vocalizations in order to maintain appropriate vocal output (Tumer and Brainard, 2007; Andalman and Fee, 2009; Sober and Brainard, 2009). However, in the absence of auditory feedback, vocalizations slowly deteriorate (Konishi, 1965; Cowie et al., 1982; Nordeen and Nordeen, 1992).

Fortuitously in songbirds, the neural circuitry that supports vocal learning, production and maintenance is composed of discrete, interconnected, and song-dedicated nuclei. One group of nuclei is critical for song production, and a cortico-basal ganglia-thalamo-cortical loop (the anterior forebrain pathway; AFP) is necessary for song learning. Within Area X, a nucleus that contains striatal and pallidal cell types (Farries and Perkel, 2002), *FoxP2* is dynamically regulated both by singing and the social context in which song is sung, as follows: In adults, expression is reduced following two hours of singing alone (undirected song; UD) relative to the robust levels observed following two hours of female-directed singing (FD; male courting a female) or in males that do not sing. In both adults and juveniles, the more the male sings alone, the lower its FoxP2 levels (Teramitsu and White, 2006; Miller et al., 2008; Teramitsu et al., 2010; Hilliard et al., 2012b; Chen et al., 2013; Shi et al., 2013; Thompson et al., 2013). Interestingly, when juvenile birds are deafened, singing-driven downregulation of *FoxP2* is no longer correlated with how much the bird sings (Teramitsu et al., 2010), suggesting that *FoxP2* levels are calibrated by auditory feedback to guide sensorimotor learning.

Interfering with behavior-linked *FoxP2* levels using viral-mediated knockdown or overexpression interferes with juvenile song learning such that birds are unable to properly imitate their memorized auditory template (Haesler et al., 2007; Heston and White, 2015; Burkett et al., 2018). Together, these data indicate that behavior-linked regulation of FoxP2 is critical for song development as young birds engage in trial-and-error learning to adaptively sculpt their vocalizations. In adults, knockdown of *FoxP2* prevents social context-dependent alterations to song (Murugan et al., 2013), suggesting that inappropriate *FoxP2* expression also impairs the precision of crystallized song.

To reveal whether FoxP2 participates in active song maintenance, we prevented behavior-driven downregulation of FoxP2 by overexpressing FoxP2 in Area X of adult male zebra finches and deafened a subset of them, similar to manipulations that demonstrated a key role for the AFP in adult song plasticity (Brainard and Doupe, 2000). A simple prediction was that high FoxP2 levels would promote song stereotypy, as is observed following performance of FD song or singing quiescence. However, we observed that constitutively high FoxP2 accelerated song deterioration in deafened birds. We also analyzed song produced in two social contexts (UD and FD), and conducted female preference tests to determine whether the resultant high vocal variability in FD song was behaviorally-meaningful.

Materials & Methods

Subjects

All animal use was in accordance with NIH guidelines for experiments involving vertebrate animals, approved by the University of California Los Angeles Chancellor's Institutional Animal

Care and Use Committee, and consistent with the American Veterinary Medical Association guidelines. Birds from our breeding colony were housed in climate-controlled rooms inside of cages and/or aviaries. A 14:11 lights on:lights off cycle was maintained; 30 minutes of dawn and dusk lighting was simulated each morning and evening. Birds had unlimited access to food, grit, and water, and were provided nutritional supplements (e.g. spray millet, green vegetables, and calcium supplements) and environmental enrichments (e.g. a variety of perches, swings, mirrors and water baths).

Experimental Timeline

Twenty-five male zebra finches (>120 days post hatch (dph), mean age = 153 dph) were recorded in sound attenuation chambers for a minimum of two weeks ('PRE') prior to injection of adeno-associated virus (AAV), serotype 1, driving expression of zebra finch FoxP2 or of GFP (Surgery A; FoxP2-AAV = 13, GFP-AAV = 12, mean age = 178 dph). We used AAV constructs previously described (Heston and White, 2015; Burkett et al., 2018), and followed those surgical procedures with the exception that 500nl of virus was injected per hemisphere.

At ~30 days following viral injection (range: 21-40 days), birds were re-recorded for two days ('POST'). All birds were then subjected to a second surgery (Surgery B, mean age = 208 dph), in which half of the birds were deafened via cochlear extraction (n = 12) and half were sham-deafened (n = 13) as described by Teramitsu et al. (2010). Birds were intermittently recorded for the following five months; songs were analyzed at 6, 14, 25, and 60 days (D06, D14, D25, D60) after deafening, and on the day of sacrifice (DOS). Time points were chosen to coincide with when changes to songs might be expected based on prior studies (e.g. D06 -

(Horita et al., 2008). Birds were sacrificed ~185 days following AAV injection (min = 182 days, max = 200 days). Birds were sacrificed by decapitation following 2 hours of undirected singing and brains were rapidly extracted and frozen by liquid nitrogen. A timeline for these experiments in schematized in Figure 1A.

Overexpression validation

Verification of targeting and over-expression of zebra finch *FoxP2* mRNA for all birds in which behavior was analyzed was done using *in situ* hybridization (data not shown) as described by Teramitsu and White (2006) and by RT-qPCR on tissue punches as described by Burkett et al. (2018). *FoxP2* expression was quantified relative to *Gapdh* (delta Ct).

To specifically assess FoxP2 protein levels following viral injection, two adult males were each injected with 500nl FoxP2-AAV in Area X of one hemisphere, and with 500nl GFP-AAV in the other. This approach allowed us to control for any difference in FoxP2 levels that are a result of dynamic behavioral regulation or inter-bird differences. After three weeks, males sang alone for two hours in the morning and were then sacrificed by rapid decapitation. Brains were extracted, flash frozen on liquid nitrogen and cryosectioned (Leica Microsystems, Bannockburn, IL) in the coronal plane at a thickness of 30µm. Tissue punches of Area X were made using a 20-gauge Luer adapter (BD, Sparks, MD) at a depth of 1mm as in Miller et al. (2008). Western blotting was also as described in Miller et al. (2008). Expression levels of FoxP2 in Figure 1C are presented and quantified as percent change in the AAV-FoxP2 hemisphere relative to the AAV-GFP hemisphere.

A second group of males (n = 15, mean age = 163 dph) was used to verify persistent overexpression of FoxP2 across the experimental time-course and to coincide with the time points in which song behavior was analyzed (e.g. Fig. 1A experimental time course D14 post-deafening corresponds to Fig. 1D post-injection day 35 in the AAV expression time course validation). Of these, 12 birds received 500nl of AAV-FoxP2 to each Area X after which 3 were sacrificed at each time point (20, 35, 45, and 80 days post-surgery); 3 birds (180 dph) served as uninjected controls. At each time point, birds were rapidly decapitated in the morning before any song had been produced, and brains were extracted, frozen on liquid nitrogen, and stored at -80˚C until use. Tissue punches from Area X and the adjacent ventrostriatal pallidum (VSP) were homogenized in 100µl Qiazol (Qiagen) and total RNA was extracted using the Direct-Zol MicroRNA Prep Kit (Zymo Research). 100ng RNA was reverse-transcribed to cDNA (Applied Biosystems High Capacity RNA-to-cDNA kit, #4387406) for qPCR (as described above). The delta-delta Ct method (Livak and Schmittgen, 2001) was used to calculate fold-changes in the expression FoxP2, the D1 and D2 dopamine receptors (D1R and D2R, respectively), as well as the dopamine biosynthetic enzyme tyrosine hydroxylase (TH), relative to Gapdh in Area X compared to VSP. Primer sequences were designed for zebra finch D1R (112bp), D2R (206 bp), and TH (107 bp) using the NCBI Primer Design Tool, and were validated using melt curve analysis and standard curves. Primers sequences were: D1R – FOR: CCGGGAGGACATTACAGTTTAG; D1R – REV: TGCAGTTCCACCCGTATTTAG; D2R – FOR: CCCAGCAGAAGGAGAAGAAAG; D2R – REV: CTCGATGTTGAAGGTGGTGTAG; TH- FOR GCACCCTGAAGAGCTTGTAT; TH-REV: CAGCTGAGGGATGTTGTTCT.

