Note to the Author:
Please confirm the Continued style fixed for the Figure 1 caption.

FoxP2 and vocalization

Stephanie A. White
Department of Integrative Biology and Physiology
University of California, Los Angeles, CA, USA

From a linguist's point of view, the ability to vocalize new sounds may not seem to be a critical component of language. Yet when this ability is impaired, the social and emotional consequences for the affected individual can be severe, as evidenced by those suffering from developmental or injury-related speech disorders. How are we to understand this vocal learning trait, and where should it be placed within a framework for language evolution? Here, I argue that studying the supporting brain pathways that are affected in vocal learning disorders is a good place to start. Since such study is largely limited to noninvasive methods in humans, investigating other animals that possess this rare trait paves the way for a comparative analysis of the molecular, cellular, and synaptic bases of vocal production learning, including human speech. This kind of inquiry can highlight shared evolutionary pathways as well as key detours.

1. Introduction

Not only is language unique to humans, language is unique. No other behavior exhibits the same suite of seemingly conflicting features, including being innately predisposed yet highly dependent on social experience; requiring practice yet remaining unstereotyped; amenable to rapid-fire interchange yet infinitely expressive. When language is (artificially) deconstructed into separate subcomponents, some similarities to other behaviors emerge. In this chapter, I focus on vocal production learning. This subcomponent can be likened to the fine art of learning to draw or paint. Both skills arise through what scientists call "procedural learning," to distinguish them from "declarative learning." The latter can be accomplished through the verbal or written transfer of information. In contrast, no one can tell you how to ride a bike or how to produce a new word or painting. Rather, these acts depend on trial-and-error learning along the lines of "practice makes perfect." Both rely on sensory input to guide them and on feedback to perfect them. In the case of vocal learning, the main (but not sole) sensory domain is hearing, whereas in painting, vision is key. In both, heightened control of the muscles that participate in creating the new sound or image must be trained. Once such control is mastered, the creative possibilities abound for the skilled speaker or artist.

Regardless of whether or when vocal learning arose as a bona fide subcomponent in the evolution of human language, it can be argued, as Nottebohm has, that once this ability was in place, it enabled the development of "an open-ended system of sounds that can be used… for the further development of language" (Nottebohm & Liu 2010, p. 3). Thus, while not the most unique subcomponent of language, vocal learning in the hominid lineage could have supported and/or reinforced language evolution. Investigation of how the brain accomplishes this special sensorimotor skill, which is shared among only a handful of animal groups, has already revealed a surprising number of developmental, anatomical, and molecular commonalities.

2. Vocal production learning: What is it, who does it, and how do you know?

How can one prove that a given species, such as *Homo sapiens*, is innately capable of producing its vocal communication signals rather than having to learn them? What would be a definitive experiment? "Take some disenfranchised children off to an island and raise them without talking with them," you say? Though shocking, Akbar the Great (1542–1605) did just that (see Cohen, this volume). As the third Mughal emperor, he commanded that the infants be reared by mute nuns and then, at 12 years of age, be returned to court for analysis. There, the children failed to make any decipherable utterances, even though Akbar had astutely convened judges from many lands in order to detect any rare languages that the children might have produced. While today such experiments are legally and morally prohibited, they serve to illustrate one general approach for gathering such proof in nonhuman animals.

In modern times, a new group of animals has made the list of vocal production learners (Poole et al. 2005). In this case, careful acoustic analysis has overcome the limitations of studying a surprisingly low number of subjects, namely two. These two were Mlaika and Calimero, African elephants living in captivity, where they were noticed to produce atypical sounds. Scientists compared their vocalizations to those of other elephants, including Asian ones, and to other noises. Although Mlaika made some "normal" sounds that overlapped in length and frequency (perceived as pitch) with other African elephants' sounds, she also uttered a distinct set of longer and higher vocalizations that overlapped with the sounds made by trucks, recorded from the nearby highway. Calimero's vocalizations, on the other hand, were intermediate in length and frequency between those of his own species and those typical of Asian elephants, with whom he had been housed as a juvenile. Together, the stories of Akbar and the elephants illustrate the types of experiments that test for vocal production learning. Animals genetically endowed with this ability but that are raised in impoverished environments, where they are deprived of hearing their own species, fail

to develop normal vocalizations and/or imitate uncharacteristic but more abundant sounds which they are not hard-wired to produce.

A more severe test of the innateness of a species' vocalizations is the deprivation of all sounds that occurs with deafness. Deafness early in development is the most devastating because it delivers two blows to vocal learning: (1) it eliminates the imitative model by preventing the learner from hearing others of its species. Equally devastating, (2) the learner cannot hear his or her own vocalizations, preventing the auditory feedback necessary for vocal imitation and refinement. Loss of hearing in adulthood causes more subtle speech deficits that accumulate over time by preventing the speaker from continuously monitoring and updating his/her speech quality. Although we are not readily conscious of such monitoring, it can be experimentally revealed in adults with normal hearing by using headphones to deliver playback of their speech while they are speaking. If the speech is played back with a slight delay it can artificially induce the speaker to stutter (Lee 1950).

Due to practical considerations, not all animals have been rigorously tested for vocal production learning, but many nonhuman primates have. Our closest relative, the chimpanzee (*Pan troglodytes*), from whom we diverged some 6 million years ago, shares 95% of our DNA sequences (this number jumps to 99% for gene coding regions; more on this topic below), yet none of our vocal learning capacity (Pollard 2009). For example, a young female chimpanzee named Vicky was raised by her keepers in their home as if she were a human child. After six and a half years of training, she was only capable of uttering the distinguishable words *mama, papa,* and *cup* (Wallman 1992). This abysmal level of verbal output does not mean that chimpanzees are incapable of understanding language. Indeed, Vicky and additional chimpanzee subjects exhibit significant language comprehension (Terrace et al. 1979). And today, the bonobo (*Pan paniscus*) known as Kanzi demonstrates an impressive ability to both understand and "talk" with human caretakers when trained to point to pictograms in order to express himself, rather than to vocalize (see Savage-Rumbaugh, this volume). It is just that, based on their peripheral and central anatomy, chimpanzees and other nonhuman primates lack the physical capacity for developing the specialized control of the muscles necessary for noninnate vocal output. Along the vocal tract, these include the larynx, pharynx, tongue, teeth, and lips, as well as the muscles of respiration. All of these muscles are controlled by motoneurons in distinct regions of the brainstem. When the motoneurons fire, the muscles that they contact contract. What appears crucial is the next step back in the pathway leading to the motoneurons. In humans, but not in nonhuman primates, neurons in the motor cortex directly innervate laryngeal motoneurons. This neural connectivity or "wiring pattern" appears necessary for producing learned vocalizations, but is dispensable for innate vocal patterns, which depend upon a separate pathway (Jurgens 2009). Whether elephants possess the

crucial direct connection between motor cortex and motoneurons in their vocal control pathway is as yet unknown.

In addition to humans and elephants, the short list of animals demonstrated to be vocal production learners is currently limited to songbirds, parrots, and hummingbirds, which are in separate taxonomic orders (raising the hypothesis that the trait emerged independently three times in the avian lineage), and certain species of marine mammals and bats. Of these, the learned song produced by songbirds is the best characterized and exhibits significant parallels to human speech (Jarvis 2004). Shared features include the facts that both are learned through social interactions with conspecifics, both occur naturally and spontaneously within the organism's own species-characteristic behavior, and, as outlined above for humans, both depend upon auditory experience. As will be detailed below, learned birdsong also shares developmental, anatomic, and genetic components with speech.

