UCLA
UCLA Electronic Theses and Dissertations

Title
FoxP2 and Basal Ganglia Function in Zebra Finch Vocal Motor Learning and Control

Permalink
https://escholarship.org/uc/item/4km6n3ph

Author
Heston, Jon

Publication Date
2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

FoxP2 and Basal Ganglia Function in Zebra Finch Vocal Motor Learning and Control

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Neuroscience

by

Jonathan Heston

2016

© Copyright by

Jonathan Heston

2016

ABSTRACT OF DISSERTATION

FoxP2 and Cortico-Basal Ganglia Function in Zebra Finch Vocal Motor Learning and Control

by

Jonathan B. Heston

Doctor of Philosophy in Neuroscience

University of California, Los Angeles, 20126

Professor Stephanie Ann White, Chair

The capacity for language is one of the most complex and least understood phenomena in neuroscience. Because language is unique to humans which are not amenable to invasive or interventionist experiments, research has focused on studying the neural basis of language subcomponents in other species. One such subcomponent is vocal learning, defined as the ability to acquire new sounds via imitation, and is present in a handful of animal groups including songbirds. The behavioral, anatomical and molecular parallels between human speech learning and zebra finch (*Taeniopygia guttata*) song learning make songbirds such as the zebra finch a compelling model system for studying learned vocalizations.

My dissertation asks related questions of the molecular and physiological basis of vocal learning through the lenses of the transcription factor FoxP2 and cortico-basal ganglia function, respectively. Mutations in the transcription factor Forkhead box protein 2 (FoxP2) are associated with a specific language disorder in humans. In both humans and songbirds FoxP2 is enriched in the basal ganglia and, in zebra finch, FoxP2 levels are decreased as a function of vocal practice

within the song-dedicated nucleus known as Area X. This downregulation is accompanied by an acute increase in vocal variability as well as the bidirectional regulation of thousands of genes many of which are known targets of FoxP2 in humans. Here, I test whether online regulation of FoxP2 is necessary for vocal learning and whether FoxP2 downregulation is causally related to the enhanced variability following vocal practice. To this end, I pioneered the use of AAV driven overexpression of a biologically relevant gene- a first in the zebra finch- and found that preventing FoxP2 downregulation led to poor song learning and disrupted the acute increase in vocal variability. Based on these findings I suggest that dynamic behavior-linked regulation of FoxP2, rather than absolute levels per se, is critical for vocal learning.

Next, to gain insight into the physiological changes that accompanied the FoxP2 regulation, I asked what physiological manipulations could recapitulate the FoxP2 dependent transition to high variability. To that end, I pioneered the use of designer drugs exclusively activated by designer receptors (DREADDs) -another first in the songbird system- to bidirectionally affect cell excitability. Because of its known role as positive regulator of variability I also focused my attention on the cortical song control nucleus lateral magnocellular nucleus of the nidopallium (LMAN). I found that bidirectional manipulations of either LMAN or basal ganglia Area X lead to bidirectional regulation of vocal variability with higher levels of LMAN activity leading to higher levels of variability whereas higher levels of Area X led to lower levels of variability. These effects on variability where limited to rendition-to-rendition and millisecond timescale moment-to-moment variability but not higher order syntax variability. Based on these observations I suggest that the cortex and striatopallidum reciprocally control vocal variability and that FoxP2-dependent decrease in neuronal excitability may account for practice dependent increases in variability.

The dissertation of Jonathan Heston is approved.

Walter Helmut Metzner

Bennett G. Novitch

Thomas J. O'Dell

Stephanie Ann White, Committee Chair

University of California, Los Angeles

2016

Table of Contents

List of Figures

Chapter 2 is a version of Heston, Jonathan B., and Stephanie A. White. "Behavior-Linked FoxP2 Regulation Enables Zebra Finch Vocal Learning." *The Journal of Neuroscience*35.7 (2015): 2885-2894. This work was supported by Grants T32MH019384, T32NS058280, UCLA Edith Hyde Fellowship (J.B.H.), RO1MH070712, and Tennenbaum Center for the Biology of Creativity (S.A.W.). I 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.

Chapter 3 is the most current version of manuscript in preparation for submission.

Appendix I is a version of Chen, Q., Heston, J.B., Burkett, Z.D., White, S.A. (2013). Expression analysis of the speech-related genes FoxP1 and FoxP2 and their relation to singing behavior in two songbird species. Journal of experimental biology, 216(19), 3682-3692. Thanks to Dr. Julie E. Miller for constructive comments on the manuscript. We also thank Lily Sung for assistance in preparation of brain sections.

Appendix II is a version of Fuxjager, Matthew J., Jonathan B. Heston, and Barney A. Schlinger. "Peripheral androgen action helps modulate vocal production in a suboscine passerine." *The Auk* 131.3 (2014): 327-334.

VITA

2008	Bachelors of Science, Neurobiology and Behavior
	Minor in Anthropology
	University of California, Irvine
2008-2009	Research Assistant
	University of California, Irvine
	Department of Pediatrics
2011-2012	Molecular and Cellular Neurobiology Training Grant
	University of California, Los Angeles
2012-2013	Edith Hyde Graduate Fellowship
	Department of Integrative Biology and Physiology
	University of California, Los Angeles
2014-2015	Neural Microcircuits Training Grant
	University of California, Los Angeles

Publications and Presentations

Villégier AS, Gallager B, **Heston J**, Belluzzi JD, Leslie FM. (2010) Age influences the effects of nicotine and monoamine oxidase inhibition on mood-related behaviors in rats. Psychopharmacology 208(4): 593-601

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(19): 3682-92.

Fuxjager MJ, **Heston JB**, and Schlinger BA (2014). Peripheral androgen action helps modulate vocal production in a suboscine passerine. *The Auk.* 131(3):327-334

Heston JB & White SA (2015). Behavior-Linked FoxP2 Regulation Enables Zebra Finch Vocal Learning. J. Neurosci. 35(7):2885–2894

Heston JB, Simon J, Day, NF, Coleman MJ**,** White SA. Reciprocal Scaling of Vocal Variability by an Avian Cortico-Basal Ganglia Circuit (in submission)

Day NF, Kim CY, **Heston JB**, White SA. Beyond sensorimotor learning: FoxP2 overexpression affects the maintenance of learned vocalization in adult zebra finch. (in preparation)

Burkett ZD, **Heston JB**, White SA. Foxp2 isoform specific regulation of zebra finch vocal learning (manuscript in preparation)

Chapter I: Introduction

Songbirds as a model system for understanding human speech learning

As a social species the ability to communicate is critical for our well-being. Indeed, language and communication deficits are one of the core symptoms of autism spectrum disorder and any treatment of this disease will likely rely on an understanding of the neurobiological basis of language. While language is a complex behavior unique to humans, this behavior can be broken down into subcomponents that are present and can be studied in other species. One such sub-component is vocal learning, the ability to produce novel sounds through imitation. While this behavior is relative rare- humans are the only primate species with this ability (Knornschild et al., 2010; Stoeger et al., 2012; Fitch, 2012)- it is present in a handful of animal groups including elephants, whales, cetaceans such as dolphins and whales, and several orders of birds that include hummingbirds, parrots, and songbirds. Of these groups, birdsong displays several parallels with human speech that provide validity to their use as model system for understanding human vocal learning. These parallels can broadly be divided into behavioral, anatomical, and genetic parallels.

Behavioral parallels

The behavioral parallels between human speech learning and zebra finch song learning are evident in a shared pattern of critical periods (Doupe and Kuhl, 1999). Both human speech learning and song learning can be divided into two learning periods: sensory and sensorimotor learning. In humans, babies undergo an early sensory learning period during which the baby is listening to and forming a sensory representation of the language specific set of phonemes that he or she will eventually imitate. In songbirds, this early phase of song learning is referred to as sensory acquisition and in a normally reared zebra finch begins at ~20 days (d) and continues

until ~65 d. During this period the young zebra finch male is listening to the song of its tutor and building a sensory representation of the song that it will eventually imitate.

These early sensory periods are followed by a sensorimotor period during which both the baby and the young zebra finch begin making vocal utterances, which are distinct from innate crying or begging calls present in both sexes, and, through a process of trial and error, refine those utterances to match the template acquired during the sensory learning period. This period is characterized not only by a gradual increase in similarity to the tutor's song, but also a gradual stabilization or decrease in the variability of song. Like the sensory period, this phase of learning is critically dependent on auditory feedback (Iyengar and Bottjer, 2002).

In humans this early speech learning forms the basis for higher order forms of language learning that will continue for years to come. In contrast, song learning in zebra finch culminates at crystallization when the song learning is complete, the bird has reached sexual maturity, and the song remains essentially stable for the rest of the bird's life. This crystallization may seem to break from the parallels between speech and song as language is learned throughout life. It should be noted that the capacity for acquiring a new language without an accent diminishes after puberty. Moreover, in both humans and songbirds, maintenance of vocalization requires ongoing hearing suggesting that in both species there is some degree of residual plasticity.

Anatomical parallels

The zebra finch brain contains two pathways that coordinate learned vocalization. One is the posterior vocal motor pathway consisting of the cortical nuclei HVC (used as a proper name) and the robust nucleus of the arcopallium (RA), which sends projections to motor neurons in the dorsomedial part of the intercollicular nucleus and to the tracheosyringial part of the hypoglossal motor nucleus. This pathway is necessary for the execution of learned vocalizations and lesions

to either of these nuclei either abolishes or disrupts song (Leonardo and Fee, 2005; Simpson and Vicario, 1990; Yu and Margoliash, 1996).

The second pathway is the anterior forebrain pathway (AFP), a song-dedicated basal-ganglia-thalamo-cortical circuit, consisting of the basal ganglia nucleus Area X, the thalamic nucleus dorsolateralis anterior thalami pars medialis (DLM), and the cortical nucleus lateral magnocellular nucleus of the anterior nidopallium (LMAN) which sends bifurcating axons to both Area X as well as RA (Vates and Nottebohm, 1995) allowing this pathway to affect vocal output. The AFP serves important roles in both learning and modification of song (Scharff and Nottebohm; Kojima et al, 2013) but also in generating variable motor output both during learning and adulthood. Both of these roles are explored in this dissertation.

Importantly these circuits bear strong similarity to those controlling procedural learning, including vocal learning, in mammals as shown in Figure 1-2.

Neuromolecular parallels

The final series of parallel are the assortment of neuromolecular parallels that includes a shared reliance on FoxP2 expression, as described below. In addition to FoxP2 at least two other speech and language related genes, FoxP1 and CNTNAP2, both of which interact with FoxP2, have been shown to be enriched in song control regions (Teramitsu et al., 2004; Panaitof et al., 2010) and shown to be critical for normal song development (Q Chen, UCLA dissertation, 2015)

Songbirds as a model system for understanding FoxP2 function

The transcription factor FoxP2 has been linked to language after it was discovered that a point mutation causes a specific language impairment in a cohort known as the KE family. The

key impairment exhibited by effected family members to is verbal dyspraxia or difficulties with controlling the movement of orofacial muscles (Vargha-Khadem et al., 1995; Alcock et al., 2000a). These deficits are thought to be central in origin as orofacial movements are normal in non-verbal tasks such as chewing and laughing. Moreover, affected family members show deficits in non-verbal tests of linguistic fluency although these may be secondary to the verbal communication deficits.

Despite the high specificity of the impairment to language, FoxP2 is expressed in a surprisingly large number of tissues. For instance, in mice FoxP2 and its closest subfamily member FoxP1 play a role in the development of the heart and lungs. Within the central nervous system FoxP2 is expressed in a number of regions including the basal ganglia, the cortex, and the developing spinal cord (Teramitsu et al., 2004; Rousso, et al. 2012). Imaging studies, however, suggest a central role of the basal ganglia in mediating the language deficits as both functional and structural abnormalities in affected family members that appear to be largely limited to the caudate and putamen, a portion of the striatum that constitutes the input layer of the basal ganglia (Varga-Khadem et al., 1998; Liégeois et al., 2003).

One approach to understanding the deficits observed in the KE family has been to examine transgenic mice carrying the single amino acid substitution observed in the KE family, known as the R552H mutation. These mice show numerous abnormalities in basal ganglia function and basal ganglia-dependent behaviors. For example, they exhibit deficits in rotarod learning, a striatal-dependent motor learning task. In vivo recordings of these mice during learning reveals that they have elevated firing rates of medium spiny neurons in the dorsolateral striatum compared to wildtype littermates. Moreover, ex vivo slice recordings reveal deficits in striatal long-term depression (LTD), a cellular correlate of striatal-dependent motor learning,

while cerebellar LTD is intact (Groszer et al., 2008) despite FoxP2 being expressed in both regions.

In contrast, mice carrying a humanized version of FoxP2 that differs from the mouse version by three amino acids have been genetically engineered. These mice showed a phenotype essentially opposite that of R552H mice: enhanced rotarod learning and striatal LTD. Further, recent evidence suggests that FoxP2 plays a special role in altering the balance between goal-directed or associative learning and habitual learning, mediated by distinct striatal sub-compartments, with the humanized version of FoxP2 preferentially favoring the latter (Schreiweis et al., 2014). The results suggest the human version of FoxP2 has a role in supporting the kind of habitual processing or automaticity that could potentially underlie human language learning.

In sum these results all provide strong evidence that FoxP2 is playing a critical and specific role in enabling striatal-dependent motor learning and its cellular correlate striatal LTD. However, on trying to connect findings from mice to the vocal motor phenotype observed in humans researchers have focused on ultrasonic isolation calls made by young pups separated from their dams. While R552H and $FoxP2^{Hum}$ both show altered ultrasonic vocalizations, this behavior is innate, rather than learned, and bears more similarity to crying than it does any learned vocalization underlying human language. My lab and others, therefore, have taken advantage of the previously described parallels between zebra finch song learning and human vocal learning to study FoxP2 in the context of vocal learning.

In both humans and songbirds FoxP2 is enriched in the basal ganglia including Area X (Teramitsu et al., 2004). Interestingly, while FoxP1, 2 and 4 were thought to function primarily during development in the formation of tissues, FoxP2 is enriched in Area X and the outlying

striatopallidum throughout the life of the zebra finch (Haesler et al., 2004; Teramitsu et al., 2004). Knockdown of FoxP2 in Area X of juvenile males leads to inaccurate song learning indicating a post-developmental role for FoxP2 (Haesler et al., 2007; Murugan et al., 2012).

Adding complexity to this role, two hours 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; Shi 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). Moreover, this process is tightly regulated by microRNAs which are increased by undirected singing and, which in turn, downregulate FoxP2 (Shi et al., 2013). Together, these observations suggest that FoxP2 downregulation is a tightly regulated process that leads to molecular and functional consequences that could be important for learning.

To directly test this idea, I used viral driven overexpression to prevent FoxP2 downregulation. In Chapter II I describe efforts to characterize virus according to their suitability for this task. In Chapter III I use AAV driven overexpression to constitutively elevate FoxP2 at the onset of the sensorimotor period for song learning and show that this manipulation disrupts song learning and behavior-dependent changes in vocal variability (Heston and White, 2015). In the closing chapter I suggest a model of FoxP2 function and means for testing this model. In the Appendix I provide evidence that song-dependent changes FoxP2 expression and in vocal variability

Songbirds as a model system for understanding vocal variability

The results of Chapter II suggest that one function of FoxP2 downregulation is alter Area X function in a manner that promotes enhanced variability. To better understand the means by which this shift occurs I explore the neural mechanisms that control vocal motor variability. Motor variability is an essential component of any trial-and-error learning based model of motor learning. Indeed, a growing body of evidence suggests individual differences in the degree of motor variability positively predict the capacity for motor learning in both humans (Wu et al., 2014) and zebra finch (Sober and Brainard, 2012). In the vocal domain, motor variability has been hypothesized to be a necessary to overcoming foreign accents (Simmonds, 2015). While the circuits that evaluate and reinforce motor patterns are well established to reside in the dopaminergic midbrain and corticostriatal synapses, respectively, the neural mechanisms that generate variable motor output are less clear.

Zebra finch offer a premier model system for understanding the neural mechanism of motor variability. In addition to the AFP's previously described role in vocal learning and plasticity this circuit is critical for generating motor variability. Lesioning or silencing LMAN, the output nucleus of the AFP, leads to decreased vocal variability (Scharff and Nottebohm, 1991; Olveczky et al., 2005; Kao and Brainard, 2006; Hampton et al., 2009; Stepanek and Doupe, 2010). Conversely, electrical stimulation or pharmacologically exciting LMAN leads to an increase in vocal variability (Kao and Brainard, 2006). Moreover, LMAN activity is decreased during stereotyped directed singing (Hessler and Doupe, 1999; Kao et al., 2008). Together these suggest that LMAN is a positive regulator of vocal variability, but because LMAN is the output of a larger circuit it is unclear whether variability is generated in LMAN de novo or whether it is generated in conjunction with or even inherited from other circuit elements.

One other attractive possibility is that Area X, which receives monosynaptic input from and gives multisynaptic input to LMAN, is a critical node of the variability generating circuit. Indeed, studies in both mammals (Barnes et al., 2005; Sheth et al., 2011) and zebra finch (Woolley et al., 2014) have identified neural correlates of motor variability in the basal ganglia. Lesions studies, however, have produced surprisingly mixed results and thus are inconclusive. While lesions to Area X do eliminate burst firing in LMAN (Kojima et al., 2013) they either have no effect on variability (Fee and Goldberg, 2011) or lead to only a transient decrease in variability (Ali et al., 2015; Kojima et al., 2013). This is in contrast to studies that suggest that Area X plays a critical role in mediating acute social context (Murugan et al., 2013; Woolley et al., 2014) and practice-dependent (Heston and White, 2015) shifts in variability.

Taking a cue from these dynamic naturalistic shifts in behavior, Chapter IV reexamines the role of LMAN and Area X in variable motor output by acutely, bidirectionally, and independently manipulating neural activity in these two brain regions using DREADDS.

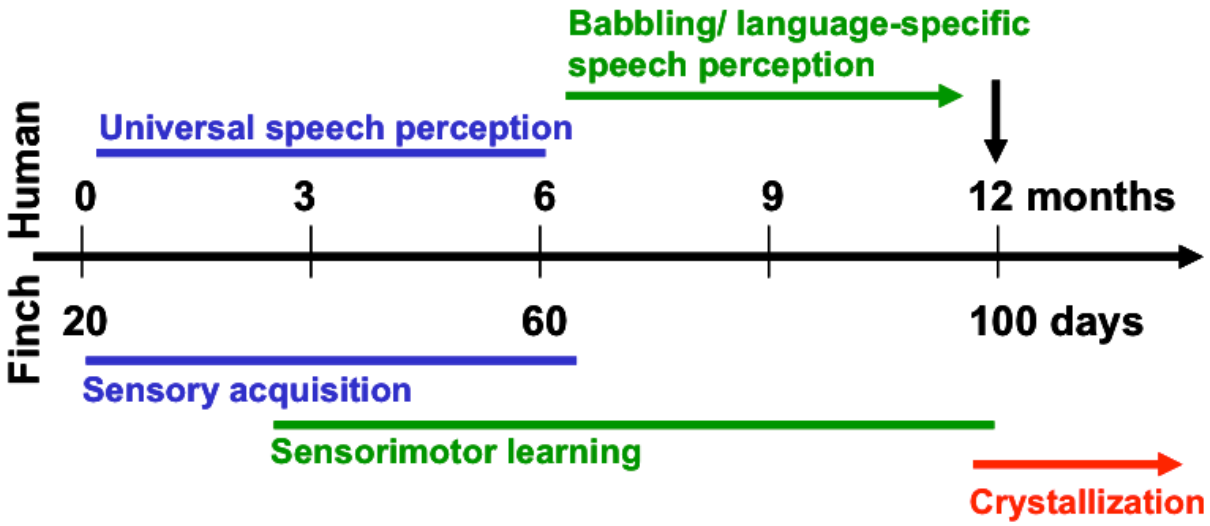


Figure 1-1. ***Timeline for human (top) and zebra finch (bottom) vocal development***. In both, learning starts with a perceptual phase (blue). Later, both begin vocal practice (green). In zebra finches, this sensorimotor phase culminates at ~100d, when the song 'crystallizes' (red). Adult song is actively maintained, similar to speech. Adapted from Doupe and Kuhl, 1999.

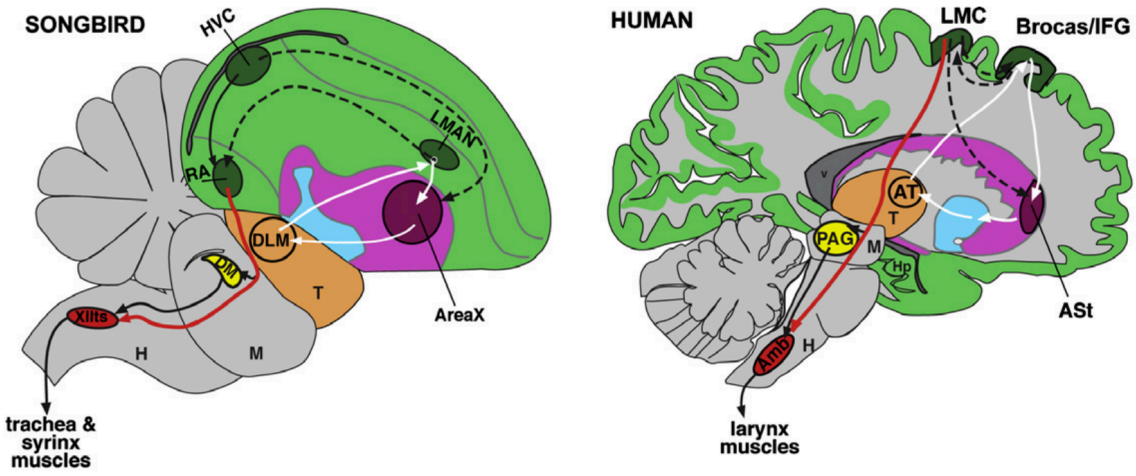


Figure 1-2. ***Comparison of avian song circuit (left) and human cortico-basal ganglia loops (right)***. Cortical regions are shown as green, basal ganglia regions as purple, and thalamic regions as orange.

Left) The zebra finch brain contains two pathways that coordinate learned vocalization. One is the posterior vocal motor pathway consisting of the cortical nuclei HVC (used as a proper name) and the robust nucleus of the arcopallium (RA), and brain stem motor neurons that control the syrinx (the avian equivalent of the larynx) and the diaphragm. This pathway is necessary for the execution of learned vocalizations (Leonardo and Fee, 2005; Simpson and Vicario, 1990; Yu and Margoliash, 1996). The second pathway is the anterior forebrain pathway (AFP), a song-dedicated basal-ganglia-thalamo-cortical circuit, consisting of the basal ganglia nucleus Area X, the thalamic nucleus DLM, and the cortical nucleus LMAN which sends bifurcating axons to both Area X as well as RA (Vates and Nottebohm, 1995) allowing this pathway to affect vocal output.

Right) Human vocalization involves a cortico-cortical connection from Broca's and inferior frontal gyrus (IFG) to the laryngeal motor cortex that sends a direct projection to motor neurons

in the nucleus ambiguous in the brainstem which control the larynx. This pathway interacts with a cortico-basal ganglia-thalamo-cortical circuit that involves the basal ganglia subregion in the anterior striatum, the thalamus, and sends a projection back to Brocas/IFG.

Adapted from Arriaga and Jarvis (2013)

Chapter II. Viral strategies for transduction of Area X neurons

Abstract

The ability to alter gene expression of neurons, either to affect expression of endogenous molecules or to introduce the expression of exogenous molecules is one of the most powerful tools for interrogating nervous system function. One aim of my dissertation research requires overexpression of FoxP2 in Area X to test the hypothesis that dynamic FoxP2 regulation plays an important role in song learning. One strategy is to accomplish this is to exploit the natural ability of viruses to insert genetic material into cells in order to drive expression of experimenter-defined genes of interest in the central nervous system. To date, however, the use of such viral strategies in the zebra finch song system is limited. To identify a virus suitable for the transfection of Area X neurons, I tested a panel of various viral types, serotypes, promoters and viral cores. Here I identify a subset of these viruses that are suitable for overexpressing genes in Area X.

Introduction

One of the most powerful tools to understand the nervous system is the ability to directly manipulate the genes expressed by the brain. Knocking down, knocking out or overexpressing an endogenously expressed gene can allow one to gain insight into its function. The ability to overexpress exogenous molecules such as fluorophores or activity manipulating molecules such as opto- or chemo-genetics have been critical in mapping neural circuits and understanding the relationship between neural activity and brain function. To date, however, the use of such strategies in the avian nervous system has been limited.

One avenue of gene manipulation common in other model systems is transgenesis, globally altering. The major hurdle in generating a transgenic bird is the fact that birds are born in eggs which provides a physical barrier against accessing the developing embryo, a necessary step in creating transgenic animals. Several studies have overcome this hurdle by opening the egg, injecting lentivirus to induce transgenesis in the proto-gametic cells of the developing embryo, and then closing the egg an allowing it to hatch. So far this strategy has been used to generate zebra finch that overexpress GFP (Agate et al., 2007), CREB (Abe et al., 2015), and mutant Huntingtin protein (Liu et al., 2015). This strategy, however, has an almost prohibitively low success rate (1-2%) to make this a viable option for overexpressing FoxP2 in Area X. Transgenesis has several other limitations that make it non-ideal even if it had a higher success rate. First, this strategy lacks spatial specificity and would lead to elevated gene levels globally including numerous off target regions and cell types. Second, transgenesis lacks temporal specificity and thus FoxP2 would be elevated throughout the life of the bird, including during development, instead of the post-developmental period that is of interest. Finally, continuation of

any transgenic line will be problematic because any manipulation that leads to song learning errors would likely lead to a low mate success rate given the fact that song is a mating behavior.

While the use of other transgenic methods including CRISPR may eventually overcome these hurdles, viral mediated gene transfer offers a more viable method for addressing the function of FoxP2 regulation in Area X. Viral strategies have several advantages over transgenic ones. First, stereotaxic viral injections have a success rate of nearly 100% contrasting with the low success of current avian transgenic strategies. Second, viruses allow for spatial specificity as the virus is expressed only in the region where it has stereotaxically been injected. And third, virus allows temporal specificity, as it will only become expressed following the delivery of the virus.

For these reasons, a viral strategy was pursued. To date the use of target stereotaxic inject of virus in the song system has been limited. The first use of this technique was to knockdown FoxP2 in Area X using a lentivirus that expressed shRNA targeted to FoxP2 (Haesler et al., 2007). Interestingly, this approach proved to be unsuccessful in overexpressing FoxP2 (C Scharff, personal communications). In another use of lentivirus, Matsunaga et al., (2011) overexpressed the cell adhesion molecule cadherin7 in RA and found that it led to song learning deficits. It should be noted that that this manipulation was done in RA of Bengalese finch, which as the output to entire song circuit may be the song nucleus most sensitive to dysfunction. Moreover, the deficits were modest in magnitude and required that a large amount of data be segregated due to low transfection rate. Roberts et al., (2012) used both AAV and HSV to overexpress channelrhodopsin in HVC. Unfortunately, no data was shown on the neural specificity, transfection rate, or pattern of transfection of these viruses. Of note, viruses similar in design and source to the viruses used in this article are now being used by this lab to retrogradely

transfect afferents in the song system. Because of the low number of studies which have used stereotaxic viral injection and the absence of any which have used virus to overexpress a biologically relevant gene in Area X, I sought identify that a virus that could be used for this purpose.

To that end, a large panel of viral vectors was tested for their suitability in overexpressing genes in the song system. The primary impetus for testing these viral vectors was to test the hypothesis that FoxP2 downregulation is necessary for zebra finch vocal learning by constitutively elevating it in Area X neurons. To do so would require a virus that fulfills the following criteria and for the following reasons:

(a) The virus and associated promotor should transduce only neurons. *Rationale:* FoxP2 is expressed only in neurons.

(b) The virus should transfect a large number of neurons. *Rationale:* I am trying to test the effect of FoxP2 overexpression on behavior. Previous work in both zebra finch and mammals suggests that 20% should be the lower limit of transfection rate

(c) Transfected cells should express the transgene at a high level (I will operationally refer to this as "strong" transduction to distinguish from the previous requirement for large numbers of neurons being transfected). *Rationale:* FoxP2 downregulation after 2 hours is quite dramatic and potentially becomes even lower after the more extended vocalizations that a zebra finch normally does. This downregulation could only be countered with a strong virus.

(d) The virus should not retrogradely transfect afferent neurons. *Rationale:* Behavioral regulation of FoxP2 expression in the song system is limited to Area X and not its two afferent nuclei, HVC and LMAN.

(e) Viral expression should not spill over into adjacent reasons. *Rationale:* Similar to the retrograde requirement, the interpretation of a behavioral phenotype should not be confounded by off-target over expression.

(f) The virus should not damage the tissue or cause neurotoxicity. *Rationale:* Any effect of FoxP2 overexpression will ideally be studied in an intact circuit.

(g) The virus should provide persistent expression. *Rationale:* The song learning period is ~60 days and I required a virus that was able to drive expression of a transgene for this long a period.

In this chapter I tested a large number of virus to identify a subset that meets these criteria.

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).

Stereotaxic neurosurgery.

Adult male zebra finches were anesthetized with 2% isoflurane and placed in a custom-built avian stereotax (Herb Adams Engineering). The head was held at a 40-45° 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 μm, 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.

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).

Results

FoxP2 does not overlap with a marker for pallidal neurons

A large amount of research seeking to understand the role of FoxP2 in mediating speech and language deficits has focused on its expression function in striatal medium spiny neurons (MSNs). This is because many abnormalities in both humans and mice that carry FoxP2 mutations (Belton et al., 2003; Liégeois et al., 2003; Groszer et al., 2008) are limited to the striatum and MSNs are not only the principal cell type of the striatum but also the only striatal neuron type that expresses FoxP2. Interestingly, however, prior work had shown that early in human development FoxP2 is enriched in the internal segment of the globus pallidus (Teramitsu et al., 2004). More recent work in rodents has revealed a unique population of neurons in the external segment of the globus pallidus that send an ascending projection back to the striatum, have unique electrophyiological properties in vivo, and play a special role in stopping behaviors (Mallet et al., 2012; Mallet et al., 2016). Thus FoxP2 has the potential to be expressed in either Area X pallidal cell population and play an important role in controlling behavior.

To better understand whether to be mindful of a potential pallidal population of FoxP2-positive cells in selecting a virus, I examined whether there was any evidence of FoxP2expression in Area X pallidal neurons. Rochefort et al., (2007) claimed that pallidal neurons in Area X do not express FoxP2 but the basis of this argument is tenuous. They showed a) that pallidal neurons do not undergo neurogenesis, and b) new born neurons express FoxP2. These two pieces of evidence still allow for the existence a non-neurogenic population of FoxP2-positive pallidal neurons. As a second method, I tested whether there was any overlap between FoxP2 and the pallidal marker Lant6 (Reiner and Carraway, 1987). I found a complete non-

overlap between these two markers (Fig. 2-1). Thus, an AAV that tended not to transfect pallidal neurons is preferable over one that did.

Lentivirus introduction

The first category of virus that I explored first was lentivirus, a class of vectors that is derived from the human immunodeficiency virus. Lentivirus is a retrovirus and functions by incorporating its DNA into the genome of its host cell and as a consequence the expression of this DNA is permanent in both dividing and non-dividing cells. The amount of DNA that can be delivered is relatively large (7-8 kb) and cloning capacity is not a major issue in using this virus. Moreover, lentivirus has been shown to have a high degree of neural tropism and be effective in preferentially transfecting neurons.

In addition to these advantages, one major reason for selecting this virus was because at the time of test it was the only virus that had been used to alter expression of a transgene in the song system. In that case, the researchers used a lentivirus bearing an shRNA construct targeted to FoxP2 in order to knockdown FoxP2 in Area X during sensorimotor learning (Haesler et al., 2007). Since then several other publications have used lentivirus to manipulate FoxP2 and other song system genes (Matsunaga et al., 2011; Murugan et al., 2013).

Lentivirus results

The lentiviruses which were the major focus of my efforts were two viruses obtained from the UCLA viral vector core which used the human synapsin promotor (hSyn1) which had been demonstrated to transduce neurons exclusively (Kugler et al., 2003; Shevtsova et al., 2004). The experimental virus had the design hSyn1-FoxP2-IRES-GFP whereas the control virus had the design hSyn1-IRES-GFP (Fig 2-2a). The major problem with these viruses was that they

transfected too few neurons. In addition, they expressed very low levels of GFP that could only be detected using an anti-GFP antibody and tyramide signal amplification (TSA; Fig 2-2). In fact, GFP was barely detectable in hSyn1-FoxP2-IRES-GFP injected tissue that had undergone TSA amplification. Because the FoxP2 and GFP were translated separately by virtue of the IRES sequence, GFP levels do not necessarily reflect the expression levels of FoxP2. Nevertheless, the results were not promising and this was abandoned.

In addition to these lentiviruses, several other viruses, each of which was produced from the UCLA lentiviral core, were tested. One of these was a lentivirus obtained from the Feldman lab (UCLA), which used the synapsin promoter to express channelrhodopsin covalently attached to eYFP (ChR2-eYFP; described in more detail in Pagliardini, et al., 2011). Several similarities and differences exist between this and the previous lentiviruses. The major similarities are that both this and previous lentiviruses were produced by the same viral core and used the same promoter. One major difference was that the design of the construct was such that the molecule of interest, ChR2, was covalently attached to the fluorophore, eYFP that was used to detect expression of the virus. Thus one would expect a 1:1 ratio of both molecules, which may not be the case in viruses that utilize an IRES sequence. Second, as consequence of being covalently attached to ChR2 the fluorophore was attached the cell membrane rather than being free to fill the cytosol. Because neurons in general, and particularly small neurons such as MSNs, have high surface area-to-volume expression levels of this virus could potentially appear greater than would a comparably transfected cell which fills the intracellular compartment. And finally, this virus was ultra-centrifuged to increase the titer. The titer was unknown but because it was obtained from the same viral core as the previous lentivirus it would almost certainly be at a higher titer than previous non-centrifuged lentivirus.

Upon initial inspection, no transduction was apparent (not-shown), but expression was revealed using an anti-GFP antibody and further enhanced by TSA amplification (Fig. 2-3). This virus clearly was able to transduce Area X neurons, but it a) transfected a relatively diffuse population of cells and, b) the level of transfection was so low that several amplification steps were necessary to reveal a signal, and c) both of these issues would have likely been worse had the virus not been ultra-centrifuged. For these reasons and others related to the previous lentiviruses, the syn1-lentivirus approach was abandoned.

One final lentivirus tested was a PGK-GFP virus obtained directly from the UCLA Viral Core. This was the least successful of all the lentiviruses tested as this virus created a large hole at the injection site and most of the signal appeared due to scar tissue rather than neurons (Fig. 2-4).

Lentivirus conclusions

Lentivirus has the strength of persistent expression and, with an appropriate promoter it may be capable of strongly transducing avian neurons. Barring the PGK-lentivirus obtained from the viral core, there was no evidence of neurotoxicity or off-target transfection. Lentivirus also never appeared to travel retrogradely nor spill into adjacent tissues, thus allowing focal transfection of brain nuclei. Its major flaw was its inability to transfect large numbers of neurons. One possible strategy to overcome this hurdle would be to ultracentrifuge the virus and concentrate it to a titer suitable for transfecting a large number of neurons. Even this, however, might not be viable strategy as the Syn-ChR2-eYFP virus obtained from the Feldman lab was ultra-concentrated to a high titer and while this did improve transfection rate it did appear insufficient to reach levels of transfection we felt comfortable going forward for behavioral testing. On the other hand, others have used lentiviruses in zebra finch, including in Area X, and

have not only reported higher levels of transfections and have altered song behavior so these conclusions may better be limited to lentivirus obtained from the UCLA Viral Core rather than lentivirus generally. As will be clearly demonstrated with AAV, the source of the virus can be a critical but underappreciated variable in selecting a virus. According to Dr. Rachael Neve, director of the MIT Viral Core, the company Cyagen makes high quality lentivirus. It may be worth evaluating lentivirus from this source.

An additional lesson learned in the foray into lentivirus is that the synapsin promoter does not appear to work well in vivo and requires amplification to detect transfected cells. One possible explanation is that zebra finch lack the synapsin 1 gene (Warren et al., 2010) and thus its promoter may be non-functional or lead to only low levels of gene expression. This conclusion may, however, be somewhat erroneous as Dr. Julie Miller (U Arizona) has evidence that the construct used in vivo does indeed drive high levels of transgene expression in electroporation transfected zebra finch primary neuronal cultures (personal communication).

AAV introduction

The second type of class of viruses tested was adeno-associated virus (AAV). Its DNA incorporation is episomal but expression can persist for years in non-dividing cells. Based on an informal survey of the literature, AAV seems to be by far the most commonly used virus for manipulating behavior.

AAV results

The first AAV tested was a panel of viruses from the University of Pennsylvania (U Penn) viral core that used the CMV promoter to drive expression of GFP. As a group, all of these viruses failed in meeting the full set of criteria and several of them failed in very interesting ways. For example, when injected into Area X, AAV2/5 drove expression throughout most of

the brain with one notable exception- Area X itself. Indeed, the outline of Area X is very clearly seen as the negative space created by off target GFP expression (Fig. 2-5). In another bird injected with this virus, a salt-and-pepper pattern of expression was observed throughout the brain. Again, low levels of expression were seen at the site of injection, but unlike the other bird injected with this virus it did not create the negative image of Area X. Instead, expression levels were uniformly low in the striatopallidum. Thus, it is unclear whether the "negative image" pattern is a reproducible effect that could be harnessed in interesting ways (i.e. comparing the FoxP2 overexpression in Area X vs the outlying striatopallidum) or whether this was due to idiosyncrasies of that particular bird and that particular surgery. In either case, AAV2/5 showed a pattern that was not suitable for the project outlined here.

Other AAV from this panel failed in less interesting and perhaps less telling ways. For example, AAV2/1 showed bilateral lesions at the site of injections and GFP+ cells were broadly distributed across the brain (Fig. 2-6). A similar pattern was observed with AAV2/rh10 (Fig. 2-7) although it showed a somewhat lower transfection rate. AAV2/8, showed a similar issue with lesions at the injection (Fig. 2-8) site but unlike AAV2/1 GFP+ cells appeared limited to the region immediately surrounding the injection/lesion site.

