A blueprint for vocal learning: auditory predispositions from brains to genomes

David Wheatcroft and Anna Qvarnström

Animal Ecology/Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

Memorizing and producing complex strings of sound are requirements for spoken human language. We share these behaviours with likely more than 4000 species of songbirds, making birds our primary model for studying the cognitive basis of vocal learning and, more generally, an important model for how memories are encoded in the brain. In songbirds, as in humans, the sounds that a juvenile learns later in life depend on auditory memories formed early in development. Experiments on a wide variety of songbird species suggest that the formation and lability of these auditory memories, in turn, depend on auditory predispositions that stimulate learning when a juvenile hears relevant, species-typical sounds. We review evidence that variation in key features of these auditory predispositions are determined by variation in genes underlying the development of the auditory system. We argue that increased investigation of the neuronal basis of auditory predispositions expressed early in life in combination with modern comparative genomic approaches may provide insights into the evolution of vocal learning.

1. Song learning in birds: a model system for understanding the cognitive basis of vocal learning

Language is a fundamental characteristic of human intelligence and sociality, but its evolutionary origin and cognitive basis remain mysterious, in large part because humans are the only extant apes to learn their vocalizations. This has led to a long-term use of animal models to understand the cognitive basis of vocal learning. The parallels between human language and animal vocal learning have been best explored in the oscine songbirds, a species-rich lineage of birds in which vocal learning is thought to be expressed by most species [1]. Like human language, song learning in birds generally occurs during a developmentally early sensitive period during which birds acquire socially relevant sounds from adult tutors [2]. The song learning process can be divided into two phases: (i) an earlier sensory phase during which a young individual imprints on a tutor’s vocalizations [3] and (ii) a later sensorimotor phase during which the individual acquires vocal elements by comparing its own ongoing vocal production against the imprinted model [4]. In some species, such as the zebra finch, Taeniopygia guttata, these phases overlap, whereas in many others, they are separated in time, implying that imprinting can form long-term auditory memories that bias a young bird’s auditory system to recognize and prefer similar sounds in the future (i.e. a learning ‘template’; [5]).

Songs vary greatly across species, and producing the right songs is critical for interactions with mates and rivals, meaning that any mechanism that focuses imprinting onto conspecific songs is likely to be under positive selection. A variety of mechanisms work to focus the range of sounds that young birds hear [6], memorize [7,8] and sing [9] towards socially relevant conspecific forms and away from other sounds they are exposed to during development. Auditory predispositions, which influence the early sensory phase by stimulating memorization of songs...
Table 1. Examples of auditory preferences for conspecific song and the song feature triggering responses.

<table>
<thead>
<tr>
<th>scientific name</th>
<th>common name</th>
<th>song feature</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fringilla coelebs</td>
<td>chaffinch</td>
<td>syntax</td>
<td>[10]</td>
</tr>
<tr>
<td>Lonchura striata</td>
<td>white-rumped munia</td>
<td>note form</td>
<td>[11]</td>
</tr>
<tr>
<td>Luscinia megarhynchos</td>
<td>nightingale</td>
<td>syntax</td>
<td>[12]</td>
</tr>
<tr>
<td>Melospiza georgiana</td>
<td>swamp sparrow</td>
<td>note form</td>
<td>[13,14]</td>
</tr>
<tr>
<td>Melospiza melodia</td>
<td>song sparrow</td>
<td>note form and syntax</td>
<td>[14]</td>
</tr>
<tr>
<td>Spizella passerina</td>
<td>chipping sparrow</td>
<td>self-matching</td>
<td>[15]</td>
</tr>
<tr>
<td>Serinus canarius</td>
<td>canary</td>
<td>song pitch</td>
<td>[16]</td>
</tr>
<tr>
<td>Taeniopygia guttata</td>
<td>zebra finch</td>
<td>unknown</td>
<td>[17]</td>
</tr>
<tr>
<td>Zonotrichia atricapilla</td>
<td>golden-crowned sparrow</td>
<td>unknown</td>
<td>[18]</td>
</tr>
<tr>
<td>Zonotrichia leucophrys</td>
<td>white-crowned sparrow</td>
<td>introductory whistle</td>
<td>[19]</td>
</tr>
</tbody>
</table>

aAuditory predispositions were demonstrated by comparing the number of chirps produced by wild nestlings raised by their parents in response to conspecific versus starling, *Sturnus vulgaris*, song.