With the goal of discovering the biological bases for vocal production learning and relating these to language evolution, an important advantage of certain songbird species is that they readily breed in the laboratory where they can be reared under controlled experimental conditions. Moreover, the brain pathways that support song learning and production are easily identified, especially in species such as the zebra finch in which only males learn to sing (i.e. their courtship songs, which are then listened to and selected for by female zebra finches). Congruent with the sexually dimorphic behavior in this species, the underlying neuroanatomical pathways are also sexually dimorphic (Nottebohm & Arnold 1976). Only males possess the full suite of interconnected brain regions that support song. These structures are dedicated to song learning and production, presenting excellent targets for the manipulation of brain circuits related to vocal learning without disrupting other cognitive processes. This is not the case in humans, nor in other vocal learning species studied thus far. Due to these unique features, songbirds such as the zebra finch provide an advantageous animal model to identify the molecular, cellular, and synaptic bases for vocal production learning.

2.1 Parallel vocal developmental programs

Similar to humans, songbirds learn their vocalizations best early in development. Learning involves two critical periods that can be distinguished by the source of the auditory input required for normal development. In the first critical period, termed "sensory acquisition," young songbirds listen to and memorize the song of an adult tutor. In zebra finches, sensory acquisition begins around the time of fledging (~20 days post-hatching) and ends by 65 days (Immelmann 1969), at which time a normally reared finch will become refractory to learning additional songs (White 2001). A second critical period known as "sensorimotor learning" occurs when young

birds begin to produce new sounds and to use auditory feedback of their own vocalizations to perfect a match to the memorized model. The onset of this process has been likened to human infant babbling (Doupe & Kuhl 1999). As sensorimotor learning progresses, the previously rambling and variable song becomes increasingly stereotyped such that by sexual maturation, which occurs at ~100 days in zebra finches, the song is sung relatively unchanged throughout adulthood.

The stereotyped nature of adult zebra finch song appears to contrast with the less limited capacity of human vocalizations. However, a broader comparison of vocal learning in the >4,000 species of songbirds to human speech reveals shared developmental constraints, as well as relative openness to experiential input throughout life, coupled with ongoing dependence on hearing. Specifically, the degree of vocal flexibility in mature songbirds varies with the species. Mockingbirds, for example, are capable of learning new songs throughout their lives. Even in zebra finches, mature song is not fixed but rather requires continuous auditory feedback in order to be maintained, as described above for human speech. The so-called "crystallized" song of zebra finches nonetheless deteriorates in birds deafened in adulthood (Brainard & Doupe 2000; Nordeen & Nordeen 1992). Also like speech, mature birdsong can be disrupted in normal hearing birds exposed to abnormal auditory feedback (Andalman & Fee 2009; Cynx & Von Rad 2001; Sober & Brainard 2009). On the human side, although there are clearly some "mockingbirds" among us, the ability to learn new languages without an accent is generally best accomplished prior to puberty (Doupe & Kuhl 1999).

2.2 Anatomical parallels

In 2004, the cell groups and fiber tracts of the avian brain were renamed in accordance with data that had accumulated prior to and since the publication of the stereotaxic atlas of the pigeon brain (Karten & Hodos 1967; Reiner et al. 2004). The new nomenclature corrects previous erroneous assumptions about the origin of avian neural tissue and the limitations of avian intelligence, and reinforces the similarity between avian and mammalian circuits. As a result, birds, including songbirds, are now acknowledged to possess a substantial amount of cortex, in addition to basal ganglia. The basal ganglia were previously thought to form the bulk of the avian telencephalon and to account for the overly instinctual behaviors of birds – another erroneous assumption. Along with a substantial cortex, certain avian species are now recognized to possess more sophisticated cognitive capacities than those exhibited by the domesticated, flightless, non-vocal-learning chicken, most familiar to humans. Even the microcircuitry within the primary avian auditory cortex has been found to comprise radial columnar arrays virtually identical to those of the mammalian auditory cortex (Wang et al. 2010).

Within the brains of songbirds, but not in non-vocal-learning birds, distinct subregions of the cortex, basal ganglia, and thalamus are dedicated to song learning

and production. Outside of these subregions, the cell types are similar to those found within, but the functions of the neurons are diverse and ill defined. This special feature whereby neurons dedicated to vocal production learning are grouped together within a given brain region, greatly facilitating their anatomical and functional identification, thus far appears limited to avian vocal learners. In songbirds, these brain regions and their interconnections are collectively referred to as the song circuit.

The song circuit consists of two component pathways: a vocal motor backbone, referred to as the posterior vocal pathway (in the back of the brain), and the anterior forebrain pathway (toward the front). The former is required for learned vocal production throughout the life of the bird and includes the nucleus known as the HVC (this name reflects a convention in which the acronym is currently used as its proper name), a subset of whose neurons project to the robust nucleus of the arcopallium (RA). RA projection neurons, in turn, synapse directly onto brainstem motoneurons of the tracheosyringeal nucleus (McCasland 1987; Nottebohm et al. 1976). Importantly, this neuroanatomical pathway comprises a direct projection from the cortex to the motoneurons controlling muscles used for vocalization, described above as a critical feature of vocal production learners. In this case, the cortical region RA directly contacts the motoneurons that control the syrinx, or song organ.

Like the posterior vocal pathway, the anterior forebrain pathway also begins with the HVC, where a separate subset of neurons innervates the basal ganglia nucleus known as area X. Area X projection neurons synapse in the dorsolateral medial thalamus, whose neurons then project to the lateral magnocellular nucleus of the anterior nidopallium (LMAN). LMAN projection neurons join the posterior and anterior pathways via their synapses in the RA (and they also project back to area X; Bottjer et al. 1989; Okuhata & Saito 1987; Scharff & Nottebohm 1991; Sturdy et al. 2003). The anterior forebrain pathway thus forms a loop between cortex, basal ganglia, and thalamic structures, and back to cortex, and resembles cortical-basal ganglia loops in humans that are important for the initiation of movements and procedural learning (Barnes et al. 2005; Bottjer & Arnold 1997; Graybiel et al. 1994).

Given that the posterior vocal pathway controls learned vocal output, what is the importance of the anterior forebrain pathway that feeds into it? The short answer to this question is "change." Beginning in 1984 (Andalman & Fee 2009; Bottjer et al. 1984; Brainard & Doupe 2000; [illegible] et al. 2005; Olveczky et al. 2005; Scharff & Nottebohm 1991; Williams & Mehta 1999) and continuing until the present, a set of elegant experiments has systematically demonstrated that the anterior forebrain pathway is required for any modifications to song, whether it be an improvement in vocal output or a deterioration. Thus, the posterior pathway can be viewed as the "command" module for learned vocal output (e.g. "sing this!"), while the anterior forebrain pathway can be seen as providing the signal for changing song, which is critical for the trial-and-error aspect of procedural learning, here in the vocal domain. As we will see below, the

anterior forebrain pathway remains important after a song is learned even in species that sing stereotyped songs such as the zebra finch.

The similarity between mammalian basal ganglia loops and song circuitry extends beyond anatomical connectivity to the identity of the cell types that make up each region, and to the neurochemicals that modulate their function. In terms of neuronal phenotypes, area X is now known to be composed of both striatal and pallidal neurons whose properties exhibit striking similarities to mammalian, including primate, basal ganglia neurons (Farries & Perkel 2002; Goldberg et al. 2010; Goldberg & Fee 2010; Reiner et al. 2004). With regard to neuromodulation, area X receives dense dopaminergic input (Bottjer 1993; Lewis et al. 1981), which modulates the excitability of medium spiny neurons via dopamine receptors (Casto & Ball 1994; Ding & Perkel 2002). Similar inputs to the mammalian striatum are critical for motor learning and reward (Balleine et al. 2009). Dopaminergic inputs to area X are differentially activated during singing, depending on the social context in which it takes place. When a male sings to a female zebra finch, his dopamine levels rise in area X and his song is more precise (Hara et al. 2007; Leblois et al. 2010; Yanagihara & Hessler 2006). When the male practices his song alone, dopamine levels are lower and songs, though still stereotyped, are more variable. Thus, social interactions modulate song circuit function by regulating dopamine release into area X, very likely during learning (Kojima & Doupe 2011), but also in maturity. These observations about the role of the basal ganglia in songbird vocal learning suggest that we should look for similar roles of the human basal ganglia in speech development, and conversely, to determine how dysfunction in this pathway impairs speech.

