VOCAL INTERACTION

Motor cortical control of vocal interaction in neotropical singing mice

Daniel E. Okobi Jr.[1,2,3]*, Arkarup Banerjee[1,2,3]*, Andrew M. M. Matheson[1,3], Steven M. Phelps[4], Michael A. Long[1,2,3]†

Like many adaptive behaviors, acoustic communication often requires rapid modification of motor output in response to sensory cues. However, little is known about the sensorimotor transformations that underlie such complex natural behaviors. In this study, we examine vocal exchanges in Alston's singing mouse (*Scotinomys teguina*). We find that males modify singing behavior during social interactions on a subsecond time course that resembles both traditional sensorimotor tasks and conversational speech. We identify an orofacial motor cortical region and, via a series of perturbation experiments, demonstrate a hierarchical control of vocal production, with the motor cortex influencing the pacing of singing behavior on a moment-by-moment basis, enabling precise vocal interactions. These results suggest a systems-level framework for understanding the sensorimotor transformations that underlie natural social interactions.

Adaptive behavior often requires adjusting action in response to a rapidly changing environment. Elucidating the mechanisms of these sensorimotor transformations has become a central focus of systems neuroscience, as researchers use simple and elegant behavioral tasks to explore the behavioral responses of traditional model species to sensory cues (*1*–*3*). Ultimately, however, we would like to understand such transformations in natural contexts; among such contexts, perhaps none is more challenging or interesting than social behavior. During social interactions, an animal must dynamically modulate complex actions in response to the changing behavior of a conspecific. For example, during conversation, we listen to the words of another person, interpret them, and respond appropriately (*4*). Indeed, acoustic exchanges are promising foci for the study of sensorimotor transformations that underpin social behavior. These exchanges are common across taxa, including insects (*5*, *6*), amphibians (*7*, *8*), birds (*9*–*11*), and mammals (*12*–*16*); they serve a variety of essential social functions, including male-male competition and mate selection; and they require dynamic interaction as signalers avoid temporal overlap with one another (*17*).

Despite the ubiquity of acoustic interactions in the natural world, there are few existing models within neuroscience. Among mammals, for example, laboratory mice produce elaborate frequency-modulated vocalizations (*18*) but fail to exhibit robust turn-taking behavior (*19*). In contrast, marmoset pairs call antiphonally (*14*, *20*), but the time scale of these interactions is relatively slow (3 to 5s) (*4*, *20*, *21*). In Alston's singing mouse (*Scotinomys teguina*), we find a robust and rapid countersinging (~500 ms) that resembles the subsecond latencies of both conditioned sensorimotor transformations in laboratory settings (*22*) and the timing of vocal turn-taking evident in human conversation (*4*). We employ a range of techniques for manipulating neural dynamics to pinpoint a motor cortical locus that works hierarchically within the song production pathway to enable precise vocal interactions between conspecific pairs.

S. teguina is a small (~12 to 15 g), highly vocal neotropical rodent native to the cloud forests of Central America (*23*–*26*) and is related to the genus *Peromyscus* and other New World rodents. Their family (Cricetidae) includes voles and hamsters and is in the same superfamily (Muroidae) as house mice and the Norwegian rat (*27*). Both male and female *S. teguina* produce vocal sequences consisting of a series of discrete frequency-modulated elements strung together, with characteristics that change predictably as the vocalization progresses (*24*) (Fig. 1, A to D, and movie S1). Following the convention of previous studies (*23*, *25*), we refer to each vocal episode as a "song" and individual components as "notes." We visualize this trend by plotting the duration of each note as a function of its onset time within the song: The song trajectory plot (Fig. 1D) provides a succinct representation of this motor sequence. We found that trajectory plots were highly stereotyped across renditions from individuals recorded in acoustic isolation (Fig. 1E). This degree of motor precision is reminiscent of vocalizations produced by a range of evolutionarily distant species (*28*, *29*) but stands in stark contrast to the variable acoustic structure of ultrasonic vocal sequences produced by laboratory mice (*18*, *30*, *31*).

We next examined whether the acoustic characteristics of *S. teguina* vocalizations are modulated by social context, as observed in other taxa (*32*). To investigate this, we staged a social encounter by relocating a male subject (a "recruit") into a testing room occupied for at least 1 week by another male (a "resident"). The two mice were held in adjacent chambers with acoustic but not visual access to each other. In this configuration, recruit males altered their singing in two ways. First, recruits vocalized four times as often in the social context [social (day 2): 20.4 ± 4.8 songs/hour; mean ± SEM unless stated otherwise] as in isolation [alone (day 1): 4.7 ± 0.8 songs/hour, alone (day 3): 4.4 ± 0.7 songs/hour] (Fig. 1, E and F). Second, the variability of song trajectory plots increased significantly when recruits could hear the resident mouse (Fig. 1E). Consistent with this observation, we found that song duration variability was higher in the social context [social (day 2): 2.7 ± 0.3 s] than in isolation [alone (day 1): 1.5 ± 0.1 s, alone (day 3): 1.4 ± 0.2 s] (Fig. 1G).

To examine the fine structure of vocal interactions between male *S. teguina*, we simultaneously recorded the songs of both the resident and recruit mice in the social condition. We found extensive temporal coordination of singing behavior within vocal pairs (Fig. 2 and movie S2). Whereas exchanges could be initiated by either male, they typically ended with a recruit's song (Fig. 2, A to E). Surprisingly, this asymmetry was observed across all recruit-resident pairs (n = 8) and was preserved for the entire ~24-hour social session (89 ± 10 interaction bouts per pair) (Fig. 2, B and E). For the remainder of this study, we restricted our analysis to the songs of the recruit mice to focus on this sensory-evoked vocal response. By aligning the interaction bouts to the songs of the resident mouse, we found that the recruit mouse precisely times his vocal onset to coincide with the end of the resident's songs (Fig. 2B, left). This observation was robust across all pairs (average response latency = 0.81 ± 0.18 s) (Fig. 2C; n = 8). To estimate the amount of countersinging that one would expect by chance given the amount of singing observed in the social condition, we shuffled the song times and quantified the likelihood of such "spurious countersinging" to be nearly an order of magnitude less (Response probability$_{Data}$ = 0.69 ± 0.09, Response probability$_{Shuffled}$ = 0.07 ± 0.02, P < 0.01, Wilcoxon signed rank test). The recruit's response probability distributions were significantly sharper when interaction bouts were aligned to the end of the resident's songs rather than the start (jitter$_{end\text{-}aligned}$ = 2.94 ± 0.64 s, jitter$_{start\text{-}aligned}$ = 5.19 ± 0.43 s, P < 0.05, Wilcoxon signed rank test; fig. S1), suggesting that the recruit mouse uses the end of the resident's song as a sensory trigger. Additionally, the recruit mouse often

[1]NYU Neuroscience Institute, New York University Langone Health, New York, NY 10016, USA. [2]Department of Otolaryngology, New York University Langone Health, New York, NY 10016, USA. [3]Center for Neural Science, New York University, New York, NY 10003, USA. [4]Department of Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: mlong@med.nyu.edu

stopped vocalizing immediately after the resident mouse started singing (Fig. 2B, right, and 2D). Thus, recruit males were capable of actively timing their vocalization onsets and offsets to avoid acoustic overlap with the resident (Fig. 2E), giving rise to turn-taking dynamics similar to those observed during human conversation (*4*). Furthermore, the recruit's response precision correlated with the degree of social engagement, as quantified by countersinging probability (Fig. 2F) and the increase in song duration variability across social contexts (Fig. 2G), suggesting that active participation in an orderly vocal exchange contributed to these changes (Fig. 1). This result is consistent with recent findings demonstrating that context can influence the timing of vocal turn-taking in other species (*33*, *34*).

We next sought to explore the neural mechanisms contributing to countersinging. As a first step, we characterized the biomechanics of song production by examining the motor elements that make up a song. Singing resulted in a rapid cycle of inhalation and exhalation (fig. S2), a stark contrast from laboratory mice whose vocalizations are strongly coupled to ongoing sniffing activity (*35*). In singing mice, phonation is coupled to exhalation and jaw movements; electromyography (EMG) confirmed that individual vocalizations were produced during the exhalation phase and were preceded by robust flexion of the jaw muscle (digastricus) (Fig. 3A). The correlation between song production and jaw movement—similar to that previously observed in rats (*36*)—allowed us to use EMG activity to probe the relationship between specific brain centers and song-related musculature. In previous studies, stimulation of motor cortical centers in primates resulted in vocal fold adduction (*37*), suggesting a possible involvement of the motor cortex in vocalization. We used intracortical microstimulation (ICMS) over a large portion of the anterior cortex to identify areas leading to flexion of song-related musculature. The minimum current that reliably elicited a fixed EMG activity threshold (Fig. 3, B and C, and fig. S3) was used to define a functional hotspot that maps to the anterolateral aspect of the motor cortex (Fig. 3C, right), which corresponds to the orofacial motor cortex in *Mus musculus* (*38*). We therefore refer to this region as the orofacial motor cortex (OMC).

What is the functional role, if any, of the OMC on song production? Although OMC stimulation can elicit electrical activity in song-relevant musculature, this does not necessarily imply that the OMC can influence song production. To address this directly, we carried out a series of perturbations during singing in the alone condition, beginning with bilateral electrical stimulation of the OMC. Strong stimulation resulted in song truncation, whereas milder stimulation (200 to 500 μA) often produced brief pauses (range: 638 to 1448 ms), with songs resuming once stimulation ended (Fig. 3D). The precise stereotypy of alone *S. teguina* songs (Fig. 1E) provides an ideal opportunity to distinguish between two possible experimental outcomes. First, the song could resume at the expected point in the sequence, accounting for the time delay (outcome 1; Fig. 3E), consistent with the hypothesis that the vocal patterning is primarily driven by a pathway independent of the OMC. In nonhuman primates, for example, there is a vocal motor stream that begins in the cingulate cortex and acts via the periaqueductal gray (*39*). An alternative outcome of our experiment is that the song could resume at the same point in the motor sequence where it had paused (outcome 2; Fig. 3E), suggesting that the pathway leading from the OMC to vocal musculature is capable of sculpting the structure of song. For every trial, we used the 10 notes preceding the perturbation to estimate the note durations that would be expected in an uninterrupted song. We then compared the actual note duration with these predicted values and found that song typically resumes at the same point in the sequence where it had paused (Fig. 3F, outcome 2). Across the population, note durations after song resumption were significantly more similar to outcome 2 than outcome 1 in 58 out of 61 trials across four animals (Fig. 3G). These results refute the hypothesis that an OMC-independent pathway shapes song patterning in *S. teguina*.