Song recording and analysis

Undirected (UD) song was collected across the entirety of the experiment by housing animals singly within a sound attenuation chamber. Although animals were moved to social housing in between experimental time points, each bird was recorded within the same isolation chamber for the duration of the experiment. All reasonable attempts were made to record a given bird using the same microphone and recording devices/settings, with occasional differences in the quality of recordings between time points. Sounds were acquired using Shure SM58 microphones (Nile, IL), digitized using a PreSonus Firepod or AudioBox (44.1 kHz sampling rate, 24 bit depth; Baton Rouge, LA) and recorded using Sound Analysis Pro (SAP) 2011 software (Tchernichovski et al., 2000).

Songs were analyzed at the level of the motif as well as the syllable, each of which were hand-segmented using custom-written Matlab code (Tumer and Brainard, 2007). Motifs were identified as repeated units of song composed of multiple syllables. Introductory notes were included in all analyses to assess any effect of stuttering following deafening (Horita et al., 2008; Kubikova et al., 2014). Canonical and non-canonical renditions of motifs were included in the analyses in order 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. Motif similarity as well as the phonology and syntax of syllables were compared to PRE vocalizations at each subsequent time point (see Figure 1A), as specified below:

Motif Similarity – The 'Similarity Index' (Mandelblat-Cerf and Fee, 2014) quantified how well birds imitated their PRE motifs. Twenty motifs, collected from songs produced on two consecutive mornings, that were sung within one week prior to Surgery A (PRE) were compared against 30 song bouts for each day included in the analysis (PRE1, PRE2 – morning of Surgery

A, POST1, POST2 – morning of deafening, D06, D14, D25, D60, DOS). Of note, PRE1 and POST1 were dates immediately preceding PRE2 and POST2.

Syllable Similarity - The first ~450 syllables of each analysis time point were segmented within Matlab using an amplitude threshold, grouped into syllable clusters, and assigned an arbitrary label using the semi-automated clustering algorithm VoICE (Burkett et al., 2015). All spectral features were calculated using Sound Analysis Tools (SAT; http://soundanalysispro.com/matlab-sat) in Matlab. We quantified both syllable similarity to PRE using custom-written Matlab code derived from the 'Similarity Batch' function of SAP 2011 (Tchernichovski et al., 2000; Burkett et al., 2015). To calculate syllable similarity over time, 30 renditions of each syllable at each time point were compared to 30 renditions of that syllable produced during PRE. 'Syllable Similarity' was represented by the mean of these 900 comparisons, and normalized to the mean of PRE vs PRE1 and PRE vs PRE2 to account for day-to-day variability within a bird. Higher scores indicate greater similar to songs produced before viral (Surgery A) or auditory manipulation (Surgery B).

Spectral Variability – For each bird and time point, the coefficient of variation (CV) was calculated using the first 40 renditions of each syllable for the following acoustic features: Entropy, Entropy Variance, Duration, Pitch Goodness, Pitch, and Frequency Modulation (FM). All acoustic features were calculated using SAT. To assess how these syllables changed relative PRE, the mean CV Effect Size (CV ES) for each bird was calculated by averaging the CV ES of all syllables. The CV ES for each syllable was determined using the following formula: CV ES = $(CV_{\text{Time Point}} - CV_{\text{Pre}}) / (CV_{\text{Time Point}} + CV_{\text{Pre}})$.

196 *Syllable Preservation* – We calculated both the number of syllables that remained in a bird's
197 motif and the number of syllables that were added to a bird's motif following deafening. First,
198 the 'core syllables' of a motif were identified as syllables that were present in >60% of a bird's
199 motifs before deafening. An average Syllable Preservation percentage was calculated by dividing
200 the total number of core syllables present each day by the total syllables produced on that day.
201 For example, a Syllable Preservation score of 0.95 indicates that 95% of the syllables produced
202 that day were syllables integral to a motif.

203 *Syntax analysis* - For each bird and time point, we created a transition probability matrix from
204 strings of identified syllables. Transition probability matrices of PRE vs. each time point were
205 correlated and included syllables that were omitted or introduced following deafening. A
206 similarity score of 0 reflects no relationship to PRE sequencing, whereas a score of 1 indicates an
207 exact match to PRE sequencing (Miller et al., 2010; Burkett et al., 2015).

208 *Social Context* – We elicited female-directed (FD) song from male birds (n = 13 birds, n = 23
209 syllables) prior to and following viral overexpression of GFP or FoxP2. A rotation of 6 female
210 zebra finches was used to prompt courtship song over the course of two hours. Females were
211 placed in the cage with the male for 10 minutes at a time, removed, and replaced with another
212 female. All interactions were video recorded and visually monitored to verify that males were
213 directing their songs to a female. To assess variability in pitch, the fundamental frequency was
214 measured for syllables containing harmonic elements in both UD and FD song epochs. The
215 coefficient of variation (CV) of the fundamental frequency (FF) was calculated using 25
216 pseudorandomly-selected renditions of each syllable in each context. Syllables that did not

exhibit the characteristic decrease in CV_{FF} during FD song (Kao and Brainard, 2006) in the PRE condition were excluded from all analyses (n = 8).

Female Preference

To determine if FoxP2 overexpression influenced courtship song quality, sexually-naïve females were used to assess preference for songs produced before and after viral injection. Mature female finches (n = 35; 100-120 dph) were selected from female-only group housing and moved to individual cages within sound attenuation chambers. Cages (38cm x 25cm x 28cm) were outfitted with two static perches and two 'switch' perches that lowered when the bird landed on them. Switch perches were made by securing a 6cm red pipe cleaner to a miniature switch requiring minimal force (Cherry D429-R1ML) and were placed on the back wall of the cage, each 4cm from the side walls and 15cm from the ground. A vertical plastic barrier (12cm) was placed in the middle of the cage to create separate, but connected, areas of the cage, and to impede spurious motion from one side to the other. A single speaker (Pioneer Electronics) was placed behind the barrier. Activation of a switch resulted in sound playbacks. Playbacks were controlled using the 'Operant Conditioning' module of SAP 2011 with a NI USB-6501 (National Instruments, Austin, TX). A schematic of the female preference testing cage is shown in Figure 5C.

Stimuli - Playbacks consisted of sound files containing 2-5 motifs. Five representative song files were generated for each of the four social contexts (UDPre, UDPost, FDPre, FDPost) and were selected for playback in a random order by SAP2011. All songs were unfamiliar to females; none had interacted with any of the males whose song was presented during the trials. Females were

trained to associate perch activations with sound playbacks using Isolate song and FD song. 'Isolate song' is produced by birds raised in the absence of a tutor and is not preferred by females (White, 2001).