2.3 The KE family: A case study in disrupted vocal production learning

The first single mutation to be linked to a language disorder occurs in the gene encoding the transcription factor known as FOXP2 (Balter 2001; Fisher 2006; Lai et al. 2001). Transcription factors affect the expression of suites of other genes by binding to regulatory regions in the noncoding portion of their targets and either increasing or decreasing their transcription. The FOXP2 discovery arose from the study of a British family known as the KE family (Hurst et al. 1990), half of whom suffer from developmental dyspraxia, a deficit in the control of complex sequential movements of the orofacial muscles including those used in speaking. Peripheral control of these same muscles appears unimpaired, and innate behaviors such as suckling, chewing, and blinking are normal. These observations indicate that the problem lies within the brain rather than between motoneurons and their muscle targets – a proposition that has been confirmed by brain imaging studies.

Magnetic resonance imaging reveals that affected family members have altered amounts of gray matter relative to their unaffected counterparts in cortical and basal

ganglia regions (Belton et al. 2003; Watkins et al. 1999). These findings are consistent with the known role of other Forkhead-type transcription factors in driving embryogenesis of different organs during development. In this case, FOXP2 likely participates in the structural differentiation of brain regions. Following development, their altered structure contributes to their dysfunction. Accordingly, functional neuroimaging of the KE family reveals abnormal activation of these regions only in affected members during verbal fluency tasks (Liegeois et al. 2003). As can be imagined, the KE family has undergone extensive testing to determine the full range of their language deficits. Discussion of the complete syndrome is beyond the limits of this chapter and the interested reader is referred to Vargha-Khadem et al. (2005). It is important to acknowledge here that the phenotype is not limited to language, as affected KE family members have a significantly lower, albeit overlapping, verbal and performance IQ compared with unaffected members. In general, deficits are greater for language production than comprehension. Accordingly, assessment of core deficits, namely tasks in which affected family members' performance is poorer than and nonoverlapping with the performance of unaffected members, identified the accuracy and consistency of speech (Vargha-Khadem et al. 2005). Meanwhile, their ability to name objects is unimpaired. Thus, of the three components of language described by Hauser et al. (2002) (i.e. recursion, conceptual-intentional, and sensory-motor), sensory-motor control of speech is the most clearly affected.

In 2001, the genetic basis of the KE family disorder was shown to lie within the FOXP2 coding sequence (Lai et al. 2001). Afflicted KE family members share a point mutation on one allele for *FOXP2* that results in a substitution of amino acid 553 from an arginine to a histidine. This change occurs in the DNA binding domain of the protein, critical for its gene regulatory role. Indeed, x-ray crystallography-derived structural models of the protein show that residue 553 is intimately associated with the DNA during binding (Stroud et al. 2006). While extremely rare, individuals within other families have now been identified who exhibit strikingly similar symptoms to those described for the KE family. In these distinctive cases, disruption of the *FOXP2* gene has been consistently demonstrated (Macdermot et al. 2005; Zeesman et al. 2006). Taken together, this body of work firmly establishes that mutations restricted to the *FOXP2* gene alone can produce a profound and complex disorder of human language.

3. From gene to phenotype: How to connect them?

On the one hand, knowing how a specific genetic mutation produces a change in protein structure that results in altered brain morphology and a fully characterized language disorder would seem to form a startlingly complete picture of things. On the other, this set of observations reveals only the edges of a glimpse into the biological

basis, and thus the evolutionary origins, of language. To paint a fuller picture, the intervening molecular, cellular, and circuit effects of altered FOXP2 must be filled in. This requires carefully controlled physiological experiments using in vitro preparations and animal models (White et al. 2006). For starters, as a transcription factor, FOXP2 by itself is ineffectual and can only exert its function on brain tissues indirectly, through regulation of its target genes. Thus, we need to know what those genes are – a topic we will return to below – and how their altered levels impact language development. Given the significant parallels between songbirds and humans in vocal production learning and its underlying circuitry, which includes brain regions affected in the KE family phenotype, songbirds present a relevant animal model for exploring FoxP2. Thus, shortly after the discovery of the FOXP2 link to language, my colleagues and I examined *FoxP2* mRNA in zebra finch brains and compared the expression pattern in hatchlings with that in the human embryonic brain. We found strong expression in the basal ganglia and thalamus as well as in the cortex of both species, consistent with a role for this Forkhead transcription factor in forming these neural structures during embryonic development (Ferland et al. 2003; Haesler et al. 2004; Lai et al. 2003; Takahashi et al. 2003; Teramitsu et al. 2004). The similar expression pattern provided a "green light" to continue testing FoxP2 function in birds, with the goal of applying what we find to other vocal production learners, including humans.

3.1 Beyond brain structure: FoxP2 as a plasticity gate

In addition to its role in forming neural structures that are later used in vocal production learning, FoxP2 appears to have ongoing functions within these structures, including during learning and in the mature organism. In zebra finch song circuitry, FoxP2 expression persists into adulthood. Importantly, the adult expression is not simply a developmental vestige, but is under active regulation, as FoxP2 mRNA and protein rapidly decrease in area X of the striato-pallidum when adult birds sing (Miller et al. 2008; Teramitsu & White 2006). This "online" regulation, precisely in the striato-pallidal subregion dedicated to song and precisely when birds engage in singing, strongly implicates the molecule in the postorganizational function of this structure.

This idea is supported by the work of Haesler and colleagues, who developed a lentivirus bearing short interfering hairpin RNA (shRNA) constructs designed to knock down FoxP2 levels in the zebra finch brain. The virus was injected bilaterally into area X of 23-day-old male finches to test whether this would interfere with sensorimotor learning (Haesler et al. 2007). Control birds received injections of virus encoding an shRNA that did not target any zebra finch genes. All juveniles underwent normal tutoring, and multiple features of their song learning were assessed. Strikingly, at maturity, birds that had received the FoxP2 knock-down construct exhibited less

precise copying of their tutors' songs than did the controls. The decreased similarity included omissions, repetitions, and abnormally variable durations of syllables. This groundbreaking work represents the first case of genetic interference in songbirds resulting in documented changes to their song. Although, conceivably, altering the expression of any major transcription factor in cells that control song might result in song abnormalities, the fact that FOXP2 is vital for normal human language is consistent with the idea that the imprecise copying in FoxP2 knock-down birds reflects its specific contribution to vocal production learning.

Findings from my own laboratory complement Haesler et al.'s results. Briefly, behavioral states known to naturally lower FoxP2 in area X also give rise to more variable songs. To test this, we carefully analyzed songs of young birds that were behaviorally manipulated to achieve high vs. low levels of FoxP2 (Miller et al. 2010), using software designed to analyze zebra finch song (Tchernichovski et al. 2000). On one day, birds were allowed to sing for two hours in order to drive down levels of FoxP2 in area X, and then their subsequent songs were recorded (designated S–S, for singsing). On the next day, the same birds did not sing for two hours, which we know from our previous work leaves FoxP2 levels high in area X. The birds were then allowed to sing (designated NS-S, for nonsinging followed by singing) and those songs were recorded. S–S versus NS-S days were counterbalanced across birds to preclude any effect of order. The songs sung under each condition were then compared. The results firmly support the model, as follows: after song practice, a time coincident with low area X FoxP2 levels, vocal variability is high in both phonological (spectral features of syllables) and sequential (syllable order) domains. By contrast, when the same birds refrain from singing – coincident with high area X FoxP2 – their songs become more stable, which could reflect reinforcement of optimal motor patterns. Examples that illustrate this effect are shown in Figure 1A for spectral (phonological) features of song and Figure 1B and C for sequential features.