One additional virus from U Penn that was not from the original panel, failed but ultimately proved very useful in identifying the elements of a successful virus. This was an AAV2/1 driving expression of CB7-GFP. The CB7 promoter is a CMV, Chicken Beta Actin fusion promoter similar to the more commonly used CAG promoter. It should also be noted that while this virus was obtained from the U Penn Viral Core, it was ordered several years later than the previously described U Penn AAV so it is unclear whether the differences between this and previous viruses are due to important differences in the viral construct or non-trivial differences

in the way this viral core produces its virus. This virus fulfilled every requirement except for the fact that it traveled retrogradely. This can be seen in Fig. 2-9 in which there is strong GFP expression in Area X, but also in cell bodies in HVC and LMAN both of which project to Area X. Moreover, there were even GFP+ terminals in RA which, like Area X, receives synaptic input from HVC and RA. It should be noted, however, that these GFP+ neuron terminals are likely to come exclusively from LMAN as the Area X projecting LMAN neurons send bifurcating axons to RA, whereas HVC sends projections to Area X and RA using discreet cell populations.

Finally, I tested a panel of viruses from Virovek, a private company based out of Hayward, CA. Each of these viruses used a CMV promoter to drive expression of GFP, the same design of the AAV tested from U Penn, and several viruses of the same serotype as the U Penn panel. Thus, these viruses were of the same construct design serotypes but differed only in where they were produced. By and large, the entire panel of Virovek AAV met the most if not all of the criteria. All of the viruses showed strong transduction and the ability to confine expression to Area X. None of them showed any sign of transfection of off-target cells types, neurotoxicity, nor retrograde trafficking. The primary difference was in the number of cells transfected by the virus. AAV1 (described in detail in Chapter III) and AAV5 (Fig. 2-10) transfected the largest number of cells with no clear difference between the two. The viruses from Virovek which did fail, such as AAV2/9 (Fig. 2-11) tended to do so by having a diffuse pattern of expression which was neither confined to Area X --but not brain wide like some of the U. Penn AAV-- and had low transfection rates.

As a final exploration of AAVs ability to overexpress genes one final permutation was tested. Having settled on AAV1 from Virovek with a CAG promoter as the means by which to overexpress FoxP2 in Area X, I tested whether the limited cloning capacity of AAV could be

worked around and could express two molecules, FoxP2 and GFP, with a single virus by using a p2A sequence. This peptide sequence, 18-22 amino acids long, would sit between FoxP2 and GFP in a large GFP-p2A-FoxP2 fusion protein, but would recruit a self-cleaving mechanism that would split it into a GFP molecule with about 10 amino acids of the p2A on its C-terminal and and FoxP2 which had about 12 amino acids on its N-terminal.

The results using this virus were very poor. First, essentially no GFP was detected in a bird unilaterally injected with the virus (Fig. 2-12). This was in contrast to easily detectable levels of expression of a virus expressing only GFP in the opposite hemisphere. Moreover, there was behavioral evidence that FoxP2 was not being overexpressed either. As will be demonstrated in Chapter III, overexpression of FoxP2 leads to poor song imitation. In contrast, a bird bilaterally injected with GFP-p2A-FoxP2 virus showed relatively normal song learning (Fig. 2-12). Together these results suggest that neither FoxP2 nor GFP were being overexpressed by this virus.

AAV conclusions

Depending on the serotype, promoter, and viral core, AAV can fulfill all of the required criteria. The U Penn viral core had a number of issues, but AAV from Virovek had a number of advantages. Several viruses transfected a large number of neurons with the exception of a few serotype and none transfected non-neuronal cell types, an observation consistent with AAV's high neuronal tropism in mammals. AAV also did not appear to be neurotoxic with the exception of a subset of AAV obtained from U Penn.

Virovek's AAV was the most promising and thus ultimately selected for overexpressing FoxP2 in Chapter III. AAV2/1 and AAV2/5 were clearly the best of the group and there was no major difference apparent. Ultimately, AAV2/1 was chosen, in part, because in rodents while

both AAV2/1 and AAV2/5 have equal transduction efficiency in the striatum, AAV2/1 has a considerably lower efficiency in the pallidum. Given the lack of evidence of FoxP2 expression in Area X pallidal neurons (Fig. 2-1), a tropism against these neurons was viewed as an advantage.

Additionally, the U Penn virus that went retrograde could prove to be important for circuit mapping, circuit manipulation, or selective targeting anatomically defined cell populations. For example, a retrograde acting AAV could be injected into either Area X or RA and another AAV carrying a CRE-dependent gene could be injected into HVC. This would allow for selective targeting of Area X projecting neurons in HVC.

HSV introduction

The final class of viruses tested was herpes simplex virus (HSV). Like lentivirus and AAV, HSV has a high degree of neural tropism. It also has by far the largest cloning capacity of the three (Neve et al., 2005). The major and well-characterized limitation of HSV is that its expression is only transient at the site of injection. More specifically, there are two types of HSV both of which show different patterns of transient expression. Short term HSV (ST-HSV) has been described to show only transient expression at the site of injection, showing expression at several hours post-delivery which is gone by 8-10 days post-delivery. In contrast, long-term HSV (LT-HSV) shows the same pattern of transient expression, but also retrogradely transfects afferent neurons and expression in these cells is thought to be permanent (personal communication, R. Neve). While neither expression pattern is suitable for constitutively elevating FoxP2 during development, the ST-HSV pattern was a viable option for the DREADDS project described in Chapter IV. Moreover, it could be suitable for acutely elevating FoxP2 levels for electrophysiological experiments proposed in the concluding Chapter

HSV results

The first HSV tested was a LT-HSV which expressed GFP using the non-specific IE 4/5 promoter. When tested at full strength, this virus caused very obvious cell death (2-13). Moreover, there was evidence of retrograde trafficking (data not shown). I found, however, both issues could be mitigated if not eliminated by diluting the virus with saline before injecting the virus (Diluted HSV will be used in Chapter IV and images will provided which show that it does not lesion tissue). To test whether the large cloning capacity of HSV could be taken advantage of to overexpress two molecules (FoxP2 and GFP) we obtained a HSV which overexpressed FoxP2 with the IE 4/5 promoter and GFP with a CMV promoter. This virus showed high levels of co-expression of both genes (Fig. 2-14).

HSV conclusions

HSV showed several desirable qualities including a high transfection rate, strong transduction of neurons, and the ability to transfect a brain region without spilling over into adjacent regions. In some cases this virus did go retrograde or cause neurotoxicity, but diluting the virus can mitigate both of these issues.

Discussion

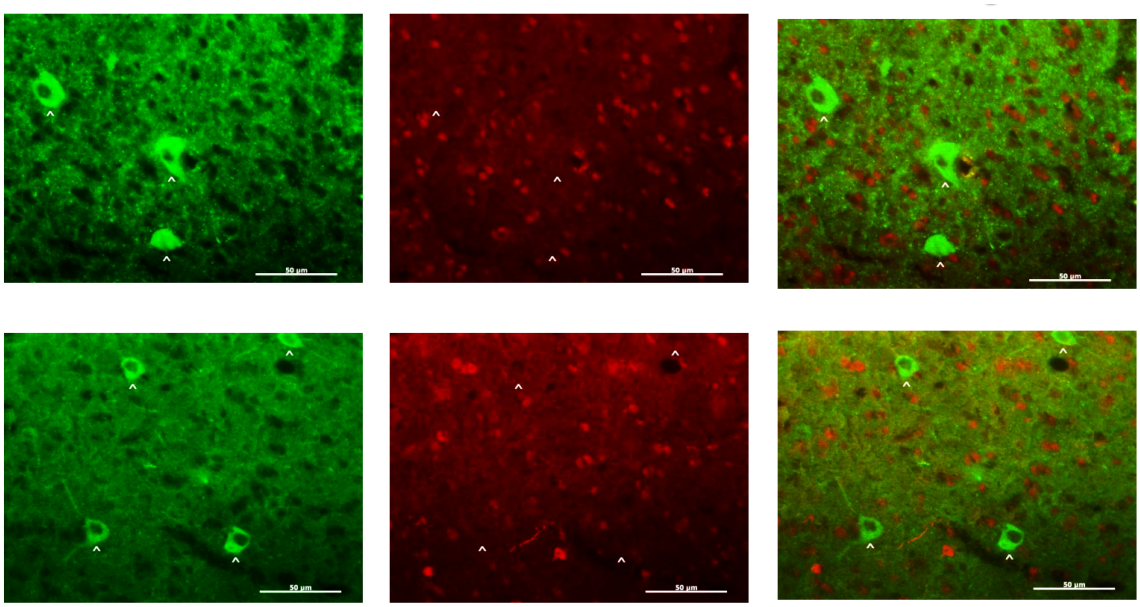
Here I qualitatively describe efforts to characterize various viruses in terms of their suitability for overexpressing genes, FoxP2 in particular, in Area X. Ultimately Virovek AAV1 using a CAG promoter was chosen as the most suitable for this task. As a more general principal, however, each virus has its advantages and disadvantages for other projects. As a guide to choosing a virus to transfect and transduce Area X neurons I suggest the Venn diagram shown in Fig 2-15 featuring three requirements that may be desired of a virus for a particular project: high transfection rate, high cloning capacity, and persistent expression. Lentivirus, AAV, and HSV each fulfill two and only two of these criteria. Lentivirus has a high cloning capacity and persistent expression, but a low transfection rate. This virus may be most suitable when a large transgene must be expressed, but where a behavioral phenotype is not required. One such application is in electrophysiological or neuromorphological experiments designed to study cell autonomous effects. AAV transfects large numbers of neurons and does so in a persistent manner, but its cloning capacity can be a limiting factor. AAV is useful in obtaining a behavioral effect but may not be up to the task for expressing large genes or multiple genes. HSV transfects large numbers of neurons and has a large cloning capacity, but its expression is transient. Moreover, neurotoxicity and retrograde trafficking are a concern when using this virus but both appear to be mitigated if not eliminated by diluting.

In addition to these concerns there are further practical concerns. AAV is biosafety level one, whereas lentivirus and HSV are biosafety level two. On the other hand, Virovek, the only AAV source I recommend for use in zebra finch, has a variable turnaround rate for viral

production which in our experience has ranged from ~1-8 months, whereas the turnaround rate from HSV from the MIT Viral Core is much faster and on the order of weeks.

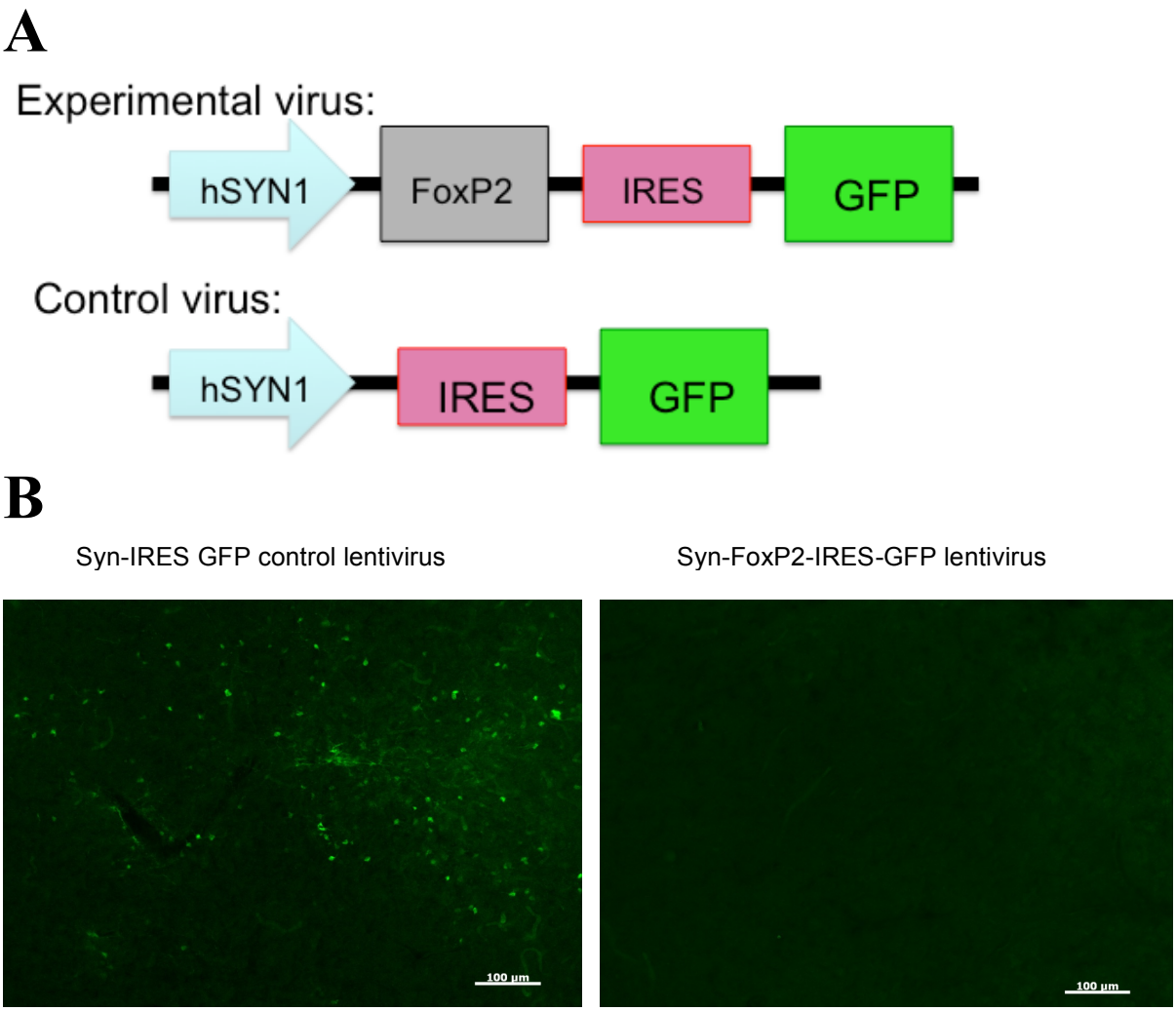
One final question that I think is worth addressing at some point is whether the CRE-lox system works in birds. In theory it should and could be harnessed to use a combinatorial strategies to transfect discreet cell populations. It should not, however, be taken for granted that this will work as my efforts here showed that such viral tools must be empirically validated in zebra finch.

Fig. 2-1: FoxP2 does not overlap with the pallidal cell marker Lant6



No overlap was detected between the pallidal marker Lant6 (green) and FoxP2 (red). This suggests that pallidal cells do not express FoxP2 and a virus should preferably not target these cells. Scale bar = 50 uM

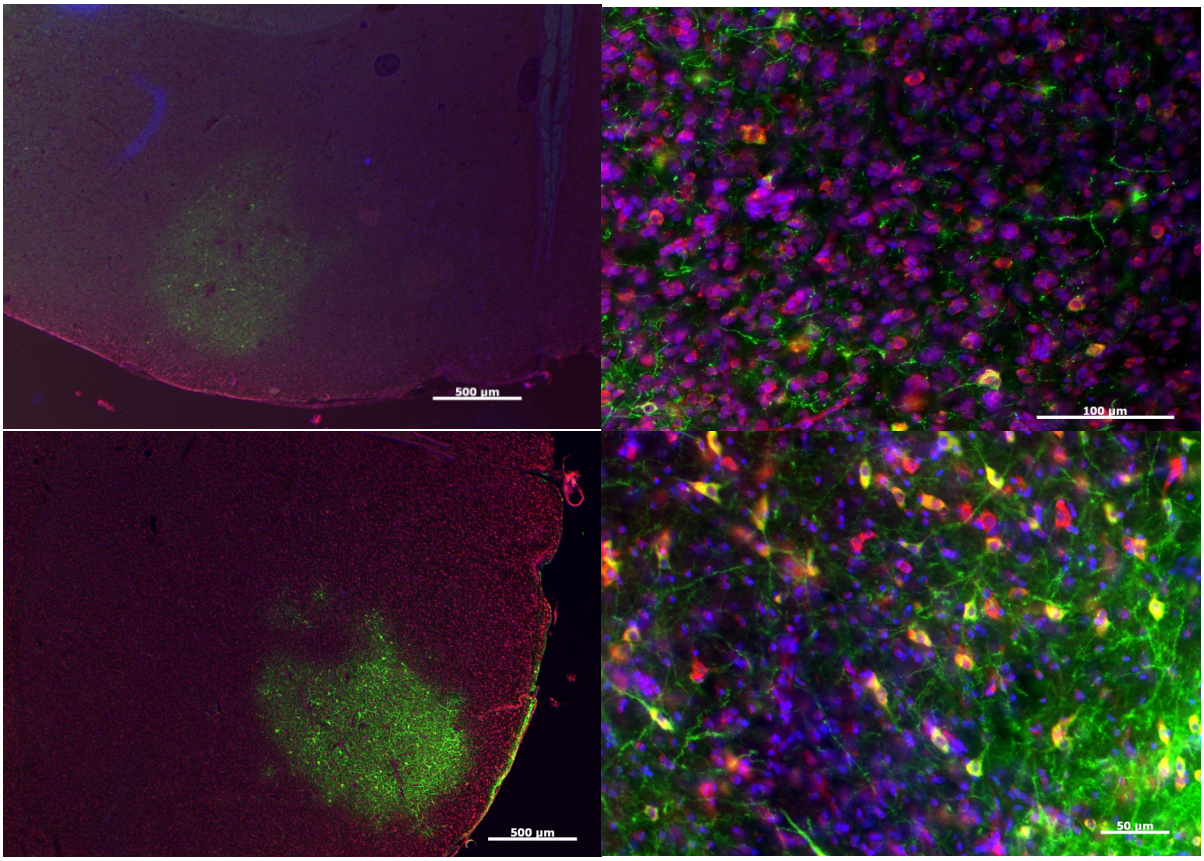
Fig. 2-2: Characterization of UCLA Syn1-FoxP2-ires-GFP lentivirus.



A) Design of lentiviral constructs. Both the experimental and control viruses used the hSyn1 promoter to achieve neuron specific expression. The experimental virus used this promoter to drive expression of FoxP2-ires-GFP, which should then be translated to FoxP2 and GFP. The control virus used the hSyn1 promoter to drive expression of the ires-GFP, which should then be translated to GFP.

B) The control virus showed sparse expression only sparse expression of GFP that could only be detected using a GFP antibody. Very little GFP was injected in the contralateral hemisphere injected with the experimental virus. Scale bar = 100 uM

Fig. 2-3: Characterization of ultra-centrifuged UCLA hSyn1 lentivirus obtained from the Feldman lab.



Shown here is a photomicrograph from a bird injected with hSYn lentivirus from the Feldman lab. GFP is shown as green. The tissue was counterstained for NeuN (red) and DAPI (blue). This coloring scheme is used throughout this chapter except where noted. No GFP was detected without an anti-GFP antibody.

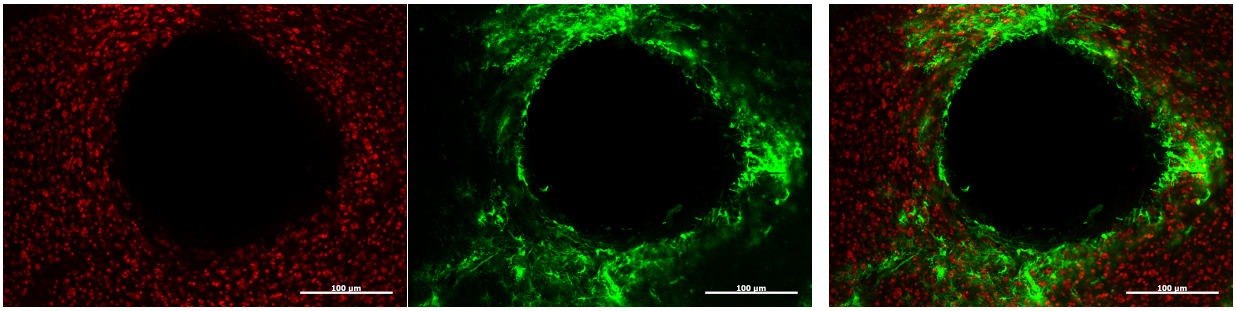
(Top left) GFP positive signal from a bird injected with this lentivirus. This section was stained using an anti-GFP antibody. This was able to reveal low level GFP transfection. Scale bar= 500 µm

(Top right) Higher magnification image from this same section allows quantification of transfection rate. Despite the presence of GFP signal, the infection was sparse with few GFP+ cell bodies.

(Bottom left) TSA amplification reveals higher levels of GFP than a standard IHC. Thus, TSA amplification was used to validate any virus that appeared to show a low transfection rate. Scale bar= 500 μm.

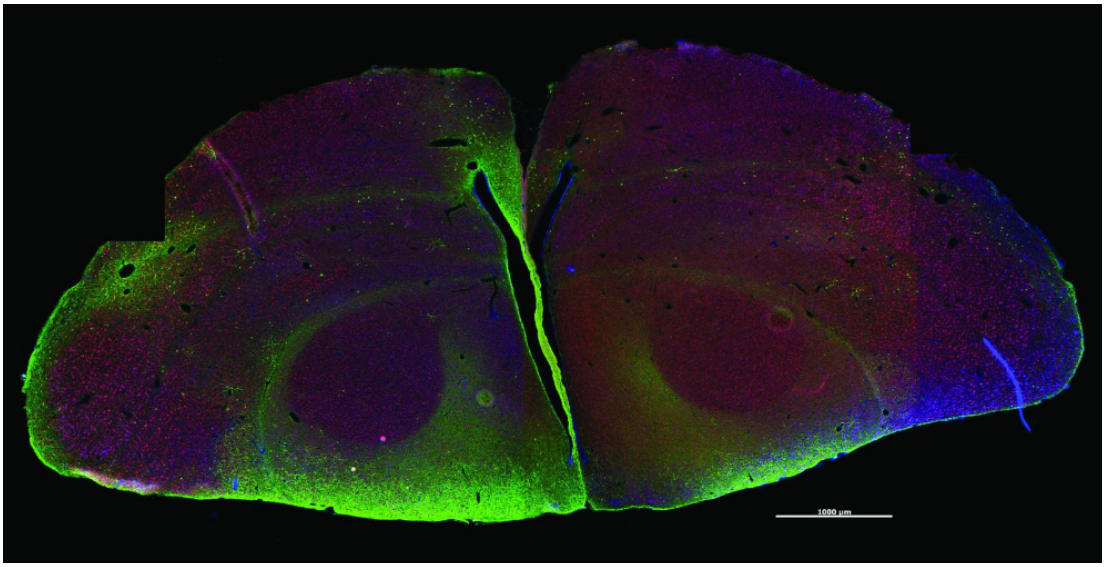
(Bottom right) TSA amplification of GFP reveals a higher transfection rate than was observed with a standard IHC. Nevertheless, the transfection rate appears fairly sparse. Scale bar= 100 μm.

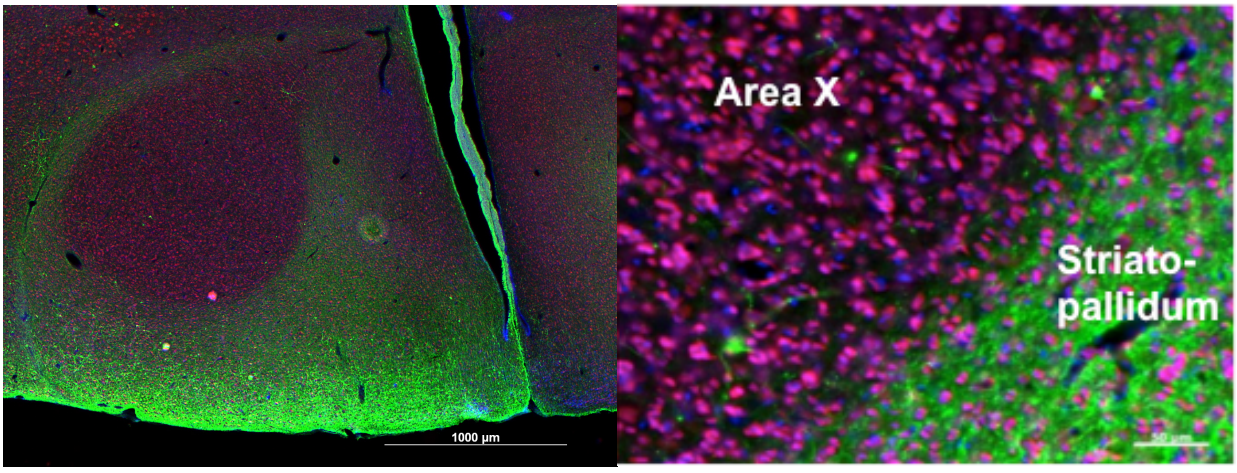
Fig. 2-4. UCLA Viral Core PGK-lentivirus causes neurotoxicity and transfects scar tissue.



A large lesion was observed at the injection site. The pattern of GFP expression around the injection site seemed to indicate off target transfection of non-neuronal cells. Scale bar= 100 µm.

Fig. 2-5. U Penn AAV2/5 transfects large portions of the brain except for Area X



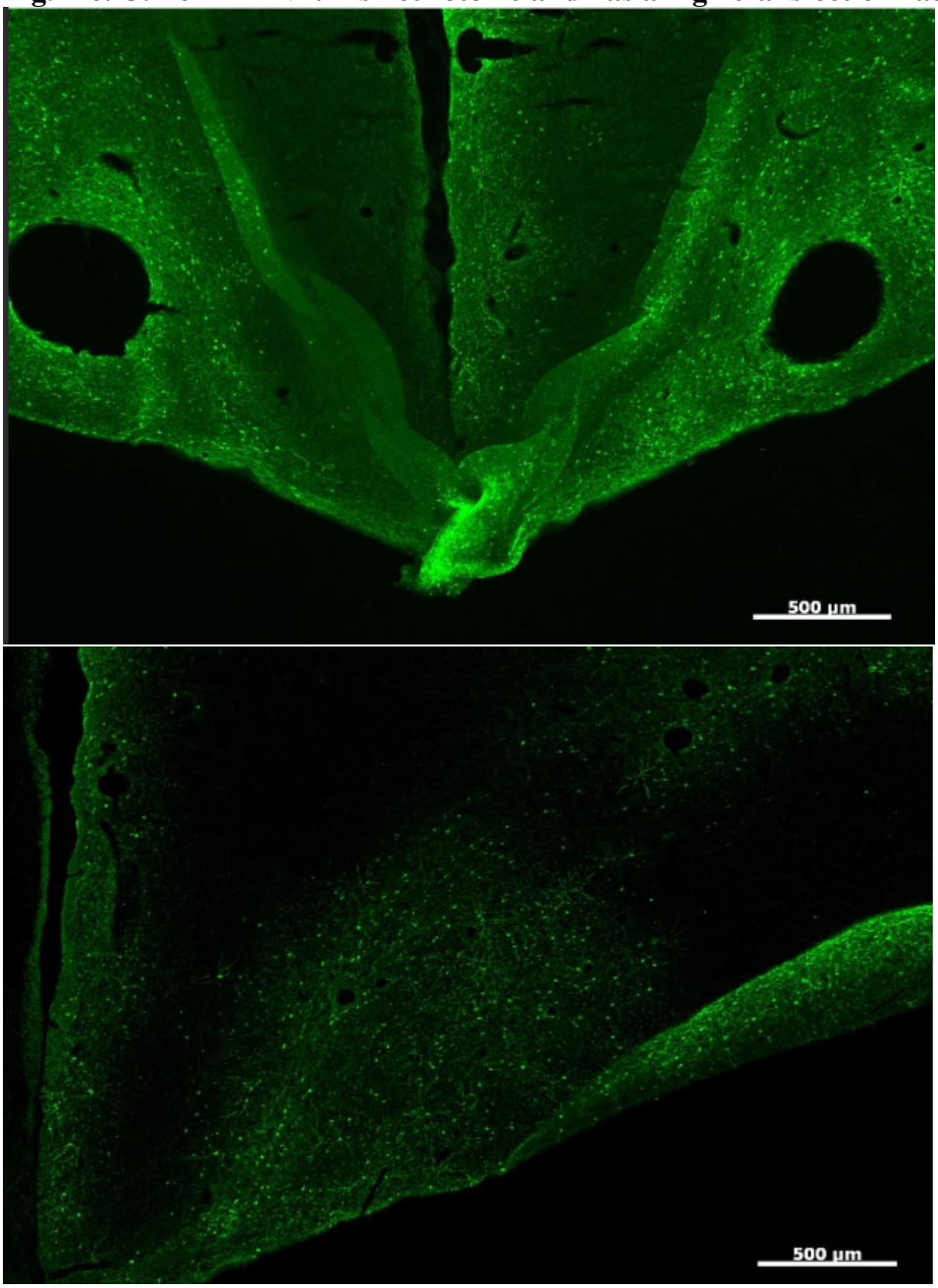


(Top) Low magnification image shows the extent of viral spread throughout the brain. Interestingly the large number of regions transfected did not include Area X. Scale bar= 1000 µm.

(Lower left) High levels of transfection were seen throughout the striatopallidum, but Area X clearly be identified as the region not transfected by virus. Scale bar= 1000 µm.

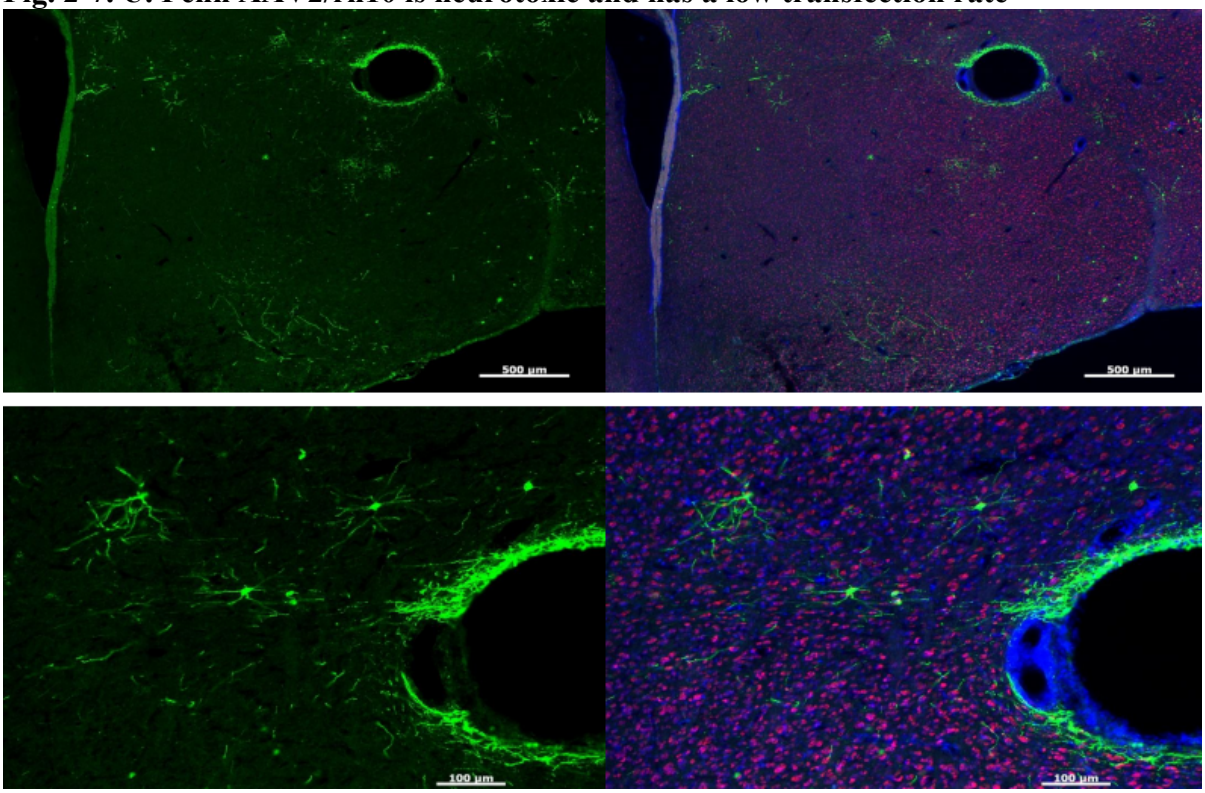
(Lower right) At higher magnification a sharp delineation between untransfected Area X and the virus transfected outlying striatopallidum is evident. Scale bar= 50 µm.

Fig. 2-6. U. Penn AAV2/1 is neurotoxic and has a high transfection rate



(Top) Large bilateral holes are seen at the site of the injection. These holes are surrounded by large numbers of GFP positive cells are seen through out this section. Scale bar= 500 µm. **(Bottom)** Dense numbers of GFP positive cells are seen throughout the brain including this striatopallidal region caudal to the injection site. Scale bar= 500 µm.

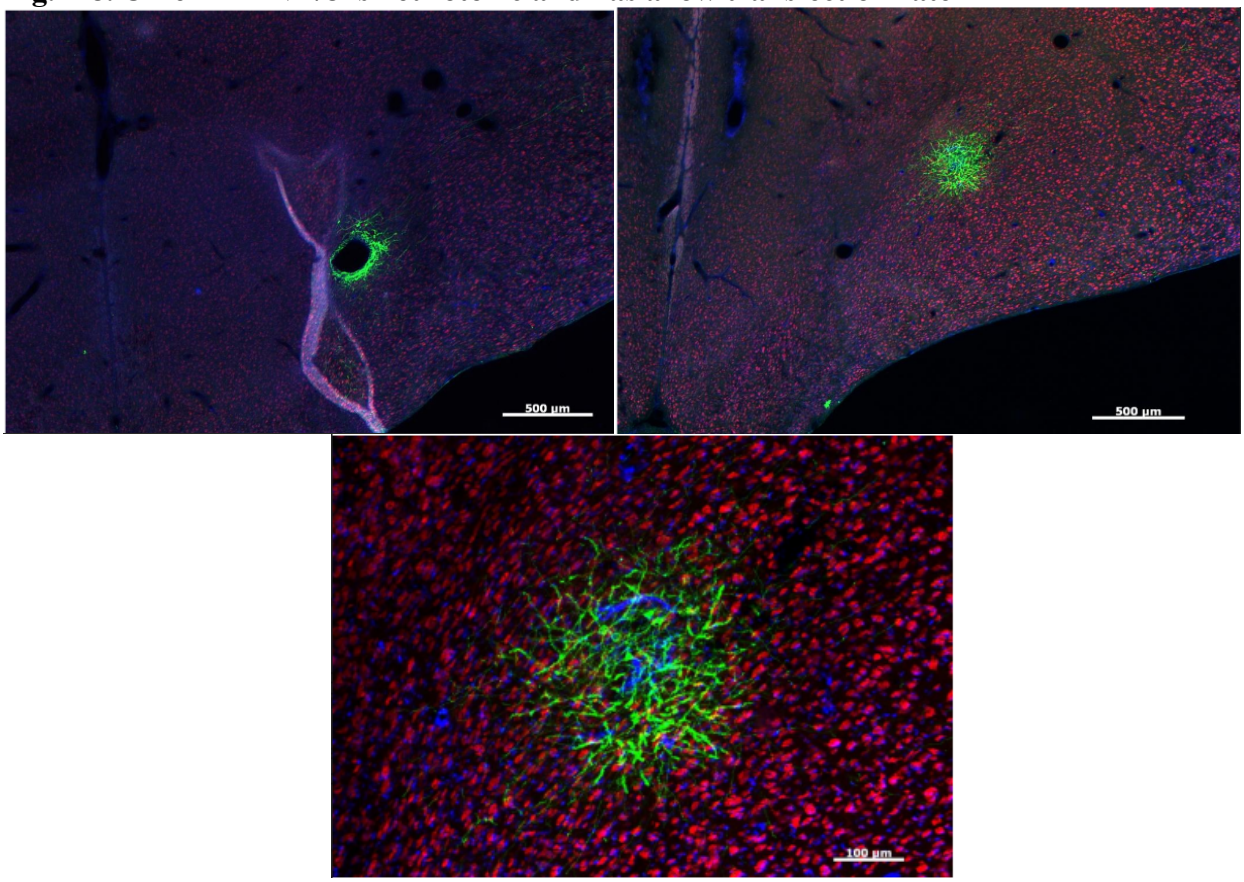
Fig. 2-7. U. Penn AAV2/rh10 is neurotoxic and has a low transfection rate



(Top) The injection site was not quite accurate and was at the striato-nidopallial border. Nevertheless, this virus showed some major flaws that disqualified it as being useful for overexpressing genes in the song system. The major flaw was that it created a large lesion at the injection site. In addition, the pattern of expression was non-ideal. There were high amounts of GFP surrounding the lesion site and it was unclear whether these were neurons or transfected scar tissue. In addition, this virus transfected neurons that were quite distal to the injection site and rather densely transfecting a limited region of the brain it showed a salt-and-pepper pattern of transfection that spanned large portions of the brain. Scale bar= 500 µm.

(Bottom) The sparse salt-and-pepper pattern of transfection is evident in these higher magnification images. Scale bar= 100 µm.

