Published in final edited form as:

Curr Biol. 2017 September 11; 27(17): 2677–2683.e3. doi:10.1016/j.cub.2017.07.059.

Evolution of vocal diversity through morphological adaptation without vocal learning or complex neural control

Sarah M. Garcia[1,6,*], **Cecilia Kopuchian**[2,3], **Gabriel B. Mindlin**[4], **Matthew J. Fuxjager**[5], **Pablo L. Tubaro**[3], and **Franz Goller**[1]

[1]Dept. of Biology, University of Utah, 257 South 1400 East, Salt Lake City, Utah 84112, USA

[2]CECOAL (Centro de Ecología Aplicada del Litoral) CONICET, Corrientes, Argentina

[3]División Ornitología, Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia' MACN-CONICET, Avenida Ángel Gallardo 470, Ciudad Autónoma de Buenos Aires, Argentina

[4]Depto. Física, FCEyN, Universidad de Buenos Aires, C. Universitaria, Pab I, Buenos Aires, Argentina

[5]Department of Biology, Wake Forest University, Winston-Salem, NC 27109, USA

SUMMARY

The evolution of complex behavior is driven by the interplay of morphological specializations and neuromuscular control mechanisms [1–3], and it is often difficult to tease apart their respective contributions. Avian vocal learning and associated neural adaptations are thought to have played a major role in bird diversification [4–8], whereas functional significance of substantial morphological diversity of the vocal organ remains largely unexplored. Within the most species rich order, Passeriformes, 'tracheophones' are a suboscine group that, unlike their oscine sister taxon, does not exhibit vocal learning [9] and is thought to phonate with tracheal membranes [10, 11] instead of the two independent sources found in other passerines [12–14]. Here we show tracheophones possess three sound sources, two oscine-like labial pairs and the unique tracheal membranes, which collectively represent the largest described number of sound sources for a vocal organ. Birds with experimentally disabled tracheal membranes were still able to phonate. Instead of the main sound source, the tracheal membranes constitute a morphological specialization, which, through interaction with bronchial labia, contributes to different acoustic features such as: spectral complexity, amplitude modulation, enhanced sound amplitude. In contrast, these same features arise in oscines from neuromuscular control of two labial sources [15–17]. These findings are supported by a modeling approach and provide a clear example for how a morphological

[*]Correspondence: Sarah.Garcia@utah.edu.
[6]Lead Contact

Author Contributions

S.M.G., C.K, M.J.F. and F.G. conducted the experiments. S.M.G., G.B.M. and F.G. performed analysis. G.B.M. developed the model. All authors wrote and edited the manuscript.

There are no conflicts of interest.

adaptation of the tracheophone vocal organ can generate specific, complex sound features. Morphological specialization therefore constitutes an alternative path in the evolution of acoustic diversity to that of oscine vocal learning and complex neural control.

Keywords

syrinx; tracheophone; suboscine; functional morphology

RESULTS AND DISCUSSION

The avian vocal organ, the syrinx, shows remarkable morphological diversity across taxa [12], yet its significance for acoustic behavior is poorly understood. Notably, the syringes of oscines are quite similar [13], yet oscines are considered to have the broadest range of acoustic features in their vocal repertoires. This broad range has been attributed in major part to vocal learning and complex neural control [4–6]. In contrast, the syringes of their sister group, suboscines, show remarkable morphological diversity [14]. In particular, tracheophones have a unique set of membranes on the dorsal and ventral surfaces of the trachea directly above the tracheal bifurcation, the *membranae tracheales* (*MT*) [11, 14]. These tracheal membranes, in addition to the unique skeletal elements (the *processi vocales*) found on their sides, constitute a synapomorphy for the tracheophone clade [14]. Based solely on anatomical inference, these membranes have been considered the main sound source of the tracheophone syrinx, although the presence of bronchial sources has been recognized [10, 14]. However, here we show that sound generation in tracheophones relies on more complex morphology than commonly assumed. The presence of three vibratory sources gives rise to diverse sound features and may have played an important role in the diversification of this clade.

In order to investigate potential sound sources, we used a fiberscope inserted into the trachea (Figure 1) to visualize the syrinx while we induced phonation by injecting air into the air sac system [18]. In all six species investigated (*Thamnophilus caerulescens*, variable antshrike; *Cercomacroides tyrannina*, dusky antbird; *Furnarius rufus*, rufous ovenbird; *Syndactyla rufosuperciliata*, buff-browed foliage-gleaner; *Lepidocolaptes angustirostris*, narrow-billed woodcreeper; *Certhiaxis cinnamomeus*, yellow-chinned spinetail; Figure S1), the *MT* were drawn into the trachea upon pressurization, and were induced to oscillate under certain flow conditions (Movies S1 and S2). In addition, all species possess a labial sound source in each bronchus, which seem to be the main sound generators for vocalizations in most tracheophones. We used two approaches to verify this conclusion. First, we visualized the syrinx, and once the *MT* began to oscillate, we pushed the fiberscope deeper so as to block their vibrations and see past them. After pushing past the *MT*, we were able to see the bronchial labia, and they continued to oscillate and produce sound (Figure 1, Movie S1). Second, we prevented oscillations of the *MT* by either rupturing the ventral *MT*, or reducing elasticity by applying tissue adhesive to the ventral *MT* (these manipulative data are unavailable for *C. cinnamomeus*, and so this species is excluded from quantitative results). Despite these treatments, all species except the dusky antbird were still able to phonate spontaneously. These treatments resulted in specific acoustic changes, which indicate that in

the intact tracheophone syrinx, all three sound sources can interact to produce complex acoustic features (Figures 2 and 3).

Consistent across all species, the intact syrinx produced low frequency sounds, characterized by vibrations with pulse-like time-waveforms, giving rise to complex harmonic content of sounds. Disabling of the *MT* resulted in three consistent changes. First, pulse-like quality decreased as indicated by the increased duty cycle (Figure 4). This decrease in pulse-like quality was found in all 4 species for which data after manipulation of the membrane are available. Second, sound amplitude significantly decreased (Figure 4D). Lastly, sound frequency showed a consistent increase after membrane disabling (Figures 2, 3, and S2), though the degree of increase varies between species.

