Brain & Language 115 (2010) 21–28

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Genes and vocal learning

Stephanie A. White

Department of Physiological Science, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Accepted 12 October 2009
Available online 13 November 2009

Keywords:
FoxP2
KE family
CNTNAP2
CASPR2
Songbird model
Fragile X

ABSTRACT

Could a mutation in a single gene be the evolutionary lynchpin supporting the development of human language? A rare mutation in the molecule known as FOXP2 discovered in a human family seemed to suggest so, and its sequence phylogeny reinforced a Chomskian view that language emerged wholesale in humans. Spurred by this discovery, research in primates, rodents and birds suggests that FoxP2 and other language-related genes are interactors in the neuromolecular networks that underlie subsystems of language, such symbolic understanding, vocal learning and theory of mind. The whole picture will only come together through comparative and integrative study into how the human language singularity evolved.

1. Introduction

In 1988, Noam Chomsky pondered, 'Perhaps at some time hundreds of thousands of years ago, some small change took place, some mutation took place in the cells of prehuman organisms. And for reasons of physics which are not yet understood, that led to the representation in the mind/brain of the mechanisms of discrete infinity, the basic concept of language and also of the number system' (Chomsky, 1988). Today, the idea that such a change was restricted to a single molecule and occurred solely in the hominid lineage, referred to as the Grammar Gene' theory, is deemed an extreme position (Bishop, 2009). Rather, the brain system underlying language is likely made up of subsystems, forms of which exist in other taxonomic groups. Language could have arisen in humans as a consequence of the unique intersection of these subsystems. A subsystem identified in non-humans could represent the homologous component in language. Alternatively, if the subsystem emerged at a point that does not feed into the hominid lineage, it could represent convergent or parallel (homoplasous) evolution whereby similar selection pressures drive parallel instances of similar biological solutions (as we, and others (e.g. Jarvis, 2004), have argued for birdsong and speech). Whether homologous or homoplasous, the good news is that this viewpoint opens the door for studying subsystems of language at the biological level using a comparative approach.

This chapter examines one language subsystem, namely the capacity for vocal learning, and the genes expressed in the central nervous system that are hypothesized to contribute to this ability. I focus on human speech and birdsong and define the vocal learning subsystem of language as the experience-dependent modification of one's vocal motor output with the goal of mimicking other members of one's species (conspecifics) or of creating new sounds. Deafness in humans and experiments in animals teach us that vocal learners must hear and attend to the vocalizations of conspecifics (with some exceptions: Feher et al., 2009; Kroodsma et al., 1997; Leitner et al., 2002) and hear their own vocal output in order to produce effective vocal communication signals (for review see Doupe and Kuhl (1999)). Comparison of these sounds to evaluate the match sets the stage for those neural changes that enable adaptive modifications of vocal output. In humans, vocal learning drives the development of speech.

Of course not all animals have been rigorously tested for the vocal learning ability. Tests of vocal learning often rely on deprivation of acoustic inputs during development and evaluation of how closely subsequent vocal output approximates normal song. One measure of how well song develops under these circumstances is whether it serves as an effective communication signal in conspecific interactions. Tests include: deafening early in development which prevents both hearing of others and of self; rearing in the absence of conspecific vocalizations which only prevents the former; and transient distortion of auditory feedback of the animal's own vocal output, affecting only the latter. Non-invasive methods include determining whether changes in vocal output during normal development are more substantial than those expected due to physical maturation of the vocal apparatus (such as the larnyx; Fitch, 1997) or are uncharacteristic of the species-specific behavior. By a majority of these tests, passerine birds of the oscine suborder, known as songbirds, are vocal learners. In addition to humans and songbirds, the short list of animals demonstrated to possess this ability is confined to: parrots and hummingbirds which are in

---

E-mail address: sawhite@ucla.edu

0093-934X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2009.10.002

separate avian orders (raising the hypothesis that the trait emerged independently three times in the avian lineage); certain marine mammals including harbor seals, dolphins and cetaceous whales; elephants; and certain bat species. As outlined in the introduction to this volume (see contribution by Brenowitz, Perkel & Osterhout, this issue) songbird species such as canaries (*Serinus canaria*), white crowned sparrows (*Zonotrichia leucophrys*), Bengalese finches (*Lonchura domestica*, also known as society finches) and zebra finches (*Taeniopygia guttata*) have relatively short generation times and are amenable to laboratory life, making them extremely practical species in which to conduct controlled studies of the biological basis for vocal learning.

