Of Mice, Birds, and Men:

The Mouse Ultrasonic Song System and Vocal Behavior

by

Gustavo Arriaga

Department of Neurobiology
Duke University

Date:_____

Approved:

_____
Erich D. Jarvis, Supervisor

_____
Nell B. Cant, Chair

_____
Stephen Nowicki

_____
Timothy E. Holy

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the Department of
Neurobiology in the Graduate School
of Duke University

2011

ABSTRACT

Of Mice, Birds, and Men: The Mouse Ultrasonic Song System and Vocal Behavior

by

Gustavo Arriaga

Department of Neurobiology
Duke University

Date:_____

Approved:

_____
Erich D. Jarvis, Ph.D., Supervisor

_____
Nell B. Cant, Chair

_____
Stephen Nowicki

_____
Timothy E. Holy

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Neurobiology in the Graduate School
of Duke University

2011

Abstract

Mice produce many ultrasonic vocalizations (USVs) in the 30 – 100 kHz range including pup isolation calls and adult male songs. These USVs are often used as behavioral readouts of internal states, to measure effects of social and pharmacological manipulations, and for behavioral phenotyping of mouse models for neuropsychiatric and neurodegenerative disorders; however, little is known about the biophysical and neurophysiological mechanisms of USV production in rodents. This lack of knowledge restricts the interpretation of data from vocalization-related experiments on mouse models of communication disorders and vocal medical conditions. Meanwhile, there has been increased interest in the social communication aspect of neural disorders such as autism, in addition to the common disorders involving motor control of the larynx: stroke, Parkinson's disease, laryngeal tremor, and spasmodic dysphonia. Therefore, it is timely and critical to begin assessing the neural substrate of vocal production in order to better understand the neuro-laryngeal deficits underlying communication problems.

Additionally, mouse models may generate new insight into the molecular basis of vocal learning. Traditionally, songbirds have been used as a model for speech learning in humans; however, the model is strongly limited by a lack of techniques for manipulating avian genetics. Accordingly, there has long been strong interest in finding a mammalian model for vocal learning studies. The characteristic features of accepted

vocal learning species include programming of phonation by forebrain motor areas, a direct cortical projection to brainstem vocal motoneurons, and dependence on auditory feedback to develop and maintain vocalizations. Unfortunately, these features have not been observed in non-human primates or in birds that do not learn songs. Thus, in addition to elucidating vocal brain pathways it is also critical to determine the extent of any vocal learning capabilities present in the mouse USV system.

It is generally assumed that mice lack a forebrain system for vocal modification and that their USVs are innate; however, these basic assumptions have not been experimentally tested. I investigated the mouse song system to determine if male mouse song behavior and the supporting brain circuits resemble those of known vocal learning species. By visualizing activity-dependent immediate early gene expression as a marker of global activity patterns, I discovered that the song system includes motor cortex and striatal regions active during singing. Retrograde and anterograde tracing with pseudorabies virus and biodextran amines, respectively, revealed that the motor cortical region projects directly to the brainstem vocal motor nucleus ambiguus. Chemical lesions in this region showed that it is not critical for producing innate templates of song syllables, but is required for producing more stereotyped acoustic features of syllables. To test for the basic components of adaptive learning I recorded the songs of mechanically and genetically deaf mice and found that male mice depend on auditory feedback to develop and maintain normal ultrasonic songs. Moreover, male mice that

display natural strain specific song features may use auditory experience to copy the pitch of another strain when housed together and stimulated to compete sexually. I conclude that male mice have neuroanatomical and behavioral features thought to be unique to humans and song learning birds, suggesting that mice are capable of adaptive modification of the spectral features of their songs.

Dedication

This dissertation is dedicated to all the teachers that told me I'd never amount to nothing; to all the people that stood around the lab bench that I was hustling in front of; reviewers that rejected my manuscripts when I was just trying to make some data to feed my P.I.; and all the nerds in the struggle. It's all good baby, baby.

And to the Fam: La Generala, Mega Man, Adri & Ivan & Bam-Bam, Israel & Gio-Gio for always being there, The Fat Man in a Little Suit for keeping the hope of a comfy couch alive, The Little Girl for keeping a brother alive, Magnum 'Trouble' Vice-Versa for putting up with and sticking by me, The Mad Iranian & The Brain for being rad, Black Beauty the Boozie for keeping it crunked & classy, Simorgh Nemo for reminding me to look up at the stars from time to time, The Hippies & Suri for always dreaming big, CBNC for keeping me mostly sane, and of course Big Joker, Little Joker, and Joker from Around the Block.

Contents

List of Figures

List of Supplementary Files

Supplementary Audio 1: Example of a normal adult BxD mouse song (audio corresponds to sonogram of USVs in Figure 1).

Supplementary Audio 2: Example of adult BxD mouse song 1 week before sham brain lesion surgery (audio corresponds to sonogram of USVs in Figure 26a).

Supplementary Audio 3: Example of adult BxD mouse song 3 weeks after sham brain lesion surgery (same mouse as Supplementary Audio 2; audio corresponds to sonogram of USVs in Figure 26a).

Supplementary Audio 4: Example of adult BxD mouse song 1 week before lesions in laryngeal M1 (audio corresponds to sonogram of USVs in Figure 26b).

Supplementary Audio 5: Example of adult BxD mouse song 3 weeks after lesions in laryngeal M1 (same mouse as Supplementary Audio 4; audio corresponds to sonogram of USVs in Figure 26b).

Supplementary Audio 6: Example of a normal hearing-intact juvenile BxD mouse song on the first day of female-elicited singing (audio corresponds to sonogram of USVs in Figure 31).

Supplementary Audio 7: Example of a mechanically deafened juvenile BxD mouse song on the first day of female-elicited singing (audio corresponds to sonogram of USVs in Figure 31).

Supplementary Audio 8: Example of adult BxD mouse song 1 month before sham deafening surgery (audio corresponds to sonogram of USVs in Figure 34a).

Supplementary Audio 9: Example of adult BxD mouse song 1 month before deafening by cochlear removal (audio corresponds to sonogram of USVs in 34b).

Supplementary Audio 10: Example of adult BxD mouse songs 8 months after sham deafening surgery (same mouse as Supplementary Audio 8; audio corresponds to sonograms of USVs in Figure 36a).

Supplementary Audio 11: Example of adult BxD mouse songs 8 months after deafening by cochlear removal (same mouse as Supplementary Audio 9; audio corresponds to sonograms of USVs in Figure 36b).

Supplementary Audios 12 & 13: Examples of normal adult C57 mouse song (both sonograms from the same mouse; audio corresponds to sonograms of USVs in Figure 41a).

Supplementary Audio 14 & 15: Examples of congenitally deaf CASP3 mouse song (both sonograms from the same mouse; audio corresponds to sonogram of USVs in Figure 41b).

Supplementary Movies 1 – 10: Z-stacks through 7 μm (0.5 μm step size) of brainstem at the level of nucleus ambiguus. BDA-filled axons (black) and puncta can be seen in close contact with the soma and primary dendrites of laryngeal motoneurons back-filled with CTb (brown).

Acknowledgements

I thank my advisor Erich Jarvis for his support and guidance throughout the years. I thank the members of my Thesis Committee, Professor Nell Cant and Dean Steve Nowicki of Duke University and Professor Tim Holy of the University of Washington for critical discussions throughout the years I conducted the dissertation research and for critical reading of the dissertation. I thank Dr. Toshio Terashima of Kobe University, Japan, for teaching the laryngeal surgery technique, and Dr. Lynn Enquist of Princeton University for supplying PRV-Bartha. I thank Dr. Simon E. Fischer at the Wellcome Trust Centre for Human Genetics, University of Oxford, UK for providing the FoxP2-R552H line of mice. Research was supported by an NIH pioneer award DP1 OD000448 and an HHMI investigator award to my advisory Erich D. Jarvis, and my NSF pre-doctoral fellowship award.

Chapter 1. Introduction

Many animals use species-specific sounds to exchange information with conspecifics about external and internal conditions such as predator detection, group affiliation, social status, and emotional or reproductive states. Laboratory mice (*Mus musculus*) and rats (*Rattus norvegicus*) participate in a significant amount of communication using ultrasonic vocalizations (USVs) produced at frequencies ranging from 30 - 100 kHz (Constantini and D'Amato, 2006; Portfors, 2007). These vocalizations are referred to as ultrasonic because they are inaudible to humans, but they are well within the hearing range of the vocalizing animals. Traditionally, two types of USVs have been studied extensively in laboratory rodents as measures of internal states: pup isolation calls (Noirot and Pye, 1969; Noirot, 1972; Sales and Smith, 1978; Elwood and Keeling, 1982; Hahn et al., 1987; Hofer and Shair, 1992; Brudzynski et al., 1999; Branchi et al., 2001; D'Amato et al., 2005; Wöhr et al., 2008a; Ise and Ohta, 2009), and adult USVs in aversive or rewarding conditions (Knutson et al., 2002; Brudzynski, 2007; Burgdorf et al., 2007; Wöhr et al., 2008b; Brudzynski, 2009). Pup ultrasonic isolation calls are produced under conditions of cold or separation from the dam and are frequently used as an index of anxiety (Olivier et al., 1994). Reliable elicitation of isolation calls by quantifiable stimuli, and a well characterized developmental trajectory have made pup USVs a useful tool for testing the effects of anxiogenic or anxiolytic compounds (Fish et al., 2000; Dirks et al., 2002; Fish et al., 2004) and for phenotyping mouse models of

neuropsychiatric disorders associated with deficits in vocal communication (Scattoni et al., 2009). Similarly, the well-characterized 22kHz and 50kHz calls of adult rats have been studied extensively as behavioral assays of negative and positive affective states, respectively (Knutson et al., 2002; Brudzynski, 2007; Burgdorf et al., 2007; Brudzynski, 2009). The predictive power of rat USV behavior with regard to affect is considered so reliable that emission rates of USV production are used as a selective breeding phenotype for generating strains as models of high anxiety and depression (Harmon et al., 2008).

Relative to rats, less is known about USV behavior in adult mice. Mouse USVs do not appear to signal affect but are used primarily during non-aggressive social encounters and may facilitate social interactions (Gourbal et al., 2004; Moles et al., 2007; Portfors, 2007). The most well characterized adult mouse USVs are those produced by males in a mating context. Males of many strains produce long bouts of USVs during courtship of a female and after copulation (Nyby, 1983; Gourbal et al., 2004; Constantini and D'Amato, 2006; Portfors, 2007). Male courtship USVs are sexually selective, and pheromones present in female urine are a strong and sufficient trigger (Guo and Holy, 2007). In two-choice experiments females responded with approach behavior preferentially to adult male USVs over pup isolation calls (Hammerschmidt et al., 2009; Musolf et al., 2010), and spent more time with vocalizing males (Pomerantz et al., 1983). Although the general occurrence of male mouse USVs has been known for decades, the

spectro-temporal and syntactic features of male courtship USVs were only recently analyzed in depth.

Holy and Guo showed that courtship USVs from different males contain identifiable syllable types produced in regular temporal patterns that differed between individuals (Holy and Guo, 2005). Moreover, the long strings of syllables they analyzed sounded remarkably similar to bird songs when the pitch of the USVs was shifted to the human audible frequency range. After observing the complexity of mouse USVs and their striking similarity to birdsongs, I wondered what the neural substrate for USV production is and whether mice might share central control mechanisms for vocalization with vocal learning species like songbirds. The generally accepted list of vocal learning species includes three clades of birds (songbirds, parrots, hummingbirds) and several clades of mammals (humans, cetaceans [dolphins and whales], bats, elephants, and pinnipeds [sea lions and seals]) (Janik and Slater, 1997; Jarvis, 2004; Schusterman, 2008). The ability to modify the spectral and syntactic composition of vocalizations is a rare trait that serves as a critical substrate for human speech (Marler, 1970; Doupe and Kuhl, 1999; Jarvis, 2004). This complex learned social behavior is well studied in humans and songbirds. Songbirds have been useful models because they display a capacity for vocal mimicry using a process similar to human speech acquisition (Marler, 1970; Doupe and Kuhl, 1999) and some species are easy to breed and study in the laboratory. Underlying the vocal learning process in both humans and song learning birds are specialized

forebrain circuits so far not found in species that produce only innate vocalizations, despite decades of searching for them (Jarvis, 2004; Jürgens, 2008). Even closely related non-human primate species lack the behavioral and neural elements classically associated with a capacity for vocal learning (Janik and Slater, 1997; Hammerschmidt et al., 2001; Jürgens, 2008). Like non-human primates, mice are assumed to be vocal non-learners (Jarvis, 2004; Enard et al., 2009; Fischer and Hammerschmidt, 2010), but I am not aware of this being experimentally tested.

1.1 Vocal Communication

Many animals communicate by broadcasting species-typical acoustic signals including insects, frogs, birds, and mammals; however, not all of these sounds are vocalizations. I will refer only to sounds that are produced by the vocal organ as vocalizations. The vocal organ in birds and mammals is the syrinx and larynx, respectively. Gross laryngeal anatomy is well conserved among mammals, as described in studies that found the organization of the mouse larynx is very similar to that of humans (Harrison, 1995; Thomas et al., 2009). Most of the laryngeal cartilages and muscles are similarly positioned in both species. Evidence supporting a laryngeal USV source comes mainly from laryngeal nerve transection and electrophysiology studies. Premotor signals to the larynx are transmitted via the superior and recurrent laryngeal nerves whose shared root is the brainstem nucleus ambiguus (Amb). Bilaterally severing the recurrent laryngeal nerve abolished pup and adult USVs (Roberts, 1975;

Nunez et al., 1985). Electrical recordings in anesthetized rats showed that a majority of the Amb motoneurons recorded displayed little background activity and tonic bursts tightly coupled to and preceding sound production by 46 ms (Yajima and Hayashi, 1983). Similar results were obtained for extracellular recordings in awake Southern pigtailed macaques (*Macaca nemestrina*), with bursts in Amb associated with variations in vocal output preceding vocalization by 100-200 ms (Yajima and Larson, 1993). Recent unpublished observations indicate that the explanted mouse larynx is capable of producing sounds displaying the non-linear dynamics characteristic of natural USVs (Berquist et al., 2010); however, these sounds were in the human audible spectrum and it remains unclear if they depend on vibrations of the vocal folds or a whistle mechanism. An example recording of naturally produced male mouse song shifted into frequencies audible to humans can be heard in Supplementary Audio 1. A sonogram representing 1 second from the same USV bout is shown below (Figure 1).

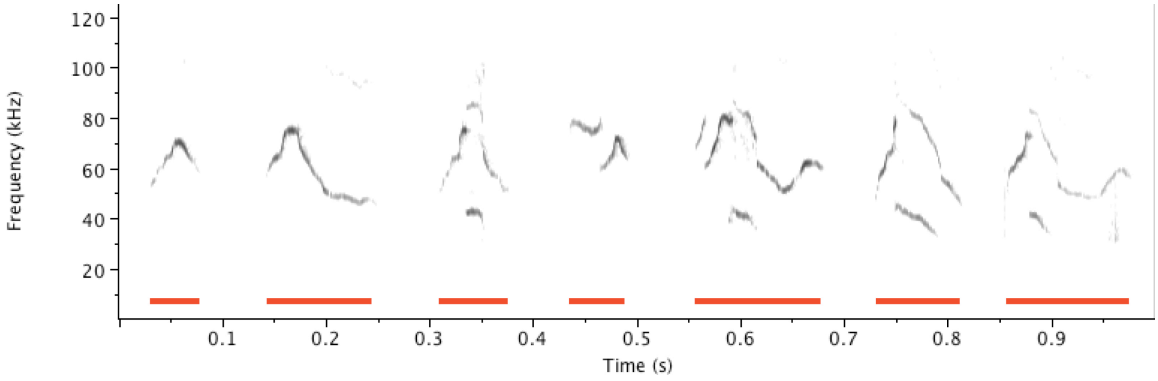

Figure 1: Sonogram of characteristic ultrasonic courtship vocalizations from an adult male BxD mouse produced in response to presentation of female urine. Multiple distinct sound units are produced in long bouts, but only one second is shown so that the frequency contours and nonlinearities of individual units can be resolved. Each unit is marked on the sonogram by a red line. The sonogram was generated from Supplementary Audio 1.

1.1.1 Types of Vocalizations

Vocalizations can take many forms, the parameters of which are often heavily determined by the production and perceptual capabilities of the sender and receiver of an acoustic signal. Example spectrograms of a spoken human sentence, song of a male zebra finch (*Taeniopygia guttata*), call of a ringdove (*Streptopelia risoria*), predator alarm call of a vervet monkey (*Cercopithecus aethiops*), and courtship USV of a male mouse are shown in Figure 2. Unlike human speech, non-human primate calls, and zebra finch songs, adult male USVs are typically comprised of whistle-like notes at frequencies ranging from 35 to 100 kHz. Many species produce a diverse repertoire of vocalizations that can include calls, songs, laughs, and cries. Certainly, courtship USVs are not the only vocalizations mice produce; therefore, I will consider how they can be distinguished from other acoustic signals including pup USVs and adult calls.

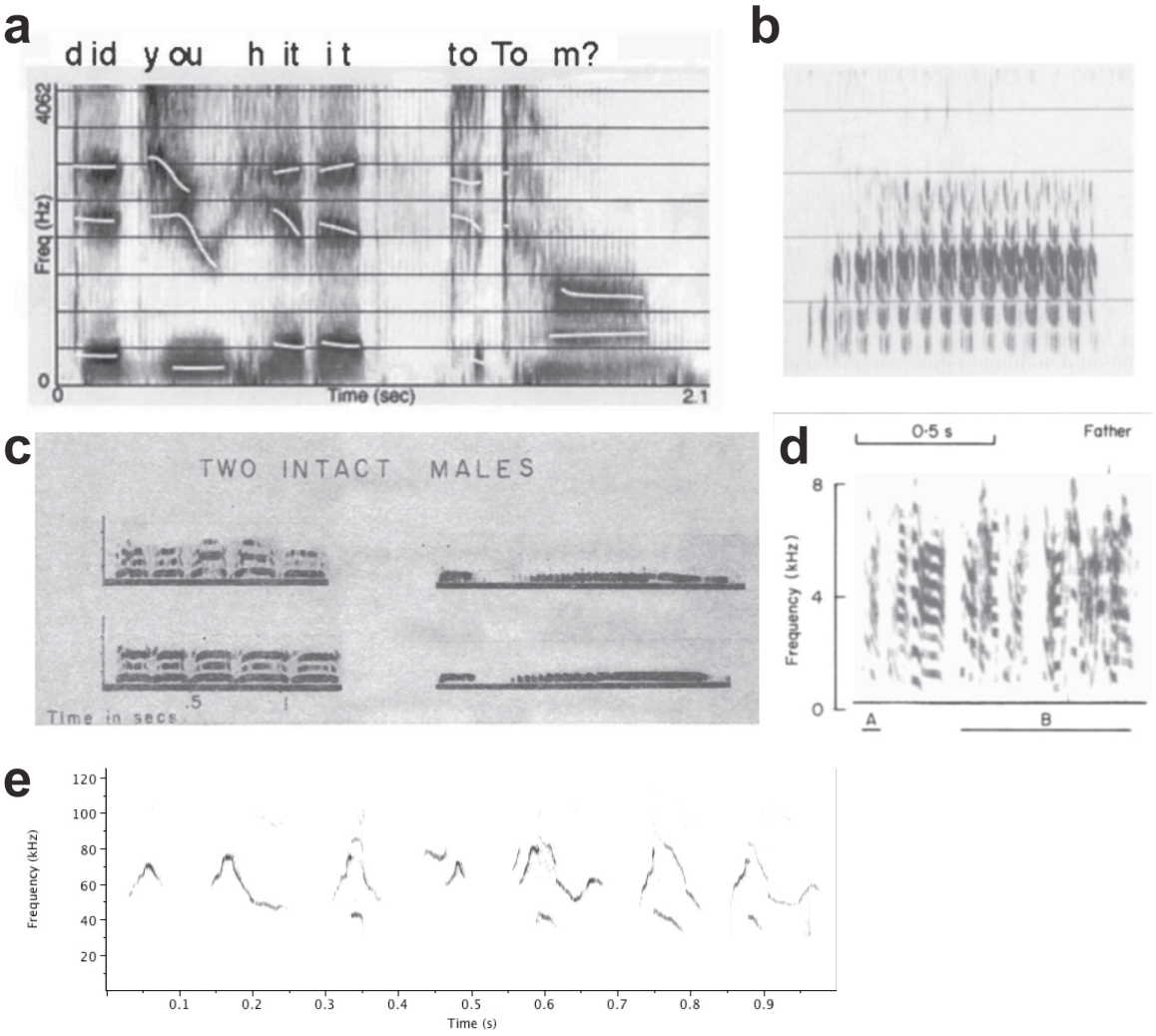

Figure 2: Sonograms of species-typical vocalizations produced by (a) humans (Doupe and Kuhl, 1999), (b) vervet monkeys (Seyfarth and Cheney, 1986), (c) ringdoves (Nottebohm and Nottebohm, 1971), (d) songbirds (Böhner, 1990), and (e) mice.

1.1.1.1 Calls

I will define calls as continuous, distinct, and reproducible units of sound temporally isolated from other sound units by intervals of silence (Doupe and Kuhl, 1999) although calls often contain multiple subunits called notes. Calls are used for communication and may acquire semantic content with experience (Seyfarth et al., 1980; Marler, 2004; Kaplan, 2008). For example, vervet monkeys produce a large repertoire of innate calls that can acquire referential meaning specifying the nature of a predator threat (Seyfarth et al., 1980). Chickens (*Gallus gallus domesticus*) also produce alarm calls to warn of specific predators, and they call selectively in the presence of conspecifics (Gyger et al., 1987; Karakashian et al., 1988). This has been taken as evidence that some calls can be emitted intentionally and are not just an automatic readout of internal state. Additionally, a recent study found that Putty-nosed monkeys (*Cercopithecus nictitans*) form larger vocal units from call combinations that acquire semantic content independent of the content of the individual calls (Arnold and Zuberbühler, 2008). These monkeys were shown to combine 'pyow' and 'hack' calls that normally signify the presence of a leopard or eagle, respectively, into a sequence containing 1-4 'pyows' followed by 1-4 'hacks'. The resulting sequence did not specify the presence of either predator, but instead signaled females to respond with a group progression. In addition to semantic content, the spectral content of some calls can also be learned. For example, when young zebra finch males were cross-fostered to Bengalese finch (*Lonchura striata*

domestica) tutors they learned to produce long calls different from those of normally reared siblings and biological fathers (Zann, 1985).

Mice produce a variety of calls (Whitney and Nyby, 1983). Some, like distress calls, are audible, and others, like courtship vocalizations, are purely ultrasonic. Still others have audible and ultrasonic components. As the term 'ultrasonic' suggests, a broad call category can be easily distinguished simply because USVs are produced in a high frequency range above the upper limit of human hearing, typically around 22 kHz. This frequency range-specific category can be broken down further by considering the vocalization-eliciting stimulus, social context of production, or the age of the vocalizing animal. For example, adult female mice will produce ultrasonic calls for a brief period upon encountering a female intruder (Gourbal et al., 2004; Moles et al., 2007). The most commonly studied ultrasonic mouse calls are those produced when a pre-weaning pup is separated from the mother (Noirot, 1972; Sales and Smith, 1978). A nursing female mouse will respond reliably to pup isolation calls by locating the sound source and retrieving the pup to her nest (Noirot, 1972; D'Amato et al., 2005; Uematsu et al., 2007). Thus, the response of the animal receiving the acoustic signal may provide another feature for differentiating vocalizations produced in overlapping frequency ranges.

1.1.1.2 Syllables

Syllables are reproducible single acoustic units structurally similar to calls in that they can be made up of multiple notes. Although syllables are structurally similar to

calls, I will distinguish them from calls by patterns of usage. Calls are typically produced in isolation or in short bursts and often obtain semantic content on their own, as previously described. Syllables, however, derive their classification from being included in a larger unit representing a longer series of rapidly produced vocalizations. A reproducible series of syllables with a relatively fixed order is labeled a 'motif'. By clustering units into motifs, an animal with a repertoire of only a few syllables can generate a wide variety of larger communication units. In this classification scheme syllables can be void of specific meaning themselves, and they would not necessarily serve a communication function if produced in isolation. This distinction is not always entirely clear for units of sound that can serve a dual function. For example, the long call of male zebra finches can function alone as a contact call or be incorporated into a motif that is reproduced in song bouts (Zann, 1990). In this case, the same unit could be labeled a call or a syllable depending on the context of production.

Adult mouse USVs feature reproducible sound units that I have grouped into general types by their spectral morphology. Most of these units are frequently produced in long sequences containing different types, and there is some evidence that simple motifs can be identified in these sequences (Holy and Guo, 2005). I will call these recurring units of adult male USVs 'syllables' because they are grouped into non-random series, rarely produced in isolation, and there is no evidence that they serve a communication function individually.

I identified the 8 most commonly occurring syllable types from our recordings of adult male B6D2F1/J (BxD) and C57BL6/J (C57) USVs (Figure 3). The first major morphological distinction between syllable types is the presence or absence of an instantaneous 'pitch jump' separating notes within a syllable. Under this classification scheme syllable Type A is the morphologically simplest note type, in that it doesn't contain any pitch jumps. Other researchers have split this category into sub-types according to the trajectory of the fundamental frequency, syllable duration, or the presence of harmonics (Scattoni et al., 2008a; Fischer and Hammerschmidt, 2010); however, I agree with an earlier classification scheme that considers these syllables as different parts of a simple sinusoidal waveform and groups them accordingly (Holy and Guo, 2005). For syllables containing pitch jumps, each jump marks the end of one note and the beginning of the next note. Two-note sequences can be identified by a single upward or downward pitch-jump. Similarly, more complex syllables are identified by the series of upward and downward pitch jumps occurring as the fundamental frequency varies between notes of higher and lower pitch.

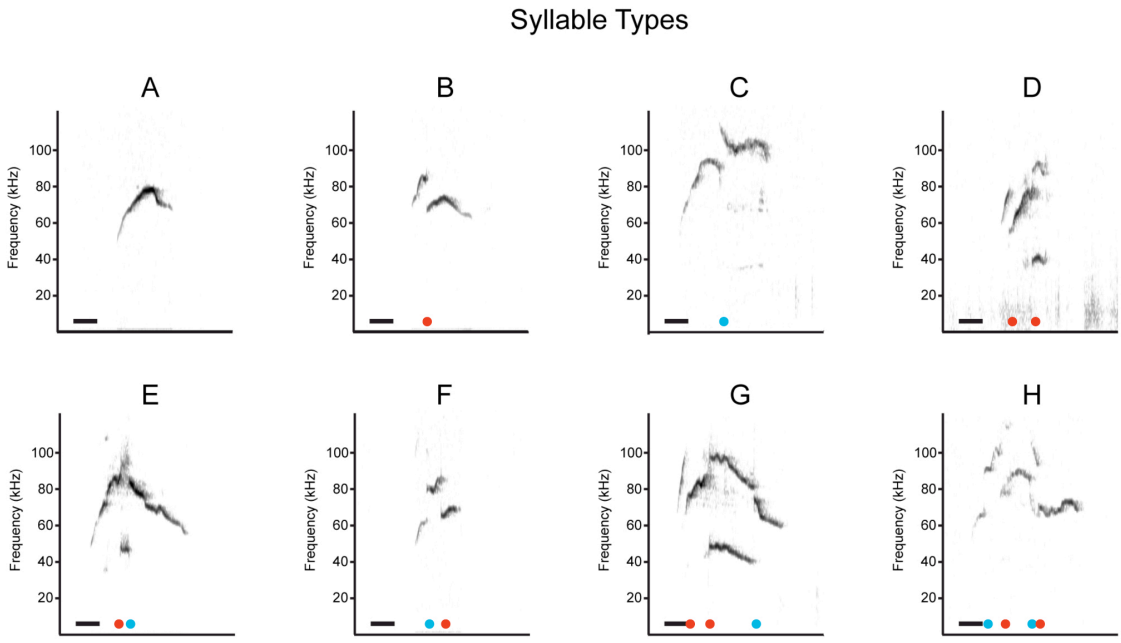

Figure 3: Examples of syllables from courtship vocalizations of adult male BxD mice. Eight major syllable classes can be distinguished by the series of notes and the corresponding sequence of instantaneous jumps (>10kHz) in the dominant pitch (Holy and Guo, 2005). The simplest syllables (Type A) are characterized by a single note with no pitch jumps. Two-note syllables (Types B & C) can be classified by a single 'Down' or 'Up' pitch jump. Common three-note syllables (Types D-F) follow one of the following sequences: 'Down-Down', 'Down-Up', or 'Up-Down'; the fourth possible pitch jump combination 'Up-Up' is rarely observed. Commonly observed four- and five-note syllables (Types G & H) follow 'Down, Down, Up' and 'Up, Down, Up, Down' pitch jump sequences, respectively. Since a jump is defined based on the instantaneous peak frequency, the harmonics in some notes are not considered for classification purposes. Blue dots mark 'Up' jumps, and red lines mark 'Down' jumps. Scale bar: 20 ms.

1.1.1.3 Songs

A song is set of vocalizations, often elaborate, delivered periodically. Songs may be produced spontaneously or in response to an external stimulus such as the presence of a conspecific. Songs typically contain multiple syllable types, or categories of reproducible vocalizations distinct from other vocalizations comprising the song. To distinguish a series of syllables in a song from a succession of calls I will apply the *sensu strictissimo* definition used previously and borrowed from Broughton (Broughton, 1963):

'a sound of animal origin which is not both accidental and meaningless'

containing,

'a series of notes, generally of more than one type, uttered in succession and so related as to form a recognizable sequence or pattern in time,'

produced in,

'a complete succession of periods or phrases'

Holy and Guo's analysis of the spectro-temporal features of male courtship USVs demonstrated convincingly that these vocalizations satisfy all conditions required for classification as song (Holy and Guo, 2005). Visually, the song-like quality of male mouse courtship USVs can be appreciated in spectrograms of longer sequences (Figure 4). Acoustically, when the pitch of courtship USVs is shifted to the audible spectrum they sound very similar to birdsong in both temporal and melodic structure

(Supplementary Audio 1). I will, therefore, now refer to male mouse courtship USVs simply as 'mouse songs'.

The behavioral responses of conspecifics also provide clues that mouse songs are distinct from calls. Recent evidence shows that female mice distinguish male songs from pup isolation calls (Hammerschmidt et al., 2009). Given the choice, females selectively approach the source of the songs instead of the source of the isolation calls. Preference for male songs is striking given that pup calls are considered a very strong and reliable stimulus, and the frequency ranges of the two signals overlap significantly. Moreover, there was a slight tendency to prefer the songs of non-kin males, further suggesting that individual songs are distinguishable and could serve an important social and reproductive function.

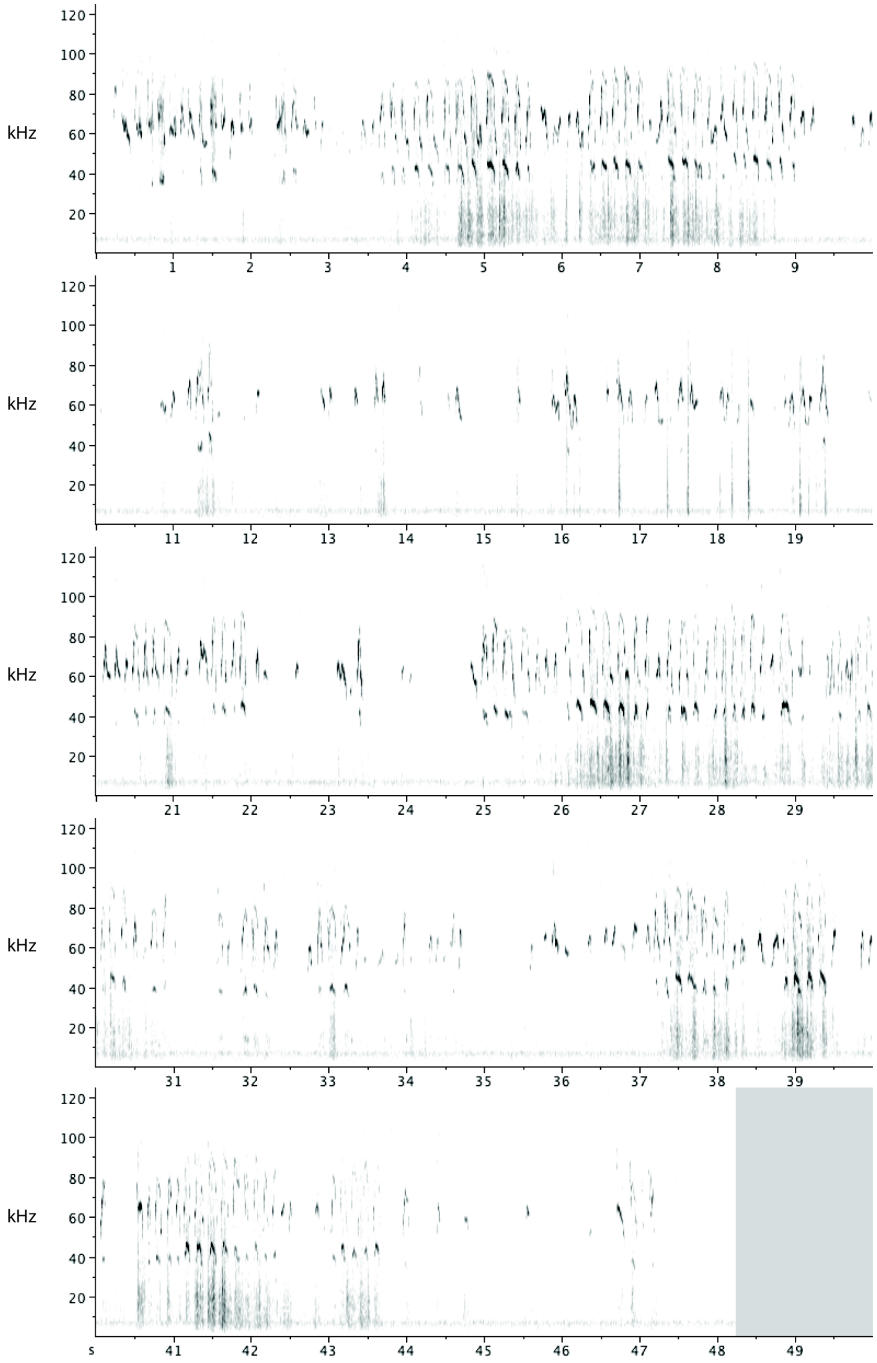

Figure 4: Song bout of an adult BxD male lasting 47 seconds and containing 264 syllables.

1.2 Vocal Learning

Many types of learning are possible within the framework of vocal communication systems. Thus, it is important not only to determine what learning capabilities are present in the mouse vocal system, but also to distinguish which types of learning are most relevant to studies of speech learning in humans. Three types of learning related to vocal communication systems have been addressed previously: comprehension learning, usage learning, and production learning (Janik and Slater, 1997; Janik and Slater, 2000; Egnor and Hauser, 2004; Jarvis, 2004; Schusterman, 2008; Abbott et al., 2009). Comprehension and usage learning both concern learning about vocalizations, and form complimentary components of what is called context learning. Production learning concerns the learning of vocalizations by experience-dependent modification of acoustic signals and is, therefore, considered the most relevant to the study of human speech (Janik and Slater, 1997; Jarvis, 2004). I will use the term 'vocal learning' to refer to comprehension learning specifically, because it results in spectral rather than contextual changes in vocalizations.

1.2.1 Comprehension Learning

Comprehension learning is a learning strategy characterized by the ability to associate a particular sound with an appropriate behavioral response or objects in the environment (Jarvis, 2004). Comprehension learning capabilities are broadly distributed among many species. For example, dogs (*Canis lupus familiaris*) can be trained to learn to

respond to the human word 'sit'; however, the vocal part of this training process is restricted to the act of correctly identifying the word through auditory learning. Learning in this case does not extend to the production (i.e. vocal motor) component. Dogs don't learn to produce the word 'sit' by adaptively modifying motor commands to achieve the required sequence of laryngeal and respiratory patterns. However, some motor behaviors can be associated with a learned auditory cue. For example, the dog's typical response to the verbal command 'sit' is the motor act of sitting on the hindquarters.

1.2.2 Usage Learning

Usage learning is a behavioral strategy by which animals learn to produce vocalizations in a specific social or environmental context. A well-studied example of usage learning is the alarm call repertoire of vervet monkeys produced in response to specific predator threats. An eagle (*Polemaetus bellicosus*) in the sky, a leopard (*Pantheru pardus*) in the trees, and a python (*Python sebae*) on the ground will elicit different species-typical calls, and a young vervet monkey must learn through experience when each call is appropriate to use (Seyfarth et al., 1980); however, the spectral content of alarm calls is thought to be innately determined (Seyfarth and Cheney, 1986). Learning is restricted to the context or 'when' of production, but the 'how' is inflexible. Usage learning and comprehension learning are often intimately linked. For example, it is critical that a young vervet monkey learn not only which call to produce in response to

each predator, but also learn the appropriate predator-specific defensive behavior to produce upon hearing each call. The leopard-specific call triggers retreat into the trees, and the eagle-specific call causes listeners to hide in the dense bush (Seyfarth et al., 1980). The learned association of auditory cues with effective predator defense strategies is similar to the training of a dog's behavior to verbal commands. Additionally, there is evidence from putty-nosed monkeys that some non-human primates can learn to produce novel combinations of calls, and the combination acquires semantic content independent of the individual calls used (Arnold and Zuberbühler, 2008). Although the semantic content of calls and syntactical operations on those calls can be learned, the spectral content is fixed in the examples discussed.

1.2.3 Production Learning

Production learning is the process of learning to modify the spectral content of vocalizations. I adopt this definition of vocal production learning which excludes changes in the amplitude and duration of vocalizations because they rely on control of respiration patterns rather than control of the laryngeal musculature of the vocal organ (Janik and Slater, 1997). In this context, the most dramatic and well-studied examples of vocal production learning are song learning in birds and speech learning in humans. Birdsong and speech share many features: auditory acquisition of learning templates, dependence on auditory feedback for learning and maintenance of learned vocalizations, temporally restrictive critical periods for learning, and specialized

forebrain networks for vocal control (Doupe and Kuhl, 1999; Jarvis, 2004). Because of these important similarities, songbirds have become the dominant neuro-ethological animal models for vocal learning studies (Brainard and Doupe, 2002). One consequence of the intense focus on the songbird model is a situation where the meaning of the term 'vocal learning' has been restricted to refer exclusively to learning vocalizations *de novo* with reference to an externally acquired model, as occurs for birdsong and speech learning. Certainly, this type of vocal mimicry is the most relevant to study for modeling and understanding the process of human speech acquisition; however, I believe this represents an overly restrictive definition of vocal learning that ignores many other factors and strategies that can be used to adaptively modify the spectral content of vocalizations. Discussing white-crowned sparrows (*Zonotrichia leucophrys*) that produce songs despite having been raised in social isolation, Masakazu Konishi highlighted the categorical problem presented by reducing learning to mimicry alone:

> Isolate songs are sometimes referred to as "innate songs." Interspecific differences in isolate songs are likely to be due to genetic differences between species. The source of instruction for isolate songs must be internal, because the bird hears no song to copy. These reasons justify the definition of an innate song as above. However, this practice of naming innate songs creates an interesting logical dilemma, when one considers how song learning is defined. As pointed out above, auditory feedback control of voice is regarded as learning. Because the development of innate songs requires auditory feedback, the innate songs are learned! By this logic only the sound patterns that develop without auditory feedback can be called innate. (Konishi, 1985)

The process of generating an isolate song without previous instruction, or adding novel parts to a tutored song has been called improvisation (Konishi, 1964; Janik and

Slater, 1997; Kroodsma et al., 1997; Marler, 1997). Improvisation is one of the simplest ways that animals may change their vocalizations without explicit need for a tutored model. Using improvisation an animal could rely on internal preference or the response of conspecifics to guide the learning process. Therefore, it is important to evaluate the relative roles of improvisation and imitation in any vocal learning species. In some experiments, songbirds like the grey catbird (*Dumetella carolinensis*) often failed to copy song models and routinely generated normal songs with novel elements not present in the template (Kroodsma et al., 1997). More strikingly, when the abnormal song of a socially isolated adult zebra finch was used as the tutor template, the tutored juveniles modified the song to more closely match a more typical finch song (Fehér et al., 2009). Accumulation of corrective improvisations over 5 generations was sufficient to transform the isolate song to a normal-sounding zebra finch song. Preferential learning by improvisation was performed even though all the birds should be perfectly capable of mechanically reproducing the isolate songs heard.

In some vocal learning species determination of what is worth learning is left up to individuals other than the one learning. For example, non-singing female cowbirds (*Molothrus ater*) exert a strong sexual selection on male song development by selectively reinforcing song variants with their wing displays (West and King, 1988). The effect of female selection is so strong that both tutored and untutored males developed different songs depending on the preferences of co-housed females from different sub-species

(King, 1983). Experiments with Pacific walruses (*Odobenus rosmarus divergens*) demonstrated that the preferences of human trainers could also reinforce novel vocal behavior (Schusterman and Reichmuth, 2008). Using a contingency learning paradigm, walruses were rewarded with fish when a vocalization was judged by the human trainer to be significantly different than the preceding vocalization. Under stimulus control, sounds in the existing repertoire were elaborated with pitch and contour changes, and several novel vocalizations emerged that had not been heard before.

It is clear that mimicry is not the only viable strategy for vocal production learning, and the spectral content of vocalizations can be modified by improvisation reinforced by individual preference or the selective distribution of social and food reinforcements. Indeed, different strategies could have been necessary for different species to transition from generating exclusively innate sounds to generating novel sounds. For these reasons, I will attempt to avoid the logical conundrum described by Konishi (Konishi, 1985) by accepting as 'vocal learning' the development of any vocalizations that depend on auditory feedback for the development or maintenance of spectral content. Under this definition it is the reliance on auditory feedback to guide the trajectory of sound development that is important. Less important is whether the trajectory results in convergence toward or divergence from an external model, the emergence of internal preferences, or acquisition of a social or food reward.

1.2.4 Animal Models of Human Speech

The study of spoken language encompasses many aspects including semantics, grammar and syntax, prosody, lexicology, and phonology; however, for the purposes of this study the most important features are those that pertain to the vocal motor component of language: phonology, prosody, and syntax.

1.2.4.1 The Songbird Model

Birdsongs are composed of ordered strings of syllables with precise timing and sequence structure, and therefore provide a fitting model for the motor production aspect of language. As mentioned previously, there are numerous developmental parallels between birdsong and human speech that have made songbirds the most widely used animal model of human speech production and learning. The most critical similarities between the two vocal behaviors include the following traits (Marler, 1970; Doupe and Kuhl, 1999): 1) a reliance on learning by mimicry of previously experienced sounds; 2) an innate perceptual predisposition for particular types of sounds, typically species-specific; 3) dedicated networks for vocal production learning; and 4) critical periods for proper development of vocalizations.

The songbird model has yielded a tremendous amount of information regarding the ecology, ontology, and neurobiology of birdsong (Zeigler and Marler, 2008). In spite of such progress, however, significant barriers remain to linking birdsong and speech at a more fundamental level. For example, it has been determined that premotor activity at

the highest level of the vocal pathway in the avian song system follows a sparse code consisting of a sequence of temporally precise spike bursts in distinct populations of premotor neurons (Hahnloser et al., 2002). The sparse time code sets the tempo of the song and is translated into a temporally precise sequence of ensemble burst patterns in the downstream arcopallial premotor nucleus (Fee et al., 2004; Leonardo and Fee, 2005). By linking various synfire chains representing discrete acoustic elements, the sparse timing code could generate reliable phonological patterns at the level of notes or syllables, and stereotyped syntactical patterns at the level of chunks or motifs. Unfortunately, it is unknown if similar premotor circuit dynamics underlie speech programming in the human brain. Although simple synfire chains provide one possible strategy for encoding the serial order of articulations at the level of phonemes, they are thought to be inadequate for encoding meaningful language units at the level of words and sentences (Pulvermüller, 2002). Thus, the synfire model studied in songbirds has potentially limited applicability to the basic understanding of human speech processing and development of treatments for many speech disorders that affect more than the serial order of phoneme articulation.

Significant obstacles have been encountered in the songbird system in two other areas: 1) difficulty establishing the nature of auditory feedback and instructive error signals (Leonardo, 2004); 2) lack of efficient methods for generating transgenic songbirds (Agate et al., 2009) or directly manipulating genes *in vivo* (Wada et al., 2006; Haesler et

al., 2007; Schulz et al., 2010). In the case of auditory feedback, the well-characterized, and easily accessed auditory cortex of rodents may provide a useful tool for investigating processing of instructive auditory information. With regard to genetic manipulation, direct manipulation of gene expression in live animals has been achieved using lentivirus-mediated RNA interference knockdown of FoxP2 in the songbird basal ganglia (Haesler et al., 2007; Schulz et al., 2010). Knockdown resulted in reduced spiny neuron dendritic spine densities and poor imitation of tutor song; however, the experience of multiple songbird laboratories, including our own, suggests that successful lentiviral delivery of genetic constructs in the avian brain is stochastic. As gene manipulation tools in songbirds continue to be developed, there are already well-established protocols for genetic manipulation in rodents that may provide a useful tool for molecular investigations of vocal communication systems. The importance of a genetically tractable model for vocal communication studies continues to grow as more associations are being discovered between gene variants and language disorders like specific language impairment, speech sound disorder, dyslexia, and autism (Ramus and Fisher, 2009; Fischer and Hammerschmidt, 2010).

1.2.4.2 Towards a Mouse Model of Vocal Communication

The search for a suitable mammalian model for the study of speech and vocal learning has continued for decades and has traditionally focused on non-human primates (Deacon, 1989; Ghazanfar and Hauser, 1999; Jürgens, 2002; Egnor and Hauser,

2004; Jürgens, 2008). The characterization of primate vocal systems led to the discovery of two critical differences from the human speech system: 1) no reliance on auditory experience for learning of species-specific calls in primates (Egnor and Hauser, 2004); 2) the absence of a direct motor cortical vocal pathway in primates (Jürgens, 2008). After extensive study, the general agreement is that non-human primates lack both reliance on auditory experience for early vocal development and direct cortical control of acoustic output; however, some researchers still argue that these species can provide a model for some of the basic processes of vocal production and perception that serve as the substrates for spoken language (Ghazanfar and Hauser, 1999; Egnor and Hauser, 2004).