Preference Testing - For each trial, females participated in two phases of testing: 'Silence' and 'Playback'. During the silence phase (2 hours), beginning at lights on, we determined a perch bias (PP – preferred perch; UP – nonpreferred perch) by observing the number of activations on each of the perches in the absence of auditory stimuli. FD song was always paired with the perch that received fewer perch activations to counteract the perch bias. Females (n = 16) that were unable to overcome their perch preference to demonstrate a song preference for FD song were excluded from further testing. A trial was excluded from analysis if the female failed to activate each perch 5 times during each of the silence and testing phases. Each male was tested by a minimum of 5 females who were tested on both PRE and POST songs. Song sets were grouped relative to 'Surgery A', such that females only heard PRE or POST songs in a given trial (e.g. UDPre vs. FDPre). Females were tested a minimum of 3 times on each set of songs (min = 3, max = 6, average = 3.5). A preference score, taking into account the perch bias during the 'Silence' phase, was calculated using the following formula:

$$\text{Preference Score} = \frac{[Playback_{FD} - Playback_{UD}]}{[Playback_{FD} + Playback_{UD}]} - \frac{[Silence_{UP} - Silence_{PP}]}{[Silence_{UP} + Silence_{PP}]}$$

A Preference Score > 0 indicates preference for FD song; negative values indicate a preference for UD song.

Experimental Design and Statistical Analysis

The criterion for statistical significance was set at alpha = 0.05. All significance levels were calculated as 2-tailed except for cases in which we had prior experimental expectation of the outcome. Such cases are noted the text. Prism 8 (GraphPad) was used to perform all statistical tests. A D'Agostino and Pearson normality test was performed on each data set to determine normality. To calculate statistically-significant effects over time, metrics from each time point were compared to PRE using a Kruskal-Wallis one-way ANOVA within each of the four groups (e.g. FoxP2-Hear, GFP-Deaf). Details for all statistical tests are included in either the results or figure legends.

Code Accessibility

Custom-written Matlab code (NFD) for the generation of syllable similarity scores using the Similarity Module is adapted from Burkett et al. (2015).

Results

Overexpression of FoxP2 in Area X of adult zebra finch males

Adeno-associated virus serotype 1 (AAV1) and the CAG promoter were used to drive overexpression of FoxP2 or GFP (Figure 1B) in the song dedicated striatal nucleus, Area X, in adult zebra finch males. This viral construct has been previously used to elevate FoxP2 levels in Area X of young songbirds, which resulted in vocal learning deficits (Heston and White, 2015; Burkett et al., 2018). To validate expression in adults, first, Western blot analysis of protein from two birds demonstrated that within each bird, FoxP2 was elevated in Area X of the hemisphere injected with AAV-FoxP2 relative to that injected with AAV-GFP (Figure 1C; see Methods). Second, *FoxP2* mRNA was quantified using *in situ* hybridization (data not shown; see Burkett et

al., 2018) and qRT-PCR as follows: In the cohort of unrecorded birds that were used to assess the time-course of FoxP2 over-expression, high Area X FoxP2 levels persisted in all animals for ≥80 days following injection compared to age-matched uninjected animals (Figure 1D; Mann-Whitney $p = 0.0002$, one-tailed; uninjected $n = 3$ vs. injected $n = 11$). To improve the clarity of Figure1D, data from one bird in the 20d group that received AAV-FoxP2 was removed for having FoxP2 expression 2 SD *greater* than the mean (ddCt = 8.74, mean with bird = 3.65, mean without bird = 1.11). Inclusion of that data point would not alter the direction of the reported changes. Importantly, these animals were sacrificed without having sung, as FoxP2 mRNA and protein levels vary depending on how much a bird sings (Teramitsu and White, 2006; Miller et al., 2008; Teramitsu et al., 2010). Among the birds whose behavior was analyzed for this study and who were permitted to sing for 2 hours prior to sacrifice, FoxP2-injected animals showed an increase in *FoxP2* expression compared to GFP-injected animals (Figure 1E; $p = 0.04$; one-tailed unpaired t-test, FoxP2 $n = 13$; GFP $n = 11$). Interestingly, separation of these two groups (FoxP2-injected and GFP-injected) into hearing and deaf subgroups suggests that this increase is largely driven by the FoxP2-deafened animals (Figure 1F). This trend toward an increase in the FoxP2-deaf animals (one-way ANOVA: $F(3,20) = 2.14$, $p = 0.127$) is not due to less singing as the average time spent singing did not differ among the four groups (mean, seconds ± SEM: FoxP2-Hear – 226.8 ± 32.8; FoxP2-Deaf – 268.0 ± 50.4; GFP-Hear – 301.0 ± 75.0; GFP-Deaf – 373.5 ± 76.9; one-way ANOVA: $F(3,17) = 1.115$, $p = 0.370$).

FoxP2 overexpression positively correlates with dopaminergic markers D1R and TH

To further validate our viral manipulation, we predicted that overexpression of FoxP2 would change the expression of specific markers in Area X. Prior work shows that knocking

down FoxP2 in Area X leads to diminished expression of certain dopamine markers, including D1R (Murugan et al., 2013). We found that *D1R* (Spearman's r = 0.62; p = 0.016, n = 15 pairs) and *TH* (Spearman's r = 0.60; p = 0.026, n = 15 pairs) were positively correlated with FoxP2 expression (Figure 1G). *D2R* expression levels were not correlated with *FoxP2* expression (Spearman's r = 0.153, p = 0.58, n = 15 pairs; Figure 1H), consistent with a previous study that identifies co-localization of Foxp2 with D1R, but not D2R in mouse striatum (Fong et al., 2018).

Undirected song quality and sequencing is unaffected by FoxP2 overexpression in hearing adults

Overexpression or knockdown of FoxP2 in Area X during sensorimotor learning impairs vocal learning (Haesler et al., 2007; Heston and White, 2015; Burkett et al., 2018). However, no role for FoxP2 in the maintenance of adult vocalizations, such as crystallized song, has been described. Overall, the songs produced following AAV-FoxP2 were visually similar to songs sung before surgery (Figure 2A, B). To check for any subtle alterations to song, we examined syllable and motif similarity produced 3 weeks after surgery (POST1, POST2) to syllables and motifs produced prior to surgery (PRE). As a proxy for syllable 'quality', syllable similarity scores were calculated using MATLAB code (Burkett et al., 2015). A set of PRE syllables from two days just prior to surgery was compared against a set of syllables produced the morning before AAV injection. POST syllables from two consecutive days >20 days following surgery were combined to compare against the same set of PRE syllables. No differences in syllable similarity (AAV-GFP = 12 birds; AAV-FoxP2 = 13 birds) were detected for either group PRE vs. POST (AAV-GFP: p = 0.278, two-tailed paired Wilcoxon; AAV-FoxP2: p = 0.677, two-tailed paired Wilcoxon; Figure 2B).

Motif-level analyses were also performed to detect overall changes to song structure, including spectral quality and sequencing. The Similarity Index (Mandelblat-Cerf and Fee, 2014) was used as an unbiased metric to compare all songs performed following AAV injection and/or sham deafening surgeries to PRE song (Figure 2C). Five PRE motifs were selected and scored against 20 bouts produced by each individual at each time point [POST1, POST 2, and Days (D) 06, 14, 25, 60 after sham or deafening surgeries, and the day of sacrifice (DOS)]. A two-way ANOVA indicated a significant main effect of time, $F(7,78) = 3.15$, $MS = 0.033$, $p = 0.006$. No significant main effect was detected for group [$F(1,78) = 0.057$, $MS = 0.0006$, $p = 0.815$], nor for interaction between group and time [$F(7,78) = 0.230$, $MS = 0.002$, $p = 0.977$]. Post hoc analyses using Sidak's multiple comparisons test showed that similarity scores at DOS for hearing birds were significantly different from PRE for both the AAV-GFP and AAV-FoxP2 groups (GFP: $p = 0.023$; FoxP2: $p = 0.043$); no other time points significantly differed from PRE.