bAuditory predispositions were demonstrated by comparing the number of hops juveniles made in response to conspecific versus white-crowned sparrow song.

containing species-typical sounds, have been particularly well studied (see below and table 1). However, very little is known about how these learning biases work, the neural structures involved, or the identity of genes underlying them. In this review, we summarize the evidence for learning predispositions and argue that comparative approaches using whole and partial genome sequencing on a wide variety of songbird species [20,21] and techniques to explore the associations between genetic and phenotypic variation between and within species offer new opportunities to reveal the neural and genetic basis of learning predispositions.

2. Guiding the imprinting process through auditory predispositions

Pioneering song playback experiments by Thorpe [22] on chaffinches, *Fringilla coelebs*, demonstrated that the tendency of young birds to learn the songs of their own species is at least in part guided by auditory biases that lead to a preference for species-specific songs. These song predispositions were shown to develop independently of previous exposure to adult songs through experiments performed on swamp sparrows, *Melospiza georgiana*. Juvenile swamp sparrows reared in isolation from adults selectively learnt conspecific song elements broadcasted interspersed among song elements from the closely related song sparrow *Melospiza melodia* [13]. Similar auditory predispositions, which guide the imprinting process underlying the formation of auditory memories, have subsequently been revealed in a wide variety of songbird species (table 1).

Auditory predispositions for conspecific songs are often assumed to be at least partly genetically determined, meaning that genetic variation across species leads to species-specific predispositions that are independent of early acoustic experience with conspecifics [23,24]. However, auditory predispositions, including those termed ‘innate’, have almost always been demonstrated in juvenile songbirds removed from both parents as young nestlings or raised by their mothers alone (examples in table 1). There are at least two caveats with these experiments. First, the work of Colombelli-Négrel et al. [25] demonstrates that songbirds can memorize specific, non-song vocalizations even at an embryonic stage, suggesting that some song learning may have occurred before young were separated from their parents. Second, even if nestling songbirds are prevented from memorizing the songs of a conspecific male, it is possible that being raised by a non-singing female or even hearing the begging calls of nest-mates can prime the auditory system to preferentially respond to songs containing similar features [26]. Marler and Peters’ [13] study is one of the few to specifically address these issues by isolating birds from adult sounds at an early embryonic stage. They demonstrated that such isolated nestlings maintained their species-specific auditory predispositions. This important result and subsequent work in which researchers either cross-fostered eggs across species [27] or used eggs that were naturally ‘fostered’ into the nest of another species by a non-songbird brood parasite [28] have demonstrated that auditory predispositions and learning tendencies can develop independently of experience with conspecifics. However, similar experiments on additional songbird species are needed to resolve the issue more generally.

In order for auditory predispositions to focus learning onto conspecific songs, the auditory systems of juvenile songbirds must be tuned into relevant features that allow them to recognize their own species’ vocalizations and discriminate against those of other species. There is significant empirical support for this idea. For example, juvenile white-crowned sparrows, *Zonotrichia leucophrys*, are stimulated to imprint on songs preceded by a species-characteristic whistle [19], swamp sparrows learn syllables containing multiple rapid frequency sweeps that are typical of their songs, whereas song sparrows also rely on the species-relevant tempo and organization of syllables during the imprinting process [14]. Table 1 gives additional examples. These examples demonstrate that auditory predispositions are not only widespread, but also adapted to trigger imprinting in response to the salient features of that species’ songs. However, little is known about the neural basis of these auditory preferences, and even less is known about the underlying genes. Does the variability in the song features that trigger imprinting across species imply a corresponding variation in the brain regions underlying auditory perception or in the developmental tuning of regions to sounds with species-specific properties? How are song properties encoded? What are the genes that give rise to these different developmental programmes?
3. Species-specific tuning of auditory perception