Together, these discoveries raise the hypothesis that FoxP2 plays a postorganizational role in vocal production learning by acting as a "plasticity gate." Behaviorally driven down-regulation of FoxP2 during song learning and adult song practice enables vocal variability. Conversely, high FoxP2 levels appear to promote organization of neural tissues during early development and may also reinforce optimal motor patterns during song learning and adult maintenance. More generally, cycles of practice and performance may improve a motor skill by altering expression levels of molecules that limit plasticity but promote reinforcement/stabilization. At first pass, this hypothesis is consistent with the lack of speech accuracy described for affected KE family members (Vargha-Khadem et al. 2005). It is important to note, however, that the KE phenotype arises from both organizational and postorganizational effects of the mutation, which is present from conception onward. Therefore, it is impossible to tease apart which of their deficits are due to abnormal development of brain structures and which are due to abnormal function of the gene throughout life.

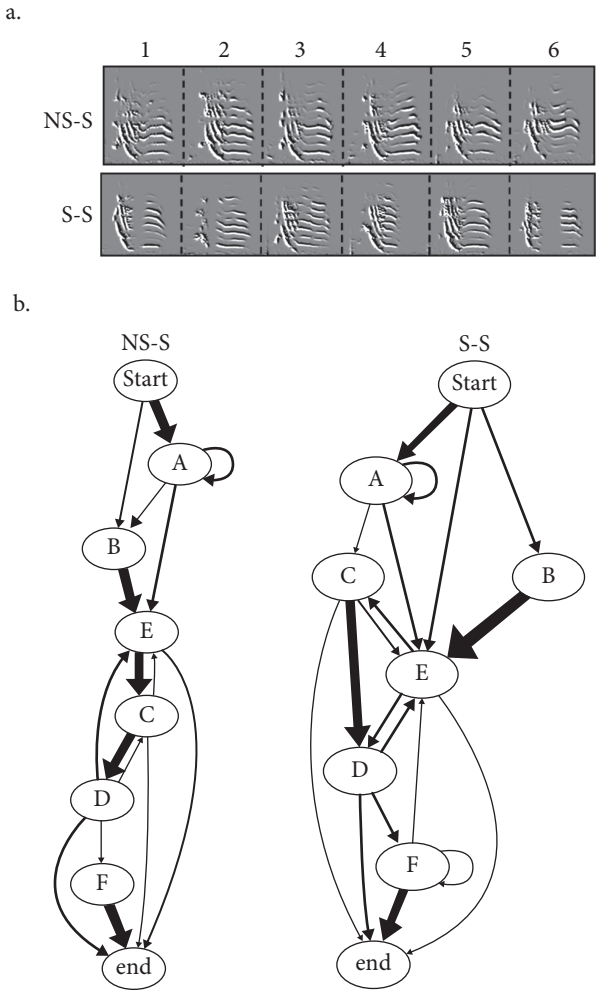

Figure 1. Phonological and sequential features are more variable under behavioral conditions known to decrease area X FoxP2 levels

A. Six renditions (1–6) of the same syllable from one bird are shown for two different days. On the first day (S–S), the bird was allowed to sing for two hours in the morning, which is known to decrease FoxP2 levels in area X. Subsequent songs contained these six syllables, which show much more variability than those shown below. This second set of syllables is from the second day (NS-S), when the bird did not sing for two hours, conditions under which FoxP2 levels remain high. Subsequently, the bird sang these more stable renditions.

B. Markov chain: an example of the possible transitions for one bird in the NS-S and S–S conditions. Letters denote syllables. Line thickness corresponds to probability; for example, in the NS-S condition, syllable E transitions 83% to syllable C (thick line), whereas a thinner line represents a 16% probability that E will end the motif; by contrast, in the S–S condition, syllable E transitions to syllable C 50% of the time, to syllable D 43% of the time, and ends the motif 7% of the time. In the NS-S condition, syllable F occurs infrequently compared to the S–S condition.

(*Continued*)

c.

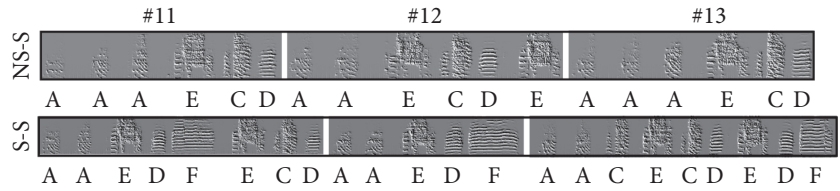

Figure 1. (Continued)
C. Exemplar of three consecutive motifs from the same bird in the NS-S and S–S conditions. Motifs occur at the same chronological order in the selected 20 motifs analyzed (#11, 12, and 13 out of 20). Individual syllables are identified by letter. In the NS-S condition, syllable A typically transitions to itself or to syllable E, and syllable C transitions most frequently to syllable D. By contrast, in the S–S condition, A also transitions to C (#13) and C to E (#13) as well. In the S–S condition, syllable F is observed (#11, 12, 13) and follows syllable D while in NS-S, syllable D transitions to E (#12) or ends the motif (#11, 13). Adapted from (Miller et al. 2010).

3.2 Other genes: FoxP family members

The plasticity gate hypothesis for FoxP2 arose largely through songbird research and remains untested in humans. Indeed, only if and when we are able to perform *FOXP2* gene therapy in humans would such a test be possible, due, in part, to the issue raised above about distinguishing organizational from postorganizational effects. A separate hypothesis arising from songbird research, however, has now been confirmed for human language. Collaborative work between my laboratory and that of Dr. Daniel Geschwind (UCLA) revealed that *FoxP1* and *FoxP2* share remarkably similar expression patterns in human and zebra finch brains (Teramitsu et al. 2004). Our observation that *FoxP1* is expressed in a sexually dimorphic pattern within zebra finch song circuitry led us to hypothesize that, like FOXP2, FOXP1 plays a role in vocal production learning and could underlie language-related disorders. Remarkably, this prediction has been borne out through the discovery of multiple human cases in which *FOXP1* mutations are associated with language deficits, accompanied by more global changes in cognitive abilities (Carr et al. 2010; Hamdan et al. 2010; Horn et al. 2010; Pariani et al. 2009). In several of these cases, the only gene shown to be disrupted is *FOXP1*, pinpointing it as an additional molecule critical for normal language development.

3.3 Genes downstream of FOXP2

Since FOXP2 is a transcription factor, its role in speech and language must be mediated by regulation of its target genes. Thus, we and others have hypothesized that *FOXP2* is not "*the* gene" for language, but rather represents an entry point into a network of molecules important for language (reviewed in Fisher & Marcus 2006;

Hilliard & White 2009). Finding downstream targets and identifying their function promises to elucidate the neuromolecular basis of language and disorders in which language is affected, such as specific language impairment (SLI) and autism spectrum disorder (ASD). Several exciting approaches have been taken to identify FOXP2 gene targets, focusing on those in humans. Two studies utilized a technique known as ChIP-chip – for chromatin immunoprecipitation followed by arraying on a microchip – that assures that identified genes are directly regulated by FOXP2. In both, an antibody against FOXP2 protein was used to specifically detect and isolate FOXP2 while doing its job of regulating transcription, that is, while FOXP2 was bound to the DNA regulatory sequences in the promoters of its target genes. These targets were then identified using promoter microarrays. In one study, human fetal lung, inferior frontal cortex, and basal ganglia tissues were used to identify target genes (Spiteri et al. 2007), with eight co-occurring in the two brain areas, but not in the lung. The different suites of genes regulated by FOXP2 depending upon the tissue help to explain the brain-specific functions of FOXP2. In the other study, human neuronal-like cell lines were similarly tested and revealed 119 targets (Vernes et al. 2007), with significant overlap with those identified in the former work. These studies do not represent a complete list of FOXP2 neural targets – not all neuronal cell types nor all known promoters were available. Such limitations will undoubtedly decrease with technological advances, promising a more complete picture of human FOXP2 targets.