Fig. 2-8. U Penn AAV2/8 is neurotoxic and has a low transfection rate

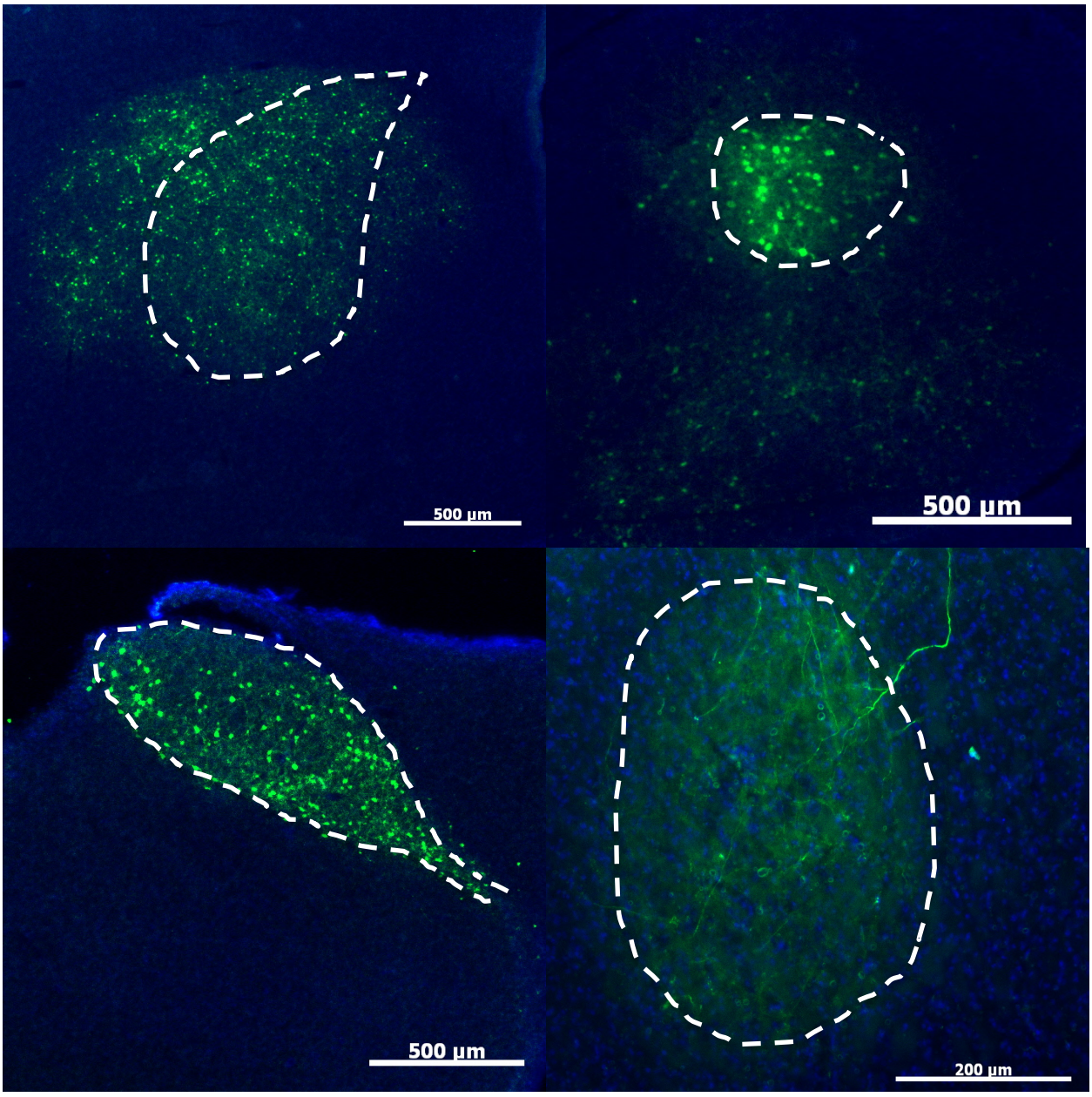


(Top left) U Penn AAV2/8 caused a lesion at the injection site which was surrounded by GFP positive cells. Scale bar= 500 µm.

(Top right) Slightly rostal to the injection site the lesion was less apparent, but the it was clear this virus transfected a very limited region of the brain. Scale bar= 500 µm.

(Bottom) At higher magnification it was clear that even at the most densely transfected region there were very few GFP+ cell bodies. Scale bar= 100 µm.

Fig. 2-9. U Penn AAV2/1-CB7-GFP retrogradely injects Area X afferent nuclei.



(Upper left) High numbers of GFP-positive cells were seen at the injection site in Area X. Scale bar= 500 µm.

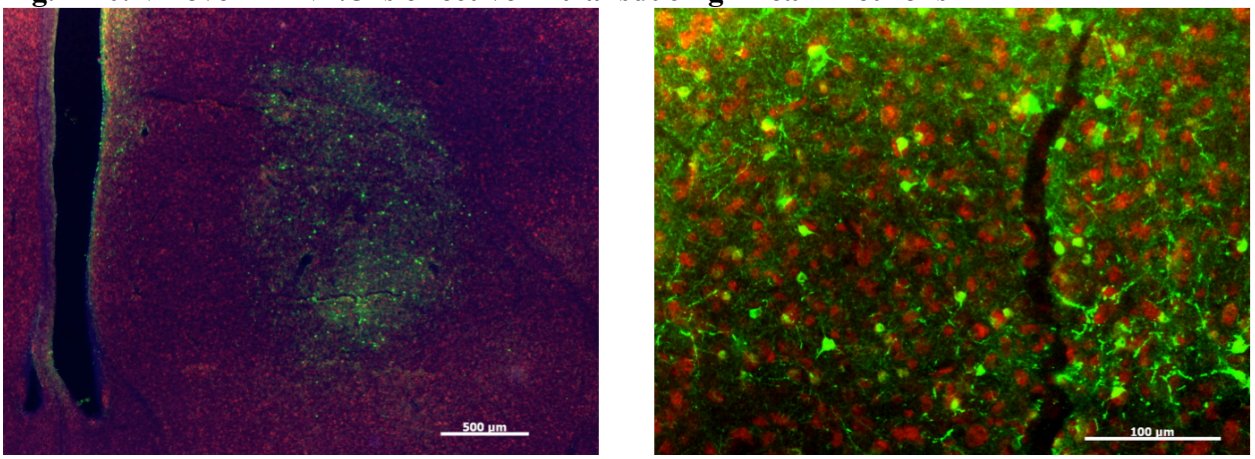
(Upper right) High numbers of GFP-positive cells were also identified in LMAN. Because LMAN is so near Area X and the injecting electrode passes through LMAN on its way to Area

X, it was unclear whether this off-target expression was the result of spillover or of retrograde transfection. Scale bar= 500 μm.

(Lower left) Spillover was ruled out by expression of GFP in HVC which is ~7 mm dorsal-caudal to Area X. Moreover, the near perfect outlining of this nucleus is best explained by retrograde transfection. Scale bar= 500 μm.

(**Lower right)** Retrograde transfection was further confirmed by expression of GFP+ fibers in RA, which does not receive input from Area X but does receive input from LMAN and HVC. Scale bar= 200 μm.

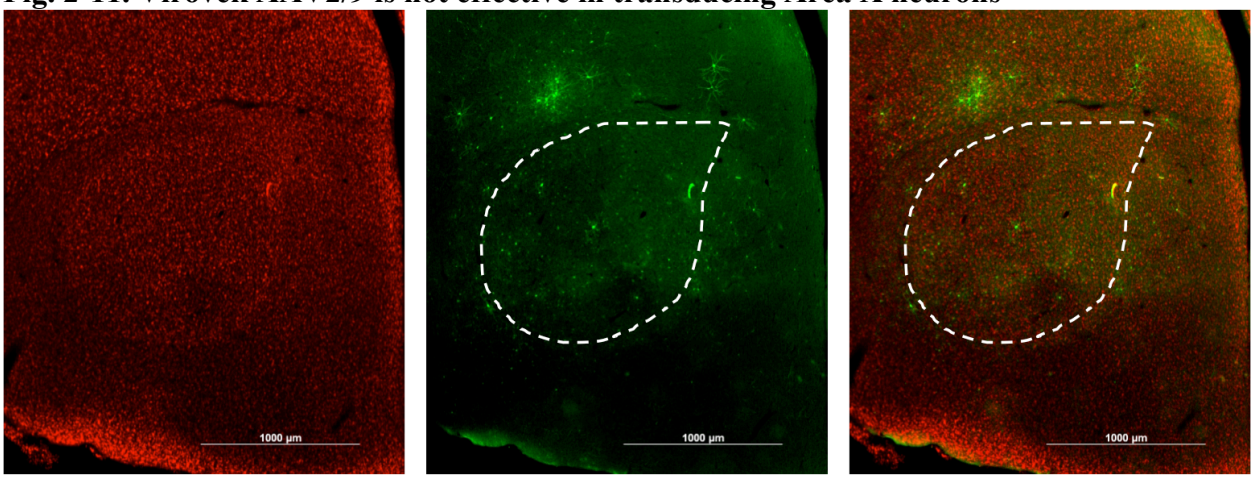
Fig. 2-10. Virovek AAV2/5 is effective in transducing Area X neurons



Left) This low magnification image shows that AAV2/5 can transfected a limit region of the brain. Some transfection is seen dorsal of the striato-nidopallial border, but this seems to be the result of mis-targeting of the virus rather than an inability of this virus to limit transfect to a well-defined region. Scale bar= 500 µm.

Right) This high magnification image shows that AAV2/5 has a high transfection rate. Scale bar= 100 µm.

Fig. 2-11. Virovek AAV2/9 is not effective in transducing Area X neurons



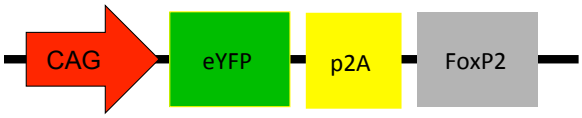
Left) Area X is clearly visible at the injection site

Middle) U Penn AAV2/9 injection in Area X leads to GFP+ cells that are diffusely scattered around the injection site.

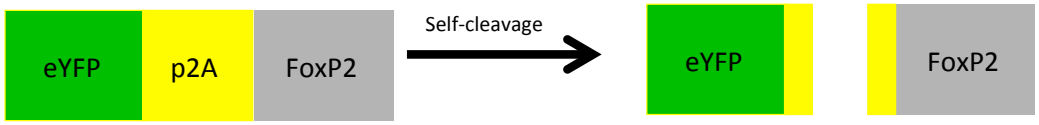
Right) Merged image.

Fig. 2-12. A p2A sequence is ineffective in expressing either GFP or FoxP2

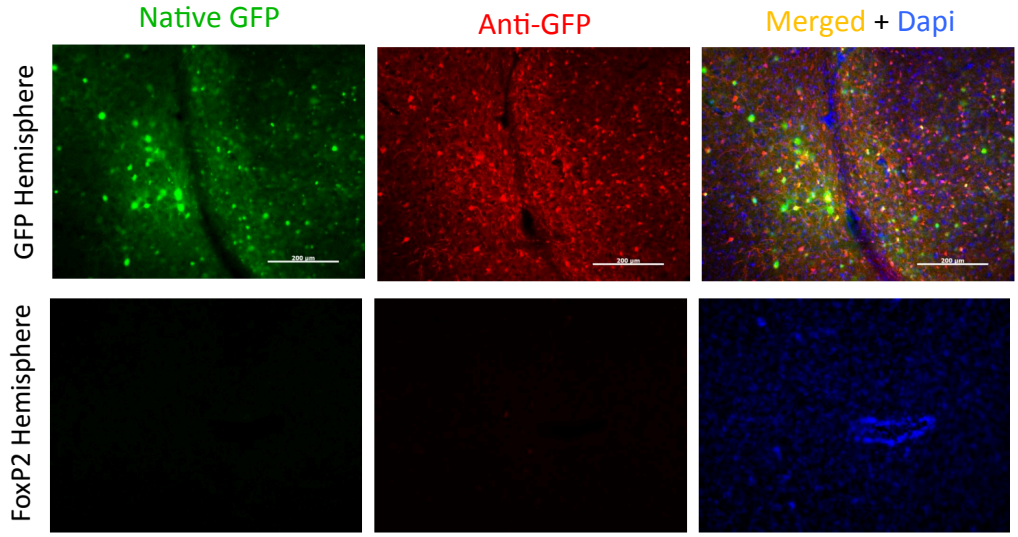
A.



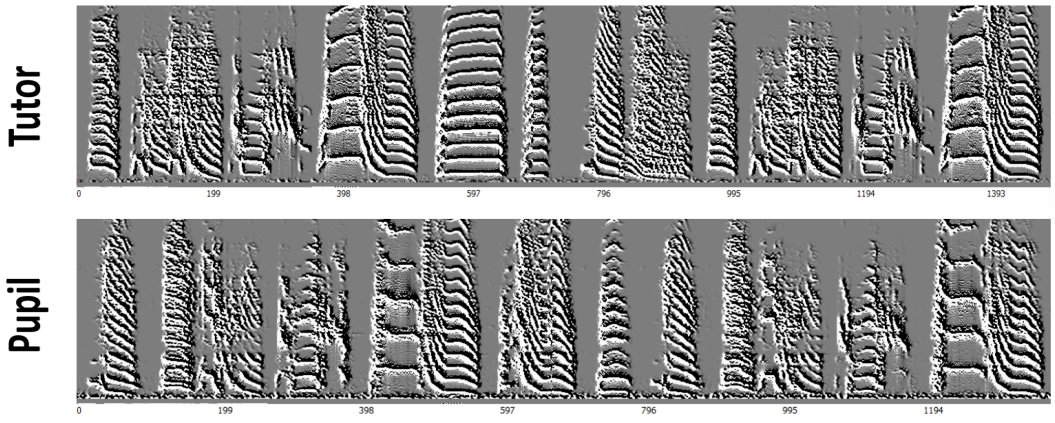
B .



C.



D.



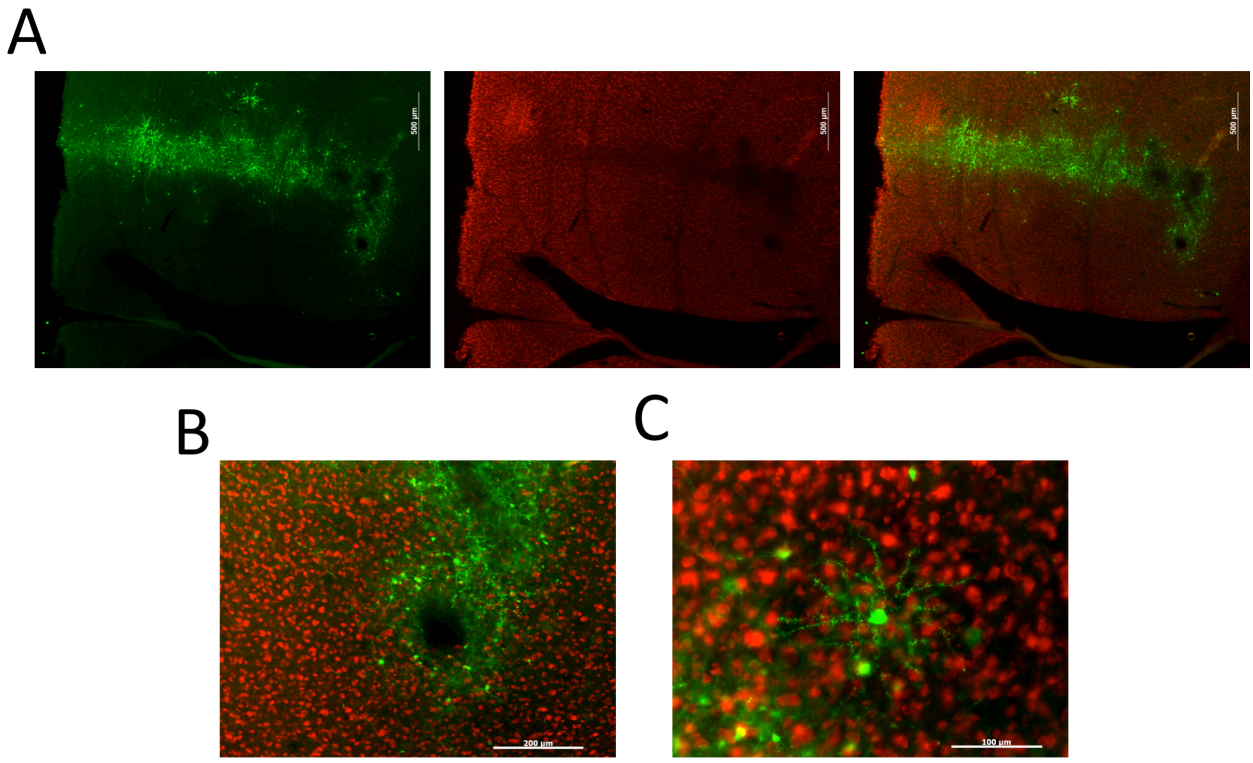
A) Schematic of virus that utilizes a p2A.

B) In theory, a single gene product should be translated which creates a GFP-p2A-FoxP2 fusion protein that is then self cleaved in separate GFP and FoxP2 molecules which carry residual p2A peptides on their C- and N-termini, respectively.

C) The GFP viruses injected hemisphere showed high levels of native GFP (green) which was further amplified using an anti-GFP antibody (red). In contrast no GFP, native or amplified, was visible in the contralateral hemisphere injected with the GFP-p2A-FoxP2 virus.

D) A young bird injected with bilaterally with the GFP-p2A-FoxP2 virus showed normal song learning and thus lacked the song learning impairments typical of birds overexpressing only FoxP2 as shown in Chapter III.

Fig. 2 -13. HSV effectively transfects Area X but can cause neurotoxicity and retrograde transfection.



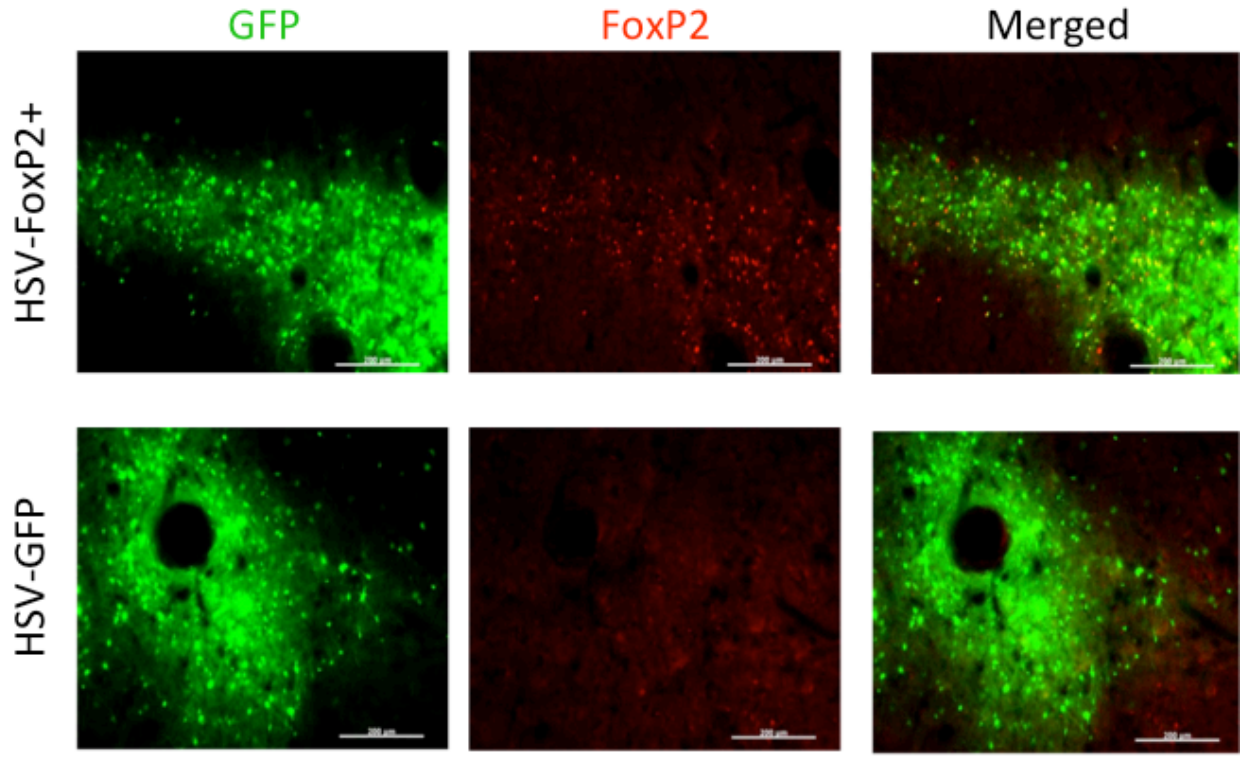
A) High levels of GFP are seen along the injection track. This was not confined to Area X which was the injection site. Along the length of this transfected area, the darkened are made it clear that there was some level of neurotoxicity

B) Like some of the earlier viruses, HSV causes a hole at the injection site, but this is smaller than the large hole created by viruses such as the PGK lentivirus and AAV 2/8.

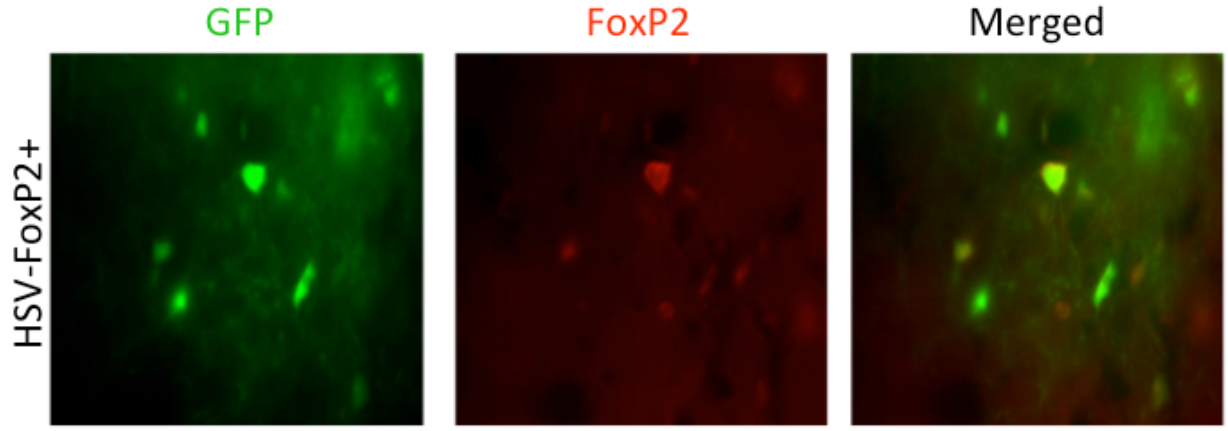
C) This virus was capable of transfecting cells that clearly appeared to be MSNs.

Fig. 2-14. HSV is very effective in overexpressing FoxP2.

A)



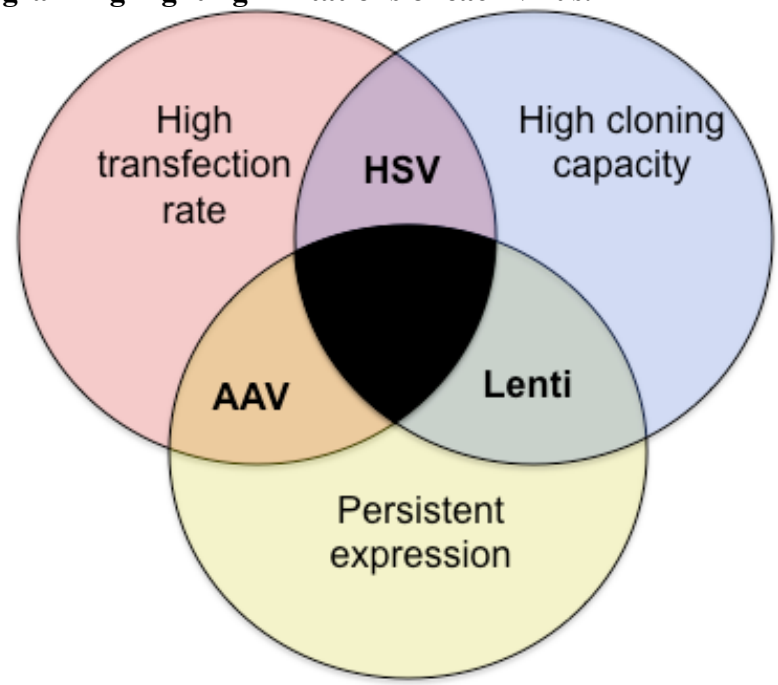
B)



A) A bird was unilaterally injected with HSV-FoxP2+ in one hemisphere and HSV-GFP in the other. While GFP is observed in both hemispheres, elevated levels of FoxP2 are detected only in the FoxP+ hemisphere and only in the region showing GFP.

B) The tight correspondence between GFP and FoxP2 in the HSV-FoxP+ hemisphere is clear at higher magnification

Fig. 2-15. Venn Diagram highlighting limitations of each virus.



This Venn diagram highlights the strengths and limitations of each viral time as they relate to specific requirements necessary for a given scientific application. Those requirements are: 1) a high rate of transfection, 2) a high cloning capacity and, 3) persistent expression. Each viral type fulfills two and only two requirements.

Chapter III. Behavior-linked FoxP2 regulation enables zebra finch vocal learning

Abstract

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.

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 post-developmental 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. 2-1). 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.

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 2-1f. At 18 d young birds were moved to sound attenuation chambers (Acoustic Systems) along with both parents and any clutchmate 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.

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 (±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 μm, 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 β 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 2.24E+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 μm for perfused brains and 20 μm for fresh frozen brains. Verification of targeting and overexpression of zebra finch *FoxP2* mRNA following injection of the FoxP2-expressing virus (Fig. 2-1B) 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 μm 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 μl of ice-cold modified RIPA lysis buffer with protease inhibitors using a hand-held homogenizer, mixed with an equal volume of 2× Laemmli loading buffer (Bio-Rad) containing 0.1% β-mercaptoethanol and stored at −80°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. 2-1e) 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 inFigure 1*e* 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 an 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 where 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. 2-1a). Both the FoxP2+ and GFP-expressing AAV1 viruses transfected 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. 2-1b) and FoxP2 protein expression (Fig. 2-1e). 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 ± 5.5% of the NeuN-positive cells also expressed GFP and, of the total number of transfected cells, 96.7 ± 1.7% were NeuN-positive (n = 4 birds; Fig. 2-1c), indicating a robust level of neuron-specific expression. Additionally, we found a high degree of overlap between GFP and endogenous FoxP2 (Fig. 2-1d), 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 transduce pallidal-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 micropunches 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 1*f*. 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. 2-2a). 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. 2-2b). 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. 2-2c). No difference in syntax accuracy was detected (Fig. 2-2d), 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. 2-2e). Across our dataset, syllable identity scores were much lower in FoxP2+ birds compared with GFP controls (Fig. 2-2f). 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 Euclidian 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 durational error than the corresponding syllables imitated by their GFP control brother (Fig. 2-2g). 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. 2-3a). 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. 2-3b). This suggests that 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. 3*c*). 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. 2-4*a*). Another way of analyzing vocal variability is to examine the coefficient of variation for specific features. Consistent with the self-similarity results, we were unable to detect any difference in CV values for any feature (Fig. 2-4*b*). Finally, the adult songs of FoxP2+ and GFP

control birds did not differ in syntax entropy (Fig. 2-4*c*), 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. 2-5*a*). We have previously shown that this behavioral manipulation both decreases Area X FoxP2 mRNA and protein levels and increases vocal variability 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. 2-5*a*,*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. 2-5c,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 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. 2-5*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. 2-5*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 overexpression 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. 2-4).

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. 2-1b) and FoxP2 protein (Fig. 2-1e) 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., 2008) including in young, as well as adult birds (Teramitsu 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 effect on the developmental progress toward the bird's own final song. Although it could be argued that both outcomes are a nonspecific consequence of altering the activity of any major transcription factor, there is precedence for specific but non-opposing 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 et al., 2015). 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 (Iacoboni 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

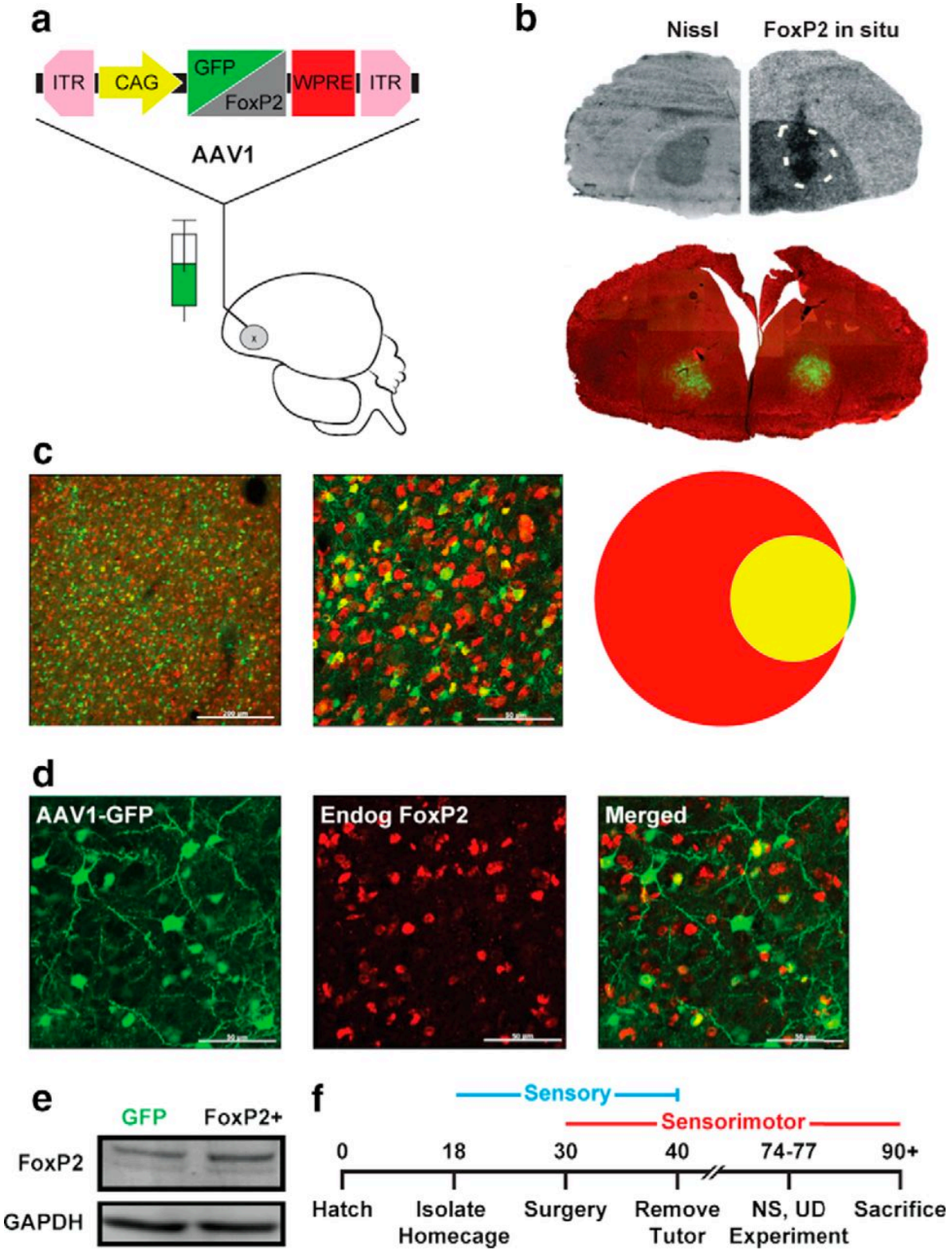
variability observed here is an effect of retarded vocal imitation rather than an intrinsic effect of FoxP2 overexpression *per se*. Ravbar et al. (2012) have shown that variability decreases as syllables hone in on their target syllable in the model song. A corollary observed here is that variability 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 gadd45α results in comparable decreases in dendrite complexity *in vitro* (Sarkisian and Siebzehnrubl, 2012).

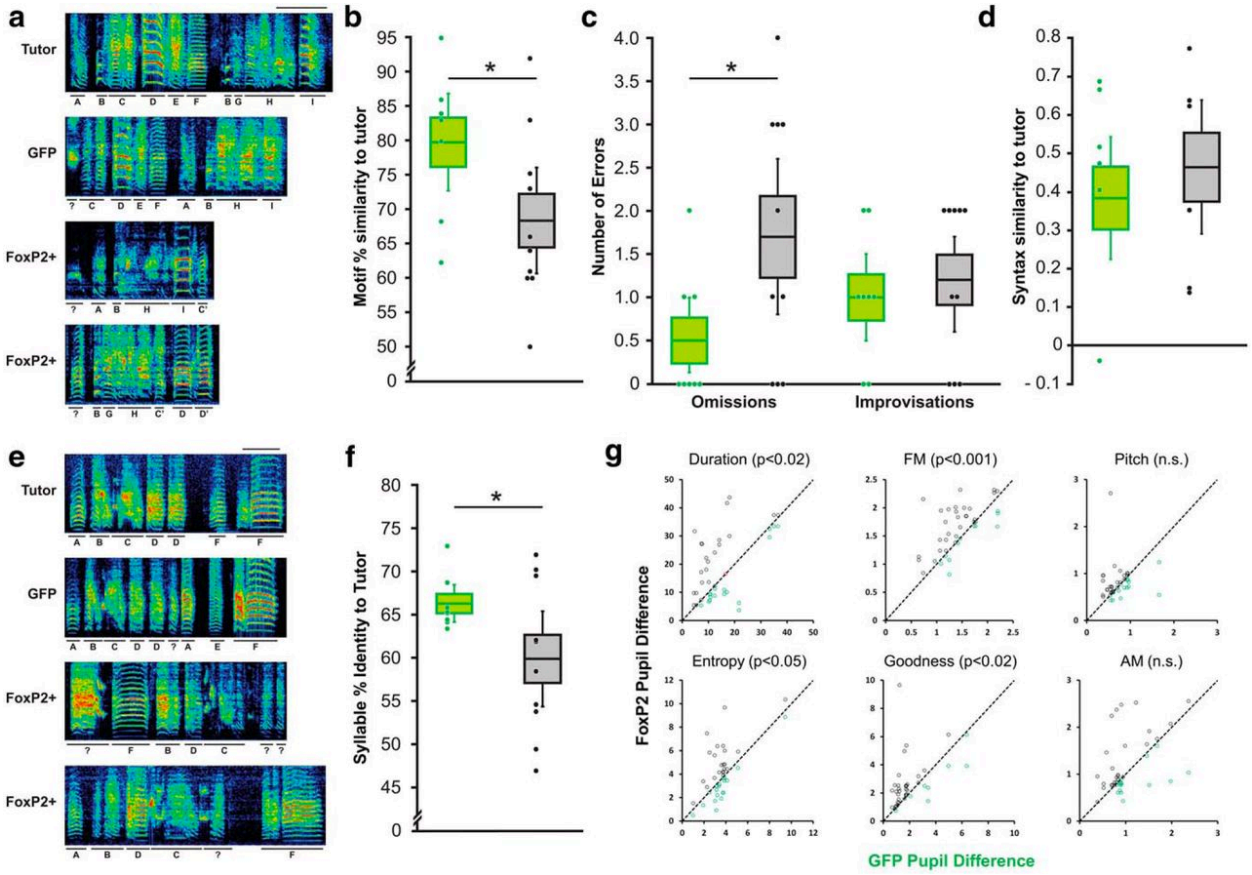
Together with our work, these studies support the dichotomy between permissive molecules, for which constitutively high (or low) levels are able to support learning versus gating molecules, which require dynamic transitions between high and low expression levels to switch between states of plasticity 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.

Figure 3-1. 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.

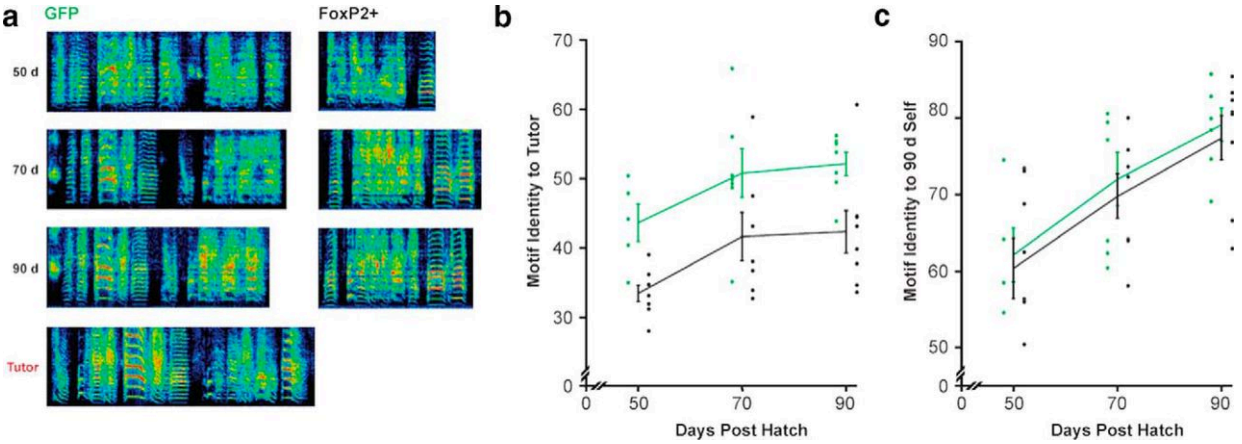
Figure 3-2: FoxP2 overexpression during the sensorimotor critical period disrupts vocal learning.



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 = 8GFP/10FoxP2+, 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 = 8GFP/10FoxP2+, unpaired one-tailed bootstrap). In contrast, GFP and FoxP2+ pupils exhibit similar levels of improvised syllables ($p = 0.3040$, n = 8GFP/10FoxP2+, 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 = 7GFP/8FoxP2+, 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 = 8GFP/10FoxP2+, 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).

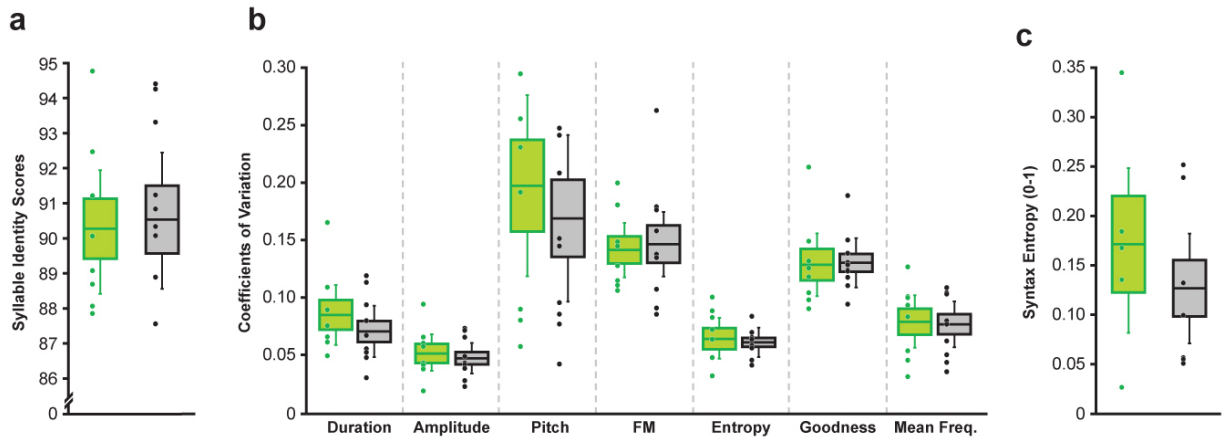
Figure 3-3: FoxP2 overexpression leads to imitation deficits that emerge early despite similar developmental trajectories.