Although our stimulation results functionally connect the OMC to the behavioral output, they do not elucidate the nature of this interaction. Previous reports suggest that most mammalian vocal communication does not involve the motor cortex and that subcortical structures are sufficient for this behavior (*39*, *40*). To isolate the contributions of local neuronal dynamics in the OMC from those of downstream structures, we used mild focal cooling of the OMC during song production. Manipulating neural circuits

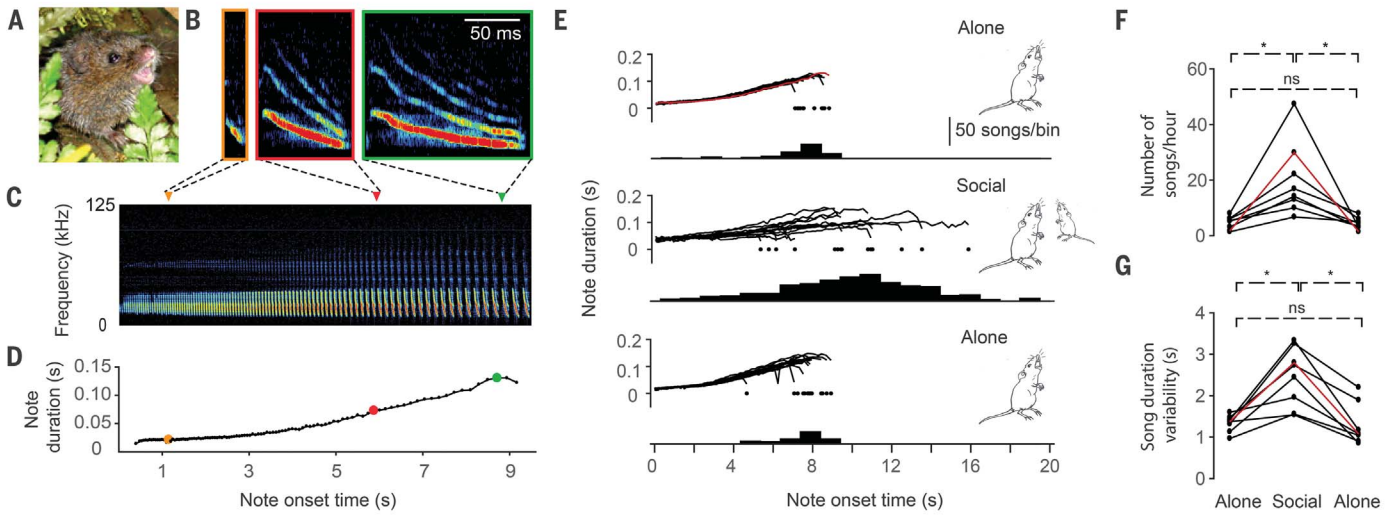

Fig. 1. Social context modulates vocalizations in *S. teguina*. (**A**) An adult *S. teguina* in its natural habitat. (**B**) Spectrograms of three example notes from one individual. Frequency range: 0 to 125 kHz. (**C**) Spectrogram of a full *S. teguina* advertisement song. The colored arrows denote the onset time of the three corresponding notes from (B). (**D**) Trajectory plot in which individual note durations are displayed as a function of their onset times in the song, with colored circles indicating notes from (B). (**E**) Trajectories from one male *S. teguina* in different social contexts (n = 15 songs per condition). The vocal stereotypy exhibited during isolated singing (top and bottom) is significantly altered during social interaction (middle). Individual dots represent the duration of each displayed song, and the histogram quantifies the durations for all songs produced in a given context (day 1 alone: n = 57 songs; day 2 social: n = 388 songs; day 3 alone: n = 50 songs). The red line is the same trajectory plotted in (D). (**F** and **G**) The number of songs per hour (F) and the song duration variability (G), defined as the standard deviation of the song duration distributions, significantly increase during the social context (n = 8 animals). Red lines represent the example mouse from (E). Asterisks indicate a significant difference between conditions (*P < 0.01, Wilcoxon signed rank test; n.s., not significant).

with temperature has emerged as a useful experimental tool for maintaining behaviorally relevant activity while selectively slowing these dynamics (*41*–*45*). We predict three possible outcomes of this manipulation. If song timing is exclusively governed by subcortical structures, as expected in standard rodent models (*35*), then the control and cooled song trajectories should completely overlap (Fig. 3H, model 1). Alternately, if OMC dynamics exclusively dictate the temporal structure of song, then cooling should lead to the dilation of vocal behavior on all time scales (i.e., note duration and song length) (Fig. 3H, model 2), as evident in both birdsong (*41*) and human speech (*42*). If motor control of the song is shared between the OMC and subcortical regions, then cooling may alter some temporal properties while preserving others (Fig. 3H, model 3). One possibility is that cooling may change the slope of the song trajectory, a parameter we observe to be socially modulated (Fig. 1). To test these models, we used a custom-built Peltier device capable of rapidly and reversibly cooling the OMC (fig. S4). Cooling strongly affected song timing by monotonically increasing the overall song duration (Fig. 3I). In contrast, cooling did not affect running speed (fig. S5), a behavior unlikely to require substantial cortical involvement (*46*, *47*). OMC cooling resulted in a shallower song trajectory that took longer to unfold (Fig. 3, I to K; n = 10 animals). We found that cooling decreases the slope of the song trajectory ($Slope_{control}$ = 0.013 ± 0.001; $Slope_{cooling}$ = 0.009 ± 0.001, P < 0.002, Wilcoxon signed rank test) as well as the time for the song trajectory to surpass an arbitrary threshold ($Threshold_{control}$ = 4.23 ± 0.16 s; $Threshold_{cooling}$ = 4.88 ± 0.2 s, P < 0.002, Wilcoxon signed rank test). These changes demonstrate that the OMC contributes significantly to song patterning, thereby ruling out model 1 (Fig. 3H). In addition, a closer examination of song acoustic structure revealed that the distribution of individual note durations did not change with cooling ($Length_{control}$ = 68.1 ± 1.5 ms; $Length_{cooling}$ = 67.8 ± 1.5 ms, P = 0.92, Wilcoxon signed rank test), which is inconsistent with the model that the OMC solely determines all aspects of song timing (Fig. 3H, model 2). Instead, we find that the OMC shapes song progression without influencing the structure of individual notes. Neither the starting nor the ending note durations change as the result of cooling, but it takes longer for this progression to occur, which is accomplished by increasing the total number of notes produced ($Note\ number_{control}$ = 44.9; $Note\ number_{cooling}$ = 48.9, P < 0.01, Wilcoxon signed rank test). Therefore, these data suggest a hierarchy of motor timing control (Fig. 3H, model 3), with the OMC being capable of exerting moment-by-moment control over the pacing of a subcortically generated song sequence.

In our initial experiments, we observed that social interaction profoundly changed song progression (Fig. 1E) and that this song variability was driven by the degree of social engagement during vocal interactions (Fig. 2, F and G). Our