Strikingly, the prominent amplitude modulations in spontaneously generated distress calls of the rufous ovenbird (*F. rufus*) were absent after ventral *MT* rupture (Figure 2). The low frequency component representing the vibration of the *MT* was strongly reduced after the membrane was disabled with various approaches (membrane rupture, membrane gluing, and fibroscopic prevention of vibration) for both spontaneous distress calls (Figures 2A and 3A) and induced phonation (Figure 3B). This strongly suggests that the *MT* produce the low-frequency vibrations that drive the change in amplitude, which is superimposed on the higher-frequency labial oscillations. This proposed mechanism for the generation of amplitude modulation involves interaction of the three sound sources. To confirm that the amplitude modulations are not produced by neuromuscular control, we denervated the syringeal muscles by transection of the tracheosyringeal nerves, and amplitude modulation clearly persisted (Figure S3).

To better understand these interactions, we used a mathematical model of the syrinx [19] to generate sound with two different oscillators representing one set of labia and the *MT*. If the frequencies of the two sources are sufficiently different, the lower frequency of the *MT* modulates the higher frequency of the labial sound source (Figure 2C, “pulsations”) as seen in the pronounced amplitude modulation in the spontaneous calls of the rufous ovenbird (Figure 2A). If the frequencies are more similar to each other, the two sources become “locked” and the resulting vibrations are of lower frequency than if the labial source were to vibrate on its own (Figure 2C, “locking”). This is consistent with the observed shift to higher frequencies after membrane disabling across species (Figure 3), as well as reduced pulse-like quality of the vibrations (Figure 4) [20, 21]. The pulsatile oscillations arise from the nonlinear interaction of two oscillators, the MT and the labia. We assume that the labial oscillators can be represented by a harmonic oscillation of higher frequency. For this reason, when the MT is disabled, the remaining labial sources are free to oscillate at their natural, higher frequency and with more harmonic character [22]. In the other species, frequency also significantly increases, but the increase is smaller than that found in *F. rufus*. This further illustrates the diversity by which the MT affects sound features in different tracheophones. In other words, the MT and bronchial labia of *F. rufus* likely vibrate at significantly disparate frequencies (as supported by data from the model and the presence of amplitude modulation) which therefore results in a large shift to higher frequencies after MT manipulation, while the variable antshrike shows a much smaller shift to higher frequencies, indicating the MT and bronchial labia likely vibrate at more similar frequencies.

In our simulations, we found that sounds generated with locked sound sources tend to be of higher amplitude than those generated with only one sound source. Consistent with that observation, the amplitude of spontaneous distress calls of individuals, for which data are available, decreased markedly after the membrane was disabled (Figure 4D). To quantify sound amplitude before and after membrane disabling in spontaneously generated distress calls, we expressed amplitude as signal amplitude (Volts) relative to background noise (Volts) to account for different recording settings. This reduction in sound amplitude was consistent for all manipulations (note: the *L. angustirostris* individual did not produce spontaneous distress calls after manipulation, and so amplitude data are unavailable) and can therefore not have resulted from altered pressure differentials across the syrinx after membrane rupture. Lastly, simulations of the model indicate that the transition between pulsations and locking (Figure 2C) occurs with certain changes in pressure. This may explain in part the diversity with which tracheophones seem to use the *MT* (Figures 3 and 4), since depending on the natural pressure produced by the bird, they may or may not produce sounds with both pulsations or locking, and may instead produce one or the other.

The *MT* and their interactions with the bronchial sound sources in the tracheophone syrinx constitute a morphological solution to generating loud sounds, amplitude modulations, pulse-like quality, and low frequency sounds. Ultimately, these acoustic features could be attributed to the increased resistance across the tracheophone syrinx when the *MT* are engaged as they cause a narrowing of the trachea. While oscines achieve some of these acoustic features through neuromuscular control of their two sound sources [20, 21, 23–26], suboscines do not have the same level of neuromuscular control [12, 28] and instead likely rely on intrinsic, morphological qualities such as the *MT* to achieve these acoustic features. Our data support this, and consistently indicate that the intact *MT* enable the generation of higher amplitudes, increased pulse-like quality, and lower frequencies. In addition, different tracheophone species seem to use the *MT* to different effect, as not all species in this study produce prominent amplitude modulations. Of those that do (*F. rufus, L. angustirostris*), some may rely on the bronchial labia to do so (*L. angustirostris*; modulations persist after manipulation); and most strikingly, the dusky antbird (*C. tyranina*) is unable to produce any vocalizations (induced or spontaneous) after ventral *MT* manipulation. In the case of the dusky antbird, the loss in *MT* function may reduce the resistance across the syrinx beyond a critical threshold, resulting in its inability to phonate. While seemingly disparate, these data highlight the diversity with which the *MT* are used in the tracheophone syrinx.

This differential use of the *MT* in tracheophones suggests that reliance on morphological specialization to achieve diversity may be inherently limiting. While there could be energetic benefits due to the lack of dedicated specialized central nervous system structures, oscines are constrained by morphology to a certain extent [29]. In contrast, a larger number of acoustic features for a given morphology may be possible through complex neuromuscular control. In tracheophones, and suboscines in general, innovation of an acoustic feature likely needs to be accompanied by specific morphological changes, and so the scope and timescale of evolutionary change may be narrower. In other words, generating new acoustic features via evolution of a new morphological structure (such as the *MT*) may take longer than generating such features through plastic neural control of existing structures. In addition, different acoustic variables appear to be linked through specific morphological adaptations,

and therefore might not be able to change independently. A common morphological origin of specific acoustic features will therefore limit independent diversification and constrain its evolutionary pace.

Despite the presence of three sound sources, there is no evidence in our data or the vocal repertoires of tracheophones that their syrinx is used to generate multi-voiced sounds as seen in the two-source oscine syrinx. Instead, even the two labial sound sources appear to interact and "lock". This is in contrast to oscines, in which the two labial sound generators are used to simultaneously generate independent tones [15, 23, 30], while the two sources can also be "locked" during generation of pulse tones [20, 29]. Locking of oscillators in tracheophones, as in other multi-source vocal organs, facilitates generation of a variety of acoustic features, including pulse tone sounds with rich harmonic content, amplitude modulation and increased sound amplitude.

The amplitude modulation seen in *F. rufus* (up to 650 Hz) represents a unique production of amplitude modulation not seen in oscines. To produce high-frequency amplitude modulations, oscines typically use the two voices to generate difference tones [25, 29]. In addition to difference tones, oscines can use muscular control of the syrinx (gating) to produce amplitude modulations up to 260 Hz [31, 32]. Neither mechanism appears to be used by *F. rufus*; bilateral transection of the tracheosyringeal nerve does not impair amplitude modulation indicating they are not using gating, nor are there any known instances of multi-voice phenomena in *F. rufus*. This, coupled with the data resulting from experimental manipulations, strongly suggests that the *MT* represent a passive morphological solution to producing amplitude modulation, which in oscines generally requires complex neural control.