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 = 5GFP/8FoxP2+, unpaired one-tailed bootstrap), 70 d ($p = 0.0398$, n = 7GFP/7FoxP2+, unpaired one-tailed bootstrap), and 90 d ($p = 0.0082$, n = 8GFP/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

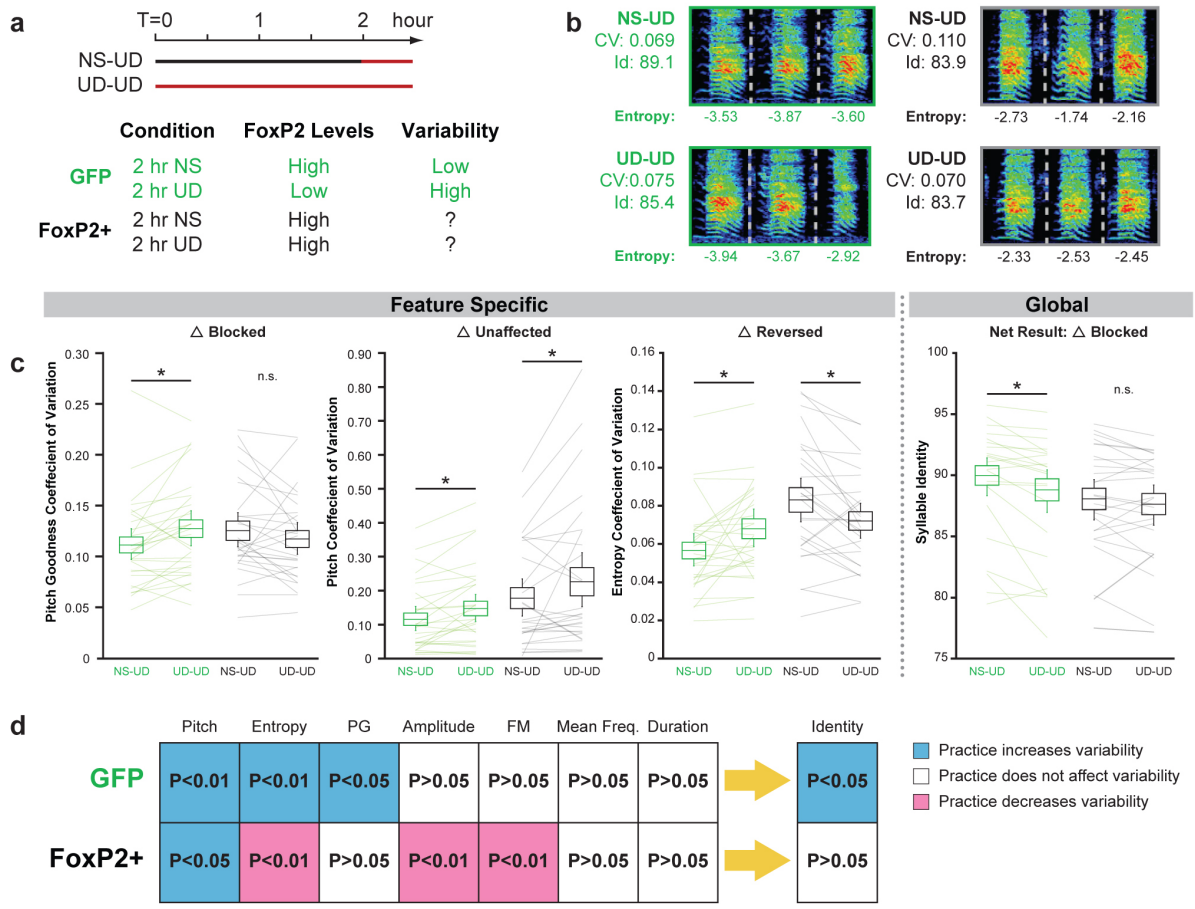
groups at the 50, 70, or 90 d time points ($p = 0.7517$, 0.6381, 0.6366, respectively; unpaired one-tailed bootstrap).

Figure 3-4: FoxP2 overexpression does not affect variability at adulthood.



a) Overall rendition-to-rendition variability of syllables, as measured by syllable identity, was not significantly different between groups (p = 0.8489, n = 8GFP/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 = 8GFP/10FoxP2+, unpaired two-tailed bootstrap). **c)** FoxP2 overexpression did not affect syntax variability (p = 0.8489, n = 7GFP/8FoxP2+, unpaired two-tailed bootstrap).

Figure 3-5. 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.

Acknowledgements

This work was supported by Grants T32MH019384, T32NS058280, UCLA Edith Hyde Fellowship (J.B.H.), RO1MH070712, 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.

Chapter IV. Reciprocal scaling of vocal variability by an avian cortico-basal ganglia circuit

Summary

Learned motor skills rely on cortico-basal ganglia loops but the relative contributions of each region to motor exploration is not well understood. We used chemogenetic interventions to separately interrogate cortical and basal ganglia sub-regions of zebra finch song control circuitry. Bi-directional manipulation of each node produced mirror image changes in vocal control with cortical activity promoting song variability and basal ganglia activity promoting song stability. The zebra finch song control system offers an opportunity to interrogate these mechanisms. The zebra finch brain contains a vocal dedicated cortico-basal ganglia loop known as the anterior forebrain pathway (AFP) which sends excitatory output to a premotor pathway required for song production as well as providing excitatory input to Area X, making the AFP a loop. Within the AFP, increasing activity within the cortical lateral magnocellular nucleus of the anterior nidopallium (LMAN) increases vocal variability whereas lesioning LMAN or decreasing its activity decreases variability. These observations suggest that LMAN functions as a variability injector but because it is the output of a larger circuit it is unclear whether this variability is intrinsic to LMAN or inherited from elsewhere.

One alternative source of variability is the basal ganglia song control nucleus, Area X, which is composed of both striatal and pallidal cell types and sits upstream of LMAN in the circuit. Lesions of Area X yield transient or inconsistent effects on song variability, questioning the role of this basal ganglia node in motor variability. Other evidence indicates that Area X mediates changes in vocal variability. For example, the instantaneous stabilization of song that occurs when a male bird performs to a female requires D1 dopamine receptor activation in Area X. Indeed, social context-induced changes in Area X spiking activity suggest it is the neural nexus

mediating this transition rather than inheriting these changes from LMAN (Woolley et al., 2014). Moreover, manipulations of the language-related gene *FoxP2* within Area X interfere with rapid social context-dependent (Murugan et al., 2013) and slower practice-dependent (Heston and White, 2015) changes in variability.

These observations raise the possibility that Area X contributes to the control of vocal exploration which can be revealed on a relatively short time scale. Here, we test this hypothesis using viral-driven expression of designer receptors exclusively activated by designer drugs (DREADDs) to transiently, bidirectionally, and independently alter the activity of LMAN or Area X neurons and measure the effect on song.

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 breeding colony, and housed in climate-controlled rooms inside cages and aviaries with a 13:11 lights on:lights off cycle including half hours of dawn and dusk. Birds had unlimited access to food, grit and water and were provided nutritional supplements and environmental enrichments.

Behavior

Adult birds (100+ days post-hatch) were moved to sound attenuation chambers (Acoustic Systems, Austin, TX) and allowed to acclimate for several days. Thereafter, procedures differed depending on whether the bird received HSV or AAV. For HSV, birds underwent song recording following pre-surgery administration of saline on day one and CNO on day two. On day three they were stereotaxically injected with HSV as described below. Following surgery, birds recovered in sound attenuation chambers for one to two days, and then underwent post-surgery behavioral test days .For AAV, birds first underwent viral surgery as described below. Three to four weeks later the bird was habituated by administration of saline. The follow day it was administered either saline or CNO and then two days later as administered CNO, or vice versa.

Stereotaxic neurosurgery was performed as in Heston and White (2015) with the following modifications:

LMAN, which was not targeted in Heston and White, was targeted at a site +5.15mm anterior, +1.6-7 lateral to the birfurcation of the midsagital sinus at a depth of 2.0 m.

The viral loads used here were 500 nl per injection site for both viruses.

Viruses

AAV- Custom designed AAVs (serotype 1) were produced by Virovek (Hayward, CA; Heston and White, 2015) containing either iDREADDs4 or eDREADDs3 driven off of the α-CaMKII promoter (AddGene). The α-CaMKII promoter was chosen because, based on other species, it was predicted to express exclusively in LMAN excitatory neurons (Jones et al., 1994; Nathanson et al.,2009) and Area X MSNs given that these are the only neurons in this nucleus which express α-CaMKII (Hein et al., 2007). In addition, vectors encoding DREADDs were obtained from the UNC Vector Core (University of North Carolina, Chapel Hill, NC). The following vectors were used: AAV5-CaMKIIα-HA-hM3D(Gq)-IRES-mCitrine (2 Area X injected birds) and AAV5-CaMKIIα-HA-hM4D(Gi)-IRES-mCitrine (2 Area X injected birds).

HSV – All HSV was obtained from the MIT Viral Core stock virus catalog. Behavioral experiments were conducted on birds that were injected with the following viruses each of which drive expression off of the IE 4/5 promoter: ST HSV-hM4Di-mCherry, ST HSV-hM3Dq-mCherry and ST HSV-mCherry. All HSV was diluted to 60-75% with saline before injection into the brain to avoid neurotoxicity and retrograde trafficking.

CNO and Saline injections

Both before and after surgery, birds were administered once with clozapine N-oxide (CNO, 150 uL at 0.1 mg/ml, i.p.; given ~15 g average weight of an adult zebra finch this is equivalent to 1 mg/kg body weight; Sigma–Aldrich, St. Louis, MO) and once, on a different day, with saline (150 uL). After each injection, the bird was prevented from singing for one hour post-injection to allow CNO activation of DREADDs. Sounds were recorded and analyzed as previously described (Heston & White, 2015).

Intrasyllable variability analysis – Analysis of intrasyllable fluctuations in FF was limited to flat or near flat syllables or sub-syllable elements. A custom software program was used to track FF across a syllable. A region of interest (ROI) was defined within the flat portion of a syllable (minimum of 20 milliseconds long) and the FF was tracked at each millisecond. Three measures were used to quantify variability across the ROI. The first was to measure the intrasyllable CV of and was obtained by calculating CV of all the 1 ms bins across the ROI. The second measure, which is operationally referred to as sweep, is proportional to the cumulative change in FF across the syllable. First, each FF string was demeaned and divided by the intrasyllable median FF so that each 1 ms time bin was expressed as a percent deviation about the median. Sweep was defined as $\Sigma(X-[X-1])^2$ where X represents the percent deviation at a given 1 ms time bin and X-1 is the percent deviation at the preceding bin. The final measure of intrasyllable variability was template variability which compared an individual syllable's fluctuations around a typical or template version of that syllable. This was calculated by obtaining 20 renditions and calculating the median percent deviation about the intrasyllable median at each time bin. The result of is a median trajectory of the FF trace and formed a template against which each individual trace

could be measured. Template variability was defined as $\Sigma(X-X_t)^2$ where X is the percent deviation of at and X_t is the percent deviation of the template. These three measures were obtained for 20 renditions of a syllable and was represented by the median of those 20 renditions.

Histological methods - To examine the efficacy in targeting and expression of viral injections, birds 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 and was carried out using immunohistological methods described by Miller et al., (2008). Brains were sectioned on a crysostat (Leica Microsystems, Bannockburn, IL) at a thickness of 30 μm.

In vivo electrophysiology recording

In vivo multiunit recordings were performed as described in Williams et al., (2012)

Spontaneous and song-evoked activity was recorded in LMAN. For each recording, 20 to 40 repetitions of the birds own song playback were played with a 10 second inter-stimulus interval. After each recording session, electrolytic lesions (+10 μA for 10 seconds) were made at the LMAN recordings site to enable histological confirmation of the recording location.

Female preference testing

Songs were evaluated by a cohort of 8 female birds that had been group housed in sound attenuation chambers and were not paired bonded. Each session consisted of a 10 minute baseline period followed by five 15 minute playback-response periods. During each of the five playback-response periods a female heard only one male. Playback of this males song included three playback periods of one song context and two of the other context (e.g.. 3 Sal and 2 CNO

or 3 directed and 2 undirected, etc) in a pseudorandom order. Each playback-response period consisted of 5 minutes of song playback in which bouts recorded in a single behavioral context were at 10 second intervals. The five minute playback period was followed by a 10 minute response period in which no songs were played back. The number of calls made by the female during the 10 minute baseline and response periods were counted. Changes in singing behavior evoked by each song time were quantified using the formula (post - pre)/(post + pre + 2). From this, the median change in calling behavior was calculated yielding a calling index for each song type. From this the CNO- or directed song-induced change was calculated as (CNO index – Sal index) or (Directed index – Sal index). Sessions were excluded if the female failed to call in three or more of the six sessions or failed to call at least 30 times across the six epochs (1 baseline and 5 response periods). Each male's song was evaluated by four females and the CNO-induced change in female response was derived from the median of those four sessions.

Statistics - Resampling statistics were used throughout our analysis. The unpaired test begins by calculating the difference in-group means. This value represents the test statistic, M. We then created pseudo data sets 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-Ms that was as large or larger than the actual M was determined and this number divided by 10,000. This value reflects the reported p-value.

Our first statistical test compared the modulation index of iDREADDs4 vs eDREADD3 for each measure using a two-tailed test. If a significant effect was found a one-tailed test was used to compare each group to control.

Results

Adult birds were injected bilaterally in either LMAN or Area X with a virus which expressed either eDREADDs3 that increases neural activity or iDREADDs4 that decreases neural activity (Armbruster et al., 2007) following administration of the DREADDs ligand clozapine-N-oxide (CNO). For the majority of birds, an AAV that utilized the α-CaMKII promoter was used in order to drive receptor expression in LMAN excitatory projection neurons (Jones et al., 1994; Nathanson et al., 2009) or Area X medium spiny neurons (MSNs; Hein et al., 2007). A minority of birds were injected with a herpes simplex virus (HSV) that utilized a non-specific promoter. We verified that these viruses targeted relevant cell types in each nucleus (Supp. Fig. 4-1) and that they could be used to alter neural activity in the predicted manner (Fig 4-1b). Having validated our methodology, we turned to analyzing the behavioral effects. Song syllable variability was measured following systemic administration of CNO on a given day and compared to that following saline administration on another in order to obtain an effect size. Control song syllables were obtained from birds injected with a DREADDs construct but that were administered saline (n=14) and birds injected with virus encoding mCherry alone (n=3). Three types of variability were assessed: rendition-to-rendition (intersyllable) differences in the way a syllable was performed, moment-to-moment (intrasyllable) fluctuations in fundamental frequency (FF), and syllable sequencing (syntax).

As a first behavioral validation of our methodology, we examined rendition-to rendition song syllable variability which LMAN is well established to positively regulate (Woolley and Kao, 2015). As predicted, activation of eDREADDs3 (n=6) in LMAN projection neurons led to an increase in syllable variability whereas activation of iDREADDs4 (n=6) led to a decrease

across numerous measures of intersyllable variability. This effect did not differ between HSV and AAV-CaMKII injected birds (Supp. Fig. 4-2) and altered the variability of syllable features but not the means of those features.

Intrasyllable variation as fluctuations of FF were also measured. This type of variability is thought to represent ultrafast motor exploration and can be used to guide learning on a millisecond timescale (Charlesworth et al., 2011) but its neural underpinnings remain unexplored. Similar to intersyllable variability, LMAN positively and bidirectionally regulated this type of variability. Activation of eDREADDs3 increased FF fluctuations whereas activation of iDREADDs4 decreased each of three different measures of intrasyllable variability. Interestingly this effect was driven primarily by HSV injected birds (Supp. Fig. 4-3) There was no effect of LMAN manipulations on syntax variability.

These analyses were then applied to Area X for which a role in variability is less clear. To our surprise, Area X negatively regulated variability in a manner reciprocal to LMAN. Activation of iDREADDs4 (n=9) on Area X MSNs increased numerous measures of intersyllable variability whereas activation of eDREADDs3 (n=7) increased intersyllable variability (Fig. 4-2), albeit these changes were limited in magnitude compared to LMAN manipulations. This effect did not differ between viral types (Supp. Fig. 4-4). Moreover, unlike LMAN, Area X manipulations did affect one mean feature: activation of eDREADDs3 decreased syllable entropy (they became more harmonic) whereas activation of iDREADDs4 led to a non-significant increase in entropy (data not shown).

Interestingly, activation of iDREADDs4 led to enhanced intrasyllable variability, but activation of eDREADDs3 was without effect on this measure. Thus, in terms of intrasyllable variability, Area X acted as a unidirectional variability suppressor. As with LMAN, this effect

was driven primarily by HSV injected birds (Supp. fig. 4-5). Similar to LMAN, no effect on syntax variability was detected.

To determine whether these changes were of potential ethological relevance, we tested whether female zebra finches could detect DREADDs induced alterations in song variability. Female calling was measured following playback of the songs of males used in the above experiments. It was first established that females could detect the difference between the stereotyped song of a given male singing in the presence of a female (termed directed song) and his more variable song recorded when he sang alone (undirected), as previously demonstrated (Woolley and Doupe, 2008). Female calling following song playbacks increased with the magnitude of the vocal variability in the male song. On the basis of this result, it was predicted that females would call more following song playbacks from variability increasing manipulations which recapitulate the lower stereotypy of undirected song (excitation of LMAN, inhibition of Area X MSNs) and call less following variability decreasing manipulations (inhibition of LMAN, excitation of Area X). Indeed, females called more following playback of songs from males with LMAN excited, although there was no effect of song playbacks from males with LMAN inhibited. Also in line with our prediction, playback of songs from males with Area X inhibited or excited song led to increased and decreased female calling, respectively.

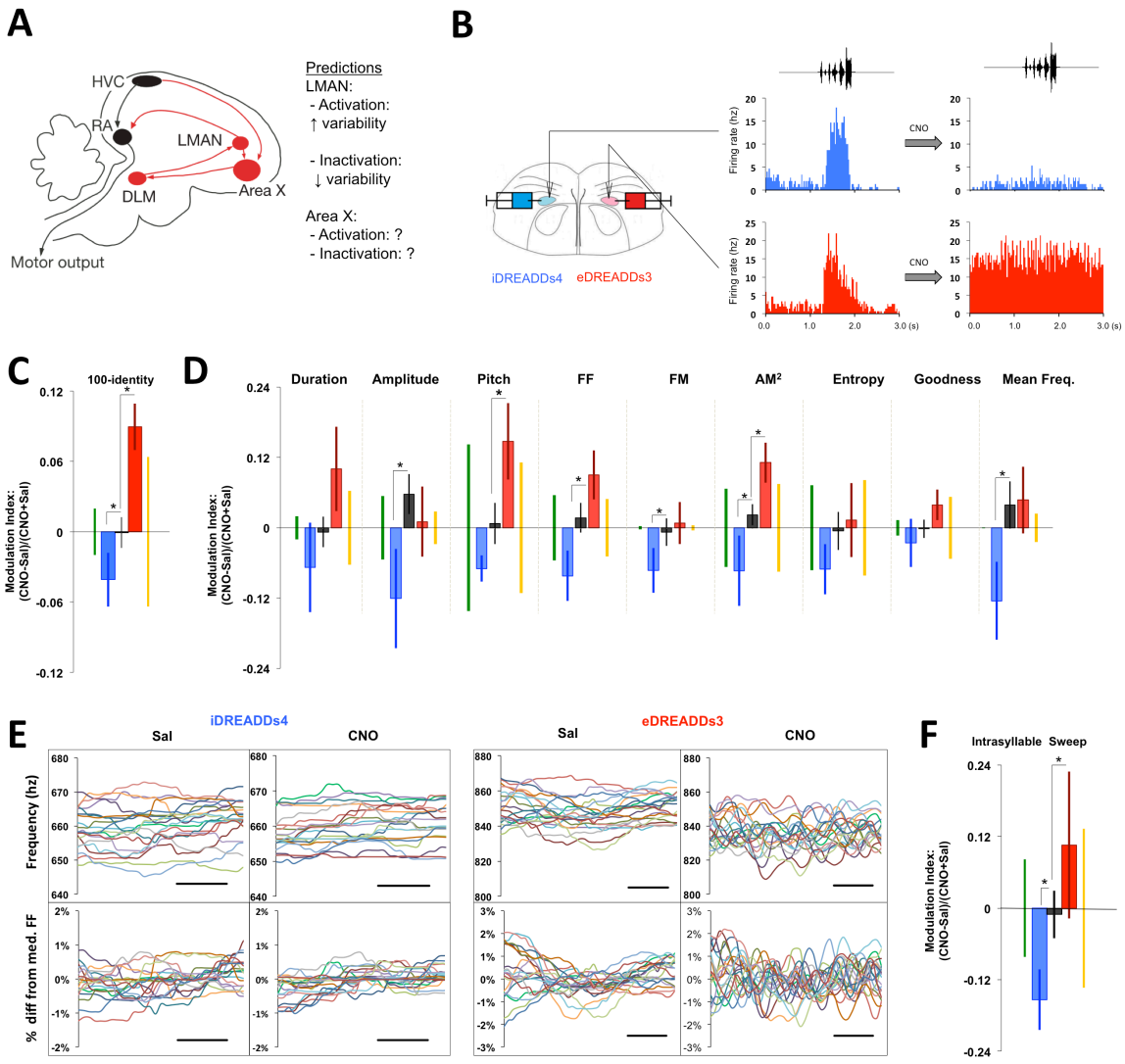
In sum, our findings indicate that LMAN and Area X can act as dual regulators of vocal variability. These results reinforce the notion that LMAN positively scales vocal variability, and they provide novel evidence implicating Area X as a reciprocal regulator of variability. Moreover, both nuclei can act at the millisecond timescale suggesting a role in the moment-to-moment execution of motor commands. A second finding is that Area X regulates variability in a manner opposite to LMAN. This result was somewhat unexpected given that, like LMAN, Area

X is thought to fire more during variable undirected singing relative to directed song (Woolley and Doupe, 2014), and that it receives monosynaptic excitatory input from LMAN. On the other hand, this finding is consistent with the observation that D1 receptor activation in Area X increases MSN excitability and decreases intersyllable variability (Ding and Perkel, 2002; Leblois and Perkel, 2012). In rats, outright silencing of the dorsolateral striatum increases the variability of execution of a well-learned motor task (Rueda-Orozco and Robbe, 2015), suggesting that the output of the striatum promotes motor stability rather than motor exploration.

One hypothesis that could reconcile these disparate observations is that LMAN co-engages Area X as a negative feedback mechanism to adaptively temper and control variability. Thus, Area X would be more active when LMAN is highly active, and therefore song more variable, yet manipulations of Area X by itself would reveal it as a variability suppressor.

Finally we show that female birds can detect DREADDs induced changes in male song in a manner consistent with how they respond to naturally occurring changes in vocal variability. This suggests that the changes induced here in the vocal motor system capture some of the natural changes that occur in male song variability and give insight into the manner that these changes endogenously occur.

Fig 4-1. Validation of methodology in LMAN confirms it is a positive regulator of vocal variability



(A) The zebra finch brain contains two interconnected pathways controlling vocal motor output. One is the posterior pathway consisting of the cortical nuclei HVC (used as a proper name) and the robust nucleus of the arcopallium (RA), and brain stem motor neurons that control vocal output. This pathway is essential for the execution of learned vocalizations. The second pathway is the anterior forebrain pathway (AFP), a song-dedicated basal-ganglia-thalamo-cortical circuit, consisting of HVC, the striato-pallidal nucleus Area X, the dorsolateral thalamic nucleus (DLM),

and the cortical nucleus LMAN which sends bifurcating axons to both Area X as well as RA, thereby allowing this pathway to affect vocal output. Area X is predominantly composed of striatal medium spiny neurons and also contains pallidal interneurons and projection neurons. It is well established that LMAN positively regulates vocal variability but it has been unclear whether Area X plays a role in controlling variability.

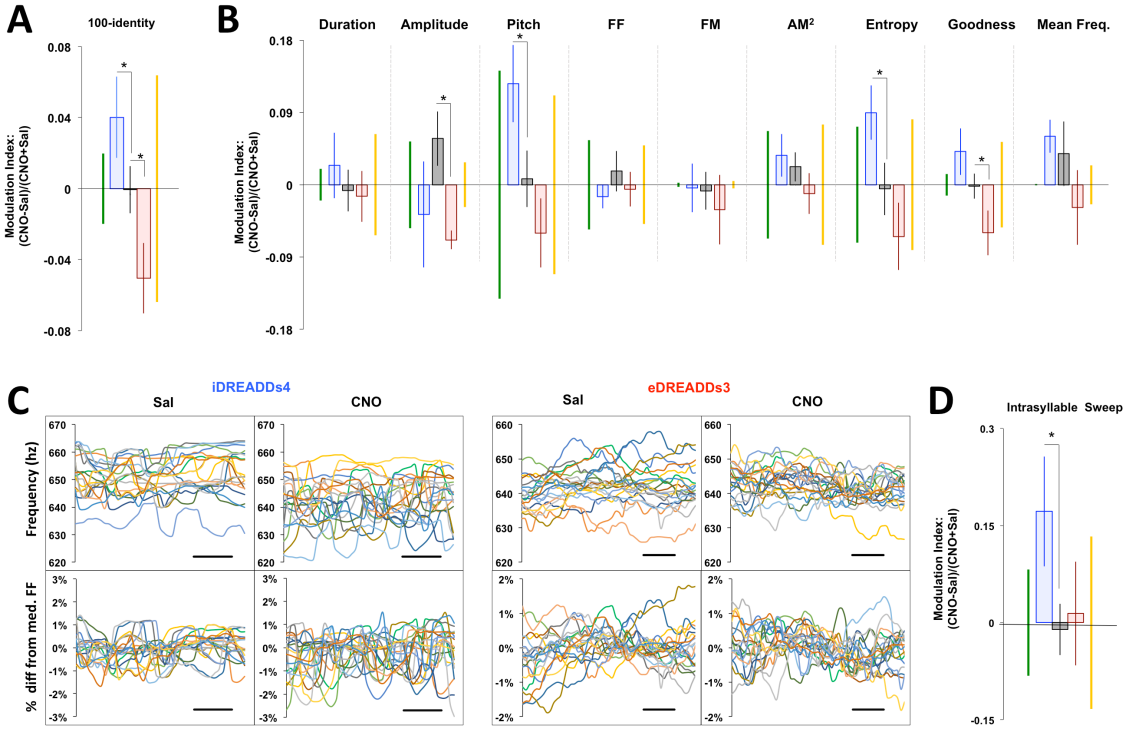
(B) LMAN neural activity in response to a bird's own song playback was recorded in one bird that was injected with iDREADDs4 in Area X of one hemisphere and eDREADDs3 in the other. Following administration of CNO, neural activity is attenuated on the iDREADDS4 side and enhanced on the eDREADDS injected side.

(C) Intersyllable variability of song syllables as measured by syllable identity decreases or increases following inhibition (via iDREADDs4; red) or excitation (via eDREADDs3; green) of LMAN, respectively. Increases in variability are upward on the y axis for this and all other measures of variability. Ethologically relevant windows of variability are indicated by bars which show the average range over which each measure varies as a function of social context (green) or practice (orange). See methods for further explanation. Asterisks indicate $p<0.05$ relative to controls using one-tailed unpaired resample test throughout.

(D) CNO-induced activation of iDREADDs4 versus eDREADDs3 in LMAN produces bidirectional changes in intersyllable variability across multiple feature specific measures of variability. For each measure shown here, in E and F below, and similar panels in Figure 3, ethologically relevant windows of variability are indicated by blue (social context-dependent) or orange (practice-induced) bars which flank values for iDREADDs4 (blue), control (black) and eDREADDs3 (red), respectively.

(E) Traces show raw (top) and demeaned FF (bottom) contours from representative flat harmonic syllables. Inhibition of LMAN via CNO-induced activation of iDREADDs4 leads to a convergence and flattening of demeaned FF traces relative to the saline condition. Intrasyllable fluctuations in FF are thereby reduced. Conversely, excitation of LMAN via CNO-induced activation of eDREADDs3 leads to a divergence and increase in the fluctuations.
(F) Similar to intersyllable variability, positive, bidirectional control of intrasyllable variability by LMAN as measured using intrasyllable sweep.

Fig 4-2. Area X negatively regulates intersyllable and unidirectionally affects intrasyllable variability



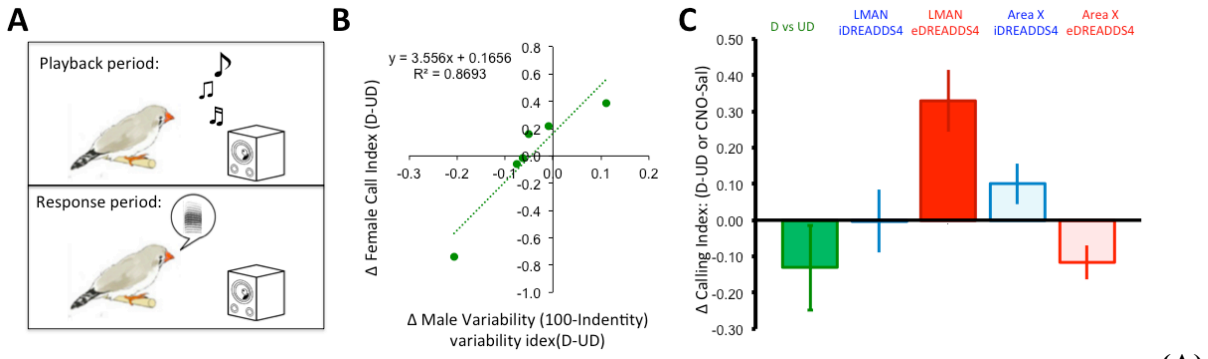
(A) In contrast to LMAN, intersyllable variability of song syllables as measured by syllable identity increases or decreases following inhibition or excitation of Area X, respectively.

(B) CNO-induced activation of iDREADDs4 versus eDREADDs3 in Area X produces bidirectional changes in intersyllable variability across multiple feature specific measures.

(C) Inhibition of Area X via CNO-induced activation of iDREADDs4 led to an increase of intrasyllable FF fluctuations. Excitation of Area X via CNO-induced activation of eDREADDs3 had no apparent effect on these fluctuations.

(D) Similar to rendition-to-rendition variability, CNO-induced activation of iDREADDs4 in Area X led to enhanced intrasyllable variability. Unlike previous measures, however, CNO-induced activation of eDREADDs3 was without effect.

Fig 4-3. DREADDs-induced changes in song can be detected by female zebra finches



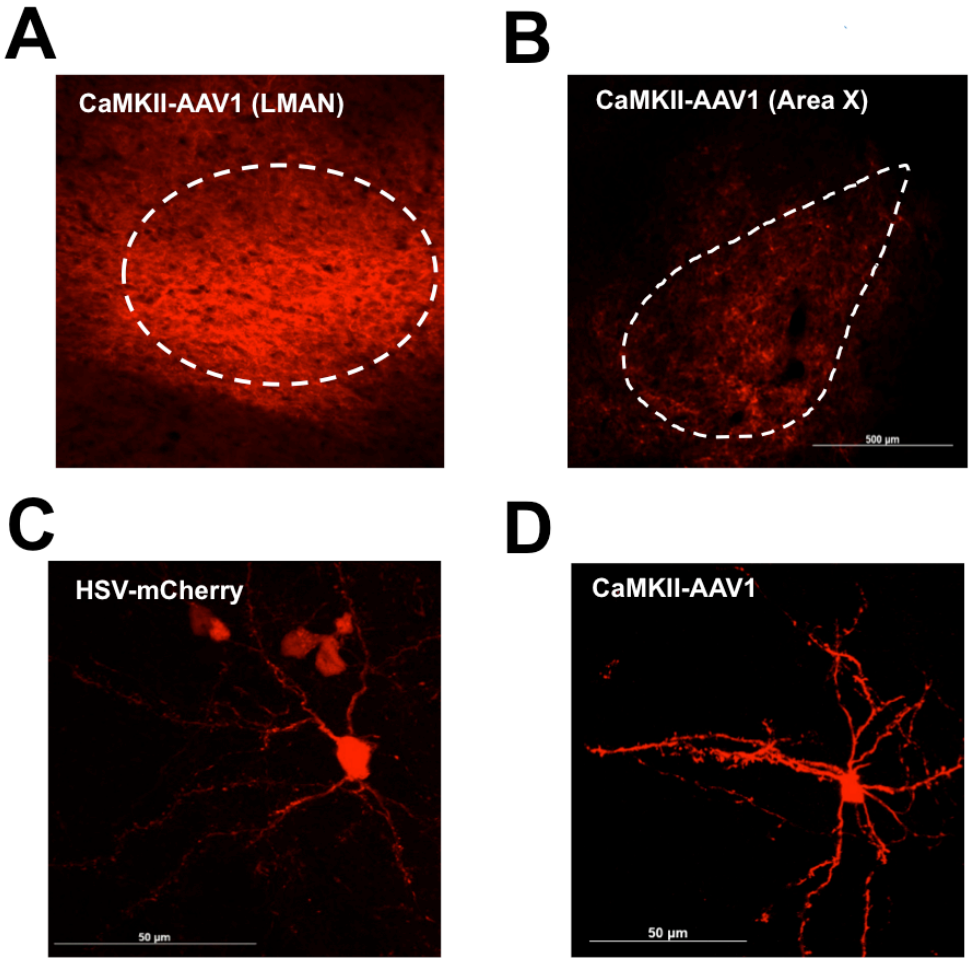
(A)

Female birds heard playbacks of directed versus undirected songs or of songs from DREADDS-injected males administered saline versus CNO. Songs from DREADDs injected males were presented in 5 blocks of 30 playbacks. Each playback was followed by a ten-minute period of silence during which females calls were recorded.

(B) Female call amount changes following playback of directed vs undirected song positively correlated with male vocal variability changes in those two conditions.

(C) Females called more after playback of songs from males in which LMAN was excited but no change was observed when LMAN was inhibited. In contrast, females called more following playback of songs from males in which Area X was inhibited and less when Area X was excited.

Supplemental Fig. 4-1. Viral expression of HSV and CaMKII-AAV in LMAN and Area X

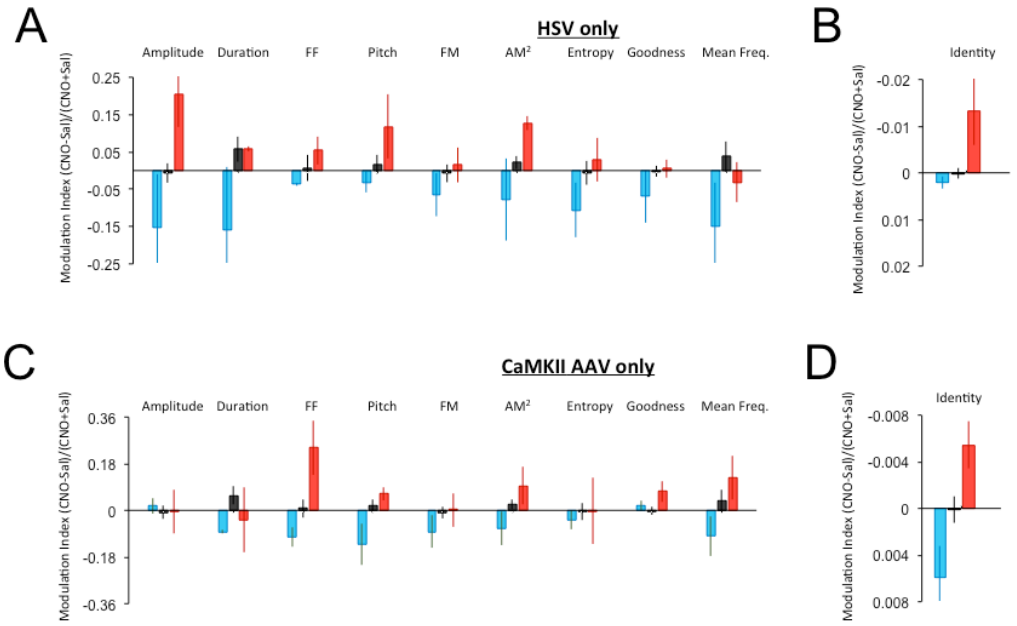


(A) Viral expression of mCherry can be observed in LMAN following injection of CaMKII-AAV into this site.

(B) Viral expression of mCherry can be observed in Area X following injection of CaMKII-AAV into this site.