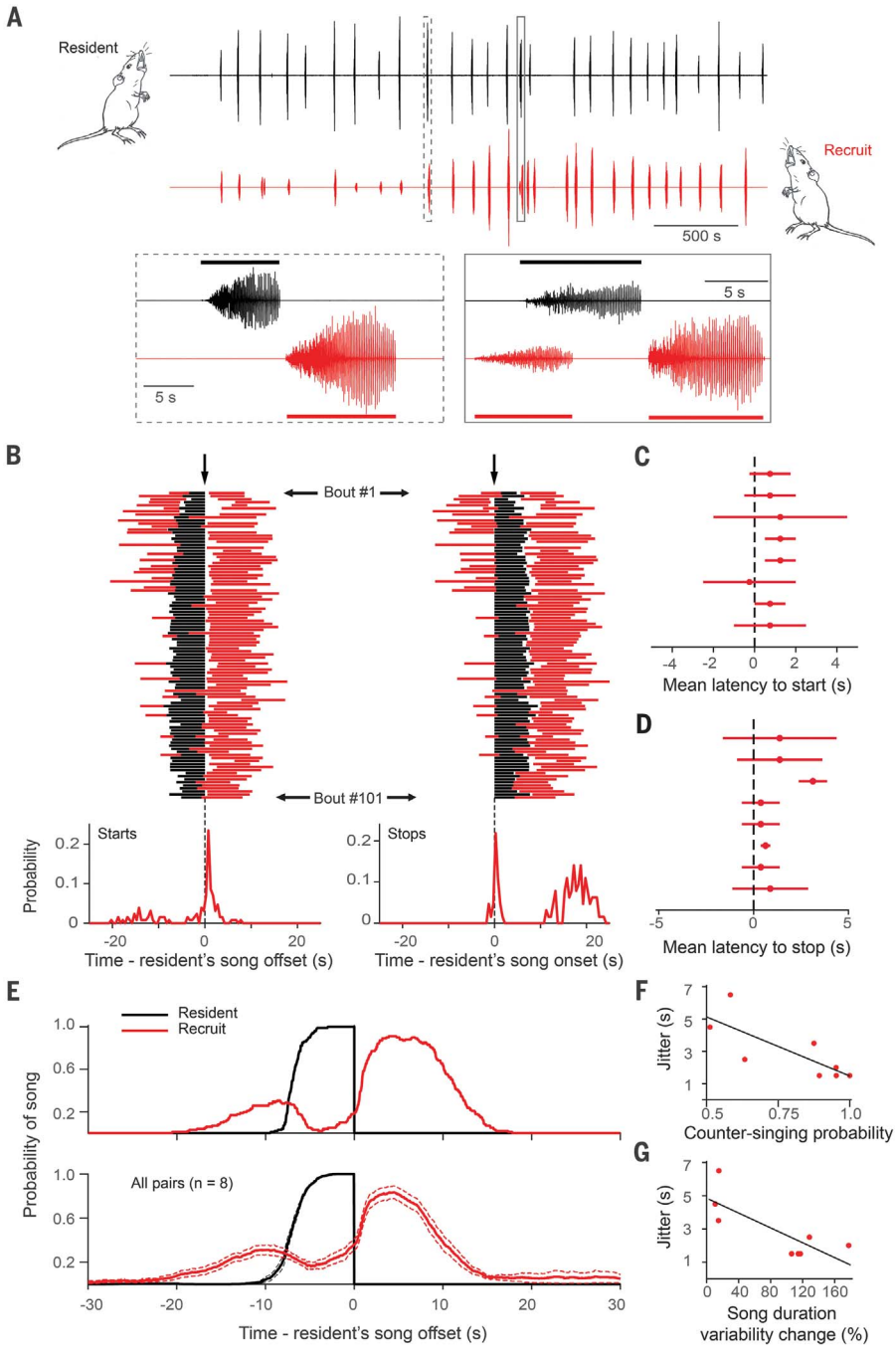

Fig. 2. Temporal coordination of vocal interactions between conspecific males. (**A**) One hour of continuous audio recordings from two interacting males. Two typical interactions are shown in detail: one initiated by resident mouse (black) and another by a recruit (red). (**B**) All vocal interactions (n = 101 interactions for this example pair) over a 24-hour period aligned to either the end (left) or the beginning (right) of the resident's songs. The corresponding start and stop probability distributions for the recruit's song are plotted below. (**C**) Summary of mean start latencies across all pairs (n = 8). For each, the circle represents the mean latency of the recruit mouse's song with respect to the offset of the resident's song, with horizontal line indicating song initiation jitter (full-width at half maximum of the probability distribution). (**D**) Mean stop latencies across all pairs with respect to onset of the resident's song. (**E**) Probability of song occurrence at any given time point aligned to the end of the resident mouse's song for the pair featured in (A) (top) as well as for all pairs (bottom), showing active avoidance of song overlap between conspecifics. In the bottom plot, dashed lines represent the SEM. (**F** and **G**) Song initiation jitter is negatively correlated with countersinging probability (F) (Pearson's correlation, r = −0.78, P < 0.05) as well as the degree of song duration variability change from the alone condition (G) (Pearson's correlation, r = −0.79, P < 0.05). Each dot represents the behavior of a single recruit mouse.

stimulation and cooling studies suggest that the OMC is well positioned to contribute to such social coordination by altering song structure. Therefore, we proceeded to test this prediction by reversibly inhibiting the OMC with muscimol (a $GABA_A$ agonist). Preliminary injections of a high muscimol dose (100 mM, 100 nl) in the motor cortex led to gross movement abnormalities as well as the complete abolishment of singing behavior for an extended period of time (>4 hours). Such nonspecific motor deficits were not evident when we lowered this concentration to 10 mM (100 nl), a common dosage (*46*, *48*, *49*). Each animal was tested for both experimental conditions: muscimol (OMC inactivation) and saline

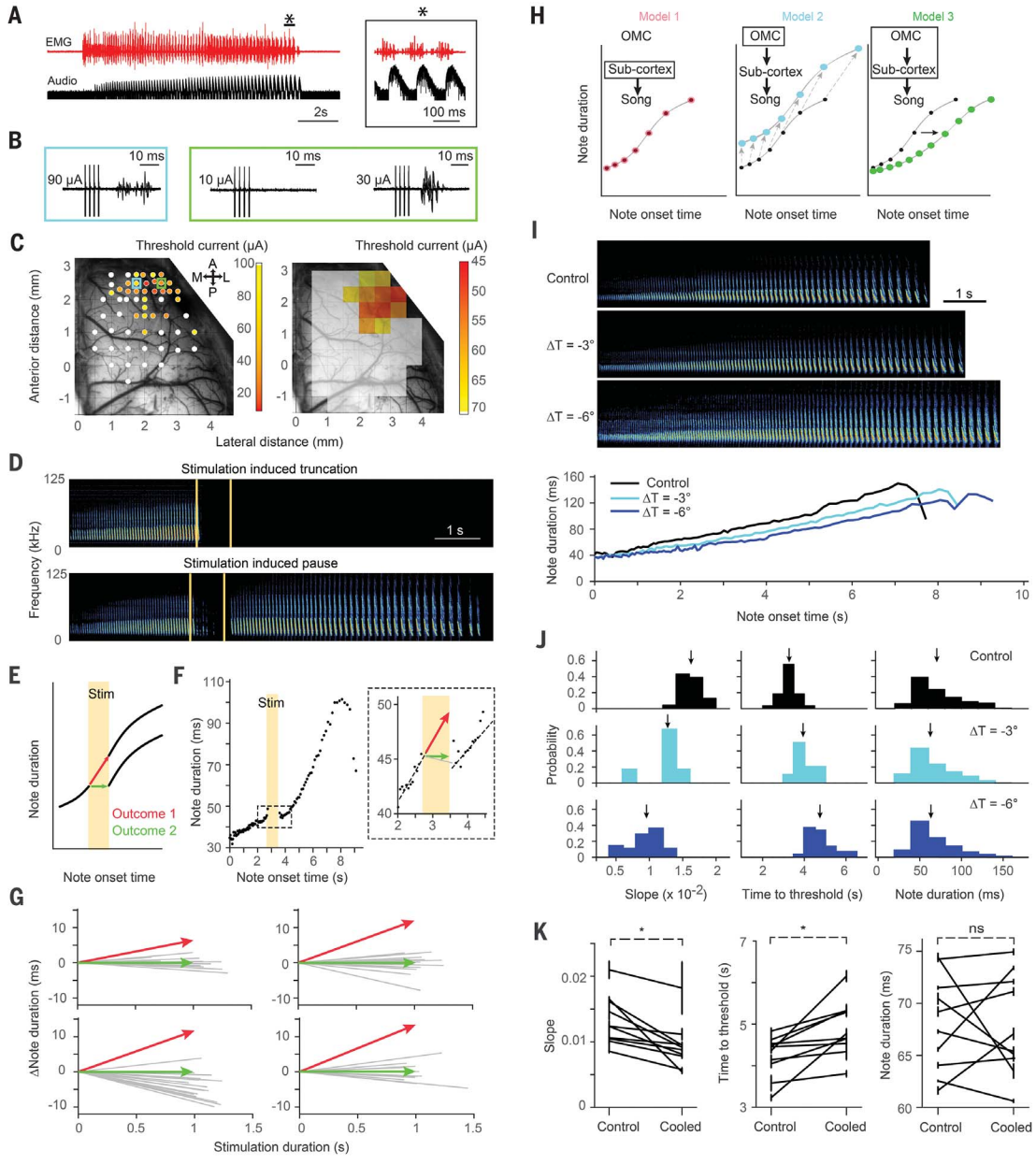

Fig. 3. Hierarchical control of song timing. (**A**) Electromyograph from the digastricus muscle and simultaneous raw audio (log amplitude) of one advertisement song, showing increased muscle activity immediately before vocalization of individual notes. The inset shows three notes (marked by an asterisk) and accompanying EMG activity in greater detail. (**B**) ICMS of two different loci elicits short-latency EMG activity. The simulation artifact (four parallel lines) is truncated for clarity. (**C**) The minimum amount of current needed to elicit a significant (statistical significance, as defined in the methods) EMG response (threshold current) from each ICMS site is color coded for one example mouse (left) and across the population (right, $n = 5$ mice), revealing a "hotspot" on the anterolateral portion of the motor cortex, henceforth referred to as the orofacial motor cortex (OMC, right). The ICMS locations for examples in (B) are indicated by cyan and green squares. A, anterior; P, posterior; M, medial; L, lateral. (**D**) Example spectrograms from one individual in which song was truncated (top) or paused (bottom) in different trials by a 200-μA electrical stimulation of the OMC. Yellow lines indicate the onset and offset of electrical stimulation. (**E**) Two possible outcomes for song resumption after a brief electrical stimulation–induced pause. (**F**) Full trajectory of an example song before and after electrical stimulation–induced pause, with dots indicating the duration of each note. The inset is an expanded view of the peristimulation period. Dashed black lines are the estimated note duration slopes. Expected trajectories under outcomes 1 and 2 are depicted as red and green lines, respectively. The gray line indicates the actual change in note duration after song resumption. (**G**) Summary data for all paused songs in each animal. The majority of trajectories are consistent with outcome 2 ($n = 12$ of 14, 14 of 15, 18 of 18, and 14 of 14 trials). (**H**) Proposed effects of OMC cooling on song trajectory if OMC activity does not affect song timing (left), if OMC exclusively controls song timing (middle), or if OMC and subcortical structures share this control (right). (**I**) Spectrograms and trajectories of example songs during baseline and cooling sessions. Cooling of the OMC lengthens song durations by decreasing the rate of change of note duration (slope) during song. (**J** and **K**) Summary for all songs during the control ($n = 27$ songs) and cooling periods (−3°C: $n =$ 10 songs, −6°C: $n = 32$ songs) for mouse C4 (J), as well as the mean ± SEM values of the entire population (K) ($n = 10$ animals). Arrows in (J) denote the mean values of each distribution. Cooling resulted in a decrease in the slope of the song trajectory (left) and an increase in the time needed to reach a threshold note length of 75 ms (middle) without changing the duration of individual notes (right). Asterisks indicate a significant difference between conditions (*$P < 0.01$, Wilcoxon signed rank test).

(control; n = 6 mice) (Fig. 4, A and B). In both conditions, we found that five of six individuals produced spontaneous songs and that the rate of spontaneous singing was not significantly influenced by this manipulation [Fig. 4C (bar graphs); control, 4.1 ± 1.51 songs/hour; muscimol, 2.1 ± 1.2 songs/hour, P = 0.31, Wilcoxon signed rank test].

We next used playback to evaluate whether the OMC mediated social influences on singing behavior. In control (saline-injected) animals, song playback led to an increase in the amount of singing as well as song duration variability, as expected in a social countersinging context (Fig. 4, C and G; $P < 0.05$, Kruskal-Wallis test). In contrast, muscimol-injected animals did not sing more songs in response to playback (Fig. 4, C and G; P = 0.81, Kruskal-Wallis test), suggesting that the OMC affects context-dependent modulation of song rate, a phenomenon we have observed during natural social encounters (Fig. 1F). Similarly, the probability of eliciting a countersinging response was significantly greater in the control condition than in the OMC-inactivated condition for each mouse (Fig. 4, C to E; $P < 0.05$, binomial test) as well as across the population (Fig. 4F; n = 5, saline: 0.59 ± 0.13; muscimol: 0.09 ± 0.05; $P < 0.05$, one-sided Wilcoxon signed rank test). Using a permutation test, we found that this difference in response probability could not be explained by our observed changes in song rate across conditions (fig. S6). Moreover, in cases where residual singing behavior remained after muscimol injection to the OMC (Fig. 4C), we observed an increase in mean response latency of 2.2 ± 0.9 s relative to that of saline-injected controls (Fig. 4, D and E; $P < 0.05$, Wilcoxon rank sum test). These data demonstrate that the OMC is critical for rapid vocal responses to playback; such responses must be driven by sensorimotor coupling rather than by more general changes in motivation.

