Leading Edge
Review

Cell

Insights into the Neural and Genetic Basis of Vocal Communication

Genevieve Konopka[1,*] and Todd F. Roberts[1,*]
[1]Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX 75390-9111, USA
*Correspondence: genevieve.konopka@utsouthwestern.edu (G.K.), todd.roberts@utsouthwestern.edu (T.F.R.)
http://dx.doi.org/10.1016/j.cell.2016.02.039

The use of vocalizations to communicate information and elaborate social bonds is an adaptation seen in many vertebrate species. Human speech is an extreme version of this pervasive form of communication. Unlike the vocalizations exhibited by the majority of land vertebrates, speech is a learned behavior requiring early sensory exposure and auditory feedback for its development and maintenance. Studies in humans and a small number of other species have provided insights into the neural and genetic basis for learned vocal communication and are helping to delineate the roles of brain circuits across the cortex, basal ganglia, and cerebellum in generating vocal behaviors. This Review provides an outline of the current knowledge about these circuits and the genes implicated in vocal communication, as well as a perspective on future research directions in this field.

Introduction

The insights discussed in this Review have been largely attained through the study of developmental disorders affecting speech and analysis of neuronal circuits in songbirds and mice. Genetic screens of individuals with inherited forms of speech disorders, like verbal dyspraxia, stuttering, and some types of autism, have allowed for the identification of a number of genes (*FOXP2*, *CNTNAP2*, *FOXP1*, *GNPTAB*, *GNPTG*, *NAGPA*) involved in speech and/or social-cognitive development that can now be studied using animal models (Konopka and Roberts, 2016; Lepp et al., 2013). Of these, the transcription factor *FOXP2* has been the most intensively studied. Mutations of *FOXP2* in humans are associated with an inherited verbal dyspraxia, a speech disorder that results from difficulties in controlling orofacial muscles. The study of FOXP2 is now providing significant insights into the underpinnings of vocal motor learning and the development of neuronal circuits.

Songbirds have long been the predominant model for studying the neural circuit mechanisms for vocal learning (Doupe and Kuhl, 1999; Mooney et al., 2008). Like human speech, birdsong is learned during a developmental sensitive period and requires early sensory exposure to a vocal model (song tutor) and auditory feedback for its normal development and maintenance. Studies in songbirds have revealed a well-delineated neural circuit spanning from the cortex to the brainstem that is necessary for song learning and song production. The organization of this song circuit is similar to the core cortical and basal ganglia circuits involved in speech (Doupe and Kuhl, 1999; Jarvis, 2004). In addition, knockdown of the transcription factor FoxP2 in songbirds disrupts song development in a manner similar to disruptions seen in human speech development, indicating analogous circuit and gene regulatory mechanisms for song and speech (Fisher and Scharff, 2009; Haesler et al., 2004, 2007; Lai et al., 2001; Murugan et al., 2013). Despite these important behavioral and neurobiological parallels between birdsong and speech, studies in songbirds have been limited by the lack of methods for efficiently and precisely editing the avian genome; however, the recent development of transgenic songbirds (Abe et al., 2015; Agate et al., 2009; Liu et al., 2015; Scott et al., 2010), advances in viral vector methods and gene editing tools (Betley and Sternson, 2011; Heidenreich and Zhang, 2016; Roberts et al., 2010, 2012), and the sequencing of the avian genome (Warren et al., 2010; Zhang et al., 2014) all promise to enrich the continued use of songbirds in the study of speech disorders.

The genetic accessibility of mice and the wide range of molecular and genetic tools available for studying the mouse brain provide a powerful platform for examining how genetic disorders affect the central nervous system and how genes implicated in speech and social/cognitive disorders impact neuronal circuit development and synaptic function. Mice exhibit both neonatal calls as well as adult vocalizations pertinent to social interactions (Scattoni et al., 2009). However, it should be appreciated that, unlike speech and birdsong, vocal behaviors in mice are not learned from social models using auditory feedback. For instance, deaf mice can develop normal vocalizations (Portfors and Perkel, 2014). This lack of vocal learning limits the use of mice for modeling speech development. However, their vocalizations still allow the study of motor and auditory brain circuits involved in vocal communication (Holy and Guo, 2005).

Overall, it is important to note that comparing vocalizations among humans, songbirds, and mice will always be challenging. While there is significant conservation of brain structures and genes among these divergent species, human language, characterized by speech and sign-based forms of communications in deaf communities, has a level of complexity and abstraction that may well be unique and thus difficult to model. Furthermore, vocal behaviors in mice and some species of songbirds are sexually dimorphic and sensitive to sex steroids, further

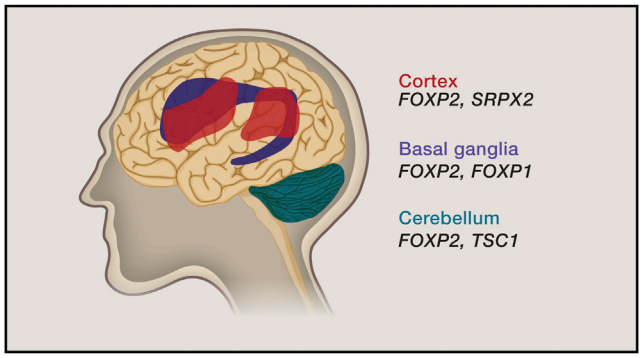

Figure 1. Schematic of the Neural Circuits and Genes Implicated in Vocal Communication