(C-D) Both neurons transfect neurons in Area X that morphologically appear to be MSNs

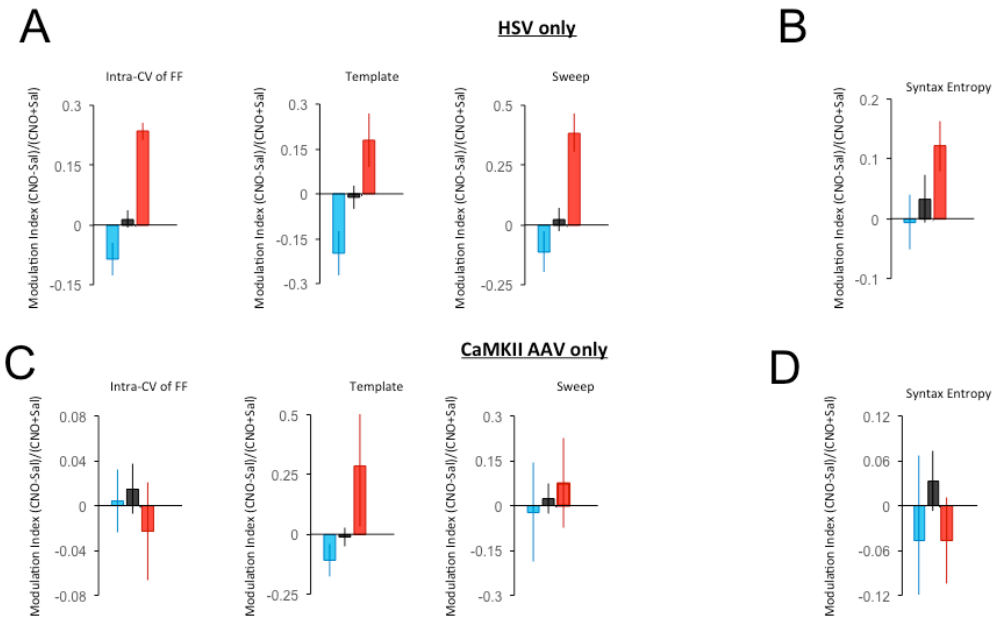
Supplemental Fig 4-2. Intersyllable effects of LMAN manipulations are independent of viral type.



LMAN injected birds were separated by viral type.

(**A-D)** HSV and AAV CaMKII injected birds show the variability injector pattern on measures of intersyllable variability that was seen in combined data.

Supplemental Fig 4-3. Intrasyllable and syntax effects of LMAN manipulations are dependent on viral type.



LMAN injected birds were separated by viral type.

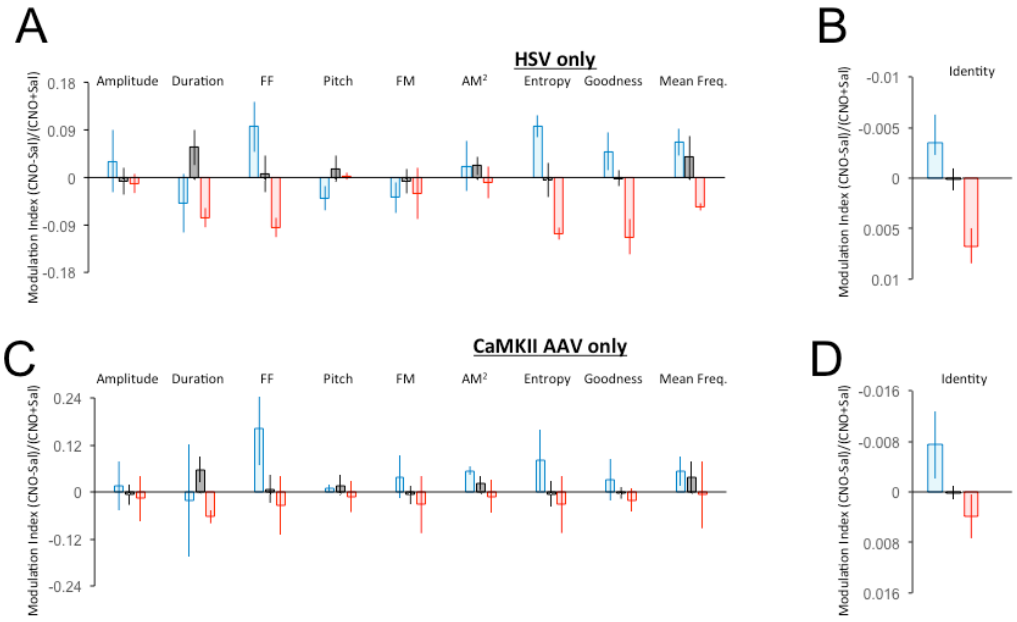
(A) HSV injected birds showed the intrasyllable variability injector pattern observed in the combined data.

(B) HSV injected birds showed a syntax entropy variability injector pattern observed not apparent in the group data.

(C) CaMKII –AAV injected birds showed a semblance of a variability injector pattern but it was clearly weaker than HSV injected birds.

(D) No effect on syntax entropy was observed in CaMKII-AAV injected birds and this might have masked any effect in HSV injected birds.

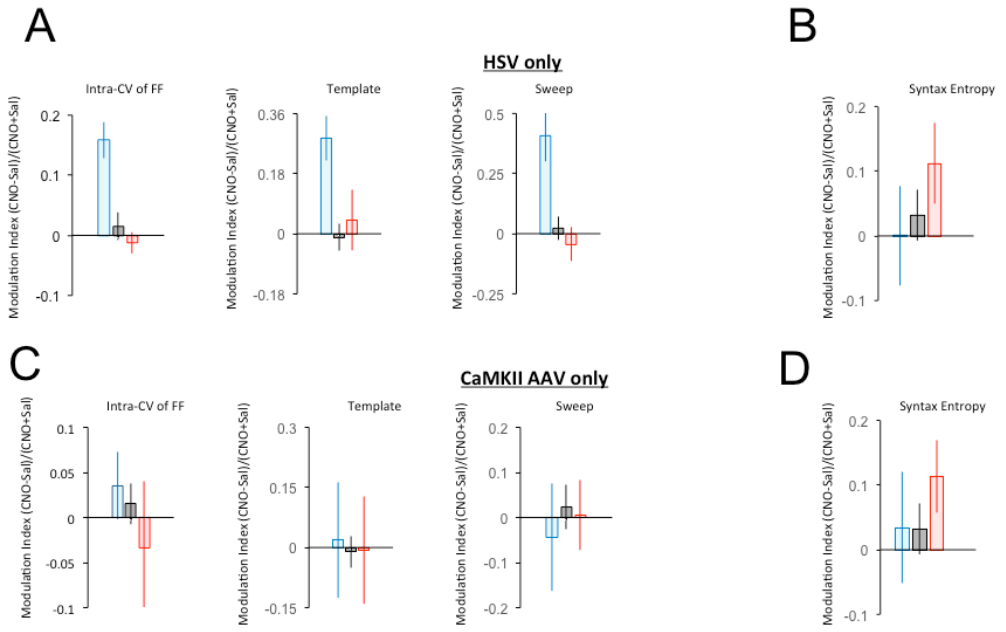
Supplemental Fig 4-4. Intersyllable effects of Area X manipulations are independent of viral type.



Area X injected birds were separated by viral type.

(A-D) HSV and AAV CaMKII injected birds show the stability injector pattern on intersyllable variability that was seen in combined data.

Supplemental Fig 4-5. Intrasyllable and syntax effects of Area X manipulations



Area X injected birds were separated by viral type.

(A) HSV injected birds showed the intrasyllable stability injector pattern observed in the combined data.

(B) No effect was observed on syntax entropy in HSV injected birds.

(C) No effect of Area X manipulation on intrasyllable variability was detected in CaMKII-AAV injected birds

(D) No effect on syntax entropy was observed in CaMKII-AAV injected birds.

Chapter V. Conclusions

The broad aim of my dissertation research has been to use viral manipulations of the zebra finch song circuit to interrogate the importance of FoxP2 singing-driven downregulation and changes in cortico-basal ganglia function in vocal motor learning and control. Following an introductory Chapter I, in Chapter II, I describe efforts to characterize different virus types according their usefulness in delivering genes to Area X. In Chapter III, I describe my investigations of FoxP2 and provide evidence that online FoxP2 regulation plays a critical role in learning and practice-dependent modulation of variability. In Chapter IV I used DREADDs mediated manipulations of neural activity to provide a candidate neural mechanism by which practice could lead to increased variability, and give insight into how FoxP2 overexpression could disrupt this change. Finally, in the appendix I present collaborative work that shows FoxP2 practice-dependent downregulation and changes in variability are present in at least one other species of songbird, suggestion that principles gleaned here on probably not unique to zebra finch. In the appendix I show also show that in addition to the central mechanism of vocal control that I explore throughout my dissertation, that peripheral control of vocal integrity is also very important.

Based on these experiments, I propose a model in which vocal practice leads to FoxP2 downregulation which in turn leads to diminished Area X medium spiny neuron (MSN) firing and is behaviorally manifested as enhanced vocal variability (Fig. 5-1) The decreased neuronal firing could be due to decreased intrinsic excitability, as was manipulated in the DREADDs experiments, but could also be the result of decreased synaptic strength either pre- or post-synaptically.

Currently the most plausible explanation for this decreased activity is a DARPP-32-associated mechanism. DARPP-32 is the downstream integrator of both D1- and D2-receptor

signaling. D1-receptor activation leads to enhanced MSN excitability by increasing DARPP-32 activity (Onn et al., 2003). Because D1-receptor activation is both necessary and sufficient to mediate the switch to high variability undirected song Leblois and Perkel, 2012), DARPP-32 by necessity plays a critical role in maintaining vocal stereotopy.

Several observations link FoxP2, vocal variability, and vocal practice to DARPP-32. First, DARPP-32 is a transcriptional target of human FoxP2 (Vernes et al., 2011) and viral knockdown of FoxP2 in Area X abolishes DARPP-32 levels and social context-dependent changes in variability as well as decreases D1R expression (Murugan et al., 2013). This manipulation also makes the AFP unresponsive to D1-agonists and renders the bird unable to switch to low variability directed song. Moreover, iDREADDs3 activation in Area X, which I showed leads to enhanced variability, utilizes the very same G_i protein mediated signaling cascade that D2 receptors use to dephosphorylate DARPP-32 (Bateup et al., 2008). It is thus possible that the DREADDs-mediated effect observed here works through this same mechanism. Together, these observations indicate that singing leads to downregulation of a molecule that promotes DARPP-32 levels and that the behavioral consequences of vocal practice can be recapitulated by activating a signaling pathway that leads to decreased DARPP-32 function. If accurate, this mechanism has implications for the cell's ability to undergo plasticity as DARPP-32 is necessary for postsynaptic forms of both striatal potentiation and depression.

While this model links together the findings of Chapters III and IV, two major questions remain. First, does Area X actually undergo a FoxP2-dependent change in neuronal firing? Second, what function does FoxP2 downregulation and its potential decrease in neuronal activity serve?

Does Area X actually undergo a FoxP2-dependent change in Area X MSN firing and excitatory drive?

The first question is actually two separate questions: 1) are there in fact practice-dependent changes in neurophysiology that accompany extended vocal practice and can explain the acute increase in vocal variability, and 2) if these changes exist, do they depend on FoxP2 downregulation? For each question I suggest in vivo and ex vivo means for addressing them.

Ex vivo approaches

The ex vivo approach to address these sub-questions is to use acute slice physiology. One could overexpress FoxP2 and GFP or GFP by itself using the HSV described in Chapter II and measure features of cellular physiology. The experimental schemes for doing so are highlighted in Fig. 5-2. On the day of the experiment, the bird would be allowed to sing in order to allow endogenous practice related changes in physiology to occur or to be disrupted by the overexpression of FoxP2. One could then patch onto transfected cells or use field recordings to record from GFP rich regions of Area X. The following findings would support the model put forth earlier: 1) GFP-only transfected cells would show low excitatory drive as indicated by low levels of cell excitability, presynaptic excitatory drive, or postsynaptic excitatory drive whereas FoxP2-GFP transfected cells will show high excitatory drive, 2) GFP transfected cell properties will not differ from their neighboring untransfected cells whereas FoxP2-GFP cells be more excitable than their untransfected neighboring cells, 3) excitatory drive will negatively correlate with the amount of singing in GFP and untransfected cells whereas it will positively correlate in FoxP2-GFP cells.

An alternative approach would be to compare cell physiology between birds after two hours of UD singing vs two hours of non-singing (NS). This approach has the technical

advantags that it does not require a surgery and all cells recorded would be “experimental cells”. Moreover, this approach could be used to assess any practice dependent changes in LMAN function in the same slice, an important concern given that the acute increase in variability following vocal practice could also be explained by increased LMAN activity. The disadvantages of this approach are that observed changes in physiology would not necessarily be linked to changes in FoxP2 levels.

Both approaches, especially the second one, rely on carrying in vivo changes in physiology forward to the ex vivo slice. A first glance this seems unlikely given the large number of changes in cell physiology that accompany sectioning a brain and preparing slices for ex vivo recording (Ho et al., 2004). On the other hand, a large number of studies have used similar approach including in the basal ganglia (Yin et al., 2009) and in the zebra finch brain (Huang et al, 2008). Nevertheless, this concern is valid and may complicate any ex vivo approach.

In vivo approaches

An in vivo approach towards addressing these sub-questions is to examine singing-related neural activity in Area X after UD singing vs after NS. Probably the optimal method is to examine neuronal spike activity from a bird implanted with a chronic electrode in Area X. This technique may be beyond the means of many labs so instead I suggest using immediate early gene (IEG) expression as a proxy for neural activity, a technique well within the scope of many labs. One could examine song induced IEG expression in a group of birds who had spent the previous two hours singing and a group that had spent the previous two hours not singing. If there is a practice-dependent decrease in neural activity then one might expect lower levels of IEG induction in the bird that had spent the previous two hours singing.

The obvious problem with this approach as stated is that the two hours of singing would itself lead to gene induction that would obscure later waves of gene expression. This issue could potentially be circumvented by using Cellular Compartment Analysis of Temporal Activity by Fluorescence In Situ Hybridization (catFISH). This technique uses subcellular imaging of the mRNA for an IEG (typically *arc* could in principal be used with other IEGs such as *c-fos* or *zenk/egr-1*) and takes advantage of the fact that mRNA is slowly transported from the nucleus to the cytosol. Thus the subcellular compartment in which the mRNA is found marks the relative timing of neural activity. Neurons active shortly before the animal is sacrificed are marked by the presence of intense intranuclear mRNA foci whereas cytoplasmic RNA signal is used to identify neurons active ~30 to 45 min earlier. Therefore if one separates the two hours of UD singing or NS by ~30 mins from a brief ten min bout of singing then one can examine the neural activity induced by the brief bout of song by focusing exclusively on cells expressing intense intranuclear mRNA foci. An outline of this experiment is shown in Fig. 5-3

In this experiment I predict that birds that have spent the previous two hours singing will have less cells expressing intranuclear IEG mRNA foci than birds that have spent the previous two hours not singing. Further, I suggest that the levels of intracellular foci will negatively correlate with amount of singing in the two hour period and potentially with the levels of cytoplasmic IEG mRNA as the latter would be a proxy for the neural activity induced during the two hour period of singing.

If this approach works and a practice-dependent change in intranuclear foci is detected, then a role for FoxP2 down regulation could be test by measuring whether or not the observed pattern is affected by FoxP2 overexpression. One would need to overexpress FoxP2 along with GFP to identify on a cell-by-cell basis whether a particular neuron is overexpressing FoxP2. As

outlined in Chapter II, because of its large cloning capacity and because this would not be a long-term experiment, HSV is ideally suited for this task. The control virus would be HSV that expresses GFP by itself or along side an inert molecule such as lacz. The catFISH experiment outlined above could be repeated as above but in bird injected bilaterally with one of the viruses or even unilaterally with each virus. The expectation is that if FoxP2 downregulation is necessary for the proposed decrease in firing then FoxP2 overexpressing cells would be more active and express more intranuclear foci than GFP transfected cells as well as untransfected neighboring cells.

What function does FoxP2 downregulation serve?

In regards to the second big question, what function FoxP2 downregulation and its potential decrease in neuronal activity serve, I suggest two possibilities: it primes the circuit to undergo learning and plasticity or it limits learning and plasticity.

In the first scenario, FoxP2 acts a plasticity repressor and its downregulation leads to changes in target gene expression and cell physiology that in some way promote or prime Area X for learning related plasticity. In support of this interpretation, Deregnaucourt et al, (2005) showed that the peak learning rate occurs after a period of ~two hours of morning singing, a time point at which we have shown is associated with low FoxP2 expression. The FoxP2-dependent increase in variability following vocal practice is also telling. As mentioned earlier, studies in songbirds and humans indicate that variability positively predicts the rate, accuracy and even capacity for learning (Deregnaucourt, 2005; Tumer and Brainard, 2007; Sober and Brainard, 2014; Wu et al, 2014). Moreover, recent evidence in humans suggests that motor variability and

the existence of sensorimotor mismatches are necessary for the engagement of reconsolidation processes and the enhancement of motor skill performance (Wyms et al., 2016).

In a second scenario, FoxP2 is a plasticity promoter and its downregulation serves to limit further plasticity. In this scenario, the bird sings and undergoes learning early in the day leading to a downregulation of FoxP2 and as FoxP2 levels decrease so does the ability of the circuit to undergo learning related plasticity. There is some support for this idea. From a behavioral perspective, while Deregnaucourt et al., (2005) do show that learning rate peaks after about two hours, they also show that after more extended vocalizations, singing no longer leads to increases in similarity to tutor suggesting that the circuit becomes implastic as the bird sings more. From a neural plasticity standpoint, normal FoxP2 function in mice is necessary for corticostriatal LTD and motor skill learning (Groszer et al., 2008) thus it is possible that there is a tandem decrease in synaptic strength and the molecule necessary for that decrease that plasticity.

At the core of this conundrum is whether FoxP2 promotes or represses learning and, conversely, whether low FoxP2 levels repress or promote learning, respectively. The fact that knockdown and overexpression, two opposing manipulations, lead to convergent learning phenotypes (Haesler et al., 2007; Heston and White, 2015) speaks to the difficulty of disentangling these two possibilities in a learning paradigm during which numerous cycling episodes occur and a delicate balance between plasticity and consolidation must be achieved for normal learning. There are a few experiments that may be able to disentangle these possibilities, both relying on changes in auditory feedback.

One way to address this relies on the paradigm of pitch contingent auditory feedback (PCAF; Andalman and Fee, 2009). In this paradigm, the fundamental frequency of a flat

harmonic syllable can be increased or decreased by delivering rapid white noise feedback to renditions of a syllable that fall below or above, respectively, an experimentally controlled frequency threshold. This learning depends on Area X (Ali et al, 2013) and, critically, can occur in a single day. Thus, one can test whether learning is affected FoxP2 levels per se and disentangle this from the overall requirement for FoxP2 cycling. So for example, the ability of a bird to undergo this learning could be evaluated when the white noise feedback program is engaged after two hours of UD singing (when FoxP2 levels are low) and again on a separate day after two hours of NS (when FoxP2 levels are high). If low levels of FoxP2 facilitate learning and plasticity then the bird should undergo more rapid learning following two hours of singing as compared to after two hours of NS, whereas the exact opposite would be expected if FoxP2 downregulation promotes stability and consolidation. Moreover, if this differential learning rate is indeed identified then one could test whether FoxP2 is the critical molecule mediating this shift by overexpressing FoxP2 in Area X and repeating this experiment. Under these conditions, if FoxP2 is critical for switching between low and high plasticity then by preventing the downregulation one would prevent this switch. This experiment is outlined in more detail in Fig. 5-4

There are at least three potential confounds that I foresee but I believe can be circumvented. One potential confound is that because the variability of song differs following two hours of UD versus two hours of NS and variability predicts learning by more effective exploration of motor space, there may be a difference in learning due to variability rather than a change in the internal plasticity state of the circuit. This could be circumvented by negatively reinforcing all renditions that fall on one side of the median or the other, rather than using some escape target that is a certain distance from the mean. Under these conditions enhanced

variability would have little effect because half of all renditions would be reinforced regardless of the amount of variability.

A second confound is that the bird might undergo different amounts of learning but as a result of differing amounts singing rather than different plasticity states in the two conditions. In this scenario, the bird might be more vocally exhausted after singing for two hours and will sing less after the learning epoch has begun, leading to lower levels of learning. One way to circumvent this is to calculate the average amount of learning per motif similar to the way Deregnaucort et al., (2005) calculated record gain in similarity to tutor, thus normalizing the data by the amount of singing. Alternatively, one could measure the total amount of learning that has occurred by a certain number of iterations of the target syllable. So for example, one could measure the amount of learning that has occurred by the 2,000th iteration of the target syllable rather than the total amount of learning that has occurred by the end of the day.

A final potential confound is that the bird may learn more in the second learning episode regardless of whether it is post NS or post UD as a simple consequence of savings, the residual effect of previous learning (Ebbinghaus, 1885). To address this confound, one should first establish whether this does indeed occur by simply testing a bird under the same condition twice. If savings, as expressed by enhanced learning on the second episode, are not apparent then one need not be too concerned with this possibility. If savings are, however, apparent perhaps this could be mitigated by having the two learning episodes distinguished by the distinct in the direction learning (i.e. if the first learning episode pushed FF upward then the second one will push it downward and vice versa). This may mitigate and savings confound, but could be used even if one fails to detect any evidence of savings.

Another way to better understand FoxP2 function is to examine maladaptive plasticity following the loss of auditory feedback. When a zebra finch is deafened it undergoes a gradual deterioration of song quality known as decrystallization (Brain and Doupe, 2000). This form of maladaptive plasticity is eliminated by lesions to LMAN (Brainard Doupe, 2000) or Area X (Kojima et al., 2013) suggesting it relies on plasticity in the AFP. Unlike this song learning, this form of learning is not thought to reflect matching a targeting or template, but is instead aberrant drift of vocal motor patterns (Brainard and Doupe, 2000). It therefore, may not rely on the delicate balance of plasticity and consolidation that is necessary for template matching in song learning. Thus, increasing plasticity will increase the rate of this drift and decreasing plasticity will decrease the rate of drift.

This experiment is currently underway in our laboratory and preliminary results indicate that FoxP2 overexpression accelerates decrystallization. One interpretation of this finding is that FoxP2 overexpression leads to a hyperplastic state. This maybe due to enhanced D1R-DARPP-32 signaling that is maintained in FoxP2+ birds is shut down in a song-dependent manner in control birds.

The link between physiology and function

The preceding section presented two independent questions and provided several experiments for independently addressing them. In reality, however, the answers obtained to address one question speak to the questions presented in the other. For example, if the physiological experiments support the proposal that vocal practice leads to a FoxP2-dependent decrease in MSN firing, then it would suggest that FoxP2 downregulation serves to decrease plasticity as most forms of plasticity require spiking and excitability promotes plasticity in numerous brain regions including the basal ganglia (Hopf et al., 2010). Conversely, for the same

reason, finding that FoxP2 downregulation leads to increased plasticity would suggest that the model of decreased MSN spiking is probably incorrect.

Loss of an adaptive signal or gain of a maladaptive signal?

Another unresolved question is whether the deficits in FoxP2+ birds are the result of the absence of an adaptive/signal or the presence of a maladaptive/disruptive signal. In human Parkinson's disease (PD) patients and rodent PD models, surgical ablation of large portions of the internal globus pallidus, the output of the basal ganglia, is an effective treatment for striatal-associated disorders such as PD and dystonia. This has led to the conclusion that it is "better to block [basal ganglia] output completely than allow faulty signals from the [basal ganglia] to pervert the normal operations of motor areas that receive [basal ganglia] output" (Turner and Desmurget, 2010). One way to resolve whether FoxP2 overexpression in Area X leads loss of an adaptive signal or gain of a faulty signal out would be to compare the effects of unilateral GFP or FoxP2 overexpression to unilateral lesions. The GFP bird would be expected to show normal song learning and normal practice-induced increases in variability. The unilaterally lesioned bird may or may not show impairments in either of these behaviors. If overexpression leads to the loss of an adaptive signal, then FoxP2+ animals should behave like lesioned animals because in both cases that adaptive signal would be lost. If, on the other hand, FoxP2 overexpression leads to the generation of a faulty signal then unilateral FoxP2 overexpression should be worse than a unilateral lesion.

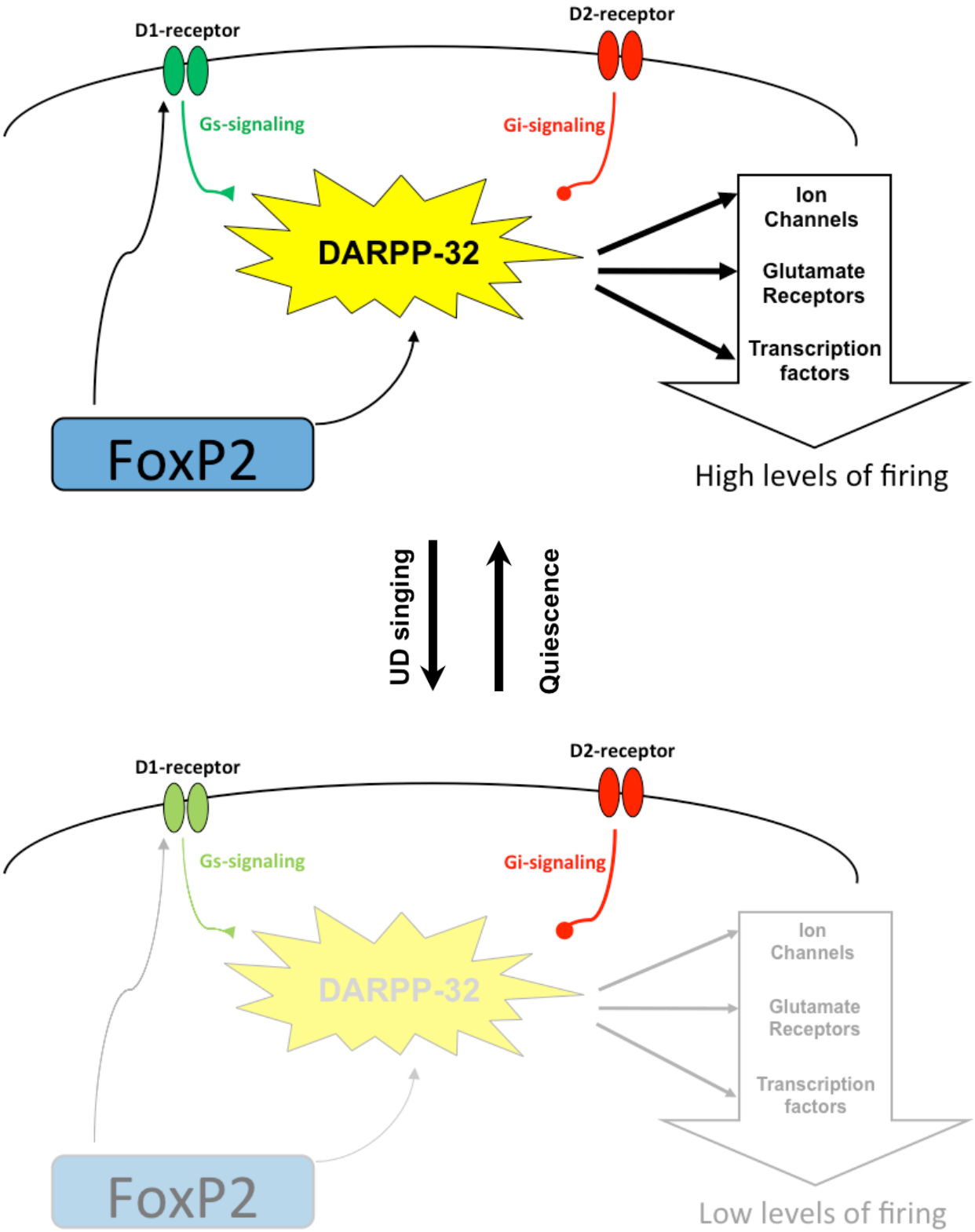
Concluding remarks

An overarching goal of biomedical research is to better understand human neurological conditions and to give insight into their treatment. To that end, I used zebra finch as a model system for understanding the genetic and physiological mechanisms underlying speech and

language disorders. My conclusions are that vocal practice-dependent dynamic regulations of FoxP2 is necessary for vocal learning and that one function of FoxP2 regulation could be to alter basal ganglia output to enhance vocal variability. If a similar mechanism underlies human speech learning then my findings present both challenges and opportunities for treating genetic and non-genetic speech 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. Instead, more sophisticated methods such as the use of CRISPR will be necessary to actually replace faulty genes and allow for their dynamic 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.

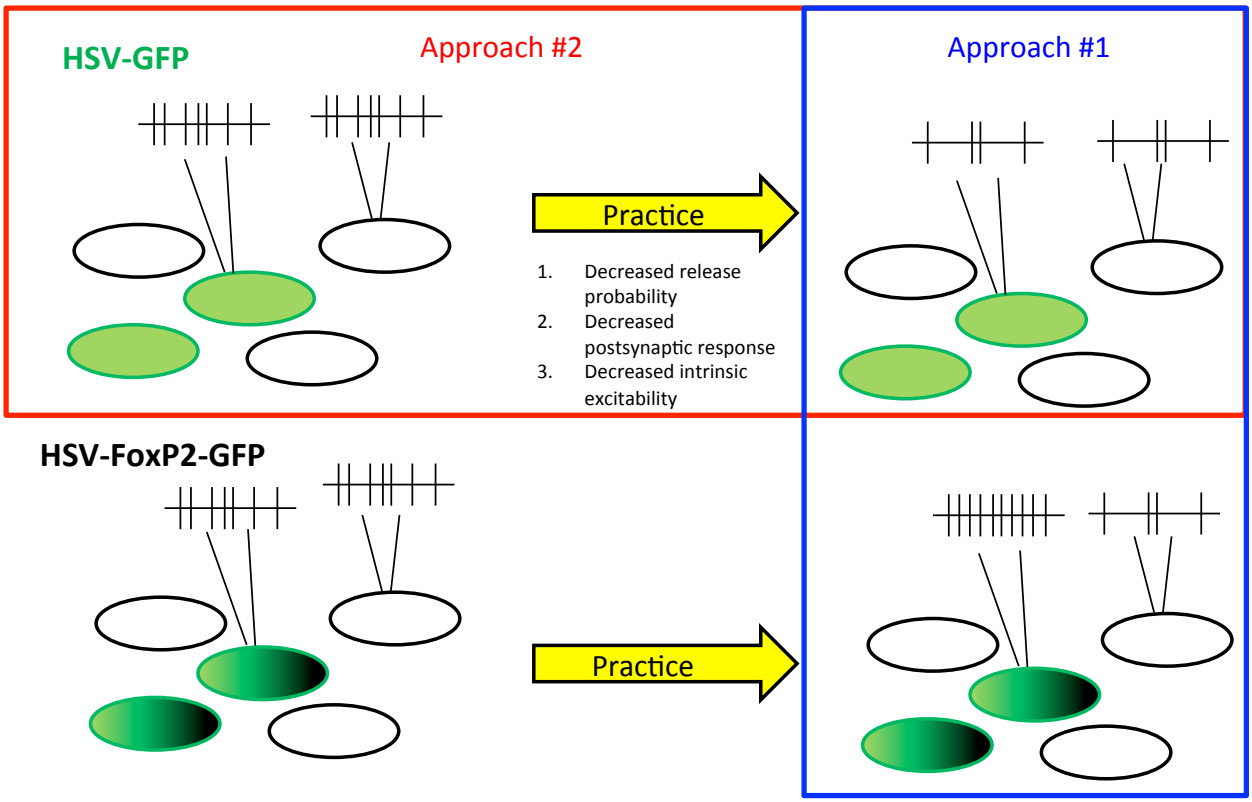
Whether such analogous changes in gene expression exist in humans remains to be determined and because of the difficulties in assaying gene expression in a living human this question will likely never be addressed directly. My findings, however, suggests two different routes by which this possibility can indirectly be assessed. The first route would be to look at human vocal variability following extended vocalization and following prolonged silence. Millisecond timescale fluctuations in flat harmonic sounds, which human speech is rich in, might provide a particularly sensitive and behaviorally relevant measure of circuit function. A second route would be to use human neuroimaging or other techniques for recording neural activity in humans to determine whether there are any practice dependent changes in basal ganglia circuits that accompany extended human vocalization, particularly during language learning. Evidence of increased vocal variability or decreased striatal activity would support my model of FoxP2 function as it applies to humans and bolster the validity of zebra finch as a model system for understanding human language learning.

Fig. 5-1. Model of FoxP2-dependent changes in MSN function.



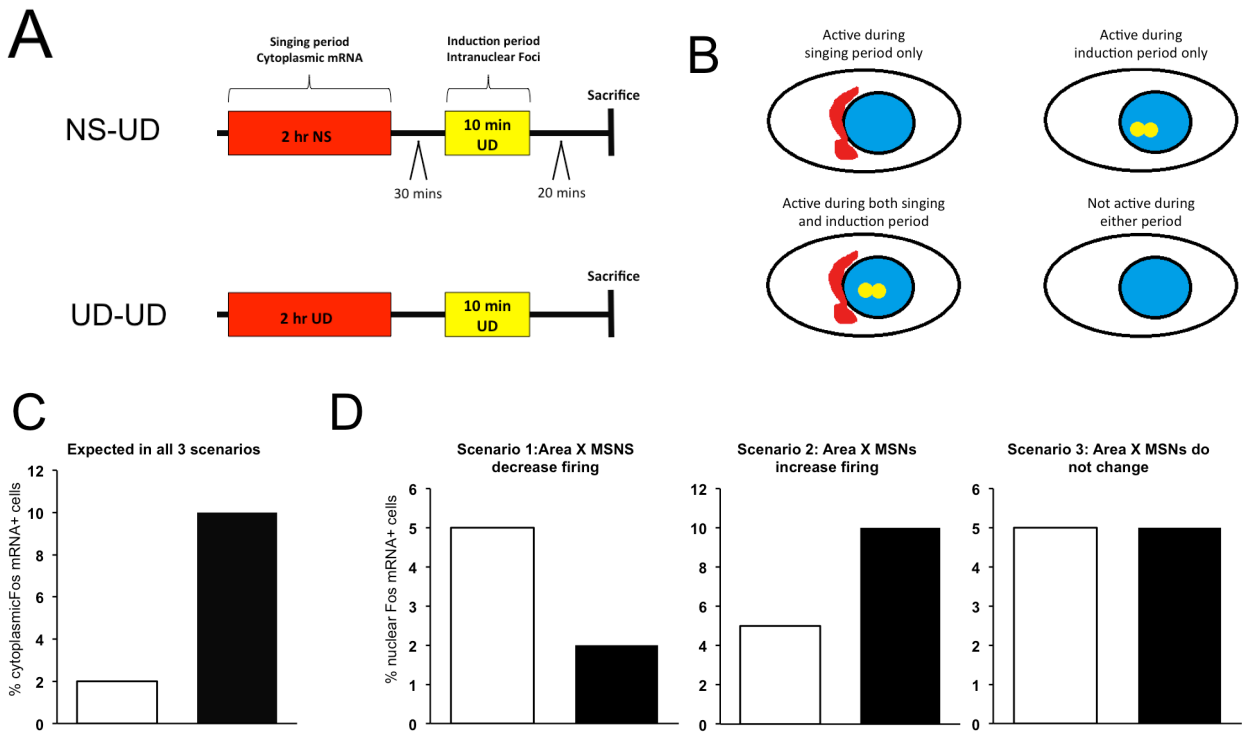
In my model Area X exists in two distinct FoxP2-dependent states. (Top) In the FoxP2 high state there are high levels of D1R and DARPP-32 as these both are thought to be FoxP2 targets in MSNs. High levels of both of these molecules enhances DARPP-32 function which through its effects on ion channels, glutamate receptors, and transcription factors leads to high levels of MSN firing. Based on my DREADDS work this would be predicted to lead to low levels of vocal variability. (Top) The FoxP2 low state has low levels of D1R and DARPP-32 levels or function as these both are thought to be FoxP2 targets in MSNs and FoxP2 knockdown decreases levels of both molecules. Decreased levels of DARPP-32 and the receptor that triggers its activation, synergistically act to decrease DARPP-32 function. This leads to low levels of neuronal excitability. Based on my DREADDS work this would be predicted to lead to high levels of vocal variability. UD singing or the lack thereof can trigger switches between the FoxP2 high to low or low to high states, respectively.

Fig. 5-2. Two approaches to interrogating FoxP2- and practice-dependent changes in Area X function ex vivo.



Two ex vivo approaches to interrogating FoxP2- and practice-dependent changes in Area X function. Approach #1 (blue) tests this by comparing GFP overexpressing cells to FoxP2-GFP overexpressing cells and to uninjected cells following vocal practice. If the hypothesis is true then GFP overexpressing cells will show low excitatory drive that will not differ from neighboring untransfected cells and will negatively correlate with the amount of singing before being sacrificed. In contrast FoxP2-GFP overexpressing cells are expected to show elevated excitatory drive that should be greater than neighbor untransfected cells and positively correlate with the amount of singing before being sacrificed. Alternatively approach #2 (red) could be used to assess practice- but not necessarily FoxP2-dependent changes in neurophysiology. In this scheme brain slices will be cut after NS or UD. Neurons in the UD animals are expected to show lower excitatory drive than those in NS animals.

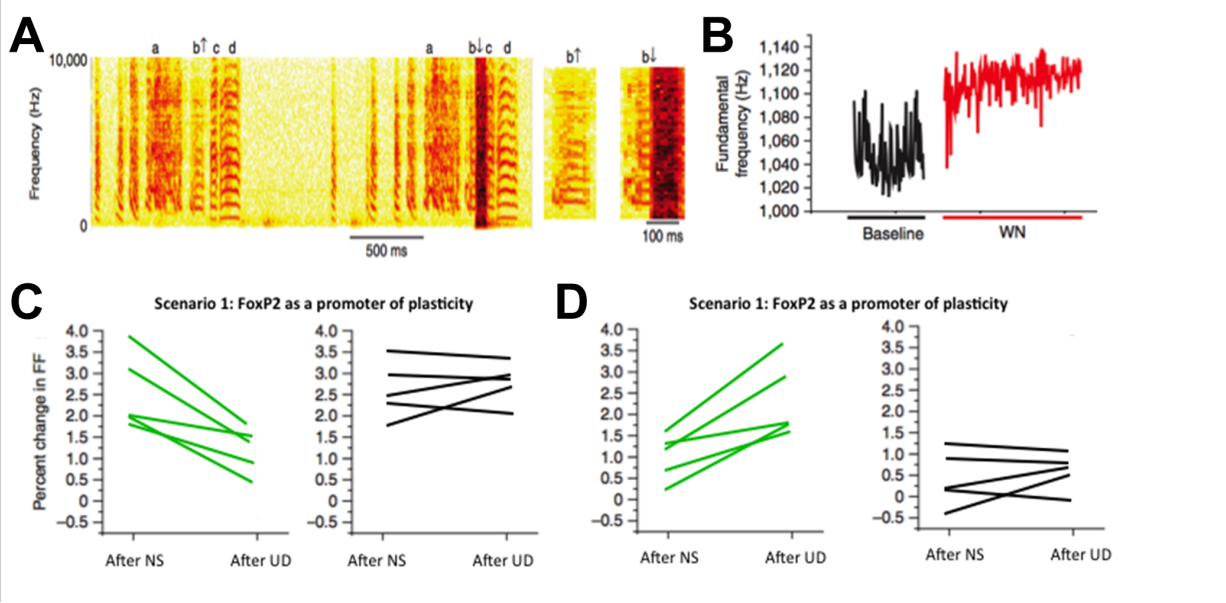
Fig. 5-3. Using catFISH to test for physiological changes following undirected singing.