Finally, we examined the sequencing of syllables using a weighted syntax score (Figure 2D). As with overall similarity, we saw no differences between groups or within groups at any time point (two-way ANOVA: main effect for time, $F(7,79) = 1.60$, $MS = 0.029$, $p = 0.148$; main effect for group, $F(1,79) = 0.373$, $MS = 0.007$, $p = 0.543$; interaction, $F(7,79) = 0.159$, $MS = 0.003$, $p = 0.992$). The variability of syntax scores in the AAV-FoxP2 group can be attributed to two animals whose syntax was variable from the onset of behavioral analysis (PRE - PRE comparisons were 0.49 and 0.46, compared to the other five animals in the group whose scores were all >0.90; all animals in the AAV-GFP group had >0.95 PRE similarity scores).

FoxP2 overexpression hastens deafening-induced song deterioration

Crystallized zebra finch song is characterized by highly stereotyped sequences of syllables and low phonological variability. Given that we might not observe obvious differences in song following overexpression of FoxP2 due to the relative stability of the behavior, we deafened a subset of birds who received AAV-GFP and AAV-FoxP2 to eliminate auditory feedback, a manipulation that causes degradation of vocalizations (Nordeen and Nordeen, 1992; Woolley and Rubel, 1997). This manipulation allowed us to test whether or not FoxP2 overexpression alters deafening-induced song deterioration. Behavioral variability is correlated with singing-induced downregulation of FoxP2 juvenile finches (Miller et al., 2008), whereas highly-stereotyped female-directed singing is correlated with robust expression of *FoxP2* mRNA (Teramitsu and White, 2006). Thus, one hypothesis was that preventing FoxP2 downregulation would stabilize song, reducing its rendition-to-rendition variability, and delay song deterioration following the removal of auditory feedback. In contrast, we observed that deafening coupled with FoxP2 overexpression accelerated the deterioration of adult song.

Representative spectrograms from two deafened siblings show that the brother who received AAV-FoxP2 had more profound alterations to his song (Figure 3A, B). To quantify this change, we performed motif/bout level similarity scoring to PRE song at four time points following deafening (Days (D) 06, 14, 25, and 60) and on the day of sacrifice (DOS). A two-way ANOVA confirmed that both time ($F(7,70) = 11.64$, $MS = 0.246$, $p < 0.0001$) and group ($F(1,70)$, $MS = 0.102$, $p = 0.031$) were significant main effects (interaction: $F(7,70) = 1.163$, $MS = 0.025$, $p = 0.335$). Within the groups, compared to PRE, Sidak's multiple comparisons tests were significant for AAV-GFP-deafened animals at DOS ($p = 0.0007$) and for AAV-FoxP2-deafened animals at D14, D25, D60, and DOS ($p = 0.0006$, 0.0002, < 0.0001, and <0.0001, respectively; Figure 3B). A post hoc Sidak's multiple comparisons test showed that at no time point did

groups differ from one another. Values for GFP-Deaf and FoxP2-Deaf groups showed the greatest separation from each other at D14 (mean ± SEM: GFP - 0.870 ± 0.092, FoxP2 – 0.644 ± 0.047; $p = 0.265$) and D25 (mean ± SEM: GFP - 0.840 ± 0.059, FoxP2 – 0.653 ± 0.110; $p = 0.302$); p-values at all other time points were $p > 0.8$.

Early-onset song deterioration in adult males overexpressing FoxP2 without auditory feedback could be the result of spectral degradation and/or changes in song sequencing. To distinguish between these, we quantified the effect of deafening on the CV of acoustic features in all groups. Deafened animals overexpressing FoxP2 showed greater variability in three spectral features at earlier time points relative to deafened GFP animals (Figure 4A). At D25, Entropy ($p = 0.025$), Entropy Variance ($p = 0.004$), and FM ($p = 0.04$) were more variable in FoxP2-Deaf birds compared to GFP-Deaf birds (two-way ANOVA with Sidak's test for multiple comparisons). Additionally, Entropy Variance was significantly more variable on DOS ($p = 0.04$) in FoxP2-deaf birds. However, GFP-Deaf birds, compared to FoxP2-Deaf birds, did not show a significant increase in the variability of any spectral feature at any time point. No statistically significant differences were observed for any spectral feature at any time point in the two groups of hearing animals (a two-way ANOVA was performed for each spectral feature between hearing groups over time; none were significant). Finally, we examined the presence/absence of each syllable following deafening and the sequencing of song syllables. We observed that deafened AAV-FoxP2 animals dropped syllables from their motifs more rapidly than AAV-GFP deafened animals (Figure 4B); however, the percentage of dropped syllables was not significant between groups (two-way ANOVA: Group – $F(1,61) = 3.017$, $MS = 0.050$, $p = 0.087$; Time – $F(6,61) = 4.39$, $MS = 0.072$, $p = 0.0010$; Interaction – $F(6,61) = 0.27$, $MS = 0.004$, $p = 0.949$). Over the course of recording, Sidak's post hoc test showed that both GFP-

deaf and FoxP2-deaf animals had significantly fewer syllables at PRE vs. DOS (p values = 0.045 and 0.0073, respectively). Lasting syntactical changes were present as early as D14 in AAV-FoxP2 animals compared to the later onset of these changes at D60 in AAV-GFP animals (Figure 4C). Together, these results indicate that a combination of spectral and sequencing alterations underlie the acceleration of deafening-induced song deterioration in animals overexpressing FoxP2 in Area X.

Female-directed song is more variable following FoxP2 overexpression

Syllables with harmonic elements are sung with less rendition-to-rendition variability during female-directed song than undirected song (Kao et al., 2005). Knockdown of FoxP2 within Area X of adult zebra finches abolishes this social context-dependent change in vocal variability, as measured by the coefficient of variation (CV) of the fundamental frequency (FF) (Murugan et al., 2013). We calculated the CV of FF in the harmonic elements of syllables in hearing birds to determine if FoxP2 overexpression alters rendition-to-rendition variability in female-directed song (Figure 5A). As expected, prior to overexpression of FoxP2 or GFP, harmonic elements were performed with a significantly lower CV during female-directed song compared to undirected song (AAV-GFP: UD Pre vs FD Pre - $p = 0.0002$, $n = 12$ syllables (6 birds), one-tailed Wilcoxon matched-pairs signed rank test; AAV-FoxP2: UD Pre vs FD Pre - $p = 0.0001$, $n = 13$ syllables (7 birds), one-tailed Wilcoxon matched-pairs signed rank test). However, after FoxP2 overexpression, the CV of harmonic elements in FD song was no longer significantly different from UD renditions ($p = 0.064$, one-tailed Wilcoxon matched-pairs signed rank test; Figure 5B). AAV-GFP birds continued to perform FD song with lower variability than UD song (one-tailed Wilcoxon matched-pairs signed rank test, $p = 0.0002$, $n = 12$ syllables from

412 6 birds). We compared the mean number of introductory notes, the mean bout duration, and
413 mean motif duration in both PRE and POST songs (UD and FD). We did not find any differences
414 in these metrics following virus injections (data not shown).

