Morphological basis for the evolution of acoustic diversity in oscine songbirds

Tobias Riede and Franz Goller

Department of Biology, University of Utah, Salt Lake City, Utah 257 S 1400 E, USA

Acoustic properties of vocalizations arise through the interplay of neural control with the morphology and biomechanics of the sound generating organ, but in songbirds it is assumed that the main driver of acoustic diversity is variation in telencephalic motor control. Here we show, however, that variation in the composition of the vibrating tissues, the labia, underlies diversity in one acoustic parameter, fundamental frequency ($F_0$) range. Lateral asymmetry and arrangement of fibrous proteins in the labia into distinct layers is correlated with expanded $F_0$ range of species. The composition of the vibrating tissues thus represents an important morphological foundation for the generation of a broad $F_0$ range, indicating that morphological specialization lays the foundation for the evolution of complex acoustic repertoires.

1. Introduction

The relationship between structure and function plays an important role in the evolution of behaviour. For vocal behaviour, the structure of the vibrating tissue is critical for determining a range of acoustic features [1–3], but this relationship has only been investigated in detail in humans [4,5]. One of the most important acoustic features of animal vocalizations is fundamental frequency ($F_0$). The large range of $F_0$ across species or within the vocal repertoire of individual species may carry important information about the sender and therefore play a major signalling role in the context of natural and sexual selection [6]. However, little specific information is available about the various factors involved in variation of frequency range both within and between species.

In avian interspecific comparisons, $F_0$ is correlated with body size, but a substantial portion of the variation is not explained by this relationship [7]. However, explaining the $F_0$ range is important because it defines the boundaries for sound frequency of a vocal repertoire and thus poses limits on spectral diversity. Because vocal diversification has played an important role in speciation and the ecological success of birds, it is important to gain a thorough understanding of how frequency range is determined in birds and thus contributes to these processes.

Airflow-driven vibration of tissue is the main mechanism of sound production in birds and mammals, including humans. The vibrating tissues of the mammalian larynx and the avian syrinx are composed of extracellular matrix [8,9]. Its composition of collagen and elastin fibres and hyaluronan ensures that the tissue can engage in self-sustained vibrations and, at the same time, can withstand the physical forces of vibration [10]. The mechanical properties of these vibrating tissues influence the oscillation frequency and, thus, $F_0$ of the generated sound [11]. The tension of the vibrating tissue can be adjusted by muscle action and is therefore controlled by the vocal motor areas in the brain [12,13]. However, neural control can only adjust $F_0$ within the physical limits, which are determined by the morphology and composition of the vibrating tissue. Previous work in songbirds [14] and mammals [15] showed that mechanical differences of labia are associated with vocal differences between species. In vivo experiments demonstrated the role of labia in $F_0$ determination and the role of muscles in labia adjustment [16–19], and computational modelling illustrated how muscle control and labial dynamics contribute to generate diverse features of sound [20]. Despite this progress, we still lack a thorough understanding of how morphology and neural control interact during the evolution of acoustic signals [21,22].
To address this gap in our knowledge, a more detailed understanding of the extracellular matrix composition of the syringleal labia is required [7]. The goal of this study was therefore to investigate the relationship between microstructure of the vocal organ and F0 in oscine songbirds. The oscine syrinx contains two independently controlled sound sources whose F0 ranges do not overlap fully [16–20]. Oscines acquire song, the most complex vocalizations of their acoustic repertoire, through vocal learning. It is thought that the forebrain control circuits associated with learned vocal behaviour enabled the evolution of increased acoustic diversity [23]. Neural control of the dual sound source undoubtedly plays a major role in acoustic versatility. However, the morphology of these two sound sources must present a mechanical basis for generating diverse sound features. Morphological complexity was investigated in two complementary ways to address the question of how design of the vibrating tissue (labia) is related to the vocal repertoire of a species. First, we expect that an increased morphological asymmetry between the two sound sources is associated with a large F0 range. In a second approach, we investigate the morphology of the labia, because data from human vocal folds indicate that the evolution of a layered structure increases versatility of vocal folds as a sound source [8]. Acoustic energy conversion is tightly linked to how labia oscillate inside the syrinx [19]. We therefore expect that species with a wider frequency range demonstrate more complex labial morphology.