A) Experimental scheme. Birds will be allowed to sing for two hours (UD-UD) or be prevented from singing for two hours (NS-UD). Following that, both sets of birds will be prevented from singing for 30 mins while IEG mRNA induced during that two hours is exported from the nucleus to be translated in the cytoplasm. After this occurs both sets of birds will be allowed to sing for 10 mins and then sacrificed 20 minutes. This time frame will allow *fos* mRNA induction but will not give it enough time to leave the nucleus. The *fos* mRNA induced here will provide a readout of how much the neurons are active in the UD induction period. B) Four patterns of IEG expression and what they indicate about that neurons activity. Neurons which were active only during the initial 2 hour NS or UD period will show only cytoplasmic *fos* mRNA. Neurons active only during the second ten min UD period will show only nuclear *fos* mRNA. Neurons active during both periods will show both cytoplasmic and nuclear *fos* mRNA. Neurons which are not active during either period will show no *fos* mRNA. C) The total number of cells expressing

cytoplasmic *fos* (with or without nuclear *fos*) is predicted to be higher in the birds that spent two hours singing. This should be the case regardless of whether and how Area X alters physiology as a consequence of singing. It should be noted that this result is not necessarily a positive control and is not critical to the experiment. D) There are three different scenarios that would indicate three different physiological states following singing. If singing leads to a decrease in Area X MSN firing, as predicted, then one would expect lower levels of nuclear fos mRNA in the birds that had spent two hours singing. If singing leads to an increase in Area X MSN firing then one would expect higher levels of nuclear fos mRNA in the birds that had spent two hours singing. If singing does not affect Area X physiology then one would expect no difference between birds that had spent two hours singing and those that did not.

Fig. 5-4: Resolving a FoxP2 role in learning using pCAF



(Adapted from Mehaffey and Doupe, 2015)

A) Specific syllables (denoted by letters) could be targeted with white noise (WN) contingent on the FF of each rendition of the target syllable. In this example, syllable b is targeted and any rendition of b lower than the FF threshold are negatively reinforced using WN. B) At baseline there is some degree of variability in the FF of syllable b, but over the course several hours of pCAF the FF of syllable b is increased to the upper end of this distribution. C). If FoxP2 promotes plasticity then the amount of learning will be less after UD in GFP injected or uninjected birds when FoxP2 levels- and thus the ability to undergo plasticity- are somewhat depleted. In this case FoxP2+ birds would be expected to show elevated levels of plasticity that would not differ between conditions. D) If FoxP2 represses plasticity then the amount of learning in GFP or uninjected birds will be greater after UD, when FoxP2 levels are low and the circuit is primed for plasticity. In this case FoxP2+ birds would be expected to show diminished levels of plasticity that would not differ between conditions.

Appendix I. Expression analysis of the speech-related genes *FoxP1* and*FoxP2* and their relation to singing behavior in two songbird species

Abstract

Humans and songbirds are among the rare animal groups that exhibit socially learned vocalizations: speech and song, respectively. These vocal-learning capacities share a reliance on audition and cortico-basal ganglia circuitry, as well as neurogenetic mechanisms. Notably, the transcription factors Forkhead box proteins 1 and 2 (FoxP1, FoxP2) exhibit similar expression patterns in the cortex and basal ganglia of humans and the zebra finch species of songbird, among other brain regions. Mutations in either gene are associated with language disorders in humans. Experimental knock-down of FoxP2 in the basal ganglia song control region Area X during song development leads to imprecise copying of tutor songs. Moreover, FoxP2 levels decrease naturally within Area X when zebra finches sing. Here, we examined neural expression patterns of FoxP1 and FoxP2 mRNA in adult Bengalese finches, a songbird species whose songs exhibit greater sequence complexity and increased reliance on audition for maintaining their quality. We found that FoxP1 and FoxP2 expression in Bengalese finches is similar to that in zebra finches, including strong mRNA signals for both factors in multiple song control nuclei and enhancement of FoxP1 in these regions relative to surrounding brain tissue. As with zebra finches, when Bengalese finches sing, FoxP2 is behaviorally downregulated within basal ganglia Area X over a similar time course, and expression negatively correlates with the amount of singing. This study confirms that in multiple songbird species, FoxP1 expression highlights song control regions, and regulation of FoxP2 is associated with motor control of song

INTRODUCTION

The importance of the FOXP subfamily of transcription factors in the brain was not clear until FOXP2 was identified as the monogenetic locus of a speech and language abnormality. Half of the members of a British pedigree, known as the KE family, suffer from a rare communication disorder. Affected members share a single mutation in FOXP2 that causes a severe impairment in the selection and sequencing of fine orofacial movements (Lai et al., 2001; Vargha-Khadem et al., 1998). In addition to articulatory problems, affected individuals have profound deficits in production and comprehension of word inflections and syntactical structure (Alcock et al., 2000; Watkins et al., 2002). The phenotype resulting from its mutation indicates that FOXP2 is linked to neural pathways underlying speech and language. FOXP1 is the closest forkhead family member to FOXP2, with which it shares high similarity at the amino acid level (68% identity and 80% similarity between the two human sequences). FOXP1 can heterodimerize with FOXP2 and can repress transcription of similar groups of genes (Li et al., 2004; Shu et al., 2001; Wang et al., 2003).

FOXP1 is also associated with speech and language through multiple cases (Carr et al., 2010; Hamdan et al., 2010; Horn et al., 2010). For example, a patient with a genetic deletion restricted to FOXP1 exhibits difficulties with verbal expression resembling the phenotype of affected KE family members (Pariani et al., 2009). Besides humans (Homo sapiens), no taxon of primates is capable of substantially modifying its vocal repertoire in response to experience. Moreover, most laboratory animals, including rodents, do not learn a substantial portion of their vocalizations (Kikusui et al., 2011; Arriaga et al., 2012; Mahrt et al., 2013). In striking contrast, thousands of songbird species share the trait of vocal learning with humans, enabling comparison of brain–behavior relationships among these taxa. Zebra finches (Taeniopygia guttata

Reichenbach 1862) are a well-studied songbird species in which song learning is sexually dimorphic: juvenile males learn their courtship songs from adult male conspecifics (tutors) whereas females do not produce learned songs. Zebra finch song is composed of notes, syllables, motifs and bouts. Notes are the smallest unit of song and are defined as a region of a syllable that maintains a temporally continuous frequency pattern. Syllables are composed of one or more notes bounded by a brief period of silence. Motifs are repeated sequences of syllables lasting ~1s with multiple motifs in succession organized in a bout. Bouts are composed of several motifs bounded by a longer period of silence (Brenowitz et al., 1997; Price, 1979).

Male, but not female, zebra finches possess the full and interconnected suite of cortico-basal ganglia nuclei that underlies song learning and production. Song control circuitry includes the anterior forebrain pathway (AFP), which is important for song learning in juveniles and song maintenance and plasticity in adults, and the posterior descending pathway, which is required for song production (Scharff and Nottebohm, 1991; Brainard and Doupe, 2000; Kao et al., 2005). Neurons in the HVC (acronym used as a proper name), a premotor vocal control nucleus, directly project to the robust nucleus of the arcopallium (RA) (Nottebohm et al., 1976; Nottebohm, 2005) and indirectly project to the RA through basal ganglia nucleus Area X, the medial nucleus of the dorsolateral thalamus, and the lateral magnocellular nucleus of anterior nidopallium (LMAN) in the AFP. The AFP is homologous to basal ganglia-thalamo-cortical circuit loops in mammals. Area X shares many features characteristic of the mammalian striatum and pallidum, including cell types and connectivity (Gale and Perkel, 2010).

Songbirds and humans also share neurogenetic mechanisms that underlie their vocal learning capacities. FoxP1 and FoxP2 exhibit similar expression patterns in the cortex and basal ganglia of humans and zebra finches (Teramitsu et al., 2004). Knock-down of FoxP2 in Area X

of juvenile zebra finches leads to imprecise copying of the tutor song, suggesting that FoxP2 is involved in the normal process of vocal learning (Haesler et al., 2007). Moreover, Area X FoxP2 is behaviorally and socially regulated. Non-singing zebra finches have high levels of Area X FoxP2 that decline acutely when males practice their songs alone (termed undirected singing) in the morning, but not when they sing to females (directed singing) (Teramitsu and White, 2006; Hilliard et al., 2012). The downregulation of FoxP2 during undirected singing is particularly robust in juvenile zebra finches undergoing sensorimotor learning: the more they practice, the lower their Area X FoxP2 levels. Interestingly, hearing is required to maintain this negative correlation (Teramitsu et al., 2010). Moreover, coincident with decreased FoxP2, vocal variability increases after 2h of undirected singing in both juvenile and adult zebra finches (Hilliard et al., 2012; Miller et al., 2010). These observations have led us to hypothesize that singing-driven decreases in Area X FoxP2 levels promote vocal variability and motor exploration whereas high levels promote song stabilization (Miller et al., 2010).

Here, we further test the relationship between learned vocal behaviors and FoxP1 and FoxP2 gene expression by examining another songbird species, the Bengalese finch (Lonchura striata domestica Linnaeus 1758), in which song learning and song control circuitry are also sexually dimorphic, but whose song exhibits features that are distinct from zebra finch song. Adult male zebra finches sing a linear song sequence and thus exhibit a very simple birdsong 'syntax', whereas male Bengalese finches generate songs with greater syntactical complexity (Okanoya, 2004). After deafening, the songs of Bengalese finches degrade faster than those of zebra finches (Okanoya and Yamaguchi, 1997; Woolley and Rubel, 1997), indicating a greater reliance on audition for their song maintenance. These observations suggest that singing-driven decreases in Area X FoxP2 levels might be more robust in Bengalese finches than in zebra

finches. As a consequence, increases in song variability following song practice might be evident in adult male Bengalese finches. We therefore tested the following hypotheses: (1) Bengalese finches and zebra finches share similar FoxP1 and FoxP2 gene expression patterns; (2) FoxP2 mRNA is behaviorally regulated in male Bengalese finches; (3) downregulation of FoxP2 within Area X is correlated with the amount of undirected singing in both species; (4) the singing-driven regulation of FoxP2 within Area X of Bengalese finches is more profound than in zebra finches; and (5) vocal practice promotes song variability in adult male Bengalese finches.

MATERIALS AND METHODS

We conducted in situ hybridization on brain tissue from Bengalese finches and zebra finches of both sexes under different behavioral conditions to investigate FoxP gene expression patterns, the time course of downregulation of FoxP2, and the relationship between amount of singing and FoxP2 levels within Area X. A separate group of adult male Bengalese finches was used to investigate song variability following two different behavioral conditions known to alter Area X FoxP2 levels (Fig. A1-1).

Animals and tissues

All animal use was in accordance with National Institutes of Health guidelines for experiments involving vertebrate animals and approved by the Institutional Animal Care and Use Committee of the University of California, Los Angeles. Adult male and female zebra finches and Bengalese finches (age >120days) were taken from our breeding colony (where they were kept under a 13h:11h light:dark cycle). After behavioral monitoring (see below), birds were decapitated for collection of brains, which were rapidly extracted and frozen on aluminum floats on liquid nitrogen, then stored at –80°C until use.

Riboprobe preparation and in situ hybridization analysis

FoxP genes are highly conserved among such disparate avian species as zebra finches and chickens (FoxP1: zebra finch versus chicken, identities=95%; FoxP2: zebra finch versus chicken, identities=97%). Although the genome of Bengalese finches is not yet available, the similarity of their FoxP genes to zebra finch sequences is expected to be even higher based on their closer phylogenetic relationship. We therefore used riboprobes directed against zebra finch FoxP1 and FoxP2 (Teramitsu et al., 2004) to detect these transcripts in both species. The FoxP1 probe was designed to hybridize to the coding region upstream of the zinc finger domain of zebra

finch FoxP1, corresponding to 661–998bp of human FOXP1 relative to the start codon. The FoxP2 probe was designed to hybridize to 1870–2127bp of the zebra finch FoxP2 relative to the start codon. pCR4-TOPO vector (Invitrogen, Carlsbad, CA, USA) with zebra finch FoxP cDNA fragments was used for in vitro transcription to generate sense and antisense RNA probes labeled with [33P]UTP (Perkin-Elmer, Foster City, CA, USA) using the Riboprobe Combination System-T3/T7 (Promega, Madison, WI, USA).

Frozen brains were cryosectioned in either the sagittal or coronal plane at 20μm and adjacent sections were mounted onto 25×75mm slides (Superfrost, Fisher Scientific, Pittsburgh, PA, USA) in a manner that created seven replicate sets. One set was stained with thionin to enable identification of neuroanatomical structures. The adjacent four sets were exposed to the FoxP1 sense, FoxP1 The Journal of Experimental Biology 216 (19) antisense, FoxP2 sense or FoxP2 antisense probes. In situ hybridizations were performed and signals from different brain regions were quantified as previously described (Teramitsu et al., 2004; Teramitsu and White, 2006; Teramitsu et al., 2010). Sections of Bengalese finches were run aligned with sections of zebra finches from the same behavioral conditions to enable direct comparisons. Preliminary analysis of Bengalese finch sections revealed that: (1) the distinct expression patterns between brains exposed to either FoxP1 or FoxP2 antisense probes were as expected based on prior studies, (2) signals from antisense probes were robust whereas those from sense probes were negligible, and (3) signals were consistent across adjacent brain sections. These results provide confidence that riboprobes designed from zebra finch cDNA also specifically detect FoxP1 and FoxP2 in Bengalese finch brain.

Behavioral monitoring and sound recording

Birds were housed individually in sound attenuation chambers (Acoustic Systems, Austin, TX, USA) for 2–3days prior to the behavioral experiments to enable acclimation to the recording environment. Sounds were recorded using Countryman EMW omnidirectional lavalier microphones (Countryman Associates, Menlo Park, CA, USA) and digitized using a PreSonus Firepod (44.1kHz sampling rate, 24bit depth; Baton Rouge, LA, USA). Recordings were acquired using Sound Analysis Pro (SAP) 2011 software (Tchernichovski et al., 2000).

Behavioral experiments were conducted between 08:00 and 11:00h, starting at lights on. For FoxP gene analysis, birds were killed following the completion of different behavioral paradigms, which are illustrated in Fig.1A and described as follows. Female birds were left alone and undisturbed inside the chamber for 2h after lights on. Non-singing males (referred to as NS; Fig.1A) were also left alone for 2h after lights on, but with the door to the chamber ajar. If they appeared to attempt to sing, they were distracted by the experimenter. Those that sang more than five motifs despite the experimenter's presence were excluded from this group. Of note, we previously found that the non-singing paradigm did not lead to detectable changes in zebra finch stress levels as measured by serum corticosterone values (Miller et al., 2008). In addition, Area X gene expression patterns from birds that were distracted from singing by an experimenter clustered together with patterns from birds that sang very little by their own volition. This suggests that singing behavior – and not the absence or presence of the experimenter – is the more crucial determinant of gene expression in Area X (Hilliard et al., 2012). Males singing undirected song (referred to as UD; Fig.1A) were allowed to sing alone inside the chamber for a pre-determined period of time – 1, 1.5 or 2h after the first song in the morning. For analysis of song variability, a separate set of birds was used for which the behavioral conditions are illustrated in Fig.1B. One group of male birds (N=6) was kept from singing for 2h and then

allowed to sing undirected songs. Songs sung during the subsequent 20min (termed NS-UD songs) were analyzed. On another day, the same group of male birds was allowed to sing undisturbed for 2h, and then songs that were sung in the subsequent 20min (termed UD-UD) were analyzed.

Quantification of the amount of singing

Audio files generated by SAP were edited with Audacity 1.3 Beta (http://audacity.sourceforge.net) by manual removal of cage noise and calls, leaving only songs. In our previous study on zebra finches, the amount of singing was quantified by counting the number of motifs (Teramitsu and White, 2006). However, there is considerable variability in phonology and macroscopic song structure both within and between the two songbird species studied here (Fig.A1-2A). The greater syntactical variability in Bengalese finch song makes it challenging to identify their motifs (Fig.A1-2B). Moreover, the length of the motifs varies among different Bengalese finches and between the two songbird species. To minimize error and avoid introducing bias by manually identifying song motifs, we used SAP to automatically measure the length of each song syllable. Syllables were segmented using experimenter-derived amplitude thresholds in SAP, and then run through the 'Feature Batch' module, which computes the duration of each syllable in the batch. The total amount of singing was then defined as the sum of the durations of all syllables identified for a given time period.

Quantification of FoxP gene expression

For semi-quantitative and quantitative analyses, optical density (OD) measurements of FoxP signals were obtained from digitized images of autoradiograms using Adobe Photoshop 7.0

(Adobe Systems, San Jose, CA, USA). First, to provide a qualitative comparison of gene expression levels across brain regions, OD values from each region were calculated from multiple sagittal sections of the brains of one 2h UD Bengalese male, one 2h NS Bengalese male (shown in Fig.A-13) and one Bengalese female (shown in Fig.A1-4). All OD values were normalized to those from a nidopallial area of the same section that did not contain any song control nuclei. Values are reported in Table1. For quantitative analysis of Area X FoxP2 levels, OD values from within Area X were normalized to those from the ventral striato-pallidum (VSP), as previously described (Teramitsu and White, 2006). To determine the statistical significance of the Area X FoxP2 levels, a resampling procedure was employed as follows: 10,000 hypothetical data sets of the same size were resampled from The Journal of Experimental Biology 216 (19) the actual normalized OD values and the amount of singing in the experiments. For each resampled data set, a slope of the linear regression of these variables (OD versus amount of singing) was calculated, generating a distribution of 10,000 slopes for each species. A correlation was determined to be significantly negative if the upper and lower boundaries of the 95% confidence interval for the distribution of slopes were negative.

Syllable identification and clustering

All syllable clustering and sequence analysis was performed in the freely available R programming language (http://www.r-project.org) using custom-written clustering and syntax entropy scripts, available at the White laboratory website (https://www.ibp.ucla.edu/ research/white/code.html). To group syllables in an unbiased fashion and subsequently calculate syntax entropy, a hierarchical clustering and automated tree-trimming algorithm was utilized. Raw acoustic recordings from the first 20min following NS or UD for each bird were subjected to SAP's 'Feature Batch', using experimenter-derived amplitude thresholds to segment syllables.

A number of filtration steps were then applied to the 'Feature Batch' output to identify song syllables from non-song sounds (wing flaps, cage noise, etc.) captured by the recording software. An initial filtration step implemented userdefined duration thresholds above and below which all sounds were removed from the data set. Next, a maximum inter-syllable interval was determined by the experimenter for all remaining prospective syllables in the data set. Syllables that fell below this inter-syllable interval were grouped into prospective motifs/bouts. A filter to remove all motifs/bouts composed of two and/or three syllables was then applied. WAV files representing each motif/bout were generated and presented to the user for visual inspection, at which point motifs consisting of calls or non-song sounds in the recordings were removed from the data set if present. Finally, individual WAV files for all remaining syllables were generated.

Individual WAV files for both behavioral sessions for each animal were run against themselves in SAP's 'Similarity Batch' module in an M × N symmetric similarity batch. Upon completion of the batch, the product of the similarity and accuracy score for each syllable–syllable comparison was calculated and stored in a square matrix with rows and columns representing individual syllables and the elements of the matrix representing the product of the similarity and accuracy scores for a given syllable–syllable comparison. A distance matrix was then created by calculating the Euclidean distance between the product of similarity and accuracy scores for all syllable–syllable pairs. This distance matrix was used as the input to a hierarchical clustering function in the WGCNA R package (Langfelder and Horvath, 2008), generating a dendrogram. Branches of the dendrogram were then pruned using the dynamic tree-trimming algorithm, also in the WGCNA R package, a novel method for detecting clusters within hierarchical trees by considering the shape of the branches when trimming them into groups (Langfelder et al., 2008). Upon completion of cluster detection, each cluster was

described by an ‘eigensyllable’, defined as the first principal component of the cluster as determined by singular value decomposition. The Pearson correlation between all module eigensyllables was then computed and clusters whose eigensyllables correlated above a user-defined threshold (in this case, 0.75) were merged, generating the final number of clusters/syllable types in each bird’s song.

Final inspection of cluster homogeneity was performed by visual inspection of syllable spectrograms within each cluster. Syllables inappropriately assigned to a cluster were manually reassigned.

Syntax entropy

The syllable syntax, defined as the sequence in which the bird orders its syllables, was determined based on syllable cluster assignment in the preceding step. Syntax entropy was then calculated as described in Miller et al. (Miller et al., 2010). A string-based approach was utilized for syntax analysis, as motifs were often difficult to identify in Bengalese finch songs. 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. A resampling paired t-test was utilized to assess the significance of change in syntax entropy scores between behavioral conditions for all birds as a group.

Similarity, accuracy, identity and syllable acoustic features

Upon completion of clustering, syllables within each cluster were divided into NS-UD and UD-UD groups. All syllable types that did not have at least 20 renditions sung in both behavioral contexts were removed from consideration in analysis of acoustic features. The range

in the number of renditions for the remaining syllables that were analyzed was 55–762. A bootstrap one-way ANOVA was performed on similarity, accuracy and identity scores and all acoustic features within each bird to determine whether syllables were independent of one another. For all acoustic measures, the between-syllable difference P-value was less than 0.05, thus syllables were treated as independent of one another.

Resampling two-way ANOVAs were performed for each acoustic measure using syllables and behavioral condition as the two independent factors. F-statistics were generated for the actual data set and then compared with a distribution of 10,000 F-statistics calculated by resampling the original data under assumption of the null hypothesis to determine whether a syllable effect, a behavioral effect and/or an interaction between the two variables were present for each measure.

RESULTS

FoxP1 expression in Bengalese finch brain

FoxP1 mRNA signals indicated high expression levels in the densocellular part of the hyperpallium, the mesopallium, the striatopallidum and the dorsal thalamus in both male (Fig. A1-3) and female (Fig. A1-4) Bengalese finches. In the basorostral pallial nucleus (Bas) and song control nucleus LMAN, FoxP1 expression was lower than in the surrounding nidopallium region regardless of sex (Fig. A1-3C, Fig. A1-4). In contrast, sexually dimorphic FoxP1 expression was observed in song control nuclei HVC, RA and striato-pallidal Area X, as the signals were greater in these nuclei relative to the respective surrounding brain tissue only in male Bengalese finches (Fig. 3-3C). In females, signals were similar across these sub-regions (Fig.4). FoxP1 did not appear to be regulated by undirected singing in male Bengalese finches. Expression patterns from sagittal sections containing multiple song control regions were broadly similar between the 2h NS and UD groups (Fig. 3-3C). A semi-quantitative summary of these observations is presented in Table1. Coronal sections from a separate set of birds were used to focus on Area X and LMAN (Fig. 3-5), but again, no behavioral regulation of FoxP1 was observed

FoxP2 expression in Bengalese finch brain

FoxP2 signals were lightly and uniformly distributed in cortical areas whereas they were robust in the striato-pallidum, the dorsal thalamus and the Purkinje cell layer of the cerebellum in both male (Fig. A1-3) and female (Fig. A1-4) Bengalese finches. No sexual dimorphism of FoxP2 expression was observed in any of the song control nuclei except for Area X. FoxP2 expression within Area X in female Bengalese finches was similar as that of the surrounding

striatopallidum (Fig. A1-4). FoxP2 expression in Area X of male Bengalese finches has reported to be lower than the surrounding striatopallidum (Haesler et al., 2004). However, the behavioral condition of the birds used in that experiment was not specified. Here we present evidence that FoxP2 within Area X is comparable to or slightly higher than in the surrounding striato-pallidum in 2h NS Bengalese finches but lower than in 2h UD Bengalese finches (Fig. A1-3D). A semi-quantitative summary of these observations is presented in Table1.

Behavioral regulation of FoxP2 within Bengalese finch

Area X In zebra finches, FoxP2 expression levels decline specifically within Area X when males engage in 2h of UD singing in the morning (Hilliard et al., 2012; Teramitsu and White, 2006; Teramitsu et al., 2010). To determine whether similar singing-driven changes occur in a related songbird species with distinct song features, we examined FoxP2 expression in Area X of male Bengalese finches, in parallel with that in zebra finches, and compared levels between UD and NS conditions. To confirm the behavioral regulation of FoxP2 suggested in Fig. A1-2D, additional 2h NS and 2h UD male Bengalese finches were killed and brain tissues were sectioned coronally to display Area X bilaterally in the same section. The additional time points of 1h UD and 1.5h UD groups were utilized to track the time course of downregulation of FoxP2 within Area X during singing. We found that Area X FoxP2 levels were significantly downregulated at the 1.5h UD and 2h UD time points for both species (Fig. A1-6).

FoxP2 levels within Area X in 2h UD Bengalese finches were significantly higher than those found in 2h UD zebra finches ($P<0.01$). In order to interpret this difference, we measured the amount of singing in both groups. We found that zebra finches in our study sang more than Bengalese finches did (means ± s.e.m., Bengalese finch 351±53 s versus zebra finch 758±166 s,

Kruskal–Wallis nonparametric test, P=0.040). Thus, the difference in FoxP2 levels between 2h UD Bengalese finches and 2h UD zebra finches could reflect the difference in the amount of singing. To explore this possibility, the relationship between the amount of singing and FoxP2 levels was further examined.

Correlation between FoxP2 levels and amount of singing

Area X FoxP2 levels were negatively correlated with the amount of singing in both zebra and Bengalese finches, as illustrated by the negative slope of the linear regression lines that were fit to the data from each species (zebra finch: P0.05, see below), indicating that, contrary to our prediction, Bengalese finch FoxP2 levels within Area X are not more responsive to singing than those in zebra finches.

Song variability after vocal practice

Songs that were sung by adult male Bengalese finches in the 20min period immediately following a 2h period of UD singing (UD-UD) were compared with those sung following 2 h of non-singing (NS-UD). One expectation is that there would be no difference between the behavioral conditions, based on prior work in zebra finches in which a difference was only observed in juveniles (Miller et al., 2010). The other expectation is that variability after UD-UD singing would be increased relative to the NS-UD conditions, based on the overall greater variability in Bengalese song and its strong dependence on hearing. In line with a majority of our predictions, we found that for many phonological and sequential measures of song variability there were no differences between the two conditions. However, on certain measures, a slight decrease in variability was observed in the UD-UD condition relative to the NS-UD condition, in contrast to our predictions. To describe syllable variability, we examined the average

withingroup similarity, accuracy and syllable identity (similarity × accuracy/100) of all syllables within a cluster analyzed as a function of behavioral condition. The variability of syllable identity (P=0.034; Fig. 3-8) was lower in the UD-UD condition, reflecting similar trends in similarity (P=0.080) and accuracy (P=0.075). We next examined the mean coefficient of variance (CV) for all syllables within a cluster. Again contrary to our predictions, the CV was lower in the UD-UD condition for individual syllable features of pitch goodness (P=0.0002), Wiener entropy (P=0.004) and mean frequency (P=0.017; Table2). A two-way ANOVA revealed that there was no effect of behavioral condition on the mean values for any of these features. Finally, we utilized entropy-based methods similar to those of Miller et al. (Miller et al., 2010) to measure syntax variability, investigating all syllables produced during the 20min following each behavioral condition using a string-based analysis described in that study. The results indicate no significant difference in syntax entropy between the two behavioral conditions (average NS-UD entropy=0.185, average UD-UD entropy=0.168; P>0.05), similar to our prior findings in adult zebra finches.

DISCUSSION

Sexually dimorphic expression of FoxP1 in songbirds

In line with our expectations, the brain expression patterns of FoxP1 and FoxP2 in Bengalese finches are broadly consistent with those previously described in zebra finches (Teramitsu et al., 2004), including strong mRNA signals for both factors in multiple song control nuclei and enhancement of FoxP1 in HVC and Area X relative to surrounding brain tissue. One apparent difference between the two species was in the arcopallial song control region, the RA. FoxP1 in the RA of female zebra finches is higher relative to the surrounding brain tissue (Teramitsu et al., 2004), but this enhancement was not prominent in coronal sections of a female Bengalese finch brain (data not shown) and was not detected in sagittal sections of another female Bengalese finch (Fig. A1-4). Whether this is a true species difference is unclear because we were unable to detect RA in the Nissl-stained female Bengalese finch sections, despite its visibility in sections from male brains subjected to the same staining conditions (Fig. A1-3). As previously reported in zebra finches (Teramitsu et al., 2004), the RA of male Bengalese finches exhibited FoxP1 signals that were slightly higher than those of the surrounding arcopallium. Projection neurons of the RA synapse directly onto the motor neurons that innervate the muscles of phonation, similar to direct projections of layer V motor cortical neurons onto laryngeal motor neurons in humans, and are thought to enable the capacity for vocal learning (Jürgens, 2009; Arriaga et al., 2012). In the spinal cord, FOXP1 plays a crucial role in defining the columnar identity of motor neurons at each axial position, as well as organizing motor axon projections (Rousso et al., 2008). Similarly, FoxP1 may organize the RA cortical motor neuron projection to syringeal and respiratory motor neurons in songbirds.

With regard to other telencephalic song control regions, enhanced expression of FoxP1 in the HVC and Area X in male, but not female, Bengalese finches mirrors the zebra finch expression pattern. There is no evidence for singing-driven regulation for FoxP1 expression in either adult Bengalese or zebra finch brains (Figs A1-3, A1-5, Table A1-1). The sexually dimorphic expression of FoxP1 in song control areas (HVC, male RA, Area X), together with the speech and language deficits associated with its mutation in humans (Carr et al., 2010; Hamdan et al., 2010; Horn et al., 2010; Pariani et al., 2009), suggest that FoxP1 plays a role in the formation of song circuitry dedicated to singing behavior.

The expression of FoxP1 within the LMAN and Bas in Bengalese finches is low relative to the surrounding tissue, and does not exhibit sexually dimorphic patterns or singing-driven regulation. The Bas is involved in feeding and oral-manipulative behaviors other than vocalization and does not anatomically connect to the vocal control system in songbirds (Wild and Farabaugh, 1996). Because both male and female finches engage in oral movements related to feeding behavior, it is not surprising that FoxP1 levels in the Bas are similar in both sexes. In contrast, the LMAN plays a key role only in male song learning and maintenance (Bottjer et al., 1984; Brainard and Doupe, 2000), yet FoxP1 mRNA expression was not sexually dimorphic in this nucleus. Further investigation may determine whether the FoxP1 protein exhibits sexual dimorphism in the LMAN, as differences between transcriptional and translational levels have been observed for other transcription factors in song control circuitry (Whitney and Johnson, 2005). Although FOXP1 mutations in humans are accompanied by language disorders, the impact of FoxP1 on song learning and production remains to be determined. Given that we did not observe behavioral regulation of FoxP1 in either species, it seems likely that its role may be

in promoting the developmental differentiation of neural structures, consistent with the general role of Fox transcription factors during embryogenesis (Carlsson and Mahlapuu, 2002).

Behavioral regulation of FoxP2 in songbirds

Unlike FoxP1, FoxP2 expression in male songbirds was not enriched in the HVC or the RA, and appeared similar to levels in the HVC and RA of female brains (Figs. A1-3, A1-4, Table A1-1). In Area X, FoxP2 was slightly higher or comparable to levels in the adjacent VSP in NS adult male songbirds. FoxP2 expression is enhanced in the striato-pallidum of hatchling zebra finches and increases in Area X during development (Teramitsu et al., 2004). This observation, together with the structural deficits in the basal ganglia of affected KE family members, is consistent with a role for FoxP2 in contributing to the structural organization of basal ganglia regions critical for vocal learning. Post-embryogenesis, Area X FoxP2 levels are downregulated after undirected singing in juvenile and adult zebra finches (Teramitsu and White, 2006; Teramitsu et al., 2010). Lentiviral-mediated FoxP2 knockdown in Area X of juvenile zebra finches results in inaccurate copying of the tutor song (Haesler et al., 2007). Together, these findings suggest that FoxP2 is involved not only in forming neural structures for vocal learning during embryogenesis, but also in the ongoing use of such structures during vocal learning and adult song maintenance, including in adult male Bengalese finches.

Correlation between Area X FoxP2 levels and undirected singing in two species of songbird

We investigated the time course over which FoxP2 levels are first observed to decrease in Area X during singing in both Bengalese and zebra finches. We found that levels became significantly downregulated at the 1.5h time point in both species (Fig.A1-6). Contrary to our

prediction, Area X FoxP2 downregulation in Bengalese finches was not more robust than in zebra finches. This outcome is qualified by the recognition that experimental quantification of the amount of singing is not always proportional to the time spent singing. For example, one zebra finch sang for 241s within 2h, whereas another sang for 487s within 1h. We observed a negative correlation between the amount of singing and FoxP2 levels within Area X of zebra finches, which confirms results from our prior studies (Hilliard et al., 2012; Teramitsu and White, 2006; Teramitsu et al., 2010). We now report a similar negative correlation in Bengalese finches (Fig.A1-7B). Thus, singing may promote FoxP2 mRNA degradation, possibly through miRNA regulation (Clovis et al., 2012), or inhibit mRNA synthesis following song onset. In either case, this regulation of FoxP2 takes time, only producing significant decreases 1.5h following song onset in this study (Fig.A1-6). It is difficult to disentangle the effects of time and the amount of singing on FoxP2 levels because we cannot control the amount and timing of singing once birds start. For each species, in birds that did sing similar amounts of song (Fig.A1-7), there is a trend that the longer they were given before being killed, the lower their Area X FoxP2 levels.

FoxP2 downregulation within Area X in Bengalese finches and zebra finches

When all birds are considered, the downregulation of FoxP2 did not occur on a faster time scale in Bengalese finches than in zebra finches, as demonstrated by the lack of a statistically significant difference in the slopes of regression lines plotted to the data (Fig.A1-7). The lack of a detectable difference between the two species may be due to a lack of sensitivity in the in situ hybridization. However, in pilot experiments, we compared FoxP2 levels obtained with quantitative reverse transcriptase PCR from cDNA obtained from unilateral punches of

Area X with those obtained from in situ hybridization of the remaining hemi-sections from the same bird (J. Liu, unpublished). The sensitivity was comparable across methods, indicating the suitability of our approach, which also enables us to compare our current findings with past studies that employed in situ analyses. The relationship between FoxP2 and singing in Bengalese finches may be underestimated here simply because they sang less as a group. A broader range of singing might enable detection of more subtle differences between the species. Alternatively, the dependence of FoxP2 levels on singing may indeed be similar in both species, despite differences in features of their songs.

Vocal variability after vocal practice

We previously found that in juvenile (75days of age) zebra finches, vocal practice for 2h in the morning leads to increased vocal variability (Miller et al., 2010) and that in adult zebra finches, the amount of singing is correlated with increased spectral entropy (Hilliard et al., 2012). Thus, we predicted that vocal practice might lead to increased vocal variability in adult Bengalese finches. To our surprise, we found that despite similar behavioral regulation of FoxP2 in Bengalese and zebra finches, periods of low FoxP2 are associated with slight decreases in variability of multiple features in Bengalese finch song (Miller et al., 2010; Hilliard et al., 2012). Thus, it is possible that FoxP2 downregulation may decrease vocal variability or that changes in FoxP2 levels are unrelated to changes in vocal variability in this species. Arguing against these possibilities is the observation that viral knockdown of FoxP2 in Area X is sufficient to increase variability in both juvenile (Haesler et al., 2007) and adult zebra finches (Murugan and Mooney, 2012). Multiple factors could contribute to the observed difference in these select song features, and are detailed below.