In accordance with the general requirements for vocal learning outlined above, songbirds listen to the songs of their own species as well as their own vocal output (song perception), in order to adaptively modify control of the syrinx, or song organ, and the respiratory muscles used in singing. Experimental deprivation of these auditory inputs generally causes abnormal song in adulthood, but if temporarily applied and then removed, can extend the critical periods for song learning. This phenomenon is analogous to the extension of critical periods for neural organization in the visual system after rearing in darkness (for review see Hooks and Chen (2007)). For example, young zebra finches deprived of tutoring during normal sensory acquisition exhibit an extended critical period such that they can now learn song from a tutor provided after 65 days, the normal close of that critical period (Morrison & Nottebohm, 1993). Similarly, when loud masking noise is used to temporarily deprive finches of auditory feedback from their own vocalizations, sensorimotor learning is extended. Once the noise is turned off, the birds can adaptively modify their songs at ages when normally reared birds do not (Funabiki & Konishi, 2003). In zebra finches, dramatic modifications to song end with sexual maturity at ~100 days when song becomes stable, or crystallized. However, this behavioral stability is maintained by dynamic neural activity and depends on ongoing auditory feedback, as does human speech (Cynx & Von Rad, 2001; Nordeen & Nordeen, 1992; Woolley & Rubel, 1997; for review see Brainard and Doupe (2000)).

Song perception, its constituent neural systems, the genes underlying the formation and function of these systems, and peripheral vocal control are inescapably intertwined with vocal learning. For more information on these topics the interested reader is referred to contributions by Gentner and Goller, in this volume. Other subsystems of language may be better studied in other taxonomic groups. While birdsong can convey individual and species identity and can advertise mating or territorial ownership, it is not 'compositional', i.e. no single song syllable combines with others to build meaning the way that words do. The additional capacity for symbolic content is a necessary step in moving beyond the musicality of birdsong to semantically compositional language. The semantic subcomponent of language may be better studied in non-human primates (see Seyfarth and Cheney contribution, this issue), or, among birds, in parrots (see Pepperberg contribution, this issue). Yet it is worth noting that, in addition to vocal learning, certain songbird species do possess additional subsystems potentially linked to cognitive capacities required for language such as tool use, hierarchical reasoning, and context free syntax. (See Clayton and Emery (2005) for a review of corvid cognition, and Gentner, this volume, for syntax discrimination in starlings.)

2. Strategies used to identify genes for vocal learning

Now that we have defined vocal learning and introduced key songbird species in which to study it, how do we go about identifying genes that function in the song circuit and might generally underlie this rare trait? Since the initial observation of specialized nuclei within the telencephalon of song learners (Nottebohm & Arnold, 1976; see contributions of Margoliash, Schmidt and Kirn (this volume)), two general strategies have been used to isolate and characterize genes that contribute to the formation or function of vocal learning pathways. The first focuses on song circuit neuroanatomy and hypothesizes that important molecules are those that are differentially expressed *within* song control nuclei relative to surrounding tissue. The second approach is to make an educated guess as to candidate molecules for vocal learning. Initially, selection of candidate molecules was based on studies of learning and memory in other taxa or on critical periods in visual system development in rodents (e.g. *n*-methyl D-aspartate (NMDA) receptors). More recently attention has been paid to the handful of genes linked to human language disorders, such as Fragile X syndrome and mutations in FoxP2.

For both approaches, expression of an identified molecule within song control areas can be compared across the developmental phases of song learning. As described above, the timing of these phases can be experimentally manipulated. Thus, the songbird model allows for a unique test of any observed temporal correlation between molecular expression and vocal learning. If expression of a given molecule differs within the song circuit compared to outlying regions during normal sensorimotor learning, sensorimotor learning can be experimentally delayed (Funabiki & Konishi, 2003) to test whether the expression pattern is merely correlated with chronological age, or instead is more directly associated with the learning process. A major challenge in using songbirds to functionally verify the role of genes in vocal learning is that the avian egg is not easily amenable to genetic intervention (Sang, 1994). However, the use of viruses to introduce transgenes into the egg, or into song control regions of the developing brain is increasing, a topic I shall return to later on. Below, I review the genes that have emerged from these general approaches and how further investigation in songbirds contributes to understanding of the way in which they operate in vocal learning.

2.1. Enrichment in song control nuclei

The first systematic exploration of relatively abundant molecules in songbird telencephalon was conducted by Clayton, Nottebohm and colleagues (Clayton, 1997; Clayton et al., 1988). No gene exhibited an expression pattern that was entirely restricted to song nuclei. However, this study and similar approaches in which mRNA expression levels are compared inside versus outside song control regions (e.g. subtractive hybridization, and differential display; see below) have revealed molecules that are enriched in song nuclei. Of these, the first to be experimentally pursued was α-synuclein, a molecule that was independently identified in studies of human neurodegenerative diseases (for review see Clayton and George (1999)). Point mutations, duplications and triplications in the α-synuclein gene cause a rare dominant form of familial Parkinson's Disease (Biskup et al., 2008). Although the precise cellular function for α-synuclein has yet to be determined, evidence suggests that it acts at presynaptic nerve terminals. Its link to Parkinson's Disease coupled with its regulation in the song nucleus LMAN during early stages of birdsong learning suggest that α-synuclein function is critical to neural circuits that underlie the execution of learned motor skills.

Other examples of molecules that are concentrated in various song nuclei include: the biosynthetic enzyme for retinoic acid (Denisenko-Nehrbass et al., 2000) which is enriched in X-projecting HVC neurons; insulin-like growth factor II (IGF-II; Holzenberger et al., 1997), also enriched in these same neurons; and an as-yet unidentified antigen whose expression is largely limited to song