The data presented in this study are the first to highlight the unique and previously unknown functions of the tracheophone syrinx. These data confirm all three sound sources of the tracheophone syrinx and the prominent role of the bronchial membranes in sound production. To date, this is the largest documented number of sound sources for a vocal organ. The *MT* clearly play a role in generating specific acoustic features (loud amplitude, low frequency, amplitude modulation, and pulse-like quality), and constitute an elegant and passive morphological solution to generating these features without complex neuromuscular control as seen in oscines.

The results of this study illustrate that diversification of acoustic features in the vocal repertoires of birds occurred via different adaptations. Whereas suboscines in the absence of vocal learning evolved morphological mechanisms, oscines show much less morphological diversification of the syrinx, but produce similar acoustic effects through diverse neuromuscular control. The latter mechanism permits production of a broad range of acoustic features by one syringeal morphology, whereas the path using morphological diversification is more restrictive to specific sound characteristics. However, the evolution of sophisticated neural control may have required concurrent evolution of vocal learning, which in turn necessitated major changes in central control, such as the evolution of a dedicated forebrain circuitry [33, 34]. Nonetheless, each adaptive mechanism facilitated the evolution

of vocal diversity between species and, thus, may have played a significant role in the remarkable radiation in suboscines and oscines.

STAR Methods

Contact for reagent and resource sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sarah M. Garcia (Sarah.Garcia@utah.edu).

Experimental model and subject details

Individuals—These experiments used wild caught, male birds in the breeding state with non-regressed testes. The species investigated were: *Thamnophilus caerulescens*, variable antshrike; *Cercomacroides tyrannina*, dusky antbird; *Furnarius rufus*, rufous ovenbird; *Syndactyla rufosuperciliata*, buff-browed foliage-gleaner; *Lepidocolaptes angustirostris*, narrow-billed woodcreeper; *Certhiaxis cinnamomeus*, yellow-chinned spinetail. Please see Figure S1 for their dendrogram.

Netting—Individuals were captured with mistnets in November of 2014 and 2015 at the Estación Biológica de Corrientes (EBCo) in Corrientes, Argentina (27,55095° S 58,68441° W), under permits issued by the Fauna Province Direction and the Provincial Natural Reserve Area. Dusky antbird (adult males) individuals were captured in March of 2015 in Gamboa, Panama with permission from the Autoridad Nacional del Ambiente and the Smithsonian Tropical Research Institute (STRI). These institutions approved the experimental procedures used on these individuals (SE/A-6-15, Gamboa; 17-11-2014, 20-10-2015, Corrientes). All experiments were conducted at the respective field sites.

Husbandry—Individuals were processed within 3 hours of capture. During this time, they were individually kept in low-light conditions to limit stress. They were not fed so as to prevent emesis during surgery. Individuals that did not readily produce spontaneous distress calls (see below) were assigned to experiments of induced phonation (see below).

Method Details

Distress calls before and after manipulations—Distress calls were elicited before and after syringeal manipulation (nerve cut, vMT rupture, vMT immobilization) by simple handling. They were recorded with an Audiotechnica AT 8356 microphone and Marantz (PMD 660) recorder while birds were held ~30 cm away from the microphone.

Nerve cut—After spontaneous distress calls were recorded, we removed a segment of either the left or right tracheosyringeal branch of the hypoglossal nerve[21]. To gain access to the nerve, a small amount of Cetacaine topical anesthetic gel was applied to the skin surface, and a small incision was made above the trachea. We dissected the tracheosyringeal nerve free of connective tissue on one side and removed a ~2 mm segment of the nerve. We temporarily closed the incision with tissue glue (Vetbond), and then repeated the recording of spontaneous distress calls. After distress calls were recorded, we removed a segment of

the tracheosyringeal nerve from the other side of the trachea, resulting in a bilateral transection of the nerve and repeated the recording of spontaneous distress calls.

v*MT* rupture—After spontaneous distress calls were recorded, we removed a rectangular section of the ventral *Membrana trachealis* (v*MT*) of the syrinx. We did this by applying a small amount of Cetacaine on the skin and the interclavicular air sac membrane prior to making incisions to gain access to the distal end of the trachea. We then removed a ~2×2 mm rectangle of the v*MT*. After removal of the v*MT*, we closed the air sac and skin with tissue adhesive. After a short recovery period (5–10 min), we again recorded spontaneous distress calls.

v*MT* gluing—After distress calls were recorded, we applied tissue glue to the v*MT* of the syrinx using the same approach to gain access to the syrinx as described above. Subsequently, the v*MT* was covered with tissue adhesive as well. We then closed the air sac and skin and again recorded spontaneous distress calls.

Observation and membrane block with a fiberscope—We anesthetized individuals with a mixture of Xylazine (5μl)/Ketamine(25μl), and inserted a flexible cannula into a thoracic air sac. We then fixed the cannula to the rib cage with a suture, and sealed the insertion site with tissue glue. After this, we cut an opening (~5mm^2) in the trachea (2–3 cm above the syrinx) through which we passed the fiberscope (Hawkeye Precision Boreoscope, 0.9 mm diameter). Holding the fiberscope steady just above the *MT*, we recorded video (Sony XC 505 camera) and audio during induced phonation by applying pressure through the air sac cannula [18]. Since all experiments were conducted in the field, we were not able to measure air sac pressure. We gently compressed the chest cavity during induced phonation to facilitate pressurization of the air sacs and therefore phonation. This process was repeated with the fiberscope tip both above and below the *MT* when possible, allowing us to block the *MT* from engaging.