415 Multiple measures of mRNA expression reveal that Area X *FoxP2* levels are lower in
416 adult males following undirected song than following production of highly-stereotyped female-
417 directed song (Teramitsu and White, 2006). One prediction based on these observations was that
418 preventing FoxP2 downregulation by overexpression may result in songs with lower rendition-
419 to-rendition variability than is typically present in undirected song. In contrast to this idea, there
420 were no song features in which FoxP2 overexpression reduced vocal variability (Figure 4A). The
421 CV of the fundamental frequency (FF) of harmonic elements within syllables did not change in
422 undirected song following AAV-GFP (two-tailed Wilcoxon matched-pairs signed rank test, $p =$
423 0.083, $n = 11$ syllables) or AAV-FoxP2 (two-tailed Wilcoxon matched-pairs signed rank test, $p =$
424 0.094, $n = 13$ syllables) injection.
425

426 *FoxP2 overexpression tempers females' preference for female-directed song*

427 Female preference for male song is inversely correlated with song variability (Woolley
428 and Doupe, 2008; Chen et al., 2017; Heston et al., 2018). To determine whether increased
429 variability of FD song induced by FoxP2 overexpression was perceived by conspecifics, and thus
430 of potential ethological relevance, we tested whether female zebra finches altered their behavior
431 in response to more stereotyped (AAV-FoxP2 PRE FD) or variable (AAV-FoxP2 POST FD)
432 songs. We used a perch-hop paradigm (Figure 5C; see Materials and Methods) to measure
433 sexually-naive females' preferences for songs performed under different social (UD vs. FD) and

viral (PRE vs. POST; GFP vs. FoxP2) conditions. We accounted for each female's bias for activating a specific perch by calculating an effect size for perch preference ([Perch 1 - Perch 2]/[Perch 1 + Perch 2]) when no playbacks were presented ('Silence') vs. when playbacks of either FD or UD song were paired with perch activations. (Notably, FD song playbacks were always paired with the lesser preferred perch during the 'Silence' testing period.) To obtain a 'Preference Score' (Figure 5D), the effect size of the 'Silence' testing period was subtracted from the effect size of the 'Playback' testing period (preference scores > 0 indicate a preference for FD song). The median Preference Score from at least 5 females was calculated between subjects for each male.

As expected, females demonstrated a preference for FD song compared to UD (Preference Score > 0; $p = 0.0006$, two-tailed one-sample t-test, $n = 10$ male birds). Overall, we found that while females still preferred FD song to UD song sung by AAV-FoxP2-injected males, their preference for those songs was diminished relative to songs sung prior to AAV injection ($p = 0.051$, one-tailed paired t-test, $n = 5$ male birds). The preference for FD song following AAV-GFP surgery was unchanged ($p = 0.182$, one-tailed paired t-test, $n = 5$ male birds).

Discussion

The transcription factor FoxP2 is critical to the proper development of learned vocalizations used for social communication in both humans and zebra finch songbirds. Here, we provide novel evidence that the maintenance of learned vocalizations in adulthood relies upon auditory-dependent regulation of striatopallidal FoxP2. In juvenile finches, the shared behavioral outcomes that follow FoxP2 overexpression or knockdown suggest that song learning is

456 dependent upon behavior-driven regulation of FoxP2 in the striatopallidal song-dedicated
457 nucleus Area X; having too much, or too little, results in similar deficits (Haesler et al., 2007;
458 Heston and White, 2015; Burkett et al., 2018). Behavior-driven FoxP2 regulation also occurs in
459 adults (Teramitsu and White, 2006; Miller et al., 2008; Hilliard et al., 2012a; Shi et al., 2013;
460 Thompson et al., 2013), which motivated us to test for a possible role forFoxP2in the
461 maintenance of learned vocalizations. We confirmed that, in hearing birds, Area X FoxP2 levels
462 affect the precision of courtship song (Murugan et al., 2013). Going further, our data suggest that
463 the auditory feedback required to maintain adult song may do so, in part, through regulation of
464 Area X FoxP2 levels. Together, these findings indicate that appropriate behavioral regulation of
465 FoxP2 is not only critical for juveniles who are in the process of song learning, but also for adult
466 animals who require ongoing auditory feedback to properly maintain their song.

467 An experimental strength offered by adult zebra finch song is its robustness,
468 characterized by marked stability across song renditions throughout the lifespan. This provides
469 an easily quantifiable behavior for assessing the effects of mechanistic interventions. Such
470 behavioral stability may reflect a fixed nature of its biological underpinnings. Indeed, a historical
471 assumption was that AFP song control nuclei were unnecessary for adult song maintenance since
472 limited-to-no changes in song were detected following lesions of these areas in adults. This was
473 in marked contrast to the profound effects on learning observed after lesioning these regions in
474 juveniles or the dramatic loss in learned vocal output that follows lesions of nuclei in the vocal
475 motor pathway at any age (Bottjer et al., 1984; Scharff and Nottebohm, 1991).

476 Subsequent landmark experiments unveiled an ongoing role for the AFP in adult song
477 maintenance by combining two interventions, i.e. by assessing changes to song following both

lesioning and deafening (Brainard and Doupe, 2000). In birds with an intact AFP, deafening resulted in song degradation, as previously shown (Nordeen and Nordeen, 1992; Brainard and Doupe, 2001; Horita et al., 2008; Nordeen and Nordeen, 2010). Strikingly, lesions of the AFP prevented deafening-induced song deterioration (Brainard and Doupe, 2000; Kojima et al., 2013). Thus, this 'double-insult' methodology unveiled the normal role of the AFP in song maintenance by actively generating vocal variability in adults (Woolley and Kao, 2015). By analogy, here we tested the role of FoxP2 in adult maintenance by introducing a genetic 'lesion', i.e. by blocking natural behavior-linked FoxP2 cycling in Area X through viral-driven overexpression. Similar to lesions of the AFP, we detected fairly subtle effects of our genetic insult in hearing birds, consistent with the robust stability of adult song. Likewise, disruptions to cortico-striatal circuits in humans and rodent models induce more prominent deficits during learning than during execution of well-learned skills (Graybiel, 2008; Kawai et al., 2015). In striking contrast, when the genetic insult was paired with deafening, it accelerated song decrystallization, revealing a role for behaviorally-regulated FoxP2 expression in ongoing song maintenance.

It is important to note that overexpression of FoxP2 does not simply recapitulate the effect of lesioning Area X in adult finches. While both chemical and genetic insults to Area X result in few changes in the songs of hearing birds, the experimental outcomes diverge in deafened animals. Chemical lesions of Area X prevented deafening-induced song deterioration (Kojima et al., 2013) whereas our genetic manipulation accelerated song degradation. These results extend our prior observation that hearing regulates Area X FoxP2 expression during sensorimotor learning (Teramitsu et al., 2010). In both deafened and hearing juvenile finches,

FoxP2 was downregulated following two hours of undirected singing, indicating that FoxP2 expression is primarily regulated by motor activity. However, FoxP2 expression and amount of singing were not correlated in deafened juveniles as they were in hearing juveniles. This suggests that while motor behavior is sufficient to decrease Area X FoxP2 levels, auditory feedback is necessary to properly calibrate its expression. Additionally, a notable trend in the deafened-FoxP2 injected animals was an increase in FoxP2 expression relative to other groups, despite singing similar amounts of song prior to sacrifice (Figure 1F). This suggests that the lack of auditory feedback was insufficient to proportionally lower FoxP2 as observed in the FoxP2-hearing animals. Molecular regulators of FoxP2 such as POU3F2 (Atkinson et al., 2018), miR-9 and miR-140-5p (Shi et al., 2013) have been identified. Thus, it will be important to determine how sensory feedback affects the regulation of these molecules and, in turn, FoxP2 in the coordination of complex motor tasks.