In this study, we examined vocal interactions between pairs of *S. teguina* to test a range of hypotheses concerning the neural mechanisms underlying complex sensorimotor interactions. Using four complementary lines of evidence (intracortical microstimulation, stimulation-induced vocal arrest, focal cooling, and pharmacological inactivation), we define a region of the motor cortex (the OMC) that influences vocalization and mediates rapid vocal interactions. Whereas previous studies have used immediate early genes or electrophysiological approaches to suggest cortical involvement in nonhuman primate communication (*50*–*54*), our study represents the first direct demonstration of cortical dependence of precise vocal interactions in a mammal. Specifically, we have shown that the motor cortex is required for adaptive countersinging but not for song production itself. Additionally, our cooling results demonstrate that the motor cortex is capable of dynamically adjusting the pacing and duration of song sequences, consistent with the changes in these same parameters during social interactions. This finding provides evidence for recent proposals that the

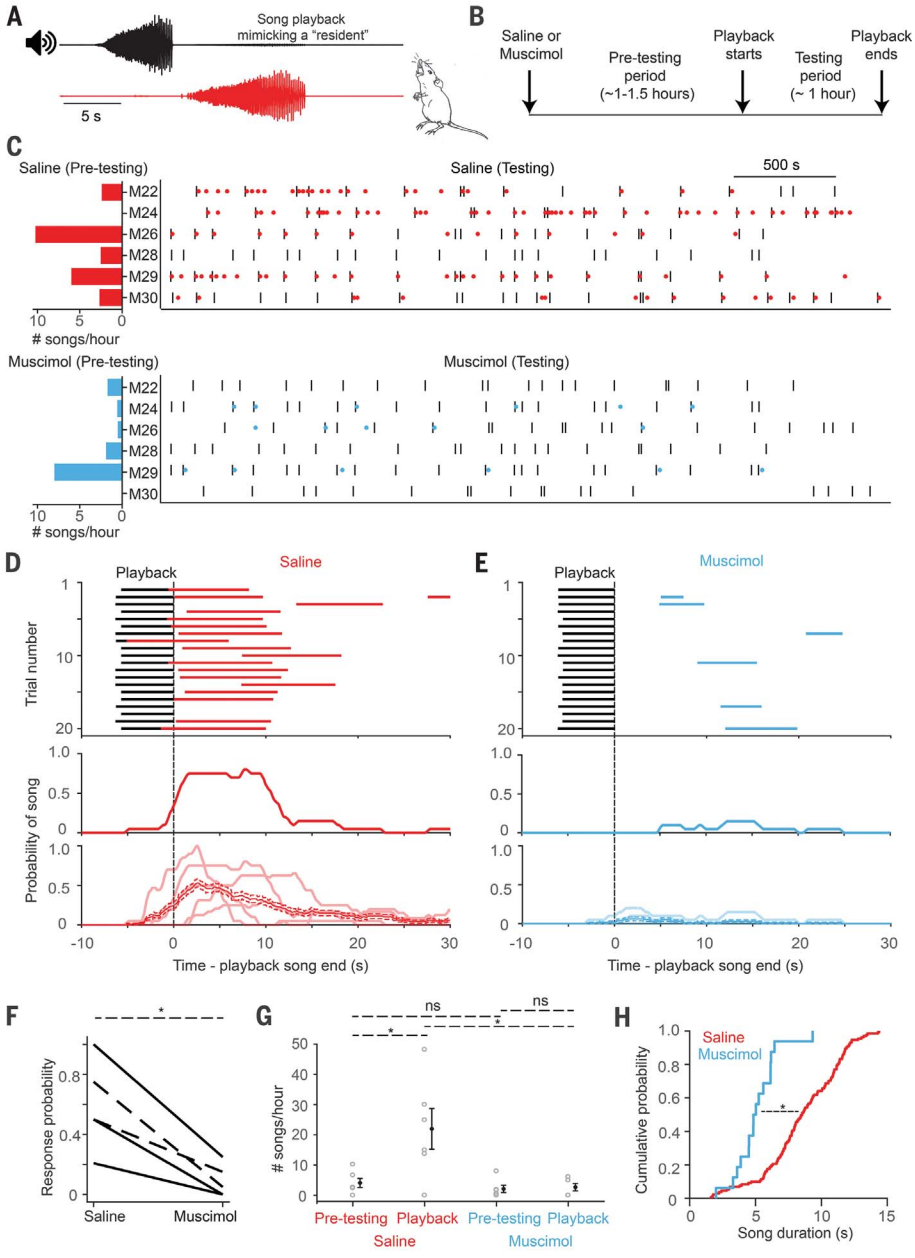

Fig. 4. The OMC is required for countersinging. (**A**) Countersinging response to audio playback of a conspecific male song. (**B**) Design of experimental paradigm. (**C**) Song raster plot of all trials with either saline (top) or muscimol (bottom); each row denotes a different session. Ticks represent playback from a loudspeaker, and colored dots represent *S. teguina* songs. Bar plots indicate the total number of spontaneous songs per hour during the pretesting (nonsocial) period for each animal. (**D** and **E**) Compared with saline (control) injections, dosing with muscimol eliminates a precise countersinging response [top and middle: mouse M29; bottom: entire population (n = 5 mice, mean ± SEM)]. Mouse M28 was excluded because he did not countersing in either condition. (**F**) Countersinging response probability significantly decreases with muscimol treatment compared with saline dosing (*$P < 0.05$ for each animal, n = 5 mice, binomial test). Dashed lines represent cases where the muscimol session preceded the saline session. (**G**) In saline-injected animals, the total number of songs per hour significantly increases during the playback period compared with the pretesting alone period (*$P < 0.05$, n = 6 mice, Kruskal-Wallis test) and the playback period after muscimol inactivation (*$P < 0.05$, n = 6 mice, Kruskal-Wallis test). This increase of song rate during the playback condition was absent upon OMC inactivation with muscimol (P = 0.8068, n = 6 mice, Kruskal-Wallis test). Gray circles represent individual animals; black circles denote mean and SEM. (**H**) Song durations during the playback period are significantly higher for saline-dosed mice compared with muscimol-injected animals (saline: 8.55 ± 0.25 s; muscimol: 5.1 ± 0.41 s, *$P < 0.00001$, Wilcoxon rank sum test).

motor cortex informs subcortical structures to appropriately respond to unexpected sensory stimuli (*47*, *55*) and is consistent with the idea that cortical control may be required for volitional vocal production in primates (*56*). In *S. teguina*, this executive role of the motor cortex may be bolstered by integrating information from other regions, potentially related to factors such as past history and social status. Future studies in which neural activity is monitored during countersinging will help to further refine our understanding of OMC's contribution to this behavior.

The hierarchical control mechanism that appears to underlie countersinging in *S. teguina* features functionally distinct regions responsible for vocal production and coordination. By segregating the vocal motor pathway from cortical control, the structure of the individual notes remains tightly constrained, thus conveying context-invariant information, perhaps related to individual identity (*57*). A similar organizing principle appears in other taxa as well (*6*, *58*, *59*). For instance, cricket stridulation is controlled by a command neuron upstream from central pattern generators (*60*). In songbirds, specific pallial regions are necessary for precise vocal timing of innate calls that are likely to originate subcortically (*61*). These examples of hierarchical control across the animal kingdom suggest a common algorithm that may mediate a wide variety of social interactions.

There has been a recent emphasis on understanding brain function through the lens of complex, ethologically relevant behaviors (*62*, *63*). Here we present *S. teguina* as a new rodent model for investigating neural mechanisms underlying vocal communication with a socially modulated, tractable, and cortically dependent behavior. Moreover, countersinging itself can be temporally segregated into distinct sensory and motor epochs (Fig. 2 and movie S2). Such segregation offers an enormous experimental advantage by recapitulating the organization of task structure typically engineered into standard laboratory sensorimotor paradigms (*1*–*3*) and will allow for the incorporation, testing, and extension of existing hypotheses for analogous brain regions.