2. Material and methods

(a) Study animals

Eight songbird species were selected to maximize variation in body size, and to maximize variation in song F0. All sound recordings were made in the same populations from which specimens for morphological investigations were collected. Zebra finches (Taeniopygia guttata) were bred at the University of Utah. Other species (see electronic supplementary material, table S1) were captured in the Salt Lake Area, Utah, USA between March and July 2008–2011. Data from zebra finches, white-crowned sparrows (Zonotrichia leucophrys) and European starlings (Sturnus vulgaris) were in part collected for previous studies [9,24,25]. The birds were deeply anaesthetized with a Ketamine/Xylazine mixture (Sigma-Aldrich, 1:1 by volume) and perfused intracardially with PBS. Only males were used in this study.

(b) Tissue preparation

The syrinx was excised and fixed for 3 days in 10% neutral-buffered formalin, decalcified for 8 h, processed and embedded for coronal sections in paraffin [9,23]. Adjacent 5 μm sections were exposed to one of the following stains: haematoxylin and eosin (H&E); elastica van Gieson stain (EVG); Masson’s trichrome stain (TRI). Micrographs were taken with a digital camera (Axiocam HRc, Carl Zeiss, Germany) combined with an Axioplan microscope (Axioplan, Carl Zeiss, Germany) and computer software (AXIOPHOTEN v. 4.0, v. 4.6.3.0., Carl Zeiss, Germany).

(c) Labial morphology

Labial composition and morphology of eight songbird species (see electronic supplementary material, table S1) were examined with established histological techniques. Individual fibrous protein components were visualized and quantified. Based on stained tissue, we constructed a schematic cross-section through the medial labium, showing the abundance of collagen and elastin fibres as well as the most prominent orientation of the two fibre types (figure 1a,b). Five basic extracellular matrix designs were found and described as layers (layer numbers are arbitrary; electronic supplementary material, table S2). Percentages for fibrillar proteins refer to positively stained areas in EVG and TRI stains, respectively, analysed with IMAGEJ (1.41o; NIH, USA):

- layer 1: less than 10% collagen, more than 60% elastic fibres, elastic fibres in crano-caudal orientation;
- layer 2: less than 5% elastic fibres, more than 60% collagen fibres, collagen fibres randomly oriented;
- layer 3: less than 5% elastic fibres, more than 60% collagen fibres, collagen fibres dorsoventrally oriented;
- layer 4: less than 5% elastic fibres, more than 60% collagen fibres, collagen fibres crano-caudally oriented; and
- layer 5: less than 5% elastic fibres, 30–60% collagen fibres, collagen fibres randomly oriented.

Labial cross-sectional areas were measured at five different levels of equal distance along the ventral to dorsal axis. The volume of labia was estimated from labial cross-sectional area and total length of labia using equation (2.1):

\[ \text{Vol} = \sum_{i=1}^{5} (\text{area} \times \text{length}) \]

where Vol is the volume of a labium, estimated from the sum of products of cross-sectional areas and the distance (length) between two sections of the syrinx.

(d) Analysis of the vocal repertoire

F0 was quantified every 20 ms for songs and calls using a pitch-tracking module (PRAAT software, v. 5.2.12). Results were visually confirmed (figure 2a,b). The F0 data were represented as histograms (0–9 kHz, 100 Hz resolution). Mean F0 (F0-mean) was calculated by averaging all 100-Hz frequency bins in the histogram, thus weighting different frequency bins according to the rate of their occurrence. The F0 limits within the song of a species were calculated in two ways. First, F0-range-A was calculated as the sum of the 100-Hz-bins. Second, F0-range-B was estimated from the F0 limits by calculating the difference between the minimum and maximum F0. While F0-range-B considers the upper and lower boundaries within which sound is produced, F0-range-A captures the actually used spectral range covered by a vocal organ.