The hastening of deafening-induced song deterioration and increase in phonological and sequencing variability following FoxP2 overexpression, suggests that 1) auditory feedback is critical for the proper function of FoxP2 to precisely control mature vocalizations and 2) dysregulated FoxP2 increases song variability. Indeed, similar to knockdown of FoxP2 in Area X of adult zebra finches (Murugan et al., 2013), we observed an increase in the acoustic variability of female-directed song, indicating that FoxP2 may mediate an adult's ability to generate appropriate behavioral responses to salient social cues. This is consistent with the result that either overexpression or knockdown of FoxP2 impairs song copying in juvenile finches (Haesler et al., 2007; Heston and White, 2015). Together, these convergent findings suggest that interfering either by overexpression or by knockdown of FoxP2 produces similar behavioral

522 outcomes in adults, as in juveniles. Our data also strengthen a model in which self-regulation of
523 FoxP2 by sensory and motor cues enable song variability that is necessary for ongoing
524 refinement of learned vocalizations.

525 Social-context driven changes to song variability have been associated with dopamine
526 modulation in Area X (Sasaki et al., 2006; Leblois et al., 2010; Leblois and Perkel, 2012;
527 Murugan et al., 2013). In particular, the marked stability of female-directed song depends on
528 activation of D1 receptors (Leblois et al., 2010). We found that *FoxP2* expression positively
529 correlates with *D1R* expression (Figure 1G) and increases the rendition-to-rendition variability of
530 the fundamental frequency of syllables containing harmonic stacks. In our study, dopamine
531 receptor transcript levels were assessed prior to the onset of singing and in the absence of any
532 females. Thus, changes in dopamine marker levels may not correlate with physiological changes
533 that occur when birds are actively singing or when in the presence of females. This difference in
534 experimental protocol may account for our findings relative to previous reports that show low
535 acoustic variability following D1R receptor antagonism in Area X (Leblois et al., 2010).

536 The mechanisms that reinforce optimal motor patterns within cortico-basal ganglia
537 circuits increasingly implicate a critical role for dopamine (Schultz et al., 1997; Graybiel, 2008;
538 Murugan et al., 2013; Gadagkar et al., 2016; Hoffmann et al., 2016; Xiao et al., 2018). FoxP2 is
539 linked to intracellular dopaminergic signaling to influence vocal variability (Vernes et al., 2011;
540 Murugan et al., 2013), but it remains untested as to whether mechanisms that alter signal
541 propagation in the AFP following FoxP2 knockdown are the same as those that may accompany
542 FoxP2 overexpression. Elucidating the interaction between FoxP2 and dopaminergic signaling,
543 particularly given that the ventral tegmental area (VTA) receives afferents from multiple
544 auditory regions (Mandelblat-Cerf et al., 2014) and dopamine signaling encodes performance

errors during singing (Gadagkar et al., 2016), will be essential in understanding its ongoing role in song maintenance and modulation during social communication. Additional experiments will also be necessary to determine whether afferents from HVC, a critical conveyer of auditory input to the AFP (Roy and Mooney, 2007; Gale and Perkel, 2010), calibrate expression of Area X FoxP2 despite evidence that HVC does not transmit error-related signals (Hessler and Doupe, 1999; Kozhevnikov and Fee, 2007) or receive auditory signals (Hamaguchi et al., 2014) during singing.

Our study provides insight into how FoxP2 may influence social communication between conspecifics and identifies FoxP2 as necessary for the execution of precise motor behaviors. We used females to demonstrate that the increase in vocal variability following FoxP2 overexpression has functional consequences. Females prefer stereotyped song with low rendition-to-rendition variability (Woolley and Doupe, 2008; Dunning et al., 2014; Chen et al., 2017). The decrease in female preference for FD song following FoxP2 overexpression is consistent with the observed increase in vocal variability in those songs. Using females to identify whether experimenter-induced changes to male song promote or impede song quality can thus tease out ethologically-relevant manipulations to song.

Within neural circuits that control behavior, the FoxP2 transcription factor can coordinate the activation or repression of hundreds to thousands of genes, affecting a variety of molecular mechanisms (Vernes et al., 2011; Hilliard et al., 2012a; Chen et al., 2016). Gene co-expression patterns within Area X shift across the critical period from song learning to song maintenance (Burkett et al., 2018), suggesting that individual genes, including FoxP2, can differentially contribute to a variety of behaviors, including both learning in juveniles and maintenance in adults. Although no differences in gene expression of transcription factors have been identified

568 in the cortical song motor pathway following deafening (Mori and Wada, 2015), we predict that
569 auditory deprivation will influence gene expression patterns in the avian striatum. Thus, in the
570 future, it will be necessary to identify how FoxP2 overexpression in the presence or absence of
571 auditory feedback alters gene co-expression. Such experiments may illuminate how FoxP2
572 orchestrates the molecular microcircuitry necessary for song maintenance, and, by extension,
573 human speech.
574
575 **Figure Legends**
576
577 **Figure 1. Adeno-associated viral construct drives overexpression of FoxP2 in adult male zebra**
578 **finches.** A) Timeline of experimental manipulations. The song-dedicated striatal nucleus Area X was
579 bilaterally injected with an adeno-associated viral (AAV) construct ('Surgery A') to drive overexpression
580 of GFP (control) or FoxP2. To remove auditory feedback in half of the birds, 'Surgery B' was performed
581 ~20 days following 'Surgery A'. Songs were analyzed (vertical lines) at two time points directly prior to
582 each surgery, and at 6, 14, 25, 60 days post deafening (e.g. D06, D14, etc.), and on the morning of
583 sacrifice (DOS). B) Schematic of the AAV construct used to drive expression either GFP or FoxP2 using
584 the CAG promoter. C) Protein levels of FoxP2 appear higher in hemispheres injected with FoxP2
585 compared to hemispheres injected with GFP in the same bird. D) In hearing birds used for evaluating the
586 time-line of *FoxP2* over-expression, RT-qPCR confirms augmented levels at 20, 35, 45, and 80 days post-
587 injection (Equivalent to 'Surgery A' time point in panel A) relative to uninjected controls (U). Fold
588 change values are normalized to the mean of the controls. E) Across all birds used for behavioral analysis,
589 *FoxP2* expression levels (delta Ct; mean ± SEM) are higher in FoxP2- vs. GFP-injected birds ($p = 0.042$),
590 on the morning of sacrifice (DOS) ~6 months after 'Surgery A'. F) Delta Ct values of *FoxP2* levels on
591 DOS and time spent singing for each group shows that FoxP2-Deaf birds trend toward higher FoxP2