The auditory systems of all bird species, including both species that learn their vocalizations, such as songbirds, and those whose vocalizations are innate (i.e. non-vocal learners), share the same basic features [29,30], namely an ascending auditory pathway, involved primarily in sensory perception (illustrated in figure 1), and a descending auditory pathway (not illustrated). An experiment by Long et al. [31] in non-vocal learning species demonstrated that neural structures, possibly those connected by these pathways, are tuned to the frequency–temporal characteristics of conspecific songs, suggesting that field L may be involved in filtering out non-conspecific vocalizations from the song system [42].
developmentally tuned to respond to species-specific vocalizations. Parts of developing neural tubes from embryonic quails were transplanted onto embryonic chickens and when their auditory preferences were later tested after hatching, some of the resulting chimeras responded more strongly to quail vocalizations. Although Long et al. [31] were unable to determine precisely which regions of the brain were chimeric, these results demonstrate that species-specific neural development underlies auditory perception. In the songbirds, a variety of structures throughout the songbird brain have been suggested to encode auditory predispositions for conspecific songs, both in the auditory pathways shared with all birds and in a specialized and highly interconnected set of brain nuclei necessary for memorization and production of song (i.e. the song system [32]).

Electrophysiological experiments on the brains of young songbirds have demonstrated that neurons in brain regions involved in the perception and production of songs are sensitive to typical elements of conspecific vocalizations, even without previous song exposure. In juvenile white-crowned sparrows and zebra finches, neurons in the HVC (formerly known as the high vocal centre), part of the song system, respond strongly to species-typical sounds (table 1; [33,34]). However, innate auditory predispositions are unlikely to be a result of neural development in the HVC. First, in contrast to auditory regions in the thalamus, neuronal growth and connectivity in the HVC and other song system nuclei in the telencephalon begin relatively late in development, around day 10 in zebra finches [35], but auditory predispositions for conspecific song have been expressed as early as day 6 in golden-crowned sparrows, Zonotrichia atricapilla [18], likely before the onset of significant HVC growth. Similarly, the connectivity of the HVC suggests that its primary role is linking auditory feedback and memories with production, rather than being the source of neural representation of song [36]. Second, although both experiments tested juveniles isolated from song or other species-specific vocalizations at a relatively young age (white-crowned sparrows: greater than 4 days old; zebra finches: 21 days old), recent experiments suggest that some acoustic learning and, therefore, tuning of the auditory system, can occur even earlier during the embryonic stage [25]. Both these lines of evidence suggest that the source of innate auditory predispositions is more likely to be found in regions directly or indirectly projecting to HVC (i.e. further ‘down’ the ascending auditory pathway depicted in figure 1).

Both the HVC and the caudomedial nidopallium (NCM), a region in the avian forebrain implicated in encoding of behaviourally relevant conspecific songs [37], receive neuronal input from nuclei in the auditory forebrain and the ascending auditory pathway (figure 1). For example, field L2, the primary receptive area of thalamic input in the telencephalon, is sensitive to the frequency–temporal characteristics typical of conspecific songs [38]. Field L2 may therefore be important in filtering behaviourally relevant sounds, such as song, into the song system (box 1). In zebra finches, field L undergoes developmental changes during the memorization phase of song learning [49], which may be important in achieving selectivity to conspecific song-like sounds [50]. Limited experiments on zebra finches and domesticated songbird breeds suggest that, in principle, species-specific developmental programmes of field L and/or other regions in the ascending auditory pathway and the inner ear may provide a basis for innate auditory preferences. For example, neurons in field L are selective for conspecific song in female zebra finches raised without exposure to song [51]. In a domesticated breed, the Bengalese finch, Lonchura striata var. domestica, more relaxed song learning predispositions, relative to its wild progenitor, correlate with higher expression levels of neurotrophic factors in the auditory regions field L and NCM [11]. Neurotrophic factors regulate neural differentiation, growth and connectivity, suggesting that expression levels might influence the strength, or bounds, of auditory predispositions. Lastly, in domesticated canaries, Serinus canaria, mutations in the ear hair cells of the Belgian waterschager strain reduce perception of high frequency sounds, which biases them to imprint on and produce low frequency sounds [6].

Auditory predispositions are likely to be represented in the ascending auditory pathway or auditory periphery outside of the songbird song system and, potentially, distributed throughout a number of regions in the brain that, together, influence auditory perception at a developmentally early age [34,52]. How species-specific developmental programmes tune the auditory system during the earliest stages of neural differentiation is unknown [41], in part because development of the auditory system is the result of a complex interplay between genetic and learned factors [51]. Although the examples of domesticated finches and canaries may be special cases owing to artificial selection regimes, they suggest that auditory predispositions may rapidly evolve in response to changes in selection on songs. We suggest that a fruitful way forward in uncovering the neural substrate of innate auditory predispositions for conspecific songs would be to compare closely related species where auditory predispositions may be evolving to track incipient changes in songs themselves.