Mathematical modeling of the syrinx—The model used to simulate the time traces displayed in Figure 2C reads as:

$$\frac{\mathrm{dx}_1}{\mathrm{dt}} = y_1$$

$$\frac{\mathrm{dy}_1}{\mathrm{dt}} = 3\gamma_1(10\mu_1 - {x_1}^2)y_1 - {\gamma_1}^2(x_1 + 0.8{x_1}^3)$$

$$\frac{\mathrm{dx}_2}{\mathrm{dt}} = y_2$$

$$\frac{dy_2}{dt}=\mu_3\gamma_2(10\mu_2 - {x_2}^2)y_2 - {\gamma_2}^2(x_2+0.8{x_2}^3)$$

with parameters (μ_1, μ_2, μ_3), for the sound in the left of Figure 3C,

$$\mu_1=0.1+0.25\ \sin(\pi t/1.8) - 0.25x_2$$

$$\mu_2=-0.1+0.25\ \sin(\pi t/1.8)+0.1x_1$$

$$\mu_3=6.0$$

and, for the sound in the right of Figure 3C:

$$\mu_1=0.1+0.25\ \sin(\pi t/1.8) - 0.25x_2$$

$$\mu_2=-0.5$$

$$\mu_3=6.0$$

The equations above rule the dynamics of the tracheal-bronchial labia (x_1), and the *MT* (x_2) [35], and were modeled in C++. We assume that the two trachea-bronchial sound sources placed at the junctures of the bronchi are synchronized, and therefore they are treated as a unique sound source. The gain parameters μ_1, μ_2 represent the transluminal pressure at the labia and *MT* respectively. The last term in the expression defining each μ_i describe the modulations to the transluminal pressures, which are induced by the airflow fluctuations due to *MT* and labial oscillations respectively. In this model, the parameters γ_i are time scale factors, and in our simulations take the values (γ_1, γ_2) = (5000,4000). The synthetic sound is generated by linearly superposing the contributions of both sound sources.

If the parameters of the two oscillators are very different, their dynamics will not lock, and the sound will consist of a high frequency oscillation (the labial oscillation), modulated by the *MT* oscillation (which for the value of μ_3 in our simulations, is pulse-like). If the two frequencies are similar, the two types of sound sources will couple, and the sound will consist of a high frequency fluctuation (at a lower frequency than the one at which the labia would oscillate in absence of the *MT*). With the choice of μ_3 = 6, the labial and *MT* oscillations will present, for small values of the pressure μ_1, dynamics sufficiently different so that pulsations arise (first part of the sound displayed at the left of Figure 3C). As the pressure μ_1 is increased, the two oscillators lock (second part of the same sound). Disabling

the MT ($\mu_2 = -0.5$) turns the MT oscillator off. Now the sound is generated by the labial oscillations, at their autonomous frequency.

Quantification and statistical analysis

Frequency extraction—All recorded phonations were analyzed in Praat (software by P. Boersma and D. Weenink), using manual pitch extraction of individual sound pulses. When amplitude modulation was present, the fundamental frequency was extracted in addition to the amplitude modulation frequency. For histogram generation (Figure 4), frequency was extracted every 5ms and placed in 100 Hz bins. Bins were then normalized relative to the bin with the most occurrences. Frequency range was calculated by visual inspection of spectrograms (Xeno-canto.org) and power spectra of the lowest and highest frequencies. We performed a Kolmogorov-Smirnov test on the distribution of frequencies before and after manipulation in Figure 3 to test for statistical significance.

Quantification of duty cycle—The duty cycle was measured as the period of the fundamental frequency of the phonation, and the pulse duration was defined as the time passed from the initial peak to first trough of the waveform. Therefore, a purely sinusoidal waveform would have a duty cycle of 0.5. The duty cycle was calculated using 20 cycles from one high-quality recording of a representative call (note: some individuals only made a handful of calls) per individual before and after manipulation. Species had relatively consistent distress calls (DC) and induced phonation (IP) throughout, and so taking into account different call types was not necessary. Cycles were generally measured every 1–5 ms per call depending on call quality and duty cycle length. The calculated duty cycle in Figure 2D consisted of 2 DCs, 1 IP for *F. rufus*; 1 DC, 1 IP for *T. caerulescens*; 1 IP for *L. angustirostris*; 1 IP for *C. tyrannina*; and 1 DC for *S. rufosuperciliata*. For example, 3 individuals were used to calculate the duty cycle for *F. rufus*, two of which had distress calls recorded while the remaining had induced phonation recorded. Of these three, 20 cycles were measured per individual before manipulation, and 20 cycles were measured from the same individual after manipulation. Duty cycle for IP versus DC may vary depending on the species/individual (though we cannot categorically say this since we do not have IP and DC from the same individuals, as this is not possible to measure in the field), though this variation seems to be generally small. We used an unpaired t-test to test the statistical significance of the change in duty cycle after manipulation, the results of which can be found in Figure 4. The number of individuals used can be found in the figure legend, and represent the number of individuals used for each quantification (e.g., in *T. caerulescens*, phonations from 2 individuals were used to calculate duty cycle, and phonations from 1 individual were used to calculate the signal to noise ratio).

Signal to noise ratio quantification—Signal to noise ratios were calculated using distress calls, and extracted using the intensity feature in Praat. While gain on the Marantz was adjusted between individuals and individual calls to accommodate field conditions, it was held constant during an individual call. The noise (intensity listing in Praat, volts) for each individual distress call was measured just before the call. The signal was then extracted every 5 milliseconds over the duration of the distress call. The number of data points used to calculate the overall SNR therefore depended on the length of the call, and how many times

the individual produced the call. Since only distress calls were used to calculate SNR, the number of signal extractions (before manipulation, after manipulation) are as follows, with multiple numbers representing multiple individuals for a species: *F. rufus* (596, 591)(604, 585)(849, 830); *T. caerulescens* (2093, 791); and *S. rufosuperciliata* (2234, 218). We used an unpaired t-test to calculate the statistical significance of the change in SNR after manipulation, the results of which can be found in Figure 4D. As with duty cycle, the number of individuals used can be found in the figure legend, and represent the number of individuals used for each quantification (e.g., in *T. caerulescens*, phonations from 2 individuals were used to calculate duty cycle, and phonations from 1 individual were used to calculate the signal to noise ratio).

Data and Software Availability

Sound files (.wav) of phonations can be found as Mendeley Data (doi:10.17632/hrnv34rydw.1). Source code for the modeling of the syrinx can be found at Github (https://github.com/gabomindlin/tracheophones).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was in part funded by the National Science Foundation (GRFP; S.M.G.), the National Institute of Health (R01-DC-012859 and R01-DC-006876; G.B.M. and F.G.), the Gordon and Betty Moore Foundation (F.G.), the Agencia Nacional de Promoción Científica y Tecnológica (PICT 2014-2057, C.K.; 2014-2154, P.L.T.), and the National Scientific and Technical Research Council (PIP 112 201301 00803 CO; Fondo iBol Argentina D3657, C.K.). Santiago Faisal, Gabriela Roteta and staff from the Natural Resources of Corrientes Province Direction Office provided support in the permitting process of Corrientes, Argentina. Vicente Fraga, Daniel Segovia, Martín Sanchez, Conrado Holzer and staff of the Provincial Natural Parks and Reserves Area from Corrientes aided in field collection in Corrientes, Argentina. Martín Kowalewsky, EBCO staff, and Adrian Di Giacomo provided key support for Corrientes, Argentina fieldwork. Autoridad Nacional del Ambiente and the Smithsonian Tropical Research Institute provided permission for fieldwork in Gamboa, Panama. Keith Luscinski provided key data analysis support.