There are mammalian species known to be capable of vocal learning that could provide valuable insights into the neurobiology of speech. Unfortunately, most of the known learners (cetaceans, elephants, pinnipeds) are not suitable for the types of studies required, due to large body sizes and inconvenient habitats. So, none of the careful examination dedicated to the primate species has been similarly directed at the majority of the known vocal learners. There has been some detailed study of bats including characterizing the midbrain and brainstem vocal premotor systems for echolocation and communication calls (Fenzl and Schuller, 2005), testing the effects of deafening on vocalizations (Rübsamen and Schäfer, 1990), and demonstrating call convergence (Boughman, 1998). Unfortunately, bats aren't well suited for large-scale study, and advanced genetic manipulation tools have not been developed for bats. I propose that if

mice have at least the same features as the non-human primate vocal system (indirect motor cortical vocal pathway and vocal convergence) or something more similar to the known vocal learners like bats, then the laboratory mouse is well positioned as a supplementary model for vocal communication because of the added benefit of being genetically tractable. Thus, it is necessary to characterize the mouse vocal system both for general knowledge of mouse vocal communication and as a potential model for certain aspects of human speech.

1.3 Aims of the Study

To extend the knowledge of the basic conditions of USV production in mice, I examined the fundamental neural and ontological elements of adult male courtship songs with the following specific aims:

1) Neocortical premotor areas, motor cortex, and the basal ganglia in non-human mammals have not been previously shown to be active during the production of vocalizations (Jürgens, 2002; Egnor and Hauser, 2004; Jarvis, 2004). Guided by previous work describing activity in specialized vocal circuits of vocal learning birds by detecting expression levels of activity-dependent immediate early genes, I investigated whether ultrasonic singing with or without auditory feedback activates mouse motor cortex and basal ganglia.

2) To map the vocal premotor pathways, and complement the behavioral gene mapping technique used in Aim 1, I applied retrograde and anterograde neural tracers with a focus on identifying connections between the forebrain and brainstem. I injected pseudorabies virus into laryngeal muscles to retrogradely label multisynaptic vocal premotor circuits convergent on Amb. I used biodextran amines to anterogradely trace descending connections from the motor cortex to the brainstem phonatory motoneurons.

3) The human face motor cortex and an analogous song region of avian arcopallium are required for the production of learned vocalizations (Nottebohm et al., 1976; Simpson and Vicario, 1990; Terao et al., 2007). In non-human primates motor cortical regions are dispensable for production of innate vocalizations (Aitken, 1981; Jürgens et al., 1982; Kirzinger and Jürgens, 1982). To test the hypothesis that motor cortical premotor pathways are critical for song production in mice, I analyzed song behavior after placing ibotenic acid lesions in motor cortex.

4) All forms of vocal learning, including but not limited to mimicry, require auditory experience and feedback. To test the hypothesis that mice require auditory feedback for normal song production I analyzed song behavior in juveniles with conductive hearing loss, adults before and

after bilateral cochlear removal, and a strain exhibiting congenital deafness.

5) To test the hypothesis that mice are capable of adaptive vocal modification I cross-housed adult male mice of two strains with significant differences in the spectral qualities of their songs, and analyzed the post-cross songs for evidence of song modification.

Chapter 2. Brain Networks for Vocalization

The vocal pathways in birds and mammals involve subsystems that contribute differentially to the initiation of vocalizations and the spectral structure of emitted sounds. It has been proposed that two different, but converging pathways are involved in the production of learned and innate vocalizations (Jürgens, 2008). According to this division of labor innate calls are programmed by a phylogenetically older pathway, and the forebrain influences the context of calling but not acoustic structure. In contrast, control of the spectral content of learned calls has been given over to a vocal pathway driven directly by forebrain premotor structures.

All vocalizing avian and mammalian species studied to date possess a basic circuit for the gating and programming of innate vocalizations which includes midbrain premotor structures, and brainstem motoneuron pools for motor control of phonation and respiration. In both vocal learning and non-learning birds the innate vocal circuit comprises the dorsomedial nucleus of the intercollicular complex (DM) in the midbrain that projects to multiple medullary nuclei including the parabrachial region (PBr), the expiratory premotor nucleus retroambigualis (RAm), and the syringeal motoneuron pool located in the tracheosyringeal part of the hypoglossal nucleus (nXIIts) (Figure 5) (Wild et al., 1997). The role of DM in call production was demonstrated in both singing male and non-singing female Bengalese finches by electrical recordings of DM firing rates coupled to call production (Fukushima and Aoki, 2002), electrical stimulation of

DM which produced calls in awake and freely moving birds (Fukushima and Aoki, 2000; Fukushima and Aoki, 2002), and bilateral lesions of DM which eliminated distance calls (Fukushima and Aoki, 2000). Electrical stimulation also produced calls in awake, freely moving seagulls and pigeons, birds that do not exhibit vocal learning (Delius, 1971).

An analogous vocal circuit is present in mammalian brains beginning with the caudal periaqueductal gray (PAG) which projects to brainstem respiratory premotor nuclei including RAm for control of respiration, and cranial nerve nuclei including Amb which directly innervates the larynx (Figure 5) (Mantyh, 1983; Ennis et al., 1997; Jürgens, 1998; Jürgens, 2002; Jürgens, 2008). Electrical or chemical stimulation in PAG has been found to produce vocalizations in multiple species including cats, bats, and monkeys (Jürgens and Pratt, 1979; Lu and Jürgens, 1993; Zhang et al., 1994; Fenzl and Schuller, 2002; Fenzl and Schuller, 2005). The most well studied mammalian models of vocalization are the squirrel monkey (*Saimiri sciureus*) and rhesus macaque (*Macaca mulatta*). Decades of work by Uwe Jürgens and colleagues using anatomical tracing (Müller-Preuss and Jürgens, 1976; Müller-Preuss et al., 1980; Jürgens, 1982; Jürgens, 1983; Jürgens, 1984; Thoms and Jürgens, 1987; Jürgens and Alipour, 2002; Simonyan and Jürgens, 2002; Simonyan and Jürgens, 2003; Simonyan and Jürgens, 2004; Dujardin and Jürgens, 2005; Hannig and Jürgens, 2005; Simonyan and Jürgens, 2005), brain imaging (Jürgens et al., 2002), electrophysiology (Lüthe et al., 2000; Jürgens, 2002; Düsterhöft et

al., 2003; Hage and Jürgens, 2006a; Hage and Jürgens, 2006b), electrical (Jürgens and Ploog, 1970) and chemical (Lu and Jürgens, 1993) brain activation, lesions (Jürgens and Pratt, 1979; Jürgens et al., 1982; Kirzinger and Jürgens, 1985), and reversible inactivations (Siebert and Jürgens, 2003; Jürgens and Ehrenreich, 2007) has produced a detailed description of the pathways involved in controlling innate mammalian vocalizations (Jürgens, 2008). The general conclusions drawn from this body of work are as follows: 1) limbic regions regulating arousal and the drive to vocalize including the amygdala and cingulate cortex converge on the PAG; 2) the PAG serves a gating function to activate motor programs for specific calls associated with different arousal states; 3) the spectral structure of calls is primarily determined at the level of medullary premotor circuits that coordinate the activity of phonatory motoneuron pools in various cranial nerve nuclei (Jürgens, 1998; Jürgens, 2002; Jürgens, 2008). This interpretation of the pathways suggests that what is truly indispensable for vocalization is the PAG and downstream circuits of the brainstem. In accordance with this model of vocal control, lesions of limbic cortical structures in monkeys only abolish conditioned vocalizations but leave intact both spontaneous vocalizations (Sutton et al., 1974) and calls produced by PAG stimulation (Jürgens and Pratt, 1979). By contrast, damage to the PAG results in mutism (Jürgens and Pratt, 1979).

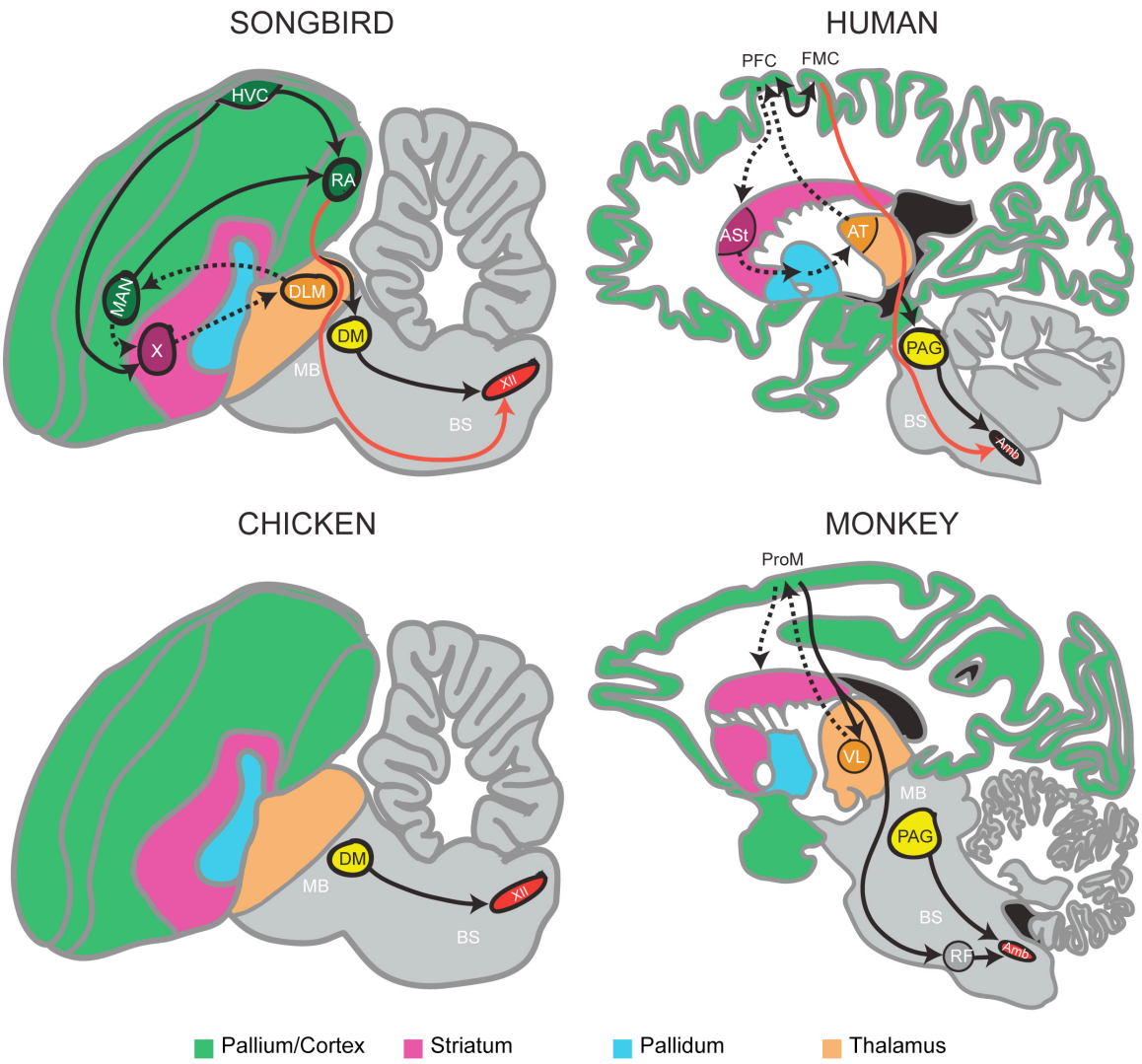

Figure 5: Summary diagrams of brain systems for vocalization in vocal learning and vocal non-learning species. All vocalizing species including monkeys and chickens possess an indirect cortical pathway to the vocal motor nucleus in the brainstem. The vocal learning species possess additional forebrain premotor circuits, including cortico-striatal-thalamic loops (dotted lines) and a direct forebrain projection to vocal motoneurons in the brainstem (red arrows: RA to XII in songbirds; Face motor cortex (FMC) to Amb in humans) (Kuypers, 1958c; Jarvis, 2004; Jürgens, 2008). The direct cortico-bulbar projection is absent in vocal non-learners such as chickens and monkeys. All diagrams show the sagittal view.

2.1 Background - Specialized Forebrain Motor Systems for Vocal Learning

In addition to the limbic-midbrain-brainstem pathway for innate vocal production, vocal-learning species have evolved cortico-striatal-thalamic loops and cortico-bulbar pathways for learning and generating novel vocalizations, respectively. Although the gross anatomy of avian and mammalian forebrains is remarkably different there are some general principles shared among all vocal learning systems. The general features of the songbird song system and human speech systems will be described separately.

2.1.1 Song System in Vocal Learning Birds

The hierarchically organized pre-motor control pathway of the oscine song system is contained within two nuclei of the caudal telencephalon and sends direct and indirect output to the vocal motoneurons of the brainstem located in nXIIts (Wild, 1997). The song premotor pathway begins with the nucleus HVC (proper name, not an abbreviation), from which a specific subset of projection neurons innervates the robust nucleus of the arcopallium (RA) (Nottebohm et al., 1976; Foster and Bottjer, 1998). These RA-projecting neurons appear to encode the timing of song via a sparse code that coordinates the bursting activity of neuron ensembles in RA (Yu and Margoliash, 1996; Hahnloser et al., 2002; Fee et al., 2004; Leonardo and Fee, 2005). RA projects to various midbrain and brainstem nuclei including DM of the innate call generating pathway, the respiratory premotor nucleus RAm, Amb, and the motoneurons of nXIIts directly

involved in controlling the vocal organ (Nottebohm et al., 1976; Wild, 1993). The downstream targets of RA make it well positioned to allow forebrain control over the activity of respiratory, laryngeal, and syringeal muscles groups during vocalization. Both HVC and RA display vocalization-specific premotor neural activity (Yu and Margoliash, 1996; Hahnloser et al., 2002; Fee et al., 2004; Leonardo and Fee, 2005), and are critical for song production throughout life (Nottebohm et al., 1976; Simpson and Vicario, 1990).

A similarly connected hierarchical vocal premotor pathway was found in the forebrain of parrots (Striedter, 1994; Jarvis and Mello, 2000) and hummingbirds (Gahr, 2000; Jarvis et al., 2000). In parrots the pathway involves analogous projections from the nucleus of the lateral nidopallium to the central nucleus of the anterior arcopallium (AAc), which projects in turn to midbrain and brainstem vocal nuclei (Striedter, 1994). Bilateral lesions to AAc, like those to RA, eliminate warble songs and the learned components of contact calls (Heaton and Brauth, 2000). In hummingbirds, a nucleus similar in location and cytoarchitecture to songbird HVC was found called HB-HVC (Gahr, 2000). HB-HVC sends descending projections to HB-RA, which resembles songbird RA and innervates nXIIts (Gahr, 2000).

Songbird HVC also provides indirect input to RA via another forebrain vocal circuit called the anterior forebrain pathway (AFP). The AFP begins with a distinct subset of HVC projection neurons that innervate a region of the anteromedial striatum

specialized for vocal learning called Area X (Nottebohm et al., 1976; Foster and Bottjer, 1998). Area X sends a GABAergic projection to the dorsolateral anterior thalamic nucleus (DLM), which projects in turn to the lateral magnocellular nucleus of the anterior nidopallium (LMAN) (Okuhata and Saito, 1987; Bottjer et al., 1989; Person et al., 2008). LMAN then projects back to Area X forming a cortico-striatal-thalamic loop specialized for vocalization (Okuhata and Saito, 1987); however, LMAN is also the output nucleus of the AFP, projecting to RA (Nottebohm et al., 1982; Bottjer et al., 1989) and allowing the AFP to modulate the ongoing activity of the direct HVC-RA premotor circuit (Kao et al., 2005). Lesions and chemical inactivation of LMAN and Area X revealed that the AFP is not strictly required for singing (Nottebohm et al., 1976), but is critical for generating the acoustic variability necessary for vocal exploration in normal song learning (Bottjer et al., 1984; Scharff and Nottebohm, 1991; Olveczky et al., 2005), social context-dependent modulation of song (Kao et al., 2005; Kao and Brainard, 2006), and experimentally-induced song deterioration (Brainard and Doupe, 2000).

A similar recurrent cortico-striatal-thalamic pathway was found in the forebrain of parrots (Durand et al., 1997). And in hummingbirds, analogous basal ganglia and cortical regions have been found to be active during song production (Jarvis et al., 2000). The connectivity between these AFP-like regions has not been established in hummingbirds except for the projection from the proposed LMAN analogue to HB-RA, similar to the oscine song system (Gahr, 2000). Thus, the general design of several

similarly arranged discrete forebrain nuclei forming a direct forebrain premotor pathway modulated by a recurrent basal ganglia loop seems to be conserved among avian vocal learners (Jarvis, 2004).

2.1.2 Vocal Motor System in Mammals

There is no data regarding the role of forebrain premotor structures in vocal production in rodents. In fact, there is quite little data on the functional organization of mouse motor cortex in general. Some information comes from an electrical micro-stimulation study, but this technique generated disjointed and overlapping representations with poor spatial resolution for broad muscle groups of the limbs and face (Pronichev and Lenkov, 1998). Motor control circuits have been mapped with good detail in the rat, especially the motor representations in the barrel cortex controlling whisker movements (Brecht et al., 2004). Additionally, forebrain contributions to vocal behavior of adult rats have been established for limbic structures connected to the dopaminergic mesolimbic system (Burgdorf et al., 2007; Brudzynski, 2009), and the infralimbic medial frontal cortex-PAG-Amb pathway (Figure 5) (Frysztak and Neafsey, 1991; Depaulis et al., 1992). Similar forebrain contributions to adult mouse USVs from either limbic or premotor systems, however, remain unstudied.

As previously described for humans, songbirds and parrots, an apparently invariant feature in vocal learning species are dedicated cortico-striatal-thalamic circuits for learning and producing vocalizations (Jarvis et al., 2000; Jarvis, 2004; Ludlow, 2005;

Jürgens, 2008; Fischer and Hammerschmidt, 2010). In humans, cortical, basal ganglia, and thalamic vocalization-related brain regions have typically been identified with functional neuroimaging techniques during speech production or brain lesion case studies (Jürgens, 2002; Ludlow, 2005). In contrast, vocalization-specific activity in vocal non-learning mammalian species has been demonstrated only in limbic, midbrain and brainstem circuits (Wild, 1997; Jürgens, 2002; Hage and Jürgens, 2006b; Jürgens, 2008). Electrical micro-stimulation of a specific motor cortical region in anesthetized macaques produces movement of the vocal folds (Hast et al., 1974) and tracing studies revealed extensive subcortical projections from this laryngeal premotor cortex to the basal ganglia, thalamus, pons and medulla (Simonyan and Jürgens, 2003); however, chemically inactivating these structures does not abolish vocal fold movements elicited by motor cortical stimulation (Jürgens and Ehrenreich, 2007), suggesting a different route to the laryngeal motoneurons. Moreover, lesions to prefrontal and primary motor cortex (Sutton et al., 1974; Aitken, 1981; Kirzinger and Jürgens, 1982) or globus pallidus (MacLean, 1978) do not produce changes in the structure of vocalizations in monkeys. Therefore, it is questionable that these structures play a role in the programming of monkey vocalization, and they may subserve other laryngeal functions in other behaviors like swallowing.

Recently, during preparation of this dissertation, two studies have claimed to find cortical activation during vocalization in marmosets (*Callithrix jacchus*) by

examining brain expression patterns of the activity-dependent immediate early genes egr-1 (Simões et al., 2010) and cFos (Miller et al., 2010). In the first study, expression levels of egr-1 were measured in prefrontal cortex of two groups of animals that heard playbacks of conspecific calls and either vocalized or remained silent (Simões et al., 2010). Higher numbers of egr-1 immunopositive cells were observed in ventral and dorsal prefrontal cortex when animals vocalized than when they remained silent. However, given the audio-motor nature of the task it is impossible to separate the relative effects of sensory processing and preparation of the motor program for vocalization. The second study attempted to distinguish between sensory, motor, and sensorimotor integration effects by including a treatment group that vocalized without hearing any conspecific playbacks (Miller et al., 2010). Strikingly, this production-only group showed the lowest amount of c-fos induction for the majority of prefrontal sites tested. The animals that showed the highest levels of induction overall were those that only heard playbacks of calls. Expression levels for the vocal production group only matched those for the vocal perception group in one area, but failed to show higher induction than the antiphonal calling group. Given the strong auditory effect observed and the relatively weak induction by vocalizing-only, it is entirely possible that the IEG expression levels observed in vocalizing groups from both studies is largely due to sensory processing of conspecific calls or the animals hearing themselves vocalize. Therefore, it remains to be determined if neocortical or basal ganglia regions associated

with vocal production in humans also participate in natural vocal production in non-human mammals.

2.2 Discovering the Mouse Song System with Behavioral Mapping of Activity-Dependent Immediate Early Genes

I used a behavioral molecular mapping technique to analyze global patterns of IEG induction to test whether forebrain regions are active during male mouse song production. The experimental design was borrowed from studies that identified seven remarkably similar forebrain song nuclei among the three independently evolved lineages of song learning birds (Jarvis and Nottebohm, 1997; Jarvis and Mello, 2000; Jarvis et al., 2000).

2.2.1 Methods – Behavioral Gene Mapping

2.2.1.1 Behavioral Conditions

Sexually experienced adult BxD males 50 – 87 days old were acclimated by placing them in a dark 15″x24″x12″ sound-attenuating recording chamber overnight. The following day males were presented with an olfactory or auditory stimulus after a period of 3 hrs with little movement and no USV production. The mice were divided into four treatment groups (n=5 per group):

1) Singing and Hearing group stimulated to sing by presentation of 200 μL of fresh urine from BALB/c females pipetted directly onto the bedding through a small covered opening on the top of the sound chamber.

2) Singing Only group similarly stimulated to sing by urine, but after undergoing bilateral cochlear removal.

3) Silent group stimulated to explore the home cage without singing by presentation of 200 μL of 10% EtOH.

4) Hearing Only group stimulated with playbacks of USVs recorded from a BxD male.

Olfactory stimuli were presented at 5 min intervals throughout the 30 min recording session to maintain exploratory and vocal behavior. USV playbacks were continuous for the 30 min session. The Silent group provided a baseline for the vocally stimulated groups. The Singing Only group allowed me to separate the effects of motor and auditory processing.

2.2.1.2 In Situ Hybridization

Immediately after the 30 min behavioral session, animals were sacrificed by decapitation without anesthesia because IEG expression can be sensitive to manipulation within 5-10 minutes of handling. Brains were removed, embedded in Tissue-Tek (Sakura Finetek), frozen on dry ice, then stored at -80°C. Coronal 12 μm sections were cut through the entire brain on a cryostat, and every other section was mounted on silanated slides in series of 10. Frozen sections were processed for *in situ* hybridization with a ^{35}S radioactively labeled riboprobe made from cDNAs for rat arc and mouse egr-1, then processed for emulsion autoradiography following a previously described protocol (Jarvis and Nottebohm, 1997). The probe for egr-1 was generated from PCR-amplified sequences of the pCMV-Sport6-egr-1 plasmid containing the full-length mouse egr-1 cDNA (3.1kb) insert from our own library (Pioneer Clone F6). The

probe for arc was generated from PCR-amplified samples of a previously published 1.5kb sequence of the rat arc cDNA (Ribeiro et al., 2007), prepared according to a previously described protocol (Jarvis and Nottebohm, 1997).

2.2.1.3 Detection and Quantification of IEG Expression

Serial coronal sections through the entire brain were assayed for behaviorally driven expression of the plasticity-related, activity-dependent IEGs arc (Bramham et al., 2009) and egr-1 (Knapska and Kaczmarek, 2004). Slides were placed under B-max X-ray film for 4 days, then dipped in photographic emulsion (NBT2, Kodak). Dipped slides were developed, counter-stained with cresyl violet, and then coverslipped with Permount mounting medium (Fisher Scientific, Pittsburgh, PA).

I visually scanned autoradiographs throughout the entire brain and found one region of the cortex and underlying striatum that showed differential IEG expression between singing and non-singing animals. Measurements of mRNA expression of both IEGs were made from photomicrographs in the anterior cingulate cortex (Cg), a region of motor cortex spanning the border between M1 and M2 (MC), somatosensory cortex (S1), primary auditory cortex (A1), and anterodorsal striatum (adSt) directly underlying MC. Regions of interest (ROIs) were defined in three serial sections for each brain area on inverted images in ImageJ (NIH, Bethesda). The mean pixel value was recorded for each region of interest (Mroi), three regions of each glass slide with no brain tissue (Mbkgnd), and control areas with no difference in the background-adjusted mean pixel

values across treatment groups (Mctrl). The control areas with no difference were: 1) the ventral striatum for cingulate cortex, motor cortex, somatosensory cortex, and anterodorsal striatum; and 2) the midbrain reticular nucleus for the auditory cortex (Figure 6). These values were used to calculate the expression score (ES) for each ROI as follows:

$$ESroi = \log10[(Mroi - Mbkgnd) / (Mctrl - Mbkgnd)]$$

Expression scores are log transformed in order to visualize comparable magnitudes for expression differences above and below silent control levels in the experimental animals.

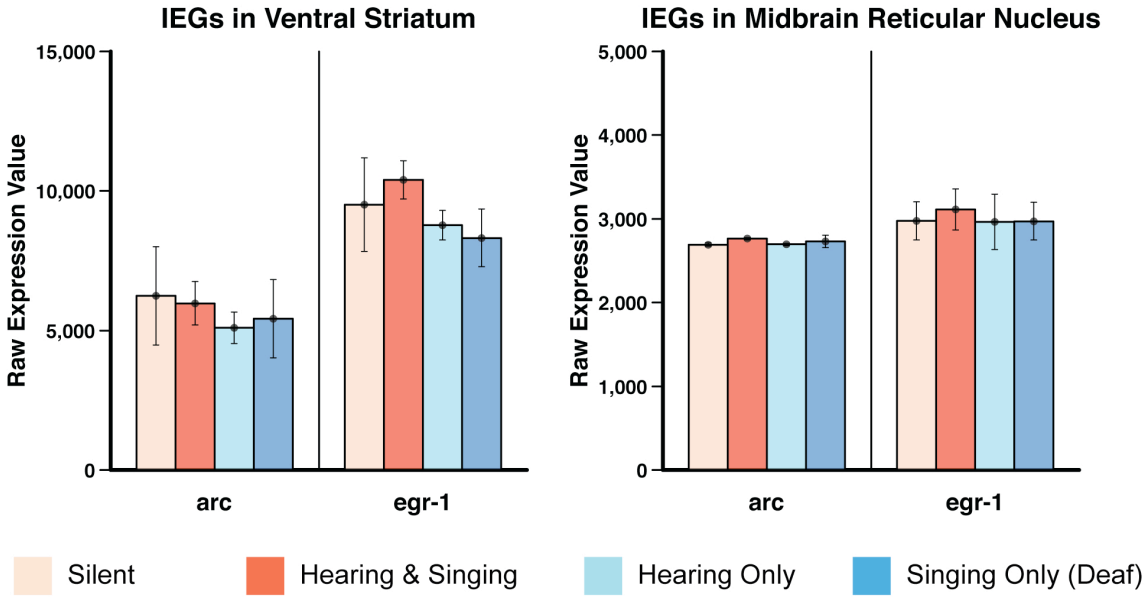

Figure 6: Raw expression measurements of arc and egr-1 in the ventral striatum and midbrain reticular nucleus, showing no mean differences among the four treatment groups (Kruskal-Wallis H-Test; n=5 per group; ventral striatum, egr-1: p=0.3, arc: p>0.5; midbrain reticular nucleus, egr-1: p>0.5, arc: p=0.070). These brain areas were used to normalize expression in other brain regions.

2.2.1.4 Song Recordings

Sounds were recorded with UltraSoundGate CM16/CMPA ultrasound microphones with a flat frequency response from 30-130 kHz and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). Detected sounds were digitized at 250 kHz, 8 bits and captured to disk as .WAV files using Avisoft Recorder USG (Avisoft Bioacoustics). Acoustic waveforms were processed using custom MATLAB-based programs (modified from code written by Timothy E. Holy, Washington University). The sonogram was computed from each waveform (256 samples/block, half-overlap), a sound intensity threshold was set to eliminate the white noise component of the signal, and frequencies outside 35-125 kHz were truncated. Syllables with duration longer than 10ms were identified by previously described methods (Holy and Guo, 2005).

2.2.1.5 Song Playbacks

A set of recorded songs from a BxD male was played at normal speed and pitch through an Avisoft ultrasound amplifier Model #70101 and speaker Model #60401 (Avisoft Bioacoustics). The same songs were used for all playback sessions. Each session was recorded and the syllable sequences were analyzed to confirm that USVs had not been produced by the listening mouse. One male sang a little spontaneously without urine stimulation and was excluded from analysis.

2.2.1.6 Video Recordings

Video recordings of mouse movements were made in the dark enclosed isolation chambers with a Speco Technologies VL-62 color IR camera. The video was saved to tape, then digitized on a Canopus ADVC110 analog-to-digital converter at full resolution, and stored to disk as MPEG video files. Digitized videos were coded for periods during which the animal was sitting still or moving throughout the home cage, blind to treatment condition, using the behavioral coding software Annotation by SaySoSoft (v1.0, http://www.saysosoft.com/). The total time during which locomotion or rotational movement was recorded was summed for each animal.

2.2.1.7 Deafening by Cochlear Removal

Five adult BxD males 77 to 87 days old were deafened by bilateral cochlear removal. The deafening surgery is described in the General Methods (Appendix A, Section A.4).

2.2.2 Results

Relative to the non-singing treatment groups, the Hearing & Singing and Singing Only vocalizing groups expressed higher levels of arc mRNA bilaterally in Cg, MC, and adSt regardless of whether or not they could hear themselves sing (Figure 7). Arc expression levels in these regions were relatively higher than more lateral parts of the cortex, suggesting a restricted increase. In A1, deafening resulted in a decrease of expression in the deeper cortical layers (Figure 7h).

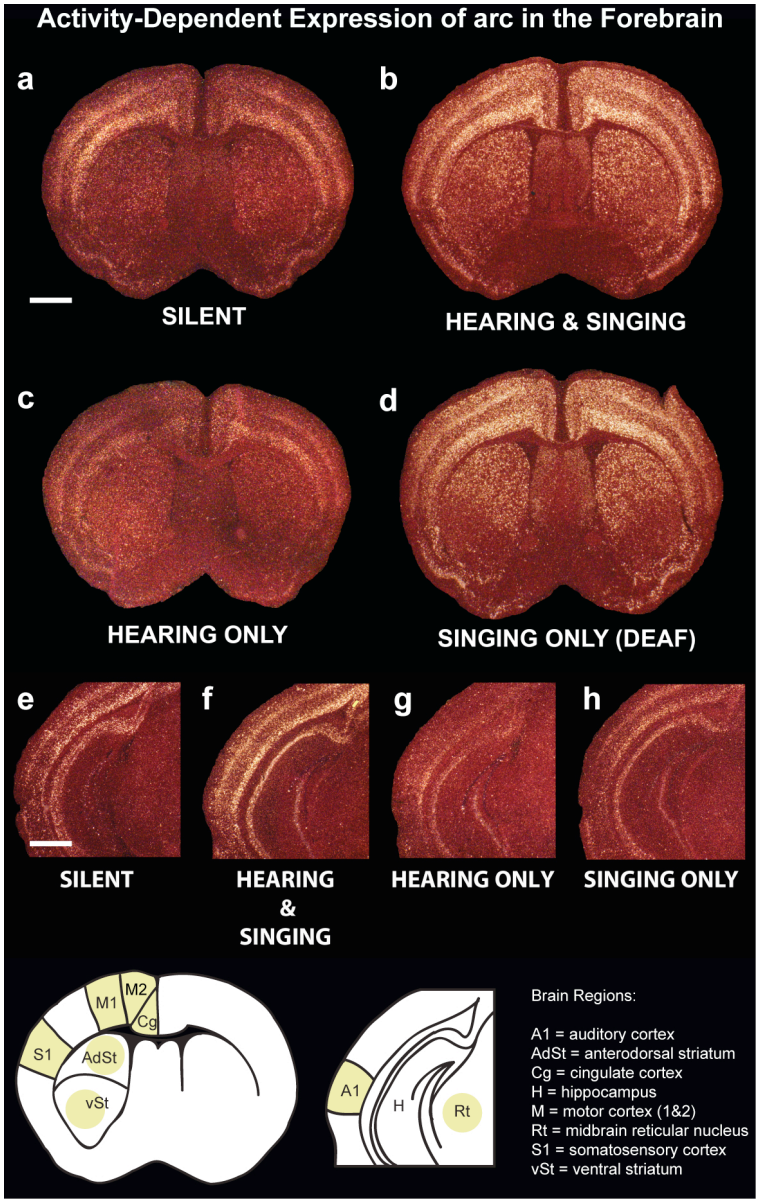

Figure 7: a-d, Dark-field images of cresyl violet stained (red) coronal brain sections at the level of motor cortex, approximately 0.2 mm rostral to Bregma, showing singing-induced arc expression (white) in the forebrain of male mice. e-h, Arc expression in the auditory cortex (one hemisphere) of animals from the same groups. Treatment groups were: a & e, Silent, did not sing; b & d, Hearing & Singing, 30 min of USVs with hearing intact; c & g, Hearing Only, heard playbacks of USVs for 30 min; d & h, Singing Only, 30 min of USVs after deafening. ROIs for each brain region are shaded in yellow. Scale bars = 1 mm.

At higher magnification it can be seen that the increased arc signal in MC is due to a higher total number arc expressing cells and higher expression levels within those cells (Figure 8).

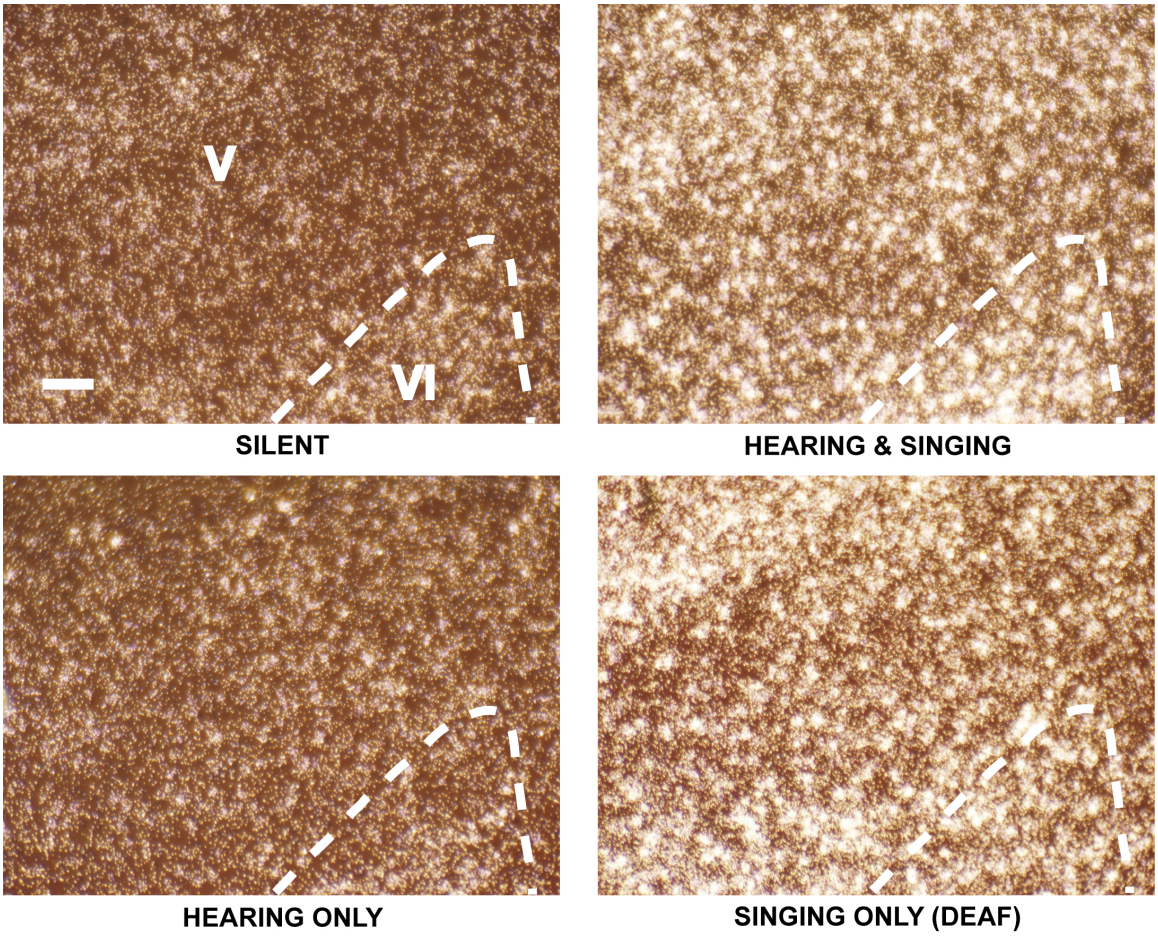

Figure 8: High magnification dark-field images near the border between MC and Cg in the same sections as in Figure 7. More cells are expressing high levels of singing-induced arc expression in cortical layers V & VI of animals producing USVs than of non-vocalizing animals. Scale bar = 100 μm.

The expression scores revealed significant mean differences in arc expression between treatment groups in Cg, MC, and adSt (Figure 9; Kruskal-Wallis H-tests; n=5 per group; p values reported on graph). For regions where significant group mean differences were detected, arc mRNA expression scores from each group were compared directly to the Silent group (Figure 9; Mann-Whitney U-tests; n=5 per group; Cg: Hearing Only, p=0.347; Hearing & Singing, p=0.009; Singing Only, p=0.016; MC: Hearing Only, p=0.076; Hearing & Singing, p=0.009; Singing Only, p=0.009; adSt: Hearing Only, p=0.458; Hearing & Singing, p=0.009; Singing Only, p=0.009). In all three regions the Hearing & Singing and Singing Only groups showed significantly higher arc expression scores than the Silent group. The corresponding relative fold change in arc expression was calculated from the mean background-corrected, normalized expression values in each region. Mean arc expression scores for the Hearing & Singing group were 0.34 ± 0.04 in Cg (1.5 fold increase), 0.36 ± 0.04 in MC (1.9 fold increase), and 0.29 ± 0.04 in adSt (2.4 fold increase). Mean arc expression scores for the Singing Only group were 0.31 ± 0.05 in Cg (1.4 fold increase), 0.32 ± 0.03 in MC (1.7 fold increase), and 0.18 ± 0.02 in adSt (2.1 fold increase). A similar amount of arc expression was observed for Hearing & Singing and Singing Only groups in Cg, MC, and adSt (Mann-Whitney U-tests; n=5 per group; Cg, p=0.347; MC, p=0.347; adSt, p=0.076), indicating high mRNA expression in this region was not caused by auditory processing. Arc expression scores did not differ between the Hearing Only and Silent groups in Cg, MC, adSt, or S1.

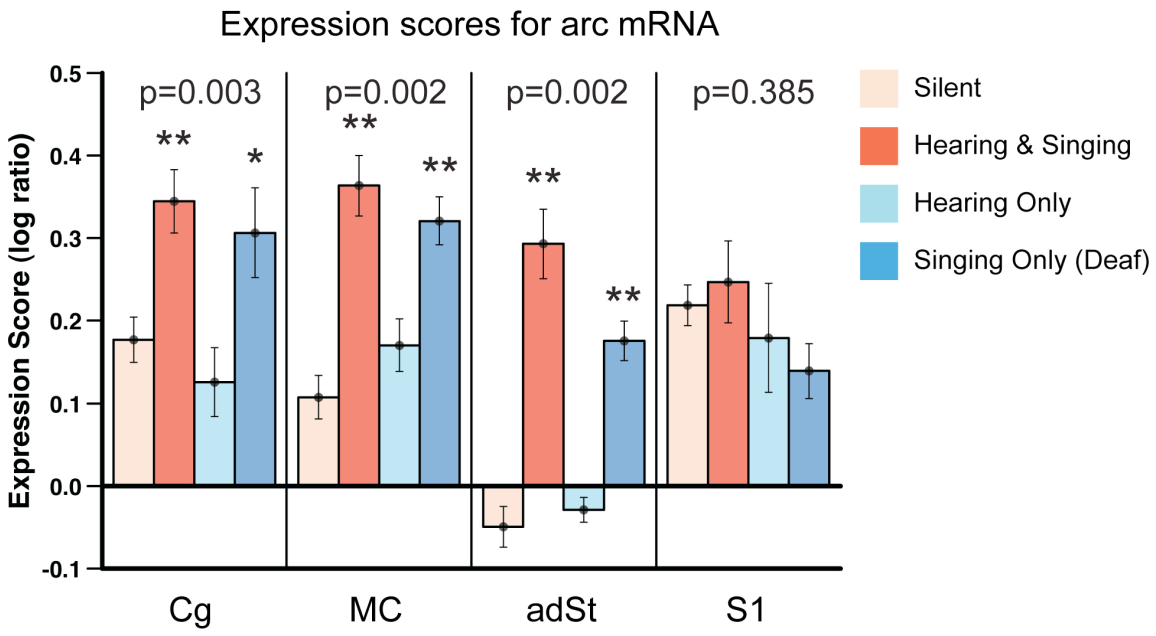

Figure 9: Arc expression scores in cingulate cortex (Cg), primary and secondary motor cortex (MC), anterodorsal striatum (adSt), and somatosensory cortex (S1) for all treatment groups. Kruskal-Wallis H-tests (n=5 per group; p values reported on graph) with post-hoc Mann-Whitney U-tests to test for differences relative to the Silent group (n =5 per group; * = p<0.05, ** = p<0.01).

Global patterns of arc and egr-1 mRNA in the forebrain were generally similar. Similar to arc, the Hearing & Singing and Singing Only groups showed high levels of egr-1 mRNA expression bilaterally in Cg, MC, and adSt (Figure 10). Unlike arc, however, the Silent group also showed high egr-1 expression levels in Cg. Visually, the expression contrast between MC and the laterally adjacent somatosensory cortex was more pronounced for egr-1, similar to differences between the two genes in songbirds (Wada et al., 2006).

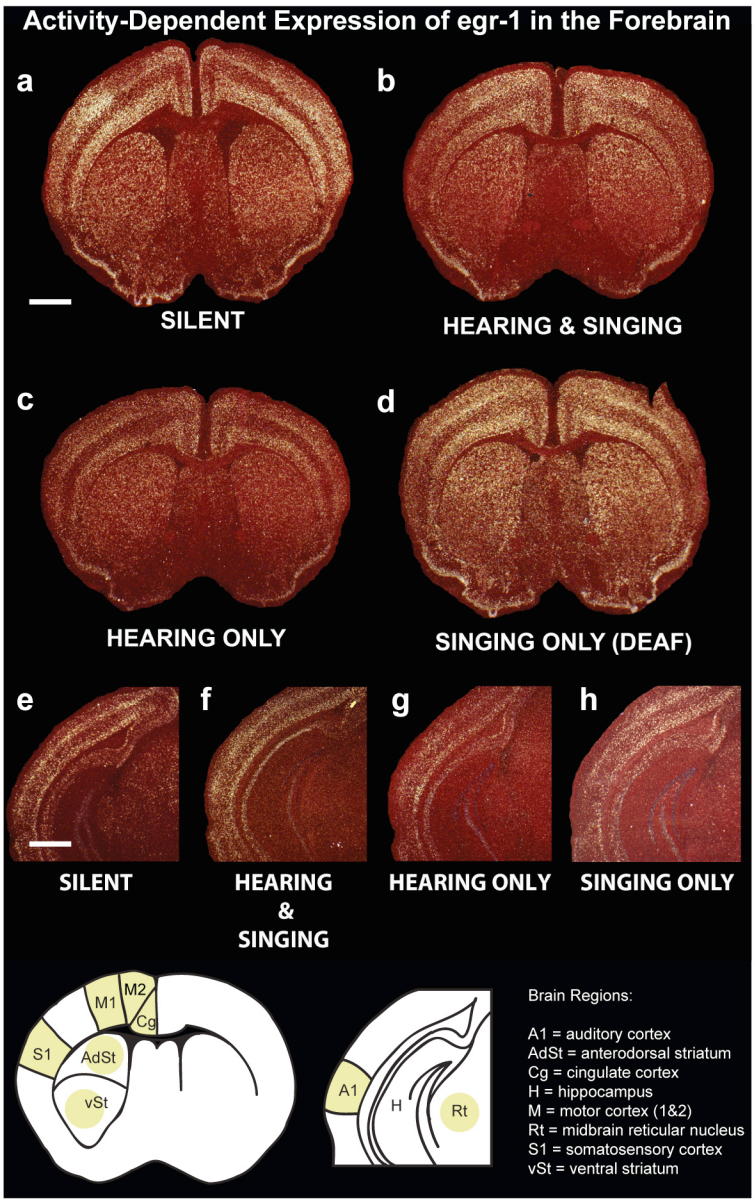

Figure 10: Dark-field images of coronal brain sections showing singing-induced egr-1 expression in the forebrain of male mice. Same treatment groups and sampling regions as in Figure 7. e-h, Auditory cortex of animals from the same groups. Scale bars = 1 mm.

Egr-1 expression scores revealed significant mean differences between treatment groups in Cg, MC, and adSt (Figure 11; Kruskall-Wallis H-tests; n=5 per group; p values reported on graph). There were no significant group mean differences in mRNA expression scores in S1 for either gene (Kruskal-Wallis H-tests; n=5 per group; arc, p=0.385; egr-1, p>0.5) confirming that the effect was not global but restricted to particular cortical regions. For those regions where group mean differences were detected, egr-1 mRNA expression scores for each treatment group were compared directly to the Silent group (Figure 11; Mann-Whitney U-tests; n=5 per group; Cg: Hearing Only, p=0.009; Hearing & Singing, p>0.5; Singing Only, p=0.465; MC: Hearing Only, p=0.251; Hearing & Singing, p=0.009; Singing Only, p=0.009; adST: Hearing Only, p=0.347; Hearing & Singing, p=0.009; Singing Only, p=0.009). Hearing & Singing and Singing Only groups only had significantly higher egr-1 expression scores than the Silent group in MC and adSt, but not Cg. Mean arc expression scores for the Hearing & Singing group were 0.14 ± 0.02 in MC (1.3 fold increase), and 0.11 ± 0.02 in adSt (1.3 fold increase). Mean arc expression scores for the Singing Only group were 0.18 ± 0.02 in MC (1.5 fold increase), and 0.10 ± 0.02 in adSt (1.2 fold increase). Because of high egr-1 expression levels in Cg for Silent animals, expression scores were significantly lower for the Hearing Only group in this region. Like arc, egr-1 displayed similar mean expression scores for the Hearing & Singing and Singing Only groups in Cg, MC, and adSt (Mann-Whitney U-tests; n=5 per group; Cg, p=0.347; MC, p=0.347; adST, p=0.076).

So, induction of mRNA for both genes appears to be independent of auditory processing in these forebrain regions.