592 expression despite singing similar amounts as other groups (mean ± SEM). G) Delta delta CT values of
593 FoxP2 levels positively correlate with dopamine markers D1R and TH, but not with D2R (H). In D-H,
594 dots represent individual birds. *Figure Contributions:* Jon Heston identified the appropriate viral
595 construct; Nancy Day performed the experiments and analyzed the data.
596
597 **Figure 2. In hearing adults, Area X FoxP2 overexpression does not alter undirected songs relative**
598 **to those of GFP control birds. A)** Representative spectrograms of song bouts from two zebra finches
599 before and after injection of an AAV that drives expression of a control GFP or FoxP2 construct. Scale
600 bars = 500ms. B) Mean syllable identity is unchanged following FoxP2 overexpression. C) Following
601 surgery, undirected songs were similar to pre-surgery songs (PRE) except at the final 6 month time point
602 (DOS) for both GFP-injected and FoxP2-injected birds. * $p < 0.05$. C) Syllable sequence (syntax
603 similarity) is not altered following injection of AAV-GFP or AAV-FoxP2. *Figure Contributions:* Nancy
604 Day performed the experiments and analyzed the data.
605
606
607 **Figure 3. FoxP2 overexpression hastens deafening-induced song deterioration.** A) Representative
608 spectrograms show deafening-induced song deterioration in two brothers who were deafened at 180 dph
609 (Surgery B) 29 days after injection of AAV-GFP (left) or AAV-FoxP2 (right) (Surgery A, Figure 1).
610 Scale bar = 500ms. B) At 14 days post-deafening, motif similarity is persistently altered in the AAV-
611 FoxP2 group (two-way ANOVA with Sidak's multiple comparisons, p values = 0.0006, 0.0002, < 0.0001,
612 and <0.0001 at D14, 25, 60, and DOS respectively, n = 6 birds). In comparison to AAV-GFP-deafened
613 birds, degradation of songs by AAV-FoxP2-injected birds is accelerated by at least 10 days. Statistically
614 significant changes to songs by AAV-GFP-deafened birds are present at DOS (two-way ANOVA with
615 Sidak's multiple comparisons, p = 0.0007, n = 5 birds). All motif similarity scores are normalized to motif
616 similarity calculated between songs collected on two days prior to AAV injection (refer to Figure 1A).

(*** p< 0.001, **** p<0.0001.) *Figure Contributions:* Nancy Day performed the experiments and analyzed the data.

Figure 4. Spectral variability and sequencing are affected by FoxP2 overexpression in deaf birds. A) Vocal variability increased more rapidly in deafened AAV-FoxP2 birds (solid black bars) than in deafened AAV-GFP birds (solid green bars) in most song features analyzed (e.g. Entropy, Entropy Variance, Duration, Pitch Goodness, FM). Positive values indicate an increase in the CV of each feature relative to PRE; negative values reflect lower variability than observed in PRE. B) Syllable omission occurs faster in FoxP2-deafened animals than in GFP-deafened animals. C) Syntax similarity (syllable sequencing; normalized to PRE) is disrupted in FoxP2-deafened animals by 14 days following deafening. (* p<0.05, ** p<0.01, *** p < 0.001, **** p<0.0001.) *Figure Contributions:* Nancy Day performed the experiments and analyzed the data.

Figure 5. Female conspecifics perceive alterations in social context-dependent song variability. A) Exemplar syllable with a harmonic element/stack. Only the 'flat' component of the syllable (indicated by dotted lines) was analyzed to determine the coefficient of variation (CV) of the fundamental frequency (FF). B) Prior to AAV injections, syllables are performed with less rendition-to-rendition variability during female-directed song compared to undirected song in both GFP and FoxP2 groups (Wilcoxon matched-pairs signed rank test, one-tailed; AAV-GFP: p = 0.0002, n = 12 syllables, AAV-FoxP2: p = 0.0001, n = 13 syllables). Following AAV injection, the CV of FF is significantly lower for female-directed syllables in GFP-injected zebra finches (Wilcoxon matched-pairs signed rank test, one-tailed, p = 0.017, n = 12 syllables), but not in FoxP2-injected animals (Wilcoxon matched-pairs signed rank test, one-tailed, p = 0.064, n = 13 syllables). (* p<0.05, *** p<0.001) C) 'Bird's eye view' schematic of the testing arena for assaying female preference. D) Female preference for FD song is reduced ($Preference_{POST} - Preference_{PRE}$) following FoxP2 overexpression compared to songs following GFP

overexpression (two-tailed t-test, $p = 0.047$, $t = 2.34$, $df = 8$, $n = 5$ male birds per group). *Figure Contributions:* Taylor Hobbs designed the female preference testing area. Taylor Hobbs and Nancy Day performed the female preference experiments and analyzed the data.

Literature Cited

Andalman AS, Fee MS (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci USA 106:12518–12523.

Atkinson EG, Audesse AJ, Palacios JA, Bobo DM, Webb AE, Ramachandran S, Henn BM (2018) No Evidence for Recent Selection at FOXP2 among Diverse Human Populations. Cell 174:1424–1435.e15.

Bolhuis JJ, Okanoya K, Scharff C (2010) Twitter evolution: converging mechanisms in birdsong and human speech. Nat Rev Neurosci 11:747–759.

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224:901–903.

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404:762–766.

Brainard MS, Doupe AJ (2001) Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. 21:2501–2517.

Brainard MS, Doupe AJ (2013) Translating Birdsong: Songbirds as a Model for Basic and Applied Medical Research. Annu Rev Neurosci 36:489–517.

Burkett ZD, Day NF, Kimball TH, Aamodt CM, Heston JB, Hilliard AT, Xiao X, White SA (2018) FoxP2 isoforms delineate spatiotemporal transcriptional networks for vocal learning in the zebra finch. eLife 7:e46610.

Burkett ZD, Day NF, Peñagarikano O, Geschwind DH, White SA (2015) VoICE: A semi-automated pipeline for standardizing vocal analysis across models. Sci Rep 5:10237.

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.

Chen Y, Clark O, Woolley SC (2017) Courtship song preferences in female zebra finches are shaped by developmental auditory experience. Proc Biol Sci 284:20170054.

672 Chen Y-C, Kuo H-Y, Bornschein U, Takahashi H, Chen S-Y, Lu K-M, Yang H-Y, Chen G-M,
673 Lin J-R, Lee Y-H, Chou Y-C, Cheng S-J, Chien C-T, Enard W, Hevers W, Pääbo S,
674 Graybiel AM, Liu F-C (2016) Foxp2 controls synaptic wiring of corticostriatal circuits and
675 vocal communication by opposing Mef2c. Nat Neurosci 19:1513–1522.

676 Cowie R, Douglas-Cowie E, Kerr AG (1982) A study of speech deterioration in post-lingually
677 deafened adults. J Laryngol Otol 96:101–112.

678 Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms.
679 Annu Rev Neurosci 22:567–631.

680 Dunning JL, Pant S, Bass A, Coburn Z, Prather JF (2014) Mate choice in adult female Bengalese
681 finches: females express consistent preferences for individual males and prefer female-
682 directed song performances. PLoS ONE 9:e89438.

683 Farries MA, Perkel DJ (2002) A telencephalic nucleus essential for song learning contains
684 neurons with physiological characteristics of both striatum and globus pallidus. J Neurosci
685 22:3776–3787.

686 Fong WL, Kuo H-Y, Wu H-L, Chen S-Y, Liu F-C (2018) Differential and Overlapping Pattern
687 of Foxp1 and Foxp2 Expression in the Striatum of Adult Mouse Brain. Neuroscience
688 388:214–223.

689 Gadagkar V, Puzerey PA, Chen R, Baird-Daniel E, Farhang AR, Goldberg JH (2016) Dopamine
690 neurons encode performance error in singing birds. Science 354:1278–1282.

691 Gale SD, Perkel DJ (2010) A basal ganglia pathway drives selective auditory responses in
692 songbird dopaminergic neurons via disinhibition. J Neurosci 30:1027–1037.

693 Graybiel AM (2008) Habits, rituals, and the evaluative brain. Annu Rev Neurosci 31:359–387.

694 Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C (2007) Incomplete and
695 inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area
696 X. Mooney R, ed. PLoS Biol 5:e321.

697 Hamaguchi K, Tschida KA, Yoon I, Donald BR, Mooney R (2014) Auditory synapses to song
698 premotor neurons are gated off during vocalization in zebra finches. eLife 3:e01833.

699 Hessler NA, Doupe AJ (1999) Singing-related neural activity in a dorsal forebrain-basal ganglia
700 circuit of adult zebra finches. Journal of Neuroscience 19:10461–10481.