Since humans are uniquely capable of language, which FOXP2 targets are uniquely human becomes of interest. The above studies used human tissues but did not show whether these same targets would also be regulated by FOXP2, for example, in our closest relative, the chimpanzee. To address this question, two additional studies have identified genes whose expression is altered specifically by the protein form of FOXP2 that exists in humans. In one, human neural progenitor cells were transduced to produce either the chimpanzee or the human FOXP2 (Konopka et al. 2009) and subsequent changes in gene expression were compared. The authors found 61 genes that were significantly upregulated and 55 genes downregulated in cells transduced with the human FOXP2 compared to those transduced with the chimp form.

Neither chimps nor zebra finches are easily amenable to transgenic approaches for altering gene expression, whereas mice are. Thus, a separate study introduced the human *FOXP2* into the endogenous form found in mice and examined the resultant changes in neuronal gene expression (Enard et al. 2009). The authors identified 34 genes whose expression differed specifically within the striatal region of the mouse basal ganglia. Medium spiny neurons are the main cell type in this area, and their dendrites – the neuronal processes upon which they receive synapses – were longer in the mice expressing the human form of FOXP2, suggesting the potential for enhanced neuronal "cross talk." In line with this, a form of synaptic plasticity thought to underlie certain forms of motor skill learning in mice was enhanced in this region. This

finding is remarkable because it directly complements a prior study in which, rather than inserting the normal human form of *FOXP2* into mice, the mice were mutagenized such that they possessed the KE family form of *FOXP2* (Groszer et al. 2008; Teramitsu & White 2008). In contrast to enhanced striatal plasticity, these mice were deficient in the very same form of synaptic change. Additionally, they exhibited deficits on the accelerating rotarod, a form of motor skill learning thought to be supported by this very type of synaptic plasticity.

What about their vocalizations? While the long answer may eventually be forthcoming, the short answer is that the type of vocalization tested thus far in all of the mice described above is an unlearned one, namely the ultrasonic cries of mouse pups when isolated from their mothers. Since mice are not capable of hearing until their second postnatal week (Ehret 1976), any alteration in these isolation calls reflects a change to an innate vocalization. As mice mature, they produce additional ultrasonic vocalizations that have even been likened to birdsong (Holy & Guo 2005). To what extent these "mouse songs" require learning is an exciting new area of intense investigation, and the experiments are following the general design discussed above for testing vocal learning in any species. At the time of writing, a first report has been published in which young mice of one strain were exposed to mature mice of a separate one, reminiscent of the developmental experience of the young African elephant who was housed with Asian elephants. In this study, normal mouse pups from one strain did not learn the songs of their foster-parents (Kikusui et al. 2011). Thus, there is as yet no murine equivalent of Calimero. As more of the mice with altered *FOXP2* genotypes reach maturity and become available for testing, new findings about their vocal output will be forthcoming. Likewise, follow-up experiments are required for the genes outlined above whose expression is altered depending on the FOXP2 isoform. Such work has begun for one of these gene targets, known as contactin-associated protein like-2 (*CNTNAP2*), and has already yielded important information about how FOXP2 connects to language uniquely in humans.

3.4 Key detour? A FOXP2 target is linked to specific language impairment and autism spectrum disorder

While FOXP2 has been linked to language in multiple cases, evidence for its role in SLI or in other developmental disorders in which language is affected, such as ASD, has been lacking (cf. Li et al. 2005; but see Peter et al. 2011). Yet, as described above, FOXP2 regulates many genes, the exact identities of which depend upon the tissue. Excitingly, a series of studies have now shown that *CNTNAP2* is implicated in developmental disorders of language, and is a direct target of FOXP2 repression in humans (Vernes et al. 2008; Whitehouse et al. 2011). This discovery arose, in part, through a modified version of the ChIP-chip technique described above. In this variant, known as ChIP-seq,

the gene targets to which FOXP2 binds are directly sequenced, rather than arrayed. One of the sequenced genes encodes CNTNAP2 (also referred to as AUTS15, CASPR2, CDFE, DKFZp781D1846, and NRXN4), a member of the neurexin superfamily of cell-adhesion molecules that, together with their binding partners, the neuroligins, have been implicated in ASD (Poliak et al. 1999). Several independent lines of evidence have converged to identify CNTNAP2 as an important modulator of diverse clinical phenotypes involving impaired language performance. *CNTNAP2* was originally linked to SLI and ASD in an Old Order Amish population that harbored an abnormal *CNTNAP2* allele. A single nucleotide deletion resulted in a frame shift and premature stop codon, producing a truncated protein that lacks its transmembrane and intracellular domains. This truncation presumably disrupts the protein's normal function. Members of the population homozygous for the mutation exhibit cortical dysplasia-focal epilepsy and symptoms of ASD and SLI (Strauss et al. 2006).

Though the truncated *CNTNAP2* described above results in a severe phenotype, less dramatic polymorphisms in the general public have been linked to ASD and SLI, using instruments of autism diagnosis, age of first word, language expression and comprehension, ability to repeat nonsense words, and reading ability (Alarcon et al. 2008; Arking et al. 2008; Newbury et al. 2010; Vernes et al. 2008). Most recently, common *CNTNAP2* variants have been shown to influence early language development even among the general population (Whitehouse et al. 2011). Since *CNTNAP2* is expressed in neurons and is associated with cognitive disorders, several groups have looked for anatomical anomalies in the brain associated with *CNTNAP2* polymorphisms. Structural MRI of affected members of the Old Order Amish revealed abnormalities in the temporal lobe and striatum, areas critical for speech and language (Strauss et al. 2006). In a separate study outside of that population, people homozygous for a risk allele of *CNTNAP2* had less white and gray matter than those bearing nonrisk alleles in several brain regions associated with ASD (Tan et al. 2010). Functional MRI has revealed altered frontal lobe connectivity associated with *CNTNAP2* risk alleles (Scott-Van Zeeland et al. 2010). Curiously, the neuroanatomical changes in humans are not mimicked in *Cntnap2* knockout mice, which exhibit typical brain morphology. In fact, pending further characterization, knockout mice display surprisingly normal anatomical, neurophysiological, and behavioral phenotypes (Poliak et al. 2003).

The difference between human and rodent Cntnap2 phenotypes may be a function of where Cntnap2 is expressed in the brain of each species. In human fetal brains, prior to myelination, *CNTNAP2* is highly enriched in the frontal cortex and otherwise restricted to the striatum and dorsal thalamus, defining key circuitry important to aspects of higher cognition, including the implicit learning essential for language development (Abrahams et al. 2007). This stands in sharp contrast to the broad transcript distribution observed in the developing brains of both rats and mice. While the jury is still out on the degree to which rodent vocalizations are learned, it is clear

that birdsong is. Intriguingly, we found that the *Cntnap2* expression pattern in zebra finch brains is more similar to the human pattern, which is not exhibited by rodents (Panaitof et al. 2010). *Cntnap2* mRNA is differentially expressed in several parts of the song circuit, including enrichment in the RA and LMAN cortical regions, relative to the surrounding areas that nonetheless contain similar cell types. In the basal ganglia song nucleus area X, there is a marked reduction in *Cntnap2* mRNA, relative to its surrounding region. Taken together, these findings support the hypothesis that CNTNAP2 plays an early developmental role in the patterning and functional specialization of circuits related to higher cognition and learned vocalizations, potentially in multiple species.