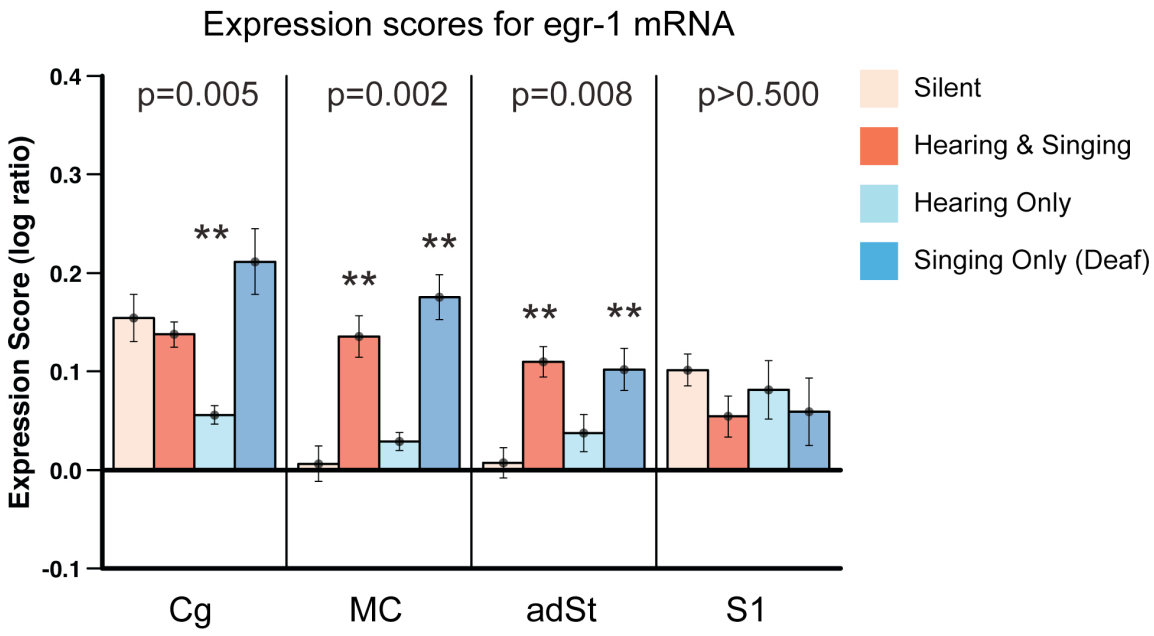

Figure 11: Egr-1 expression scores in cingulate cortex, motor cortex, anterodorsal striatum, and somatosensory cortex for all treatment groups. Kruskal-Wallis H-tests (n=5 per group; p values reported on graph) with post-hoc Mann-Whitney U-tests to test for differences relative to the Silent group (n =5 per group; * = p<0.05, ** = p<0.01).

General movement around the testing cage does not explain the differential IEG expression patterns, as there was no significant difference between treatment groups in the amount of locomotion (Figure 12; Kruskal-Wallis H-test; n=5 per group; p=0.295). Two mice in the Hearing Only group moved very little during testing; however, even those still mice showed IEG expression levels similar to Silent mice with high locomotor activity.

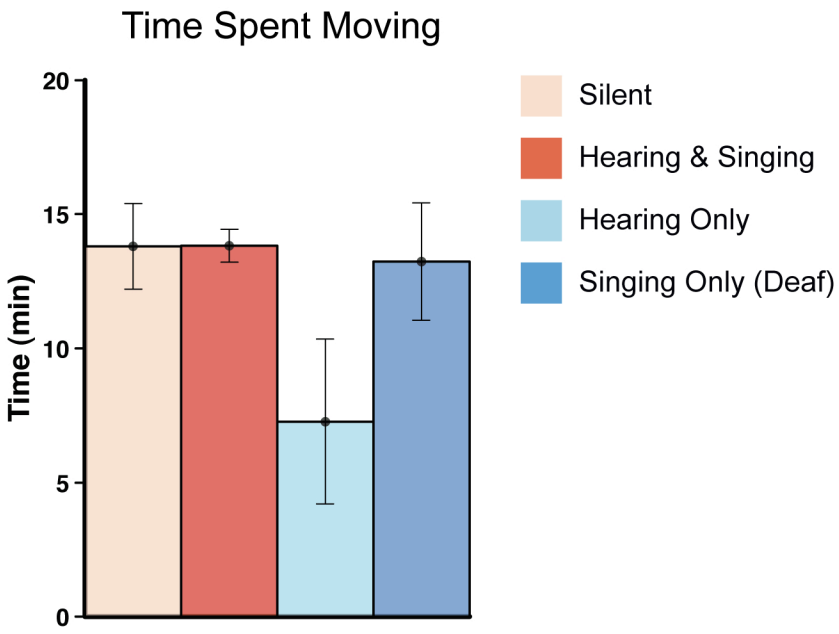

Figure 12: Total amount of locomotion during testing for each treatment group (Kruskal-Wallis H-test; n=5 per group; p=0.295).

Group mean differences in expression scores for both IEGs were also detected in A1 (Figure 13; Kruskal-Wallis H-tests; n=5 per group; p-values reported on graph). For both arc and egr-1, IEG expression scores for each treatment group were compared directly to the Singing Only (deaf) group (Figure 13; Mann-Whitney U-tests; n=5 per group; arc: Hearing Only, p=0.009; Hearing & Singing, p=0.009; Silent, p=0.009; egr-1: Hearing Only, p=0.009; Hearing & Singing, p=0.009; Silent, p=0.009). As was visually apparent, deafening resulted in significantly lower IEG expression scores for the Singing Only animals that could not hear themselves sing. This effect is in contrast to the failure of deafening to reduce high levels of IEG expression in singing-activated Cg, MC, or adSt. Egr-1 expression scores in A1 were similar across all three hearing-intact groups; however, arc expression scores for the Hearing Only group were significantly lower than the Silent and Hearing and Singing groups (Figure 13; Mann-Whitney U-tests; n=5 per group; arc: Silent vs. Hearing & Singing, p=0.117; Hearing & Singing vs. Hearing Only, p=0.009; Hearing Only vs. Silent, p=0.28; egr-1: Silent vs. Hearing & Singing, p>0.5; Hearing & Singing vs. Hearing Only, p=0.347; Hearing Only vs. Silent, p=377;).

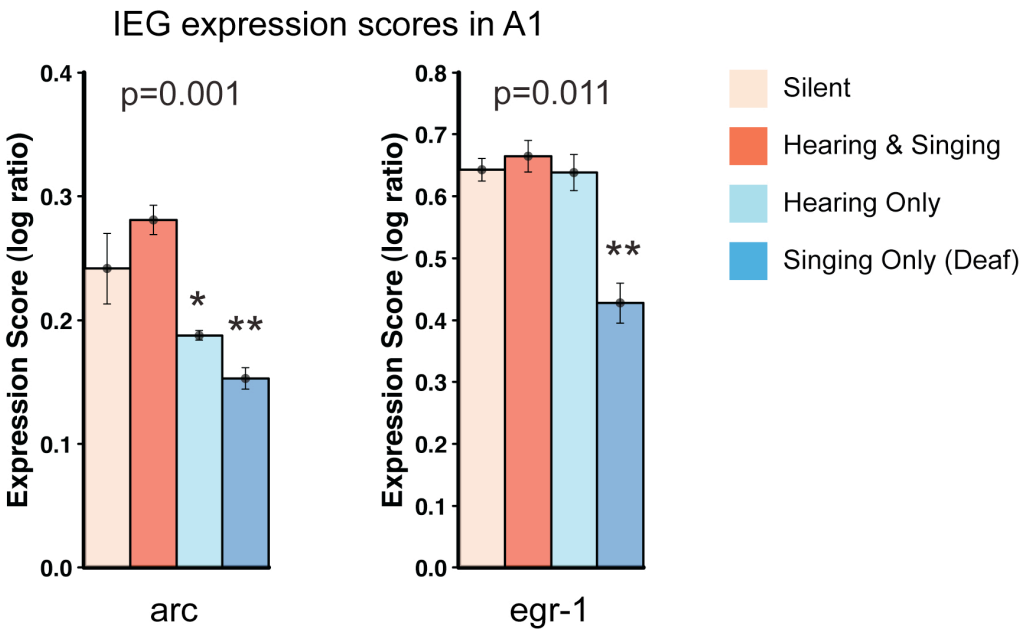

Figure 13: Expression scores for arc and egr-1 in auditory cortex for all treatment groups. Kruskal-Wallis H-tests (n=5 per group; p-values reported on graph) with post-hoc Mann-Whitney U-tests to test for differences relative to the Singing Only (deaf) group (n =5 per group; black * = p<0.05, ** = p<0.01) or relative to the Hearing Only group (n=5 per group; red * = p<0.05, ** = p<0.01).

The results of this behavioral gene mapping study show that production of ultrasonic courtship songs increases motor-driven expression of two activity-dependent IEGs in motor cortical, limbic, and striatal regions of the mouse brain. This is the first time vocal motor specific activity has been shown in the forebrain of a non-human mammal. Moreover, the IEG expression pattern was reminiscent of the results obtained by similar techniques in the brains of vocal learning birds (Jarvis and Nottebohm, 1997; Jarvis and Mello, 2000; Jarvis et al., 2000)

Chapter 3. Cortical Projections to Phonatory Motoneurons

Despite IEG activation patterns suggesting a motor cortical contribution to singing, it remained uncertain if motor cortex actually provides input to the vocal system at any level. Therefore, I first attempted to determine if motor cortex is part of a vocal premotor pathway by identifying the circuits connected to the larynx through Amb. Previous studies of neural control of vocalization in non-human primates have attempted unsuccessfully to identify a direct cortical projection to the laryngeal motoneurons in Amb (Kuypers, 1958b; Simonyan and Jürgens, 2003). The direct motor cortical vocal pathway, uniquely found in humans among mammals (Kuypers, 1958c; Iwatsubo et al., 1990), has been given particular attention by researchers studying the evolution of brain systems for vocalization:

> The most important derived feature in the human lineage regarding the ontogeny of speech appears to be the evolution of the direct pathway from the motor cortex to the motoneurons, enabling volitional control over the oscillations of the vocal folds. Together with the intricate coordination of breathing and articulation, this feature allows for the precise control over speech production. The role of the basal ganglia in the modulation of vocal behavior, in contrast, appears to be an ancestral feature. (Fischer and Hammerschmidt, 2010)

Also,

> I argue that a mutational event that caused descending projections of avian arcopallium neurons to synapse onto nXIIts or mammalian layer 5 neurons of the face motor cortex to synapse onto nucleus ambiguous [*sic*] may be the only major change that is needed to initiate a vocal learning pathway. (Jarvis, 2004)

Also,

Are there any specific anatomical substrates that correlated with the faculty of vocal learning?

One of the candidates for this question is the direct cortical-medullar pathway for articulation and breathing. In humans, a part of motor cortex directly projects to the medullary nuclei, the nucleus ambiguus, and the nucleus retro-ambiguus. This projection was absent in the squirrel monkey and Jürgens assumes that this projection exists only in humans. Similarly, there is a direct cortical-medullar pathway for articulation and breathing in the zebra finch, a species of songbirds, but a similar projection in pigeons that do not learn to sing [*sic*]. Thus, this projection exists in the species that show vocal learning while it is absent in the species without vocal learning. (Okanoya, 2004)

Also,

In humans, in contrast, direct connections of the motor cortex with the nucl. ambiguus do exist. This direct cortico-ambigual connection seems to be a very recent acquisition in hominid evolution and possibly represents one of the prerequisites for speech development. (Simonyan and Jürgens, 2003)

Also,

This discrepancy between man and monkey finds its anatomical correlate in the fact that there is a direct projection from the cortical larynx area to the laryngeal motoneurones in man but not in monkey. The human cortex thus has a more direct control on phonation than that of the monkey. This more direct control is paralleled by an advanced capability of voluntary control of fine vocal fold movements which the monkey lacks almost completely. (Kirzinger and Jürgens, 1982)

Also,

The discrepancy between human aphonia and intact monkey phonation, on the one hand, and intact phonation and defective orofacial behaviour in the monkey, on the other, is explained by the anatomical fact that there is a direct cortical projection to the nucleus ambiguus in man but not in monkey, whereas the facial and hypoglossal nuclei receive direct cortical projections in man and monkey. The lack of direct cortical control of the laryngeal motoneurones in the monkey is paralleled by a lack of volitional control of fine vocal fold movements. (Jürgens et al., 1982)

The clear underlying assumption connecting these statements is that the unique vocal learning capabilities of songbirds and humans derive from improved vocal control as a result of direct cortical control of the vocal apparatus. That is, reorganization of neural architecture was the basis and precondition for the reorganization of function. Of course, this theory is based on correlative evidence and should be approached with some caution because the neuroanatomical differences could be ancillary (Deacon, 2000). Despite this caveat, the basic assumptions have been generally accepted. Therefore, the presence of a direct motor cortical vocal pathway would greatly increase the utility and acceptance of an animal model for vocalization. Even a rudimentary projection from motor cortex to Amb could be leveraged to test various hypotheses regarding the genetic specification of this pathway and its relevance to the evolution of vocal learning.

3.1 Background - Direct Projections to Brainstem Vocal Motor Nuclei

To date, projections from motor cortex to phonatory brainstem nuclei have only been found in primates. In a comparative study of projections from the motor cortical tongue area to the hypoglossal nucleus (XII) that innervates the tongue muscles, it was observed that the strength of the projection varies between primate species (Jürgens and Alipour, 2002). Rhesus macaques have a relatively denser projection than squirrel monkeys, and saddle-back tamarins (*Saguinus fuscicollis*) have putative fibers of passage but no terminals in XII. In chimpanzees (*Pan troglodytes*) (Kuypers, 1958b) and humans (Kuypers, 1958c) projections to XII are strong. By contrast, no motor cortical projection

to XII was observed in tree shrews (*Tupaia belangeri*) (Jürgens and Alipour, 2002), cats (*Felis catus*) (Kuypers, 1958a), or rats (Travers and Norgren, 1983). Combined with the direct motor cortical vocal pathway to Amb in humans, this distribution of cortico-bulbar projections has been interpreted as a progressive increase in cortical innervation in phylogenetically newer primate species leading to improved vocal abilities:

> As the direct connection is heavier in man than in the rhesus monkey, heavier in the rhesus monkey than in the squirrel monkey (present study) and 'heavier' in the squirrel monkey than in the tree shrew (where it is non-existent), this may be interpreted as a phylogenetic trend of strengthening the cortico-hypoglossal connections towards man. This strengthening might be one of the factors responsible for the greater role tongue movements play in human vocal communication than in other mammals' communication. (Jürgens and Alipour, 2002)

This interpretation reflects the previously addressed general assumption that presence or absence of direct cortical input to phonatory motor nuclei largely determines the level of vocal abilities. Though untested, this basic assumption supports the framework in which either the presence or absence of direct cortico-bulbar projections in the mouse will be interpreted.

3.1.1 Direct Cortico-bulbar Projections in Rodents

The only demonstration of a direct motor cortical projection to cranial motor nuclei outside of primates was in the whisker movement system of rats (Grinevich et al., 2005). Rodents generate high-speed rhythmic whisker movements and possess a large cortical motor field dedicated to controlling the muscles that move the whiskers, the vibrissa motor cortex (VMC). Given the large representation of whisker muscles (45% of

rat primary motor cortex) and the importance of precise whisker movements for exploratory behavior, the authors hypothesized a direct projection from VMC to the facial motoneurons innervating whisker muscles. They imaged axons from VMC projection neurons expressing the lentivirally delivered gene for green fluorescent protein (GFP) innervating the facial nucleus (VII) and confirmed the presence of presynaptic terminals with electron microscopy (Grinevich et al., 2005). The vibrissa motoneurons were not specifically identified in this study. Despite this limitation, this was the first demonstration of a direct cortico-bulbar projection outside of primates and was motivated by the general assumption that direct motor cortical projections to motoneurons confer fine motor skills.

3.2 Discovering a Direct Cortical Projection to Nucleus Ambiguus in the Mouse

Mice have been assumed to lack a direct cortico-bulbar projection to Amb (Jarvis, 2004; Fischer and Hammerschmidt, 2010), but this has never been experimentally tested. Despite IEG activation patterns suggesting a motor cortical contribution to singing (see Section 2.2.2), it remained uncertain if motor cortex actually provides input to the vocal system at any level. Therefore, I attempted to test the hypothesis that motor cortex is part of a vocal premotor pathway by injecting the retrograde trans-synaptic tracer pseudorabies virus peripherally in the larynx of BxD males and tracing back through premotor brain pathways that converge on Amb. A recombinant pseudorabies virus strain (PRV-Bartha) expressing enhanced green fluorescent protein (eGFP) was injected

into the cricothyroid (CT) and lateral cricoarytenoid (LCA) muscles, two of the seven laryngeal muscles. PRV-Bartha only crosses functional synapses retrogradely through the sequence of synaptic connections away from the infection site (Aston-Jones and Card, 2000).

3.2.1 Methods - Retrograde Tracing of Premotor Circuits with PRV-Bartha

3.2.1.1 PRV-Bartha Injections into Laryngeal Muscles

Live virus was received from the laboratory of Dr. Lynn Enquist at Princeton University at a titer of 1×10^9 pfu/mL, aliquoted at 4 μL per tube, then stored at -80°C, and thawed immediately before injection. A total of 1 μL of PRV-Bartha was injected into the CT and LCA muscles using a Nanofil microsyringe system with a 34 gauge stainless steel needle (World Precision Instruments, Sarasota, FL). Details of the surgery are described in the General Methods (Appendix A, Section A.5).

3.2.1.2 Preparation of Brain Sections

PRV-Bartha-injected mice were sacrificed at approximately 48 (n=1), 72 hours (n=2), 96 hrs (n=8), 125 hrs (n=1), 135 hrs (n=1) or 145 hrs (n=1) post-innoculation by overdose of sodium pentobarbital, and then perfused transcardially with 0.9% saline followed by 4% paraformaldehyde in 0.1M PBS. Brains were removed, post-fixed in 4% paraformaldehyde overnight, and cryoprotected in 30% sucrose in PB until sectioned. Coronal sections were cut at 40 μm on a cryostat into 0.1 M PBS.

3.2.1.3 Immunochemical Detection of eGFP

Free-floating sections were quenched 30 min 0.3% H_2O_2 in PBS, then blocked 30 min in 0.3% PBST with NGS (VECTASTAIN Elite Kit, Vector Labs). Blocked sections were reacted for 3.5 hrs at RT in rabbit anti-eGFP (1:1000 dilution, Open Biosystems). Washed sections were incubated 1 hr in goat anti-rabbit biotinylated secondary antibody (VECTASTAIN Elite Kit) or 2 hrs at RT in donkey anti-rabbit secondary antibody conjugated to Alexa-fluor 488 (1:500 dilution, Invitrogen). Washed sections were reacted 1 hr in ABC solution (VECTASTAIN Elite Kit). Sections were then washed 3 times for 10 min in PB, and developed for 15 min in 0.05% DAB solution (Sigma, #D5905). DAB-stained sections were mounted on SuperFrost Plus slides (Fisher Scientific, Pittsburgh, PA) and dehydrated through a series of alcohol steps for 5 min each: 70%, 95%, 100%, and 100%. Dehydrated sections were cleared through two xylene washes and coverslipped with Permount mounting medium (Fisher Scientific). Sections to be imaged directly for eGFP were mounted on SuperFrost Plus slides and coverslipped with VectaShield (Vector Labs) fluorescent mounting medium.

Unless otherwise noted washes were 3 times for 5 min in PBS and reaction buffer was 0.3% PBST for all reactions.

3.2.1.4 Immunochemical Detection of Choline Acetyltransferase

For choline acetyltransferase (ChAT) detection in brainstem motoneurons, free-floating sections were blocked 1 hr in 0.3% PBST with NDS (VECTASTAIN Elite Kit).

Blocked sections were incubated overnight at 4°C in goat anti-ChAT (1:100 dilution, Millipore). Washed sections were incubated 2 hrs at RT in donkey anti-goat secondary antibody conjugated to Alexa-fluor 568 (1:500 dilution, Invitrogen). Washed sections were mounted on SuperFrost Plus slides and coverslipped with VectaShield. All washes were 3 times for 5 min in PBS and reaction buffer was 0.3% PBST for all reactions.

3.2.2 Results

No eGFP was detected at 48 hours post-injection. Beginning at 72 hours after a PRV-Bartha injection into the left CT and LCA muscles, eGFP was being lightly expressed in left Amb (Figure 14). At this survival time, expression of eGFP in the brainstem was restricted to Amb ipsilateral to the injections.

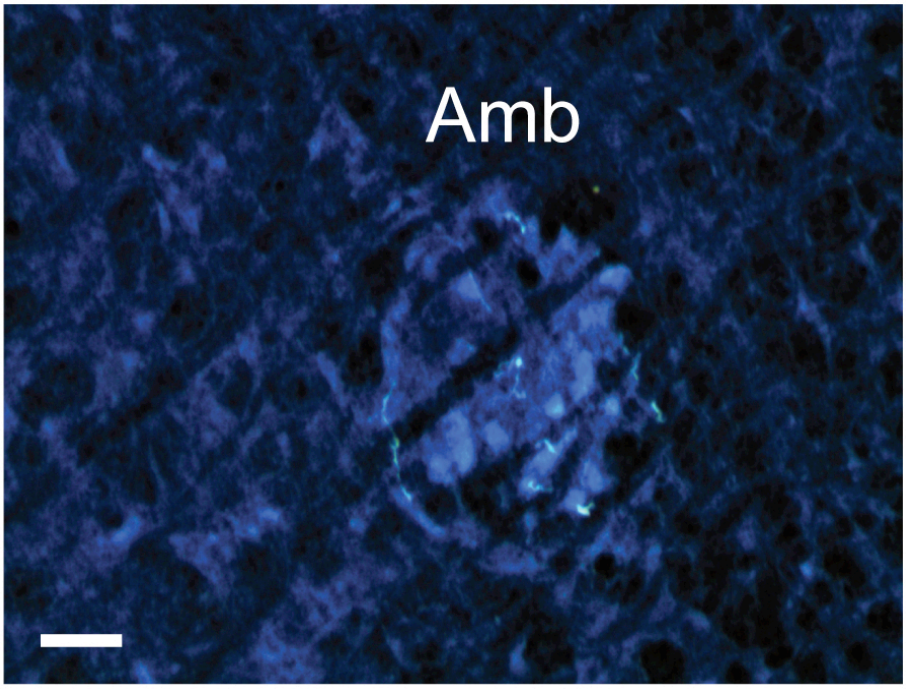

Figure 14: Neurons infected with pseudorabies virus (PRV-Bartha) expressing enhanced green fluorescent protein (whitish blue) in the left nucleus ambiguus (Amb) 72 hours after injection of PRV-Bartha into the left cricothyroid and lateral cricoarytenoid laryngeal muscles. Scale bar = 50 μm.

By approximately 96 hrs post-injection, I observed a pattern of labeling consistent with known connectivity in rodents (Van Daele and Cassell, 2009), and mammals in general (Jürgens, 2002). Robustly labeled eGFP-expressing neurons were found in the laryngeal motor center Amb (Figure 15a) but not in XII that innervates the tongue (Figure 15b).

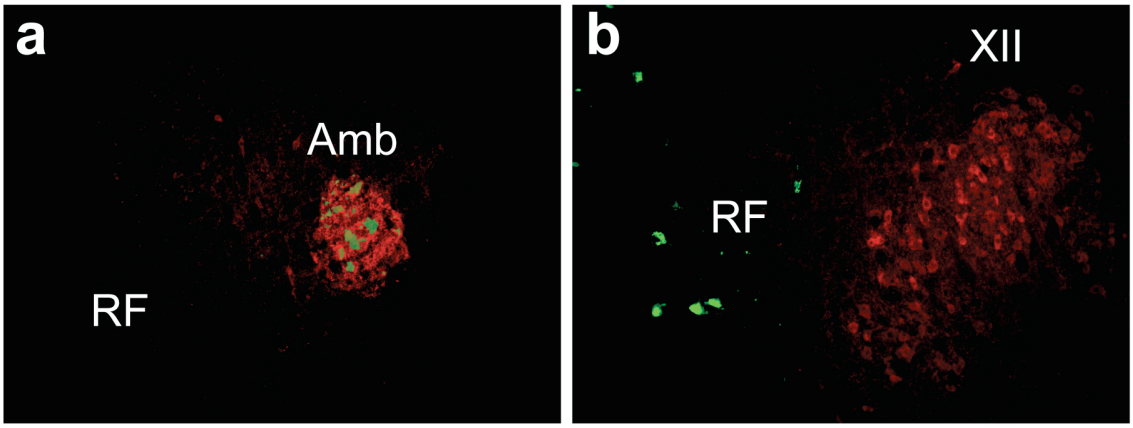

Figure 15: Neurons infected with pseudorabies virus (PRV-Bartha) expressing enhanced green fluorescent protein (green) and choline acetyltransferase immunopositive motoneurons (red) in the (a) ambiguus (Amb) and (b) hypoglossal (XII) nuclei of the mouse brainstem 94 hours after injection of PRV-Bartha into the cricothyroid and lateral cricoarytenoid laryngeal muscles.

By 96 hours post-injection, PRV-Bartha had spread retrogradely from Amb to other parts of the brain (Figure 16). Cells expressing eGFP were detected bilaterally in a set of regions in the midbrain and limbic system with known roles in the control of innate species-specific calls and respiration in mammals (Jürgens, 2002). The main regions labeled by PRV included: the medullary reticular formation (RF), PBr, spinal trigeminal nucleus (SP), and solitary nucleus (Sol) of the brainstem; periaqueductal grey (PAG) and ventral tegmentum of the midbrain; throughout the hypothalamus (HT); and the amigdalopyriform transition area (Apir), and central amygdala (CeA). Even at the longest survival times, eGFP-positive cells were not observed in the hypoglossal (XII) or facial (VII) motor nuclei, confirming that the only phonatory motoneuron pool labeled was in the ipsilateral Amb (Figure 15a).

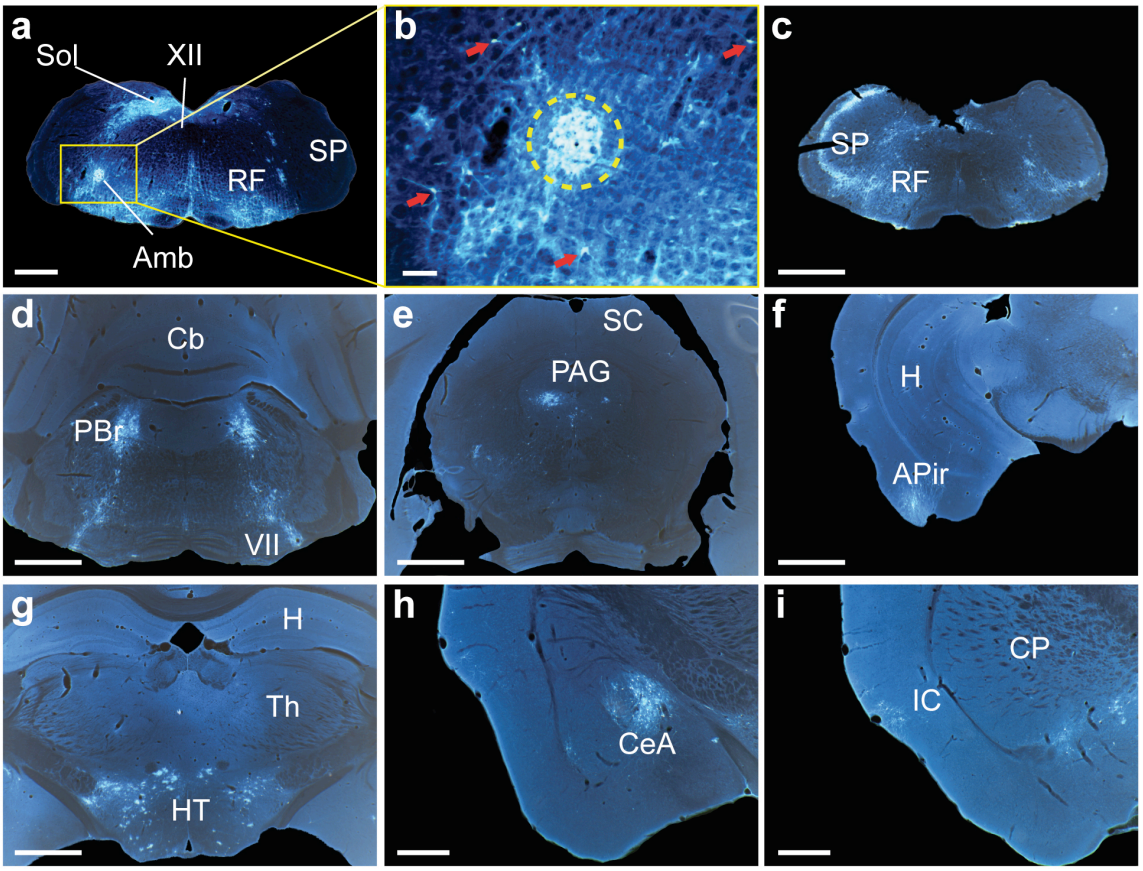

Figure 16: a-i, Psuedorabies virus (PRV-Bartha) infection in the brainstem, midbrain, and cortex in one mouse 86 hrs after injecting PRV-Bartha into the cricothyroid and lateral cricoarytenoid laryngeal muscles. a, Neurons expressing enhanced green fluorescent protein (white; colors inverted from original brightfield images) in nucleus ambiguus (Amb), the surrounding reticular formation (RF) and the solitary nucleus (Sol), but not the hypoglossal nucleus (XII). b, Amb and reticular interneurons at higher magnification (red arrows, cell bodies in RF). Scale bars: 1 mm for a; 100 μm for b; 1 mm for c-g; 50 μm for h-i.

I scanned for PRV label throughout the cortex and found that only two regions were reliably labeled at approximately 96 hours post-injection: 1) a population of layer V pyramidal neurons in M1 within the motor cortex region that exhibited robust singing-driven IEG expression (Figure 17); and 2) neurons in a small region of the insular cortex (IC) (Figure 16i). The M2 and Cg regions where singing-related IEG expression was observed were not labeled. In M1, the somata of labeled pyramidal cells were restricted to cortical layer V. These cells formed a continuous cluster centered 1.2 mm lateral to the midline, and extending rostro-caudally from approximately 0.2 mm posterior to 0.8 mm anterior to Bregma in coronal sections. Cells located laterally in IC were clustered in layer II/III. Although I did not quantify it, I did not note an obvious difference in IEG expression in IC in the behavioral gene mapping study. Therefore, it is unclear if this region of insular cortex is functionally active during song production. Both the M1 and IC regions were labeled in all mice with a survival time longer than approximately 90 hours.

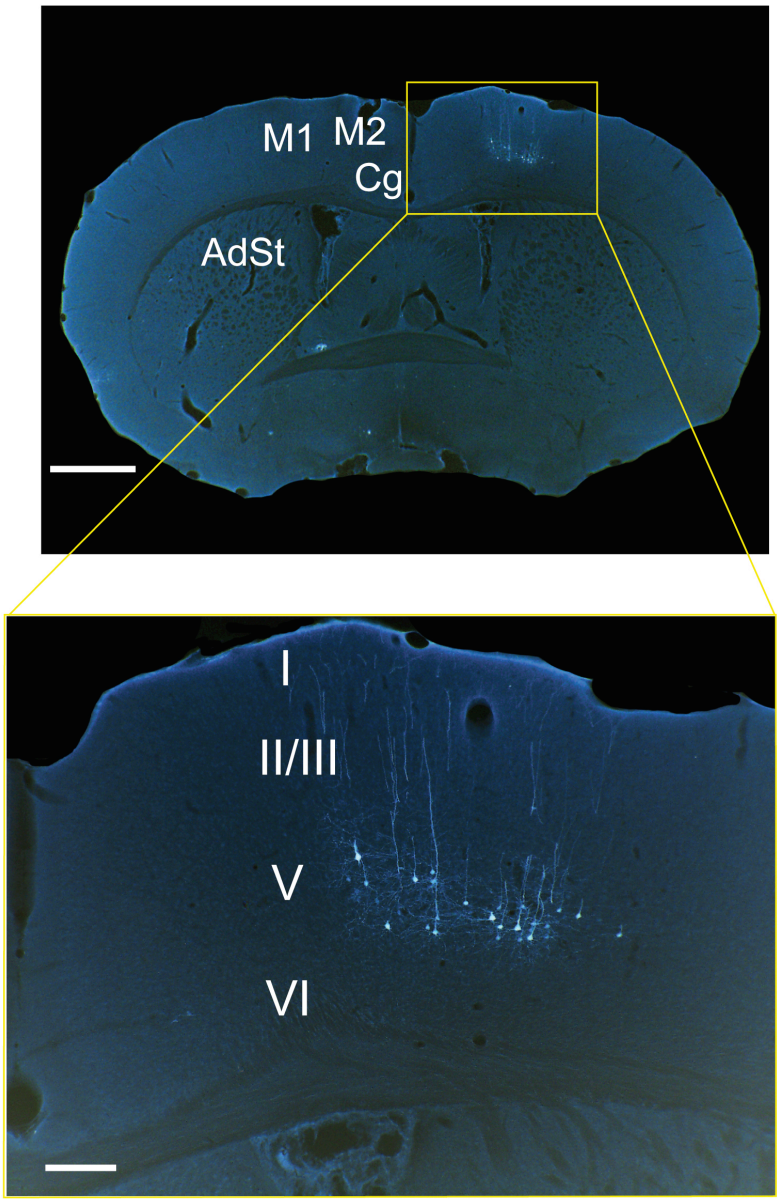

Figure 17: Pyramidal neurons expressing enhanced green fluorescent protein in cortical layer V of M1 following injection of psuedorabies virus (PRV-Bartha) into the cricothyroid and lateral cricoarytenoid laryngeal muscles. Labeled cells were not observed in cingulate cortex (Cg) or anterodorsal striatum (adSt). Scale bars: 1 mm for main panel; 200 μm for inset.

3.3 Tracing Motor Cortex Projections with BDA

The highly specific labeling of a region of motor cortex previously shown to be activated by USV production suggests that the second vocal pathway proposed by Jürgens (Jürgens, 2008), originating in the cortex and bypassing the limbic system and PAG, may also be present in mice. The multisynaptic nature of PRV tracing made it difficult to determine at what level the cortical pathway and limbic-PAG pathways converge. It is possible that they converge in the medullary reticular formation, as in monkeys (Jürgens and Ehrenreich, 2007; Jürgens, 2008); however, the relatively short latency at which label was observed in M1 suggested that perhaps they converge at an even lower level via a direct cortical projection to Amb. I used biotinylated dextran amines (BDA) to label motor cortical axons, and cholera toxin subunit b (CTb) to identify laryngeal motoneurons.

Immunochemical detection of CTb with peroxidase and DAB was chosen to label the motoneurons because it is very effective at revealing cellular processes including distal dendrites (Dederen et al., 1994). Multiple studies have used CTb to retrogradely label laryngeal motoneurons in songbirds (Wild, 1993), rhesus macaques (Vanderhorst et al., 2001), cats (Yoshida et al., 1998), and rats (Altschuler et al., 1991; Hayakawa et al., 2000; Zhang et al., 2005). In the rat CTb was injected into the CT muscle to retrogradely label motoneurons and revealed a large dendritic field extending far beyond the boundaries of Amb (Altschuler et al., 1991). The cell bodies were confined to the rostral

semicompact formation of Amb, but the dendrites of labeled cells extended ventrolaterally and dorsomedially into the adjacent RF forming a field bounded rostrally by VII and extending mostly ventrally and caudally in the sagittal plane, and extending more than 1mm medially in the horizontal plane. Therefore, a direct projection could reach the dendrites of laryngeal motoneurons without extending all the way to Amb.

To test the hypothesis that the cortical input from M1 to Amb is direct I injected BDA into the region of cortex where I observed singing-induced IEG expression (Section 2.2.2), and PRV back-traced neurons (Section 3.2.2). A subset of the BDA-injected mice was also injected with CTb into the CT and LCA laryngeal muscles. This dual tracing technique allowed me to visualize motor cortical axons as well as laryngeal motoneuron somata and dendrites from the same animals.

3.3.1 Methods – Anterograde Tracing of Cortico-fugal Projections with Biotinylated Dextran Amines

3.3.1.1 Tracer Injections in Motor Cortex

Adult BxD males greater 61 – 213 days old received tracer injections targeted to M1. General anesthesia was induced with 1% isofluorane and maintained by intramuscular injection of ketamine-xylazine (75 mg/kg ketamine; 5 mg/kg xylazine). The scalp was retracted and a small craniotomy was made over the injection site. Injections of 7.5% BDA (n=6, bilateral; n=1, right unilateral) were made through a glass micropipette using the Nanoject II microinjector (Drummond Scientific) at 4 sites 550μm from the brain surface (55 - 100 nL per site) along a track 1.2mm lateral and -0.2 to

0.4mm anterior to Bregma. Two mice received iontophoretic delivery of BDA in the same region using a 7.5 μA current for 10 – 20 min with a 5 sec half-duty cycle. After injections were complete the wound was closed with TissueMend.

3.3.1.2 Tracer Injections in Laryngeal Muscles

Six to twelve days after BDA injections, 1% CTb in saline was injected into CT and LCA muscles (n=3). Anesthesia was induced as described above. A midline incision was made from the sternum to the hyoid bone, and the portion of the sternohyoid muscle covering the larynx was removed. Five 100 nL injections were made 1 min apart into CT using a Nanofil microsyringe system with a 34 gauge stainless steel needle. After 5 min, the microinjection pipette was retracted, and the injection was repeated for LCA. A single break in the fascia was made for each muscle. After all injections were complete the wound was closed with TissueMend. Survival time after CTb injections was two to three days.

3.3.1.3 Visualizing BDA

For labeling of Biotinylated Dextran Amines (BDA) tracer, free-floating sections were quenched 30 min 0.3% H_2O_2 in PBS. Washed sections were reacted 1 hr in ABC solution (VECTASTAIN Elite Kit). Sections were then washed 3 times for 10 min in PB, and developed for 15 min in 0.05% DAB solution with nickel (DAB Substrate Kit, Vector Labs) to give a black reaction product. DAB-stained sections were mounted on SuperFrost Plus slides and dehydrated through a series of alcohol steps for 5 min each:

70%, 95%, 100%, and 100%. Dehydrated sections were cleared through two xylene washes and coverslipped with Permount.

3.3.1.4 Immunochemical Detection of CTb

For double labeling with Cholera Toxin Subunit b (CTb), following BDA detection, free-floating sections were blocked 1 hr in 0.3% PBST with NRS (VECTASTAIN Elite Kit). Blocked sections were incubated 2 hrs at RT in goat anti-CTb (1:10000 dilution, List Biological Laboratories). Washed sections were incubated 1 hr in rabbit anti-goat biotinylated secondary antibody (VECTASTAIN Elite Kit). Washed sections were reacted 30 min in ABC solution. DAB staining was as for BDA (Section 3.3.1.3), but for 3 min without nickel to give a brown reaction product. Sections were dehydrated, cleared, and coverslipped as for BDA (Section 3.3.1.3).

3.3.2 Results

I found that the singing-activated portion of M1 projects densely to adSt (Figure 18a), in the same region that displayed singing-related increase of IEG expression (Figures 7 and 10). Unilateral BDA injection revealed a cortical projection from M1 to the contralateral M1 via the corpus callosum (CC) (Figure 19). I also observed dense terminal fields in adSt bilaterally; however, the contralateral projection could be axon collaterals from retrogradely labeled neurons in contralateral M1. Cortical input to the M1 injection site was received from the ventral secondary auditory cortex (Av), where a number of BDA-positive cell bodies where observed in layers 2/3 (Figure 18c). M1 sent a

dense ipsilateral projection to the ventral lateral nucleus of the thalamus (VL) via the internal capsule (int) (Figures 18b & 19). Along the dorsolateral edge of VL a tight cluster of cortically projecting cells was retrogradely labeled, indicating a reciprocal connection between VL and M1. These two projections are likely to form part of a cortico-striatal-thalamic loop for vocalization similar to those reported in humans and song learning birds (Figure 5); however, we have not confirmed the striatal projection to globus pallidus or the pallidal projection to thalamus for this vocal circuit in mice.

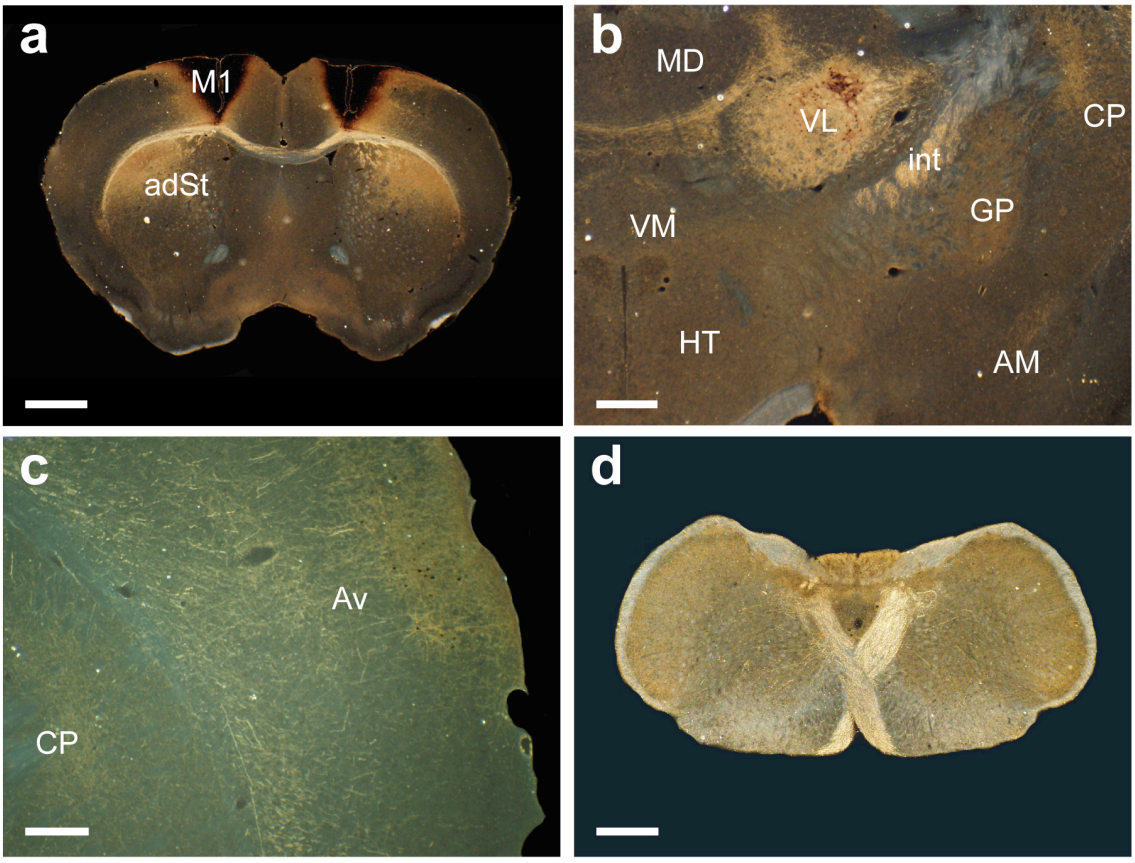

Figure 18: a, Bilateral biotinylated dextran amine (BDA) injections (black) in primary motor cortex (M1) revealed a dense terminal projection field (gold) in the underlying anterodorsal striatum (adSt). b, Cortical axons from M1 that travel through the internal capsule (int) and terminate in the ventral lateral nucleus of the thalamus (VL); A cluster of thalamic cells (dark brown) that project back to the M1 singing region is also observed in VL. Surrounding brain regions labeled for reference and orientation include the amygdala (AM), caudate-putamen (CP), globus pallidus (GP), hypothalamus (HT), medial dorsal nucleus of the thalamus (MD), and ventral medial nucleus of thalamus (VM). c, A cluster of motor cortical-projecting backfilled cells in layer 2/3 of ventral secondary auditory cortex (Av). d, BDA labeled cortico-spinal fibers in the brainstem at the level of the pyramidal decussation. Scale bars: 1 mm for a; 200 μm for b and for c; 500 μm for d.

Cortico-fugal fibers from M1 continued to descend via the genu of the internal capsule to the cerebral peduncles, where they entered the brainstem. Cortico-spinal axons labeled with BDA could be seen in the pyramidal tract all the way through the brainstem, crossing at the pyramidal decussation (Figure 18d) and exiting the caudal brainstem contralaterally to the spinal cord (Figure 19).

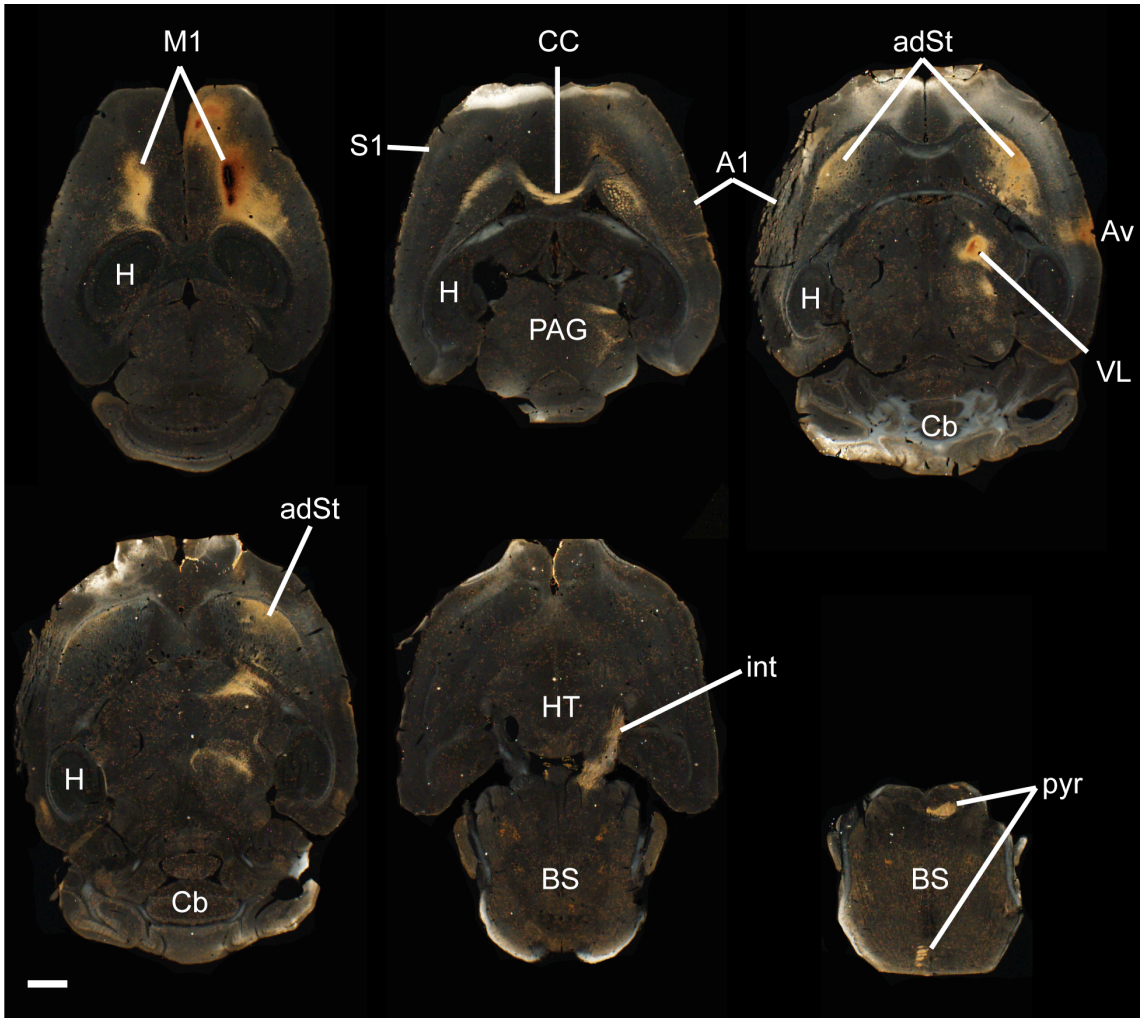

Figure 19: Horizontal sections showing the descending projections resulting from unilateral injection of biotinylated dextran amines into right primary motor cortex (M1). Axons reached contralateral anterodorsal striatum (adSt) and M1 via the corpus callosum (CC). Ipsilateral ventral secondary auditory cortex (Av) showed retrogradely labeled cells in cortical layer 2/3. Also ipsilaterally, descending axons traveled via the internal capsule to reach the ventral lateral nucleus of the thalamus (VL), and via the pyramidal tract (pyr) to reach the brainstem (BS). Surrounding regions that were not innervated but labeled for reference and orientation include primary auditory cortex (A1), cerebellum (Cb), hippocampus (H), hypothalamus (HT), and periaqueductal grey (PAG). Scale bar = 1 mm.