For each species, one song from between three to six individuals was submitted to F0 tracking. The sample size of audio recordings was limited to these relatively small numbers for two reasons. First, the relationship between syrinx morphology and song F0 must be explored with samples from the same population to account for potential differences in the learned aspects of song. Second, averaging across many individuals may allow for detailed characters of a vocal repertoire. For example, in a species whose song is composed of two non-overlapping distinct frequencies within the total range, the gap between these frequencies will disappear if individuals vary in the respective distinct frequencies.

3. Results

Marked differences between species were found in the distribution and orientation of individual fibre components within the labia (figure 1). For each species, we categorized and quantified the number of different layers according to fibre orientation and density (figure 1b) and established a schematic representation (figure 1a). For example, the labia of golden-crowned kinglets and white-crowned sparrows display a deep layer with prominent presence of elastic fibres (layer 1), followed by a layer of loose connective tissue with
lower elastin content (layer 2). The zebra finch labia consist of a homogeneous single layer (layer 2), while the labia in ruby-crowned kinglets, European starlings and yellow-headed blackbirds exhibit multiple layers (figure 1c). Although the left medial labia tend to be larger than the right medial labia, the layer structure was identical on the two sides with one exception. In European starlings, we found four layers on the left side (layers 2, 3, 4, 5) and two layers (layers 4, 5) in the much smaller labia of the right syrinx.

We estimated labial size by measuring labial cross-sectional area in serial sections of the syrinx from anterior to posterior to assess asymmetry between the left and right sound sources for each species. In all investigated species, the area of the left labia was at least somewhat larger than that of the right labia, but this lateral asymmetry varied strongly between species, especially for the medial labia. The difference in medial labia ranged from 0% in the zebra finch to a more than fourfold larger left medial labium in the European starling (electronic supplementary material, table S3).

We then explored these two datasets for correlations between morphology, F0 and F0 limits. We tested whether body mass can explain the F0-range-A, F0-range-B or F0-mean of songs in these eight species. Body size should be a good
Figure 2. Fundamental frequency ($F_0$) tracking procedure and quantification exemplified for (a) black-billed magpie and (b) ruby-crowned kinglet songs. Songs are shown as oscillograms (top panel) and spectrograms (middle panel) and the tracked $F_0$ (bottom panel). $F_0$ data points for these songs are displayed as histograms (c,d) showing the occurrence of each frequency between 0 and 9 kHz (100 Hz bins). Whereas magpie song is composed of all frequencies between 0.5 and 3 kHz, song in the ruby-crowned kinglet demonstrates two distinct distributions with maxima around 3 and 6 kHz. The two frequency ranges are not connected, and probably represent the independent contributions from the two sides of the syrinx. (e) Distributions for $F_0$ in song (bars, upward) and calls (bars, downward; stars indicate data obtained from published spectrographic images) for all eight species. For each species, one song from between three to six individuals was submitted to $F_0$ tracking and data were pooled for histograms. Species abbreviations as in figure 1.

predictor for $F_0$ since labial size scaled with body size (figure 3a), and absolute size of the labia explained $F_0$-mean (figure 3b). However, body mass differences explain only 35% of the variance in $F_0$-mean (figure 3c), which falls within the range of earlier analyses of larger datasets [26,27]. More importantly, body mass variations appear neither related to the $F_0$-range-A calculated by weighing all frequencies in the song repertoire of a species, nor the $F_0$-range-B (figure 3d,e).

Two features of the labia show a significant correlation with $F_0$-range-A. First, the degree of asymmetry of labial cross-sectional area at mid-organ level is positively correlated with the $F_0$-range-A (figure 3f). This relationship is particularly strong for the medial labia, which show more variation in lateral asymmetry than the lateral labia in these eight species (see electronic supplementary material, table S4). For $F_0$-range-B only the relationship for the medial labia was significant (see electronic supplementary material, table S4). Asymmetry should lead to different $F_0$ limits of the two sound generators with minimal overlap and, thus, expand the combined $F_0$ range that can be produced by the dual sound source. This interpretation is supported by physiological data, which show that the two sound generators contribute different frequency ranges in a number of species [16–18]. The European starling, with the largest asymmetry in this dataset, also generates the largest $F_0$-range.