701 Heston JB, Simon J, Day NF, Coleman MJ, White SA (2018) Bidirectional scaling of vocal
702 variability by an avian cortico-basal ganglia circuit. Physiol Rep 6:e13638.

703 Heston JB, White SA (2015) Behavior-linked FoxP2 regulation enables zebra finch vocal
704 learning. 35:2885–2894.

705 Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA (2012a) Molecular microcircuitry
706 underlies functional specification in a basal ganglia circuit dedicated to vocal learning.
707 Neuron 73:537–552.

708 Hilliard AT, Miller JE, Horvath S, White SA (2012b) Distinct neurogenomic states in basal
709 ganglia subregions relate differently to singing behavior in songbirds. Behrens T, ed. PLoS
710 Comput Biol 8:e1002773.

711 Hoffmann LA, Saravanan V, Wood AN, He L, Sober SJ (2016) Dopaminergic Contributions to
712 Vocal Learning. 36:2176–2189.

713 Horita H, Wada K, Jarvis ED (2008) Early onset of deafening-induced song deterioration and
714 differential requirements of the pallial-basal ganglia vocal pathway. Eur J Neurosci 28:2519–
715 2532.

716 Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-
717 dependent changes to song variability. J Neurophysiol 96:1441–1455.

718 Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain
719 circuit to real-time modulation of song. Nature 433:638–643.

720 Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, Kampff AR, Ölveczky BP
721 (2015) Motor cortex is required for learning but not for executing a motor skill. Neuron
722 86:800–812.

723 Kojima S, Kao MH, Doupe AJ (2013) Task-related "cortical" bursting depends critically on
724 basal ganglia input and is linked to vocal plasticity. Proc Natl Acad Sci USA 110:4756–
725 4761.

726 Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-
727 crowned sparrow. Z Tierpsychol 22:770–783.

728 Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the
729 zebra finch. J Neurophysiol 97:4271–4283.

730 Kubikova L, Bosikova E, Cvikova M, Lukacova K, Scharff C, Jarvis ED (2014) Basal ganglia
731 function, stuttering, sequencing, and repair in adult songbirds. Sci Rep 4:6590.

732 Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is
733 mutated in a severe speech and language disorder. Nature 413:519–523.

734 Leblois A, Perkel DJ (2012) Striatal dopamine modulates song spectral but not temporal features
735 through D1 receptors. Eur J Neurosci 35:1771–1781.

736 Leblois A, Wendel BJ, Perkel DJ (2010) Striatal dopamine modulates basal ganglia output and
737 regulates social context-dependent behavioral variability through D1 receptors. J Neurosci
738 30:5730–5743.

739 Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time
740 quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408.

741 Mandelblat-Cerf Y, Fee MS (2014) An automated procedure for evaluating song imitation.
742 Bolhuis JJ, ed. PLoS ONE 9:e96484.

743 Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS (2014) A role for descending auditory cortical
744 projections in songbird vocal learning. eLife 3:e02152–e02152.

745 Miller JE, Hilliard AT, White SA (2010) Song practice promotes acute vocal variability at a key
746 stage of sensorimotor learning. PLoS ONE 5:e8592.

747 Miller JE, Spiteri E, Condro MC, Dosumu-Johnson RT, Geschwind DH, White SA (2008)
748 Birdsong decreases protein levels of FoxP2, a molecule required for human speech. J
749 Neurophysiol 100:2015–2025.

750 Mori C, Wada K (2015) Audition-Independent Vocal Crystallization Associated with Intrinsic
751 Developmental Gene Expression Dynamics. Journal of Neuroscience 35:878–889.

752 Murugan M, Harward S, Scharff C, Mooney R (2013) Diminished FoxP2 levels affect
753 dopaminergic modulation of corticostriatal signaling important to song variability. Neuron
754 80:1464–1476.

755 Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of
756 stereotyped song in adult zebra finches. Behavioral and Neural Biology 57:58–66.

757 Nordeen KW, Nordeen EJ (2010) Deafening-induced vocal deterioration in adult songbirds is
758 reversed by disrupting a basal ganglia-forebrain circuit. 30:7392–7400.

759 Roy A, Mooney R (2007) Auditory plasticity in a basal ganglia-forebrain pathway during
760 decrystallization of adult birdsong. J Neurosci 27:6374–6387.

761 Sasaki A, Sotnikova TD, Gainetdinov RR, Jarvis ED (2006) Social context-dependent singing-
762 regulated dopamine. J Neurosci 26:9010–9014.

763 Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions
764 of various parts of the zebra finch song system: implications for vocal learning. Journal of
765 Neuroscience 11:2896–2913.

766 Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science
767 275:1593–1599.

768 Shi Z, Luo G, Fu L, Fang Z, Wang X, Li X (2013) miR-9 and miR-140-5p target FoxP2 and are
769 regulated as a function of the social context of singing behavior in zebra finches. J Neurosci
770 33:16510–16521.

771 Sober SJ, Brainard MS (2009) Adult birdsong is actively maintained by error correction. Nat
772 Neurosci 12:927–931.

773 Tchernichovski O, Nottebohm F, Ho C, Pesaran B, Mitra P (2000) A procedure for an automated
774 measurement of song similarity. Anim Behav 59:1167–1176.

775 Teramitsu I, Poopatanapong A, Torrisi S, White SA (2010) Striatal FoxP2 is actively regulated
776 during songbird sensorimotor learning. Tanimoto H, ed. PLoS ONE 5:e8548.

777 Teramitsu I, White SA (2006) FoxP2 Regulation during Undirected Singing in Adult Songbirds.
778 Journal of Neuroscience 26:7390–7394.

779 Thompson CK, Schwabe F, Schoof A, Mendoza E, Gampe J, Rochefort C, Scharff C (2013)
780 Young and intense: FoxP2 immunoreactivity in Area X varies with age, song stereotypy, and
781 singing in male zebra finches. Front Neural Circuits 7:24.

782 Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of
783 "crystallized" adult birdsong. Nature 450:1240–1244.

784 Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM, Ho J, Mombereau C,
785 Brewer A, Lowy E, Nicod J, Groszer M, Baban D, Sahgal N, Cazier J-B, Ragoussis J,
786 Davies KE, Geschwind DH, Fisher SE (2011) Foxp2 regulates gene networks implicated in
787 neurite outgrowth in the developing brain. Akey JM, ed. PLoS Genet 7:e1002145.

788 Watkins KE, Vargha-Khadem F, Ashburner J, Passingham RE, Connelly A, Friston KJ,
789 Frackowiak RSJ, Mishkin M, Gadian DG (2002) MRI analysis of an inherited speech and
790 language disorder: structural brain abnormalities. Brain 125:465–478.

791 White SA (2001) Learning to communicate. Curr Opin Neurobiol 11:510–520.

792 Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior
793 and gene expression. Brenowitz E, ed. PLoS Biol 6:e62.

794 Woolley SC, Kao MH (2015) Variability in action: Contributions of a songbird cortical-basal
795 ganglia circuit to vocal motor learning and control. Neuroscience 296:39–47.

796 Woolley SM, Rubel EW (1997) Bengalese finches Lonchura Striata domestica depend upon
797 auditory feedback for the maintenance of adult song. Journal of Neuroscience 17:6380–
798 6390.

799 Xiao L, Chattree G, Oscos FG, Cao M, Wanat MJ, Roberts TF (2018) A Basal Ganglia Circuit
800 Sufficient to Guide Birdsong Learning. Neuron 98:208–221.e5.