A subset of cortico-bulbar axons from M1 turned abruptly in the brainstem just rostral to the pyramidal decussation and entered the reticular formation at the level of Amb (Figure 20). These axons continued extending laterally through the magnocellular nucleus of the reticular formation. I observed extensive radiation of dendrites from the motoneurons innervating the CT and LCA muscles dorsomedially and ventrolaterally into the RF; however, it was not possible to differentiate between CT- and LCA-projecting motoneurons.

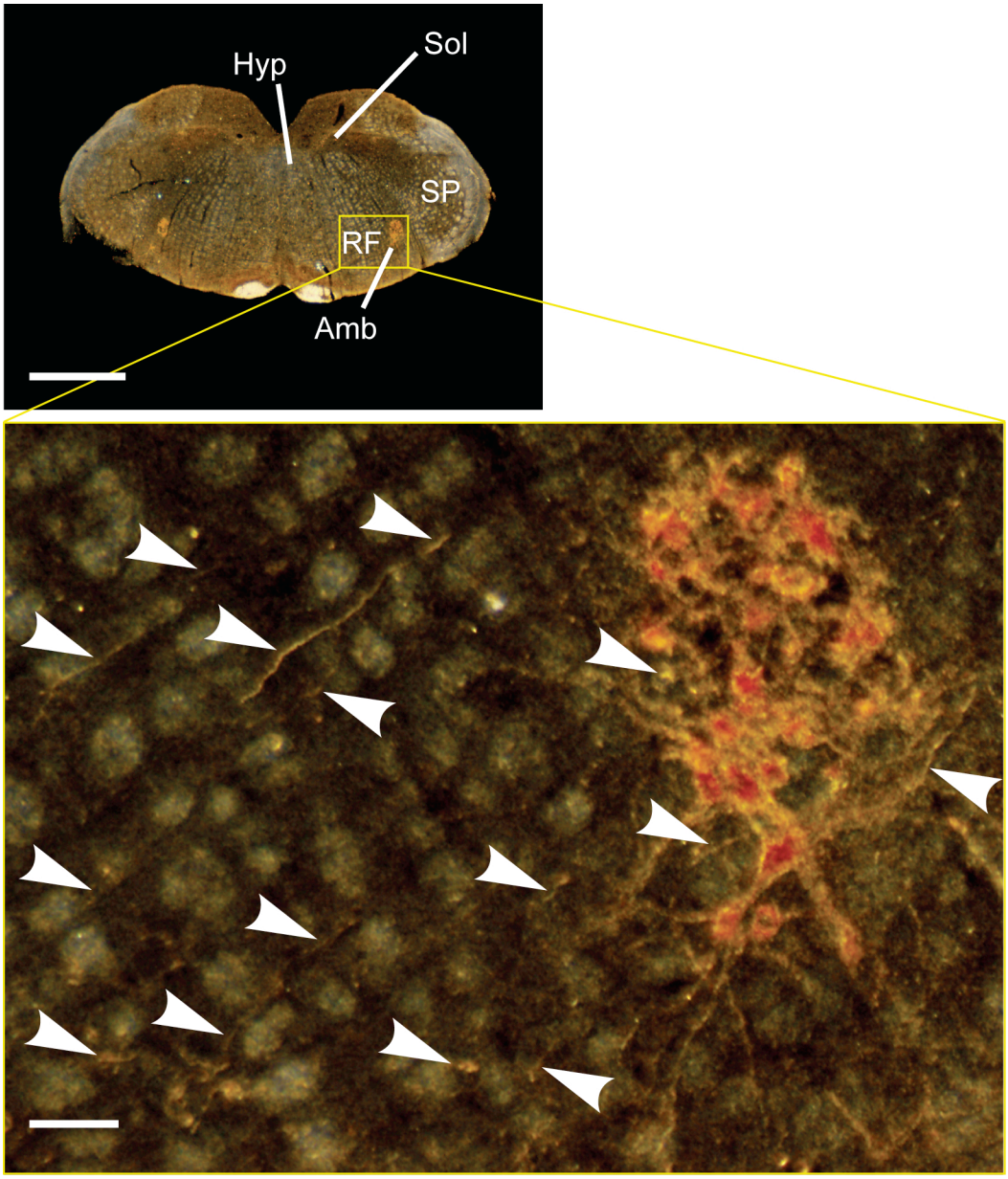

Figure 20: Fine caliber, varicose M1 axons labeled with biotinylated dextran amines (white arrow heads) and cholera toxin subunit b-labeled laryngeal motoneurons (red-orange) in the brainstem reticular formation (RF) at the level of nucleus ambiguus (Amb). Scale bars: 1 mm for main panel; 50 μm for inset.

Thus, although the cell bodies of laryngeal motoneurons are restricted to the boundaries of Amb, many synaptic connections are likely to occur in the extensive radiating dendritic field in the RF. Some axons exited the pyramidal tract and traveled dorsally along the midline then turned laterally to innervate Sol (Figure 21). Additionally, many fine caliber axons exited the pyramidal tract then dispersed throughout the RF contralateral to the M1 injection site and just medial to Amb in the region of laryngeal motoneuron dendritic arborization (Figure 22). Axons from M1 approached Amb medially, traveling along a dorsolateral trajectory. In z-stacks through the depth of multiple sections, axons were observed co-localized with dendrites from CT- and LCA-projecting motoneurons and tracking along the paths these dendrites made through the RF (Supplemental Movies 1-5). Many small black puncta can also be seen in the z-stacks dotting the area surrounding the laryngeal motoneurons, suggesting the presence of many synaptic boutons along the axons in the region of the dendritic fields. Some motor cortical axons extended laterally all the way to Amb reaching the zone where motoneuronal cell bodies were located, and terminated next to the dendrites (Figure 22) and soma of motoneurons (Figure 23).

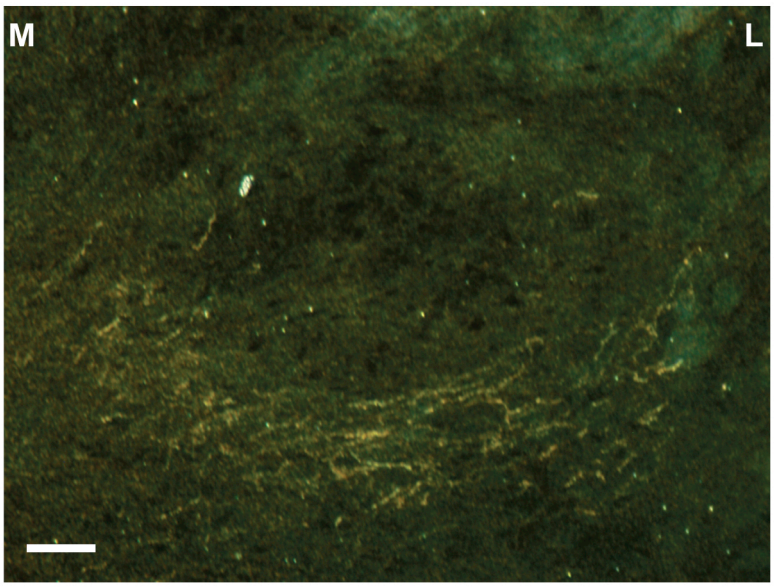

Figure 21: Motor cortical axons (white arrow heads) labeled with biotinylated dextran amines terminating in the solitary nucleus of the brainstem.

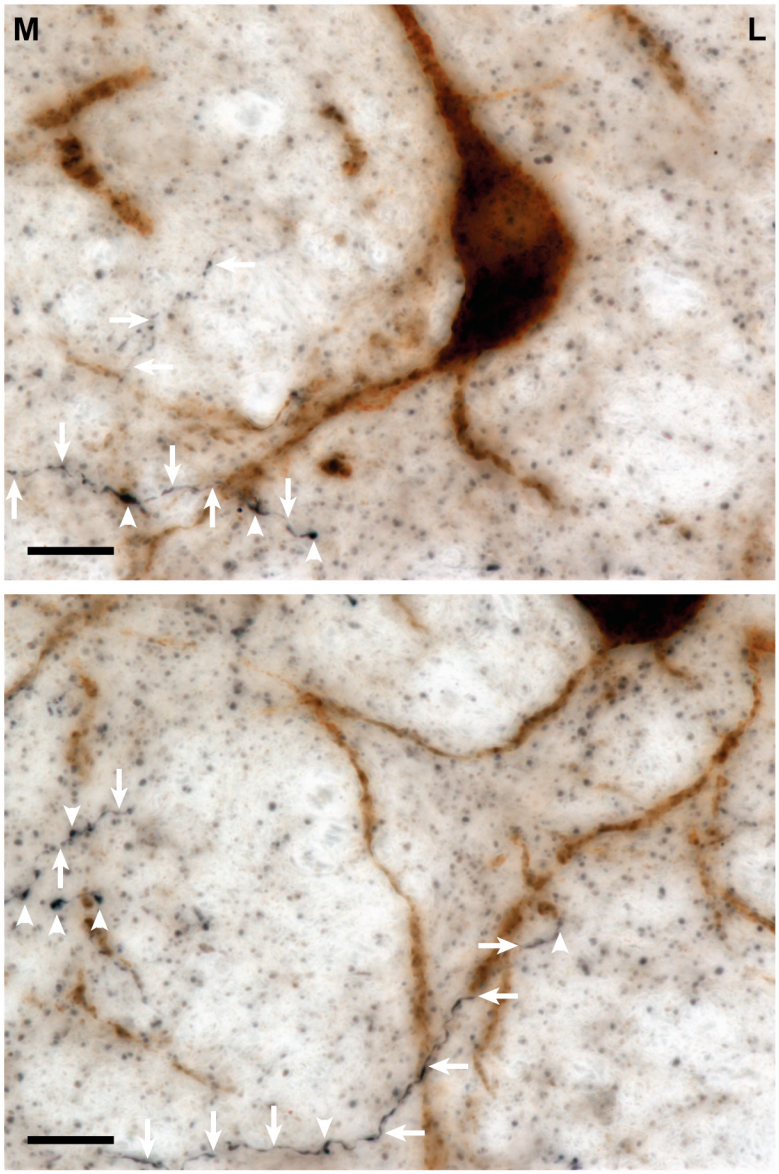

Figure 22: Motor cortical axon (white arrows) labeled with biotinylated dextran amines in the reticular formation medial to nucleus ambiguus forming boutons (white arrow heads) near the dendrites of cholera toxin subunit b-positive laryngeal motoneurons (brown). Scale bars = 10 μm.

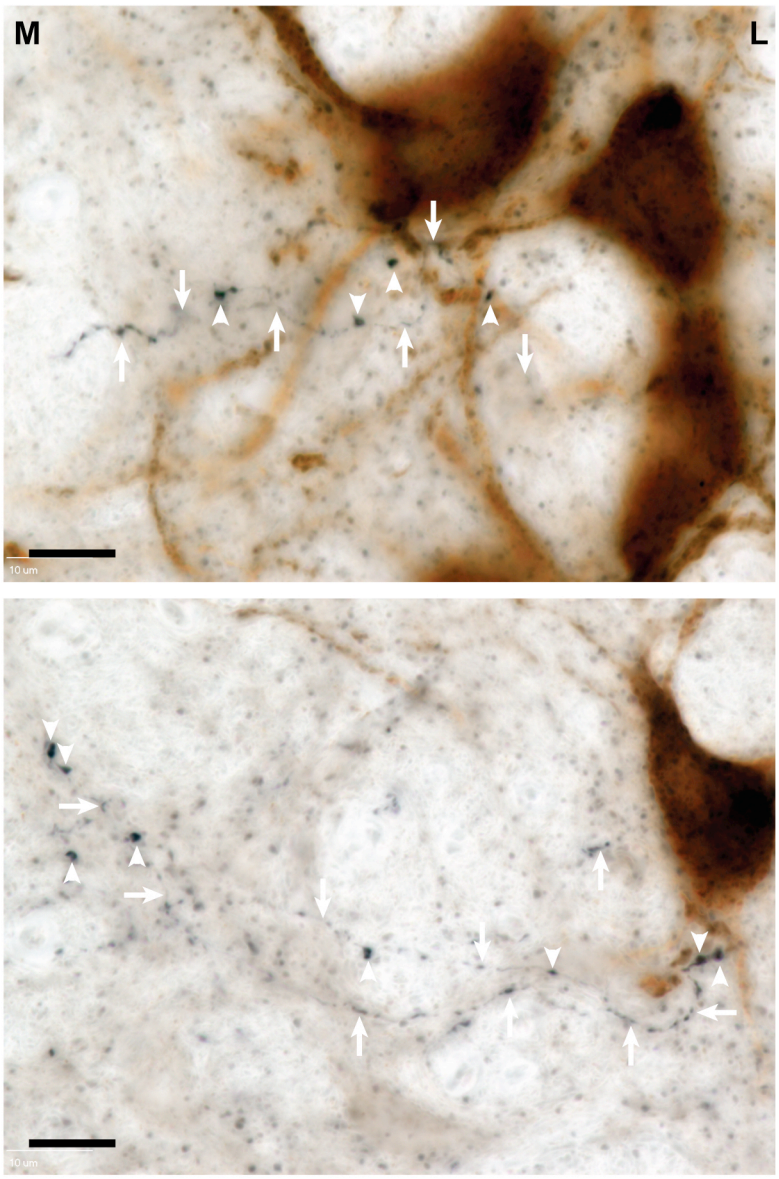

Figure 23: Motor cortical axons (white arrow heads) labeled with biotinylated dextran amines entering nucleus ambiguus and forming boutons (white arrows) near the somata of cholera toxin subunit b-positive laryngeal motoneurons (brown). Scale bars = 10 μm.

Motor cortical innervation of Amb was visually sparse relative to the dense terminal fields observed in the thalamus, striatum, and cortex; however, the only other known direct cortical projection to brainstem motoneurons, from barrel cortex to the VII is of similar density (Grinevich et al., 2005). Despite a lower density, the axons were concentrated in the area where laryngeal motoneuron dendrites are located. For example, BDA labeled axons were not observed rostral to Amb at the level of VII, nor in XII (Appendix C). Additionally, the motor cortical axons within and around Amb contact the dendrites and somata of laryngeal motoneurons. These results suggest a direct and specific projection from M1 to the laryngeal motoneurons of Amb.

The combined retrograde and anterograde tracing patterns show that mice have a cortical vocal premotor circuit that connects directly to vocal motoneurons in the brainstem and also to the anterior striatum and thalamus, similar to known circuits in humans and song learning birds (Figure 24). Previous to this discovery, a cortico-bulbar projection to vocal motoneurons was thought to be unique to humans amongst mammals (Kuypers, 1958c; Deacon, 2000; Jarvis, 2004; Jürgens, 2008; Fischer and Hammerschmidt, 2010). The results suggest that the cortical vocal pathway converges with the limbic-PAG pathway at the level the Amb, as in humans, and led me to ask whether mice require this projection to sing—the subject of Chapter 4.

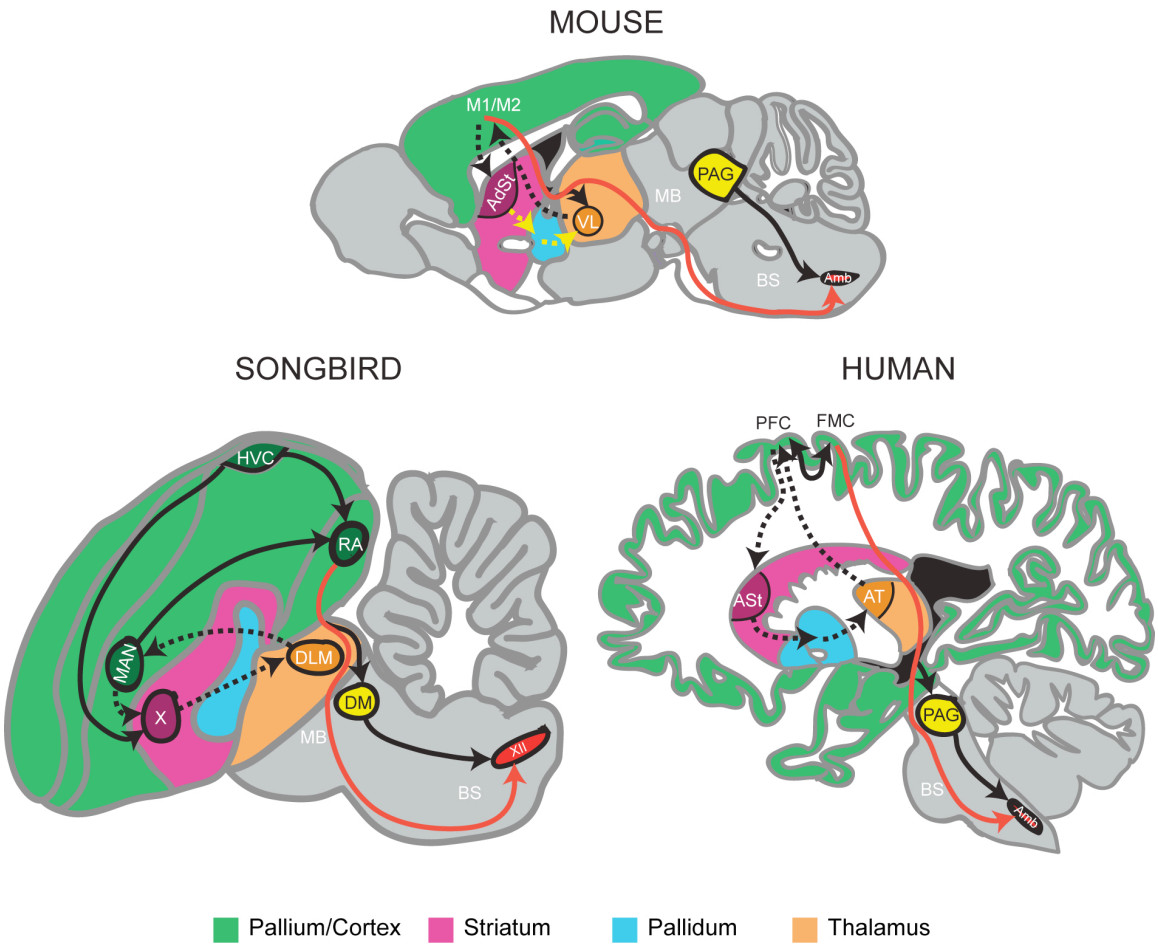

Figure 24: Summary diagram of mouse song system; yellow arrows are inferred from known connectivity of motor pathways in mammals. The songbird song system and human vocal control pathways are presented for comparison. All three vocal systems include a direct projection from cortex to the brainstem motoneurons innervating the vocal organ.

Chapter 4. Laryngeal Motor Cortex is Necessary for Producing Stereotyped Songs

In the previous chapters I reported that mice have not only the purportedly phylogenetically older limbic-midbrain pathway that controls the gating and programming of innate vocal production, but also the direct cortico-bulbar pathway that controls production of learned sounds like speech and birdsong (Nottebohm et al., 1976; Simpson and Vicario, 1990; Terao et al., 2007). Thus, I wondered if control of songs had been given over to this part of the song system in mice as it has in known vocal learning species.

4.1 Background – Brain Lesion Effects on Vocalizations

4.1.1 Effects of Forebrain Lesions on Learned Vocalizations

Innate vocalizations are thought to be controlled by the limbic-midbrain vocal pathway that involves PAG in mammals (Jürgens and Pratt, 1979), and DM in birds (Fukushima and Aoki, 2000). Lesions to PAG result in mutism in squirrel monkeys, completely abolishing all naturally produced species-specific calls (Jürgens and Pratt, 1979); however, artificial calls could still be evoked by electrical stimulation in the lateral medulla. Similarly, lesions to the analogous nucleus DM in both male and female Bengalese finches result in mutism for calls (Fukushima and Aoki, 2000). By contrast, call behavior is spared by lesions that included face motor cortex in squirrel monkeys (Kirzinger and Jürgens, 1982), the monkey homologue of Broca's area (Aitken, 1981; Jürgens et al., 1982), or the cortical song premotor pathway in songbirds (Simpson and

Vicario, 1990). While there was no discernable effect of the cortical lesions on monkey vocalizations, post-lesion calls of male zebra finches were slightly modified by the elimination of a learned component. This difference could reflect a difference in the underlying neural pathways. In monkeys, the indirect motor cortical vocal pathway bypasses PAG (Simonyan and Jürgens, 2003) and interfaces with the limbic-PAG pathway at the level of the brainstem (Jürgens and Ehrenreich, 2007). In songbirds the arcopallial vocal premotor nucleus RA projects to both the midbrain nucleus DM and the syringeal motoneurons (Wild, 1993), providing a mechanism for both driving song and directly influencing call programming.

4.1.2 Effects of Forebrain Lesions on Innate Vocalizations

Unlike innate calls, the effects of cortical lesions on speech and birdsong are severe. In humans, unilateral lesions to the cortical face area (Jürgens et al., 1982) produce transitory aphonia. Even small unilateral lesions restricted to primary motor cortex result in transitory aphonia, without aphasia, followed by severe dysarythia (Terao et al., 2007). This speech motor deficit suggests that the motor output for speech articulation depends critically on a small cortical representation in primary motor cortex. Similarly, in songbirds the premotor drive for song production is carried through the HVC-RA-nXIIts pathway. Similar to the aphonia reported in human patients, bilateral lesions targeted to RA in songbirds (Simpson and Vicario, 1990) completely abolish songs without recovery. Unilateral RA lesions in canaries lead to syllable instability

(Nottebohm et al., 1976). When HVC is lesioned or inactivated adult zebra finches lost their normal songs but reverted back to subsong (Aronov et al., 2008). Similar results were reported in budgerigars (*Melopsittacus undulatus*), with complete elimination of songs following bilateral damage to the central nucleus of AAc (Heaton and Brauth, 2000), a region analogous to RA in songbirds. Clearly, central control of production of learned vocalizations is dominated by cortical premotor pathways in both humans and vocal learning birds.

4.2 Effect of Motor Cortical Lesions on Mouse Ultrasonic Songs

To test the hypothesis that mouse ultrasonic songs are generated by the motor cortical vocal pathway I reported in previous experiments (Section 3.3.2), I made chemical ibotenic acid lesions bilaterally to the singing-activated and Amb-projecting region of MC. Songs before and after MC lesions were compared to those of mice that suffered lesions to visual cortex or a sham surgery.

4.2.1 Methods – Chemical Lesions and Song Analysis

4.2.1.1 Ibotenic Acid Injections

Adult BxD males more than 140 days old were given cortical lesions with ibotenic acid (10 μg/μL). General anesthesia was induced with 1% isofluorane and maintained by intramuscular injection of ketamine-xylazine (75 mg/kg ketamine; 5 mg/kg xylazine). The scalp was retracted and a small craniotomy was made over the injection site. Injections of ibotenic acid were made through a glass micropipette using

the Nanoject II microinjector (Drummond Scientific) into MC (Motor; n=5) at 4 sites 550μm from the brain surface (55 nL per site) along a track 1.2mm lateral and -0.2 to 0.4mm anterior to Bregma. One control group received bilateral ibotenic acid injections placed in visual cortex (Visual; n=4) at 4 sites 500 μm from the brain surface (55 nL per site) along a track 1.5 mm lateral and 2.6 to 3.2 mm posterior to Bregma. Another control group received sham surgery where a glass micropipette was inserted into the brain along the same track as the Motor group, but no injection was made (Sham; n=6). After all injections were complete, the wound was closed with TissueMend adhesive (Veterinary Products Laboratories, Phoenix, AZ).

4.2.1.2 Lesion Verification

PRV-Bartha was injected into CT and LCA muscles as described in the General Methods (Appendix A, Section A.5). PRV-Bartha-injected mice were sacrificed after a survival period of 5 days by overdose of pentobarbital sodium, and then perfused transcardially with 0.9% saline followed by 4% paraformaldehyde in 0.1M PBS. Brains were removed, post-fixed in 4% paraformaldehyde overnight, and cryoprotected in 30% sucrose in PB until sectioned. Coronal sections were cut at 40 μm on a cryostat into 0.1 M PBS, then processed for immunochemical detection of eGFP in PRV-Bartha-infected cells with DAB staining as described in Section 3.2.1.3.

I counted the number of surviving PRV-Bartha-labeled layer V pyramidal cells in M1 in seven 40 μm-thick frontal serial sections per hemisphere. Lesion sizes are

expressed as a percentage of cells eliminated from a baseline of 102 ± 16 cells (s.e.m.) counted from 7 similarly quantified unlesioned hemispheres.

4.2.1.3 Song Recordings

Each mouse was placed singly into a dark 15″x24″x12″ sound-attenuating recording chamber, and ultrasonic songs elicited by presentation of 200 μL of freshly collected female urine were recorded with UltraSoundGate CM16/CMPA ultrasound microphones and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). Details of the recording procedure are described in the General Methods (Appendix A, Section A.2).

4.2.1.3 Song Analysis

Sonograms were computed from each waveform (256 samples/block, half-overlap), and the following spectral features were calculated from the sonograms for each syllable: standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. Details of song analysis are described in the General Methods (Appendix A, Section A.3).

For the longitudinal data from before and after surgery, the mean feature value (FV) for each spectral feature was calculated for pre- and post-surgical recording epochs for all note types. I calculated the spectral feature score (SFS) as the logarithm of the normalized FV for each epoch (n) such that:

$$SFS(n) = \log10[FV(n) / FV(1)]$$

This log ratio allowed me to compare normalized longitudinal data to pre-treatment conditions across animals, which could differ in their absolute values.

4.2.2 Results

Lesions were restricted to MC (M1 and M2), and did not encroach on the medially adjacent Cg (Figure 25a). The lesions eliminated 79 – 100% of the layer V pyramidal neurons traceable with PRV-Bartha from the larynx in unlesioned brains (Figure 25b). The majority of lesions eliminated more than 95% of cells, and the average percentage of cells eliminated was 94%. In cases where the MC lesions were not complete, a few PRV-Bartha infected cells could be seen lateral to the injection site (Figure 25a).

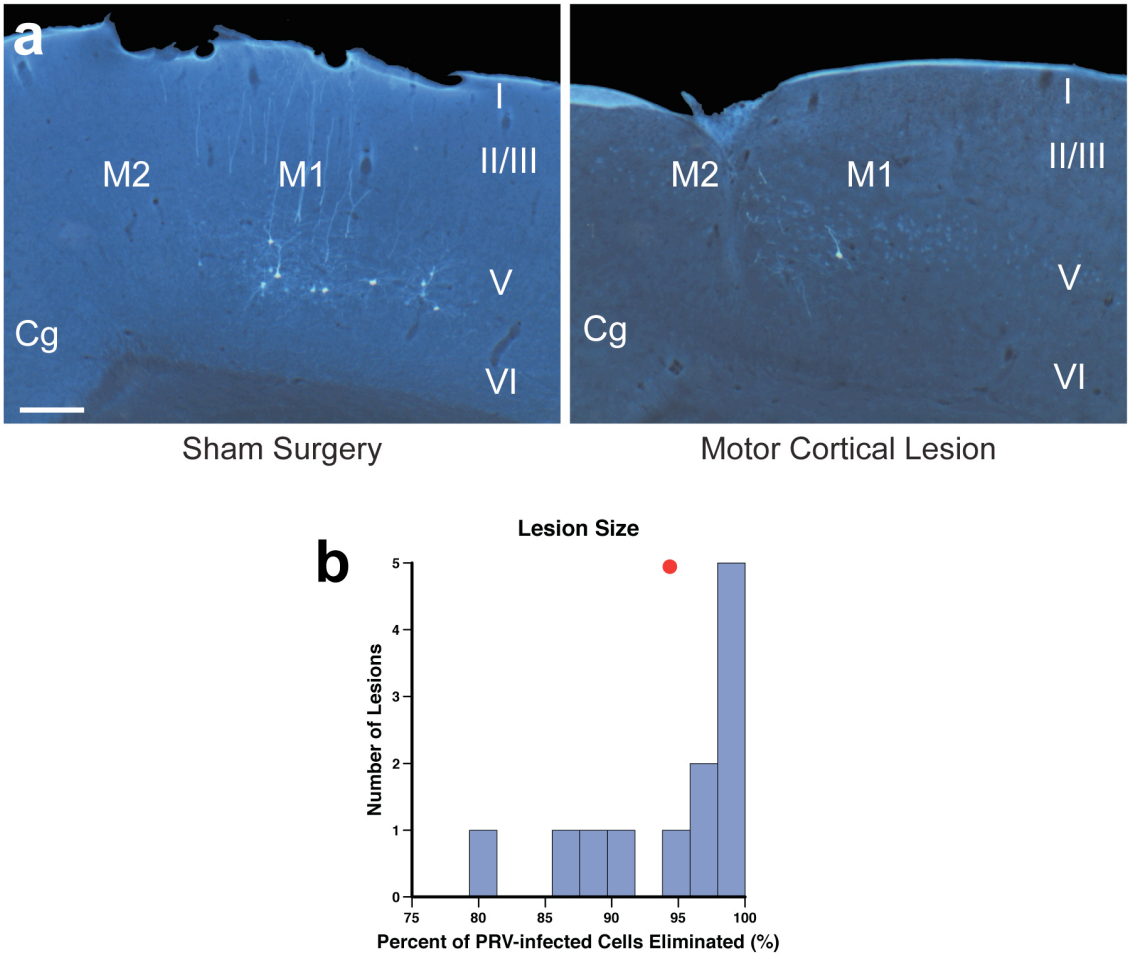

Figure 25: a, Elimination of pseudorabies virus (PRV-Bartha) back-traced premotor neurons in primary motor cortex (M1) following chemical lesions. Scale bar = 1 mm. b, Histogram representing lesion sizes as the percentage of Layer V pyramidal cells in M1 traced with PRV-Bartha that were eliminated by chemical lesions (12 cerebral hemispheres in 6 mice). Most lesions eliminated more than 85% of PRV-Bartha-labelled cells, with a mean lesion size of 94% (red dot).

I found that unlike lesions to motor cortex for humans and RA for songbirds, bilateral lesions of M1 did not eliminate production of mouse USVs (Figure 26). Visually, the songs looked relatively normal and the same individual syllable types were recognizable before and after lesion; however, syllable structure from one rendition to another looked more variable after lesion.

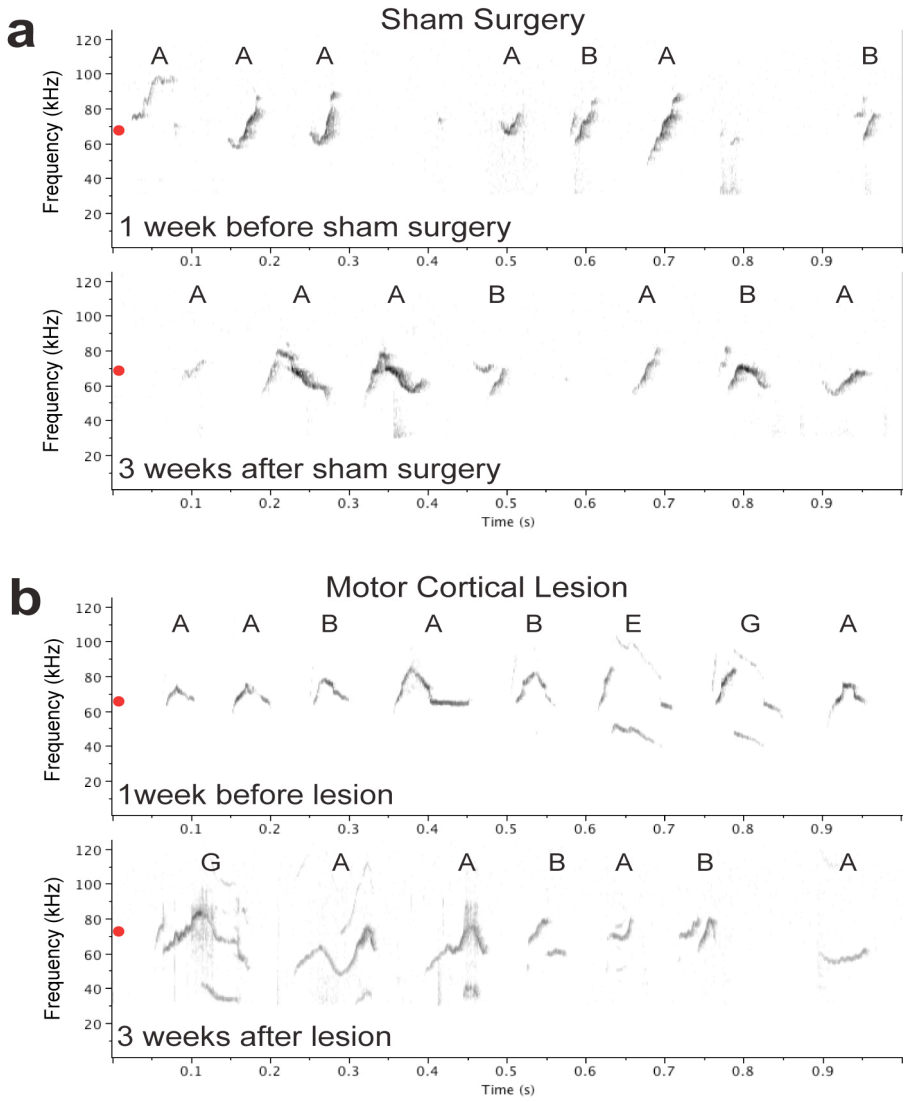

Figure 26: Sonograms representing 1 sec of male USVs before and after (a) sham surgery or (b) chemical lesion of M1. The pitch-shifted recordings corresponding to these sonograms can be heard in Supplementary Audios 2 - 5. Red dots represent the average pitch over the entire recording session for that individual animal.

To check for fine scale changes in the songs, I parsed the songs by syllable type and analyzed individual groups of syllables. Each of the 8 major syllable types described previously (Figure 3) was observed in all groups before and after surgery (Figure 27); however, syllable types I, J and K each consistently represented less than 1% of all identified syllables. The simple sinusoidal Type A syllables dominated the USV repertoire, followed by Type B syllables, then Type E in most conditions. Surgical treatment did not systematically change the general composition of the repertoire. Overall, the relative proportions of the 4 most frequently produced syllable types (A, B, E, and G) remained stable between pre- and post-operative recordings (Paired Student's t-tests; Sham: Pre, n=6; Post, n=6; A, $p=0.451$; B, $p=0.221$; E, $p=0.367$; G, $p>0.5$; Visual: Pre, n=4; Post, n=4; A, $p=0.251$; B, $p=0.224$; E, $p=0.173$; G, $p=0.204$; Motor, Pre, n=5; Post, n=5; A, $p=0.345$; B, $p>0.5$; E, $p=0.034$; G, $p=0.070$).

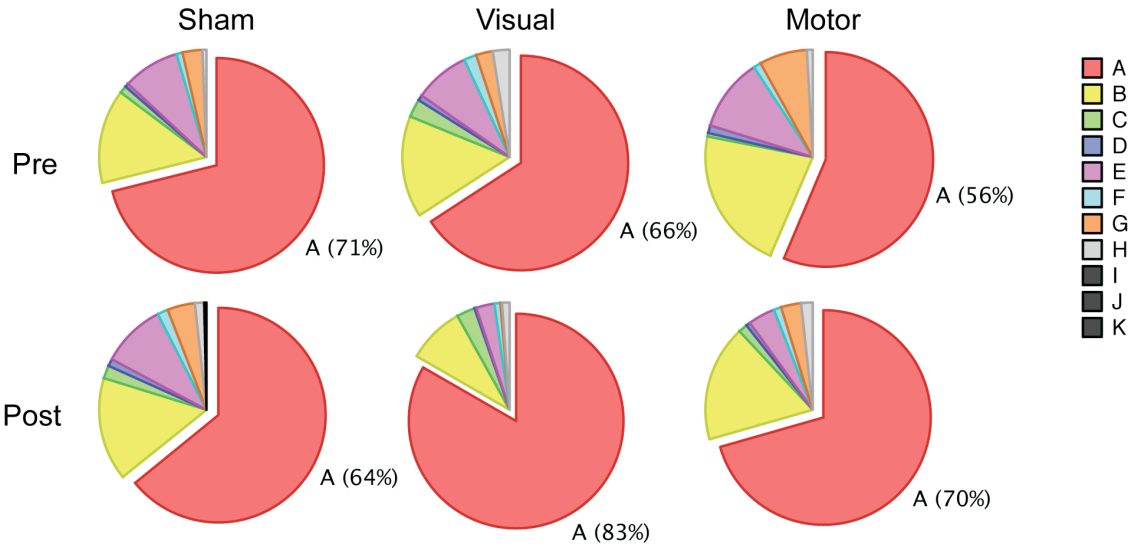

Figure 27: Composition of the repertoire of male mice that received sham surgery (Sham), lesions in visual cortex (Visual), or lesions in singing-activated motor cortex (Motor). Percentage of Type A syllables is indicated for each chart.

Significant differences between pre- and post-lesion songs were observed for some spectral features. For example, for Type A syllables group mean differences were detected in the post-operative SFS for standard deviation of the pitch (Figure 28; Kruskal-Wallis H-Test; Sham, n=6; Visual, n=4; Motor, n=5; Pre: $p=0.325$; Post: $p=0.036$). When the post-operative SFS of each of the lesion groups was compared to Sham controls, only the SFS for the Motor group was significantly higher (Mann-Whitney U-tests; Sham, n=6; Visual, n=4; Motor, n=5; Visual, $p>0.5$; Motor, $p=0.011$). This increase in standard deviation suggests that the Motor group was no longer restricting the pitch of their vocalizations to a normal frequency range. Figure 29 illustrates this loss of fine control of pitch for one mouse from the Motor group. The instantaneous pitches produced while singing Type A syllables before surgery were concentrated in a frequency range between 60 – 80 kHz; however, after sustaining a lesion to MC the distribution flattened and widened, extending more than 20 kHz on both the lower and higher ends from 40 – 115 kHz. On average, the standard deviation of pitches for the Motor group increased from 7.95 kHz to 11.56 kHz, corresponding to an average increase of 0.33 in the SFS.

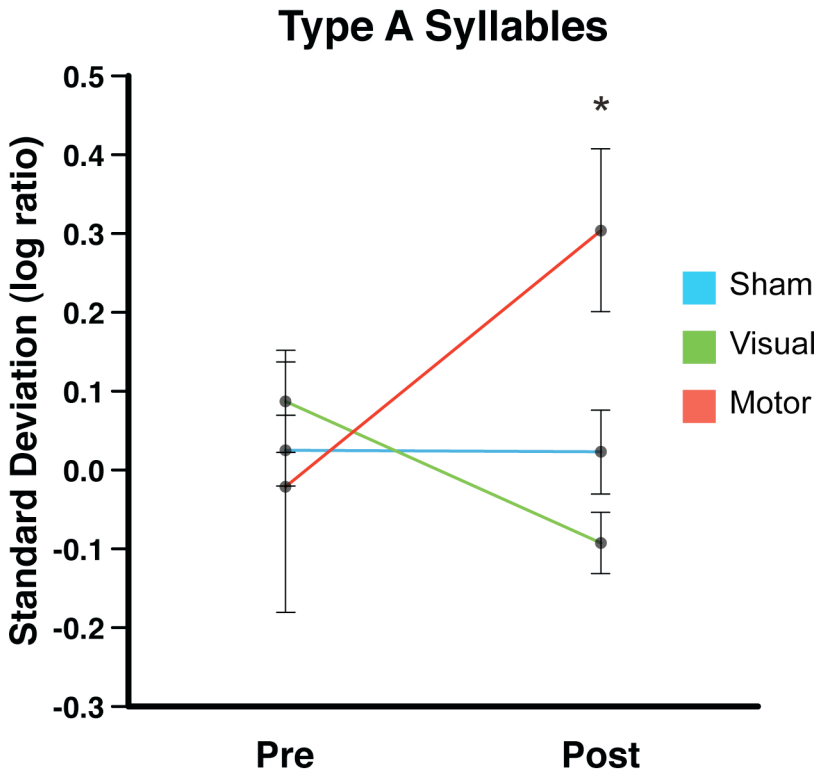

Figure 28: Spectral feature scores (SFS; expressed as log-ratio) for the standard deviation of the pitch distribution of Type A syllables before and after surgery in animals that received sham surgery (Sham), visual cortex lesions (Visual), or lesions to the singing-activated portion of motor cortex (Motor). The Mann-Whitney U Test was used to test individual group mean differences relative to Sham controls. (Sham, n=6; Visual, n=4; Motor, n=5; * = $p<0.05$).

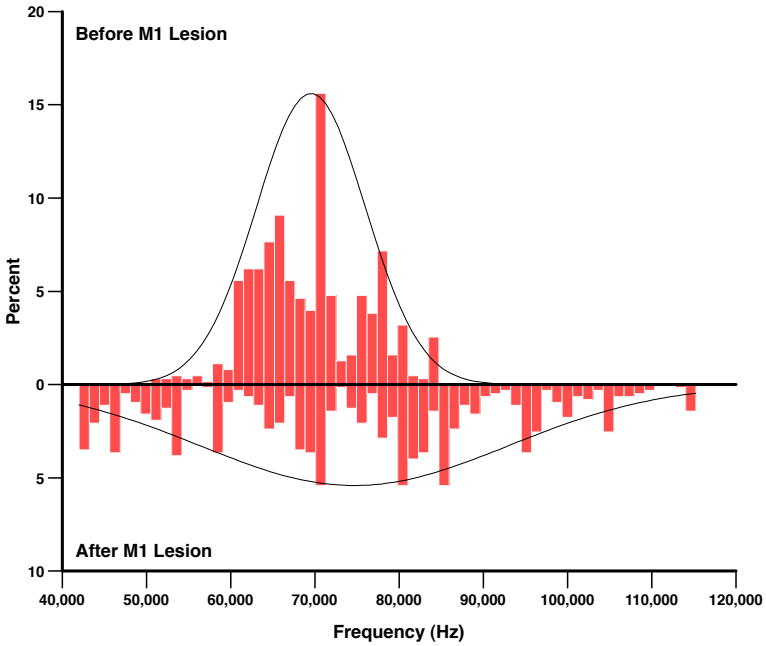

Figure 29: Histogram showing the relative occurrence of instantaneous pitches produced by one mouse while singing Type A syllables before and after receiving bilateral lesions to the singing-activated portion of primary motor cortex (M1).

To determine if other spectral features followed a similar trend I calculated the change in SFS (ΔSFS) from pre- to post-operative songs for each mouse. Significant group mean differences in ΔSFS were detected for the following 4 spectral features: standard deviation, ¾ frequency, final frequency, and frequency variance (Figure 30; Kruskal-Wallis H-tests; Sham, n=6; Visual, n=4; Motor, n=5; p-values reported on the graph). For the Sham control group mean ΔSFS did not deviate from zero for any spectral features, suggesting no post-operative changes in spectral content of USVs. For those features where group mean differences in ΔSFS were detected, individual group means were compared directly to Sham controls (Mann-Whitney U-tests; Sham, n=6; Visual, n=4; Motor, n=5; Standard Deviation: Visual, p=0.201, Motor, p=0.028; ¾ Frequency: Visual, p=0.088, Motor, p=0.068; Final Frequency: Visual, p=0.88, Motor, p=0.045; Frequency Variance: Visual, p>0.5, Motor, p=0.011). For all spectral features tested there was an overall trend towards higher ΔSFS values in the Motor group versus the Sham controls, and the increase was significant for three of the four features that showed group mean differences: standard deviation, final frequency, and frequency variance. These results indicate that USVs of the Motor group changed significantly more than those of the Sham control group, but the Visual group did not. Thus, the effect on the spectral features of song resulted specifically from motor cortex lesions and was not an artifact of surgical manipulation. The main effects observed were on

standard deviation of pitch and frequency variance, with a slight effect on the final pitch of Type A syllables.

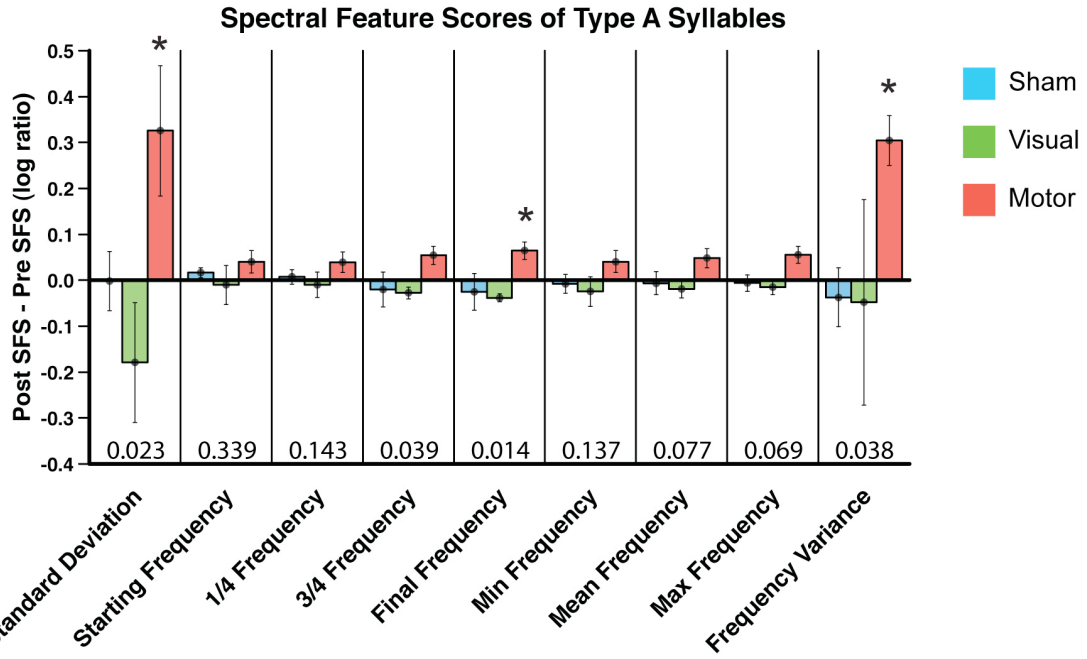

Figure 30: The change in SFS for various pitch-based features following surgery (Post-Pre) compared between treatment groups. Kruskal-Wallis H-tests (Sham, n=6; Visual, n=4; Motor, n=5; p-values reported on graph) with post-hoc Mann-Whitney U-tests to test for differences relative to the Sham group (Sham, n=6; Visual, n=4; Motor, n=5; * = $p<0.05$).