Second, the number of different layers within the labia is positively correlated with the $F_0$-range-A but not $F_0$-range-B (figure 3h and electronic supplementary material, table S4). $F_0$-mean is not significantly correlated with labial asymmetry or the number of layers (see electronic supplementary material, table S4). A layer structure provides a morphological basis for anisotropic behaviour. Wave propagation in the soft tissue of the labia is one essential component of self-sustained vibratory behaviour during sound generation [2,20], and these waves propagate differently in the soft material depending on the direction, magnitude and rate of deformation. Effectively, the inhomogeneous structure can enlarge the frequency range or support specific frequency ranges by the potential for engaging different portions of the tissue into vibration or beneficially influence the vibration characteristics [4].

4. Discussion
The combination of two morphological adaptations in the labia of songbirds lays the foundation for extending $F_0$-range and for generating the potential to produce all frequencies within this
range and shifting F0-mean presumably with the potential of greater specialization. First, the asymmetry in labial size defines the lower and upper frequency boundaries as well as the wider range of actually produced frequencies within these boundaries. Second, the inhomogeneous composition of the labia (layer structure) permits production of all frequencies throughout the range, and thus also the potential for continuous frequency modulation across the entire range. These specializations of the extracellular matrix of the vibrating tissue are therefore essential for the production of acoustic diversity and lay the foundation for neural control of sophisticated vocal repertoires.

The relationship between labial composition and used frequency range (F0-range-A) is stronger than that found for the simple lower and upper boundaries of the frequency range (F0-range-B). These results indicate that the mechanism for continuous frequency modulation and generation of frequencies within a broad range rests on the degree of inhomogeneity within the labial extracellular matrix. The data also show that quantifying used frequencies within the lower and upper frequency boundaries is a meaningful and functionally relevant acoustic descriptor of a vocal repertoire.

It remains to be tested how variability in labia design contributes to acoustic variability within species, particularly between singing males. Studies in songbirds demonstrate mixed results [28–30] for the relationship between body size and song frequency (‘dominant’ or ‘peak’ frequency was measured as frequency with the highest amplitude in spectrum). A lack of close relationship could be explained by an uncoupling between labial size and F0, which might be more dependent on species-specific labia design. Available data do not permit predictions on the relationship between F0 and body size within species. Although the labia show species-specific characteristics, individual-specific factors (e.g. hormones, stress, nutrition and hydration) are likely to influence their morphology and mechanical properties and may do so more readily than changes in body mass or skeletal body size.

The nonlinear relationship between stress and strain (figure 4) presents a complex target for frequency control for two main reasons. First, the active movements required for regulating
Figure 4. (a) Hypothetical stress–strain relationship for connective tissue as found in the labia of the songbird syrinx. The relationship between stress, strain and fundamental frequency of sound is established via the string model, which explains that oscillation frequency is determined by tissue stress, string length and density. However, the impact of stress depends on the range of operation (b). In the low-strain region, relatively large tissue deformation is necessary to achieve a given change in stress (box 1) and thereby to cover a large frequency range. The same magnitude of change in stress can be achieved in the high-strain region with much less change in tissue deformation (box 2). This nonlinear relationship permits greater stability in the low-strain region. For example, a syllable with constant F0 requires constant re-adjustment if produced in the high-strain region because tissue relaxation occurs more rapidly (i.e. energy loss due to heat) than in the low-strain region. The ‘disadvantage’ of the low-strain region is (i) less fine-control, (ii) more change in strain is required to cover a large F0 range, and (iii) less frequency flexibility (due to drag and resistance associated with larger movement amplitudes). A potential mechanism that counters the requirement for larger amplitude active movements in the low-strain region could be a multi-layer tissue design, which allows differential recruitment of tissue. The ‘differential tissue recruitment’ could cause higher stresses in the oscillating tissue aspects facilitating high oscillation frequencies, even at lower strains. Stress is the ratio of force over cross-sectional area. By reducing the volume of oscillating tissue while the same aerodynamic forces apply, greater local stresses inside the oscillating tissue should occur.

5. Caton T, Thibeault SL, Klemuk S, Smith ME. 2007 Viscoelasticity of hyaluronan and nonhyaluronan