References

1. Galis F. The application of functional morphology to evolutionary studies. Tree. 1996; 11:124–129. [PubMed: 21237779]
2. Fuxjager MJ, Schlinger BA. Perspectives on the evolution of animal dancing: a case study of manakins. Curr. Opin. Behav. Sci. 2015; 6:7–12.
3. Dickinson MH, Farley CT, Full RJ, Koehl MAR, Kram R, Lehman S. How animals move: An integrative view. Science. 2000; 288:100–106. [PubMed: 10753108]
4. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 2014; 346:1320–1331. [PubMed: 25504713]
5. Whitney O, Pfenning AR, Howard JT, Blatti CA, Liu F, Ward JM, Wang R, Audet J, Kellis M, Mukherjee S, et al. Core and region-enriched networks of behaviorally regulated genes and the singing genome. Science. 2014; 346:1256780. [PubMed: 25504732]
6. Pfenning AR, Hara E, Whitney O, Rivas MV, Wang R, Roulhac PL, Howard JT, Wirthlin M, Lovell PV, Ganapathy G, et al. Convergent transcriptional specializations in the brains of humans and song-learning birds. Science. 2014; 346:1256846. [PubMed: 25504733]

7. Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014; 346:1311–1320. [PubMed: 25504712]
8. Verzijden MN, ten Cate C, Servedio MR, Kozak GM, Boughman JW, Svensson EI. The impact of learning on sexual selection and speciation. Trends Ecol. Evol. 2012; 27:511–519. [PubMed: 22705159]
9. Touchton JM, Seddon N, Tobias JA. Captive rearing experiments confirm song development without learning in a tracheophone suboscine bird. PLoS One. 2014; 9:e95746. [PubMed: 24788343]
10. Rüppell W. Physiologie und Akustik der Vogelstimme. J. Ornithol. 1933; 81:433–542.
11. Müller, J. Über die bisher unbekannten typischen Verschiedenheiten der Stimmorgane der Passerinen. Berlin: Abhandlungen der Königlichen Akademie der Wissenschaften; 1847. p. 321-391.
12. King, AS. Functional anatomy of the syrinx. In: King, AS., McLelland, J., editors. Form and Function in Birds. Cambridge: Academic Press; 1989. p. 105-192.
13. Raikow RJ, Bledsoe AH. Phylogeny and evolution of the passerine birds. BioScience. 2000; 50:487–499.
14. Ames PL. The morphology of the syrinx in passerine birds. Bull. Peabody Mus. Nat. Hist. 1971; 37:1–194.
15. Suthers, RA., Zollinger, SA. Mechanisms of song production in songbirds. In: Zeigler, HP., Marler, P., editors. Neuroscience of Birdsong. Cambridge: University Press; 2008. p. 78-98.
16. Goller F, Riede T. Integrative physiology of fundamental frequency control in birds. J. Physiol. Paris. 2013; 107:230–242. [PubMed: 23238240]
17. Beckers GJL. Bird speech perception and vocal production: a comparison with humans. Hum. Biol. 2011; 83:191–212. [PubMed: 21615286]
18. Goller F, Larsen ON. In situ biomechanics of the syrinx and sound generation in pigeons. J. Exp. Biol. 1997; 200:2165–2176. [PubMed: 9286098]
19. Arneodo EM, Perl YS, Mindlin GB. Acoustic signatures of sound source-tract coupling. Phys. Rev. E. 2011; 83:04192.
20. Jensen KK, Cooper BG, Larsen ON, Goller F. Songbirds use pulse tone register in two voices to generate low-frequency sound. Proc. R. Soc. B. 2007; 274:2703–2710.
21. Sitt JD, Amador A, Goller F, Mindlin GB. Dynamical origin of spectrally rich vocalizations in birdsong. Phys. Rev. E. 2008; 78:011905.
22. Strogatz, SH. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. Westview press; 2014.
23. Suthers RA. Contributions to birdsong from the left and right sides of the intact syrinx. Nature. 1990; 347:473–477.
24. Suthers RA, Goller F, Hartley RS. Motor dynamics of song production by mimic thrushes. J. Neurobiol. 1994; 25:917–936. [PubMed: 7964705]
25. Goller F, Suthers RA. Role of syringeal muscles in controlling the phonology of bird song. J. Neurophysiol. 1996a; 76:287–300. [PubMed: 8836225]
26. Goller F, Suthers RA. Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. J. Neurophysiol. 1996b; 75:867–876. [PubMed: 8714659]
27. Srivastava KH, Elemans CPH, Sober SJ. Multifunctional and context-dependent control of vocal acoustics by individual muscles. J. Neurosci. 2015; 35:14183–14194. [PubMed: 26490859]
28. Amador A, Goller F, Mindlin GB. Frequency modulation during song in a suboscine does not require vocal muscles. J. Neurophysiol. 2008; 99:2383–2389. [PubMed: 18287554]
29. Riede T, Goller F. Peripheral mechanisms for vocal production in birds – differences and similarities to human speech and singing. Brain & Language. 2010; 115:69–80. [PubMed: 20153887]
30. Greenewalt, CH. Bird Song: Acoustics and Physiology. Washington, D.C: Smithsonian Institute Press; 1968.
31. Elemans CPH, Mead AF, Rome LC, Goller F. Superfast vocal muscles control song production in songbirds. PLoS ONE. 2008; 3:e2581. [PubMed: 18612467]

32. Uchida AM, Meyers RA, Cooper BG, Goller F. Fibre architecture and song activation rates of syringeal muscles are not lateralized in the European starling. J. Exp. Biol. 2010; 213:1069–1078. [PubMed: 20228343]

33. Jarvis ED, Ribeiro S, Luisa da Silva M, Ventura D, Vielliard J, Mello CV. Behaviourally driven gene expression reveals song nuclei in hummingbird brain. Nature. 2000; 406:628–632. [PubMed: 10949303]

34. Gahr M. 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. 2000; 426:182–196. [PubMed: 10982462]

35. Alonso R, Goller F, Mindlin GB. Motor control of sound frequency in birdsong involves the interaction between air sac pressure and labial tension. Phys. Rev. E. 2014; 89:032706.