The results of this experiment lead me to reject the hypothesis that mouse USVs are generated by motor cortex; however, they also indicate that the motor cortex is necessary for generating more acoustically stereotyped songs. The pitch of Type A syllables, representing the bulk of vocal output, was more variable and slightly higher than before the lesions. When slowed and pitched-shifted down to the human hearing range, the post-lesion USVs sounded less refined (Supplementary Audios 4 & 5, corresponding to sonograms in Figure 26b). These findings suggest that without motor cortical input to the song system the mice were unable to finely control acoustic features of their songs.

Chapter 5. Auditory Feedback in Vocal Production

Like input from motor cortex, auditory experience seems to be more important for the production of learned vocalizations than innate calls. In humans and songbirds auditory experience plays a critical role at multiple stages in the ontology of vocal behavior: 1) a sensory phase during which an auditory memory or 'template' is formed following exposure to an appropriate model; 2) a sensorimotor phase during which vocal output is monitored and compared to the model in a guided learning process; 3) an adult maintenance phase during which auditory feedback is used to stabilize vocal output over the long-term. I posit that the main difference between learning by imitation and improvisation is the dependence on the first stage. In imitation the model or template is acquired externally. In improvisation there is no external model against which to measure progress, so another instructive signal must guide the learning process; however, this strategy likely involves a similar mechanism of auditory self-monitoring followed by selection and retention of preferred learned features. Thus, auditory experience is critical under either learning paradigm. Accordingly, experiments testing for vocal learning have typically focused on modifying, disrupting, or removing auditory information at the various developmental phases.

5.1 Background – Auditory Feedback in Vocal Development

5.1.1 Effects of Deafening on Innate Vocalizations

It has been demonstrated in various mammalian (Talmage-Riggs et al., 1972; Romand and Ehret, 1984; Hammerschmidt et al., 2001) and avian (Konishi, 1964; Nottebohm and Nottebohm, 1971; Kroodsma and Konishi, 1991) species that innate vocalizations do not depend on auditory experience at any developmental stage. Eastern phoebes (*Sayornis phoebe*), a sub-oscine vocal non-learning species, develop normal species-specific songs after being mechanically deafened by cochlear removal before the onset of singing behavior (Kroodsma and Konishi, 1991), despite being very closely related to vocal learning songbirds. Similar results have been reported in the more distantly related ringdove (Nottebohm and Nottebohm, 1971) and chicken (Konishi, 1963). In non-human primates, neither hereditary deafness (Hammerschmidt et al., 2001) nor deafening by cochlear coagulation (Talmage-Riggs et al., 1972) has an effect on normal vocal behavior. Unsurprisingly, the less severe auditory deprivation caused by social isolation also has no effect on monkey call spectral structure (Winter et al., 1973; Hammerschmidt et al., 2001).

5.1.2 The Importance of Auditory Information for Learned Vocalizations

Learned vocalizations are susceptible to elimination or disturbance of auditory feedback at various stages in development. For example, juncos (*Junco oreganus*) require auditory feedback to develop a normal song repertoire and the fine structure of their

syllables, but do not need to hear and copy examples from other conspecifics (Marler et al., 1962; Konishi, 1964). Thus, deafening but not social isolation disrupts vocal development. In species like the junco the role of auditory experience is limited to the sensorimotor phase of vocal ontogeny, but auditory feedback is still necessary for normal vocal development. The effects of deafening are more severe in species that also require early auditory experience to form a model or template that guides learning. In songbirds early deafening in the sensory acquisition (Marler and Waser, 1977) or sensorimotor phase of song learning (Konishi, 1965) has a dramatic effect, resulting in severely degraded song characterized by a small repertoire with highly variable and unstable notes. Songbirds raised in social isolation, thus removing the possibility of external auditory influence, develop highly abnormal 'isolate song' (Marler, 1970). Taken together these findings reveal that songbirds need to hear others to learn what to mimic and themselves to practice their own copy. But songbirds continue to depend on auditory information even after learning and stabilizing normal songs. For example, Bengalese finches suffer rapid deterioration of syntax and phonology when deafened as adults (Woolley and Rubel, 1997). Even the milder treatment of disrupting auditory feedback signals in real-time without deafening is sufficient to cause a destabilization of learned song features (Leonardo and Konishi, 1999). Thus, songbirds clearly rely heavily on auditory experience throughout the entire song development process, including for maintenance and stabilization of songs learned early in life.

Human speech shares with birdsong a continued dependence on auditory information throughout life (Doupe and Kuhl, 1999). For example, early language deprivation by social isolation severely disrupts speech acquisition (Fromkin et al., 1974). In this regard, humans and some songbirds (Thorpe, 1958; Marler, 1970) are subject to sensitive periods for vocal development. Exposure to and learning of language-specific sounds must occur before a particular developmental stage after which the capacity for vocal learning is dramatically attenuated (Doupe and Kuhl, 1999). But the reliance on auditory feedback does not end when the sensitive period closes. Post-lingually deaf patients suffer a degradation of speech sounds that results in decreased control of phonation, disrupted prosody, and abnormal suprasegmental properties of sentences (Waldstein, 1990), with younger patients more strongly afflicted. Thus vocal learners seem to make use of auditory feedback to calibrate the fine phonetic control required to produce high-quality vocalizations even after the waning of a robust vocal learning ability.

5.2 Effects of Deafness on Mouse Ultrasonic Songs

I reasoned that if male mice learn any aspect of their courtship vocalizations through either imitation or improvisation, then they should require auditory information in order to practice and maintain the spectral quality of songs. To assess the role of auditory information in mouse song behavior I analyzed the songs of males that

were genetically deaf, or mechanically deafened either in early development or as adults.

5.2.1 Methods – Deafening and Song Analysis

5.2.1.1 Conductive Hearing Loss in Pups

BxD mouse pups 12 to 13 days old (n=10) were deafened by bilateral removal of the tympanic membrane and stapes. Anesthesia was induced with 3% isofluorane in oxygen. Because the pinna is still closed over the outer ear canal at this age, a retro-aural incision was made and the skin was retracted to reveal the inner ear canal. A small incision was made through the wall of the ear canal to reveal the tympanic membrane, which was punctured with a pair of fine forceps followed by removal of the stapes. The wound was closed with TissueMend (Veterinary Products Laboratory, Phoenix, AZ). Pups were sexed at weaning (postnatal day 22) and songs were recorded from the juvenile males after postnatal day 30. Attempting this deafening procedure at an earlier age lead to excessive bleeding due to tissue being too soft prior to P12.

5.2.1.2 Cochlear Removal in Adults

Adult BxD males 134-136 days old (n=5) were deafened by bilateral cochlear removal. Details of the deafening surgery are described in the General Methods (Appendix A, Section A.4).

5.2.1.3 Congenital Deafness

B6.129S1-Casp3^{tm1Flv}/J males were purchased from the Jackson Laboratory (Bar Harbor, Maine).

5.2.1.4 Song Recordings

Each mouse was placed singly into a dark 15″x24″x12″ sound-attenuating recording chamber, and ultrasonic songs were recorded with UltraSoundGate CM16/CMPA ultrasound microphones and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). Adult songs were elicited by presentation of 200 μL of freshly collected female urine. Juvenile songs were elicited by placing a sexually experienced female in the recording chamber. Details of the recording procedure are described in the General Methods (Appendix A, Section A.2).

5.2.1.5 Song Analysis

Sonograms were computed from each waveform (256 samples/block, half-overlap), and the following spectral features were calculated from the sonograms for each syllable: standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. Details of song analysis are described in the General Methods (Appendix A, Section A.3).

For the longitudinal data from adults, the mean feature value (FV) for each spectral feature was calculated for pre- and post-surgical recording epochs for all note

types. I calculated the spectral feature score (SFS) as the logarithm of the normalized FV for each epoch (n) such that:

$$SFS(n) = \log10[FV(n) / FV(1)]$$

This log ratio allowed me to compare normalized longitudinal data to pre-treatment conditions across animals, which could differ in their absolute values.

5.3 Results

5.3.1 First Songs of Deaf Juveniles

Songs were recorded from the juvenile males' first experience with a female since weaning. On the first day of female-elicited singing, males deafened on P12 (Deaf) and hearing-intact controls (Hearing) produced recognizable songs with complex notes containing characteristic pitch jumps (Figure 31).

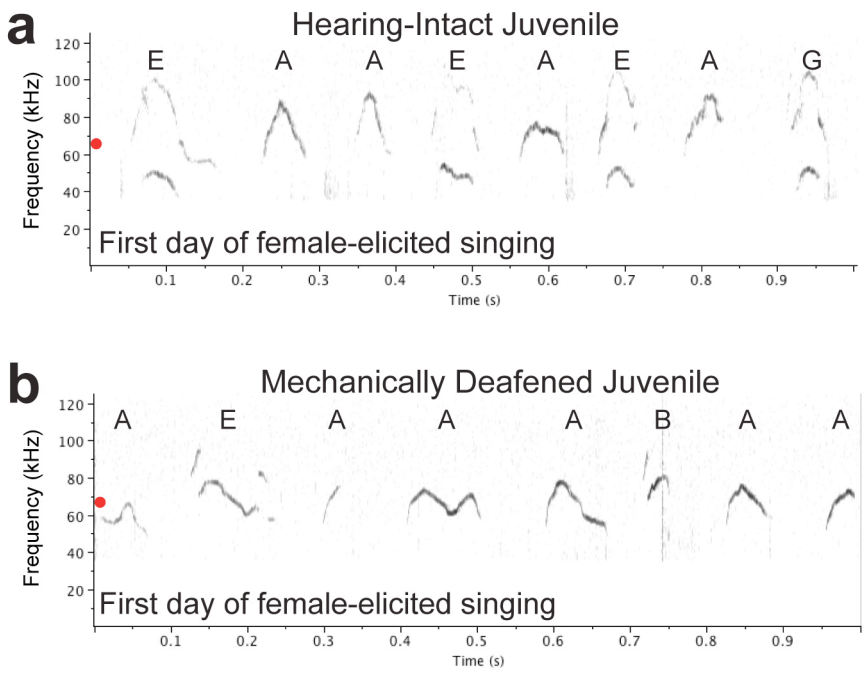

Figure 31: Sonograms representing 1 second of ultrasonic song from (a) a hearing-intact (Hearing) juvenile and (b) a P12-deafened (Deaf) littermate on the first day of singing to a female (sonograms correspond to Supplementary Audios 6 & 7). Red dots represent the average pitch over the entire recording session for that individual animal.

The songs of both Hearing and Deaf juveniles contained all 8 major syllable types. Type A syllables were significantly overrepresented in the Deaf repertoire, representing on average 78 ± 2.2% of all identified syllables in Deaf songs versus 64 ± 4.3% of syllables in Hearing songs (Figure 32; Student's t-test; Deaf, n=10; Hearing, n=8; p<001). This greater percentage of simple notes was accompanied by a significantly lower percentage of the complex three-note Type E syllables which were 4 ± 1.7% and 12 ± 1.6% of identified syllables in Deaf and Hearing songs, respectively (Student's t-test; Deaf, n=10; Hearing, n=8; p=018). The proportion of the other major syllable Type B did not change after deafening (Student's t-test; Deaf, n=10; Hearing, n=8; p=0.172). Thus, the P12-deafened juveniles were singing a higher proportion of simple notes in their songs, and singing fewer of the more complex 3-note syllables.

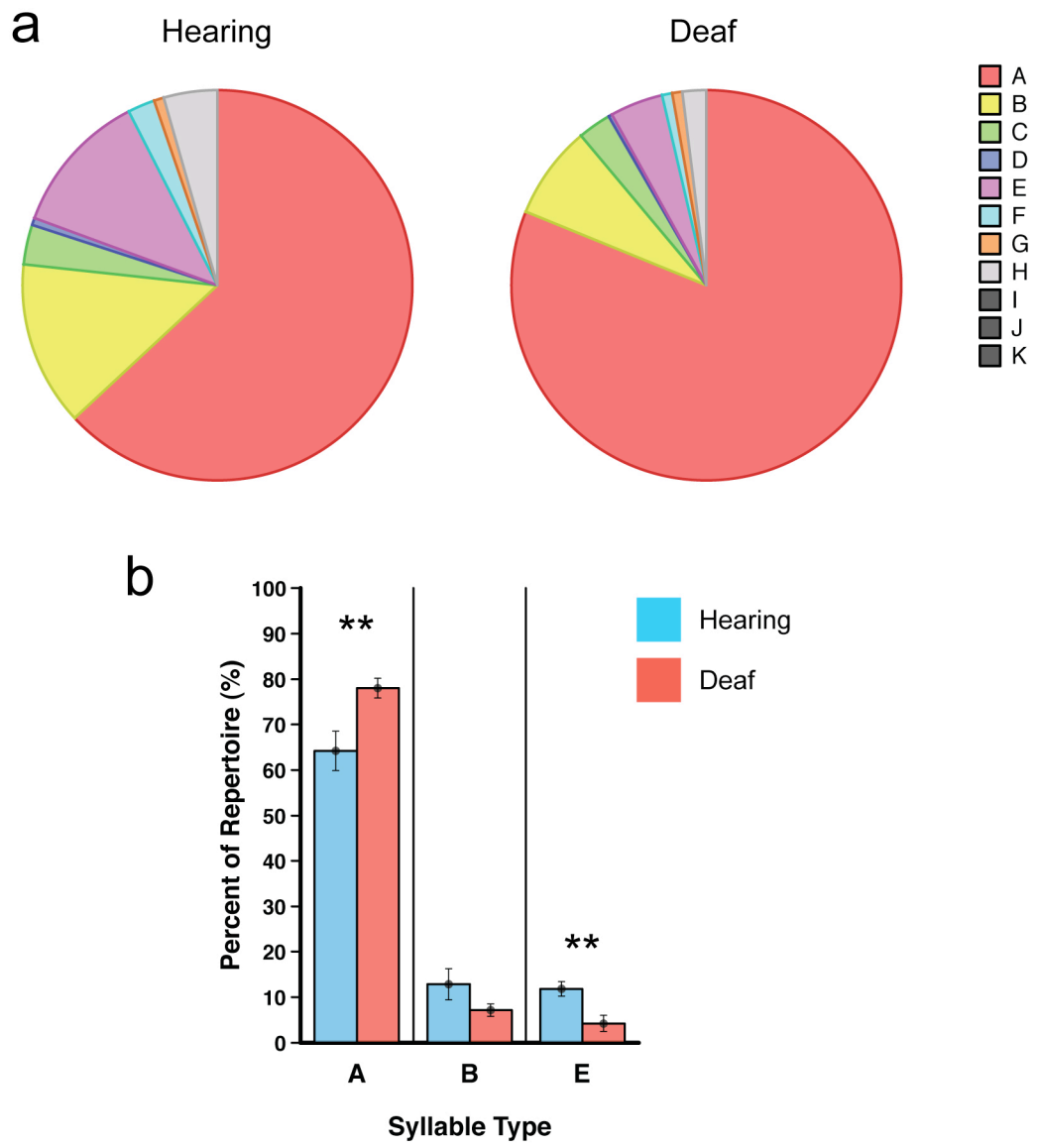

Figure 32: a, Composition of repertoire by syllable type for hearing-intact (Hearing) and P12-deafened (Deaf) juveniles. b, Type A, B, and E syllables as a percentage of all identified notes produced by Hearing and Deaf juveniles (Student's t-test; Deaf, n=10; Hearing, n=8; * = p<0.05; *** = p<0.001).

In addition to changes in repertoire composition, I found that the songs of P12-deafened juveniles also displayed a spectral abnormality. Here, I focused analysis on Type A syllables, because they comprised the bulk of the juvenile repertoire regardless of hearing condition (Figure 32a). The standard deviation of the pitch distribution for Type A syllables was significantly lower for the songs of P12-deafened juveniles (Figure 33a; Student's t-test; Deaf, n=10; Hearing, n=8; p=0.28). This decrease can be seen in the pitch distributions for Type A syllables from the two mice with the highest standard deviation in each treatment group (Figure 33b). The means of the pitch distributions were almost identical at 69.8 kHz and 69.9 kHz for songs from the Hearing and Deaf juveniles, respectively; however, the 9.4 kHz standard deviation of the pitch distribution of the Deaf juvenile's songs was 30 percent lower than the 13.4 kHz standard deviation of the pitch distribution from the Hearing juvenile's songs. On average, there was an 18.75 percent difference in standard deviation of pitch between 9.6 ±0.75 kHz in the Hearing group and 7.8 ± 0.30 kHz in the Deaf group. No other pitch-based spectral feature differed between Hearing and Deaf groups (data not shown; Student's t-tests; Deaf, n=10; Hearing, n=8; Starting Frequency, p>0.5; ¼ Frequency, p>0.5; ¾ Frequency, p>0.5; Final Frequency, p>0.5; Minimum Frequency, p>0.5; Mean Frequency, p>0.5 ; Maximum Frequency, p>0.5 ; Bandwidth, p>0.5 ; Frequency Variance, p=0.126; Spectral Purity, p=0.429).

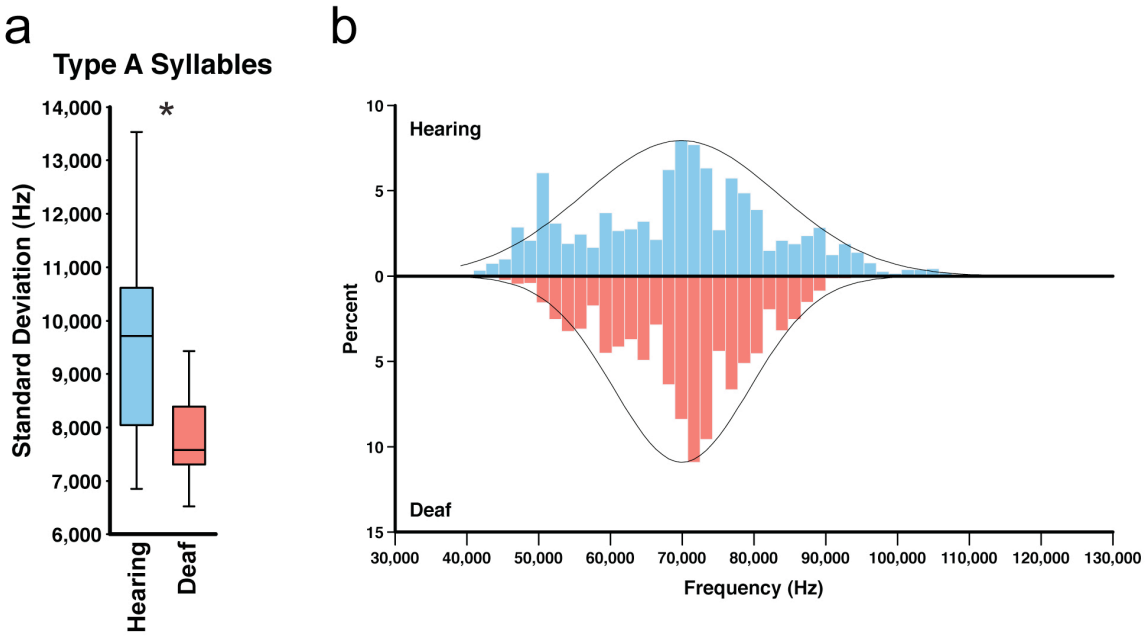

Figure 33: a, Standard deviation of the pitch distribution for Type A syllables of hearing-intact (Hearing) and P12-deafened (Deaf) juveniles. (Student's t-test; n=10, Deaf; n=8, Hearing; * = $p<0.05$). b, Distribution of all pitches produced by the Hearing and Deaf juveniles with the highest standard deviation.

These results suggest that deafening early in development had a slight effect on the very first songs produced by juveniles in response to females. Deaf juveniles produced songs that were less acoustically diverse and incorporated more simple notes than their hearing-intact littermates. The differences between deaf and hearing-intact juveniles cannot be explained by the Lombard effect, because the total acoustic power did not differ between the two groups (data not shown; Student's t-test; Deaf, n=10; Hearing, n=8; $p>0.5$).

5.3.2 The Effects of Deafening on Adult Songs

Before deafening both mechanically deafened (Deaf) and sham-operated (Sham) groups produced normal looking songs (Figure 34a & b) that included all major syllable types (Figure 31c). As in prior analyses of BxD songs, syllable Type A dominated the repertoire (Figures 27 and 32). Relative proportions of the major note types A, B, E, and G did not differ between the pre-operative songs of Sham and Deaf adults (Mann-Whitney U-tests; n=5 per group; A, $p>0.5$; B, $p>0.5$; E, $p=0.347$; G, $p=0.175$).

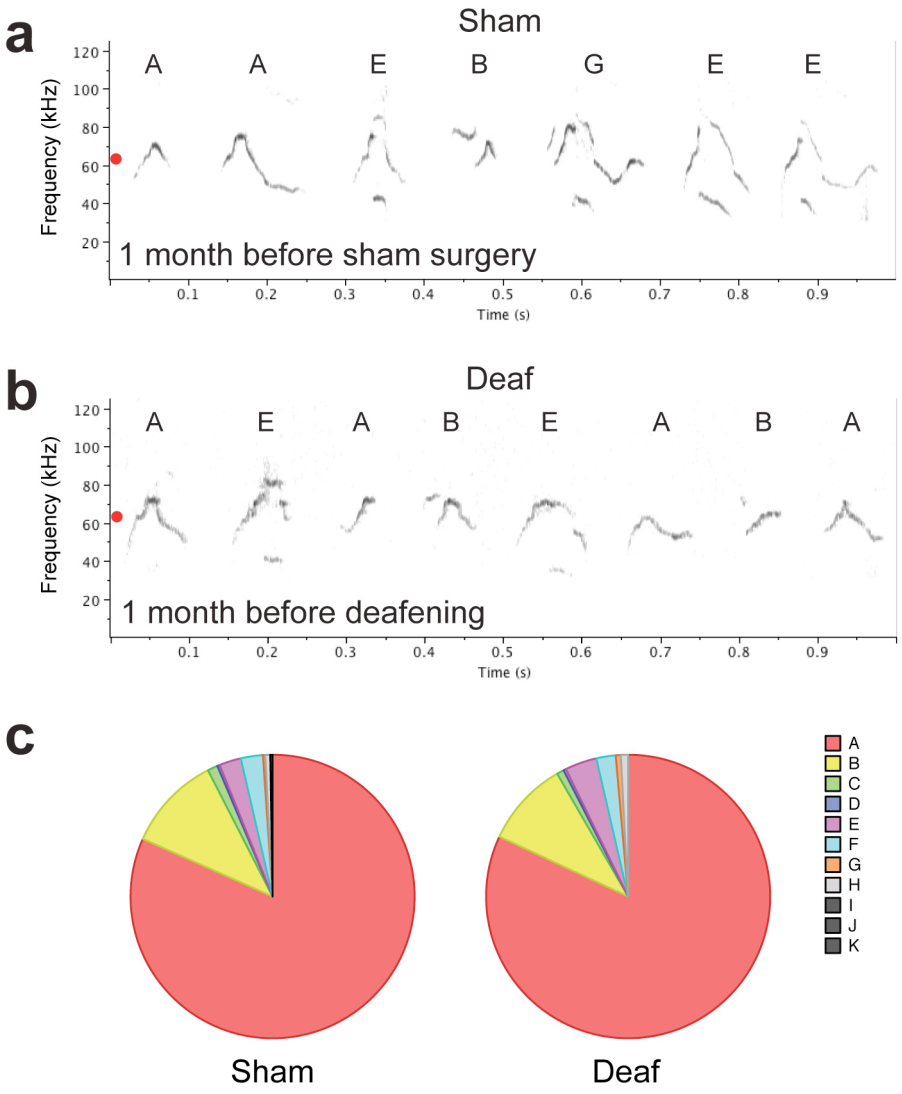

Figure 34: Sonograms representing 1 second of ultrasonic song from adult mice before sustaining (a) sham surgery (Sham) or (b) bilateral cochlear removal (Deaf). The pitch-shifted recordings corresponding to these sonograms can be heard in Supplementary Audios 8 - 9. Red dots represent the average pitch over the entire recording session for that individual animal. c, Composition of the repertoire by syllable type for Sham and Deaf mice before surgery.

In addition to having similarly composed syllable repertoires before surgery, the SFS's of Deaf and Control groups were statistically indistinguishable (Figure 35; Mann-Whitney U-tests; n=5, Sham; n=5, Deaf; p-values reported on graphs).

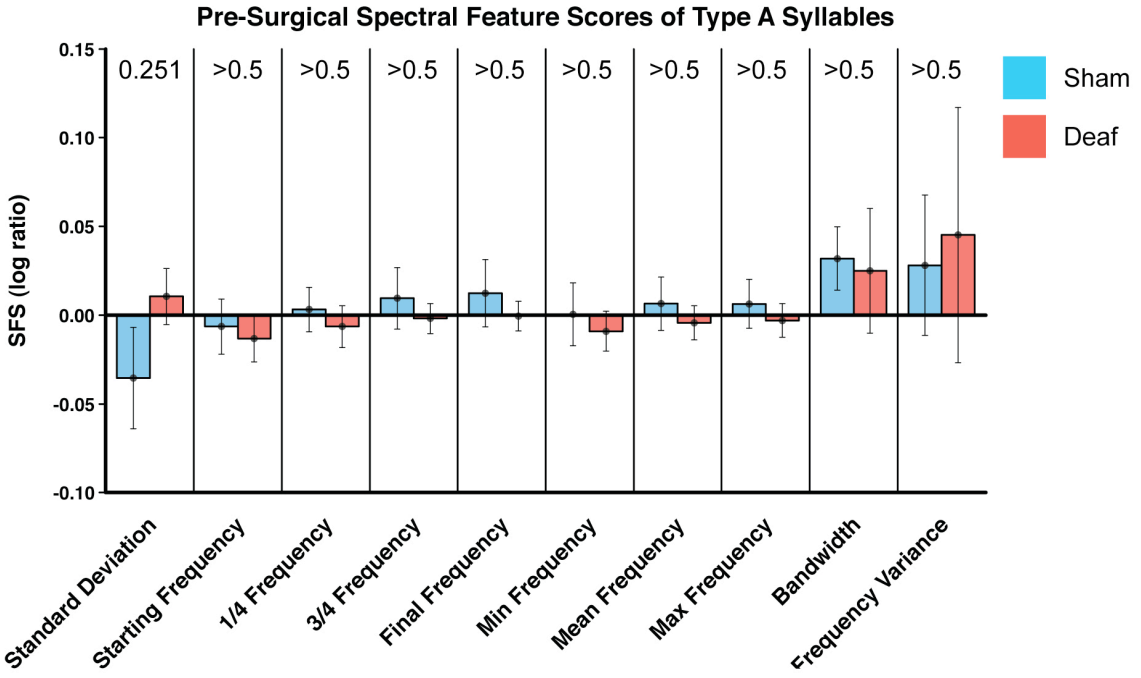

Figure 35: Pre-surgical spectral features (SFS) of Type A syllables produced by adult mice before sustaining sham surgery (Sham) or bilateral cochlear removal (Deaf) (Mann-Whitney U-tests; n=5, Sham; n=5, Deaf; p-values reported on graphs).

After surgery the post-operative songs of Sham mice appeared normal (Figure 36a). Eight months after deafening, however, the songs of the Deaf mice appeared spectrally distorted with many noisy looking syllables (Figure 36b). Though acoustically distorted, syllables of all major types were identified in the songs of both Deaf and Control groups in all post-surgery months recorded (Figure 37). Unlike deafening in early development, deafening in adulthood did not affect the composition of the repertoire. At 8 months after surgery the relative proportions of the major syllable types A, B, E, and G did not differ between Sham and Deaf groups (Mann-Whitney U-tests; n=5 per group; A, $p>0.5$; B, $p=0.175$; E, $p>0.5$; G, $p=0.175$). Although only representing a small part of the repertoire, it was primarily the complex syllables of Type B, E, and G that appeared degraded in sonograms of Deaf mice from the later recording months.

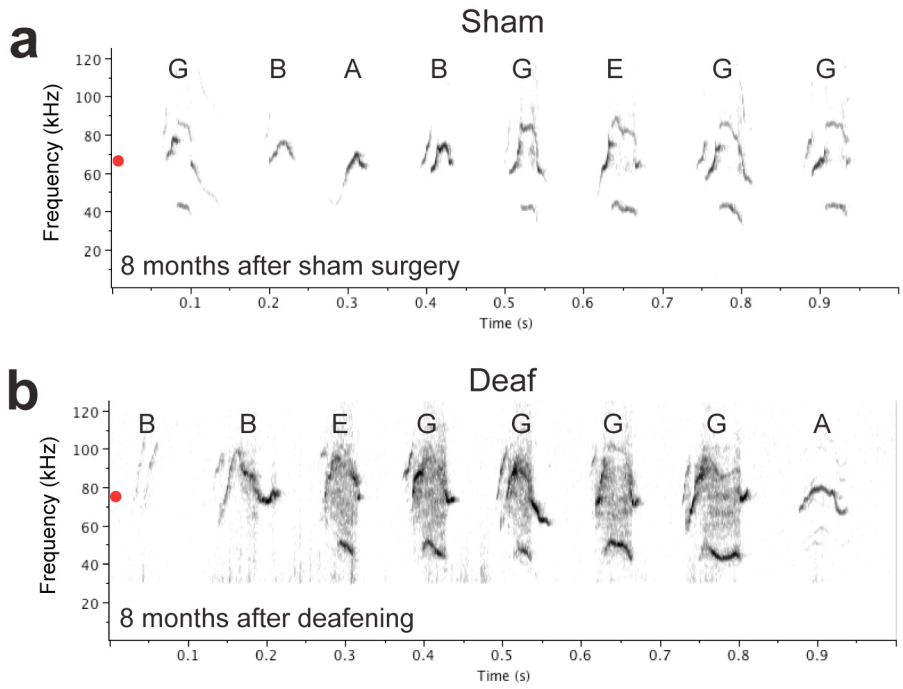

Figure 36: Sonograms representing 1 second of ultrasonic song from adult mice 8 months after sustaining (a) sham surgery (Sham) or (b) bilateral cochlear removal (Deaf). The pitch-shifted recordings corresponding to these sonograms can be heard in Supplementary Audios 10 - 11. Red dots represent the average pitch over the entire recording session for that individual animal.

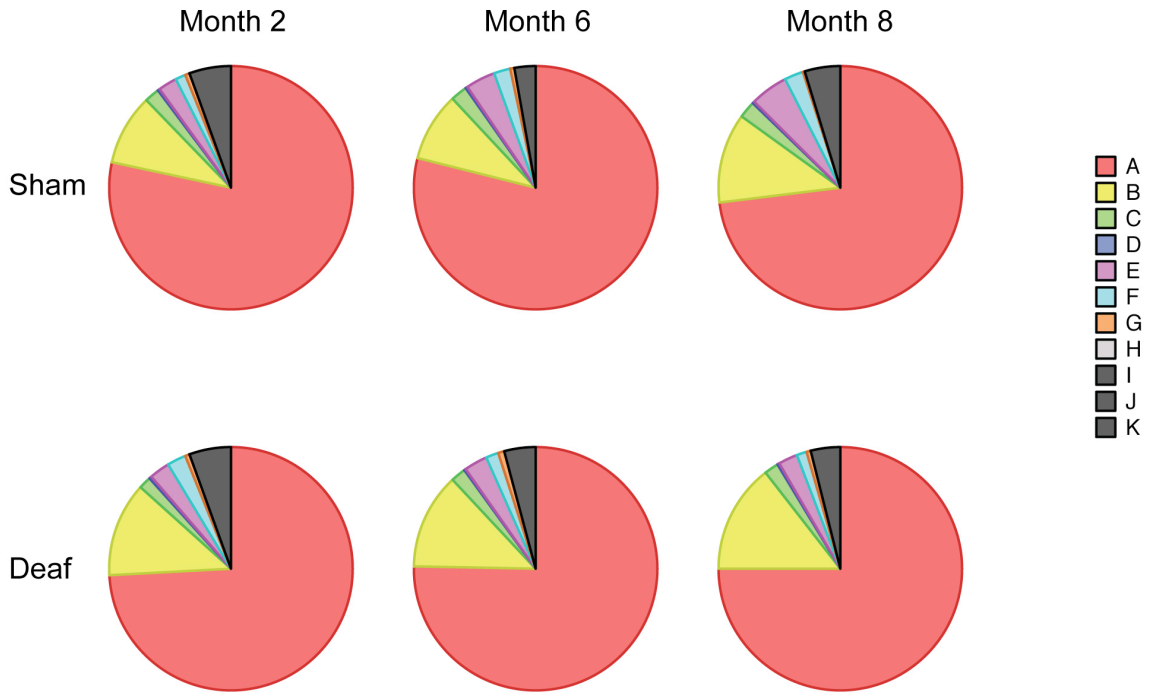

Figure 37: Composition of the repertoire by syllable type of adult mice 2, 6, and 8 months after sustaining sham surgery (Sham) or bilateral cochlear removal (Deaf).

Confirming the observation made from visual inspection of post-operative sonograms, Type A, B, and E syllables from Deaf mouse songs had significantly lower spectral purity than those from Sham songs 6 - 8 months after surgery (Figure 38; Student's t-tests; A: n=10 per group; p=0.016; B: n=10 per group; p=0.022; E: n=8, Sham; n=9, Deaf; p=0.05). Because some mice did not produce enough E Type syllables for analysis in post-operative month 6 or 8, data from both months were pooled so that every mouse would be represented in the data set. Consistent with the quantitative analysis, qualitatively the syllables sounded degraded in pitched-shifted audios (compare Supplementary Audios 10 & 11).

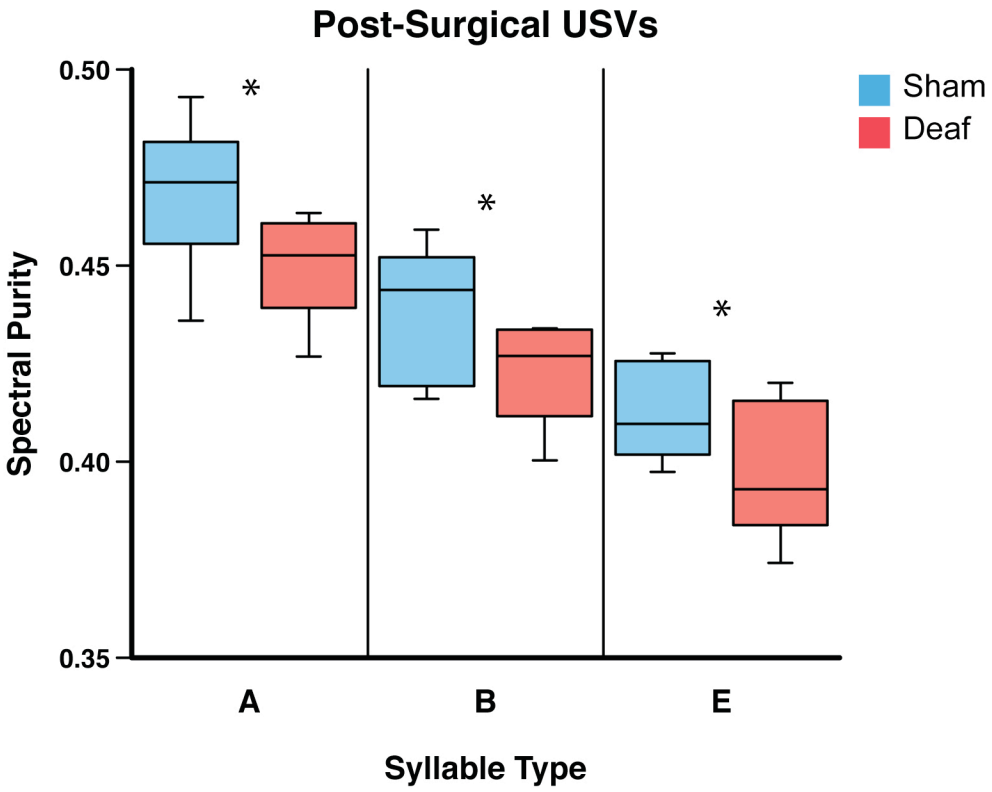

Figure 38: Spectral purity of most frequently produced syllable types A, B, and E in the late post-surgical (Months 6 & 8 post-op) songs of mice that sustained sham surgery (Sham) or bilateral cochlear removal (Deaf) (Student's t-tests; A: n=10 per group; B: n=10 per group; E: n=8, Sham; n=9, Deaf; * = p≤0.05).

Analysis of the most abundant syllable Type A revealed that some pitch-derived spectral features changed gradually over time. The mean pitch of Type A syllables in songs from the Deaf group increased progressively over the course of 8 months (Figure 39a). Two-way Mixed ANOVA revealed main effects of treatment and recording month, and a significant interaction between the two factors on the SFS for mean pitch (Two-way mixed ANOVA; n=5 per group; Factor A=Treatment, Factor B=Month; A, $F=7.17$, $p=0.028$; B, $F=4.76$, $p=0.0097$; AxB Interaction, $F=3.47$, $p=0.0319$). Comparing within-group means across all recording months showed that SFS for mean pitch of Type A syllables sung by the Deaf group was significantly higher 6 - 8 months after deafening than before deafening (Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Pre vs. Month 2, $p=0.057$; Pre vs. Month 6, $p<0.001$; Pre vs. Month 8, $p<0.001$). By 6 months after deafening, the mean pitch stopped increasing and remained stable at the elevated SFS in post-operative month 8 (Bonferroni-Dunn post-hoc test; Adjusted Alpha = 0.0083; Month 6 vs. Month 8, $p>0.5$). By contrast, the SFS for mean pitch of Type A syllables from Sham mice songs did not vary significantly between any of the months during the same period (Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Pre vs. Month 2, $p>0.5$; Pre vs. Month 6, $p>0.5$; Pre vs. Month 8, $p>0.5$; Month 2 vs. Month 6, $p=0.407$; Month 2 vs. Month 8, $p>0.5$; Month 6 vs. Month 8, $p>0.5$). Additionally, directly comparing the SFS for mean pitch between Deaf and Sham groups 8 months after surgery showed that the Deaf scores were significantly higher than those of Sham mice (Mann-Whitney U-test;

n=5 per group; p=0.028). In effect, the pitch of Deaf mice songs had increased to the point that 6 - 8 months after surgery they were reliably singing at a significantly higher frequency relative to both their own pre-deafening levels and those of hearing-intact controls. A similar gradual increase was observed in the SFS for the standard deviation of the pitch distribution for Type A syllables sung by Deaf but not Sham mice (Figure 39b; Two-way mixed ANOVA; n=5 per group; Factor A=Treatment, Factor B=Month; A, F=4.48, p=0.0672; B, F=7.28, p=0.0012; AxB Interaction, F=1.15, p=0.3499; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Deaf: Pre vs. Month 2, p=0.246; Pre vs. Month 6, p=0.008; Pre vs. Month 8, p=0.002; Sham: Pre vs. Month 2, p=0.126; Pre vs. Month 6, p=0.048; Pre vs. Month 8, p=0.028).

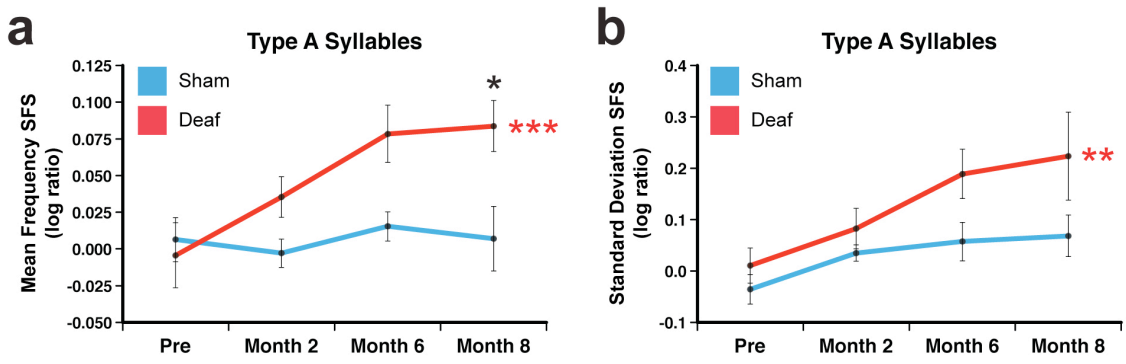

Figure 39: Mean frequency (a) and standard deviation (b) of Type A syllables from Sham and Deaf groups (expressed as spectral feature scores, SFS, a log-ratio) over an 8 month post-operative period. Repeated measures ANOVA with the Bonferroni-Dunn post-hoc test comparing within-group means across recording months (n=5 per group; Deaf: red ** = $p<0.01$, *** = $p<0.001$). Mann-Whitney U-tests comparing the mean differences between both treatment groups 8 months after surgery (n=5 per group; black * = $p<0.05$).

Several other pitch-based features showed a main effect of recording month on mean SFS, including bandwidth, starting, ¼, ¾, and minimum frequency (Figure 40; Two-way mixed ANOVA; n=5 per group; Factor A=Treatment, Factor B=Month; Starting Frequency: A, F=5.27, p=0.0509; B, F=8.86, p=0.0004; AxB Interaction, F=2.82, p=0.0601; ¼ Frequency: A, F=8.51, p=0.0194; B, F=7.25, p=0.0013; AxB Interaction, F=3.79, p=0.0233; ¾ Frequency: A, F=4.77, p=0.0604; B, F=3.44, p=0.0327; AxB Interaction, F=3.01, p=0.0501; Final Frequency: A, F=3.37, p=0.1039; B, F=2.86, p=0.0582; AxB Interaction, F=2.82, p=0.0607; Min Frequency: A, F=6.16, p=0.0380; B, F=7.86, p=0.0008; AxB Interaction, F=2.92, p=0.545; Max Frequency: A, F=6.09, p=0.0389; B, F=2.62, p=0.0741; AxB Interaction, F=3.55, p=0.0295; Bandwidth: A, F=0.24, p>0.5; B, F=3.60, p=0.0281; AxB Interaction, F=2.00, p=0.1413; Frequency Variance: A, F=0.53, p=0.4866; B, F=2.26, p=0.1068; AxB Interaction, F=1.48, p=0.2451). Additionally, mean SFS for starting, ¼, minimum, mean, and maximum frequency was significantly higher 8 months after deafening than before deafening (Figure 40; Bonferroni-Dunn post-hoc tests; n=5 per group; Adjusted Alpha = 0.0083; p-values reported on the graph). These effects on pitch were not observed in the Sham controls whose SFS did not vary significantly after surgery (Figures 39 and 40). By 8 months after surgery, the mean SFS's for starting, ¼, ¾, minimum, mean, and maximum frequency in the Deaf group had also diverged from the Sham group, but amplitude had not (Mann-Whitney U-tests; n=5 per group; Starting Frequency, p=0.028; ¼ Frequency, p=0.016; ¾ Frequency, p=0.076; Final Frequency,

p=0.076; Min Frequency, p=0.028; Mean Frequency, p=0.028; Max Frequency, p=0.047; Bandwidth, p=0.251; Frequency Variance, p=0.175; Amplitude, p>0.5).

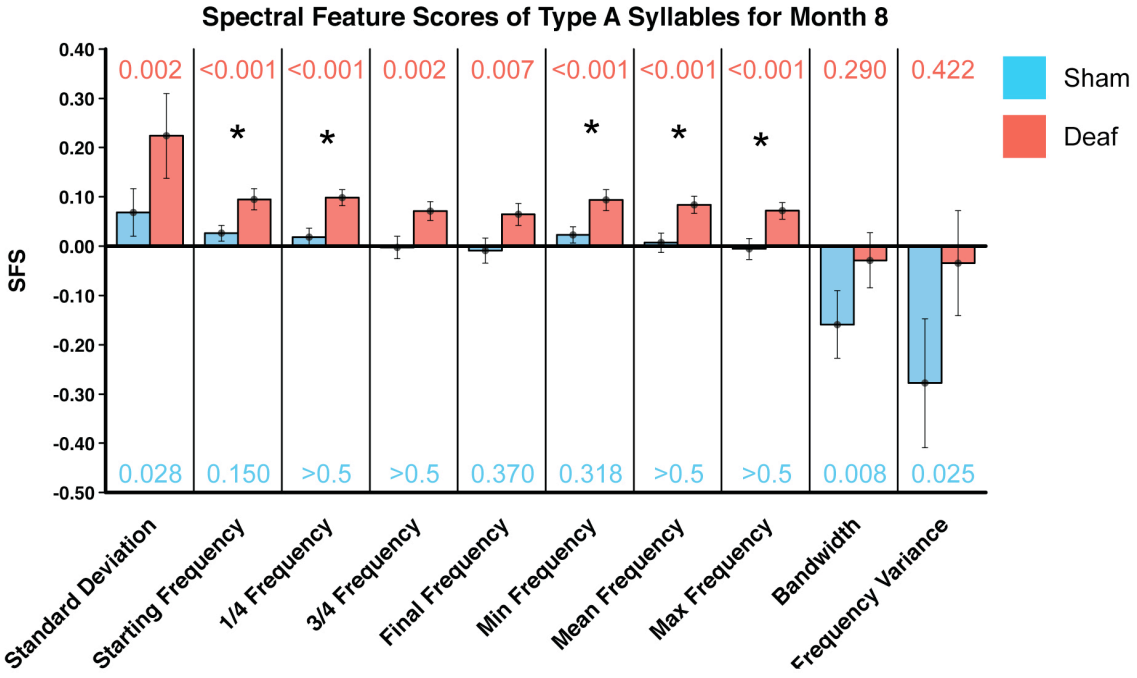

Figure 40: Post-surgical spectral features (SFS) of Type A syllables produced by adult mice 8 months after sustaining sham surgery (Sham) or bilateral cochlear removal (Deaf). The SFS for many features of Deaf mice increased relative to their own scores before deafening (Bonferroni-Dunn post-hoc tests; Pre vs. Month 8; n=5 per group; Adjusted Alpha = 0.0083; Deaf: red p-values; Sham: blue p-values). Starting, ¼, Min, Mean, and Max Frequency also diverged from the Sham group (Mann-Whitney U-tests; n=5 per group; * = p<0.05).