3.5 Looking into the dark matter

The demonstration of an interaction between FOXP2 and CNTNAP2 in humans (Vernes et al. 2008) begins to define a neuromolecular network related to language and could underlie learned vocal communication in other species. As detailed above, within a given species, FOXP2 interacts with different suites of genes in different tissues, which helps to explain how the effects of its mutation are largely restricted to the brain. Further, FoxP2 likely interacts with different suites of genes in the same tissue of different animals, as evidenced by the work comparing differential gene regulation due to human versus chimpanzee forms of FoxP2. This phenomenon can provide hints to the biological origin of the language phenotype. For example, in the case of CNTNAP2, the genetic region to which FOXP2 binds and represses transcription is located in an intron. Introns, like promoters, are noncoding regions of the DNA and can contain regulatory sequences that indicate where, when, and how much of the gene will be expressed. Two consensus sites for FOXP2 binding were found within this intronic region, namely two instances of the DNA sequence CAAATT (Vernes et al. 2008). If these sites are lacking in the *Cntnap2* of rodents, then Foxp2 may be unable to repress rodent Cntnap2 expression. This possibility fits with the restricted pattern of *CNTNAP2* expression observed in human fetal brains, which is inverse to *FOXP2* expression therein, and contrasts with the diffuse pattern of *Cntnap2* expression observed in mice and rats (Abrahams et al. 2007). Specific repression of *CNTNAP2* during human brain development could thereby enhance the functional connectivity of brain areas critical for language development (Scott-Van Zeeland et al. 2010).

Whether or not the FOXP2-CNTNAP2 connection in humans represents a key evolutionary detour, it serves to illustrate a broader point, namely that regulatory sequences in the noncoding regions of genes, in the so-called "dark matter," are important players in evolution. The human genome project has revealed that 44% of our DNA is composed of mobile transposable elements (Lander et al. 2001). It has

been speculated that short regulatory regions such as transcription factor binding sites can be present in these mobile elements and thereby produce species-specific gene expression patterns (Britten & Kohne 1968). Strong support for this scenario has recently been provided through study of the gene encoding human cathelicidin antimicrobial peptide (CAMP; Gombart et al. 2009). Specifically, the binding site for the vitamin D receptor was shown to be present in the promoter for this gene in the primate lineage, including humans, but not conserved in nonprimate mammals. Insertion of this site was mediated by a primate-specific Alu family of mobile, middle repetitive short-interspersed elements. As a result of this added regulatory region, vitamin D is able to potentiate the innate immune response in human and nonhuman primates but does not do so in other mammals.

Much attention has been paid to interspecies variation in the coding sequences of FoxP2, with the important finding that, among primates, two amino acids are unique to humans (Enard et al. 2002). Again, this discovery enabled the identification of certain genes that are uniquely regulated by the human, as opposed to the chimpanzee or mouse, form of FoxP2 (Konopka et al. 2009). In addition to coding sequences, however, alterations in the noncoding "dark matter" can give rise to important species-specific changes. Such changes are not limited to what has been uniquely added in humans, as just described for the CAMP gene (Gombart et al. 2009) and speculated on for CNTNAP2, but also to what is uniquely lacking. A recent report now demonstrates that, relative to the chimpanzee, ~500 genes have undergone human-specific deletions which are largely restricted to their noncoding regions (McLean et al. 2011). Perhaps not surprisingly, both *FOXP2* and *CNTNAP2* are on this list. To validate the potential impact of these changes, the authors demonstrated that one of the deletions removes a regulatory region of a growth arrest gene and is correlated with the expansion of specific brain regions in humans. In this way, deletions or insertions of even short pieces of DNA that happen to contain consensus sites for transcription factor binding can shape human evolutionary divergence.

3.6 Follow-through: Prioritizing genes

Given the startling amount of genomic complexity, how can we prioritize genes and gene interactions for investigation into language origins? Clearly, multiple approaches are needed, as in the case of CNTNAP2, where, despite being one of many FOXP2 targets uncovered in the ChIP-seq experiment, independent lines of converging evidence for its association with language brought it to the forefront (Vernes et al. 2008). To go beyond "one gene at a time," statistical techniques for probing correlations in gene expression are being generated and used to highlight gene interactions that are unique to brain regions that support specialized human cognitive capacities (e.g. Oldham et al. 2006). One of these techniques, known as weighted gene coexpression

network analysis (WGCNA) is at the forefront of modern tools required to analyze high dimensional data sets while avoiding the pitfalls of multiple hypothesis testing (Zhang & Horvath 2005). The approach highlights clusters of genes whose expression levels change in concert, and groups them into modules, with genes at the center of the modules being the most highly correlated, or connected – so-called "hub" genes. This methodology has an outstanding track record in predicting novel genes within highlighted pathways, such as previously unknown molecular targets in cancer (Horvath et al. 2006). It has even been fruitfully applied to clustering voxels (WVCNA), rather than genes, in fMRI data (Mumford et al. 2010).

One study has applied WGCNA to gene expression data from human fetal brains (Johnson et al. 2009). The analysis highlighted 11 hub genes as being critical for human brain development. Excitingly, four of these exact molecules have recently been shown to have undergone the human-specific loss of their regulatory DNA (McLean et al. 2011), relative to chimpanzees, while five are in the same family as genes that bear the human-specific deletions. This is another instance of the highly predictive value of WGCNA. I and my colleagues have applied this approach, for the first time, to a procedurally learned behavior, by examining suites of genes that are coregulated in songbirds during singing. We used the same paradigm described above in which birds alter their own area X FoxP2 levels as a function of how much they sing (Hilliard et al. 2012) in order to highlight genes that are coregulated with and functionally interact with FoxP2. Comparison of these to the known targets in humans reveals shared evolutionary drivers of vocal production learning, as well as molecular interactions unique to humans. The latter represent high priorities for further investigation as to their role in language origins.

4. Summary

This chapter has focused on one subcomponent of language, namely vocal production learning. We have argued that probing the neural circuitry that gives rise to this behavior, as well as what happens in cases where it malfunctions, can highlight the relevant biology upon which evolution has acted. Much progress has been made in understanding how the brain accomplishes this sensorimotor feat in the vocal domain by using a songbird animal model that is "expert" in this capacity. While birdsong and speech evolved independently, the brain appears to have found similar biological solutions to the challenge of learning to communicate vocally. Other animals offer distinct insights for other language subcomponents (cf. Zuberbühler, this volume). In all of these domains, analysis of gene interactions, largely mediated through the noncoding regions of the genome, provide even more biological fodder for evolutionary change.

Acknowledgments

The author's research is supported by a grant from the National Institutes of Health (HD065271; MH070712) and the Department of Defense (AR093327).

References

Abrahams, B.S., Tentler, D., Perederiy, J.V., Oldham, M.C., Coppola, G., & Geschwind, D.H. (2007). Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 17849–17854.

Alarcon, M., Abrahams, B.S., Stone, J.L., Duvall, J.A., Perederiy, J.V., Bomar, J.M., et al. (2008). Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *American Journal of Human Genetics, 82,* 150–159.

Andalman, A.S., & Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 12518–12523.

Arking, D.E., Cutler, D.J., Brune, C.W., Teslovich, T.M., West, K., Ikeda, M., et al. (2008). A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *American Journal of Human Genetics, 82,* 160–164.

Balleine, B.W., Liljeholm, M., & Ostlund, S.B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behavioural Brain Research, 199,* 43–52.

Balter, M. (2001). Genetics. First gene linked to speech identified. *Science, 294,* 32.

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., & Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature, 437,* 1158–1161.

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,* 194–200.

Bottjer, S.W. (1993). The distribution of tyrosine hydroxylase immunoreactivity in the brains of male and female zebra finches. *Journal of Neurobiology, 24,* 51–69.

Bottjer, S.W., & Arnold, A.P. (1997). Developmental plasticity in neural circuits for a learned behavior. *Annual Review of Neuroscience, 20,* 459–481.