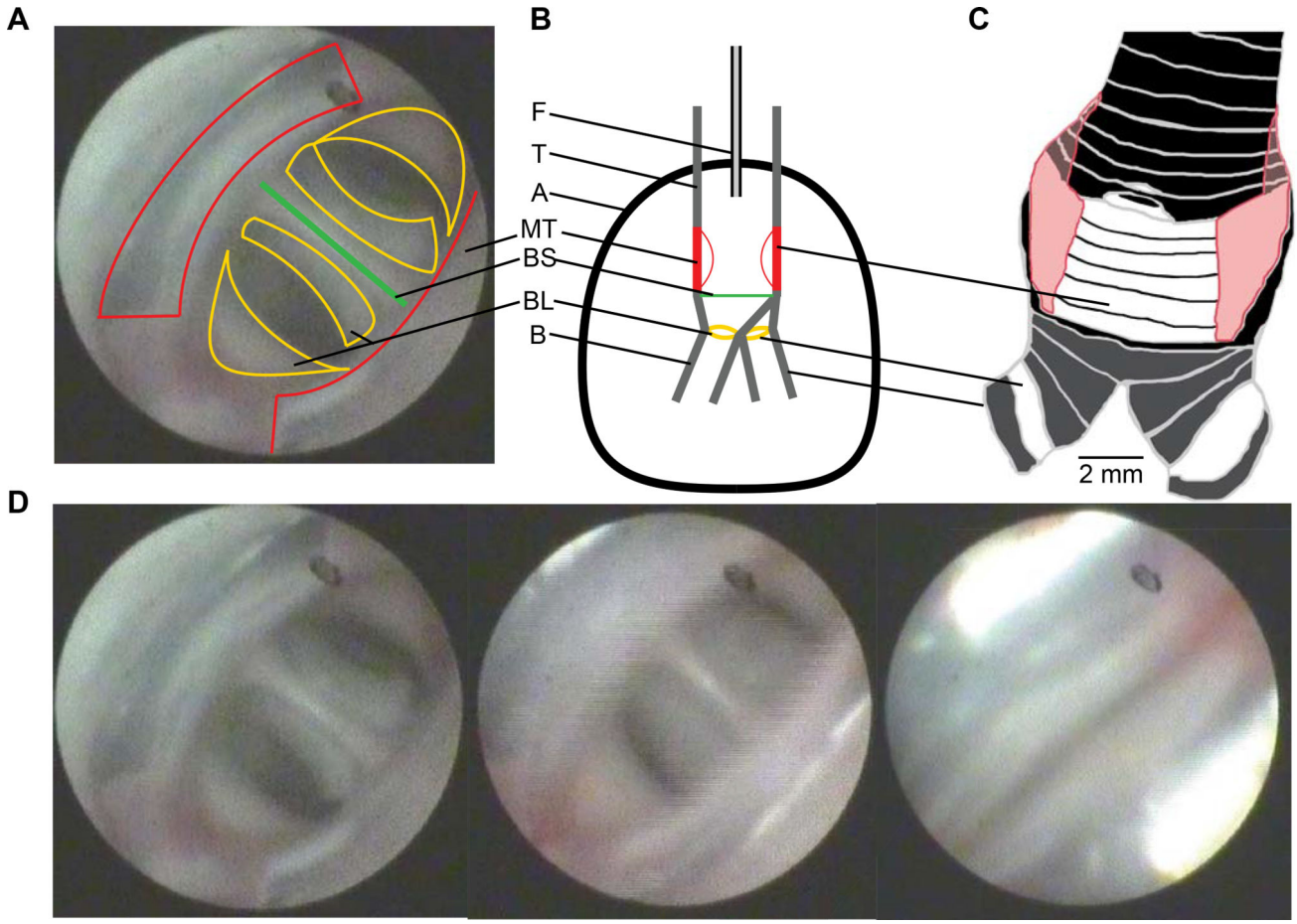

Figure 1. Fiberscopic visualization of the *F. rufus* syrinx reveals sound sources

To internally visualize the syrinx and identify sound sources, a fiber optic cable was surgically inserted into the trachea. Relevant structures are labeled and localized in an image of the view during quiet respiration (A, corresponding to D1), a schematic of a lateral view of the tracheobronchial junction (B) and a drawing of a ventral view of the *F. rufus* syrinx (C) (Ve = ventral; Do = dorsal; F = fiber optic cable; T = trachea; A = air sac; *MT* = *membrana trachealis*; BS = bronchial septum; BL = bronchial labia; B = bronchus). (D) Still shots of the syrinx during quiet respiration (1), when the bronchial sound sources are visible and vibrating (2), and when the MT are vibrating (3). See also Movies S1 and S2.