REFERENCES AND NOTES

1. W. T. Newsome, K. H. Britten, J. A. Movshon, *Nature* **341**, 52–54 (1989).
2. P. Znamenskiy, A. M. Zador, *Nature* **497**, 482–485 (2013).
3. K. Svoboda, N. Li, *Curr. Opin. Neurobiol.* **49**, 33–41 (2018).
4. S. C. Levinson, *Trends Cogn. Sci.* **20**, 6–14 (2016).
5. M. Hartbauer, S. Kratzer, K. Steiner, H. Römer, *J. Comp. Physiol. A* **191**, 175–188 (2005).
6. B. Hedwig, *J. Comp. Physiol. A* **192**, 677–689 (2006).
7. J. J. Schwartz, *Evolution* **41**, 461–471 (1987).
8. R. Zelick, P. M. Narins, *J. Comp. Physiol. A* **156**, 223–229 (1985).
9. D. J. Mennill, P. T. Boag, L. M. Ratcliffe, *Naturwissenschaften* **90**, 577–582 (2003).
10. S. L. Vehrencamp, J. M. Ellis, B. F. Cropp, J. M. Koltz, *Behav. Ecol.* **25**, 1436–1450 (2014).
11. J. Hyman, *Anim. Behav.* **65**, 1179–1185 (2003).
12. O. Behr, M. Knörnschild, O. Von Helversen, *Behav. Ecol. Sociobiol.* **63**, 433–442 (2009).
13. Y. Goll, V. Demartsev, L. Koren, E. Geffen, *Anim. Behav.* **134**, 9–14 (2017).
14. C. T. Miller, X. Wang, *J. Comp. Physiol. A* **192**, 27–38 (2006).
15. G. G. Carter, M. D. Skowronski, P. A. Faure, B. Fenton, *Anim. Behav.* **76**, 1343–1355 (2008).
16. A. A. Ghazanfar, D. Smith-Rohrberg, A. A. Pollen, M. D. Hauser, *Anim. Behav.* **64**, 427–438 (2002).
17. S. Pika, R. Wilkinson, K. H. Kendrick, S. C. Vernes, *Proc. R. Soc. B* **285**, 20180598 (2018).
18. T. E. Holy, Z. Guo, *PLOS Biol.* **3**, e386 (2005).
19. K. M. Seagraves, B. J. Arthur, S. E. R. Egnor, *J. Exp. Biol.* **219**, 1437–1448 (2016).
20. D. Y. Takahashi, D. Z. Narayanan, A. A. Ghazanfar, *Curr. Biol.* **23**, 2162–2168 (2013).
21. C. T. Miller, A. Wren Thomas, *J. Comp. Physiol. A* **198**, 337–346 (2012).
22. N. Uchida, A. Kepecs, Z. F. Mainen, *Neuroscience* **7**, 485–491 (2006).
23. J. R. Miller, M. D. Engstrom, *J. Mammal.* **88**, 1447–1465 (2007).
24. P. Campbell *et al*., *Evolution* **64**, 1955–1972 (2010).
25. P. Campbell, B. Pasch, A. L. Warren, S. M. Phelps, *PLOS ONE* **9**, e113628 (2014).
26. B. Pasch, B. M. Bolker, S. M. Phelps, *Am. Nat.* **182**, E161–E173 (2013).
27. S. J. Steppan, J. J. Schenk, *PLOS ONE* **12**, e0183070 (2017).
28. M. Konishi, *Annu. Rev. Neurosci.* **8**, 125–170 (1985).
29. G. Pavan *et al*., *J. Acoust. Soc. Am.* **107**, 3487–3495 (2000).
30. G. Arriaga, E. P. Zhou, E. D. Jarvis, *PLOS ONE* **7**, e46610 (2012).
31. G. A. Castellucci, M. J. McGinley, D. A. McCormick, *Sci. Rep.* **6**, 23305 (2016).
32. N. A. Hessler, A. J. Doupe, *Nat. Neurosci.* **2**, 209–211 (1999).
33. V. Demartsev, A. Strandburg-Peshkin, M. Ruffner, M. Manser, *Curr. Biol.* **28**, 3661–3666.e3 (2018).
34. D. A. Liao, Y. S. Zhang, L. X. Cai, A. A. Ghazanfar, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3978–3983 (2018).
35. Y. B. Sirotin, M. E. Costa, D. A. Laplagne, *Front. Behav. Neurosci.* **8**, 399 (2014).
36. T. Riede, in *Handbook of Behavioral Neuroscience*, S. M. Brudzynski, Ed. (Academic Press, 2018), vol. 25, pp. 45–60.
37. U. Jürgens, *Brain Res.* **81**, 564–566 (1974).
38. T. Komiyama *et al*., *Nature* **464**, 1182–1186 (2010).
39. U. Jürgens, *J. Voice* **23**, 1–10 (2009).
40. K. Hammerschmidt, G. Whelan, G. Eichele, J. Fischer, *Sci. Rep.* **5**, 8808 (2015).
41. M. A. Long, M. S. Fee, *Nature* **456**, 189–194 (2008).
42. M. A. Long *et al*., *Neuron* **89**, 1187–1193 (2016).
43. A. Yamaguchi, D. Gooler, A. Herrold, S. Patel, W. W. Pong, *J. Neurophysiol.* **100**, 3134–3143 (2008).
44. L. S. Tang *et al*., *PLOS Biol.* **8**, e1000469 (2010).
45. A. Pires, R. R. Hoy, *J. Comp. Physiol. A* **171**, 79–92 (1992).
46. A. Miri *et al*., *Neuron* **95**, 683–696.e11 (2017).
47. G. Lopes, J. Nogueira, G. Dimitriadis, J. A. Menendez, J. J. Paton, A. R. Kampff, bioRxiv 058917 [Preprint]. 18 May 2017.
48. G. H. Otazu, H. Chae, M. B. Davis, D. F. Albeanu, *Neuron* **86**, 1461–1477 (2015).
49. M. J. Siniscalchi, V. Phoumthipphavong, F. Ali, M. Lozano, A. C. Kwan, *Nat. Neurosci.* **19**, 1234–1242 (2016).
50. S. J. Eliades, X. Wang, *Nature* **453**, 1102–1106 (2008).
51. S. Roy, L. Zhao, X. Wang, *J. Neurosci.* **36**, 12168–12179 (2016).
52. C. T. Miller, A. W. Thomas, S. U. Nummela, L. A. de la Mothe, *J. Neurophysiol.* **114**, 1158–1171 (2015).
53. S. R. Hage, A. Nieder, *Nat. Commun.* **4**, 2409 (2013).
54. C. S. Simões *et al*., *Front. Integr. Neurosci.* **4**, 123 (2010).
55. C. L. Ebbesen, M. Brecht, *Nat. Rev. Neurosci.* **18**, 694–705 (2017).
56. S. R. Hage, A. Nieder, *Trends Neurosci.* **39**, 813–829 (2016).
57. T. T. Burkhard, R. R. Westwick, S. M. Phelps, *Proc. R. Soc. B* **285**, 20180090 (2018).
58. S. Schöneich, B. Hedwig, *Naturwissenschaften* **98**, 1069–1073 (2011).
59. J. M. Kittelberger, B. R. Land, A. H. Bass, *J. Neurophysiol.* **96**, 71–85 (2006).
60. B. Hedwig, *J. Neurophysiol.* **83**, 712–722 (2000).
61. J. I. Benichov *et al*., *Curr. Biol.* **26**, 309–318 (2016).
62. A. Gomez-Marin, J. J. Paton, A. R. Kampff, R. M. Costa, Z. F. Mainen, *Nat. Neurosci.* **17**, 1455–1462 (2014).
63. J. W. Krakauer, A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, D. Poeppel, *Neuron* **93**, 480–490 (2017).

ACKNOWLEDGMENTS

We thank D. Aronov, J. Sakata, C. Ebbersen, S. Ghosh, C. Desplan, and members of the Long laboratory for comments on earlier versions of this manuscript. We thank N. Leal for help with recording the movies. We also thank the individuals involved in caring for and overseeing the NYULMC *S. teguina* colony, especially C. Jackson. **Funding:** This research was supported by the New York Stem Cell Foundation (M.A.L.), NIH grant T32GM007308 (D.E.O.), the Simons Foundation Society of Fellows (A.B.), and the Simons Collaboration on the Global Brain (M.A.L.). S.M.P.'s efforts were funded in part by the NSF (IOS 1457350). **Author contributions:** M.A.L., S.M.P., A.B., and D.E.O. conceived the study and designed the experiments; D.E.O. and A.B. conducted the research; A.B. and D.E.O. performed data analyses; A.B., D.E.O., and M.A.L. created the visualizations; A.M.M.M. collected preliminary data for the project; A.B. wrote the initial draft of the manuscript; M.A.L, A.B., S.M.P., D.E.O., and A.M.M.M. edited and reviewed the final manuscript; S.M.P. provided the starter *S. teguina* mice for establishing the NYU colony; M.A.L. acquired funding; and M.A.L. and S.M.P. supervised the project. **Competing interests:** None declared. **Data and materials availability:** For research and educational purposes, we will share all data collected as part of this project at NeuroData.io.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/363/6430/983/suppl/DC1
Materials and Methods
Figs. S1 to S6
References
Movies S1 and S2

30 July 2018; accepted 23 January 2019
10.1126/science.aau9480

Motor cortical control of vocal interaction in neotropical singing mice

Daniel E. Okobi Jr., Arkarup Banerjee, Andrew M. M. Matheson, Steven M. Phelps and Michael A. Long

Science **363** (6430), 983-988.
DOI: 10.1126/science.aau9480

Turn-taking in singing mice

The ability to take turns is a hallmark of social interaction among animals. It occurs in many different species, from dueting birds to frogs, and is a notable part of human speech. Such rapid response requires a complex cascade of sensory and motor actions that has been difficult to characterize. Okobi *et al.* examined turn-taking in tropical singing mice, in which males interrupt, and alter, each other's songs (see the Perspective by Hage). They describe an orofacial motor cortex that mediates rapid transition from the motor cortex to the vocal motor apparatus and facilitates rapid vocal interactions.

Science, this issue p. 983 see also p. 930

Supplementary Materials for

Motor cortical control of vocal interaction in neotropical singing mice

Daniel E. Okobi Jr.*, Arkarup Banerjee*, Andrew M. M. Matheson, Steven M. Phelps, Michael A. Long†

*These authors contributed equally to this work.
†Corresponding author. Email: mlong@med.nyu.edu

Published 1 March 2019, *Science* **363**, 983 (2019)
DOI: 10.1126/science.aau9480

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/363/6430/983/suppl/DC1)

Movies S1 and S2

Materials and Methods

Animals

All procedures were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee of NYU Langone Medical Center. Animals used in the study were adult (> 3 months) male laboratory-reared offspring of wild-captured *Scotinomys teguina* from La Carpintera and San Gerardo de Dota, Costa Rica (*23*). Mice were maintained at 22 ± 3 °C with a 12:12 L:D cycle.

Behavioral recording

S. teguina were housed in individual recording chambers (Med Associates) lined with sound insulation foam (Soundproof Cow). Vocalizations were recorded using condenser microphone (Avisoft Bioacoustics CM16/CMPA) sampling at 250 kHz (digitized with Avisoft UltraSoundGate 116Hb) that were placed within home cages. In the 'social' condition, a 'recruit mouse' was transferred to nearby acoustic chamber with auditory (but not visual) access to a conspecific male ('resident mouse'). To precisely align the singing behavior of both individuals, each audio stream was additionally recorded continuously into a separate digital signal processor (RX8, PO5e:2150 PC interface card, Tucker-Davis Technologies) at 100 KHz. In some cases, we replaced the 'resident' mouse with song playback through an ultrasonic tweeter (Vifa). For these playback experiments, two male conspecific songs were used 10 times each with random intertrial intervals. To avoid adaptation, playback sessions were separated by at least four days.

To obtain noninvasive measurements of respiratory flow and tidal volumes during song and normal breathing patterns, we placed mice in a whole-body plethysmograph chamber. Respiratory flow was collected via a digital pressure transducer (Emka) and song was collected using an ultrasonic microphone (Avisoft) placed along the wall of the plethysmograph chamber. Respiratory flow and song data were acquired simultaneously and digitized using a high-speed digital acquisition card (National Instruments). In separate experiments, we monitored running speed by affixing a Cadmium-Sulfide photoresistor (RadioShack) within a running wheel (Innovive) that contained an opaque position marker at the rim of the wheel. The frequency of wheel revolutions was used to calculate instantaneous speed using an estimate of elapsed linear running distance over time.