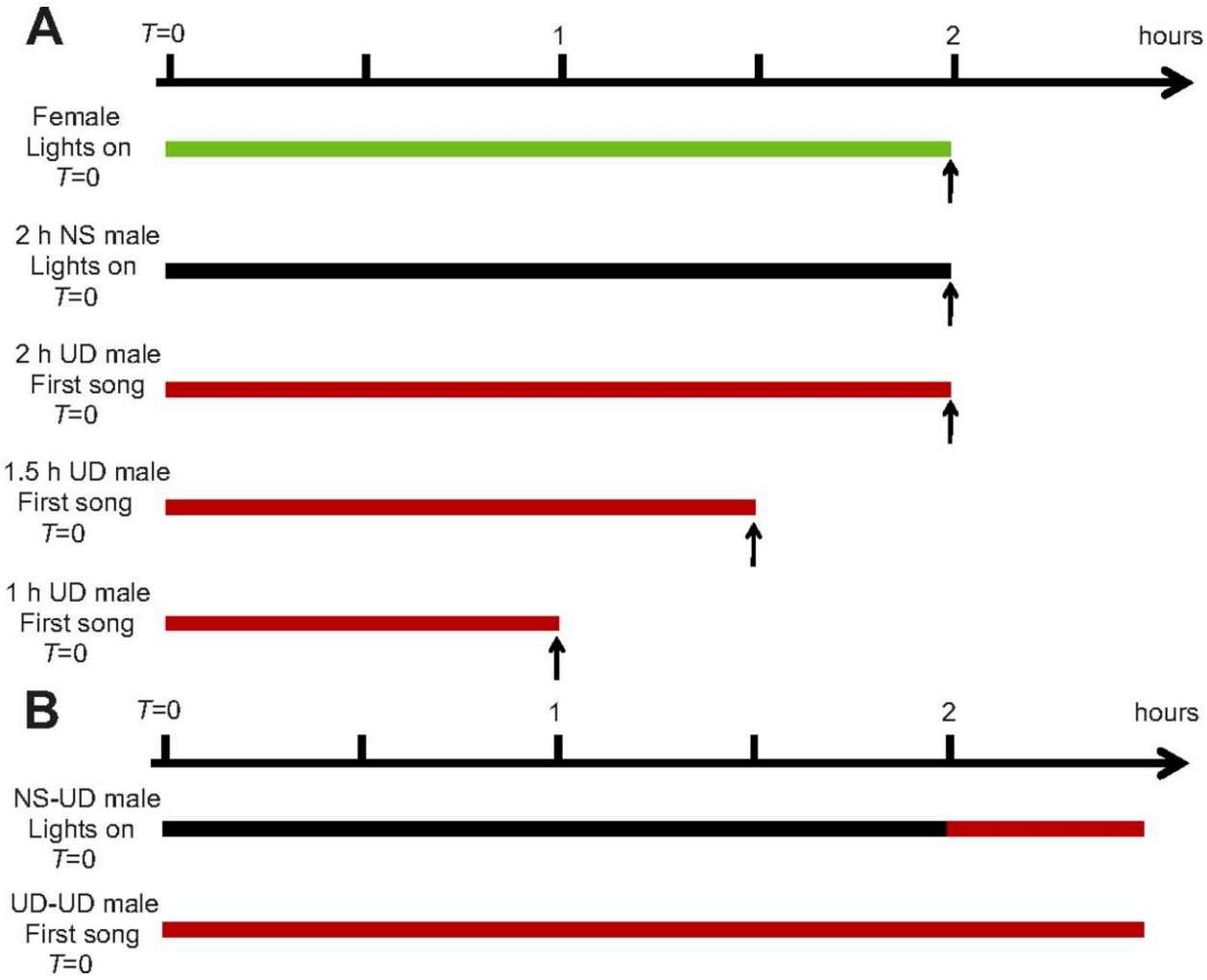
The amount of singing performed by each species could influence whether song is more or less variable in the UD-UD condition. Bengalese finches in our study sang roughly half as much as the zebra finches and the corresponding downregulation of FoxP2 is about half the magnitude. It is possible that FoxP2 levels must drop below a critical threshold in order to de-repress gene transcription and initiate molecular changes that lead to increased variability, or that the amount of singing by Bengalese finches was sufficient to downregulate FoxP2 mRNA but not the protein (Miller et al., 2008). These possibilities could be supported by examining Bengalese song after more extended bouts (~4h) of UD singing; however, this may be confounded by the fact that FoxP2 levels vary as a function of both the amount of singing and the total time allotted for singing (Fig.A1-7).

The age of the Bengalese finches used here (>300days) may present another confounding factor in our ability to detect differences in vocal variability between NS-UD and UD-UD birds. Increased song variability was previously observed to be correlated to the amount of song in younger adult zebra finches [N=18 between 120 and 200days old (Hilliard et al., 2012)]. Both Bengalese and zebra finches undergo age-related changes in vocal quality and the ability to exhibit vocal plasticity (Brainard and Doupe, 2001; Cooper et al., 2012), thus they may undergo age-related changes in how molecular microcircuits impact behavior. Further, age- and speciesrelated differences in basal vocal variability may have statistically limited our ability to detect these changes. In zebra finches, our ability to detect acute regulation of vocal variability was limited to 75-day-old juvenile birds, as 65-day-old birds and a group of six adult birds showed too much and too little variability, respectively, to derive adequate statistical power (Miller et al., 2010). A followup study found that statistical power was achieved when the number of adult zebra finches was increased to 18 UD singers with higher numbers of motifs

uttered in the 2h being correlated with increased song variability (Fig.A1-3B) (Hilliard et al., 2012).

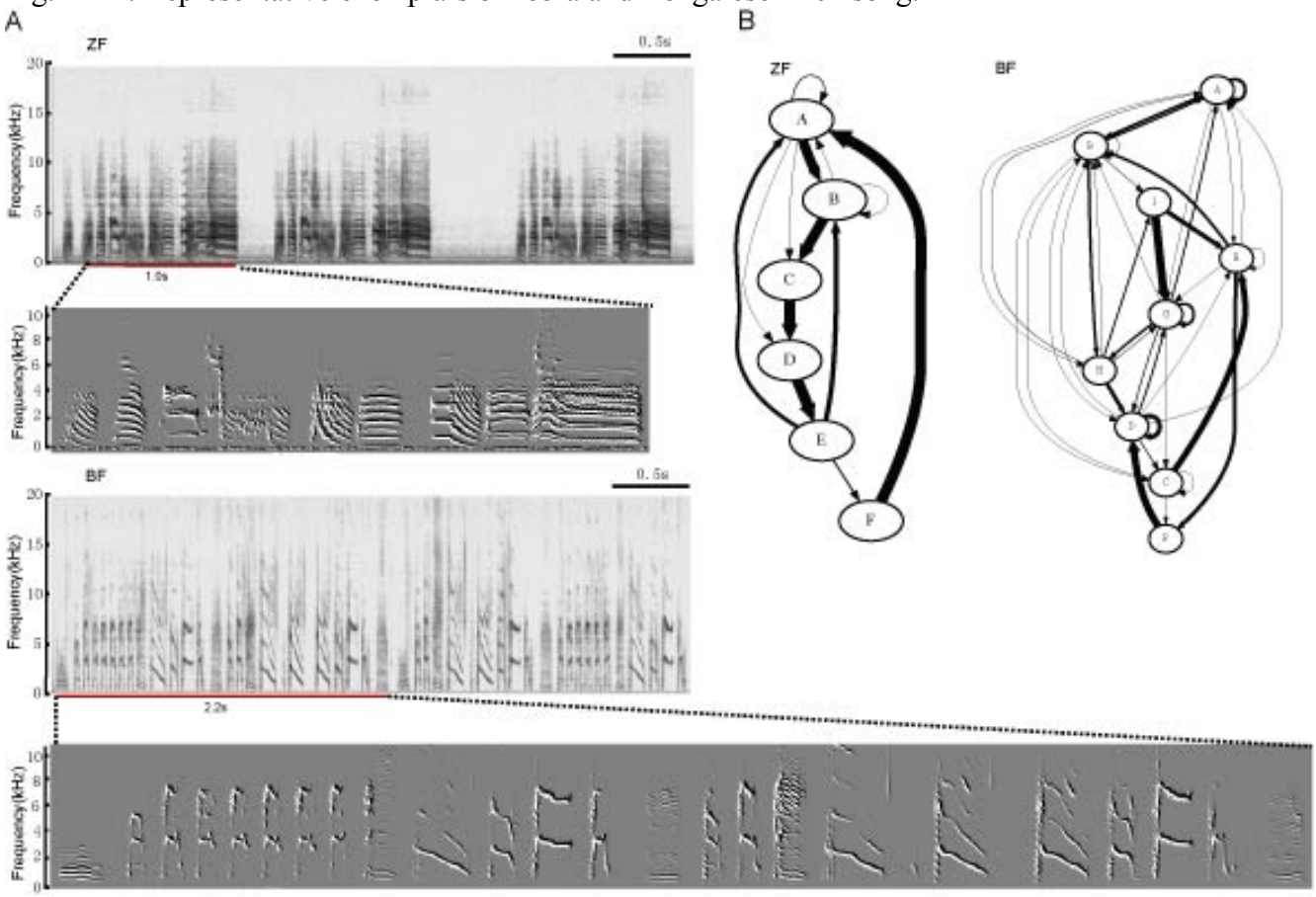
In summary, these data indicate that FoxP1 is enriched in most song control nuclei of male Bengalese finches, with the notable exception of the LMAN, similar to its expression pattern in zebra finches. No singing-driven regulation of this transcription factor was observed in either species, suggesting a sexually dimorphic role in the formation of brain structures that support vocal learning in songbirds. In contrast, FoxP2 levels in Area X do exhibit singingdriven decreases in both species, with a similar dependence on both the amount of singing and the time since song onset, with the caveat that Bengalese finches in our study sang less than zebra finches. The impact of this downregulation in zebra finches appears to be to increase vocal motor exploration, particularly during song learning and as evidenced by multiple prior studies. Here, in Bengalese finches, we did not observe a similar relationship, which could reflect a true species difference. We deem it more likely that the differences in age and amount of singing of the Bengalese finches in our study relative to the zebra finches precluded detection of this relationship. Future work in songbirds to examine protein expression of these factors as well as to genetically intervene in their expression promise to illuminate organizational versus activational functions of these molecules related to human language.

Fig. A1-1. Timelines for the behavioral groups used in this study.



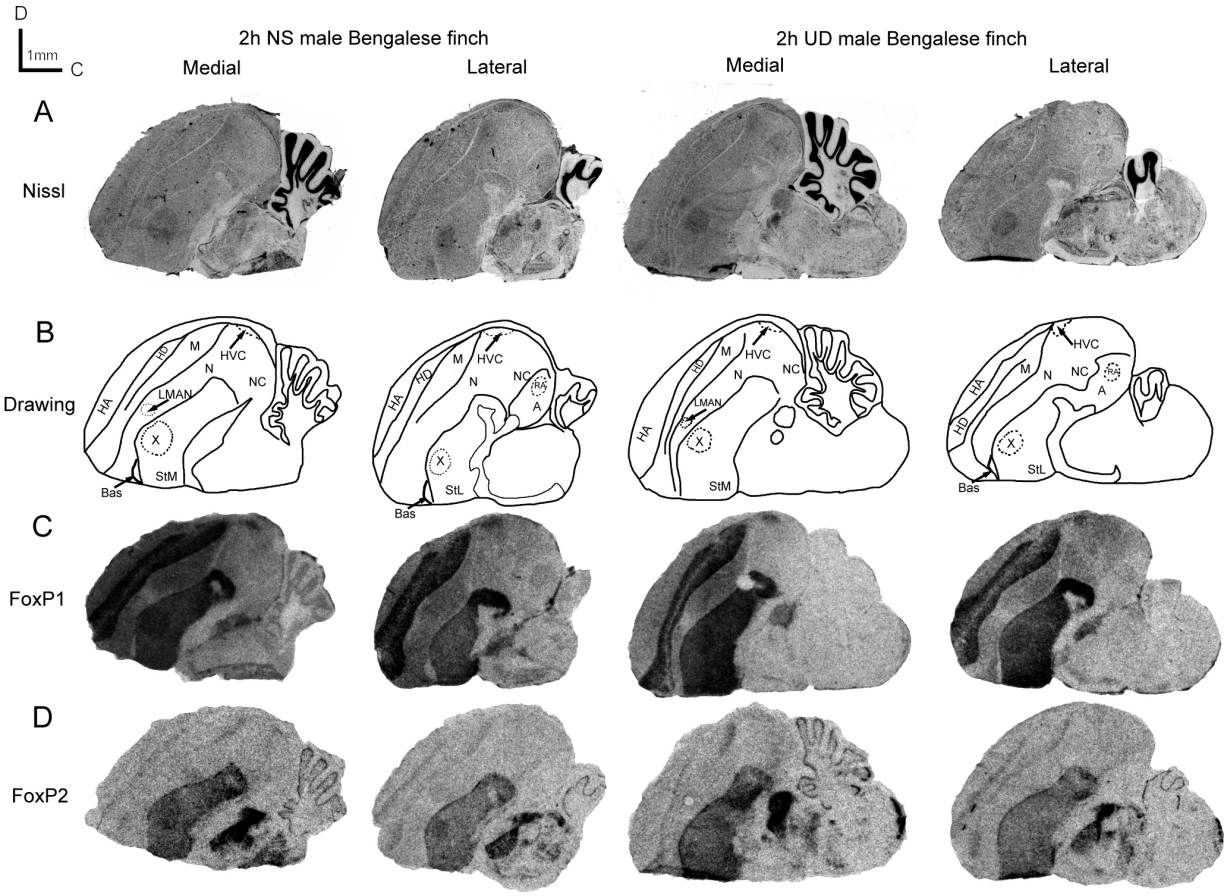
(A) Experimental design for time-course analysis of *FoxP1* and *FoxP2* behavioral regulation. On the day of the experiment, female birds remained alone in sound attenuation chambers for 2h (green bar). NS males were discouraged from singing by the experimenter sitting nearby for 2h (black bar). UD males sang alone in the isolation chamber for variable periods of time (red bars). Arrows indicate the timepoints of sacrifice. (B) Experimental design for song variability analysis. Songs sung after the 2h timepoint were analyzed for song variability. Birds were not sacrificed in this experiment.

Fig. A1-2. Representative exemplars of zebra and Bengalese finch song.



(A) Spectrograms from a male zebra finch (ZF, top) and a male Bengalese finch (BF, middle) are shown. The red bar underneath each spectrogram indicates the length of one motif. Spectral deriviatives of these motifsare shown underneath each spectrogram. (B) Markov chains generated from zebra finch and Bengalese finch songs. Letters denote syllables. Lines represent the probability of syllable transitions. Thicker lines indicate greater probabilities.

Fig. A1-3. *FoxP1* and *FoxP2* mRNA expression patterns in the adult male Bengalese finch brain.



Representative bright-field photomicrographs of *FoxP1* and *FoxP2* mRNA expression patterns in a series of sagittal sections from one 2h NS (left) and one 2h UD (right) adult male Bengalese finch brain. Both medial and lateral sections are shown to enable display of the song control nuclei investigated here. (A) Nissl-stained sagittal sections. Locations of medial and lateral sections correspond to the level of sagittal plates 6 and 11 in the zebra finch atlas of Nixdorf-Bergweiler and Bischof (2007), respectively. (B) Schematic drawings based on the Nissl stains. (C) *FoxP1* mRNA signals. (D) *FoxP2* mRNA signals. Medial sections in A, C and D were adjacent or near adjacent to one another; similarly, lateral sections were adjacent or near adjacent. Scale bar 1mm. D, dorsal; C, caudal.

Fig. A1-4 . Representative bright-field photomicrographs of *FoxP1* and *FoxP2* mRNA

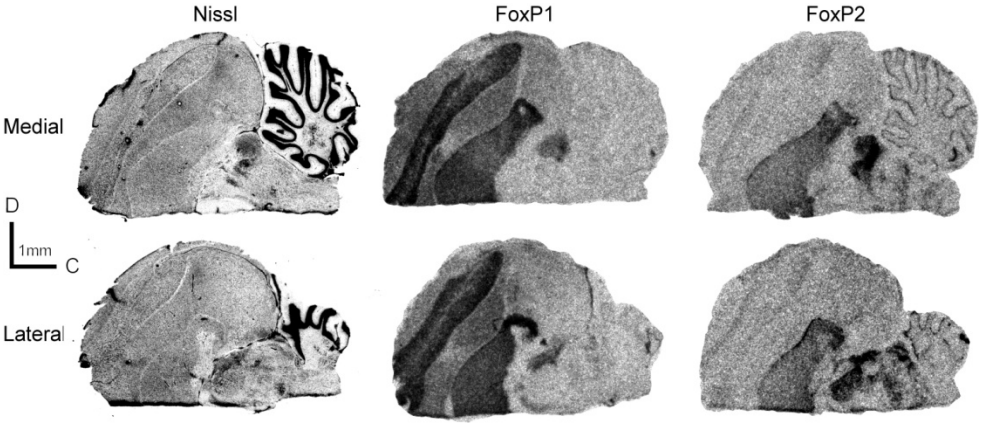


Fig. A1-4 . Representative bright-field photomicrographs of *FoxP1* and *FoxP2* mRNA expression patterns in a pair of sagittal sections from adult female Bengalese finch brain. Locations of medial and lateral sections correspond to the level of sagittal plates 6 and 11 in the zebra finch atlas of Nixdorf-Bergweiler and Bischof (2007), respectively. Medial plate shows HVC and LMAN and lateral plate shows HVC and RA in corresponding sections from male birds

Fig. A1-3. *FoxP1* and *FoxP2* mRNA expression patterns in the adult male Bengalese finch brain.

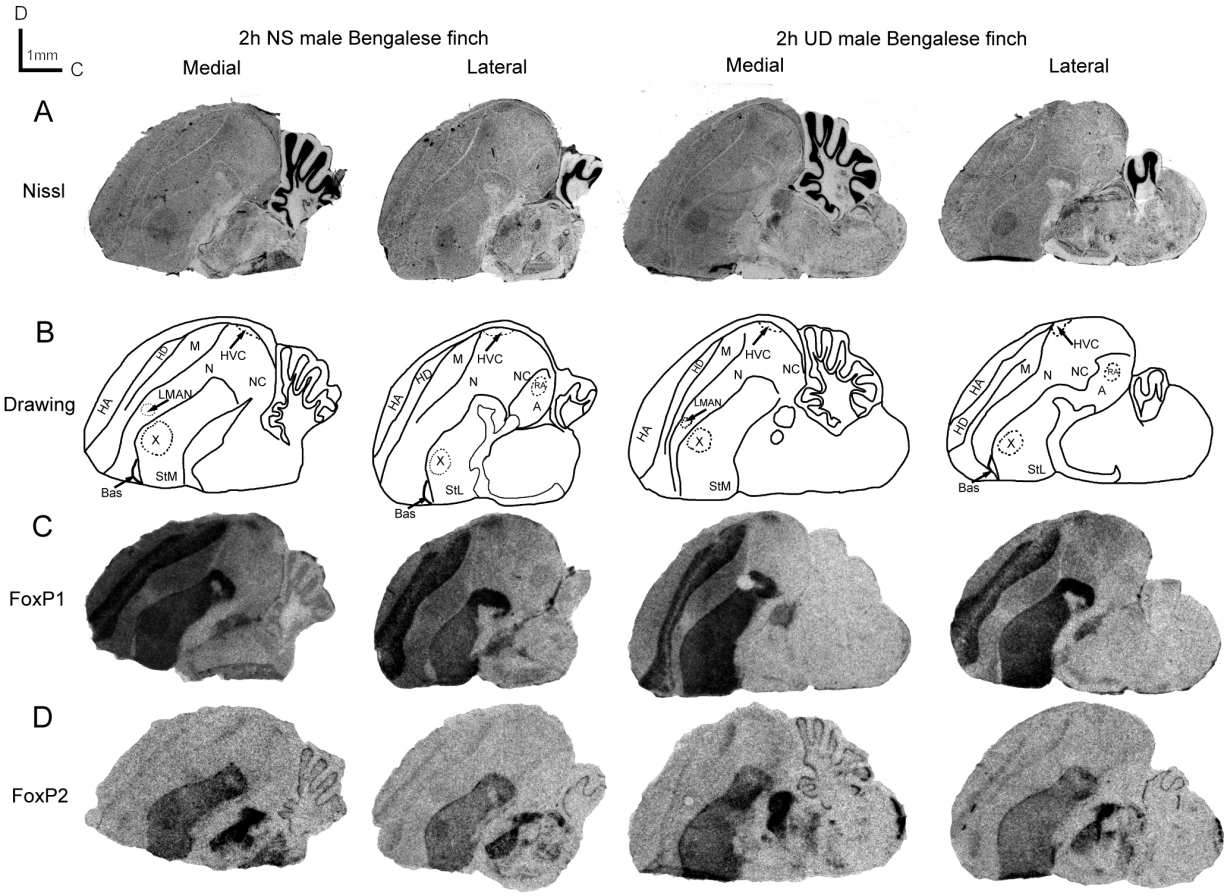
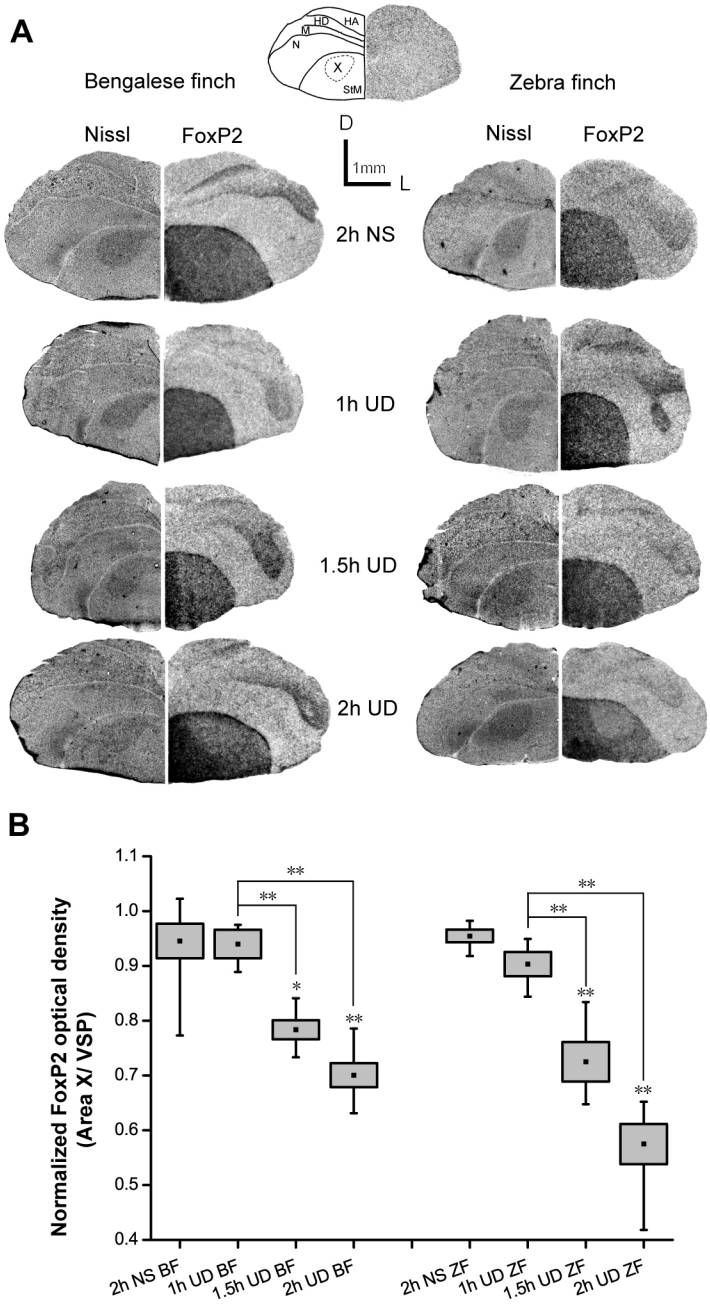


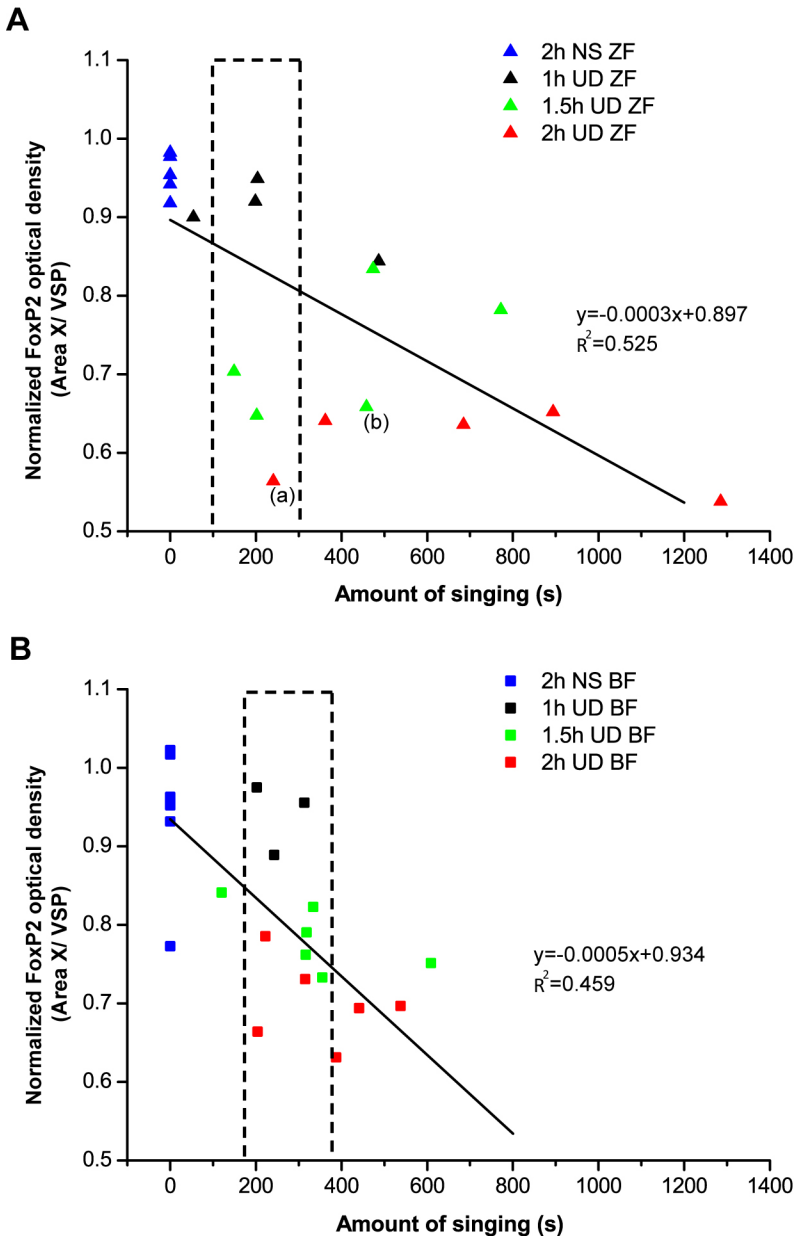
Fig. A1-2. *FoxP2* mRNA expression within Area X diminishes after birds sing undirected songs.



(A) Top - Schematic drawing based on a Nissl stained hemi-coronal section is shown together with a control hemi-section incubated with sense RNA. Beneath, representative bright-field photomicrographs of *FoxP2* mRNA expression patterns in hemi-coronal sections from Bengalese finches (left) and zebra finches (right) of different behavioral groups are shown together with corresponding Nissl-stained hemi-sections. Scale bar 1mm. D, dorsal; L, lateral. (B) Quantitative results of *FoxP2* mRNA expression level within Area X relative to VSP. Box indicates s.e.m..

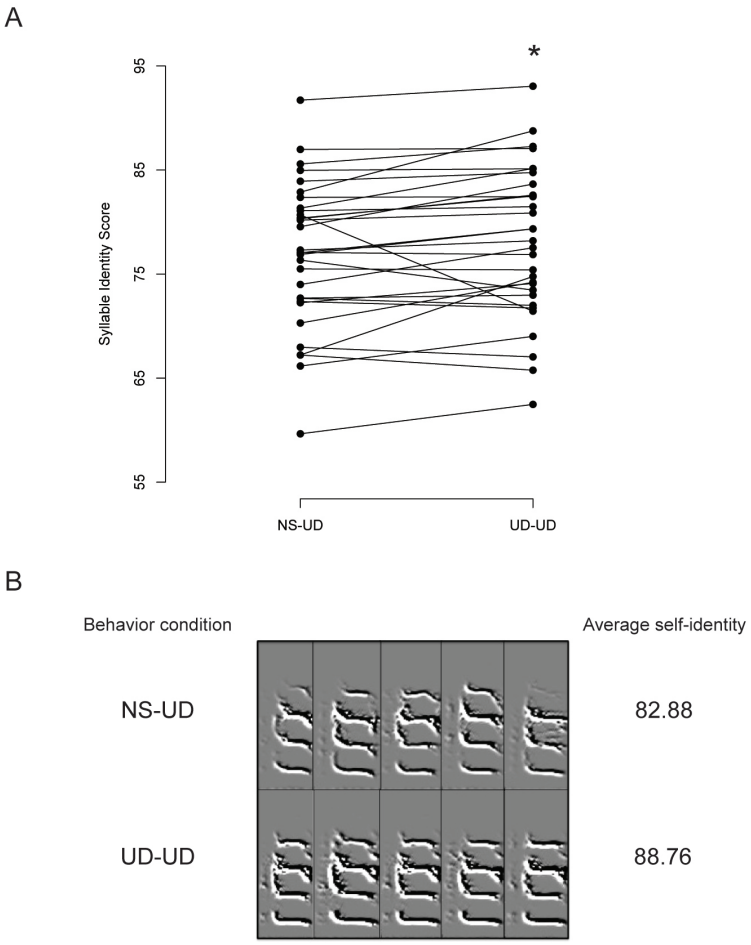
Dots in boxes indicate mean. Whiskers indicate max and min. 2h NS BF: n=7; 1h UD BF: n=3; 1.5h UD BF: n=6; 2h UD BF: n=6; 2h NS ZF: n=5; 1h UD ZF: n=4; 1.5h UD ZF: n=5; 2h UD ZF: n=6. Kruskal–Wallis nonparametric ANOVA, * $p<0.001$.

Fig. A1-3. Correlation between *FoxP2* and amount of singing.



(A) In zebra finches, *FoxP2* levels decrease as the amount of singing increases ($p<0.0002$) (B) *FoxP2* levels also decrease as the amount of singing increases in Bengalese finches ($p<0.0003$) There is no significant difference between the two regression lines ($p>0.05$). The dotted rectangle indicates data from those birds that sang similar amounts of song for each species (see Discussion).

Fig. A1-4. Behavioral changes in syllable self-identity.



(A) Paired plot of syllable cluster self-identity in NS-UD (left) and UD-UD (right) conditions. Denoted by an asterisk, the UD-UD condition had higher mean self-identity (p=0.034, 2-tailed paired bootstrap). (B) Representative spectral derivatives of five syllables from one cluster in the NS-UD (top) and UD-UD (bottom) conditions with self-identity scores reported (right).

Table A1-1. FoxP2 mRNA throughout Bengalese Finch brain

Mean OD	UD Male	FoxP1 NS Male	Female	UD Male	FoxP2 NS Male	Female
A	0.428	0.565	0.464	0.572	0.557	0.804
Area X	1.410	1.307	N/A	1.941	2.695	N/A
Bas	0.618	0.681	0.540	0.723	0.886	1.022
Dorsal thalamus	1.045	1.169	1.020	3.469	4.189	4.406
HA	0.906	0.986	0.952	1.005	0.965	0.887
HD	1.449	1.399	1.216	1.193	1.120	1.211
HVC	1.312	1.240	N/A	0.990	1.027	N/A
LMAN	0.591	0.732	0.544	0.915	1.072	1.022
M	1.448	1.396	1.232	1.238	1.199	1.288
N	1.000	1.000	1.000	1.000	1.000	1.000
RA	0.577	0.711	N/A	0.618	0.567	N/A
Striato-pallidum	1.300	1.296	1.166	2.419	2.484	2.744

Mean optical density (OD) values measured from multiple sections of one UD male Bengalese finch, one NS male Bengalese finch and one female Bengalese finch. All values are normalized to the mean value of the OD in nidopallium outside of song control areas.

Table A1-1.2 Changes in vocal variability following extended vocal practice

Mean Values	**NS-UD**	**SD**	**UD-UD**	**SD**	**p-value**
Pitch	0.290	0.251	0.282	0.272	0.7169
Frequency Modulation	0.147	0.068	0.136	0.058	0.1242
Entropy	0.127	0.053	0.120	0.055	**0.0376**
Pitch Goodness	0.226	0.121	0.193	0.110	**0.0002**
Mean Frequency	0.166	0.081	0.151	0.078	**0.0168**

Variance	**NS-UD**	**SD**	**UD-UD**	**SD**	**p-value**
Frequency Modulation	0.206	0.069	0.198	0.087	0.3800
Entropy	0.500	0.259	0.440	0.184	**0.0136**
Pitch Goodness	0.662	0.319	0.581	0.412	0.1078
Mean Frequency	0.895	0.678	0.775	0.328	0.1901

For each acoustic feature, the average CV and standard deviation (SD) for all syllable clusters within each behavioral condition is reported along with p-values generated by a 2-tailed paired bootstrap test. Significant values are shown in bold font.

Appendix 2: Peripheral androgen action helps modulate vocal production in a suboscine passerine.

Summary

Androgenic activation of intracellular androgen receptors (AR) influences avian vocal production, though this has largely been investigated at the level of the brain. We investigated the influence of predominantly peripheral AR on vocal output in wild Golden-collared Manakins (*Manacus vitellinus*). In this suboscine species, males court females by performing acrobatic displays and by producing relatively simple *chee-poo* vocalizations. To assess whether peripheral AR influences the acoustic structure of these vocal signals, we treated reproductively active adult males with the peripherally selective antiandrogen bicalutamide and then measured phonation performance. Inhibiting AR outside of the central nervous system increased the duration of the *chee* note and decreased the fundamental frequency of the *poo* note. This treatment caused no discernable change to *chee-poo* frequency modulation or entropy. Our results show that activation of peripheral AR mediates note-specific changes to temporal and pitch characteristics of the Golden-collared Manakin's main sexual call. Thus, our study provides one of the first demonstrations that androgenic action originating outside of the brain and likely on musculoskeletal targets can modulate avian vocal production.

Introduction

Androgenic hormones act via intracellular androgen receptors (AR) to influence vertebrate social behavior (Adkins-Regan 2005), and avian vocal production is a prime example of this trait. In some species, vocalizations are acoustically complex and require exquisite coordination between (1) central systems that govern sensorimotor and motor programming and (2) peripheral systems that govern the generation of sound (Schlinger 1997). In passerine birds, for example, males sing and/or call to attract mates (Catchpole and Slater 2008), and research has shown that androgens influence such behavior by changing the number of times that individuals sing and/or call (Silverin 1980,Nowicki and Ball 1989, Ketterson et al. 1992, P. G. McDonald et al. 2001, Kurvers et al. 2008), as well as the acoustic structure or makeup of these songs and/or calls (Deviche and Schumacher 1982,Groothuis and Meeuwissen 1992, Fusani et al. 1994, Galeotti et al. 1997, Cynx et al. 2005,Apfelbeck et al. 2012). However, in general, we know relatively little about where and how androgens act within the body to mediate acoustic parameters of bird songs and calls.

Most studies that have examined how androgens influence avian vocal production have focused at the level of the brain. The midbrain nucleus intercollicularis (nICO), for example, is an androgen-sensitive premotor region that regulates the calls of many species (Brown 1965, Cohen 1981, Cohen and Cheng 1982, Panzica et al. 1991). Additionally, in oscine songbirds, a higher-level song-control system underlies the learning and production of complex songs (see Jarvis et al. 2005), and androgenic hormones modulate this system in a way that presumably influences when and how songs are produced (Nottebohm 1980, Tramontin et al. 2003, Sartor et al. 2005). Despite this focus on the brain, it is also possible that androgens signal

via AR in peripheral parts of the body. Musculo-skeletal systems, for instance, can express abundant AR (Michel and Baulieu 1980,Brantley et al. 1993, Regnier and Herrera 1993, Bland 2000, Kawano et al. 2003, Monks et al. 2004,Feng et al. 2010, Wyce et al. 2010), and some of the tissues that compose these systems are essential for avian vocal production. The avian vocal organ, the syrinx, is a prime example: It not only expresses AR (Wade and Buhlman 2000, Veney and Wade 2004) but also influences acoustic parameters of vocal output, such as fundamental frequency (F_0), frequency modulation (FM) and entropy (Goller and Suthers 1996, Riede et al. 2006, Elemans et al. 2008, Secora et al. 2012). Thus, in principle, this means that androgens have the capacity to act not only on the brain, but also on the musculature and cartilaginous structures that influence vocal filtering and production (Deviche and Schumacher 1982, Fusani et al. 1994). To date, such effects have been studied infrequently.

Here, we examine how activation of AR mainly outside of the central nervous system (CNS) influences acoustic production in the Golden-collared Manakin (*Manacus vitellinus*). This suboscine passerine species inhabits Panamanian rainforests. Males regularly perform elaborate courtship displays that involve mechanical "wing-snaps" and rapid dancing routines over the forest floor (Schlinger et al. 2013). As part of this sexual repertoire, males also broadcast simple *chee-poo* calls (Figure A2-1), although such vocalizations are produced independently of physical maneuvering. Females use the *chee-poo* in choosing mates, which suggests that these calls are adaptive and that their underlying mechanisms are influenced by sexual selection (Barske et al. 2011). Thus, like many other manakin species (Durães et al. 2011), Golden-collared Manakins utilize vocalizations as an important component of their reproductive and territorial behavior.

Narrow-band spectrograph of a *chee-poo* call from a reproductively active adult male Golden-collared Manakin.

To inhibit AR primarily in the periphery, we treated reproductively active adult male Golden-collared Manakins with the antiandrogen bicalutamide (BICAL). This pharmacological agent blocks AR peripherally without affecting AR centrally (Freeman et al. 1989, Furr 1989). We have verified that BICAL acts in a peripherally selective manner in the study species, in that it significantly disrupts the expression profiles of known androgen-dependent genes peripherally but does not significantly affect expression of androgen-dependent genes in the brain (Fuxjager et al. 2013). Additionally, we have shown that, within days of treatment, BICAL decreases the rates at which males perform wing-snaps and courtship dances; BICAL administration does not, however, significantly influence the rate at which males produce *chee-poos* (Fuxjager et al. 2013). Those results suggest that inhibition of peripheral AR changes the physicality of male courtship behavior. Given that the syrinx (the avian vocal organ) of the Golden-collared Manakin expresses large amounts of AR compared with other passerine species (Feng et al. 2010), we asked, in the present study, whether blocking AR in peripheral tissues that contribute to phonation, such as the syrinx, over the same period similarly disrupts acoustic production.

To address this question, we assessed the acoustic features of the *chee-poos* recorded from those wild males treated with BICAL or control implants (from Fuxjager et al. 2013). We specifically focused on measurements of *chee-poo* pitch (F_0), degree of sound change over time (FM), and tonal purity (entropy), because these characteristics can be affected by peripheral sound-generating structures (Goller and Suthers 1996, Riede et al. 2006, Elemans et al. 2008, Secora et al. 2012). We also measured the duration of the notes within the *chee-poo*.

Methods

Experimental design

We studied reproductively active adult male Golden-collared Manakins during the height of the breeding season (February–April) at the Smithsonian Tropical Research Institute in Gamboa, Panama. Birds were captured via passive mist netting and then weighed, uniquely leg-banded for future identification, and randomly assigned to 1 of 2 treatment groups. In the first group, males (n = 6) received a time-release implant that emitted 0.25 mg day^{-1} of the peripherally selective antiandrogen BICAL for 21 days (Innovative Research of America, Florida, USA; dose = 12.5 mg kg^{-1} day^{-1}). In the second group, males (n = 6) received a control implant that was identical in every way but emitted no BICAL. Implants measured 1.6 × 5 mm (height × diameter) and were placed subcutaneously on the bird's back at the base of its neck. Implantation procedures are described in detail elsewhere (Fusani et al. 2007, Fuxjager et al. 2013). Notably, implantation is quick (~2 min) and does not complicate the birds' health or activity levels (Fuxjager et al. 2013).