5.3.3 Hereditary Deafness

Although the songs of P12-deafened mice were abnormal, they still contained recognizable elements of mouse song. Given that the P12-deafened mice only suffered conductive hearing loss, I wondered if there could have been some early auditory experience that influenced vocal development. To account for this possibility, I analyzed and compared the songs of normal hearing-intact C57 males to songs of males congenitally deaf due to loss of inner ear hair cells and spiral ganglion neurons 2 – 5 weeks after birth resulting from knockout of the caspase-3 gene (Morishita et al., 2001; Takahashi et al., 2001) on a C57 background. Caspases are a highly conserved class of cysteine-containing, aspartate-specific, pro-apoptotic proteases that cleave specific substrates in the nucleus and cytoplasm during programmed cell death (PCD) (Nicholson, 1999). Caspase-3 is involved in targeted cleavage of specific protein substrates during the penultimate step of PCD (Li and Yuan, 2008). Knockout of caspase-3 (CASP3 KO) in mice has been shown to result in strain-dependent effects on brain development (Leonard et al., 2002). On the 129X1/SvJ (129sv) background, CASP3 KO was lethal and caused severe exencephaly. By stark contrast, caspase-3-deficient C57 mice showed normal brain development. Normal brainstem development was also described for CASP3 KO on a mixed C57/129sv strain (Oppenheim et al., 2001). Normal brain development in the C57 animals can be accounted for by increased activity of caspase-7 that was also efficient at cleaving inhibitor of caspase-activated DNase and

achieving DNA fragmentation in the C57 strain (Houde et al., 2004). It is unlikely, therefore, that differences found in vocal production between wildtype and caspase-3-deficient C57 mice would be caused primarily by non-specific brain deficits. Caspase-3 has, however, been implicated in skeletal muscle differentiation (Fernando et al., 2002), and it is not known whether caspase-7 also plays a compensatory role in non-neural tissues.

Striking differences were observed between the songs of C57 and CASP3 KO mice. Many of the syllables in CASP3 KO songs were highly degraded and barely recognizable visually, but our automated detection based on pitch jumps was able to distinguish standard syllable categories (Figure 41) indicating some genetically innate component to mouse song that does not depend on auditory experience. Though severely degraded, I interpret these deteriorated vocalizations as songs, because they were observed specifically when mice were stimulated to vocalize with female urine. The gross temporal structure of CASP3 KO songs did not appear to have been affected. Songs of these deaf mice lowered to the human hearing range sounded like squawks and screams rather than whistles (compare Supplementary Audios 12 & 13 with Supplementary Audios 14 & 15).

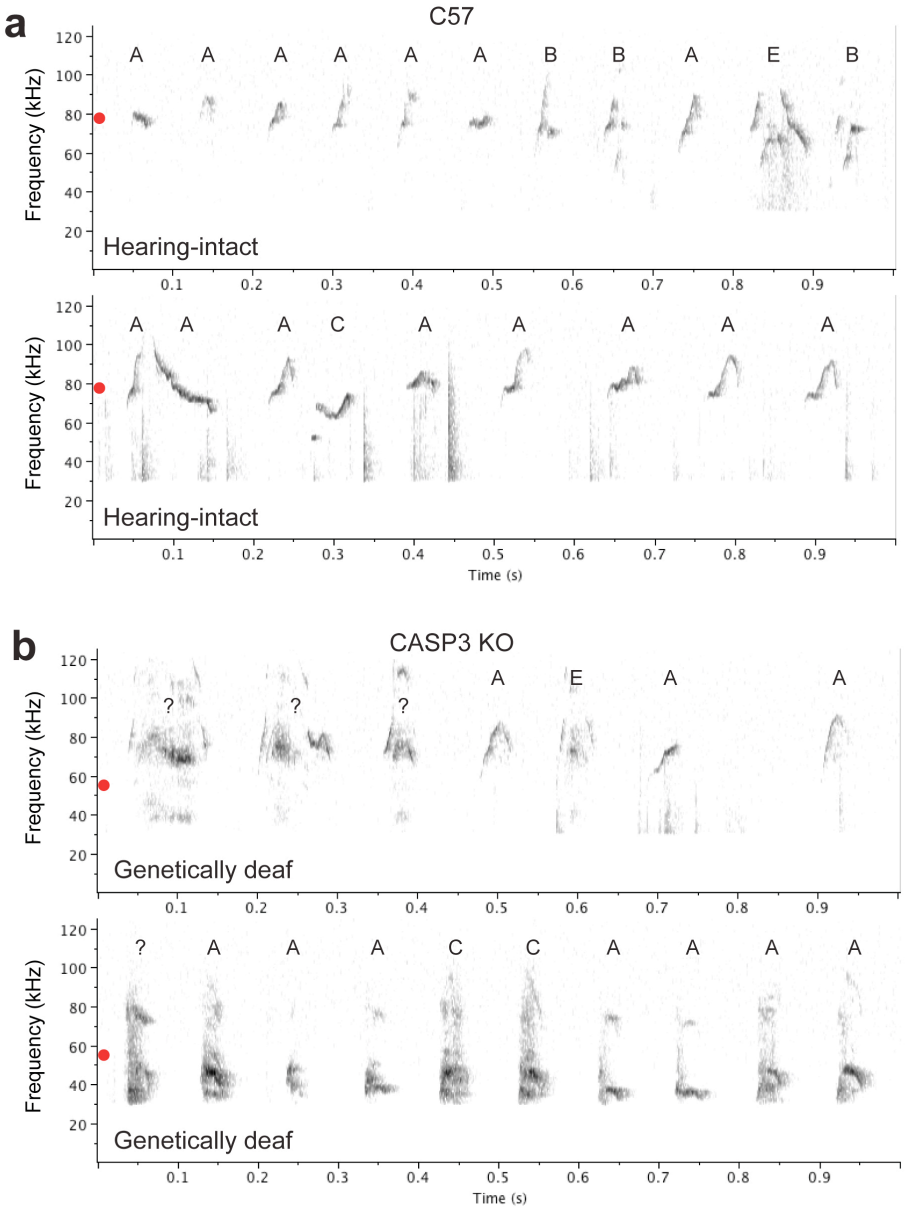

Figure 41: Sonograms representing 1 second of ultrasonic song from adult C57 wildtype (a) and caspase-3-deficient (CASP3 KO) (b) male mice. The pitch-shifted recordings corresponding to these sonograms can be heard in Supplementary Audios 12 - 15. Red dots represent the average pitch over the entire recording session for that individual animal.

Examples of all major syllable types were recognized in songs from both C57 and CASP3 KO mice (Figure 42a). Type A syllables were the most frequently produced by both strains, but significantly over-represented in the CASP3 KO repertoire similar to the effect on deafened BxD juvenile songs (Figure 42b; Mann-Whitney U-test; C57, n=8; CASP3 KO, n=6; p=0.014). On average, Type A syllables made up 78 ± 3.6% of all identified notes in CASP3 KO songs versus 59 ± 5.4% in C57 songs. Also similar to the effect in deafened BxD juveniles, the representation of some complex syllable types decreased in the CASP3 KO repertoire. Specifically, the percentage of Type B syllables was significantly lower in CASP3 KO songs than C57 songs. The percentage of Type E syllables did not differ between groups (Mann-Whitney U-Tests; C57, n=8; CASP3, n=6; B, p=0.007; E, p=0.366).

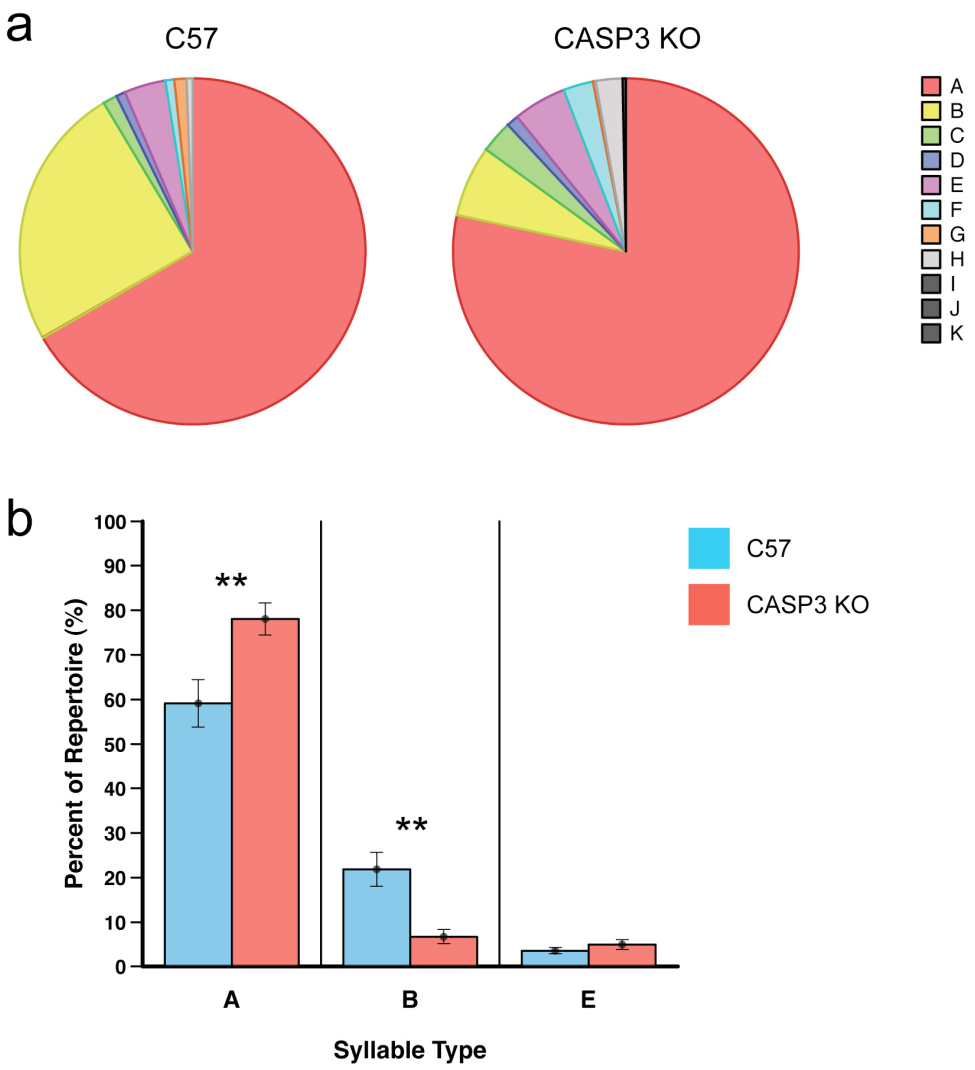

Figure 42: a, Composition of repertoire by syllable type for wildtype (C57) and caspase-3-deficient (CASP3 KO) adult male mice. b, Type A, B, and E syllables as a percentage of all identified notes produced by C57 and CASP3 KO adult males (Mann-Whitney U-Tests; C57, n=8; CASP3 KO, n=6; ** = p<0.01).

I found that all spectral features whose SFS changed after deafening in juvenile and adult BxD males also varied significantly between CASP3 KO and C57 mice (Figure 43). The magnitude of the effect was always larger in congenitally versus mechanically deaf mice, but sometimes in the opposite direction. For example, the mean frequency of CASP3 KO Type A syllables was lower than hearing-intact controls, not higher as had been observed for mechanically deafened BxD adults (Mann-Whitney U-test; C57, n=8; CASP3 KO, n=6; p=0.005). On average, there was a 17.1 kHz (23%) difference between the mean frequency of C57 and CASP3 KO Type A syllables measured at 75.8 ± 0.91 kHz and 58.7 ± 4.2 kHz, respectively. Similar to BxD males deafened as adults, the standard deviation of the pitch distribution was greater for CASP3 KO Type A syllables (7.5 ± 0.80 kHz for C57 vs. 11.7 ± 1.3 kHz for CASP3 KO). All of the pitch-based features analyzed, except for bandwidth, differed significantly between C57 and CASP3 KO groups (Mann-Whitney U-Tests; C57, n=8; CASP3 KO, n=6; Standard Deviation, p=0.28; Starting Frequency, p=0.007; ¼ Frequency, p=0.007; ¾ Frequency, p=0.002; Final Frequency, p=0.002; Minimum Frequency, p=0.007; Maximum Frequency, p=0.002; Bandwidth, p=0.302).

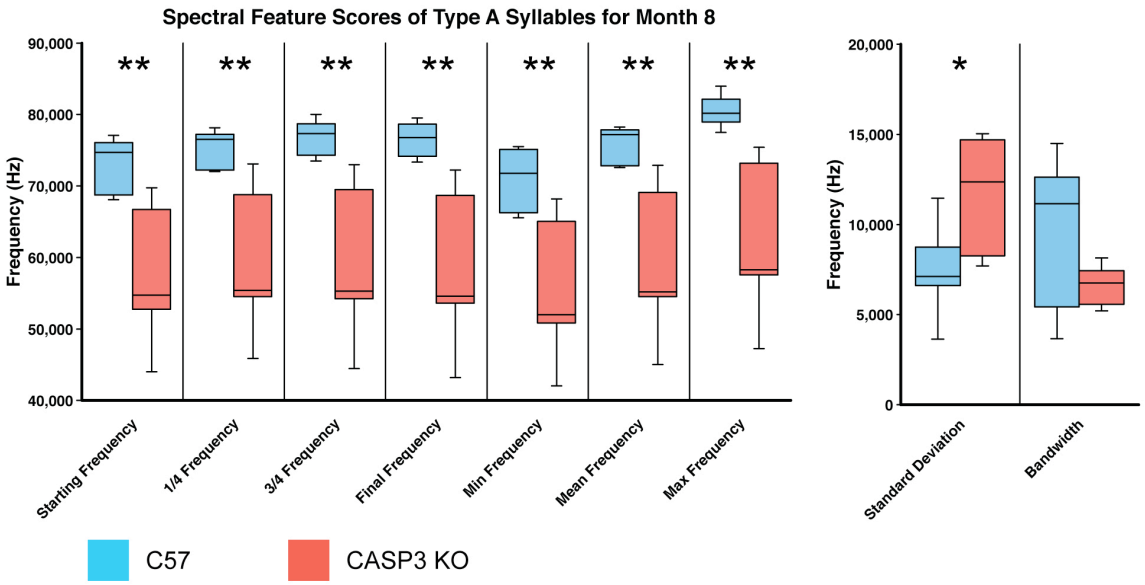

Figure 43: Spectral features of Type A syllables produced by wildtype (C57) and caspase-3-deficient (CASP3 KO) adult males (Mann-Whitney U-Tests; C57, n=8; CASP3 KO, n=6; * = p<0.05, ** = p<0.01).

The results from the CASP3 KO mice are consistent with the hypothesis that early auditory experience contributes to song development. When early experience is prevented by a complete loss of auditory transduction due to loss of inner ear hair cells, the effects on songs were stronger than those observed after conductive hearing-loss on P12. Importantly, all of the features affected by mechanical deafening in juveniles and adults were also susceptible to CASP3 KO, including repertoire composition, standard deviation of pitch distribution, various pitch-based features including mean frequency. This could suggest that the main effect on vocalization is due to loss of hearing; however, it must be noted that although no motor deficits have been reported in caspase-3-deficient C57 mice, we were not equipped to test for them in our study. I did not note overt changes in motor behaviors in CASP3 KO mice, suggesting that the changes in pitch may not be attributed to an obvious general motor deficit. But without rigorous testing of motor abilities, it remains possible that the effect was due to non-specific motor deficits caused by the effects of CASP3 KO on skeletal muscle differentiation. Mutation or knockout of two other genes, FoxP2 and alpha-synuclein, led to changes in amount of ultrasonic vocalization produced by mice (Enard et al., 2009; Kurz et al., 2010) but nothing like the dramatic spectral changes observed in CASP3 KO mice. It will be important to test other strains with hereditary early-onset hearing loss to determine if this is a general effect on all mice deprived of auditory feedback in early development.

Overall, the effects of deafening on adult song are consistent with the hypothesis that male mice rely on auditory feedback to regulate the fine structure of their songs. Type A syllables, representing most of the vocal output of each mouse, changed in pitch and degraded spectrally over the course of 6 months after deafening. Although the general effect on songs was similar in some ways to that of cortical lesions, the rate of SFS increase following mechanical deafening was comparably slower. Pitch increases and spectral degradation took months in deafened adults, whereas significant effects were observed weeks after lesions. Additionally, the pitch was more strongly affected by deafening than lesions in that more spectral features changed significantly. For the final pitch of Type A syllables, the one directly pitch-based feature that changed significantly in both experiments, the magnitude of the change was very similar. In deafened animals final frequency increased from 67.5 ± 1.4 kHz (s.e.m.) to 72.1 ± 1.0 kHz (s.e.m.), which was closely matched by an increase from 68.3 ± 0.78 kHz (s.e.m.) to 72.2 ± 1.0 kHz (s.e.m.) in cortically lesioned mice. Mean frequency, a more intuitive measure of pitch, increased more than final pitch in post-deafening adult BxD songs from 65.1 ± 0.83 kHz (s.e.m.) to 71.1 ± 1.0 kHz (s.e.m.). The 6 kHz increase in mean pitch is comparable to changes in USVs reported for deafened horseshoe bats, a confirmed vocal learning species (Rübsamen and Schäfer, 1990). Deafening by puncturing the cochlea in bats resulted in an increase between 4 and 6 kHz in the mean pitch of the constant frequency component of ultrasonic echolocation pulses. Overall, the combined effects on pitch and

spectral purity were similar in character and timing to changes in vocalizations observed in post-lingually deaf humans and mechanically deafened song-learning birds (Waldstein, 1990; Woolley and Rubel, 1997; Heaton et al., 1999; Brainard and Doupe, 2000; Watanabe et al., 2006).

Chapter 6. Cultural Transmission of Vocal Features

The reliance on auditory information for normal production of adult courtship songs was suggestive of a possible capacity for adaptive vocal modification. To date, destabilization of vocal production after deafening has only been observed in vocal learners; however, these observations remain correlative and not diagnostic. Thus, the gold standard for demonstrating vocal learning has been the transfer of vocal elements between individuals. In songbirds, this is typically observed as the intergenerational transfer of song from father to male offspring. In humans we can see the parallel process of children learning language-specific speech elements from their parents and peers. In parrots, even heterospecific sounds like human speech can be copied. While vocal learning is clearly evident in this type of overt mimicry, some researchers argue that a more limited form of vocal plasticity should be considered whereby the spectral content of existing calls of conspecifics converge (Egnor and Hauser, 2004; Tyack, 2008). While there are reports of such call convergence in adult non-human primates, much of the evidence is anecdotal or limited by small data sets and incomplete knowledge of the vocal ontogeny (Tyack, 2008). Although no totally new calls are generated, convergence does require the transfer of vocal elements between individuals and may reflect a rudimentary ability that could have been expanded to include production of novel elements. Therefore, I wondered if mice were capable of any type of adaptive vocal

flexibility, whether to the limited extent reported in non-human primates or to a greater extent reflective of full vocal mimicry.

6.1 Background – Vocal Imitation

6.1.1 Vocal Imitation in Birds

Among the songbirds some species like the junco learn without reference to external models (Marler et al., 1962) but still require their own auditory feedback to develop normal songs (Konishi, 1964). Other songbird species like the white-crowned sparrow (Marler and Tamura, 1964) mimic an externally acquired song model. Transfer can occur from father to son, between mates, and even between foster-fathers or affiliated heterospecific individuals. Importantly, imitation often falters if the requisite social conditions are not met. For example, captive white-crowned sparrows learned better when tutored with live birds than when hearing only playbacks of taped songs (Baptista and Petrinovich, 1984; Baptista and Petrinovich, 1986). The most prolific learners are capable of mimicking sounds even without the need for any social bonds. For example, mockingbirds can pick up the songs of other birds from many different species in the surrounding environment (Allard, 1939), and the lyrebird even copies non-biological sounds from its surroundings with high fidelity (Armstrong, 1963).

The budgerigar, an open-ended vocal learning species, learns both a long non-stereotyped warble song and contact calls (Brittan-Powell et al., 1997). Interestingly, in this species males imitate the calls of females upon forming a pair bond (Hile et al.,

2000). Call convergence is also likely to occur in hummingbirds. Male Anna's hummingbirds (*Aulacorhynchus prasinus*) species sing in congregations, and within an assembly each male has a song resembling that of its neighbors (Baptista and Schuchmann, 1990). Thus, vocal learning in birds is not limited to tutored mimicry and includes a wide assortment of different learning strategies including improvised self-learning, call convergence, mimicry of socially bonded conspecifics, mimicry of heterospecifics, and even mimicry of unnatural sounds.

6.1.2 Vocal Learning in Non-Human Mammals

Like in vocal learning birds, vocal learning in mammals is expressed in a variety of ways. Particularities of the learning process vary between different species, but the important point is that mammals use a wide array of learning strategies including call convergence, mimicry, and reinforced improvisation. The strongest experimental evidence for vocal learning in non-human mammals comes from bottlenosed dolphins (*Tursiops truncates*) that can modify their whistles through experience. Dolphins have been trained to imitate artificially designed sounds using reinforcement learning, and even learned to attach novel sounds to objects as labels (Richards et al., 1984). Whistle mimicry also occurred spontaneously in some cases where dolphins were given control over the rate of target sound presentation (Reiss and McCowan, 1993). Like dolphins, pinnipeds are also capable of learning new vocalizations when trained using food rewards as reinforcement (Schusterman and Reichmuth, 2008). In a contingency

learning paradigm walruses were rewarded only when they generated vocalizations significantly different from the preceding one, and over time were able to produce entirely novel sounds. African elephants (*Loxodonta africana*) are able to imitate the sounds of Asian elephants (*Elephas maximus*) and even trucks (Poole et al., 2005).

For call convergence in mammals, the strongest evidence comes from bats. The frequency of echolocation calls of young greater horseshoe bats (*Rhinolophus ferrumequinum*) correlate strongly with the calls of their mother (Jones and Ransome, 1993). Because the frequency of a mother's calls varies with her age, the correlation with her offspring's frequency is likely to result from their learning her pitch. When female greater spear-nosed bats (*Phyllostomus hastatus*) were transferred to a new social group both the residents of the group and the new members changed the spectro-temporal features of their existing screech calls to converge on a similar call (Boughman, 1998). Call convergence is also thought to occur in non-human primates based mostly on observations of within-group similarity and geographical variation in call features (Janik and Slater, 1997; Tyack, 2008); although, some experimental evidence has been reported for pygmy marmosets (*Cebuella pygmaea*) that minimized spectral differences between each other's calls when new male/female pairs were housed in a cage together (Snowdon and Elowson, 1999). Given the range of mammalian vocal learning strategies from mimicry to call convergence, it is important to test which, if any, of these abilities are also present in mice.

6.2 Call Convergence in Laboratory Mice

The presence of a direct motor cortical vocal pathway and the contribution of auditory feedback to normal vocal output are suggestive of a capability for adaptive modification. However, neither provides any clues as to how this capability might be used to actually modify naturally produced USVs. To test the hypothesis that male mice are capable of copying specific features of another male's songs, I cross-housed pairs of BxD and C57 males and analyzed their songs for transfer of species-specific spectral features.

6.2.1 Methods – Cross-Housing and Song Analysis

6.2.1.1 Housing

The BxD and C57 strains were chosen because during previous recordings I noticed that they sing in different frequency bands, with C57 songs typically being of a higher mean frequency than those of BxD males. I took advantage of the reliable difference in mean frequency to test whether strain-specific features could be transferred between genetically distinct males.

Seven cross-strain pairs were generated, comprised of 85 day old BxD and C57 males (n=7 pairs). Prior to pairing, all males were group housed with other males of the same strain in acoustic isolation from other cages, and baseline recordings were made for two weeks. A single sexually experienced BxD female was also housed with each pair, and the trio was kept in acoustic isolation from other cages. The females were

removed two days prior to each post-pairing recording session in order to collect urine for stimulating USV production. Over the course of 8 weeks, each male was recorded individually then repaired with the same cage-mate as before and a BxD female. The composition of the male pairs was stable throughout the experiment, but the identity of the female was not recorded from week to week.

6.2.1.2 Song Recordings

Each mouse was placed singly into a dark 15″x24″x12″ sound-attenuating recording chamber, and ultrasonic songs elicited by presentation of 200 μL of freshly collected female urine were recorded with UltraSoundGate CM16/CMPA ultrasound microphones and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). Details of the recording procedure are described in the General Methods (Appendix A, Section A.2).

6.2.1.3 Song Analysis

Sonograms were computed from each waveform (256 samples/block, half-overlap), and the following spectral features were calculated from the sonograms for each syllable: standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. Details of song analysis are described in the General Methods (Appendix A, Section A.3).

6.2.2 Results

Before crossing, the repertoires of BxD and C57 males were normal and displayed the typical dominance of Type A syllables (Figure 44) whose representation did not differ between the strains (data not shown; Mann-Whitney U-test; n=7 per group; p=0.277).

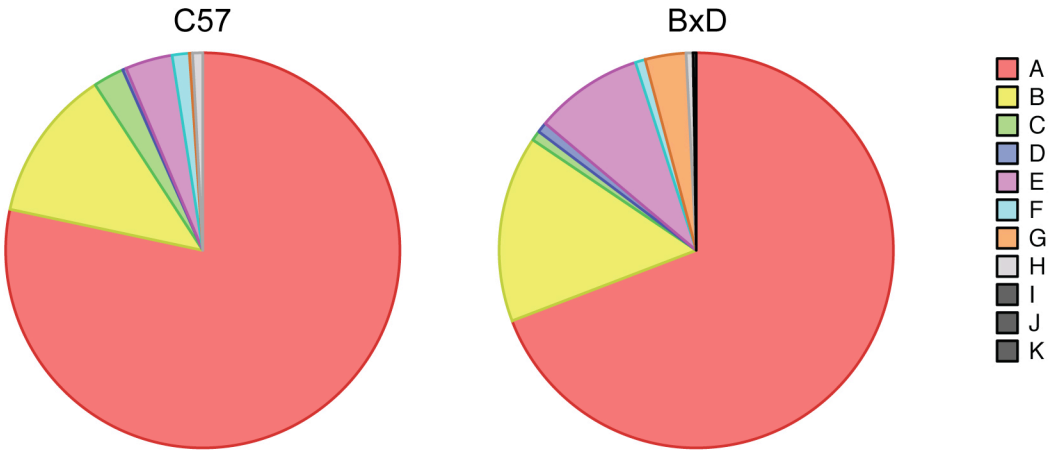

Figure 44: Composition of repertoire by syllable type for C57 and BxD males before cross-housing in multi-strain pairs.

Before crossing, the pitch of Type A syllables from seven C57 and seven BxD males segregated into two non-overlapping distributions (Figure 45; Mann-Whitney U-test; n=7 per group; Pre: p=0.002). After crossing three outcomes were possible: 1) no transference; 2) mice of both strains could change resulting in a variable individual effect, but no systematic change; 3) one strain could reliably learn the pitch of the other. Consistent with the third outcome described above, after cross-housing I observed a significant decrease in the pitch of C57 Type A syllables such that after 8 weeks of cross-housing the pitches of BxD and C57 songs were no longer statistically distinguishable (Figure 45; Mann-Whitney U-test; Week 1: C57, n=4; BxD, n=5; p=0.086; Week 2: C57, n=4; BxD, n=7; p=0.008; Week 3: C57, n=5; BxD, n=7; p=0.004; Week 4: C57, n=6; BxD, n=7; p=0.015; Week 5: C57, n=6; BxD, n=6; p=0.004; Week 6: C57, n=7; BxD, n=7; p=0.013; Week 7: C57, n=5; BxD, n=6; p=0.018; Week 8: C57, n=4; BxD, n=7; p=0.131). On average, over 8 weeks the mean pitch of C57 males had decreased from 76.8 ± 0.73 kHz to 71.9 ± 1.1 kHz, and the pitch of BxD males increased only slightly from 66.9 ± 0.90 kHz to 69.5 ± 0.88 kHz.

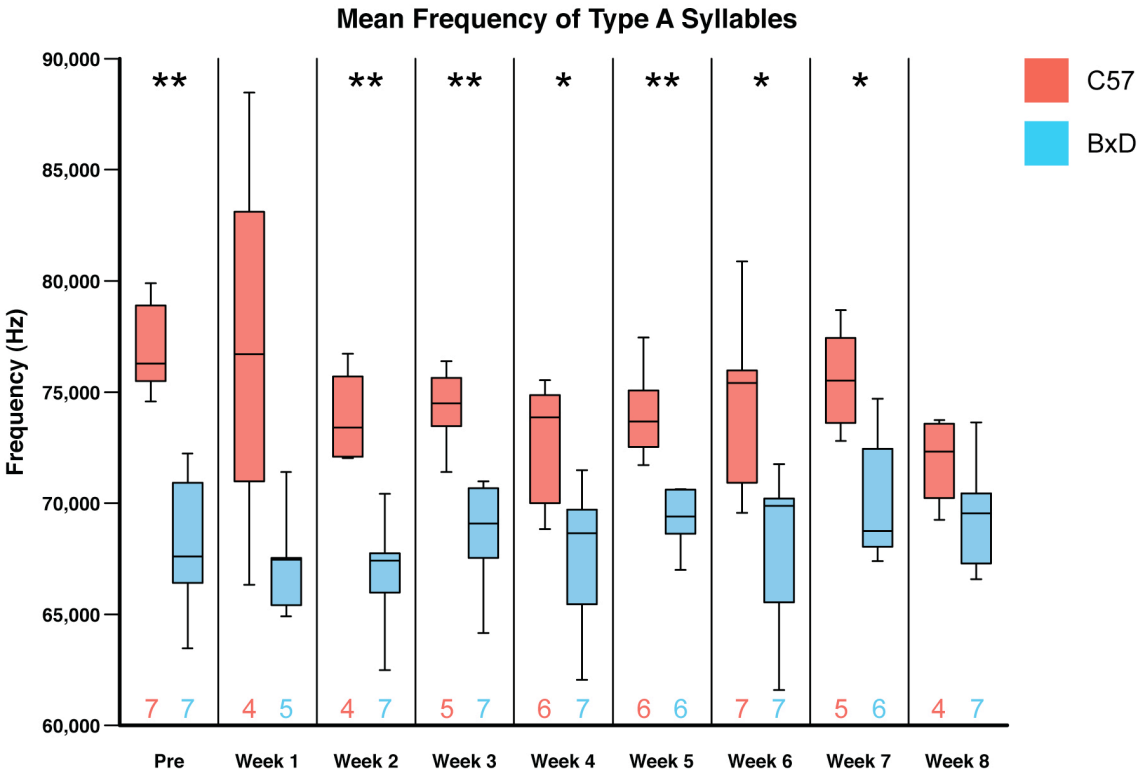

Figure 45: Mean frequency of Type A syllables from C57 and BxD males before and after cross-housing in inter-strain pairs (Mann-Whitney U-test; n's for each group are listed on the x-axis, red=C57, blue=BxD; * = p<0.05, ** = p<0.01).

In the eighth and final week of recording after crossing, only 4 of the C57 mice sang enough Type A syllables for analysis. Following these 4 mice back through previous weeks when they had all vocalized sufficiently, I confirmed that the pitch convergence in week 8 resulted mainly from changes in the songs of C57 males after crossing (Figure 46). All 4 C57 males that sang in week 8 had decreased their pitch from pre-cross-housing levels and were singing in a significantly lower frequency band as a group (Figure 46; One-way Repeated Measures ANOVA; n=4; $F=4.03$, $p=0.0450$; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Pre vs. Week 4: $p=0.039$; Pre vs. Week 6: $p=0.129$; Pre vs. Week 8: $p=0.004$). There was no significant effect on the mean pitch of BxD songs (Figure 46; One-way Repeated Measures ANOVA; n=7; $F=1.22$, $p=0.329$; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Pre vs. Week 4: $p>0.5$; Pre vs. Week 6: $p>0.5$; Pre vs. Week 8: $p=0.205$).

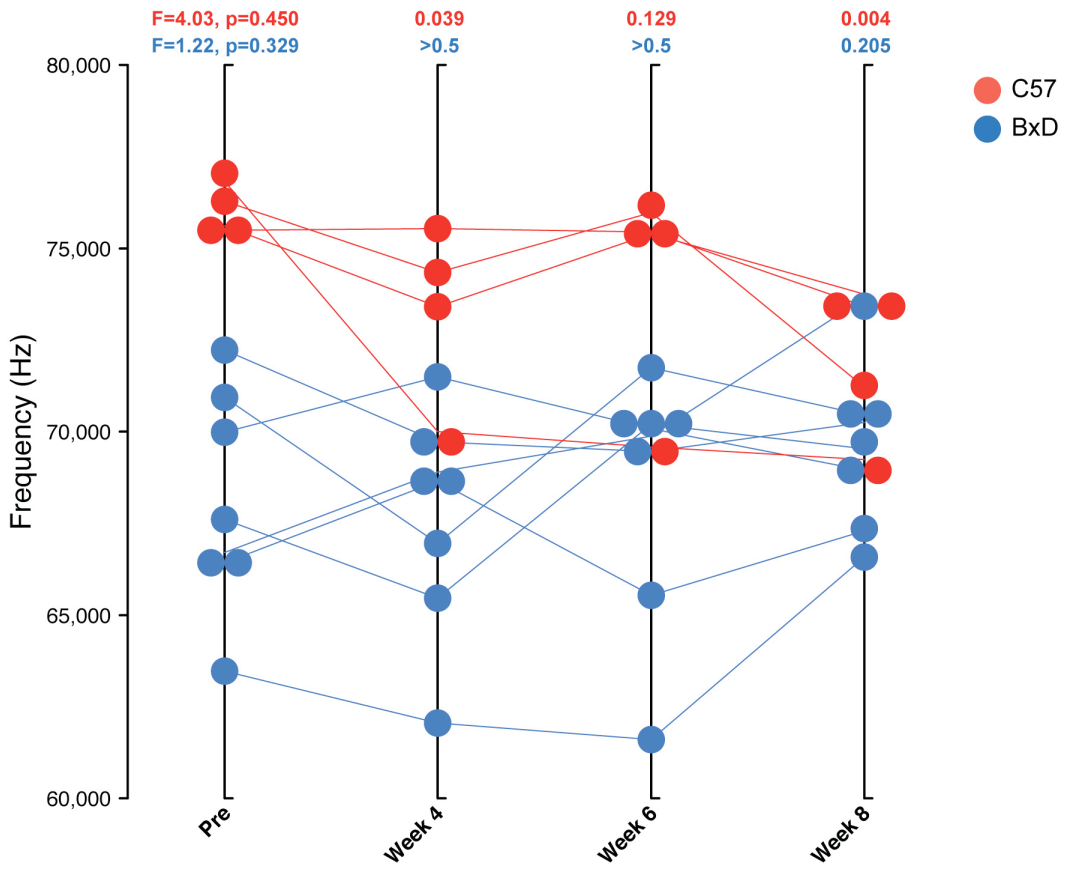

Figure 46: Mean pitch of individual C57 and BxD males before and after cross-housing in multi-strain pairs. (One-way Repeated Measures ANOVAs; C57, n=4; BxD, n=7; F and p-values are reported on the graph, red = C57, blue = BxD; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; p-vales from post-hoc tests vs. Pre are reported on the graph).

Additionally, I observed a decrease in difference in pitch between BxD and C57 males housed together (Figure 47). There was a general downward trend in the mean pitch difference among all pairs in the 8 weeks after crossing. Before crossing, the mean pitch difference between pairs was 8.6 ± 0.51 kHz. By 3 weeks after crossing the mean pitch difference had decreased significantly, and continued to decline to a global minimum difference of 2.1 ± 1.4 kHz at 8 weeks after crossing (Figure 47; Mann-Whitney U-tests; Pre, n=7; Week 2: n=4, $p>0.5$; Week 3: n=5, $p=0.028$; Week 4: n=6, $p=0.46$; Week 5: n=5, $p=0.019$; Week 6: n=7, $p=0.406$; Week 7: n=4, $p>0.5$; Week 8: n=4, $p=0.008$).

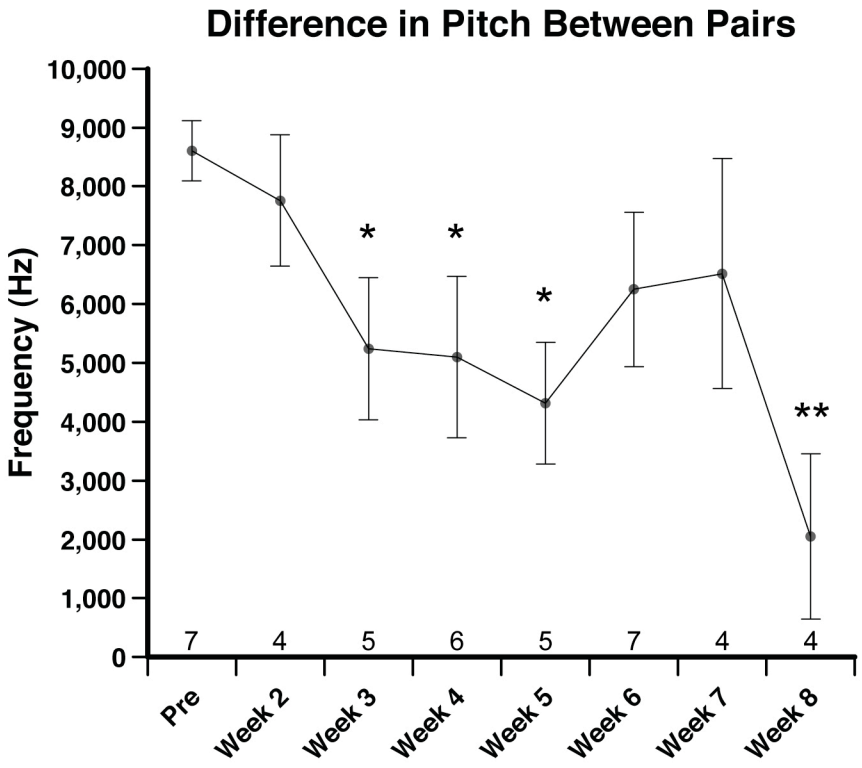

Figure 47: Mean difference in pitch between paired C57 and BxD males (Mann-Whitney U-tests; n's for each week are listed on x-axis; post-crossing weeks vs. Pre: * = $p<0.05$, ** = $p<0.01$).

Following the 4 pairs that sang sufficiently in post-cross week 8 back through previous weeks, I found that the difference in pitch had decreased for each pair (Figure 48). For these 4 pairs there was a significant effect of pairing on the difference in pitch from pre-crossing to 8 weeks post-crossing (Figure 48; One-way Repeated Measures ANOVA; $n=4$; $F=7.66$, $p=0.0076$; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; Pre vs. Week 4: $p=0.041$; Pre vs. Week 6: $p=0.041$; Pre vs. Week 8: $p<0.001$). After 8 weeks of cross-housing 3 of the 4 pairs had reduced their difference in pitch by more than 80 percent, and two of the pairs had converged to within 300 Hz of each other's pitch.

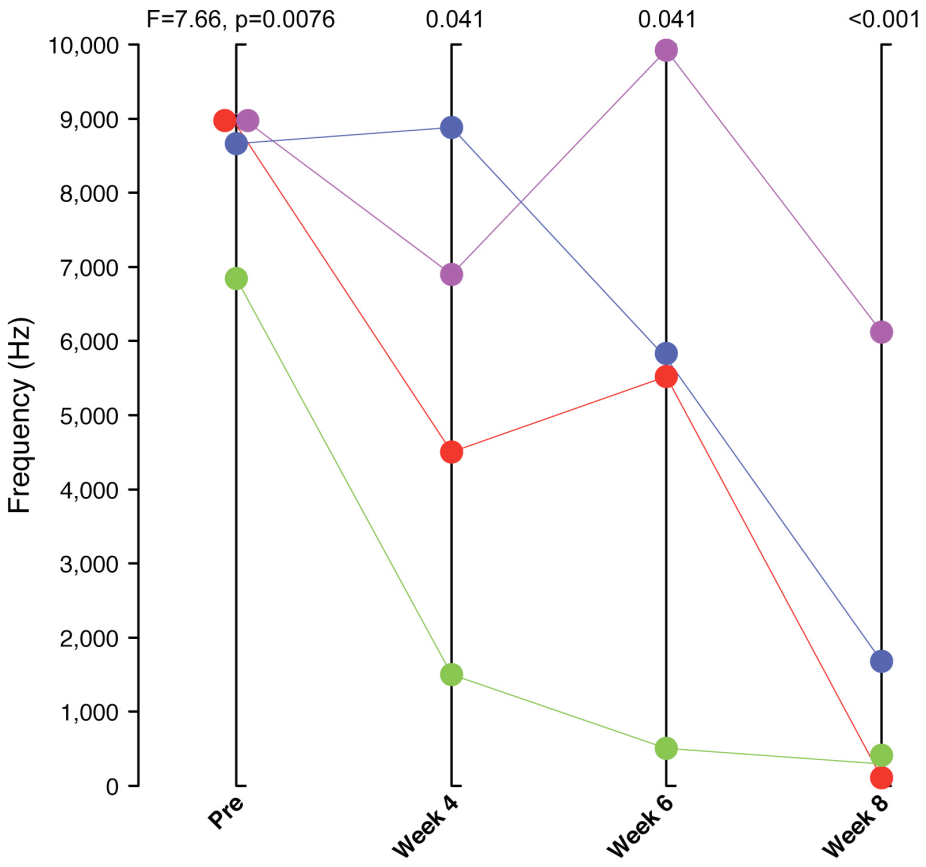

Figure 48: Difference in pitch between individual cross-housed C57+BxD pairs (One-way Repeated Measures ANOVA; n=4; F and p-values are reported on the graph; Bonferroni-Dunn post-hoc tests; Adjusted Alpha = 0.0083; post-crossing weeks vs. Pre: p-values reported on the graph).

The results of cross-housing pairs of BxD and C57 males support the hypothesis that mice are capable of copying specific features of another male's songs. Adult C57 males adaptively modified their USVs to approximate the pitch of adult males of a completely different strain. Importantly, the pitch of C57 songs both deviated from baseline levels, and converged onto the lower pitch range of BxD males. The changes observed were made to an existing note type shared between both strains. So, unlike vocal mimicry in songbirds, no novel syllables or elements were created. Therefore, the reported change is more akin to vocal convergence observed in chickadees and marmosets. It should be noted that the syllable classification scheme I used was based on a shared repertoire of the most frequently produced syllable types in BxD songs. Therefore, there could have been rare, novel, or strain-specific syllable types that were missed by this identification strategy. More careful cataloging of the entire repertoire of all mice would be necessary to confirm or deny *de novo* generation of song elements.

It is curious that C57 males changed as a group but the BxD males were relatively unaffected by crossing. One possibility is that the identity of the females co-housed with the pairs provided a selection force in the direction of the preferred range of BxD females. It would be important to test if changing the identity of the females either prevents the change in pitch or causes the BxD males to converge on the pitch of C57 males. While females could certainly provide a reinforcing stimulus for convergence, the close approximation of the BxD male's pitch in two of the four pairs

analyzed at 8 weeks post-crossing suggests that they were likely guided by auditory information. It will be important to dissect the relative contributions to pitch convergence of auditory experience, social reinforcement, and the motor cortical vocal pathway.

Chapter 7. Conclusions and Future Directions

In my dissertation research, I examined the underlying neural circuits that support production of ultrasonic courtship songs of male laboratory mice, and tested the basic capabilities of adult mice to modify the spectral content of their songs. In Chapter 2, I described limbic, motor cortical, and striatal regions that are activated during song production even in the absence of auditory feedback. In Chapter 3, I described connections of the vocal premotor network and discovered that the mouse song system includes coritical, striatal, thalamic, and brainstem circuits. Importantly, these pathways include a direct motor cortical vocal pathway to Amb. In Chapter 4, I reported that the motor cortical vocal pathway identified in Chapter 3 is not critical for the initiation of song, but necessary for maintaining a level of acoustic stereotypy. In Chapter 5, I described the role of auditory experience and feedback on song production. I found that early auditory experience was necessary for acoustic exploration and establishing a normal repertoire in juveniles, and that auditory feedback was critical for regulating the pitch and spectral purity of songs into adulthood. In Chapter 6, I reported that adult mice are capable of adopting song features from another strain. I found that adult C57 males were able to copy the pitch of BxD males with whom they shared a cage for 8 weeks in the presence of a BxD female. Taken together, my results indicate that a combination of neural and behavioral features is present in laboratory mice that had previously only been reported in humans and song learning birds.

Many aspects of vocal production described in these studies are being reported for the first time in mice. Perhaps more importantly, some are being reported for the first time in non-human mammals. The discovery of brain regions and pathways involved in song production should aid interpretation of past studies and inform the design of future studies investigating the effects of social, genetic, and pharmacological manipulation on vocal behavior. Additionally, the discovery of direct cortical input to the song system, a requirement for auditory feedback, and a capacity for adaptive vocal modification should inform studies investigating the distribution, development and evolution of the rare vocal learning trait. Each of the findings will be discussed separately along with its implications and potential avenues of further investigation.

7.1 Activity in Forebrain Regions During Vocalization

I believe the results reported in this study provide the clearest evidence for activity in non-human mammalian motor cortex during naturally produced vocalizations that is related specifically to vocal production. As discussed previously, two recent studies reported activity-dependent expression of immediate early genes in prefrontal cortex during vocalization in marmosets (Miller et al., 2010; Simões et al., 2010). Both studies failed to include a treatment group that vocalized while deaf, effectively eliminating confounding effects of auditory feedback, even though this control has been included in a similar study in songbirds (Jarvis and Nottebohm, 1997). In marmosets, without eliminating the contribution of auditory processing, which one of

the studies showed to be quite large (Simões et al., 2010), to activity in prefrontal cortex it is difficult to directly attribute the reported increases in gene expression associated with calling to the motor act itself. Because the levels of gene expression I observed in motor cortex were similar after ultrasonic song production in the presence and absence of auditory feedback, I am more confident that the activity associated with the gene expression was related to vocal motor production. Furthermore, the pattern of gene expression was confirmed with two different IEG's, permitting more confidence in the reliability and interpretation of the results.

Additional support for my interpretation of singing-induced IEG expression comes from one mouse that was intended as a non-vocalizing control that emitted a small amount of USV spontaneously without stimulation with female urine. This mouse showed a low-level IEG activation in a pattern similar to the vocalizing treatment groups (Appendix C), as would be predicted for singing-related activity. That IEG expression in the brainstem was observed in Amb but not XII of vocalizing mice (Appendix C) is also indicative of singing-related activity. To confirm song premotor activity it will be necessary to record multi- and single-unit activity from neurons in the high IEG expressing regions of motor cortex, cingulate cortex, and anterodorsal striatum. To my knowledge, no laryngeal premotor activity has been recorded directly from motor cortex in non-human mammals.