Bottjer, S.W., Halsema, K.A., Brown, S.A., & Miesner, E.A. (1989). Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *Journal of Comparative Neurology, 279,* 312–326.

Bottjer, S.W., Miesner, E.A., & Arnold, A.P. (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science, 224,* 901–903.

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

Brainard, M.S., & Doupe, A.J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature, 404,* 762–766.

Britten, R.J., & Kohne, D.E. (1968). Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science, 161,* 529–540.

Carr, C.W., Moreno-De-Luca, D., Parker, C., Zimmerman, H.H., Ledbetter, N., Martin, C.L., et al. (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,* 1216–1220.

Casto, J.M., & Ball, G.F. (1994). Characterization and localization of D1 dopamine receptors in the sexually dimorphic vocal control nucleus, area X, and the basal ganglia of European starlings. *Journal of Neurobiology, 25,* 767–780.

Cynx, J., & Von Rad, U. (2001). Immediate and transitory effects of delayed auditory feedback on bird song production. *Animal Behaviour, 62,* 305–312.

Ding, L., & Perkel, D.J. (2002). Dopamine modulates excitability of spiny neurons in the avian basal ganglia. *Journal of Neuroscience, 22,* 5210–5218.

Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22,* 567–631.

Ehret, G. (1976). Development of absolute auditory thresholds in the house mouse (*Mus musculus*). *Journal of the American Audiology Society, 1,* 179–184.

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S.M., Blass, T., Somel, M., et al. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell, 137,* 961–971.

Enard, W., M. Przeworski, Fisher, S.E., Lai, C.S., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature, 418,* 869–872.

Farries, M.A., & Perkel, D.J. (2002). A telencephalic nucleus essential for song learning contains neurons with physiological characteristics of both striatum and globus pallidus. *Journal of Neuroscience, 22,* 3776–3787.

Ferland, R.J., Cherry, T.J., Preware, P.O., Morrisey, E.E., & Walsh, C.A. (2003). Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *Journal of Comparative Neurology, 460,* 266–279.

Fisher, S.E. (2006). Tangled webs: Tracing the connections between genes and cognition. *Cognition, 101,* 270–297.

Fisher, S.E., & Marcus, G.F. (2006). The eloquent ape: Genes, brains and the evolution of language. *Nature Reviews Genetics, 7,* 9–20.

Goldberg, J.H., Adler, A., Bergman, H., & Fee, M.S. (2010). Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: Comparison to the primate internal and external pallidal segments. *Journal of Neuroscience, 30,* 7088–7098.

Goldberg, J.H., & Fee, M.S. (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. *Journal of Neurophysiology, 103,* 2002–2014.

Gombart, A.F., Saito, T., & Koeffler, H.P. (2009). Exaptation of an ancient Alu short interspersed element provides a highly conserved vitamin D-mediated innate immune response in humans and primates. *BMC Genomics, 10,* 321.

Graybiel, A.M., Aosaki, T., Flaherty, A.W., & Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science, 265,* 1826–1830.

Groszer, M., Keays, D.A., Deacon, R.M., de Bono, J.P., Prasad-Mulcare, S., Gaub, S., et al. (2008). Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Current Biology, 18,* 354–362.

Haesler, S., C. Rochefort, 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 Biology, 5,* e321.

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E.E., Lints, T., Jarvis, E.D., et al. (2004). FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience, 24,* 3164–3175.

Hamdan, F.F., Daoud, H., Rochefort, D., Piton, A., Gauthier, J., Langlois, M., et al. (2010). De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *American Journal of Human Genetics, 87,* 671–678.

Hara, E., Kubikova, L., Hessler, N.A., & Jarvis, E.D. (2007). Role of the midbrain dopaminergic system in modulation of vocal brain activation by social context. *European Journal of Neuroscience, 25,* 3406–3416.

Hauser, M.D., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298,* 1569–1579.

Hilliard, A.T., Miller, J.E., Fraley E. R., Horvath, S., & White, S.A. (2012). Molecular microcircuitry underlies the functional specification of a basal ganglia circuit dedicated to vocal learning. *Neuron*, *73*, 537–552.

Hilliard, A.T., & White, S.A. (2009). Possible precursors of syntactic components in other species. In D. Bickerton & E. Szathmáry (Eds.), *Biological foundations and origin of syntax* (pp. 161–183). Cambridge, MA: The MIT Press.

Holy, T.E., & Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biology, 3,* e386.

Horn, D., Kapeller, J., Rivera-Brugués, N., Moog, U., Lorenz-Depiereux, B., Eck, S., et al. (2010). Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits. *Human Mutation, 31,* E1851–1860.

Horvath, S., Zhang, B., Carlson, M., Lu, K.V., Felciano, R.M., Laurance, M.F., et al. (2006). Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 17402–17407.

Hurst, J.A., Baraitser, M., Auger, E., Graham, F., & Norell, S. (1990). An extended family with a dominantly inherited speech disorder. *Developmental Medicine and Child Neurology, 32,* 352–355.

Immelmann, K. (1969). Song development in zebra finch and other Estrildid finches. In R.A. Hinde (Ed.), *Bird vocalisations* (pp. 61–74). London: Cambridge University Press.

Jarvis, E.D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016,* 749–777.

Johnson, M.B., Kawasawa, Y.I., Mason, C.E., Krsnik, Z., Coppola, G., Bogdanovic, D., et al. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron, 62,* 494–509.

Jurgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice, 23,* 1–10.

Kao, M.H., Doupe, A.J., & Brainard, M.S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature, 433,* 638–643.

Karten, H.J., & Hodos, W. (1967). *A stereotaxic atlas of the brain of the pigeon* (Columbia livia). Baltimore, MD: Johns Hopkins Press.

Kikusui, T., Nakanishi, K., Nakagawa, R., Nagasawa, M., Mogi, K., & Okanoya, K. (2011). Cross fostering experiments suggest that mice songs are innate. *PLoS ONE, 6,* e17721.

Kojima, S., & Doupe, A.J. (2011). Social performance reveals unexpected vocal competency in young songbirds. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 1687–1692.

Konopka, G., Bomar, J.M., Winden, K., Coppola, G., Jonsson, Z.O., Gao, F., et al. (2009). Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature, 462,* 213–217.

Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., & Monaco, A.P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413,* 519–523.

Lai, C.S., Gerrelli, D., Monaco, A.P., Fisher, S.E., & Copp, A.J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain, 126,* 2455–2462.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature, 409,* 860–921.

Leblois, A., Wendel, B.J., & Perkel, D.J. (2010). Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *Journal of Neuroscience, 30,* 5730–5743.

Lee, B.S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America, 22,* 824–826.

Lewis, J.W., Ryan, S.M., Arnold, A.P., & Butcher, L.L. (1981). Evidence for a catecholaminergic projection to area X in the zebra finch. *Journal of Comparative Neurology, 196,* 347–354.

Li, H., Yamagata, T., Mori, M., & Momoi, M.Y. (2005). Absence of causative mutations and presence of autism-related allele in FOXP2 in Japanese autistic patients. *Brain and Development, 27,* 207–210.

Liegeois, 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,* 1230–1237.

Macdermot, K.D., Bonora, E., Sykes, M., Coupe, A.M., Lai, C.S., Vernes, S.C., et al. (2005). Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *American Journal of Human Genetics, 76,* 1074–1080.

McCasland, J.S. (1987). Neuronal control of bird song production. *Journal of Neuroscience, 7,* 23–39.

McLean, C.Y., Reno, P.L., Pollen, A.A., Bassan, A.I., Capellini, T.D., Guenther, C., et al. (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature, 471,* 216–219.

Miller, J.E., Hilliard, A.T., & White, S.A. (2010). Song practice promotes acute vocal variability at a key stage of sensorimotor learning. *PLoS One, 5,* e8592.