Birds came from a total of 7 leks, with at least 2–10 birds lek^{-1}. In 5 of these leks that contained ≥4 displaying males, we used 2 birds lek^{-1} (each of these birds was assigned to a different treatment group). In 2 of these 7 leks that contained ≤3 displaying males, we used only 1 bird lek^{-1}. In one instance, this bird was assigned to the BICAL group, and in the other instance this bird was assigned to the control group. Ultimately, we obtained data from 4 males group^{-1}, given that some males (n = 2 group^{-1}) did not *chee-poo* during the tape-recorded observational session (see below).

Bicalutamide

In vertebrates, BICAL acts as a potent antiandrogen that blocks AR exclusively outside of the CNS (Freeman et al. 1989, Furr 1989, Furr and Tucker 1996). For example, Freeman et al. (1989) injected animals with radio-labeled BICAL and found significant accumulation of radioactivity in all of the peripheral organs examined, but not within the brain. Moreover, treatment with modest amounts of BICAL (sufficient to block peripheral AR) had no effect on the androgen-dependent mammalian hypothalamic–pituitary–gonadal axis (Freeman et al. 1989, Furr 1989). As noted above, we had previously validated the efficacy of BICAL in the study species by examining central and peripheral androgen-dependent gene expression: The BICAL-treated birds appeared to be healthy and displayed the same overall activity and locomotor abilities as nontreated birds (Fuxjager et al. 2013).

After implantation, males were immediately released onto the lek from which they were captured. Each bird returned to its respective display arena, and some individuals were witnessed displaying within minutes of implantation and release.

Chee-poo *Recordings*

Each bird was observed for a 10-day period after implantation. We selected this time frame because past work had shown that BICAL inhibited display behavior beginning on the first day after treatment and through the following 10 days thereafter (these data, including the frequency of *chee-poo* production, are provided in Fuxjager et al. 2013). Each observation session lasted 30 min, occurring between 0700 and 0900 hours and between 1200 and 1630 hours, when the birds' activity levels were highest (Stein and Uy 2006, Fusani et al. 2007). Observers sat ~10 m from the display arena and provided birds with a 15-min habituation period before collecting data. In a

randomly selected subset of observation sessions over the 10-day period, *chee-poos* were tape-recorded using a Sennheiser microphone (K6 series, model ME66) and a Sony TC-D5M Professional tape recorder (sample rate = 48 KHz; 16-bit dynamic range). Sound files were digitized from the recorder using Audacity Audio Editor. Multiple *chee-poos* from each individual were recorded (range: 3–6), and only *chee-poos* that were definitively determined to be from the focal animal were used for analysis.

We used Praat Phonetics software to generate spectrograms of each *chee-poo* (window length = 5 msec; dynamic range = 70 dB), and from these we measured both the duration and the F_0 of each note. Duration was determined by selecting the beginning and end of each note with the cursor; we carefully avoided inclusion of the echo at each note's end. The F_0 for each note was computed by averaging F_0 measurements determined (via Praat) at 10-msec intervals along each note's base (fundamental) frequency band. Finally, we used the free, open-code software Sound Analysis Pro to calculate the FM and Weiner entropy of each note.

Statistical Analysis

For each acoustic variable, we used a separate general linear mixed model to examine the effects of both BICAL treatment and note (*chee* vs. *poo*). As such, BICAL treatment and note were included in each model as a fixed factor, whereas bird identity was also included as a random factor. Significant interactions were followed by a calculation of the percent change between control and BICAL treatment for each note. We were unable to collect large numbers of *chee-poos* from each focal individual across the entire 10-day observation period, so we were unable to include treatment time in our model as a fixed factor.

Results

The *chee-poo* is a 2-note call in which a *chee* note always precedes a *poo* note (Figure 1). On average, the *chee* note by itself is shorter in duration than the *poo* note. The *chee* note also has a higher F_0 and greater entropy than the *poo* note. There is, however, no discernable difference in FM between these 2 separate notes (Table 1).

Although BICAL treatment does not induce a singular main effect on any of the measured acoustic parameters (Table 1), this antiandrogen does exert note-specific effects on specific acoustic characteristics. Namely, both the duration and F_0 of the *chee-poo* are influenced by a significant BICAL × note interaction (Table 1). With respect to call duration, this effect is driven by a BICAL-induced increase in the duration of the *chee* by ~21% (~20 msec), with virtually no effect on the duration of the *poo* (Figure 2A). The magnitude of this note-specific change in duration is greater than the likely error in measuring the onset and offset of sound in relatively low signal-to-noise ratios, which in our experience is ~3 msec. Next, for call F_0, BICAL has a negligible effect on the chee but decreases the F_0 of the *poo* by >5% (~115 Hz; Figure 2B). Neither FM nor entropy is affected by a BICAL × note interaction (Table 1).

Discussion

We examined how AR influences the structure and sound features of the Golden-collared Manakin's adaptive *chee-poo* call. We found that by using BICAL to block AR in a peripherally selective manner, we changed the call's temporal and pitch characteristics. These results suggest that activation of AR populations outside of the brain and spinal cord has the ability to mediate acoustic properties of a vocal signal in this species. Our results also show that BICAL treatment does not affect other sound elements of the *chee-poo*, such as the degree of frequency change over time (FM) and/or tonal purity (entropy). Effects on entropy may, in theory, be difficult to discern, because our recordings were obtained from animals in nature, a setting which is relatively more "noisy" than a controlled laboratory environment. However, metrics such as FM should be easily detected from wild birds in field settings; the absence of any effect of BICAL on this vocal parameter therefore points to the generally selective influence of BICAL on vocal output. In other words, the blockade of AR outside of the CNS does not induce a dramatic change in the *chee-poo* itself, but instead subtly changes the way in which certain components of vocal signals are produced. Our results provide compelling evidence that peripheral activation of AR plays a role in guiding avian phonation more-or-less independently of central activation of AR.

BICAL does not affect acoustic output by suppressing an individual's health or by altering its social motivation, the latter of which is likely driven by central actions of androgen (Fusani et al. 2007). Previous studies of these birds show that BICAL treatment does not change the rates at which males broadcast *chee-poos* around their lek (Fuxjager et al. 2013) and has no effect on an individual's activity, feeding behavior, and general social arousal (Fuxjager et al.

2013). Thus, BICAL treatment appears to drive the observed effects on sound characteristics of the *chee-poo* by inhibiting androgenic activity via AR on peripheral substrates that are either directly or indirectly related to sound generation.

Peripheral Androgens and Vocal Control

Birds given BICAL not only increase the duration of *chee* notes by ~20 msec (21%), but also reduce the F_0 of *poo* notes by ~115 Hz. These findings are consistent with other work that similarly shows androgen-dependent changes in the acoustic "content" of vocal production. In male Black Redstarts (*Phoenicurus ochruros*), for instance, inhibition of androgenic and estrogenic action induced shifts of ~300 Hz in frequency parameters of aggressive song (Apfelbeck et al. 2012). Likewise, in male Zebra Finches (*Taeniopygia guttata*), long-term testosterone implantation caused a ~100 Hz decrease in the F_0 of directed sexual song (Cynx et al. 2005). Given that the effects reported in these past studies generally mirror, in magnitude, those reported here, it is tempting to conclude that the peripheral AR influences song production in numerous avian species.

Equally interesting is that the effects of BICAL on acoustic output occur within days of treatment, which is consistent with BICAL's impact on physical display behavior (Fuxjager et al. 2013). Nonetheless, this result stands in contrast to work by Cynx et al. (2005), which showed that some effects of testosterone on acoustic output require a month to emerge. This difference may result from interspecific variation in AR expression in peripheral sound-producing tissues such as the syrinx, given that Golden-collared Manakins express more AR mRNA in this organ than Zebra Finches (Feng et al. 2010). Of course, we cannot rule out the possibility that effects of

BICAL on manakin vocal production differ in other ways in response to longer-term AR blockade.

These data raise two important questions: (1) Where do androgens act in the periphery to influence vocal duration and pitch? and (2) What do androgens do to these tissues to effectuate such changes? Musculoskeletal systems that modify sound production include the syrinx, expiratory and intercostal muscles, and upper vocal tract (Wild et al. 1998, Suthers et al. 2002, Goller and Riede 2013); thus, these tissues may be substrates on which peripheral AR influences vocal production. We suspect that the syrinx is the prime organ through which this occurs, because it sits at the tracheo–bronchial junction and controls intertra-cheal sound-generating labia (Goller and Riede 2013). The muscles and labia of the syrinx contain AR, which makes these tissues susceptible to functional and/or morphological changes in response to androgenic action (Veney and Wade 2004, Feng et al. 2010). Blocking syringeal AR may therefore alter (1) the ability of the organ's musculature to appropriately control labial movement during expiration and/or (2) the structural constitution of the extracellular matrix and epithelium that make up and determine the labia's oscillatory (i.e. sound-generating) properties. Both of these tissues respond to steroid hormones, including androgens (Luine et al. 1983, Abitbol et al. 1999, Wade and Buhlman 2000, Chan et al. 2007).

These results do not exclude the possibility that activation of peripheral AR modulates central systems that regulate manakin vocal production. Stimulation of AR in skeletal muscle can influence the morphology of innervating motor neurons (Rand and Breedlove 1995), such that inhibition of peripheral AR may change the properties of the retrograde signaling that underlies muscle–CNS feedback. These effects may influence how the *chee-poo* is produced and may even

explain the change in call duration, given evidence that the brain controls this acoustic variable (Long and Fee 2008).

Functional Significance

In Golden-collared Manakins, circulating testosterone is elevated at the onset of the breeding season and activates display behavior (Schlinger et al. 2013). We suspect that elevated androgen levels act to fine-tune acoustic performance. As a consequence, BICAL treatment likely induces a peripherally specific "nonreproductive" state by blocking AR exclusively outside of the brain (Day et al. 2007,Schlinger et al. 2013). The shifts in acoustic parameters that we document here appear to deviate significantly from the apparent natural variation that otherwise exists in control birds, even though the magnitude of these changes in and of themselves is relatively small. Birds listening to the calls of treated males are therefore likely to perceive such differences in acoustic content that we document (Nelson 1988), and this may explain why information from the *chee-poo* is supposedly used both by females in choosing mates (Barske et al. 2011) and by males in competing with one another (D. B.McDonald et al. 2001).

It is more difficult to assess how acoustic content affects *chee-poo* function, because we know so little about how note duration, F_0, FM, and entropy are related to female choice and/or male–male interactions. Most studies that have attempted to manipulate these acoustic parameters did so in a way that simultaneously altered other factors intrinsic to male quality (McDonald 1989, Alatalo et al. 1990). Work on Zebra Finches has avoided this limitation and shown that gross manipulations of vocal production dramatically shape courtship success (Tomaszycki and Adkins-Regan 2005). In the case of the *chee-poo*, we expect that call duration

and F_0 similarly contain salient information that is relevant to social interactions, including the solicitation of female copulations. In particular, these features of the call may be honest indicators of male quality, given that they are guided by androgenic action, which is considered "costly" (Ketterson et al. 1992).

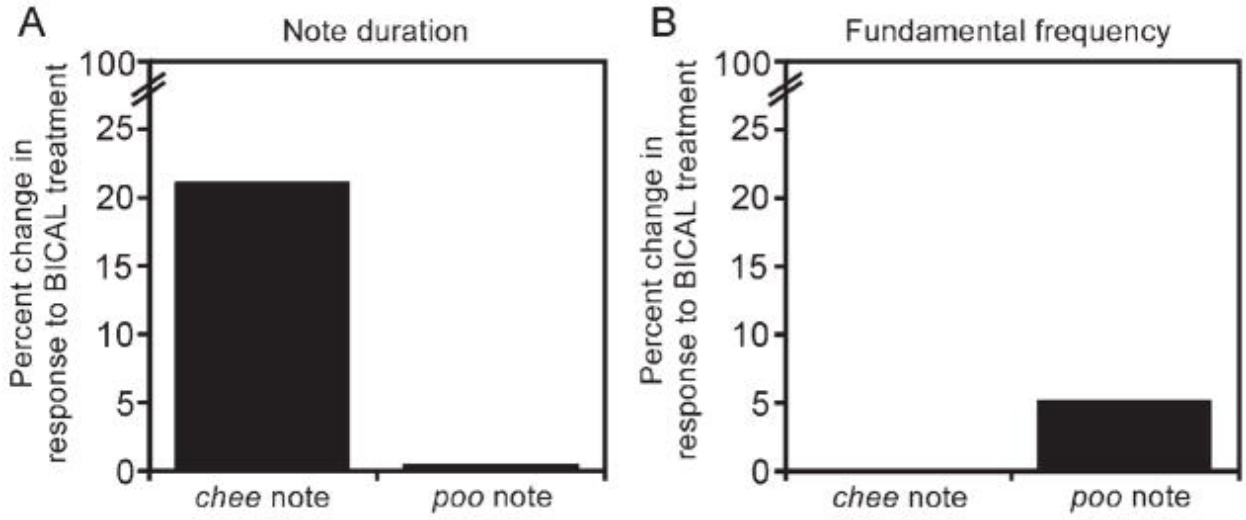
Phylogenetic Considerations

Our results highlight that androgens are capable of modulating the song of a suboscine passerine. Most work investigating the effects of sex steroids on vocal performance have utilized oscine passerine birds (Barker et al. 2004). One of the main functional characteristics that distinguish these suborders is the inability of suboscine birds, including Golden-collared Manakins (Saldanha et al. 2000), to learn songs during development and the lack of any defined song-control system in the brain (Kroodsma and Konishi 1991). In suboscines, it is likely that androgens act primarily on the midbrain nICO to drive the motor programming of call production (Cohen 1981, Cohen and Cheng 1982). It is also possible that suboscine birds rely on androgenic mediation of peripheral substrates as a means of sound control. Future work should more closely consider the contributions of peripheral and central androgenic action on avian vocal production, particularly in the suboscine avian suborder.

TABLE 1

Sound characteristics of the Golden-collared Manakin's *chee-poo* vocalization and linear mixed-model results for the effects of bicalutamide (BICAL) treatment and note (*chee* vs. *poo*). Statistical results in bold denote significant effects ($P <$ **...**

FIGURE A2-2



Note-specific changes in acoustic parameters of the *chee-poo* after bicalutamide (BICAL) treatment. (**A**) Percent change in the length, or duration, of each note. (**B**) Percent change in the fundamental frequency (F_0) of each note.

References

Abe, K., Matsui, S., and Watanabe, D. (2015). Transgenic songbirds with suppressed or enhanced activity of CREB transcription factor." *Proc. Natl. Acad. Sci. USA* **112**: 7599-7604.

Adkins-Regan, E., 2005. Hormones and Animal Social Behavior. Princeton University Press, Princeton.

Agate, R. J., Scott, B. B., Haripal, B., Lois, C., & Nottebohm, F. (2009). Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. Proceedings of the National Academy of Sciences, 106(42), 17963-17967.

Alcock, K. J., Passingham, R. E., Watkins, K. E. and Vargha-Khadem, F. (2000). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain and Language.***75**(1), 17-33.

Amador, A., Goller, F., Mindlin, G.B., 2008. Frequency modulation during song in a suboscine does not require vocal muscles. J. Neurophysiol. 99, 2383-2389.

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**, 12518-23.

Apfelbeck, B., Kiefer, S., Mortega, K.G., Goymann, W., Kipper, S., 2012. Testosterone affects song modulation during simulated territorial intrusions in male black redstarts (Phoenicurus ochruros). PLoS One 7.

Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S., & Roth, B. L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. Proceedings of the National Academy of Sciences, 104(12), 5163-5168.

Arnold, A.P., 1975. Effects of castration and androgen replacement on song, courtship, and aggression in zebra finches (*Poephila guttata*). J. Exp. Zool. 191, 309-325.

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* **320**, 630-4

Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate?. Brain and Language, 124(1), 96-116.

Arriaga, G., Zhou, E. P. and Jarvis, E. D. (2012). Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. *PLoS ONE* **7**(10): e46610.

Barker, F.K., Cibois, A., Schikler, P., Feinstein, J., Cracraft, J., 2004. Phylogeny and diversification of the largest avian radiation. Proc. Natl. Acad. Sci. USA 101, 11040-11045.

Barske, J., Schlinger, B.A., Wikelski, M., Fusani, L., 2011. Female choice for male motor skills. Proc. R. Soc. Lond., Ser. B: Biol. Sci. 278, 3523-3528.

Belton, E., Salmond, C. H., Watkins, K. E., Vargha□Khadem, F., & Gadian, D. G. (2003). Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. Human brain mapping, 18(3), 194-200.

Böhner J (1990) Early acquisition of song in the zebra finch, Taeniopygia guttata. Anim. Behav. **39**, 369–374

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* **224**, 901-3

Brainard, M. S. and Doupe, A. J. (2000). Auditory feedback in learning and maintenance of vocal behaviour. *Nature reviews. Neuroscience.* **1**(1), 31-40.

Brainard, M. S. and Doupe, A. J. (2001). Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions. *J Neurosci.* **21**(7):2501-17.

Brantley, R.K., Marchaterre, M.A., Bass, A.H., 1993. Androgen effects on vocal muscle structure in a teleost fish with intersexual and intrasexual dimorphism. J. Morphol. 216, 305-318.

Brenowitz, E., Margoliash, D. and Nordeen, K. (1997). An introduction to birdsong and the avian song system. *J Neurobiol.* **33**(5):495-500.

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**, 302-17

Burns, C.M.B., Rosvall, K.A., Ketterson, E.D., 2013. Neural steroid sensitivity and aggression: comparing individuals of two songbird subspecies. J. Evol. Biol. 26, 820-831.

Canoine, V., Fusani, L., Schlinger, B., Hau, M., 2007. Low sex steroids, high steroid receptors: increasing the sensitivity of the nonreproductive brain. Dev. Neurobiol. 67, 57-67.

Carlsson P. and Mahlapuu M. (2002). Forkhead transcription factors: key players in development and metabolism. *Dev Biol.* **250**(1):1-23.

Carr, C. W., Moreno-De-Luca, D., Parker, C., Zimmerman, H. H., Ledbetter, N., Martin, C. L., Dobyns, W. B. and Abdul-Rahman O.A. (2010). Chiari I malformation, delayed gross

motor skills, severe speech delay, and epileptiform discharges in a child with FOXP1 haploinsufficiency. *European Journal of Human Genetics.* **18**(11), 1216-1220.

Catchpole, C.K., Slater, P.J.B., 2008. Bird Song: Biological themes and variations. Cambridge University Press, Cambridge.

Chapman, F.M., 1935. The courtship of Gould's manakin (*Manacus vitellinus vitellinus*) on Barro Colorado Island, Canal Zone. Bulletin of the American Museum of Natural History 68, 472-521.

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**, 3682-92

Clovis Y. M., Enard W., Marinaro F., Huttner W. B. and De Pietri Tonelli D. Convergent repression of Foxp2 3'UTR by miR-9 and miR-132 in embryonic mouse neocortex: implications for radial migration of neurons. *Development.* **139**(18):3332-42.

Cooper B. G., Méndez J. M., Saar S., Whetstone A. G., Meyers R. and Goller F. (2012). Age-related changes in the Bengalese finch song motor program. *Neurobiol Aging.* **33**(3):564-8.

Cynx, J., Bean, N.J., Rossman, I., 2005. Testosterone implants alter the frequency range of zebra finch songs. Horm. Behav. 47, 446-451.

Deshpande M, Pirlepesov F, Lints T (2014) Rapid encoding of an internal model for imitative learning *Proc. Biol. Sci.* **281**

Deviche, P., Schumacher, M., 1982. Behavioral and morphological dose responses to testosterone and to 5-alpha-dihydrotestosterone in the castrated male japanese quail. Behav. Processes 7, 107-121.

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci*. **22**, 567-631

Ebbinghaus, H. (1885). Über das gedächtnis: untersuchungen zur experimentellen psychologie.

Elemans, C.P.H., Spierts, I.L.Y., Muller, U.K., van Leeuwen, J.L., Goller, F., 2004. Superfast muscles control dove's trill. Nature 431, 146-146.

Feng, N.Y., Katz, A., Day, L.B., Barske, J., Schlinger, B.A., 2010. Limb muscles are androgen targets in an acrobatic tropical bird. Endocrinology 151, 1042-1049.

Fiete IR, Senn W, Wang CZH, Hahnloser RHR (2010) Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron* **65**, 563-76

Forkhead Transcriptional Repressors. *Journal of Biological Chemistry.* **278**(27), 24259-24268.

Freeman, S.N., Mainwaring, W.I.P., Furr, B.J.A., 1989. A possible explanation for the peripheral selectivity of a novel non-steroidal pure antiandrogen, Casodex (ICI 176,334). Br. J. Cancer 60, 664-668.

Furr, B.J.A., 1989. Casodex (ICI - 176,334): A new, pure, peripherally-selective anti-androgen - preclinical studies. Hormone Res. 32, 69-76.

Fusani, L., Beani, L., Dessifulgheri, F., 1994. Testosterone affects the acoustic structure of the male call in the grey partridge (*Perdix perdix*). Behaviour 128, 301-310.

Fusani, L., Day, L.B., Canoine, V., Reinemann, D., Hernandez, E., Schlinger, B.A., 2007. Androgen and the elaborate courtship behavior of a tropical lekking bird. Horm. Behav. 51, 62-68.

Fuxjager, M.J., Forbes-Lorman, R.M., Coss, D.J., Auger, C.J., Auger, A.P., Marler, C.A., 2010. Winning territorial disputes selectively enhances androgen sensitivity in neural pathways related to motivation and social aggression. Proc. Natl. Acad. Sci. USA 107, 12393-12398.

Fuxjager, M.J., Longpre, K.M., Chew, J.G., Fusani, L., Schlinger, B.A., 2013. Peripheral androgen receptors sustain the acrobatics and fine motor skill of elaborate male courtship. Endocrinology 154, 3168-3177.

Gale, S. D. and Perkel, D. J. (2010). Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. *Journal of Chemical Neuroanatomy*. **39**(2), 124-131.

Galeotti, P., Saino, N., Sacchi, R., Moller, A.P., 1997. Song correlates with social context, testosterone and body condition in male barn swallows. Anim. Behav. 53, 687-700.

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**, 3343-51

Goller, F., Riede, T., 2013. Integrative physiology of fundamental frequency control in birds. Journal of Physiology-Paris 107, 230-242.

Goller, F., Suthers, R.A., 1996. Role of syringeal muscles in controlling the phonology of bird song. J. Neurophysiol. 76, 287-300.

Groothuis, T., Meeuwissen, G., 1992. The influence of testosterone on the development and fixation of the form of displays in two age classes of young black-headed gulls. Anim. Behav. 43, 189-208.

Groszer M, et al. (2008) Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol.* **18**, 354-62

Haesler S, Rochefort C, Georgi B, Licznerski 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

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**, 3164-75

Hamdan, F. F., Daoud, H., Rochefort, D., Piton, A., Gauthier, J., Langlois, M., Foomani, G., Dobrzeniecka S., Krebs M.O. and Joober R. et al. (2010). De Novo Mutations in FOXP1 in Cases with Intellectual Disability, Autism, and Language Impairment. *The American Journal of Human Genetics*. **87**(5), 671-678.

Hampton CM, Sakata JT, Brainard MS (2009) An avian basal gangliaforebraincircuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. J Neurophysiol 101:3235–3245

Hessler NA, Doupe AJ (1999) Social context modulates singing related neural activity in the songbird forebrain. Nat Neurosci 2:209–211

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**, 537-52

Ho, O. H., Delgado, J. Y., & O'Dell, T. J. (2004). Phosphorylation of proteins involved in activity-dependent forms of synaptic plasticity is altered in hippocampal slices maintained in vitro. Journal of neurochemistry, 91(6), 1344-1357.

Hopf, F.W., Seif, T., Mohamedi, M. L., Chen, B. T., & Bonci, A. (2010). The small□ conductance calcium□activated potassium channel is a key modulator of firing and long□term depression in the dorsal striatum. European Journal of Neuroscience, 31(11), 1946-1959.

Horn, D., Kapeller, J., Rivera-Brugués, N., Moog, U., Lorenz-Depiereux, B., Eck, S., Hempel, M., Wagenstaller J., Gawthrope A. and Monaco A.P. et al. (2010). Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. *Human Mutation.* **31**(11), E1851-E1860.

Huang, Y. C., & Hessler, N. A. (2008). Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. PloS one, 3(10), e3281.

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. *Science* **286**, 2526-8

Immelmann K. Song development in the zebra finch and other estrildid finches (1969) In *Bird Vocalizations*, R.A. Hinde, ed. (Cambridge: Cambridge University Press), 64–74

Jürgens U. (2009). The neural control of vocalization in mammals: a review. *J Voice.* **23**(1):1-10.

Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* **433**, 638-43

Kao MH, Wright BD, Doupe AJ (2008) Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28:13232–13247

Ketterson, E.D., Nolan, V., 1992. Hormones and life histories: an integrative approach. Am. Nat. 140, S33-S62.

Ketterson, E.D., Nolan, V., Wolf, L., Ziegenfus, C., 1992. Testosterone and avian life histories: effects of experimentally elevated testosterone on behavior and correlates of fitness in the dark-eyed junco (*Junco hyemalis*). Am. Nat. 140, 980-999.

Kikusui, T., Nakanishi, K., Nakagawa, R., Nagasawa, M., Mogi, K. and Okanoya, K. (2011). Cross Fostering Experiments Suggest That Mice Songs Are Innate. (B. Brembs, Ed.) *PLoS ONE.* **6**(3), e17721.

Kroodsma, D.E., 1984. Songs of the alder flycatcher (*Empidonax alnorum*) and willow f lycatcher (*Empidonax traillii*) are innate. Auk 101, 13-24.

Kroodsma, D.E., Konishi, M., 1991. A suboscine bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. Anim. Behav. 42, 477-487.

Kurvers, R., Roberts, M.L., McWilliams, S.R., Peters, A., 2008. Experimental manipulation of testosterone and condition during molt affects activity and vocalizations of male blue tits. Horm. Behav. 54, 263-269.

Kugler, S., Lingor, P. , Scholl, U., Volutukhin., S., and Bahr, M. (2003) Differential expression in brain cells in vivo and in vitro from AAV-2 vectors with small transcriptional units. Virology **311**: 89-95

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**, 519-23

Langfelder, P. and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *Bioinformatics*. **9**, 559.

Langfelder, P., Zhang, B. and Horvath, S. (2008). Defining clusters from a hierarchical tree: the Dynamic Tree Cut packaged for R. *Bioinformatics.* **24**(5), 719-720.

lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* **11**, 2896–2913.

Li, S., Weidenfeld, J. and Morrisey, E. E. (2003). Transcriptional and DNA Binding Activity of the Foxp1/2/4 Family Is Modulated by Heterotypic and Homotypic Protein Interactions. *Molecular and Cellular Biology,* **24**(2), 809-822.

Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. Nature neuroscience, 6(11), 1230-1237.

Liston C, Jenneteau F, Jia Z, Chao MV, Gan WB (2013) Circadian glucocorticoid oscillations promote learning-dependent synapse formation and maintenance. *Nat. Neurosci.* **16**, 698-705

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. *Nat. Neurosci.* **11**, 579-86

Liu, W. C., Kohn, J., Szwed, S. K., Pariser, E., Sepe, S., Haripal, B., Oshimori, N., Marsala, M., Miyanohara, A., & Lee, R. (2015). Human mutant huntingtin disrupts vocal learning in transgenic songbirds. Nature neuroscience, 18(11), 1617-1622.

Lu, M. M., Li, S., Yang, H. and Morrisey, E. E. (2002). Foxp4: a novel member of the Foxp subfamily of winged-helix genes co-expressed with Foxp1 and Foxp2 in pulmonary and gut tissues. *Mechanisms of development.* 119 Suppl 1, S197-202.

Mehaffey W. H., Doupe A. J. (2015). Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex. Nat. Neurosci. 18, 1272–1280.

Mallet, N., Micklem, B. R., Henny, P., Brown, M. T., Williams, C., Bolam, J. P., Nakamura, K.C. & Magill, P. J. (2012). Dichotomous organization of the external globus pallidus. Neuron, 74(6), 1075-1086.

Mallet, N., Schmidt, R., Leventhal, D., Chen, F., Amer, N., Boraud, T., & Berke, J. D. (2016). Arkypallidal Cells Send a Stop Signal to Striatum. Neuron.

Marler, C.A., Moore, M.C., 1988. Evolutionary costs of aggression revealed by testosterone manipulations in free-living male lizards. Behav. Ecol. Sociobiol. 23, 21-26.

McDonald, P.G., Buttemer, W.A., Astheimer, L.B., 2001. The influence of testosterone on territorial defence and parental behavior in male free-living rufous whistlers, Pachycephala rufiventris. Horm. Behav. 39, 185-194.

Michel, G., Baulieu, E.E., 1980. Androgen receptor in rat skeetal muscle: characterization and physiological variations. Endocrinology 107, 2088-2098.

Miller JE, Hilliard AT, White SA (2010) Song practice promotes acute vocal variability at a key stage of sensorimotor learning. *Plos One* **5**, e8592

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**, 2015-25

Murugan M, Harward S, Scharff C, Mooney R (2013) Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron* **80**, 1464-76

Nixdorf-Bergweiler, B. E. and Bischof, H. (2007). A Stereotaxic Atlas Of The Brain Of The Zebra Finch, *Taeniopygia Guttata*, With Special Emphasis On Telencephalic Visual And Song System Nuclei in Transverse and Sagittal Sections., 1-86.

Nottebohm, F. (2005). The Neural Basis of Birdsong. *PLoS Biology.* **3**(5), e164.

Nottebohm, F., 1980. Testosterone triggers growth of brain vocal control nuclei in adult female canaries. Brain Res. 189, 429-436.

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

Nowicki, S., Ball, G.F., 1989. Testosterone induction of song in photosensitive and photorefractory male sparrows. Horm. Behav. 23, 514-525.

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation inthe juvenile songbird requires a basal ganglia circuit. PLoS Biol3:e153.

Okanoya, K. (2004). The Bengalese Finch: A Window on the Behavioral Neurobiology of Birdsong Syntax. *Annals of the New York Academy of Sciences.* **1016**(1), 724-735.

Okanoya, K. and Yamaguchi, A. (1997). Adult Bengalese finches (Lonchura striata var. domestica) require real-time auditory feedback to produce normal song syntax. *J Neurobiol.* **33**(4):343-56.

Onn, S.P., Fienberg, A.A., & Grace, A.A. (2003). Dopamine modulation of membrane excitability in striatal spiny neurons is altered in DARPP-32 knockout mice. Journal of Pharmacology and Experimental Therapeutics, 306(3), 870-879.

Pagliardini, S., Janczewksi, W.A., Tan, W., Dickson, C.T., Deisseroth, K., Feldman, J.L. (2011) Active expiration induced by excitation of ventral medulla in adult anesthetized rats. J. Neurosci., 31(8), 2895-2905.

Panaitof SC, Abrahams BS, Dong H, Geschwind DH & White SA (2010) Language-related Cntnap2 is differentially expressed in sexually dimorphic nuclei essential for vocal learning in songbirds. J Comp Neurol, 518:1995-2018

Pariani, M. J., Spencer, A., Graham, J. M. and Rimoin, D. L. (2009). A 785kb deletion of 3p14.1p13, including the FOXP1 gene, associated with speech delay, contractures, hypertonia and blepharophimosis. *European Journal of Medical Genetics.* **52**(2-3), 123-127.

Price, P.H. (1979). Developmental determinants of structure in zebra finch song. *J. Comp Physiol Psychol* vol. **93** (2) pp. 260–277

Ravbar P, Lipkind D, Parra LC, Tchernichovski O (2012) Vocal exploration is locally regulated during song learning. J. Neurosci. **32**, 3422-32

Regnier, M., Herrera, A.A., 1993. Differential sensitivity to androgens within a sexually dimorphic muscle of male frogs (*Xenopus laevis*). J. Neurobiol. 24, 1215-1228.

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**, 1454-9

Rosvall, K.A., Burns, C.M.B., Barske, J., Goodson, J.L., Schlinger, B.A., Sengelaub, D.R., Ketterson, E.D., 2012. Neural sensitivity to sex steroids predicts individual differences in aggression: implications for behavioural evolution. Proc. R. Soc. Lond., Ser. B: Biol. Sci. 279, 3547-3555.

Rousso D.L., Gaber Z.B., Wellik D., Morrisey E.E. and Novitch B.G. (2009). Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. *Neuron*. 2008; **59**:226–240.

Sarkisian MR, Siebzehnrubl D (2012). Abnormal levels of Gadd45alpha in developing neocortex impair neurite outgrowth. Plos One, 7, e44207

Sartor, J.J., Balthazart, J., Ball, G.F., 2005. Coordinated and dissociated effects of testosterone on singing behavior and song control nuclei in canaries (*Serinus canaria*). Horm. Behav. 47, 467-476.

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**, 2896-913

Scharff, C. and Nottebohm, F. (1991). A comparative study of the behavioral deficits following

Schlinger, B.A., 1997. Sex steroids and their actions on the birdsong system. J. Neurobiol. 33, 619-631.

Schlinger, B.A., Barske, J., Day, L., Fusani, L., Fuxjager, M.J., 2013. Hormones and the neuromuscular control of courtship in the golden-collared manakin (*Manacus vitellinus*). Front. Neuroendocrinol. 34, 143-156.

Schulz SB, Haesler S, Scharff C, Rochefort C (2010) Knockdown of FoxP2 alters spine density in Area X of the zebra finch. *Genes Brain Behav*. **9**, 732-40

Shevtsova, Z., Malik, J., Michel, U., Bahr, M., Kugler, S. (20040. Promoters and serotypes: targeting adeno-associated virus vectors for gene transfer in rat central nervous systems in vitro and in vivo. Experimental Physiology 90.1: 53-50

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**:16510-21

Shu, W. (2001). Characterization of a New Subfamily of Winged-helix/Forkhead (Fox) Genes That Are Expressed in the Lung and Act as Transcriptional Repressors. *Journal of Biological Chemistry*. **276**(29), 27488-27497.

Silverin, B., 1980. Effects of long acting testosterone treatment on free-living pied flycatchers, *Ficedula hypoleuca*, during the breeding period. Anim. Behav. 28, 906-912.

Stein, A.C., Uy, J.A.C., 2006. Plumage brightness predicts male mating success in the lekking golden-collared manakin, Manacus vitellinus. Behav. Ecol. 17, 41-47.

Stepanek L, Doupe AJ (2010) Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. J Neurophysiol 104:2474–2486.

Suthers, R.A., Goller, F., Wild, J.M., 2002. Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. Proc. Natl. Acad. Sci. USA 99, 5680-5685.

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

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**, 3152-63

Teramitsu I, Poopatanapong A, Torrisi S, White SA (2010) Striatal FoxP2 is actively regulated during songbird sensorimotor learning. *Plos One* **5**, e8548

Teramitsu I, White SA (2006) FoxP2 regulation during undirected singing in adult songbirds. *J. Neurosci*. **26**, 7390-4

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

Tomaszycki, M.L., Adkins-Regan, E., 2005. Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. Anim. Behav. 70, 785-794.

Troyer TW, Doupe AJ (2000a) An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables *J. Neurophysiol.* **84**, 1204-23

Troyer TW, Doupe AJ (2000b) An associational model of birdsong sensorimotor learning II. Temporal hierarchies and the learning of song sequence. *J. Neurophysiol.* **84**, 1224-39

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. USA* **95**, 12695-700

Veney, S.L., Wade, J., 2004. Steroid receptors in the adult zebra finch syrinx: a sex difference in androgen receptor mRNA, minimal expression of estrogen receptor alpha and aromatase. Gen. Comp. Endocrinol. 136, 192-199.

Voigt, C., Goymann, W., 2007. Sex-role reversal is reflected in the brain of African black coucals (*Centropus grillii*). Dev. Neurobiol. 67, 1560-1573.

Wang, B. (2003). Multiple Domains Define the Expression and Regulatory Properties of Foxp1

Warren, Wesley C., et al. (2001) The genome of a songbird. *Nature* **464**. 757-762.

Watkins, K. E., Dronkers, N. F. and Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain.* **125**(Pt 3), 452-64.

Whitney O. and Johnson F. (2005). Motor-induced transcription but sensory-regulated translation of ZENK in socially interactive songbirds. *J Neurobiol.* **65**(3):251-9.

Wild, J.M. and Farabaugh S.M. (1996). Organization of afferent and efferent projections of the nucleus basalis prosencephali in a passerine, *Taeniopygia guttata. J. Comp. Neurol*. **365**(2):306-28.

Wild, J.M., Goller, F., Suthers, R.A., 1998. Inspiratory muscle activity during bird song. J. Neurobiol. 36, 441-453.

Woolley, S. M. and Rubel, E. W. (1997). Bengalese finches Lonchura Striata domestica depend upon auditory feedback for the maintenance of adult song. *J Neurosci*. **17**(16).

Wyce, A., Bai, Y.C., Nagpal, S., Thompson, C.C., 2010. Research resource: The androgen receptor modulates expression of genes with critical roles in muscle development and function. Mol. Endocrinol. 24, 1665-1674.

Wymbs, N. F., Bastian, A. J., & Celnik, P. A. (2016). Motor Skills Are Strengthened through Reconsolidation. *Current Biology*. (Online only)

Vernes, S.C., Oliver, P.L., Spiteri, E., Lockstone, H.E., Puliyadi, R., Taylor, J.M., Ho, J., Mombereau, C., Brewer, A., Lowy, E., et al. (2011). Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. PLoS Genet. 7, e1002145

Yin, H. H., Mulcare, S. P., Hilário, M. R., Clouse, E., Holloway, T., Davis, M. I., ... & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nature neuroscience, 12(3), 333-341.