The regions displaying singing-related IEG induction in mice seem larger and more diffuse than the foci of activation observed in song system nuclei of birds. For example, the mouse motor cortical area identified by behavioral gene mapping contains, but is much larger than, the region where PRV-Bartha labeling revealed laryngeal premotor neurons (compare Figures 7 And 17). The laryngeally-connected subpopulation of motor cortical layer V pyramidal cells may not be spatially restricted to a discrete region as in the songbird nucleus RA, but may be embedded within a heterogeneous region of the motor cortex that supplies other oro-facial premotor neurons. This possibility is supported by the tracer injections (Chapter 3) that labeled cortico-fugal axons with brainstem and spinal cord targets other than Amb. Vocal premotor neurons in a multi-functional motor cortical region would be consistent with the motor theory of vocal learning origin, where it is proposed that vocal forebrain regions that control vocalizations evolved by duplication and emergence from adjacent motor pathways in vocal learning species (Feenders et al., 2008). The cortical regions displaying singing-related IEG expression and not labeled by PRV-Bartha could be part of the song circuit, but have different connections than M1. For example, M2 would not be expected to make significant projections to the brainstem. Instead, I would predict that the singing-activated region of M2 projects to the singing-activated region of M1.

7.2 Descending Vocal Pathway from the Motor Cortex

The singing-associated forebrain pathways discovered in this study included brain regions and connectivity similar to cortico-striatal-thalamic loops for song learning in birds and proposed loops for speech learning in humans (Figure 5) (Lieberman, 2001; Jarvis, 2004; Jürgens, 2008). To test this idea, future experiments should test for the proposed connections between dorsolateral striatum and the thalamus, which are likely to go through the globus pallidus. Further investigation should also test whether the cortico-striatal-thalamic circuit is dedicated to vocalization as in songbirds, a hypothesis that is difficult to test in human subjects.

The direct forebrain projection to Amb in mice appears much less robust than in vocal learning birds (Wild, 1993). The analogous projection in humans also appears sparse (Kuypers, 1958c; Iwatsubo et al., 1990), suggesting that density of innervation may not be the critical factor for this complex behavior in mammals. A recent study in rats using the same PRV-Bartha back-tracing technique employed in this study did not report the same cluster of motor cortical cells I observed in this study (Van Daele and Cassell, 2009). Instead, they found very few, isolated, labeled cells in primary motor cortex at a much later survival time (more than 120 hrs after injection into laryngeal muscles). They suggest a weak and indirect connection between M1 and Amb, and propose instead that laryngeal motor cortex is located laterally in the insular cortex; however, that study did not discuss the possible implications of these findings or test for

a direct projection. The presence of a direct cortico-bulbar connection from motor cortex suggests that mice, if not rodents generally, share a neuroanatomical feature with humans not shared with our closest primate relatives.

Future studies should test whether the motor cortical axons detected in the peri-ambigual zone of dendritic arborization of laryngeal motoneurons make functional synaptic connections onto identified motoneurons. This can be accomplished with electron microscopy, a technique that was employed previously to identify the only known direct cortico-bulbar connection to brainstem motoneurons in rodents from vibrissa motor cortex to VII (Grinevich et al., 2005). Most importantly, the identification of a direct connection, even if weak, opens the possibility of studying the molecular basis for specifying this projection that is considered one of the most critical steps in the evolution of vocal learning. Identification of the genetic factors involved in developing this connection could allow for the establishment of a connection de novo in non-learning species, and perhaps the amplification and recovery of vocal learning abilities in species that already learn vocalizations.

7.3 Direct Motor Cortical Vocal Pathway Does Not Generate Ultrasonic Songs

The results of motor cortical lesions in mice were somewhere between reported effects in vocal learners (humans and songbirds) and those in non-human primates. Specifically, I found that mouse ultrasonic songs are not generated by the motor cortex like speech and learned birdsong, but they are modified by cortical input in ways that

do not seem to operate in monkeys. My interpretation of the effects in mice is that mouse USV song syllables are similar to male zebra finch long calls, which contain both learned and innate components. Lesions of the analogous vocal premotor pathway in zebra finches eliminate the modified features of calls and leave only a basic innate template (Simpson and Vicario, 1990).

Thus, mouse songs may not necessarily be entirely 'learned' or 'innate'. Previous studies of vocal learning have focused entirely on systems that are either innate and dominated by the limbic-PAG vocal pathway (suboscine song, calls of chickens and pigeons, non-human primate calls) or dominated by the direct motor cortical vocal pathway (human speech, birdsong). It is conceivable that an intermediate neural organization existed as a step from a phylogenetically older vocal system focused on the PAG to a system controlled by motor cortex capable of greater vocal flexibility and eventually generation of novel sounds. Such a system would be a valuable tool for investigating the interplay of motor commands from the cortical and limbic-PAG vocal premotor pathways.

Future studies should investigate the relative contributions of both vocal pathways in mouse USVs. For example, reversible pharmacological inactivation (Jürgens and Ehrenreich, 2007) of motor cortex could help elucidate the role of this brain region in song production throughout the lifespan of the mouse. Complementary studies using optogenetic techniques to activate molecularly identified populations of

Amb-projecting cortical neurons would allow temporally precise examination of the link between cortical activity, muscle activation, and real-time acoustic output (Aravanis et al., 2007; Ayling et al., 2009). Additionally, it will be important to test whether the direct motor cortical vocal pathway is necessary for driving the pitch convergence between strains, and whether modifying activity in this pathway can enhance the rate of convergence.

7.4 The Role of Auditory Feedback in Ultrasonic Song Production

The deleterious effects of deafening on birdsong and human speech are well documented, including changes in repertoire, sequencing, and spectral quality (Marler and Waser, 1977; Marler and Sherman, 1983; Waldstein, 1990; Okanoya and Yamaguchi, 1997; Woolley and Rubel, 1997; Doupe and Kuhl, 1999; Heaton et al., 1999; Watanabe et al., 2006), but similar results have not been obtained in monkeys (Talmage-Riggs et al., 1972; Hammerschmidt et al., 2001). The results of deafening in mice suggest that a re-evaluation of the past literature may be necessary. It is possible that closer examination of calls from deafened non-human primates could also reveal changes in fine acoustic structure. So, the relationship between auditory experience and vocal development may have to be re-examined in some species. Even with the caveat that similar effects may have been overlooked in previous studies, the results I obtained in mice are more extensive than any that have been reported to date outside of humans.

Importantly, the deafening-induced deficits in adult mice appear less dramatic than those observed in humans and vocal learning birds (Marler and Sherman, 1983; Waldstein, 1990; Okanoya and Yamaguchi, 1997; Woolley and Rubel, 1997; Doupe and Kuhl, 1999; Heaton et al., 1999; Watanabe et al., 2006); although, deafening-induced deterioration can also be subtle and slow in some vocal learning species. For example, in budgerigars, spectral abnormalities in warble songs and contact calls took several months to develop in many animals (Heaton et al., 1999). Moreover, even 16 months after deafening the spectral degradation was not as severe as that reported for songbirds and bats after as little as 1 month (Rübsamen and Schäfer, 1990; Woolley and Rubel, 1997). Thus, there appears to be significant variability in the timing and magnitude of post-deafening deterioration of vocalizations between species, and by comparison the results reported in this study for mice are not exceptionally small or slow.

Ultimately, estimations of what constitutes large or small changes by human observers are arbitrary. Future work should test whether the deafening- and lesion-induced changes in songs are behaviorally relevant in the mating context as perceived by expert listeners for whom the communication signal is designed. Do females prefer the ultrasonic courtship songs of hearing-intact males to the abnormal songs of deaf males? There is some evidence that females can distinguish songs of kin and non-kin males that would most likely be statistically indistinguishable on the spectral features I measured (Figure 35). Therefore, it is possible that even a small but reliable change in

repertoire composition, pitch, or spectral purity could carry enough information to affect behavioral choices. Additionally, future studies should test whether auditory feedback is critical for adaptive vocal modification, such as pitch convergence between strains.

7.5 Vocal Learning Capacity in Laboratory Mice

In addition to the dependence on auditory feedback, I found that mice are able to adaptively modify their vocalizations in certain social conditions. The pitch of adult male pairs from different strains converged when they were housed together in the presence of a female. There is clear evidence of convergence as a result of vocal imitation from mammalian vocal learning species including bats (Boughman, 1998) and dolphins (Watwood et al., 2004). Each of these species demonstrated a persistent change in the spectral content of vocalizations as a result of auditory experience. Bats mimicked the spectral composition of calls when transferred to a new social group, and male dolphins matched each other's whistle types after forming social alliances. In both bats and dolphins the convergence was gradual and persistent. Vocal convergence in mice also showed a gradual systematic trend, suggesting that similar processes of acoustic modification and retention may underlie the effects in mice, bats, and dolphins.

Vocal convergence of spectral content has also been proposed for non-human primate calls (Mitani and Gros-Louis, 1998; Marshall et al., 1999; Snowdon and Elowson, 1999; Weiss et al., 2001; Crockford et al., 2004; Egnor and Hauser, 2004; Tyack, 2008); however, the case is based largely on descriptive observations with little experimental

verification. For example, studies in cotton-top tamarins (Weiss et al., 2001) and chimpanzees (Marshall et al., 1999; Crockford et al., 2004) reported within-group similarities and between-group differences in calls that were not explained by genetic factors. These observations imply, but do not demonstrate, the possibility of vocal learning as a factor in generating group vocal identity. One study of chimpanzee pant-hoot calls demonstrated modulation of spectral content over a short time period (Mitani and Gros-Louis, 1998). In this case, males that vocalized together produced similar calls as a result of one male matching the call parameters of his chorusing partner; however, the changes did not extend to calls produced in isolation, indicating that this is a transient effect distinct from the long-term changes observed in mice, bats, and elephants. Thus, the result could be interpreted as an example of context learning where a similar shared variant of the pant-hoot is produced in a particular social context, rather than a persistent change to the structure of the call acquired through production learning. Moreover, it was unclear what specific acoustic features were contributing to the effect, and data were only reported for one dyad of males making it difficult to draw reliable conclusions from changes in a single animal.

Attempts have been made to directly test for convergence in non-human primates by experimentally modifying the social environment (Elowson and Snowdon, 1994; Snowdon and Elowson, 1999). When two groups of marmosets were housed in a common acoustical environment for 10 weeks, monkeys from both groups responded

with a gradual increase in peak frequency of trill calls (Elowson and Snowdon, 1994). While the results of that study demonstrated a degree of persistent vocal flexibility in response to a social change, the response was a parallel change not targeted convergence. The same researchers later conducted a more direct test for vocal convergence by analyzing the calls of males and females before and up to three years after pairing (Snowdon and Elowson, 1999). The results were inconsistent in that after pairing only 2 of the 4 pairs showed increased call similarity on one of the two spectral features measured, bandwidth and peak frequency. In those pairs with increased post-pairing similarity on bandwidth or peak frequency, only one animal changed significantly from baseline. This single male increased both mean peak frequency and bandwidth, making him more similar to his female cage-mate after pairing. In the other instances where an initial difference in mean frequency or bandwidth was eliminated neither of the animals varied significantly from their baseline levels. As with the data for convergence of chimpanzee pant-hoot calls, it is difficult to draw reliable conclusions from changes in one animal. Therefore, the more parsimonious explanation in this case is that most of the effects observed were within the natural variability of the individual monkeys.

I observed pitch convergence after cross-strain pairings in adult mice that was much more pronounced than what has been reported previously for non-human primates. Furthermore, the nature and timing of the pitch convergence of Type A

syllables was similar to what has been reported for calls in bats and dolphins. These results suggest that mice are capable of at least limited vocal learning in the form of vocal convergence of existing call types. Future work should investigate whether the learning abilities of mice extend beyond modification of templates to the generation of novel sounds or learning syllable sequences. The most convincing evidence of vocal learning would come through successful tutoring of heterospecific, artificial, or anthropogenic sounds. My initial attempt at tutoring young pups by foster parents of a different strain yielded inconclusive results (Appendix D). A related preliminary report has indicated that two strains of mice that sing different songs (C57 & BALB/c) are not able to imitate each other's songs when cross-fostered (K. Okanoya, personal communication); however, the learning paradigm used in that study did not ensure or test for vocal production by the foster father. Absence of such vocal production would prevent the young males from acquiring a model song. Additionally, generation of novel but untutored vocalizations could be attempted via contingency learning. Therefore, the possibilities of mouse song learning by true imitation or by improvisation remain open and require further study.

7.6 The Laboratory Mouse as an Animal Model for Studying Vocal Communication

The search for a suitable mammalian model for the study of speech and vocal learning has continued for decades and has traditionally focused on non-human primates (Deacon, 1989; Ghazanfar and Hauser, 1999; Jürgens, 2002; Egnor and Hauser,

2004; Jürgens, 2008). Despite general agreement that non-human primates lack key neuroanatomical and developmental requirements for speech, such as the direct cortical control of acoustic output and a reliance on auditory experience for early vocal development, some researchers argue that these species can still provide a model for the basic processes of vocal production and perception that serve as the substrates for spoken language (Ghazanfar and Hauser, 1999; Egnor and Hauser, 2004). I believe the results reported in my studies of the mouse ultrasonic song system add to the non-human primate model in the following ways:

1) Demonstrated forebrain activity related to vocal production in mice, but not clearly in monkeys.
2) Direct cortico-bulbar projection to laryngeal motoneurons discovered in mice, but not in monkeys.
3) Lesions to the direct motor cortical vocal pathway changed vocalizations in mice, but not in monkeys.
4) Removal of auditory feedback changed vocalizations in mice, but not in monkeys.
5) Found stronger evidence for vocal convergence of adult vocalizations in mice than in monkeys.

An additional benefit of using mouse models to study mammalian vocal communication is the abundance of readily available, optimized techniques for genetic manipulation. The best example to date has been the foxP2 gene that encodes a forkhead-box transcription factor. Multiple studies have linked foxP2 mutations to severe speech deficits in human patients (Vernes et al., 2006) and vocal production in mice and songbirds (Haesler et al., 2004; Shu et al., 2005; Teramitsu and White, 2006; White et al., 2006; Haesler et al., 2007; Rochefort et al., 2007; Groszer et al., 2008; Miller et al., 2008; Enard et al., 2009; Fischer and Hammerschmidt, 2010; Gaub et al., 2010). Although foxP2 is regulated during singing and development in songbirds (Haesler et al., 2004; Teramitsu and White, 2006; Miller et al., 2008), there has been little progress directly linking gene sequence, and protein function to phenotype in this system. Knockdown of foxP2 was achieved in the avian striatum and affected the quality of song learning (Haesler et al., 2007), but required the relatively inefficient technique of lentiviral injection. By comparison, multiple studies have leveraged the tools available for generating transgenic mouse models to test the effects of targeted disruptions of the foxP2 gene on neural circuits and vocal production (Shu et al., 2005; French et al., 2007; Mizutani et al., 2007; Fujita et al., 2008; Groszer et al., 2008; Enard et al., 2009; Gaub et al., 2010). It was even shown that mouse pups expressing a humanized version of the foxP2 gene produced calls with lower peak frequency (Enard et al., 2009), directly linking specific functional mutations from the human lineage to changes in vocal output.

In a preliminary experiment, I found that mutations of the mouse foxP2 gene identical to those that cause verbal apraxia in the human KE family also result in changes to the spectral characteristics of adult mouse songs (Appendix E). Moreover, foxP2 is expressed in layer V of M1 (Campbell et al., 2008) in what appears to be the same region where I observed singing-related IEG expression and laryngeal premotor neurons (Figures 7, 10 and 17). Thus, it may soon be possible to directly link specific genetic mutations, protein function in identified subsets of neurons, and neural activity in a vocal premotor pathway to features of acoustic output within the mouse song system. Future studies should investigate expression of foxP2 in identified mouse vocal circuits, and test for a possible function in establishing the direct cortico-ambigual projection in mice. Given the demonstrated role of foxP2 in birdsong learning, it would also be interesting to test if foxP2 expression levels affect vocal convergence in adult mice.

Mouse vocalizations are also beginning to be studied in models of psychiatric disorders involving communication deficits, such as autism (Jamain et al., 2008; Scattoni et al., 2008a; Scattoni et al., 2008b; Radyushkin et al., 2009; Fischer and Hammerschmidt, 2010). Some genes involved in the regulation of social behavior have also been implicated in USV production in mice including vasopressin (Scattoni et al., 2008b) and neuroligins implicated in monogenic heritable autism spectrum disorders (Jamain et al., 2008; Radyushkin et al., 2009). Additionally, an inbred strain of mice (BTBR) with social

deficits and repetitive behaviors similar to autism also displays abnormal vocal behavior (Scattoni et al., 2008a), like some autistic infants.

As a supplement to non-human primate studies of vocal communication, mouse models can clearly serve to cover the gap in understanding the molecular basis of vocal production, social communication dysfunctions, and the evolution of brain systems that form the basic substrates of speech. More work is necessary to establish how useful mouse models will be as a supplement to the songbird model in studying the process of vocal learning. This will be chiefly determined by whether the learning capabilities of mice extend beyond the limits of vocal convergence of pitch. The current framework for classifying vocal learning and non-learning species presents a dichotomous scheme whereby a species is either: 1) a vocal mimic with the associated neuroanatomical traits found shared among all vocal learning species studied to date; 2) a vocal non-learner producing innate vocalizations without the associated neuroanatomical and developmental characteristics of learners. This schema overlooks some problematic examples, such as those that develop novel vocalizations without mimicry. I propose that mice can be described at minimum as limited vocal learners with the corresponding neural circuits for adaptive vocal modification, and a requirement of auditory experience for normal vocal production.

The interpretation of mice as limited vocal learners challenges the dogmatic assumption that the distribution of the vocal learning trait among vertebrates is

dichotomous. Instead, it suggests that vocal learning/modification abilities may be expressed along a spectrum with vocal mimics and so-called vocal non-learning species at the extremes, and mice somewhere in between. Having vocal modification abilities expressed along a spectrum may make it easier to select for or against a convergent, advanced vocal imitation trait in multiple taxa. It has been proposed that positive selection for adaptive vocal modification can occur by sexual selection for more complex vocalizations, but that this could be strongly countered by negative selection by auditory systems of predators that habituate less to complex and variable vocalizations (Okanoya, 2002; Jarvis, 2004). The laboratory mice I studied have been bred in captivity for many generations, possibly allowing positive sexual selection for complex vocalizations without negative predatory pressure. It will be useful to determine if wild mice display similar vocal plasticity-associated neuroanatomical and behavioral features.

The alternative interpretation is that mice are vocal non-learners under the dichotomous classification scheme. In light of our findings, this interpretation should require a reappraisal of what strictly defines vocal learning. In contrast to current dogma, a dichotomous paradigm would require that a direct motor cortical vocal pathway, adaptive modification of vocal motor output guided by auditory experience is not sufficient to distinguish a vocal learning species. Only generation of vocalizations with reference to an externally acquired auditory template—learning through imitation—would be considered for vocal learning. I believe that such a strict definition

of vocal learning does not conform to the evidence, and instead some classically defined traits can be present with or without the advanced behavioral capacity for mimicry. Therefore, mouse models should be considered as a potential supplement to songbirds in studies of the neural architecture supporting cortical control of vocal behavior, sensorimotor integration for vocal production, and the social aspects of vocal communication.

Appendix A – General Methods

A.1 Animals

Adult males and females of the B6D2F1/J (BxD), BALB/c, C57BL/6J (C57), and B6.129S1-Casp3^{tm1Flv}/J (CASP3 KO) strains were purchased from the Jackson Laboratory (Bar Harbor, Maine). For experiments on juveniles, male pups of the BxD strain were bred in-house. For FoxP2 experiments adult male breeders of the Foxp2-R552H line (R552H) (Groszer et al., 2008) were obtained from the laboratory of Simon E. Fischer at the Wellcome Trust Centre for Human Genetics, University of Oxford, UK, and F1 male pups were bred in-house by crossing to C57 females. At 5 weeks old, males were socialized by spending at least one night with an adult female. All mice were group housed and kept on a 12 hr light/dark cycle.

A.2 Song Recordings

For audio recordings of ultrasonic songs, males were placed into a dark recording box with fresh bedding, allowed to acclimate overnight for IEG experiments or 15 min for audio only experiments, and then stimulated to sing by presenting either 200 μL of female urine directly into the bedding or a live female in the case of the juveniles. Sounds were recorded with UltraSoundGate CM16/CMPA ultrasound microphones with a flat frequency response from 30-130 kHz and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics). Detected sounds were digitized at 250 kHz, 8 bits and captured to disk as .WAV files using Avisoft Recorder USG (Avisoft

Bioacoustics). Example songs were pitched and slowed down to the human hearing range using Raven 4.0 (Cornell Laboratory of Bioacoustics).

A.3 Song Analysis

Acoustic waveforms were processed using custom MATLAB-based programs (modified from code written by Timothy E. Holy, Washington University). The sonogram was computed from each waveform (256 samples/block, half-overlap), a sound intensity threshold was set to eliminate the white noise component of the signal, and frequencies outside 35-125 kHz were truncated. Syllables with duration longer than 10ms were identified by previously described methods (Holy and Guo, 2005). The following spectral features were calculated from the sonograms for each note: Standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. The mean feature value (FV) for each spectral feature was calculated for each recording epoch for all note types. Recording epochs were: 1 month for adult deafening experiment (Section 5.3.2; an average of 4,713 syllables recorded per animal per month); 1 week for the cross-housing (Section 6.2.2; an average of 996 syllables recorded per animal per week) and cortical lesion (Section 4.2.2; an average of 3,376 syllables recorded per animal per week) experiments; and 1 day for the juvenile deafening (Section 5.3.1; an average of 7,411 syllables recorded per animal), cross-fostering (Appendix D, Section D.2; an average of 2,725 syllables recorded per animal),

and FoxP2 (Appendix E, Section E.2; an average of 403 syllables recorded per animal) experiments.

For longitudinal data in the adult deafening (Section 5.3.2) and cortical lesion (Section 4.2.2) experiments, I took the logarithm of the normalized mean value for each epoch (n) as the spectral feature score (SFS) such that:

$$SFS(n) = \log10[FV(n) / FV(1)]$$

This log ratio allowed me to compare the relative difference to the pre-treatment conditions across animals, which could differ in their absolute values.

A.4 Deafening in Adults

Anesthesia was induced with 5% isofluorane in oxygen and maintained by intramuscular injection of ketamine-xylazine (75 mg/kg ketamine; 5 mg/kg xylazine). A retro-aural incision was made and the skin and muscle were retracted to reveal the tympanic bulla. The lateral wall of the bulla was punctured to reveal the cochlea, and a pair of fine forceps was used to remove the tympanic membrane, stapes and parts of the cochlear walls until no cochlear structure was visible. The wound was closed with TissueMend (Veterinary Products Laboratories, Phoenix, AZ).

A.5 Pseudorabies Virus Injections

General anesthesia was induced with 5% isofluorane and maintained by intramuscular injection of ketamine-xylazine (75 mg/kg ketamine; 5 mg/kg xylazine). A midline incision was made from the sternum to the hyoid bone. The portion of the sternohyoid muscle covering the larynx was removed. Five 200 nL injections were made 1 min apart into each of m. cricothyroideus and m. circoarytaenoideus lateralis using a Nanofil microsyringe system with a 34 gauge stainless steel needle (World Precision Instruments, Sarasota, FL). After 5 min, the microinjection pipette was retracted, and the injection was repeated for m. cricoarytenoideus lateralis. A single break in the fascia was made for each muscle and sealed with TissueMend adhesive (Veterinary Products Laboratories) to prevent spread of virus to other tissues. After all injections were complete, the wound was closed with TissueMend.

Appendix B – List of Abbreviations

129sv – 129X1/SvJ strain

A1 – primary auditory cortex

AAc – central nucleus of the anterior arcopallium

adSt – anterodorsal striatum

AM – amygdala

Amb – nucleus ambiguus

Apir – amigdalopyriform transition area

Av – ventral auditory cortex

BDA – biodextran amines

BS – brainstem

BTBR – BTBR T+tf/J strain

BxD – B6D2F1/J strain

C57 – C57BL6/J strain

CASP3 – caspase-3

CASP3 KO – caspase-3 knockout

Cb – cerebellum

CC – corpus callosum

CeA – central nucleus of the amygdala

Cg – cingulate cortex

ChAT – choline acetyltransferase

CP – caudate-putamen

CT – cricothyroid muscle

CTb – cholera toxin subunit b

DAB – diaminobenzidine

eGFP – enhanced green fluorescent protein

DLM – dorsolateral anterior thalamic nucleus

DM – dorsomedial nucleus of the intercollicular complex

ES – expression score

FV – feature value

GFP – green fluorescent protein

GP – globus pallidus

H – hippocampus

HT – hypothalamus

IC – insular cortex

Int – internal capsule

LCA – lateral cricoarytenoid muscle

LMAN - lateral magnocellular nucleus of the anterior nidopallium

M1 – primary motor cortex

M2 – secondary motor cortex

MC – motor cortex

MD – medial dorsal nucleus of the thalamus

nXIIts – tracheosyringeal part of the hypoglossal nucleus

PAG – periaqueductal grey

PBr – parabrachial region

PCD – programmed cell death

PCR – polymerase chain reaction

PRV-Bartha – psuedorabies virus (Bartha strain)

Pyr – pyramidal tract

R552H – Foxp2-R552H line

RA – robust nucleus of the arcopallium

RAm – nucleus retroambigualis

RF – reticular formation

ROI – region of interest

S1 – primary sensorimotor cortex

SFS – spectral feature score

Sol – solitary nucleus

SP – spinal trigeminal nucleus

USV – ultrasonic vocalization

VM – ventral medial nucleus of thalamus

VMC – vibrissa motor cortex

VII – facial nucleus

VL – ventral lateral nucleus of the thalamus

XII – hypoglossal nucleus.

Appendix C – Behavioral Gene Mapping Supplement

Behavioral gene mapping (Chapter 2.1.3) also revealed singing-related arc mRNA expression in Amb (Figure 49a). Arc mRNA expression was observed in the area where laryngeal motoneurons would be expected in all of the mice from both Singing & Hearing and Singing Only groups. Unfortunately, Amb is difficult to locate reliably in Nissl stained sections, and without a molecular marker or tracer it was not possible to confirm the presence of laryngeal motoneurons in the non-vocalizing mice. However, no cell clusters expressing high levels of arc mRNA were observed in any animal from either the Hearing Only or Silent treatment groups in the region where laryngeal motoneurons would be expected. None of the mice showed high arc mRNA expression in XII (Figure 49b).

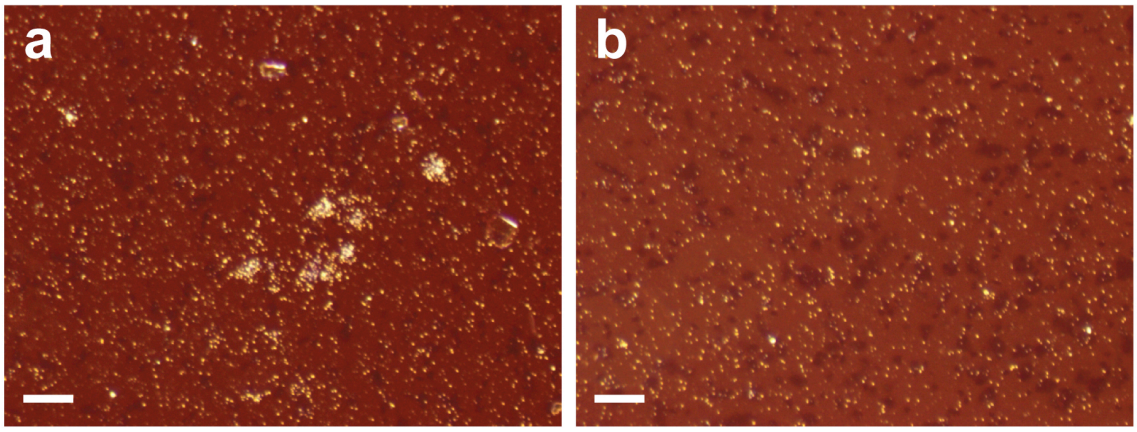

Figure 49: Dark-field images of cresyl violet stained (red) coronal brain section at the level of the brainstem showing singing-induced arc expression (white) in the forebrain of a mechanically deafened male mouse. Arc-expressing cells were observed in nucleus ambiguus (a) but not in the hypoglossal nucleus (b) from the same section.

During behavioral testing, one animal that was intended for the Hearing Only group spontaneously emitted USVs and was excluded from analysis. Brain sections from this mouse were, however, processed for *in situ* hybridization with the arc riboprobe. The general pattern of arc expression in the forebrain was reminiscent of the pattern observed in the vocalizing Singing & Hearing and Singing Only treatment groups (compare Figures 7 and 50). For example, high arc mRNA expression was observed in the cingulate cortex and motor cortex, and expression dropped sharply around the lateral border of MC. Thus, this pattern of forebrain activation is possible in singing animals even without presentation of urine as an eliciting stimulus.

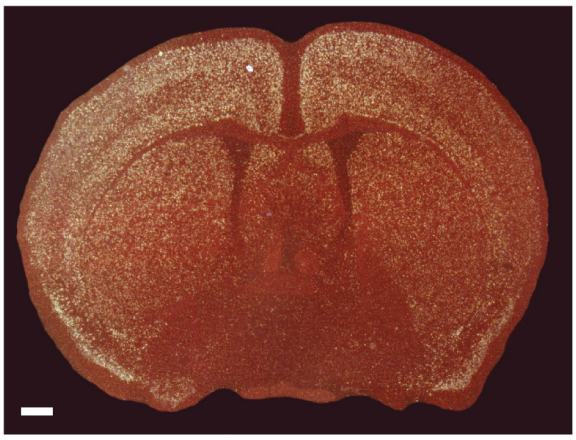

Figure 50: Dark-field image of a cresyl violet stained (red) coronal brain section at the level of the motor cortex showing singing-induced arc expression (white) in the forebrain of a mouse hearing ultrasonic playbacks. This mouse was not presented with female urine, but spontaneously produced USVs during the testing session. Low-level arc expression can be observed in the motor cortex and cingulate cortex in a pattern similar to Hearing & Singing and Singing Only treatment groups (Figure 7b & d).

Appendix D – Cross-Fostering of Mouse Pups

D.1 Methods – Cross-Fostering and Song Analysis

To test whether mouse pups could learn the songs of adult males I cross-fostered BxD and C57 pups with parents of the other strain.

D.1.1 Cross-Fostering

A C57 litter and a BxD litter were born 2 days apart, and half the litters were switched between the parents (post-natal day 2 for C57 and post-natal day 1 for BxD). The switch was made blind to the sex of the pups, and resulted in 2 C57 and 3 BxD pups being transferred. At weaning, it was discovered that only one transferred pup from each strain was male. Four BxD pups were not switched, and 3 of those were male. Three C57 pups were not switched, but none were male. The same C57 pair had a second litter a month and a half later. The entire litter of eight pups was transferred blind to sex on post-natal day 1, but to a different BxD pair than before. The recipient BxD pair had a litter of 13 born on the same day, of which 10 were transferred to the C57 pair. At weaning it was discovered that 6 of the transferred BxD pups, and 2 of the transferred C57 pups were male. In total, 7 BxD male pups were raised by the C57 pair (1 from the first cross and 6 from the second cross), and 3 C57 pups were raised by BxD pairs (1 by the first pair and 2 by the second pair). Three BxD males were raised by their biological parents. Around the same time, two other C57 pairs each produced a single male pup that was tutored by its biological parents.

Female pups were sacrificed after weaning. Male pups were left in the cage with the biological or foster parents until post-natal day 35 to extend the window for potential vocal imprinting.

D.1.2 Song Recordings

Each mouse was placed singly into a dark 15"x24"x12" sound-attenuating recording chamber, and ultrasonic songs elicited by presentation of 200 μL of freshly collected female urine were recorded with UltraSoundGate CM16/CMPA ultrasound microphones and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). Details of the recording procedure are described in the General Methods (Appendix A, Section A.2).

D.1.3 Song Analysis

Sonograms were computed from each waveform (256 samples/block, half-overlap), and the following spectral features were calculated from the sonograms for each syllable: standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. Details of song analysis are described in the General Methods (Appendix A, Section A.3).

The normal BxD group included 5 adult males recorded during a previous experiment (pre-surgical adults from Section 5.3) and the two males that served as foster fathers. The normal C57 group also included data from a previous experiment (8

wildtype C57 males from Section 5.3) and the single C57 male that served as a foster father.

D.2 Results

As in previous experiments, I focused on syllable Type A and compared the mean frequency syllables between the different strains and rearing groups. Neither BxD nor C57 males raised by their biological parents (BxD Uncrossed and C57 Uncrossed, respectively) differed from the pitch range of normal adults from the same strain (Figure 51; Mann-Whitney U-tests; BxD Uncrossed, n=3; BxD, n=7; C57 Uncrossed, n=2; C57, n=9; BxD Uncrossed vs. BxD, $p>0.5$; C57 Uncrossed vs. C57, $p>0.5$). Likewise, cross-fostered C57 males did not differ from the normal C57 group (Mann-Whitney U-test; C57 Crossed, n=3; C57, n=9; $p>0.5$). By contrast, BxD males that had been cross-fostered by C57 parents (BxD Crossed; n=7) sang in a pitch range that was intermediate between normal BxD and C57 adults. The mean frequency of Type A syllables from BxD Crossed songs [70.2 ± 0.6 kHz (s.e.m.)] was significantly higher than that from normal BxD songs [65.4 ± 1.1 kHz (s.e.m.)] and significantly lower than those from normal C57 songs [74.9 ± 1.3 kHz (s.e.m.)] (Mann-Whitney U-tests; BxD Crossed, n=7; BxD, n=7; C57, n=9; BxD Crossed vs. BxD, $p=0.018$; BxD Crossed vs. C57, $p=0.010$).

It appeared that cross-fostering by C57 parents had resulted in a significant pitch shift away from the normal BxD range towards the range of the foster strain; however, closer inspection of the data revealed that the biological father of 6 of the 7 BxD Crossed

males was also singing at a high pitch. The mean frequency of Type A syllables from the BxD father's songs was 70.5 kHz, very close to the 70.2 kHz mean frequency of the BxD Crossed group. Therefore, it was ambiguous whether the higher group mean was due to a shift towards the C57 range or genetic influence from an abnormally high-pitched BxD father. The issue was further complicated, because recordings from foster fathers were made more than a month after crossing. So, it was unclear if the high pitch of the BxD father might also have resulted from an upward shift following co-housing with younger and possibly more vocal C57 juveniles. Interestingly, just as the BxD foster father represented the upper limit of the BxD group, the C57 foster father represented the lower limit of the C57 group (Figure 51). The mean frequency of Type A syllables from the C57 foster father's songs was 67.1 kHz, very close to the range of both the BxD [65.4 ± 1.1 kHz (s.e.m.)] and BxD Uncrossed [64.2 ± 1.8 kHz (s.e.m.)] groups.

Although the observation of atypical song pitch in both BxD and C57 foster fathers after cross-fostering complicated the interpretation of the results, it also suggested the possibility that learning song features may not be limited to a unidirectional transfer from father to son. Instead, some form of learning might be possible even in adults. This was the motivation for the adult cross-housing experiment described in Chapter 6.

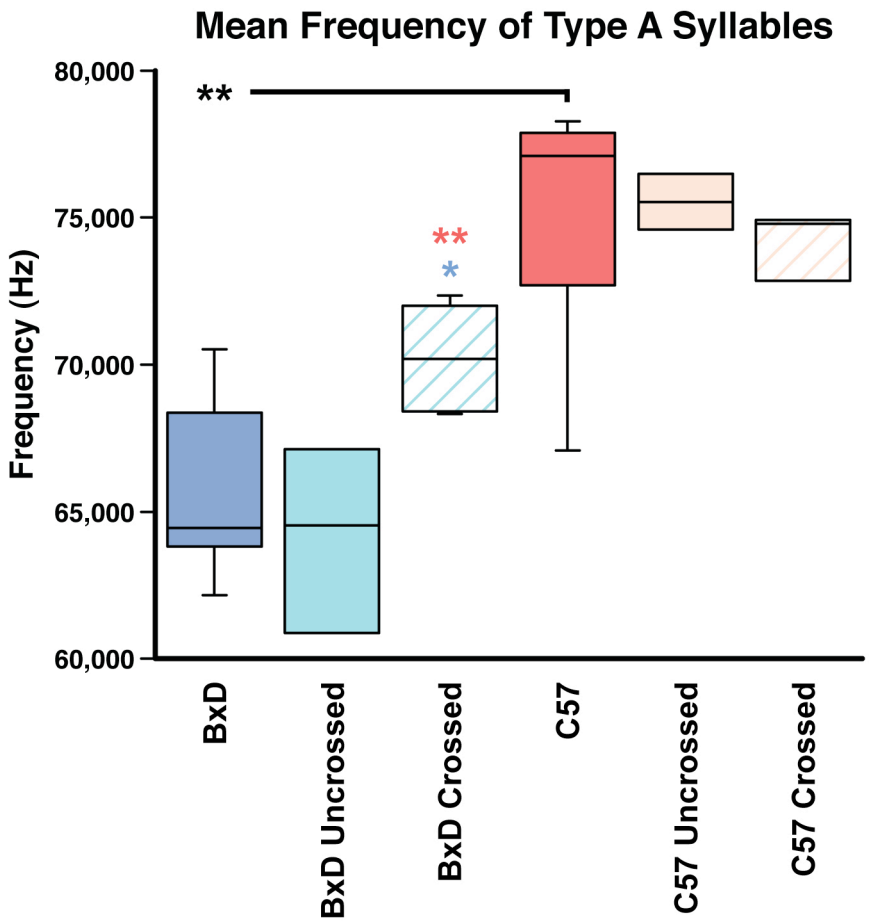

Figure 51: Mean frequency of Type A syllables from BxD and C57 male mice, including males that were raised by the biological parents (BxD Uncrossed and C57 Uncrossed) or foster parents of the other strain (BxD Crossed and C57 Crossed) until sexual maturity (Mann-Whitney U-tests comparing the mean differences between groups; BxD, n=7; BxD Uncrossed, n=3; BxD Crossed, n=7; C57, n=9; C57 Uncrossed, n=2; C57 Crossed, n=3; black ** = $p<0.01$; red ** = $p=0.01$ relative to C57 group; blue * = $p<0.05$, relative to BxD group).

Appendix E – Ultrasonic Songs of Adult FoxP2 Mutants

E.1 Methods – FoxP2 Breeding and Song Analysis

E.1.1 Breeding

R552H males carrying an arginine-to-histidine substitution (Groszer et al., 2008) identical to the R553H substitution in affected members of the KE family (Vernes et al., 2006) were obtained from the laboratory of Simon E. Fischer at the Wellcome Trust Centre for Human Genetics, University of Oxford, UK. Heterozygous R552H males were crossed to C57 females, yielding mixed heterozygous and wildtype litters. Pups were raised under normal conditions and sexed at weaning (postnatal day 22). Young males were group housed with littermates from weaning and socially experienced at 35 days old by housing with a female overnight.

E.1.2 Genotyping

Genotyping was performed after song recording and analysis. Genomic DNA was isolated from tail tissue samples using the Qiagen DNeasy Blood & Tissue Kit (Qiagen). Amplification of the FoxP2 gene was performed according to a previously described protocol (Groszer et al., 2008). Briefly, the polymerase chain reaction (PCR) and DNA sequencing was used to identify individuals carrying the R552H mutation. The following PCR primers were used:

Foxp2-R552H forward: 5′ -GTTCCTCTGGACATTTCAAC-3′

and

Foxp2-R552H reverse: 5′ -TGTGAGCATGCCTTTAGCTG-3′

The following PCR conditions were used: 94°C for 1 min (1 cycle), 94°C for 30 s, 55°C for 30 s, 68°C for 1 min (35 cycles), and 72°C for 10 min (1 cycle). The resulting 603 basepair fragment was purified and sequenced using the following primer:

5'-TAAAGCACTGTGTTGGGCTG-3'

E.1.3 Song Recording

Each mouse was placed singly into a dark 15″x24″x12″ sound-attenuating recording chamber, and ultrasonic songs elicited by presentation of 200 μL of freshly collected female urine were recorded with UltraSoundGate CM16/CMPA ultrasound microphones and an UltraSoundGate 416-200 recording interface (Avisoft Bioacoustics, Berlin, Germany). A continuous recording was made over 5 min for each animal, bling to genotype, starting at the time of urine presentation. Details of the recording procedure are described in the General Methods (Appendix A, Section A.2).

E.1.4 Song Analysis

Sonograms were computed from each waveform (256 samples/block, half-overlap), and the following spectral features were calculated from the sonograms for each syllable, blind to genotype: standard deviation of pitch distribution, starting frequency, ¼ frequency, ¾ frequency, final frequency, minimum frequency, mean frequency, maximum frequency, frequency variance, & spectral purity. Details of song analysis are described in the General Methods (Appendix A, Section A.3).

E.2 Results

Both wildtype C57 (+/+) and heterozygous R552H (R552H/+) males produced ultrasonic songs (Figure 52a). Both genotypes had poor and strong singers. Of 9 R552H/+ males recorded, 3 produced 1 - 8 total syllables over 5 min of recording. The other six R552H/+ males produced 348 ± 62 syllables. The 7 poor wildtype singers produced 1 – 10 total syllables each over 5 min, while the 7 strong singers produced 451 ± 163 syllables. Analysis was focused on the strong singers. Repertoire composition was similar in both genotypes, showing the typical dominance of Type A syllables (Figure 52b). On average, R552H/+ and wildtype littermates produced similar amounts of syllables over the 5 min recording session (Figure 52b; +/+, n=7; R552H/+, n=6; $p>0.5$).

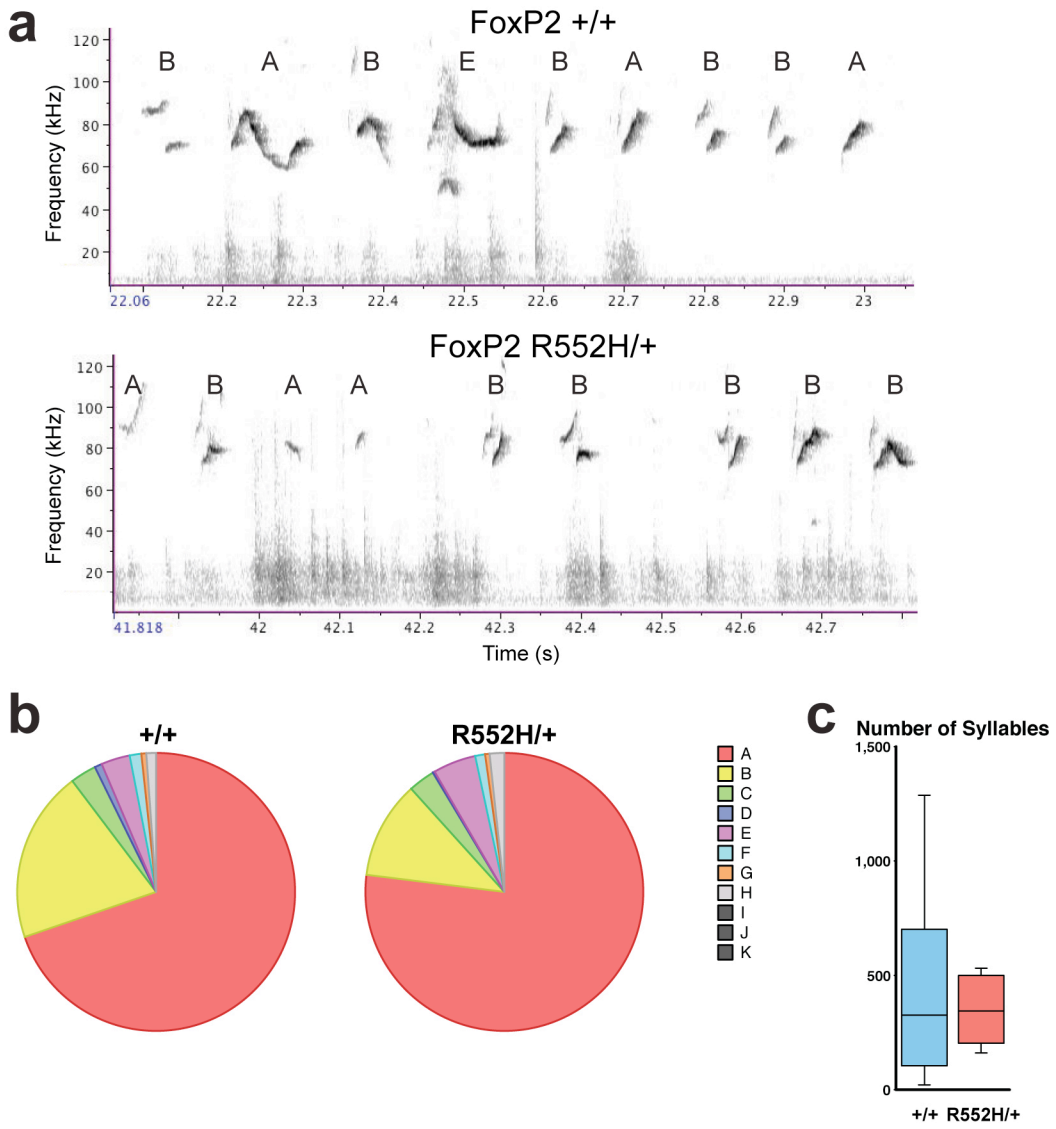

Figure 52: Sonograms representing 1 second of ultrasonic song from adult wildtype (+/+) and heterozygous FoxP2-R552H (R552H/+) males. b, Composition of the repertoire by syllable type for +/+ and R552H/+ males. c, Total number of syllables produced by +/+ and R552H/+ males during a 5 min recording session (Student's t-test, +/+, n=7; R552H/+, n=6; p>0.5).

Analysis of the most abundant syllable Type A revealed that some pitch-derived spectral features were different between wildtype and R552H/+ males. The starting, ¼, and minimum frequencies of Type A syllables in songs from the R552H/+ group were significantly higher than those in wildtype songs (Figure 53a; Student's t-tests; +/+, n=7; R552H/+, n=6; Starting Frequency, $p=0.009$; ¼ Frequency, $p=0.021$; ¾ Frequency, $p=0.403$; Final Frequency, $p>0.5$; Minimum Frequency, $p=0.28$; Mean Frequency, $p=0.220$; Maximum Frequency, $p>0.5$). There was no genotype-driven difference in the ¾, final, mean, or maximum frequency of Type A syllables. The standard deviation of the pitch distribution for Type A syllables was lower in R552H/+ males (Figure 53b; Student's t-test; +/+, n=7; R552H/+, n=6; $p=0.039$), suggesting the heterozygotes were singing over a more restricted frequency band. Accordingly, both the bandwith (Figure 53b; Student's t-test; +/+, n=7; R552H/+, n=6; $p=0.001$) and frequency variance (Figure 53c; Student's t-test; +/+, n=7; R552H/+, n=6; $p<0.001$) of Type A syllables were lower in R552H/+ males. Thus, male C57 mice heterozygous for the R552H mutation of foxP2, analogous to the mutation causing language impairment in the human KE family, show significant changes in the spectral structure of their ultrasonic songs, but the amount of song and proportion of syllables produced are not affected.