4. Uncovering the genetic basis of variation in auditory predispositions

Innate predispositions to learn conspecific songs, demonstrated in at least some songbird species, imply species-specific genetic variation in the development of the auditory system. The identity of the genes that underlie species-specific auditory perception remains unknown, in part because such genes are likely to be expressed very early in development. Research over the past 25 years has demonstrated that transcription factors regulating neural plasticity, such as ZENK, are expressed in auditory regions when birds hear song playbacks [37] and that the strength of expression is related to the accuracy of song learning [3], thereby relating gene expression with song memory formation. However, these results may offer only tentative clues into the nature of auditory predispositions. First, song-induced ZENK expression depends on previous song exposure and cannot be detected during early stages of development when auditory predispositions are known to be active [53]. Second, song-induced expression of ZENK and other related immediate early genes is highest in regions of the avian nidopallium (regions shaded in green in figure 1), but it is more likely that auditory predispositions arise via species-specific tuning of the auditory pathway that project into the nidopallium, such as in region L2 of field L.

Horita et al. [54] demonstrated that song playbacks to adult zebra finches induce expression of another immediate early gene, dusp1, in L2 and the thalamic auditory region, Ov (regions shaded in red in figure 1). The authors argue that dusp1 expression in L2 may inhibit ZENK expression in the
NCM, suggesting that song-induced expression of dusp1 could regulate memory formation, thereby serving as a marker of perceptual filtering in L2. However, Hortia et al. did not compare the expression profile of dusp1 in response to heterospecific songs, which would be necessary to test this hypothesis. Ongoing efforts have moved beyond single-gene expression studies to characterizing the suite of genes and regulatory elements that are induced by song exposure (RNA: [55], microRNA: [56]) and have suggested that patterns of gene expression may differ greatly across species [57] and developmental stages [58]. Similar studies specifically aimed at comparing song-induced transcriptomic profiles early in development across species may provide clues into the genetic basis of auditory predispositions. Some recent studies suggest that such comparative approaches may hold promise. For example, expression levels of genes involved in neural plasticity are related to variation in the vocal learning behaviour of the Bengalese finch and white-rumped munia [11]. Further, work on different canary breeds has revealed genetic loci that may underlie population differences in auditory perception and song learning biases [16]. Larger-scale approaches aimed at associating individual variation in auditory predispositions with genome-wide variation may provide insights into the genetic basis of within-species variation in song learning. Comparisons across species could progress using any detected candidate loci. Such efforts will be aided by ongoing efforts by members of the Avian Genome Consortium to sequence the genomes of all avian species [20,21]. The large number of bird species that learn their song and variability in how songs are learned will make it possible to investigate how variation in ecology, sexual behaviours, social structure and learning modes (i.e. open- versus closed-ended learners) may relate to variation in the genes guiding vocal learning in birds.

There remains much work to be done to uncover the genetic and neural basis of innate auditory predispositions across species, but such research may have far ranging consequences for our understanding of the evolution of vocal learning and its relationship with sexual selection and mate recognition. Both males and females express similar innate auditory predispositions [33], suggesting that auditory predispositions for conspecific song in both males and females are likely to have a common genetic basis [59]. A shared genetic basis for auditory predispositions in males and females would mean that the same genes guide the songs that males will learn to sing and that female will perceive as relevant signals to use when selecting their mates. An alternative, suggested by studies showing the neural substrate of song preferences in females may be different from that in males [43], is that different sets of genes underlie female and male auditory predispositions. Such genes, especially if they are shared by both sexes, may be important in the build-up and maintenance of reproductive barriers across species, as divergence in gene expression-induced developmental programmes underlying auditory perception may lead to species assortative mating. Thus, the existence of innate auditory predispositions may connect the evolution of vocal learning with sexual selection and speciation processes.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

Acknowledgements. D.W. is supported by a postdoctoral research fellowship from the National Science Foundation (www.nsf.gov) and a research grant from the Nilsson-Ehle Donationerna (www.fysiografen.se). A.Q. is funded by the Swedish Research Council.

References