Miller, J.E., Spiteri, E., Condro, M.C., Dosumu-Johnson, M.T., Geschwind, D.H., & White, S.A. (2008). Birdsong decreases protein levels of FoxP2, a molecule required for human speech. *Journal of Neurophysiology, 100,* 2015–2025.

Mumford, J.A., Horvath, S., Oldham, M.C., Langfelder, P., Geschwind, D.H., & Poldrack, R.A. (2010). Detecting network modules in fMRI time series: A weighted network analysis approach. *NeuroImage, 52,* 1465–1476.

Newbury, D.F., Paracchini, S., Scerri, T.S., Winchester, L., Addis, L., Richardson, A.J., et al. (2010). Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behavior Genetics, 41,* 90–104.

Nordeen, K.W., & Nordeen, E.J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and Neural Biology, 57,* 58–66.

Nottebohm, F., & Arnold, A.P. (1976). Sexual dimorphism in vocal control areas of the songbird brain. *Science, 194,* 211–213.

Nottebohm, F., & Liu, W.C. (2010). The origins of vocal learning: New sounds, new circuits, new cells. *Brain and Language, 115,* 3–17.

Nottebohm, F., Stokes, T.M., & Leonard, C.M. (1976). Central control of song in the canary, *Serinus canarius*. *Journal of Comparative Neurology, 165,* 457–486.

Okuhata, S., & Saito, N. (1987). Synaptic connections of thalamo-cerebral vocal nuclei of the canary. *Brain Research Bulletin, 18,* 35–44.

Oldham, M.C., Horvath, S., & Geschwind, D.H. (2006). Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 17973–17978.

Olveczky, B.P., Andalman, A.S., & Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology, 3,* e153.

Panaitof, S.C., Abrahams, B.S., Dong, H., Geschwind, D.H., & White, S.A. (2010). Language-related Cntnap2 gene is differentially expressed in sexually dimorphic song nuclei essential for vocal learning in songbirds. *Journal of Comparative Neurology, 518,* 1995–2018.

Pariani, M.J., Spencer, A., Graham, J.M., Jr., & 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,* 123–127.

Peter, B., Raskind, W.H., Matsushita, M., Lisowski, M., Vu, T., Berninger, V.W., et al. (2011). Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *Journal of Neurodevelopmental Disorders, 3,* 39–49.

Poliak, S., Gollan, L., Martinez, R., Custer, A., Einheber, S., Salzer, J.L., et al. (1999). Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K+ channels. *Neuron, 24,* 1037–1047.

Poliak, S., Salomon, D., Elhanany, H., Sabanay, H., Kiernan, B., Pevny, L., et al. (2003). Juxtaparanodal clustering of Shaker-like K+ channels in myelinated axons depends on Caspr2 and TAG-1. *Journal of Cell Biology, 162,* 1149–1160.

Pollard, K.S. (2009). What makes us human? *Scientific American, 300,* 44–49.

Poole, J.H., Tyack, P.L., Stoeger-Horwath, A.S., & Watwood, S. (2005). Animal behaviour: Elephants are capable of vocal learning. *Nature, 434,* 455–456.

Reiner, A., Perkel, D.J., Bruce, L.L., Butler, A.B., Csillag, A., Kuenzel, W., et al. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. *Journal of Comparative Neurology, 473,* 377–414.

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. *Journal of Neuroscience, 11,* 2896–2913.

Scott-Van Zeeland, A.A., Abrahams, B.S., Alvarez-Retuerto, A.I., Sonnenblick, L.I., Rudie, J.D., Ghahremani, D., et al. (2010). Altered functional connectivity in frontal lobe circuits is associated with variation in the autism risk gene CNTNAP2. *Science Translational Medicine, 2,* 56ra80.

Sober, S.J., & Brainard, M.S. (2009). Adult birdsong is actively maintained by error correction. *Nature Neuroscience, 12,* 927–931.

Spiteri, E., Konopka, G., Coppola, G., Bomar, J., Oldham, M, Ou, J., et al. (2007). Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *American Journal of Human Genetics, 81,* 1144–1157.

Strauss, K.A., Puffenberger, E.G., Huentelman, M.J., Gottlieb, S., Dobrin, S.E., Parod, J.M., et al. (2006). Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *New England Journal of Medicine, 354,* 1370–1377.

Stroud, J.C., Wu, Y., Bates, D.L., Han, A., Nowick, K., Paabo, S., et al. (2006). Structure of the forkhead domain of FOXP2 bound to DNA. *Structure, 14,* 159–166.

Sturdy, C.B., Wild, J.M., & Mooney, R. (2003). Respiratory and telencephalic modulation of vocal motor neurons in the zebra finch. *Journal of Neuroscience, 23,* 1072–1086.

Takahashi, K., Liu, F.C., Hirokawa, K., & Takahashi, H. (2003). Expression of Foxp2, a gene involved in speech and language, in the developing and adult striatum. *Journal of Neuroscience Research, 73,* 61–72.

Tan, G.C., Doke, T.F., Ashburner, J., Wood, N.W., & Frackowiak, R.S. (2010). Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of CNTNAP2. *NeuroImage, 53,* 1030–1042.

Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B., & Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour, 59,* 1167–1176.

Teramitsu, I., Kudo, L.C., London, S.E., Geschwind, D. H, & White, S.A. (2004). Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *Journal of Neuroscience, 24,* 3152–3163.

Teramitsu, I., & White, S.A. (2006). FoxP2 regulation during undirected singing in adult songbirds. *Journal of Neuroscience, 26,* 7390–7394.

Teramitsu, I., & White, S.A. (2008). Motor learning: The FoxP2 puzzle piece. *Current Biology, 18,* R335–337.

Terrace, H.S., Petitto, L.A., Sanders, R.J., & Bever, T.G. (1979). Can an ape create a sentence? *Science, 206,* 891–902.

Vargha-Khadem, F., Gadian, D.G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience, 6,* 131–138.

Vernes, S.C., Newbury, D.F., Abrahams, B.S., Winchester, L, Nicod, J., Groszer, M., et al. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine, 359,* 2337–2345.

Vernes, S.C., Spiteri, E., Nicod, J., Groszer, M., Taylor, J.M., Davies, K.E., et al. (2007). High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *American Journal of Human Genetics, 81,* 1232–1250.

Wallman, J. (1992). *Aping language*. Cambridge, UK: University of Cambridge Press.

Wang, Y., Brzozowska-Prechtl, A., & Karten, H.J. (2010). Laminar and columnar auditory cortex in avian brain. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 12676–12681.

Watkins, K.E., Gadian, D.G., & Vargha-Khadem, F. (1999). Functional and structural brain abnormalities associated with a genetic disorder of speech and language. *American Journal of Human Genetics, 65,* 1215–1221.

White, S.A. (2001). Learning to communicate. *Current Opinion in Neurobiology, 11,* 510–520.

White, S.A., Fisher, S.E., Geschwind, D.H., Scharff, C., & Holy, T.E. (2006). Singing mice, songbirds, and more: Models for FOXP2 function and dysfunction in human speech and language. *Journal of Neuroscience, 26,* 10376–10379.

Whitehouse, A.J., Bishop, D.V., Ang, Q.W., Pennell, C.E., & Fisher, S.E. (2011). CNTNAP2 variants affect early language development in the general population. *Genes, Brain, and Behavior, 10,* 451–456.

Williams, H., & Mehta, N. (1999). Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *Journal of Neurobiology, 39,* 14–28.

Yanagihara, S., & Hessler, N.A. (2006). Modulation of singing-related activity in the songbird ventral tegmental area by social context. *European Journal of Neuroscience, 24,* 3619–3627.

Zeesman, S., Nowaczyk, M.J., Teshima, I., Roberts, W., Cardy, J.O., Brian, J., et al. (2006). Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. *American Journal of Medical Genetics. Part A, 140,* 509–514.

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology, 4,* Article 17.