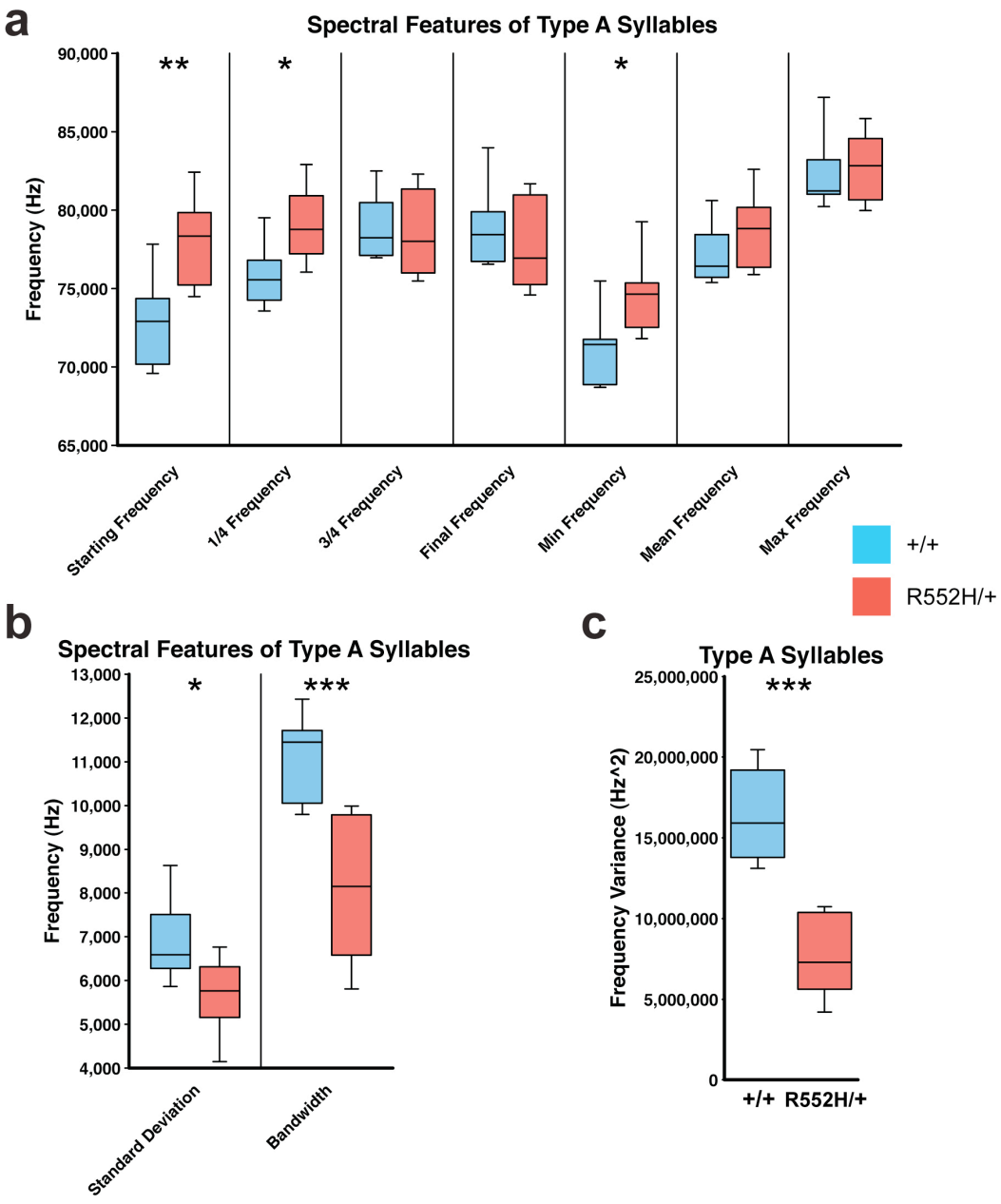

Figure 53: Spectral features of Type A syllables from ultrasonic songs of adult male wildtype (+/+) and heterozygous FoxP2-R552H (R552H/+) males. a, Instantaneous pitch-based spectral features of Type A syllables. b, Standard deviation of the pitch distribution and bandwidth of Type A syllables. c, Frequency variance of Type A syllables. (Student's t-tests, +/+, n=7; R552H/+, n=6; * = p>0.05, ** = p>0.01, *** = p>0.001).

References

Abbott SBG, Stornetta RL, Fortuna MG, Depuy SD, West GH, Harris TE, Guyenet PG (2009) Photostimulation of retrotrapezoid nucleus phox2b-expressing neurons in vivo produces long-lasting activation of breathing in rats. *J Neurosci* **29**(18):5806–5819.

Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F (2009) Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning. *P Natl Acad Sci USA* **106**(42):17963–17967.

Aitken PG (1981) Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain Lang* **13**(1):171–184.

Allard HA (1939) Vocal mimicry of the starling and the mockingbird. *Science* **90**(2338):370–371.

Altschuler SM, Bao XM, Miselis RR (1991) Dendritic architecture of nucleus ambiguus motoneurons projecting to the upper alimentary tract in the rat. *J Comp Neurol* **309**(3):402–414.

Aravanis AM, Wang LP, Zhang F, Meltzer LA, Mogri MZ, Schneider MB, Deisseroth K (2007) An optical neural interface: in vivo control of rodent motor cortex with integrated fiberoptic and optogenetic technology. *J Neural Eng* **4**(3):S143–56.

Armstrong EA (1963) A study of birdsong. Oxford University Press: London, UK.

Arnold K, Zuberbühler K (2008) Meaningful call combinations in a non-human primate. *Curr Biol* **18**(5):R202–3.

Aronov D, Andalman AS, Fee MS (2008) A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* **320**(5876):630–634.

Aston-Jones G, Card JP (2000) Use of pseudorabies virus to delineate multisynaptic circuits in brain: opportunities and limitations. *J Neurosci Methods* **103**(1):51–61.

Ayling OG, Harrison TC, Boyd JD, Goroshkov A, Murphy TH (2009) Automated light-based mapping of motor cortex by photoactivation of channelrhodopsin-2 transgenic mice. *Nat Methods* **6**(3):219–224.

Baptista LF, Petrinovich L (1984) Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Anim Behav* **32**(1):172–181.

Baptista LF, Petrinovich L (1986) Song development in the white-crowned sparrow: social factors and sex differences. *Anim Behav* **34**(5):1359–1371.

Baptista LF, Schuchmann KL (1990) Song learning in the Anna Hummingbird (Calypte anna). *Ethology* **84**(1):15–26.

Berquist SW, Ho JP, Metzner W (2010) Sound production in the isolated mouse larynx. *Society for Neuroscience Annual Meeting* **411.8**

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* **224**(4651):901–903.

Bottjer SW, Halsema KA, Brown SA, Miesner EA (1989) Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *J Comp Neurol* **279**(2):312–326.

Boughman JW (1998) Vocal learning by greater spear-nosed bats. *P Roy Soc Lond B Bio* **265**(1392):227–233.

Böhner J (1990) Early acquisition of song in the zebra finch, Taeniopygia guttata. *Anim Behav* **39**(2):369–374.

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**(6779):762–766.

Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. *Nature* **417**(6886):351–358.

Bramham CR, Alme MN, Bittins M, Kuipers SD, Nair RR, Pai B, Panja D, Schubert M, Soule J, Tiron A, Wibrand K (2009) The Arc of synaptic memory. *Exp Brain Res* **200**(2):125–140.

Branchi I, Santucci D, Alleva E (2001) Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. *Behav Brain Res* **125**(1-2):49–56.

Brecht M, Krauss A, Muhammad S, Sinai-Esfahani L, Bellanca S, Margrie TW (2004) Organization of rat vibrissa motor cortex and adjacent areas according to cytoarchitectonics, microstimulation, and intracellular stimulation of identified cells. *J Comp Neurol* **479**(4):360–373.

Brittan-Powell EF, Dooling RJ, Farabaugh SM (1997) Vocal development in budgerigars (Melopsittacus undulatus): contact calls. *J Comp Psychol* **111**(3):226–241.

Broughton WP (1963) Acoustic behavior of animals. Elsevier: Boston.

Brudzynski SM, Kehoe P, Callahan M (1999) Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Dev Psychobiol* **34**(3):195–204.

Brudzynski SM (2007) Ultrasonic calls of rats as indicator variables of negative or positive states: acetylcholine-dopamine interaction and acoustic coding. *Behav Brain Res* **182**(2):261–273.

Brudzynski SM (2009) Communication of adult rats by ultrasonic vocalization: biological, sociobiological, and neuroscience approaches. *ILAR J* **50**(1):43–50.

Burgdorf J, Wood PL, Kroes RA, Moskal JR (2007) Neurobiology of 50-kHz ultrasonic vocalizations in rats: electrode mapping, lesion, and pharmacology studies. *Behav Brain Res* **182**(2):274–283.

Campbell P, Reep RL, Stoll ML, Ophir AG, Phelps SM (2008) Conservation and diversity of Foxp2 expression in muroid rodents: functional implications. *J Comp Neurol* **512**(1):84–100.

Constantini F, D'Amato FR (2006) Ultrasonic vocalizations in mice and rats: social contexts and functions. *Acta Zool Sin* **52**(4):619–633.

Crockford C, Herbinger I, Vigilant L, Boesch C (2004) Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* **110**(3):221–243.

D'Amato FR, Scalera E, Sarli C, Moles A (2005) Pups call, mothers rush: does maternal responsiveness affect the amount of ultrasonic vocalizations in mouse pups? *Behav Genet* **35**(1):103–112.

Deacon TW (1989) The neural circuitry underlying primate calls and human language. *Hum Evol* **4**(5):367–401.

Deacon TW (2000) Evolutionary perspectives on language and brain plasticity. *J Commun Disord* **33**(4):273–291.

Dederen PJWC, Gribnau AAM, Curfs MHJM (1994) Retrograde neuronal tracing with cholera toxin B subunit: comparison of three different visualization methods. *Histochem J* **26**(11):856–862.

Delius JD (1971) Neural substrates of vocalizations in gulls and pigeons. *Exp Brain Res* **12**(1):64–80.

Depaulis A, Keay KA, Bandler R (1992) Longitudinal neuronal organization of defensive reactions in the midbrain periaqueductal gray region of the rat. *Exp Brain Res* **90**(2):307–318.

Dirks A, Fish EW, Kikusui T, van der Gugten J, Groenink L, Olivier B, Miczek KA (2002) Effects of corticotropin-releasing hormone on distress vocalizations and locomotion in maternally separated mouse pups. *Pharmacol Biochem Be* **72**(4):993–999.

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

Dujardin E, Jürgens U (2005) Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Res* **1034**(1-2):114–131.

Durand SE, Heaton JT, Amateau SK, Brauth SE (1997) Vocal control pathways through the anterior forebrain of a parrot (Melopsittacus undulatus). *J Comp Neurol* **377**(2):179–206.

Düsterhöft F, Häusler U, Jürgens U (2003) Neuronal activity in the periaqueductal gray and bordering structures during vocal communication in the squirrel monkey. *Neuroscience* **123**(1):53–60.

Egnor SER, Hauser MD (2004) A paradox in the evolution of primate vocal learning. *Trends Neurosci* **27**(11):649–654.

Elowson AM, Snowdon CT (1994) Pygmy marmosets, Cebuella pygmaea, modify vocal structure in response to changed social environment. *Anim Behav* **47**(6):1267–1277.

Elwood RW, Keeling F (1982) Temporal organization of ultrasonic vocalizations in infant mice. *Dev Psychobiol* **15**(3):221–227.

Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Favor J, Fuchs H, Gailus-Durner V, Hans W, Hölzlwimmer G, Javaheri A, Kalaydjiev S, Kallnik M, Kling E, Kunder S, Mossbrugger I, Naton B, Racz I, Rathkolb B, Rozman J, Schrewe A, Busch DH, Graw J, Ivandic B, Klingenspor M,

Klopstock T, Ollert M, Quintanilla-Martinez L, Schulz H, Wolf E, Wurst W, Zimmer A, Fisher SE, Morgenstern R, Arendt T, de Angelis MH, Fischer J, Schwarz J, Pääbo S (2009) A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell* **137**(5):961–971.

Ennis M, Xu SJ, Rizvi TA (1997) Discrete subregions of the rat midbrain periaqueductal gray project to nucleus ambiguus and the periambigual region. *Neuroscience* **80**(3):829–845.

Fee MS, Kozhevnikov AA, Hahnloser RHR (2004) Neural mechanisms of vocal sequence generation in the songbird. *Ann NY Acad Sci* **1016**(153–170.

Feenders G, Liedvogel M, Rivas M, Zapka M, Horita H, Hara E, Wada K, Mouritsen H, Jarvis ED (2008) Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. *PLoS ONE* **3**(3):e1768.

Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-type song culture in the zebra finch. *Nature* **459**(7246):564–568.

Fenzl T, Schuller G (2002) Periaqueductal gray and the region of the paralemniscal area have different functions in the control of vocalization in the neotropical bat, Phyllostomus discolor. *Eur J Neurosci* **16**(10):1974–1986.

Fenzl T, Schuller G (2005) Echolocation calls and communication calls are controlled differentially in the brainstem of the bat Phyllostomus discolor. *BMC Biol* **3**

Fernando P, Kelly JF, Balazsi K, Slack RS, Megeney LA (2002) Caspase 3 activity is required for skeletal muscle differentiation. *Proc Natl Acad Sci USA* **99**(17):11025–11030.

Fischer J, Hammerschmidt K (2010) Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication. *Genes Brain Behav* **no. doi: 10.1111/j.1601-183X.2010.00610.x(**.

Fish EW, Sekinda M, Ferrari PF, Dirks A, Miczek KA (2000) Distress vocalizations in maternally separated mouse pups: modulation via 5-HT1A, 5-HT1B and GABAA receptors. *Psychopharmacology* **149**(3):277–285.

Fish EW, Faccidomo S, Gupta S, Miczek KA (2004) Anxiolytic-like effects of escitalopram, citalopram, and R-citalopram in maternally separated mouse pups. *J Pharmacol Exp Ther* **308**(2):474–480.

Foster EF, Bottjer SW (1998) Axonal connections of the high vocal center and surrounding cortical regions in juvenile and adult male zebra finches. *J Comp Neurol* **397**(1):118–138.

French CA, Groszer M, Preece C, Coupe AM, Rajewsky K, Fisher SE (2007) Generation of mice with a conditional Foxp2 null allele. *Genesis* **45**(7):440–446.

Fromkin V, Krashen S, Curtiss S, Rigler D, Rigler M (1974) The development of language in genie: a case of language acquisition beyond the "critical period". *Brain Lang* **1**(1):81–107.

Frysztak RJ, Neafsey EJ (1991) The effect of medial frontal cortex lesions on respiration, "freezing," and ultrasonic vocalizations during conditioned emotional responses in rats. *Cereb Cortex* **1**(5):418–425.

Fujita E, Tanabe Y, Shiota A, Ueda M, Suwa K, Momoi MY, Momoi T (2008) Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. *P Natl Acad Sci USA* **105**(8):3117–3122.

Fukushima Y, Aoki K (2000) The role of the dorsomedial nucleus (DM) of intercollicular complex with regard to sexual difference of distance calls in Bengalese Finches. *Zool Sci* **17**(9):1231–1238.

Fukushima Y, Aoki K (2002) Neural function of the mesencephalic dorsomedial nucleus (DM) on distance call production in Bengalese finches. *Zool Sci* **19**(4):393–402.

Gahr M (2000) Neural song control system of hummingbirds: comparison to swifts, vocal learning (songbirds) and nonlearning (suboscines) passerines, and vocal learning (budgerigars) and nonlearning (dove, owl, gull, quail, chicken) nonpasserines. *J Comp Neurol* **426**(2):182–196.

Gaub S, Groszer M, Fisher SE, Ehret G (2010) The structure of innate vocalizations in Foxp2-deficient mouse pups. *Genes Brain Behav* **9**(4):390–401.

Ghazanfar AA, Hauser MD (1999) The neuroethology of primate vocal communication: substrates for the evolution of speech. *Trends Cogn Sci* **3**(10):377–384.

Gourbal BEF, Barthelemy M, Petit G, Gabrion C (2004) Spectrographic analysis of the ultrasonic vocalisations of adult male and female BALB/c mice. *Naturwissenschaften* **91**(8):381–385.

Grinevich V, Brecht M, Osten P (2005) Monosynaptic pathway from rat vibrissa motor cortex to facial motor neurons revealed by lentivirus-based axonal tracing. *J Neurosci* **25**(36):8250–8258.

Groszer M, Keays DA, Deacon RMJ, de Bono JP, Prasad-Mulcare S, Gaub S, Baum MG, French CA, Nicod J, Coventry JA, Enard W, Fray M, Brown SD, Nolan PM, Pääbo S, Channon KM, Costa RM, Eilers J, Ehret G, Rawlins JNP, Fisher SE (2008) Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol* **18**(5):354–362.

Guo Z, Holy TE (2007) Sex Selectivity of Mouse Ultrasonic Songs. *Chem Senses* **32**(5):463–473.

Gyger MG, Marler P, Pickert R (1987) Semantics of an avian alarm call system: the male domestic fowl, Gallus domesticus. *Behaviour* **102**(1-2):15–39.

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**(13):3164–3175.

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**(12):e321.

Hage SR, Jürgens U (2006a) Localization of a vocal pattern generator in the pontine brainstem of the squirrel monkey. *Eur J Neurosci* **23**(3):840–844.

Hage SR, Jürgens U (2006b) On the role of the pontine brainstem in vocal pattern generation: a telemetric single-unit recording study in the squirrel monkey. *J Neurosci* **26**(26):7105–7115.

Hahn ME, Hewitt JK, Adams M, Trully T (1987) Genetic influences on ultrasonic vocalizations in young mice. *Behav Genet* **17**(2):155–166.

Hahnloser RHR, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**(6902):65–70.

Hammerschmidt K, Freudenstein T, Jürgens U (2001) Vocal development in squirrel monkeys. *Behaviour* **138**(9):1179–1204.

Hammerschmidt K, Radyushkin K, Ehrenreich H, Fischer J (2009) Female mice respond to male ultrasonic 'songs' with approach behaviour. *Biol Lett* **5**(5):589–592.

Hannig S, Jürgens U (2005) Projections of the ventrolateral pontine vocalization area in the squirrel monkey. *Exp Brain Res* **169**(1):92–105.

Harmon KM, Cromwell HC, Burgdorf J, Moskal JR, Brudzynski SM, Kroes RA, Panksepp J (2008) Rats selectively bred for low levels of 50 kHz ultrasonic vocalizations exhibit alterations in early social motivation. *Dev Psychobiol* **50**(4):322–331.

Harrison DFN (1995) The anatomy and physiology of the mammalian larynx. Cambridge University Press: Cambridge, UK.

Hast MH, Fischer JM, Wetzel AB, Thompson VE (1974) Cortical motor representation of the laryngeal muscles in Macaca mulatta. *Brain Res* **73**(2):229–240.

Hayakawa T, Takanaga A, Maeda S, Ito H, Seki M (2000) Monosynaptic inputs from the nucleus tractus solitarii to the laryngeal motoneurons in the nucleus ambiguus of the rat. *Anat Embryol* **202**(5):411–420.

Heaton JT, Dooling RJ, Farabaugh SM (1999) Effects of deafening on the calls and warble song of adult budgerigars (Melopsittacus undulatus). *J Acoust Soc Am* **105**(3):2010–2019.

Heaton JT, Brauth SE (2000) Effects of lesions of the central nucleus of the anterior archistriatum on contact call and warble song production in the budgerigar (Melopsittacus undulatus). *Neurobiol Learn Mem* **73**(3):207–242.

Hile AG, Plummer TK, Striedter GF (2000) Male vocal imitation produces call convergence during pair bonding in budgerigars, Melopsittacus undulatus. *Anim Behav* **59**(6):1209–1218.

Hirasugi K, Hisa Y, Setsu T, Terashima T (2007) Ambiguus motoneurons innervating laryngeal and esophageal muscles are malpositioned in the Reelin-deficient mutant rat, Shaking Rat Kawasaki. *Acta Otolaryngol* **127**(2):213–220.

Hofer MA, Shair HN (1992) Ultrasonic vocalization by rat pups during recovery from deep hypothermia. *Dev Psychobiol* **25**(7):511–528.

Holy TE, Guo Z (2005) Ultrasonic songs of male mice. *PLoS Biol* **3**(12):e386.

Houde C, Banks KG, Coulombe N, Rasper D, Grimm E, Roy S, Simpson EM, Nicholson DW (2004) Caspase-7 expanded function and intrinsic expression level underlies

strain-specific brain phenotype of caspase-3-null mice. *J Neurosci* **24**(44):9977–9984.

Ise S, Ohta H (2009) Power spectrum analysis of ultrasonic vocalization elicited by maternal separation in rat pups. *Brain Res* **1283**(58–64.

Iwatsubo T, Kuzuhara S, Kanemitsu A, Shimada H, Toyokura Y (1990) Corticofugal projections to the motor nuclei of the brainstem and spinal cord in humans. *Neurology* **40**(2):309–312.

Jamain S, Radyushkin K, Hammerschmidt K, Granon S, Boretius S, Varoqueaux F, Ramanantsoa N, Gallego J, Ronnenberg A, Winter D, Frahm J, Fischer J, Bourgeron T, Ehrenreich H, Brose N (2008) Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *P Natl Acad Sci USA* **105**(5):1710–1715.

Janik VM, Slater PJB (1997) Vocal learning in mammals. *Adv Stud Behav* **26**(59–99.

Janik VM, Slater PJB (2000) The different roles of social learning in vocal communication. *Anim Behav* **60**(1):1–11.

Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. *P Natl Acad Sci USA* **94**(8):4097–4102.

Jarvis ED, Ribeiro S, da Silva ML, Ventura D, Vielliard J, Mello CV (2000) Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature* **406**(6796):628–632.

Jarvis ED, Mello CV (2000) Molecular mapping of brain areas involved in parrot vocal communication. *J Comp Neurol* **419**(1):1–31.

Jarvis ED (2004) Learned birdsong and the neurobiology of human language. *Ann NY Acad Sci* **1016**(749–777.

Jones G, Ransome RD (1993) Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proceedings of the Royal Society of London B* **252**(1334):125–128.

Jürgens U, Ploog D (1970) Cerebral representation of vocalization in the squirrel monkey. *Exp Brain Res* **10**(5):532–554.

Jürgens U, Pratt R (1979) Role of the periaqueductal grey in vocal expression of emotion. *Brain Res* **167**(2):367–378.

Jürgens U, Kirzinger A, Von Cramon D (1982) The effects of deep-reaching lesions in the cortical face area on phonation: a combined case report and experimental monkey study. *Cortex* **18**(1):125–139.

Jürgens U (1982) Afferents to the cortical larynx area in the monkey. *Brain Res* **239**(2):377–389.

Jürgens U (1983) Afferent fibers to the cingular vocalization region in the squirrel monkey. *Exp Neurol* **80**(2):395–409.

Jürgens U (1984) The efferent and afferent connections of the supplementary motor area. *Brain Res* **300**(1):63–81.

Jürgens U (1998) Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Naturwissenschaften* **85**(8):376–388.

Jürgens U, Alipour M (2002) A comparative study on the cortico-hypoglossal connections in primates, using biotin dextranamine. *Neurosci Lett* **328**(3):245–248.

Jürgens U (2002) Neural pathways underlying vocal control. *Neurosci Biobehav R* **26**(2):235–258.

Jürgens U, Ehrenreich L, De Lanerolle NC (2002) 2-Deoxyglucose uptake during vocalization in the squirrel monkey brain. *Behav Brain Res* **136**(2):605–610.

Jürgens U (2002) A study of the central control of vocalization using the squirrel monkey. *Med Eng Phys* **24**(7-8):473–477.

Jürgens U, Ehrenreich L (2007) The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Res* **1148**(90–95.

Jürgens U (2008) 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**(7026):638–643.

Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* **96**(3):1441–1455.

Kaplan G (2008) Alarm calls and referentiality in Australian magpies: between midbrain and forebrain, can a case be made for complex cognition? *Brain Res Bull* **76**(3):253–263.

Karakashian SJ, Gyger MG, Marler P (1988) Audience effects on alarm calling in chickens (Gallus gallus). *J Comp Psychol* **102**(2):129–135.

King AP (1983) Epigenesis of cowbird song–a joint endeavour of males and females. *Nature* **305**(704–706.

Kirzinger A, Jürgens U (1982) Cortical lesion effects and vocalization in the squirrel monkey. *Brain Res* **233**(2):299–315.

Kirzinger A, Jürgens U (1985) The effects of brainstem lesions on vocalization in the squirrel monkey. *Brain Res* **358**(1-2):150–162.

Knapska E, Kaczmarek L (2004) A gene for neuronal plasticity in the mammalian brain: Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK? *Prog Neurobiol* **74**(4):183–211.

Knutson B, Burgdorf J, Panksepp J (2002) Ultrasonic vocalizations as indices of affective states in rats. *Psychol Bull* **128**(6):961–977.

Konishi M (1963) The role of auditory feedback in the vocal behavior of the domestic fowl. *Z Tierpsychol* **20**(3):349–367.

Konishi M (1964) Effects of deafening on song development in two species of juncos. *Condor* **66**(2):85–102.

Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* **22**(7):770–783.

Konishi M (1985) Birdsong: from behavior to neuron. *Annu Rev Neurosci* **8**(125–170.

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

Kroodsma DE, Houlihan PW, Fallon PA, Wells JA (1997) Song development by grey catbirds. *Anim Behav* **54**(2):457–464.

Kurz A, Wöhr M, Walter M, Bonin M, Auburger G, Gispert S, Schwarting RKW (2010) Alpha-synuclein deficiency affects brain Foxp1 expression and ultrasonic vocalization. *Neuroscience* **166**(3):785–795.

Kuypers HGJM (1958a) An anatomical analysis of cortico-bulbar connexions to the pons and lower brain stem in the cat. *J Anat* **92**(2):198–218.

Kuypers HGJM (1958b) Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *J Comp Neurol* **110**(2):221–255.

Kuypers HGJM (1958c) Corticobular connexions to the pons and lower brain-stem in man: an anatomical study. *Brain* **81**(3):364–388.

Leonard JR, Klocke BJ, D'Sa C, Flavell RA, Roth KA (2002) Strain-dependent neurodevelopmental abnormalities in caspase-3-deficient mice. *J Neuropath Exp Neur* **61**(8):673–677.

Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* **399**(6735):466–470.

Leonardo A (2004) Experimental test of the birdsong error-correction model. *Proc. Natl. Acad. Sci. U.S.A.* **101**(48):16935–16940.

Leonardo A, Fee MS (2005) Ensemble coding of vocal control in birdsong. *J Neurosci* **25**(3):652–661.

Li J, Yuan J (2008) Caspases in apoptosis and beyond. *Oncogene* **27**(48):6194–6206.

Lieberman P (2001) Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought. *Perspect Biol Med* **44**(1):32–51.

Lu CL, Jürgens U (1993) Effects of chemical stimulation in the periaqueductal gray on vocalization in the squirrel monkey. *Brain Res Bull* **32**(2):143–151.

Ludlow CL (2005) Central nervous system control of the laryngeal muscles in humans. *Resp Physiol Neurobi* **147**(2-3):205–222.

Lüthe L, Häusler U, Jürgens U (2000) Neuronal activity in the medulla oblongata during vocalization. A single-unit recording study in the squirrel monkey. *Behav Brain Res* **116**(2):197–210.

MacLean PD (1978) Effects of lesions of globus pallidus on species-typical display behavior of squirrel monkeys. *Brain Res* **149**(1):175–196.

Mantyh PW (1983) Connections of midbrain periaqueductal gray in the monkey. II. Descending efferent projections. *J Neurophysiol* **49**(3):582–594.

Marler P, Kreith M, Tamura M (1962) Song development in hand-raised Oregon Juncos. *The Auk* **79**(1):12–30.

Marler P, Tamura M (1964) Culturally transmitted patterns of vocal behavior in sparrows. *Science* **146**(3650):1483–1486.

Marler P (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. *J Comp Physiol Psych* **71**(2, Pt 2):1–25.

Marler P (1970) Birdsong and speech development: could there be parallels? *Am Sci* **58**(6):669–673.

Marler P, Waser MS (1977) Role of auditory feedback in canary song development. *J Comp Physiol Psych* **91**(1):8–16.

Marler P, Sherman V (1983) Song structure without auditory feedback: emendations of the auditory template hypothesis. *J Neurosci* **3**(3):517–531.

Marler P (1997) Three models of song learning: evidence from behavior. *J Neurobiol* **33**(5):501–516.

Marler P (2004) Bird calls: a cornucopia for communication. In: Nature's Music: The science of birdsong (Marler P, Slabbekoorn H, eds), pp 132–177. San Diego, CA: Elsevier Academic Press.

Marshall AJ, Wrangham RW, Arcadi AC (1999) Does learning affect the structure of vocalizations in chimpanzees? *Anim Behav* **58**(4):825–830.

Miller CT, DiMauro A, Pistorio A, Hendry S, Wang X (2010) Vocalization induced cFos expression in marmoset cortex. *Front Integr Neurosci* **4**(128):1–15.

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**(4):2015–2025.

Mitani JC, Gros-Louis J (1998) Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**(8/9):1041–1064.

Mizutani A, Matsuzaki A, Momoi MY, Fujita E, Tanabe Y, Momoi T (2007) Intracellular distribution of a speech/language disorder associated FOXP2 mutant. *Biochem Bioph Res Co* **353**(4):869–874.

Moles A, Costantini F, Garbugino L, Zanettini C, D'Amato FR (2007) Ultrasonic vocalizations emitted during dyadic interactions in female mice: a possible index of sociability? *Behav Brain Res* **182**(2):223–230.

Morishita H, Makishima T, Kaneko C, Lee YS (2001) Deafness due to degeneration of cochlear neurons in caspase-3-deficient mice. *Biochem Bioph Res Co* **284**(1):142–149.

Müller-Preuss P, Jürgens U (1976) Projections from the 'cingular' vocalization area in the squirrel monkey. *Brain Res* **103**(1):29–43.

Müller-Preuss P, Newman JD, Jürgens U (1980) Anatomical and physiological evidence for a relationship between the 'cingular' vocalization area and the auditory cortex in the squirrel monkey. *Brain Res* **202**(2):307–315.

Musolf K, Hoffmann F, Penn DJ (2010) Ultrasonic courtship vocalizations in wild house mice, Mus musculus musculus. *Anim Behav* **79**(3):757–764.

Nicholson DW (1999) Caspase structure, proteolytic substrates, and function during apoptotic cell death. *Cell Death Differ* **6**(11):1028–1042.

Noirot E, Pye D (1969) Sound analysis of ultrasonic distress calls of mouse pups as a function of their age. *Anim Behav* **17**(2):340–349.

Noirot E (1972) Ultrasounds and maternal behavior in small rodents. *Dev Psychobiol* **5**(4):371–387.

Nottebohm F, Nottebohm ME (1971) Vocalizations and breeding behaviour of surgically deafened ring doves (Streptopelia risoria). *Anim Behav* **19**(2):313–327.

Nottebohm F (1972) The origins of vocal learning. *Am Nat* **106**(947):116–140.

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, Serinus canarius. *J Comp Neurol* **165**(4):457–486.

Nottebohm F, Paton JA, Kelley DB (1982) Connections of vocal control nuclei in the canary telencephalon. *J Comp Neurol* **207**(4):344–357.

Nunez AA, Pomerantz SM, Bean NJ, Youngstrom TG (1985) Effects of laryngeal denervation on ultrasound production and male sexual behavior in rodents. *Physiol Behav* **34**(6):901–905.

Nyby J (1983) Ultrasonic vocalizations during sex behavior of male house mice (Mus musculus): a description. *Behav Neural Biol* **39**(1):128–134.

Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (Lonchura striata var. domestica) require real-time auditory feedback to produce normal song syntax. *J Neurobiol* **33**(4):343–356.

Okanoya K (2002) Sexual display as a syntactical vehicle: the evolution of syntax in birdsong and human language through sexual selection. In: The transition to language (Wray A, ed), pp 46–63. Oxford, UK: Oxford University Press.

Okanoya K (2004) Functional and structural pre-adaptations to language: insight from comparative cognitive science into the study of language origin. *JPN Psychol Res* **46**(3):207–215.

Okuhata S, Saito N (1987) Synaptic connections of thalamo-cerebral vocal nuclei of the canary. *Brain Res Bull* **18**(1):35–44.

Olivier B, Molewijk E, van Oorschot R, van der Poel G, Zethof T, van der Heyden J, Mos J (1994) New animal models of anxiety. *Eur Neuropsychopharm* **4**(2):93–102.

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* **3**(5):e153.

Oppenheim RW, Flavell RA, Vinsant S, Prevette D, Kuan CY, Rakic P (2001) Programmed cell death of developing mammalian neurons after genetic deletion of caspases. *J Neurosci* **21**(13):4752–4760.

Person AL, Gale SD, Farries MA, Perkel DJ (2008) Organization of the songbird basal ganglia, including area X. *J Comp Neurol* **508**(5):840–866.

Pomerantz SM, Nunez AA, Bean NJ (1983) Female behavior is affected by male ultrasonic vocalizations in house mice. *Physiol Behav* **31**(1):91–96.

Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood SL (2005) Animal behaviour: elephants are capable of vocal learning. *Nature* **434**(7032):455–456.

Portfors CV (2007) Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J Am Assoc Lab Anim* **46**(1):28–34.

Pronichev IV, Lenkov DN (1998) Functional mapping of the motor cortex in the white mouse by microstimulation. *Neurosci Behav Physiol* **28**(1):80–85.

Pulvermüller F (2002) Synfire chains as the basis of serial order in the brain. In: The Neuroscience of Language: On Brain Circuits of Words and Serial Order ed), pp 147–158. New York, NY: Cambridge University Press.

Radyushkin K, Hammerschmidt K, Boretius S, Varoqueaux F, El-Kordi A, Ronnenberg A, Winter D, Frahm J, Fischer J, Brose N, Ehrenreich H (2009) Neuroligin-3-deficient mice: model of a monogenic heritable form of autism with an olfactory deficit. *Genes Brain Behav* **8**(4):416–425.

Ramus F, Fisher SE (2009) Genetics of language. In: The Cognitive Neurosciences (Gazzaniga MS, ed), pp 855–872. Cambridge, MA: MIT Press.

Reiss D, McCowan B (1993) Spontaneous vocal mimicry and production by bottlenose dolphins (Tursiops truncatus): evidence for vocal learning. *J Comp Psychol* **107**(3):301–312.

Ribeiro S, Shi X, Engelhard M, Zhou Y, Zhang H, Gervasoni D, Lin SC, Wada K, Lemos NA, Nicolelis MA (2007) Novel experience induces persistent sleep-dependent plasticity in the cortex but not in the hippocampus. *Front Neurosci* **1**(1):43–55.

Richards DG, Wolz JP, Herman LM (1984) Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, Tursiops truncatus. *Journal of Comparative Psychology* **98**(1):10–28.

Roberts LH (1975) Evidence for the laryngeal source of ultrasonic and audible cries of rodents. *J Zool* **175**(2):243–257.

Rochefort C, He X, Scotto-Lomassese S, Scharff C (2007) Recruitment of FoxP2-expressing neurons to area X varies during song development. *Dev Neurobiol* **67**(6):809–817.

Romand R, Ehret G (1984) Development of sound production in normal, isolated, and deafened kittens during the first postnatal months. *Dev Psychobiol* **17**(6):629–649.

Rübsamen R, Schäfer M (1990) Audiovocal interactions during development? Vocalisation in deafened young horseshoe bats vs. audition in vocalisation-impaired bats. *J Comp Physiol A* **167**(6):771–784.

Sales GD, Smith JC (1978) Comparative studies of the ultrasonic calls of infant murid rodents. *Dev Psychobiol* **11**(6):595–619.

Scattoni ML, Gandhy SU, Ricceri L, Crawley JN (2008a) Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PLoS ONE* **3**(8):e3067.

Scattoni ML, McFarlane HG, Zhodzishsky V, Caldwell HK, Young WS, Ricceri L, Crawley JN (2008b) Reduced ultrasonic vocalizations in vasopressin 1b knockout mice. *Behav Brain Res* **187**(2):371–378.

Scattoni ML, Crawley JN, Ricceri L (2009) Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neurosci Biobehav R* **33**(4):508–515.

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**(9):2896–2913.

Schulz SB, Haesler S, Scharff C, Rochefort C (2010) Knock-down of FoxP2 alters spine density in Area X of the Zebra Finch. *Genes* **9**(7):732–740.

Schusterman RJ, Reichmuth C (2008) Novel sound production through contingency learning in the Pacific walrus (Odobenus rosmarus divergens). *Anim Cogn* **11**(2):319–327.

Schusterman RJ (2008) Vocal learning in mammals with special emphasis on pinnipeds. *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication.* 41–70.

Seyfarth RM, Cheney DL, Marler P (1980) Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**(4471):801–803.

Seyfarth RM, Cheney DL (1986) Vocal development in vervet monkeys. *Anim Behav* **34**(6):1640–1658.

Shu W, Cho JY, Jiang Y, Zhang M, Weisz D, Elder GA, Schmeidler J, De Gasperi R, Sosa MA, Rabidou D, Santucci AC, Perl D, Morrisey E, Buxbaum JD (2005) Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene. *P Natl Acad Sci USA* **102**(27):9643–9648.

Siebert S, Jürgens U (2003) Vocalization after periaqueductal grey inactivation with the GABA agonist muscimol in the squirrel monkey. *Neurosci Lett* **340**(2):111–114.

Simões C, Vianney PVR, de Moura MM, Freire MAM, Mello LE, Sameshima K, Araújo JF, Nicolelis MA, Mello CV, Ribeiro S (2010) Activation of frontal neocortical areas by vocal production in marmosets. *Front Integr Neurosci* **4**(123):1–12.

Simonyan K, Jürgens U (2002) Cortico-cortical projections of the motorcortical larynx area in the rhesus monkey. *Brain Res* **949**(1-2):23–31.

Simonyan K, Jürgens U (2003) Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Res* **974**(1-2):43–59.

Simonyan K, Jürgens U (2004) Afferent subcortical connections into the motor cortical larynx area in the rhesus monkey. *Neuroscience* **130**(1):119–131.

Simonyan K, Jürgens U (2005) Afferent cortical connections of the motor cortical larynx area in the rhesus monkey. *Neuroscience* **130**(1):133–149.

Simpson HB, Vicario DS (1990) Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J Neurosci* **10**(5):1541–1556.

Snowdon CT, Elowson AM (1999) Pygmy marmosets modify call structure when paired. *Ethology* **105**(10):893–908.

Striedter GF (1994) The vocal control pathways in budgerigars differ from those in songbirds. *J Comp Neurol* **343**(1):35–56.

Sutton D, Larson C, Lindeman RC (1974) Neocortical and limbic lesion effects on primate phonation. *Brain Res* **71**(61–75.

Takahashi K, Kamiya K, Urase K, Suga M, Takizawa T, Mori H, Yoshikawa Y, Ichimura K, Kuida K, Momoi T (2001) Caspase-3-deficiency induces hyperplasia of supporting cells and degeneration of sensory cells resulting in the hearing loss. *Brain Res* **894**(2):359–367.

Talmage-Riggs G, Winter P, Ploog D, Mayer W (1972) Effect of deafening on the vocal behavior of the squirrel monkey (Saimiri sciureus). *Folia Primatol* **17**(5):404–420.

Teramitsu I, White SA (2006) FoxP2 regulation during undirected singing in adult songbirds. *J Neurosci* **26**(28):7390–7394.

Terao Y, Ugawa Y, Yamamoto T, Sakurai Y, Masumoto T, Abe O, Masutani Y, Aoki S, Tsuji S (2007) Primary face motor area as the motor representation of articulation. *J Neurol* **254**(4):442–447.

Thomas LB, Stemple JC, Andreatta RD, Andrade FH (2009) Establishing a new animal model for the study of laryngeal biology and disease: an anatomic study of the mouse larynx. *J Speech Lang Hear R* **52**(3):802–811.

Thoms G, Jürgens U (1987) Common input of the cranial motor nuclei involved in phonation in squirrel monkey. *Exp Neurol* **95**(1):85–99.

Thorpe WH (1958) The learning of song patterns by birds, with especial reference to the song of the chaffinch Fringilla coelebs. *Ibis* **100**(4):535–570.

Travers JB, Norgren R (1983) Afferent projections to the oral motor nuclei in the rat. *J Comp Neurol* **220**(3):280–298.

Tyack PL (2008) Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *J Comp Psychol* **122**(3):319–331.

Uematsu A, Kikusui T, Kihara T, Harada T, Kato M, Nakano K, Murakami O, Koshida N, Takeuchi Y, Mori Y (2007) Maternal approaches to pup ultrasonic vocalizations produced by a nanocrystalline silicon thermo-acoustic emitter. *Brain Res* **1163**(91–99.

Van Daele DJ, Cassell MD (2009) Multiple forebrain systems converge on motor neurons innervating the thyroarytenoid muscle. *Neuroscience* **162**(2):501–524.

Vanderhorst VGJM, Terasawa E, Ralston HJ (2001) Monosynaptic projections from the nucleus retroambiguus region to laryngeal motoneurons in the rhesus monkey. *Neuroscience* **107**(1):117–125.

Vernes SC, Nicod J, Elahi FM, Coventry JA, Kenny N, Coupe AM, Bird LE, Davies KE, Fisher SE (2006) Functional genetic analysis of mutations implicated in a human speech and language disorder. *Hum Mol Genet* **15**(21):3154–3167.

Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV, Horita H, Patterson MA, White SA, Scharff C, Haesler S, Zhao S, Sakaguchi H, Hagiwara M, Shiraki T, Hirozane-Kishikawa T, Skene P, Hayashizaki Y, Carninci P, Jarvis ED (2006) A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. *P Natl Acad Sci USA* **103**(41):15212–15217.

Waldstein RS (1990) Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J Acoust Soc Am* **88**(5):2099–2114.

Watanabe A, Eda-Fujiwara H, Kimura T (2006) Auditory feedback is necessary for long-term maintenance of high-frequency sound syllables in the song of adult male budgerigars (Melopsittacus undulatus). *J Comp Physiol A* **193**(1):81–97.

Watwood SL, Tyack PL, Wells RS (2004) Whistle sharing in paired male bottlenose dolphins, Tursiops truncatus. *Behav Ecol Sociobiol* **55**(6):531–543.

Weiss DJ, Garibaldi BT, Hauser MD (2001) The production and perception of long calls by cotton-top tamarins (Saguinus oedipus): acoustic analyses and playback experiments. *J Comp Psychol* **115**(3):258–271.

West MJ, King AP (1988) Female visual displays affect the development of male song in the cowbird. *Nature* **334**(6179):244–246.

White SA, Fisher SE, Geschwind DH, Scharff C, Holy TE (2006) Singing mice, songbirds, and more: models for FOXP2 function and dysfunction in human speech and language. *J Neurosci* **26**(41):10376–10379.

Whitney G, Nyby J (1983) Sound communication among adults. Charles C Thomas Pub Ltd: Springfield, IL.

Wild JM (1993) Descending projections of the songbird nucleus robustus archistriatalis. *J Comp Neurol* **338**(2):225–241.

Wild JM (1997) Neural pathways for the control of birdsong production. *J Neurobiol* **33**(5):653–670.

Wild JM, Li D, Eagleton C (1997) Projections of the dorsomedial nucleus of the intercollicular complex (DM) in relation to respiratory-vocal nuclei in the brainstem of pigeon (Columba livia) and zebra finch (Taeniopygia guttata). *J Comp Neurol* **377**(3):392–413.

Winter P, Handley P, Ploog D, Schott D (1973) Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour* **47**(3):230–239.

Wöhr M, Dalhoff M, Wolf E, Holsboer F, Schwarting RKW, Wotjak CT (2008a) Effects of genetic background, gender, and early environmental factors on isolation-induced ultrasonic calling in mouse pups: an embryo-transfer study. *Behav Genet* **38**(6):579–595.

Wöhr M, Houx B, Schwarting RKW, Spruijt B (2008b) Effects of experience and context on 50-kHz vocalizations in rats. *Physiol Behav* **93**(4-5):766–776.

Woolley SMN, Rubel EW (1997) Bengalese finches Lonchura Striata domestica depend upon auditory feedback for the maintenance of adult song. *J Neurosci* **17**(16):6380–6390.

Yajima Y, Hayashi Y (1983) Ambiguous motoneurons discharging synchronously with ultrasonic vocalization in rats. *Brain Res* **50**(2-3):359–366.

Yajima Y, Larson CR (1993) Multifunctional properties of ambiguous neurons identified electrophysiologically during vocalization in the awake monkey. *J Neurophysiol* **70**(2):529–540.

Yoshida Y, Yatake K, Tanaka Y, Imamura R, Fukunaga H, Nakashima T, Hirano M (1998) Morphological observation of laryngeal motoneurons by means of cholera toxin B subunit tracing technique. *Acta Otolaryngol* **118**(539):98–105.

Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* **273**(5283):1871–1875.

Zann R (1985) Ontogeny of the zebra finch distance call: I. Effects of cross-fostering to bengalese finches. *Z Tierpsychol* **68**(1):1–23.

Zann R (1990) Song and call learning in wild zebra finches in south-east Australia. *Anim Behav* **40**(5):811–828.

Zhang J, Yang R, Pendlebery W, Luo P (2005) Monosynaptic circuitry of trigeminal proprioceptive afferents coordinating jaw movement with visceral and laryngeal activities in rats. *Neuroscience* **135**(2):497–505.

Zhang SP, Davis PJ, Bandler R, Carrive P (1994) Brain stem integration of vocalization: role of the midbrain periaqueductal gray. *J Neurophysiol* **72**(3):1337–1356.

Zeigler HP, Marler P (2008) The neuroscience of birdsong. Cambridge University Press: New York.

Biography

Gustavo Arriaga Jr. was born on Cannon Air Force Base in New Mexico on the twelfth of September, 1981. He attended Duke University and graduated with a Bachelor of Science in Psychology with a concentration in Neuroscience in 2003. Since then he has received a number of awards and fellowships including: Duke Endowment Fellowship from Duke University in 2005; Neuroscience Scholar Award from the Society for Neuroscience in 2006; Graduate Research Fellowship Honorable Mention from the National Science Foundation in 2006; and Graduate Research Fellowship from the National Science Foundation in 2007.