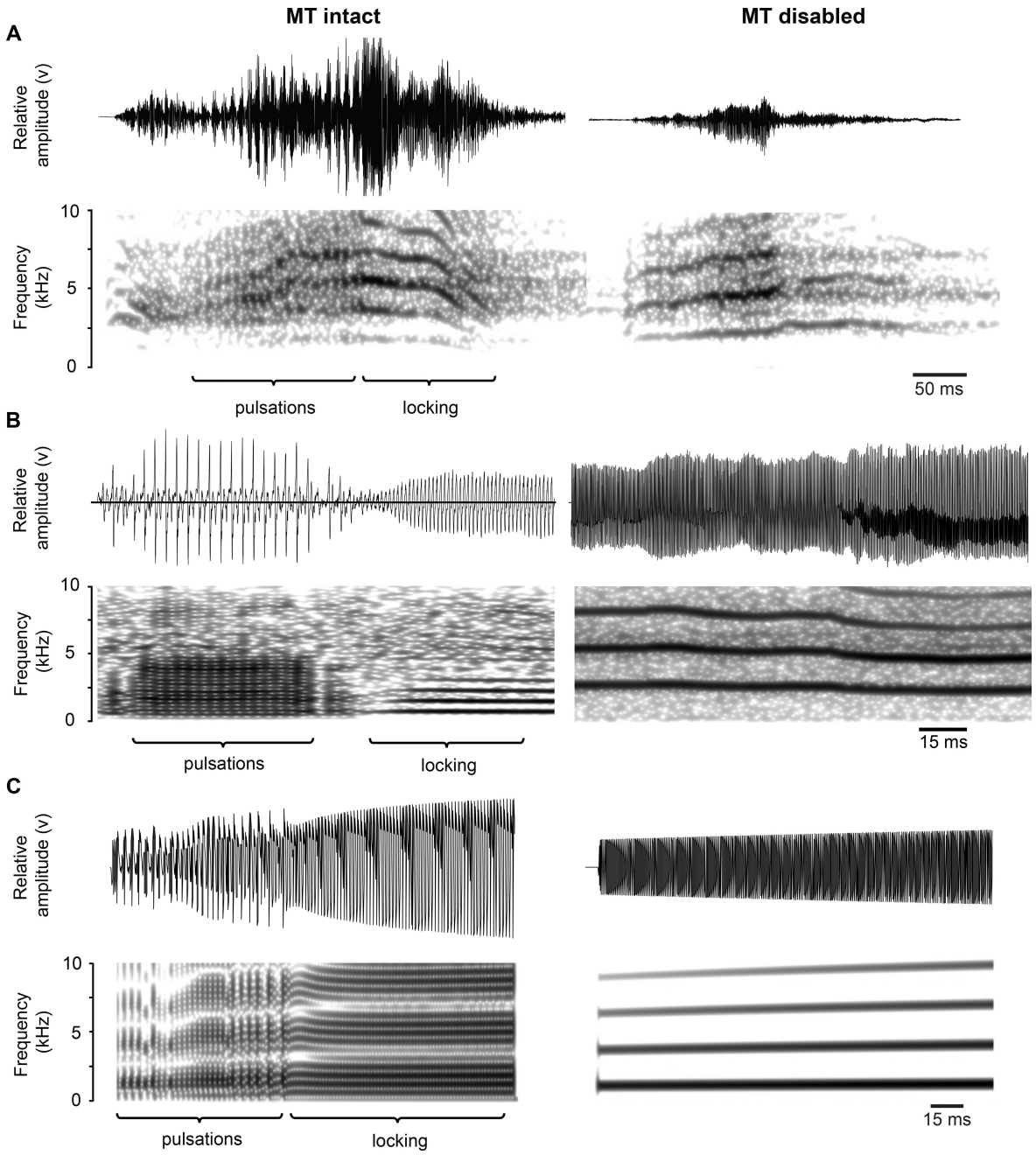

Figure 2. Manipulation of the ventral *membrana trachealis* reveals role in sound production and interactions with labial sound sources in *F. rufus*

Examples (oscillographic and spectrographic representation) of spontaneous distress calls (A), induced phonation (B), and mathematically-modeled sound (C) before (*MT* intact) and after (*MT* disabled) manipulation of the *MT* in *F. rufus*. Note the absence of the prominent, low-frequency amplitude modulations of spontaneous distress calls and the absence of the pulsations in induced sound after membrane manipulation, as well as an increase in fundamental frequency (see also Figures 3 and S3). Mathematical modeling of the *MT* and a labial sound source interacting to produce sound C) makes two clear predictions: (1) When the *MT* oscillate at a frequency disparate from that of the bronchial sound sources,

amplitude modulation will occur; (2) When the natural frequency of the *MT* is similar to that of the fundamental frequency of the bronchial sound sources, they lock and oscillate at an intermediate frequency.

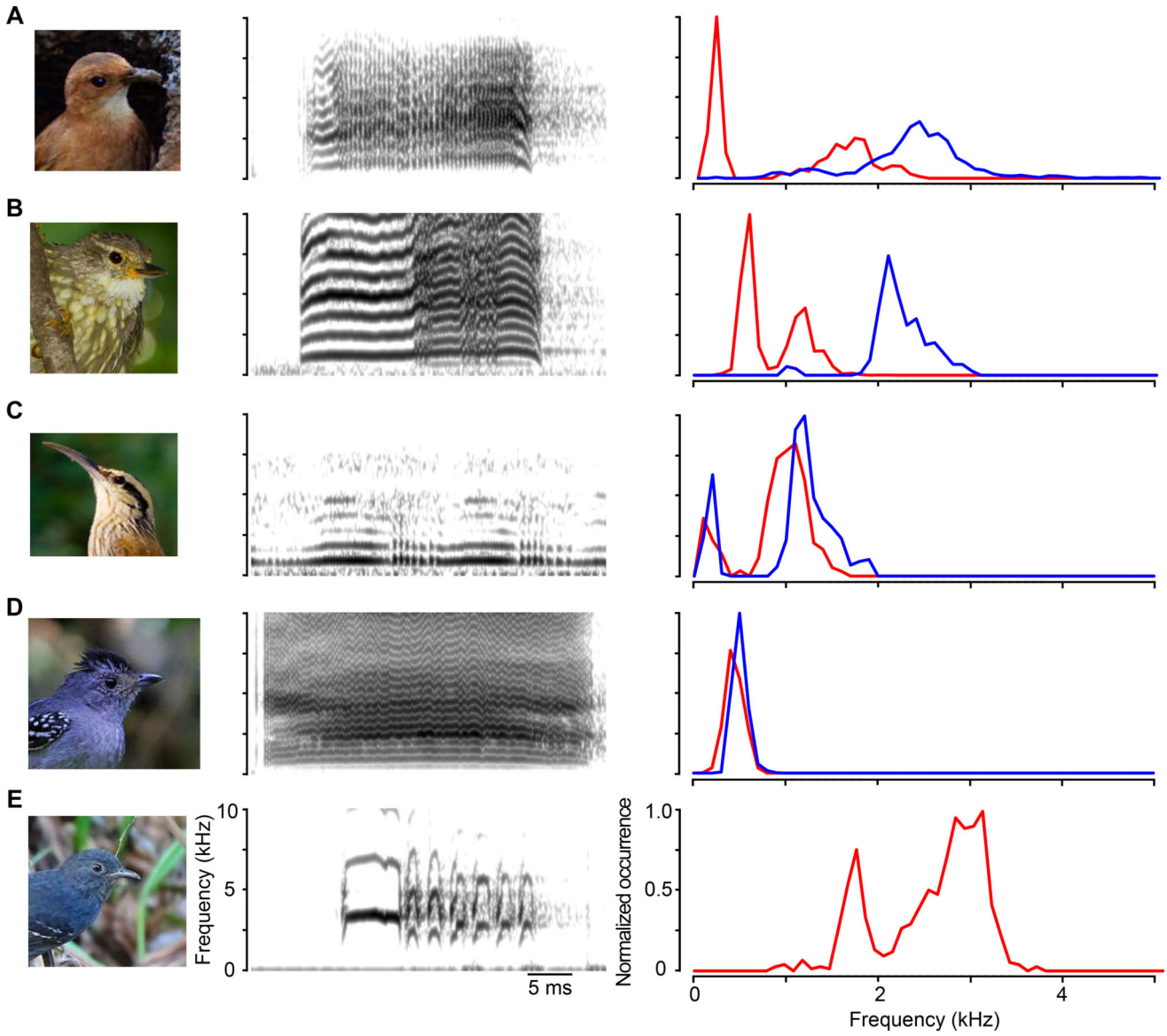

Figure 3. Sound frequency shifts after manipulation of *membranae tracheales* in tracheophones