Song analysis

We analyzed song structure using custom software (MATLAB). We first smoothed the sound waveform with a 4 ms sliding window. We then sought to identify individual notes, which typically exhibited an absolute intensity threshold corresponding to 25-40 dB below the mouse's loudest note. Exact note start and stop times were calculated based on the maximum intensity of each note, such that onsets and offsets were first and last crossings of 1% (20 dB quieter) of each note's maximum intensity. Note duration was calculated as the difference between the offset and the onset for each note. For each song, a song trajectory plot was generated in which the duration of each note of the song was as a function of its onset time. The slope of the song trajectory, calculated by fitting a linear regression line between 2 and 6 s, quantifies the rate at which note durations increase per unit time. Vocalizations shorter than 2 s were not considered as songs and a pause greater than 0.5 s (much longer than the typical inter-note durations) indicated the start of a new song. For social interaction and playback experiments, song onsets and offsets were determined using a fixed threshold after band-pass filtering (15-40 kHz) the temporally aligned audio streams.

Surgery

For all surgical procedures, mice were anesthetized with 1-2% isoflurane in oxygen and placed in a stereotaxic apparatus. Two primary devices were implanted in the brain bilaterally as part of these studies: (1) Teflon-coated .002" stainless steel bipolar (400-600 µm separation) stimulating electrodes (California Fine Wire) whose tips were positioned at a depth of 800 µm or (2) Peltier devices (Custom Thermoelectric) similar to those previously used for focal surface cooling (*36*). Devices were centered on the location of the vocal motor cortex (VMC), as determined by intracranial microstimulation (Fig. 3C). A waterproof silicone elastomer (Kwikcast, WPI) was used to seal the craniotomy, and all devices were secured to the skull with dental acrylic and Metabond cement (Parkell). Mice were treated with subcutaneous (Baytril) and topical (Neosporin) antibiotics before being released into their home cages. Manipulations were carried out once singing rates returned to presurgical levels. For the VMC inactivation experiments, 50 nL of either saline or muscimol (1 mg/ml, Sigma) was injected to VMC with a glass pipette (20-30 µm tip diameter) at depths of 400 µm and 700 µm each using an oil-based injection system (Nanoject II, Drummond Scientific, 2.5 nL/cycle, 20 cycles, 8 second interval). Mice were allowed to recover in their home-cages for 1-1.5 hours (pre-testing) before being tested in the countersinging paradigm with audio playback.

For electromyography (EMG) recordings, mice were rotated into a supine position, and an incision was made into the skin above the throat. Skin and fascia were retracted to expose the digastricus muscle bilaterally. EMG recordings were conducted using coated seven-stranded .003" stainless steel wire (A-M Systems) placed in the digastricus muscle. The distal 2-4 mm of each wire was stripped of insulation, folded over, and embedded into the muscle as a flat hook where it was held securely in the muscle using a light layer of VetBond restricted to the muscle insertion point, which was then covered with a waterproof silicone elastomer (Kwikcast, WPI) and secured with sutures. Each EMG recording was obtained as a differential signal between the two wires. For experiments that tracked EMG during song, EMG wires were routed through the fascia between the skin and muscles of the cheek, terminating on a microelectronics connector (Omnetics) affixed to the back of the skull via dental acrylic and Metabond cement (Parkell).

For the intracortical microstimulation experiments, mice were initially anesthetized with isoflurane during the implantation of the EMG recording electrodes and the preparation of the craniotomy and then transferred to ketamine-xylazine (100 mg/kg ketamine, 15 mg/kg xylazine). We used a custom bipolar concentric stimulating electrode fabricated by threading 50 µm insulated stainless steel wire (California Fine Wire) through a 10 µL Nanofil syringe with a 34-gauge beveled tip (WPI). Stimulation sites were located along a grid with either 250 or 500 µm spacing, and sites were evaluated serially along the medial-lateral axis. Occasional deviations from this grid pattern were necessary to avoid large blood vessels. For each site, stimulus trains were delivered 800 µm below the pial surface using an isolated pulse stimulator (A-M Systems). Stimulus trains consisted of four 200 µs long biphasic pulses, delivered 2.5 ms apart. EMG responses were sampled at 40 kHz (National Instruments) following stimulation at a range of amplitudes (10-100 µA, 10 µA steps). Threshold responses were defined as the lowest current at which the standard deviation of the muscle response (2.5 to 35 ms following the stimulus train) was 50% greater than baseline (2.5 to 12.5 ms preceding stimulation) (e.g., Fig. 3B and S3).

Definitions and Statistics

Countersinging probability (Fig 2F) was defined as the ratio of the number of bouts in which the 'recruit' sang within 60 s of the 'resident' mouse's song ending to the total number of times the 'resident' sang. This quantifies the degree of engagement for a 'recruit' mouse to participate in

coordinated vocalizations with the 'resident' mouse. For each 'recruit', the peak (modal value) and the full-width at half maximum of the start or stop probability distributions (Fig. 2B) were defined as the latency and the jitter respectively. Mean latency was calculated as the average across all 'recruits' (n = 8). For each 'recruit' mouse during 'alone' and 'social' sessions, song duration variability was calculated as the standard deviation of all the song durations produced in that session. Song duration variability change (Fig. 2G) was calculated as the percentage change in song duration variability between the 'social' and the 'alone' conditions, with a large value indicating a greater degree of social modulation. Response probability (e.g., Fig. 4F) for each session was defined as the proportion of playback trials in which the 'recruit' mice sang in the interval occurring after the onset of playback song and within 5 s of its end. The likelihood of 'spurious countersinging' for each recruit mouse given the average number of songs per hour alone was estimated by –

$$\text{Chance Response Probability} = \frac{\emptyset * \text{M}}{3600}$$

where, $\emptyset$ is the time-interval over which response probability is calculated (in seconds, see above) and M being the number of recruit songs/hour (Fig. 1F).

Normality of underlying distributions was not assumed and hence non-parametric hypothesis tests were used throughout the manuscript. Wilcoxon signed rank tests were performed when comparing data from each animal across two conditions (e.g., control vs cooling). Wilcoxon rank sum tests were used when data were pooled across animals before comparisons were performed. A Kruskal-Wallis test was performed when comparisons were made across more than two conditions (e.g., Fig. 4G). Since responses to playback stimuli (Fig. 4C-F) is a categorical variable that could either be true or false, for each mouse, we calculated how likely it was to observe the 'muscimol' response probability given the 'saline' responses using a binomial test. We used a permutation test to estimate the expected probability of countersinging by chance given the total number of songs in each session. To do so, the playback timepoints were randomly shuffled to generate 10000 mock playback sessions (20 songs/session). The shuffled response probability was calculated across the mock playback sessions while keeping the experimentally measured song numbers and timepoints fixed (Fig. S6A)To account for the difference between the spontaneous (pre-testing epoch) singing rates between saline and muscimol conditions, the total number of songs in the saline playback condition was downsampled by 52% (ratio of spontaneous singing rates during muscimol compared to saline). As before, the shuffled response probability distribution was calculated by bootstrapping (n = 10,000, Fig. S6B) and p-values were calculated with respect to these shuffled response probability distributions. All data in this study are presented as mean ± SEM unless otherwise indicated.

Song Perturbation: Electrical Stimulation

During VMC stimulation experiments, mice were tethered but allowed to roam freely about the cage. The chamber was small enough to ensure that mice were never more than 30 cm from the microphone. Stimulation (biphasic, 200 µs duration, 5-20 pulses at 25 Hz, 25 to 900 µA) was administered using an isolated pulse stimulator (A-M Systems) triggered by song. Changes to song were classified as either *Pauses*, defined as the cessation of singing for the duration of at least one note, or *Truncations*, defined as cases in which song production halted for at least 5 seconds, such that any resumption in song was considered an independent song attempt.

To analyze the effect of electrical stimulation of the VMC on singing behavior, we examined song trajectory plots. The 10 notes produced before stimulation were used to fit a linear pre-stimulus song trend (e.g., Fig. 3E-G). Because song occasionally resumed with quiet and sometimes intermittent short notes, song was deemed to have continued when notes as loud as the quietest pre-stimulus trend note were produced. Using this criterion, the first 10 notes following stimulation were used to fit a post-stimulus song trend. Only songs in which the point of continuation fell at least 400 ms after the beginning of stimulation were analyzed. Changes in note length between the time of stimulation and official resumption were noted and compared to an extrapolated song trend line made by averaging pre- and post-stimulus song trend lines for individual songs.

Song Perturbation: Cooling

During the course of cooling experiments, a DC power supply (Kepco) was connected to the mouse via a wire tether. Current steps were applied continuously, usually for 20-60 minute periods. We excluded from our analysis any songs recorded during the initial 5 minutes of every cooling trial in which the temperature was asymptotically approaching a steady value. For each animal, the degree of cooling was intentionally varied in a pseudorandom order to limit interactions across trials. A water-based heat sink capable of maintaining a flow rate of 10 mL/min was used to keep the Peltier device at a consistent temperature. We calibrated the extent of cooling with a miniature thermocouple (Custom Thermoelectric) placed ~1 mm below the implanted cooling probe in isoflurane anesthetized mice. As before, we quantified slope by fitting a linear regression line to relatively linear portion of the song trajectory plots (2 – 6 s). Cooling-related changes in note number were also quantified within this time epoch. Time to threshold was defined as the elapsed time in seconds between the start of the song to the beginning of the first note longer than 75 ms.

fig. S1.

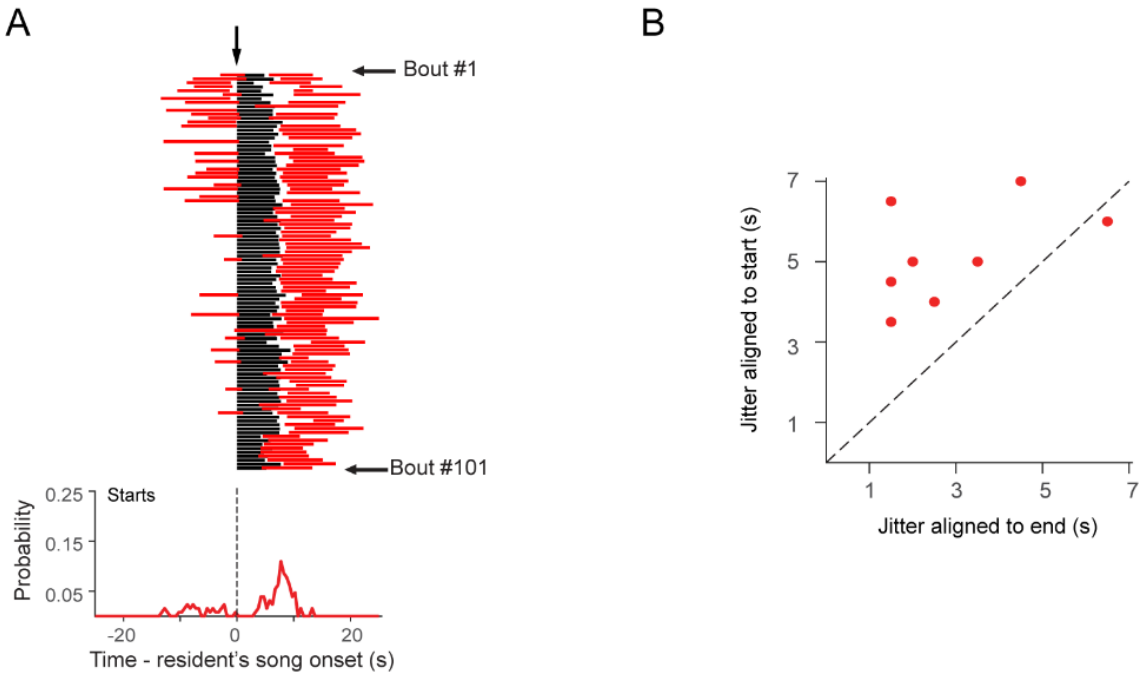

fig. S1. Determining the stimulus trigger for the recruit's song onset. (A) All vocal interactions (n = 101) for a single pair of *S. teguina* throughout a 24-hour period aligned to the beginning of resident's songs. An equivalent plot aligned to the end of the resident's songs is provided elsewhere (Figure 2B, *left*). The corresponding song onset probability distributions are plotted underneath. (B) Song initiation jitter of the recruit mouse is significantly smaller when aligned to the end of resident's song rather than the beginning (n = 8, $p < 0.05$, Wilcoxon signed rank test).

fig. S2.