The neuronal circuits implicated in learning and accurately producing vocalizations include several interconnected networks in the cortex, basal ganglia, and cerebellum. While less is known about the gene networks involved in vocal communication, genetic screens of individuals with inherited speech disorders and the continued biological examination of neuronal circuits in vocal learning species, like songbirds, provide insights into the brain and genetic mechanisms for vocal communication.

underscoring the different evolutionary trajectories associated with vocal communication. However, by focusing on brain structures associated with speech—the cortex, basal ganglia, and cerebellum—we here provide touchstones for comparing and integrating genetic and neural circuit data from songbirds and mice with data from humans (Figure 1).

Cortex

The observation that brain lesions of the inferior frontal cortex lead to a disruption in speech production (expressive aphasia) in the late 1800s heralded the study of brain functions out of the dark ages of phrenology and provided one of the first insights into the brain mechanisms for vocal communication (Dronkers et al., 2007). This work by Paul Broca was soon followed by that of Karl Wernicke, who found that lesions of the superior temporal gyrus (STG) led to a deficit in speech perception (receptive aphasia) (Mathews et al., 1994). These early descriptions, along with later accounts provided by the pioneering work of the neurosurgeon Wilder Penfield, who carried out stimulation and recording of specific neocortical areas in awake patients (known as electrocorticography [ECoG] or intracranial electroencephalography [iEEG]) (Penfield and Rasmussen, 1949), laid the basis for attributing neural mechanisms to speech and language. Modern approaches have additionally used non-invasive techniques such as magnetoencephalography (MEG), transcranical magnetic stimulation (TMS), and functional MRI (fMRI) to study speech in both patient and neuro-typical populations (for an in-depth discussion and primary references, please see Cattaneo, 2013; Chang et al., 2015; Devlin and Watkins, 2007; Poeppel, 2012; Price, 2010).

The use of these techniques has revealed that the early divisions of speech production and perception into independent cortical regions were overly simplistic (see references in Hickok et al., 2011). For example, premotor cortex may modulate speech perception, and auditory areas (e.g., STG) are thought to influence speech production. The integration of these feedback loops among speech-related cortical areas permits ongoing learning, maintenance, and refinement of speech. Interestingly, a bilateral ECoG study directly demonstrated the existence of sensory-motor integration during speech and also provided evidence for bilateral neural activity in contrast to much of the work focusing on left hemisphere lateralization of language (Cogan et al., 2014).

Researchers have recently begun to parse the neuronal substrates for perceiving and producing the basic elements of speech. Application of ECoG allowed the determination of the neural responses to specific phonemes, or units of sound, during speech perception, showing that there are discrete and localized invariant responses to specific phonemes in the STG (Mesgarani et al., 2014). In addition, MEG of the cortex was recently used to identify the timescales of linguistic structure in a study of speech perception (Ding et al., 2015). Multi-electrode recordings have also recently helped map the spatial representation of phonetic features for speech production in the ventral sensorimotor cortex (adjacent to the so-called "Broca's area" in the inferior frontal cortex) (Bouchard et al., 2013). Building upon more than a century of work, these and other studies are redefining areas of the cortex important for speech.

These insights into speech production and comprehension are pertinent to the understanding of genetic and neuropsychiatric disorders that affect speech and language. Structural imaging of individuals with *FOXP2* mutations have identified both increases and decreases in gray matter in several cortical regions associated with speech, such as the STG and the inferior frontal gyrus (Belton et al., 2003; Watkins et al., 2002). fMRI studies of some of these individuals have also found decreases and/or alterations in cortical brain activity during word and non-word repetition paradigms (Liégeois et al., 2003, 2011), suggesting that deficits in cortical function may be associated with language difficulties imposed by this mutation possibly as a consequence of altered cortico-cerebellar or cortico-striatal circuitry. Disruption of corollary discharge pathways linking motor and auditory cortical circuits are speculated to contribute to auditory hallucinations and "imaginary inner speech" in schizophrenia (Heinks-Maldonado et al., 2007; Horga et al., 2014 and references in Hugdahl, 2015). Deficits in vocal communication are also associated with autism spectrum disorders (ASD). Patients with ASD-related syndromes or the more severe diagnosis of intellectual disability often have speech delay or can even be completely non-verbal. At the functional level, a reduction in left hemispheric lateralization of language has been observed in ASD patients as well as changes in prosody, verbal fluency, and activation of non-typical language areas (Kleinhans et al., 2008 and references in Dichter, 2012). Recent fMRI work has demonstrated hypoactivation of the STG in patients with ASD who exhibit language problems, suggesting that this fMRI signature could be used as a biomarker for ASD patients who will progress to poor outcomes and presenting an opportunity for therapeutic intervention (Lombardo et al., 2015).

Studies in songbirds have provided important insights into the architecture and function of cortical circuits for vocal communication. First, cortical song circuits involved in production of learned song are separable from those involved in vocal plasticity (Aronov et al., 2008; Brainard and Doupe, 2000;