Examples of phonations (spectrograms) and frequency occurrence normalized to most common frequency before (red) and after (blue) manipulation are displayed for (A) *F. rufus*, spontaneous distress call (n=3), (B) *S. rufosuperciliata*, spontaneous distress call (n=1), (C) *L. angustirostris*, induced phonation (n=1), (D) *T. caerulescens*, spontaneous distress call (n=1), (E) *C. tyrannina*, spontaneous distress call (n=1) (see also Figure S1). All distributions after the manipulation are significantly different from those before (Kolmogorov-Smirnov, $p<0.001$). Manipulation caused immobilization of the v*MT*, and was accomplished by either rupture or glue application. Sounds after manipulation are characterized by a shift to higher frequencies, either through absence of very low frequencies indicating loss of amplitude modulation as seen in *F. rufus* (see also Figure S3), or through smaller shifts indicating loss of "locked" vibration as seen in *L. angustirostris* and *T. caerulescens* (see also Figure S2). *C. tyrannina* was unable to produce sound after

manipulation. The dotted portion of the x-axis indicates the frequency range of the vocal repertoire of each species measured from recordings found at Xeno-Canto.org. Images downloaded from various sources. See also Movie S1.

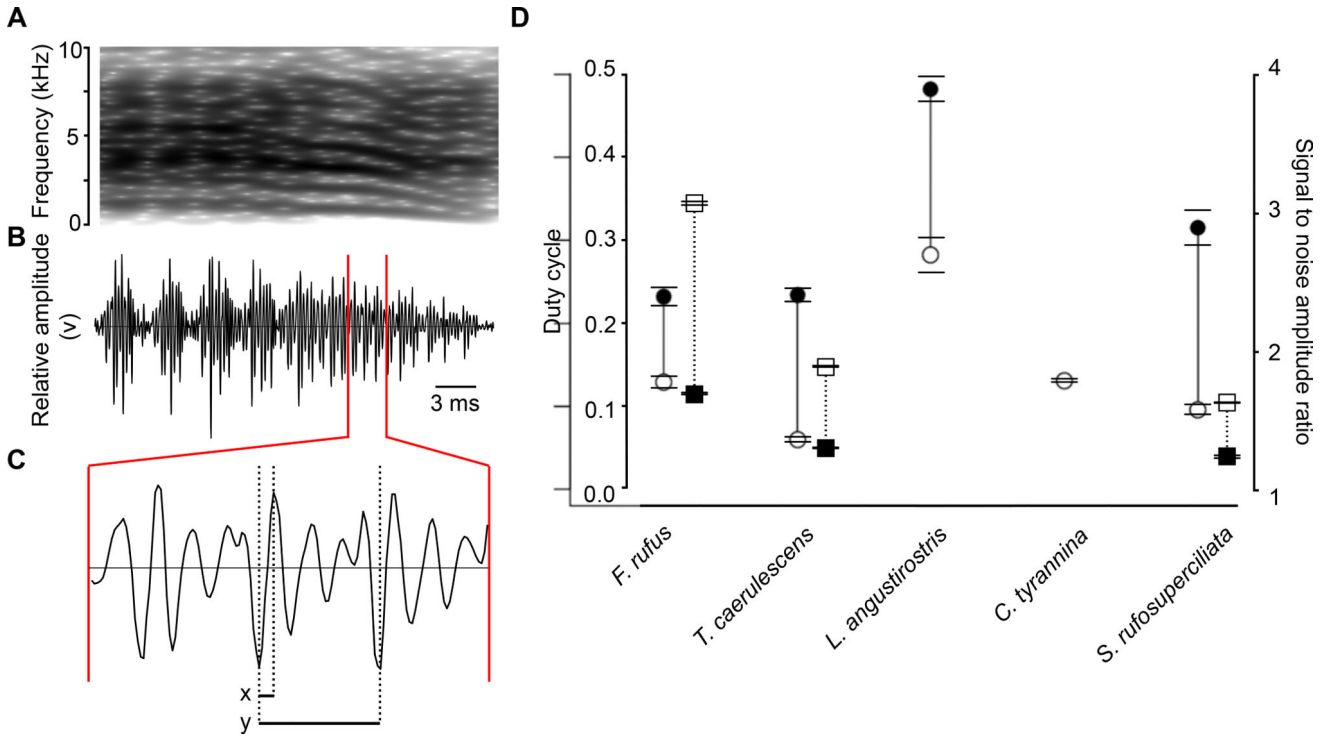

Figure 4. Manipulation of the ventral *MT* increases duty cycle and decreases sound amplitude in tracheophone species

Using an example of *F. rufus* spontaneous phonation with corresponding spectrogram (A) and oscillogram (B), we quantified duty cycle by measuring × (beginning of the period to apex of the first initial peak) and y (period) (C), and dividing x/y. We present (D) duty cycle (left y-axis, circles) and signal to noise ratio (SNR; right y-axis, squares) of phonations before (open circles and squares) and after (closed circles and squares) v*MT* manipulation was used to prevent oscillation in (n) individuals listed respectively for each parameter (n=duty cycle individuals, sound amplitude individuals) in *F. rufus* (n=3, 3), *T. caerulescens* (n=2, 1), *L. angustirostris* (n=1), *C. tyrannina* (n=1), and *S. rufosuperciliata* (n=1,1). SNR is included for species with recorded spontaneous distress calls. Changes in both duty cycle and SNR are highly statistically significant (unpaired t-test, $p<0.0001$) for 4 of 5 species; *C. tyrannina* was unable to phonate after membrane manipulation. Horizontal bars indicate standard errors for duty cycle and SNR before and after manipulation (D).