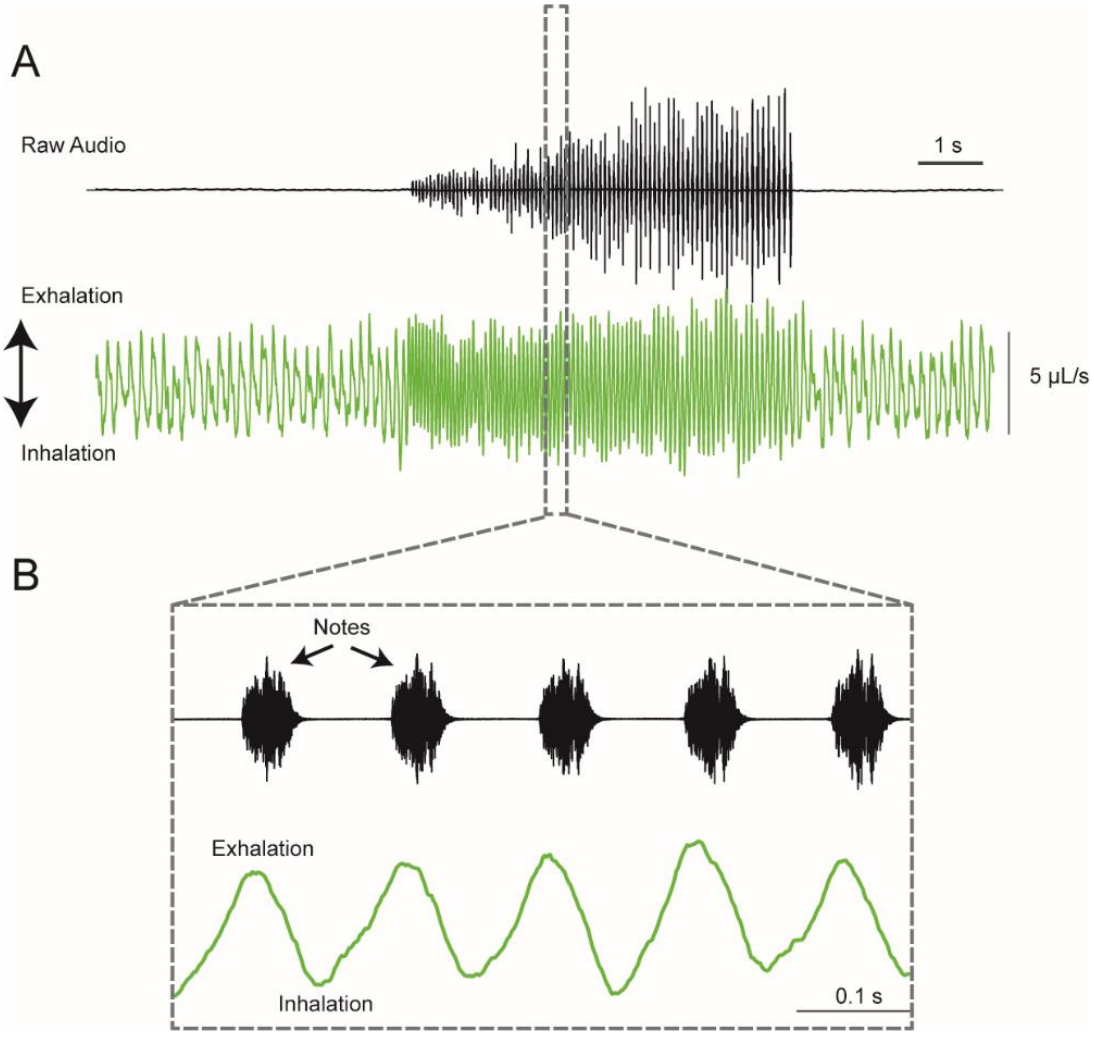

fig. S2. Relationship between respiratory rhythm and song production. (A) Song waveform *(top, black)* shown with simultaneous plethysmograph recording *(bottom, green)*. Upward deflections represent exhalation. During song, the basal respiratory rate increases considerably, matching note production. (B) An expanded view of the region highlighted in dotted box in A. The pattern of high-frequency oscillation between exhalation (notes) and inhalation (mini-breaths) is clearly visible, with each note produced during the exhalation phase of the respiratory rhythm.

fig. S3.

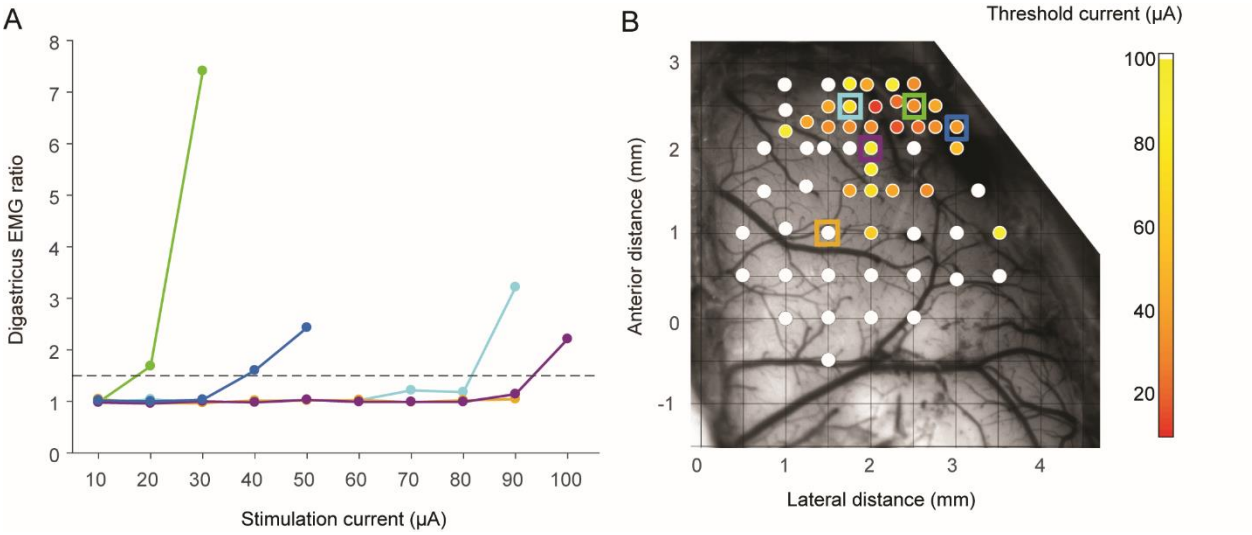

fig. S3. A full intracortical microstimulation (ICMS) map of frontal cortex for one example animal. (A) Example EMG responses (digastricus muscle) as a function of stimulation current from four different cortical locations in one *S. teguina*. The dashed line indicates the minimum current (current threshold) required to elicit a significant EMG response defined as a response 50% greater than baseline (see Materials and Methods). (B) All sampled areas for one individual were represented by colored circles overlaid at the appropriate location on a background of a *S. teguina* brain image. The color bar indicates the current threshold needed to achieve a significant EMG response. Sites where no responses were triggered are plotted in white circles. Example traces from A are indicated by colored squares surrounding the corresponding location.

fig. S4.

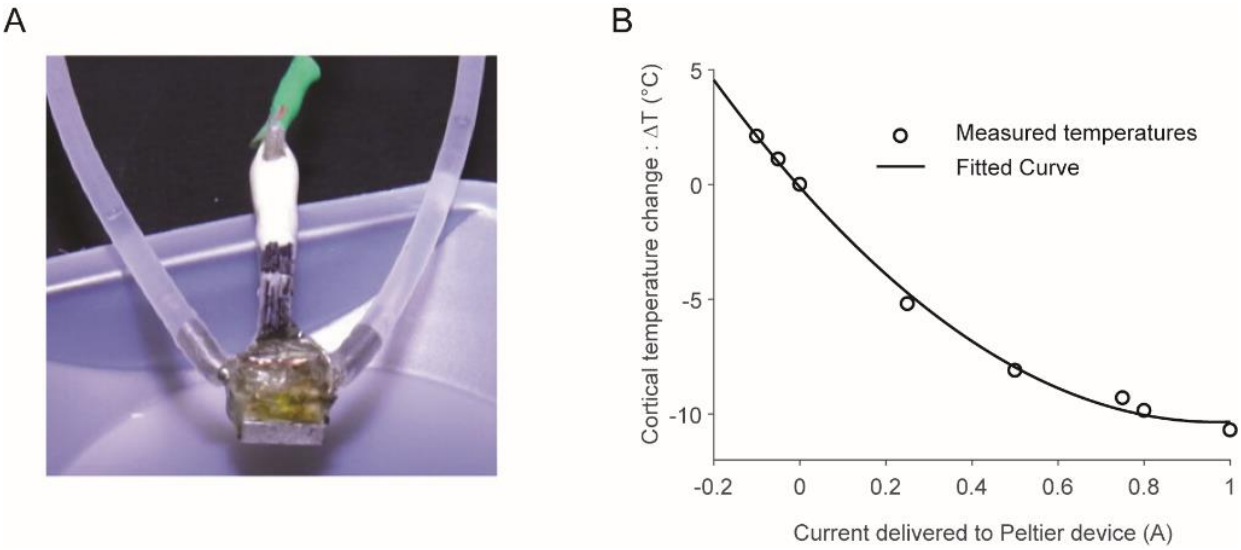

fig S4. Peltier cooling device. (A) Cooling device composed of Peltier elements, a copper heat sink with hosing to circulate water, and a microconnector (pictured attached to tether) that interfaces with the surface of the brain via a bioinert platinum sheet (long axis: 4 mm). (B) Cooling probe performance calibrated within the cortex of an anesthetized mouse. The calibration curve shows the amount of cooling achieved as a function of delivered current.

fig. S5.

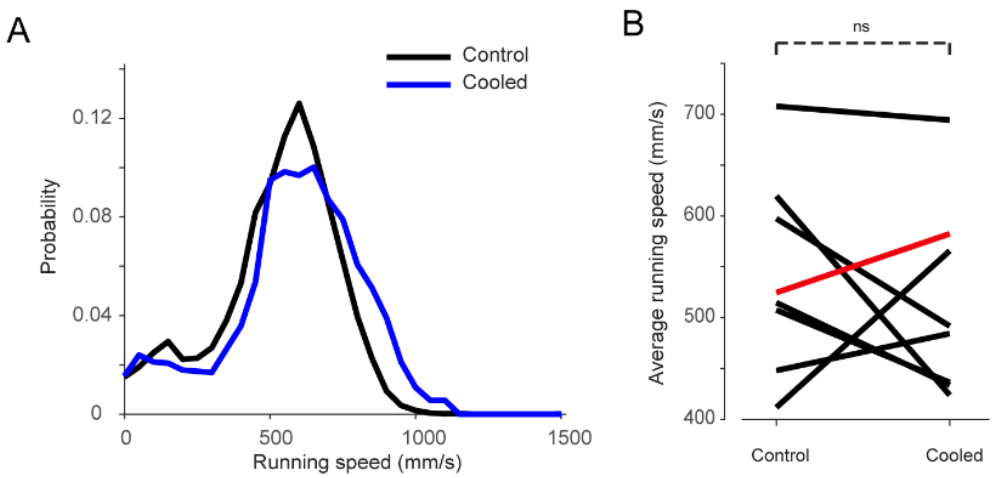

fig S5. Cooling motor cortex does not affect running speed. (A) Histogram of running speed of an example *S. teguina* during control and cooling sessions. (B) Across the population of animals tested (n = 8), there was no significant difference between the average running speed during cooling of the motor cortex (control: 541.4 ± 34.0 mm/s; cooled: 514.0 ± 33.2 mm/s, p = 0.46, Wilcoxon rank sum test). The red line denotes the exemplar mouse from panel A.

fig. S6.

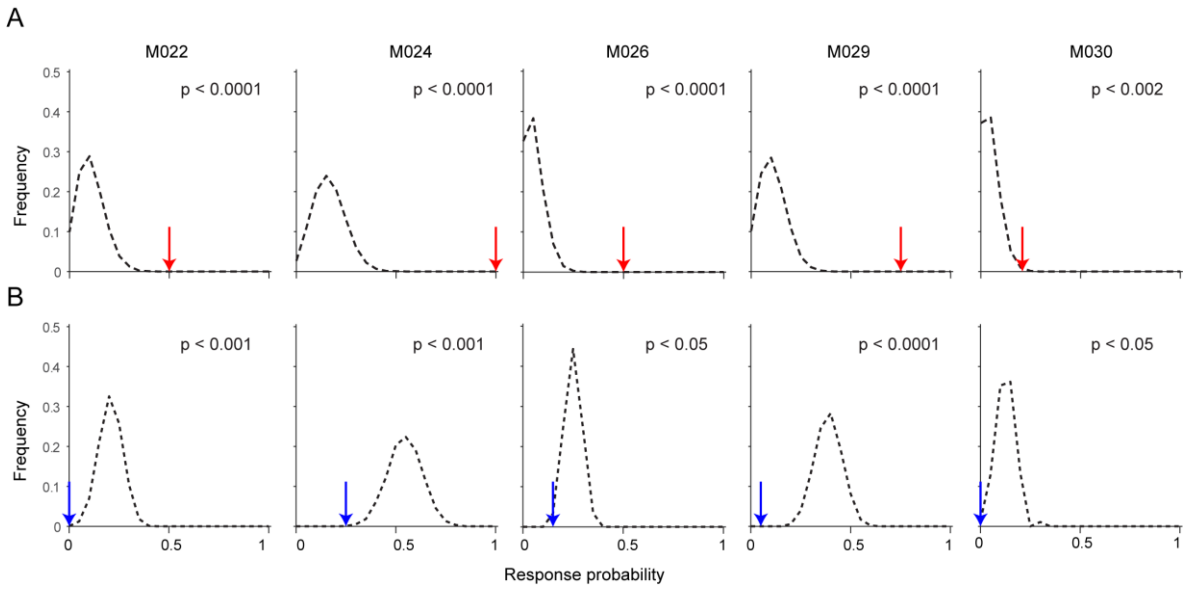

fig S6. Song rates alone cannot quantitatively account for the measured response probabilities in both saline and muscimol conditions. (A) For each animal, the playback song times were randomly shuffled (bootstrapping, n = 10,000) to generate the expected response probability distributions (*dashed line*) given the amount of singing in the 'saline' playback sessions. The observed response probabilities (*red arrows*) were significantly higher than this null distribution, rejecting the hypothesis that increased song rates during control playback sessions could simply explain the high degree of sensorimotor coupling we observed. (B) To account for a reduction in spontaneous song rates after muscimol injection, we downsampled the total number of songs during the control playback sessions to match the amount of singing in the muscimol case and again generated the expected response probability distributions (*dashed lines*). For each animal, the observed response probability after muscimol inactivation (*blue arrows*) was significantly lower than expected by chance. p-values are indicated on the panel.

Supplementary Movie Legends:

Movie S1. Spontaneous song of a male *S. teguina* mouse. Video recording of an adult male *S. teguina* in acoustic isolation spontaneously singing an advertisement song. The characteristic posture and the jaw movements accompanying each note can be observed, especially for longer notes towards the end of the song.

Movie S2. A male *S. teguina* mouse engaged in countersinging. Playback of a conspecific song from outside the homecage (the first and fainter song) elicits an advertisement song from a male *S. teguina*.

References and Notes

1. W. T. Newsome, K. H. Britten, J. A. Movshon, Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989). doi:10.1038/341052a0 Medline

2. P. Znamenskiy, A. M. Zador, Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013). doi:10.1038/nature12077 Medline

3. K. Svoboda, N. Li, Neural mechanisms of movement planning: Motor cortex and beyond. *Curr. Opin. Neurobiol.* **49**, 33–41 (2018). doi:10.1016/j.conb.2017.10.023 Medline

4. S. C. Levinson, Turn-taking in Human Communication—Origins and Implications for Language Processing. *Trends Cogn. Sci.* **20**, 6–14 (2016). doi:10.1016/j.tics.2015.10.010 Medline

5. M. Hartbauer, S. Kratzer, K. Steiner, H. Römer, Mechanisms for synchrony and alternation in song interactions of the bushcricket Mecopoda elongata (Tettigoniidae: Orthoptera). *J. Comp. Physiol. A* **191**, 175–188 (2005). doi:10.1007/s00359-004-0586-4 Medline

6. B. Hedwig, Pulses, patterns and paths: Neurobiology of acoustic behaviour in crickets. *J. Comp. Physiol. A* **192**, 677–689 (2006). doi:10.1007/s00359-006-0115-8 Medline

7. J. J. Schwartz, The function of call alternation in anuran amphibians: A test of three hypotheses. *Evolution* **41**, 461–471 (1987). doi:10.1111/j.1558-5646.1987.tb05818.x Medline

8. R. Zelick, P. M. Narins, Characterization of the advertisement call oscillator in the frog Eleutherodactylus coqui. *J. Comp. Physiol. A* **156**, 223–229 (1985). doi:10.1007/BF00610865

9. D. J. Mennill, P. T. Boag, L. M. Ratcliffe, The reproductive choices of eavesdropping female black-capped chickadees, Poecile atricapillus. *Naturwissenschaften* **90**, 577–582 (2003). doi:10.1007/s00114-003-0479-3 Medline

10. S. L. Vehrencamp, J. M. Ellis, B. F. Cropp, J. M. Koltz, Negotiation of territorial boundaries in a songbird. *Behav. Ecol.* **25**, 1436–1450 (2014). doi:10.1093/beheco/aru135 Medline

11. J. Hyman, Countersinging as a signal of aggression in a territorial songbird. *Anim. Behav.* **65**, 1179–1185 (2003). doi:10.1006/anbe.2003.2175

12. O. Behr, M. Knörnschild, O. Von Helversen, Territorial counter-singing in male sac-winged bats saccopteryx bilineata: Low-frequency songs trigger a stronger response. *Behav. Ecol. Sociobiol.* **63**, 433–442 (2009). doi:10.1007/s00265-008-0677-2

13. Y. Goll, V. Demartsev, L. Koren, E. Geffen, Male hyraxes increase countersinging as strangers become ‘nasty neighbours’. *Anim. Behav.* **134**, 9–14 (2017). doi:10.1016/j.anbehav.2017.10.002

14. C. T. Miller, X. Wang, Sensory-motor interactions modulate a primate vocal behavior: Antiphonal calling in common marmosets. *J. Comp. Physiol. A* **192**, 27–38 (2006). doi:10.1007/s00359-005-0043-z Medline

15. G. G. Carter, M. D. Skowronski, P. A. Faure, B. Fenton, Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim. Behav.* **76**, 1343–1355 (2008). doi:10.1016/j.anbehav.2008.04.023

16. A. A. Ghazanfar, D. Smith-Rohrberg, A. A. Pollen, M. D. Hauser, Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins. *Anim. Behav.* **64**, 427–438 (2002). doi:10.1006/anbe.2002.3074

17. S. Pika, R. Wilkinson, K. H. Kendrick, S. C. Vernes, Taking turns: Bridging the gap between human and animal communication. *Proc. R. Soc. B* **285**, 20180598 (2018). doi:10.1098/rspb.2018.0598 Medline

18. T. E. Holy, Z. Guo, Ultrasonic songs of male mice. *PLOS Biol.* **3**, e386 (2005). doi:10.1371/journal.pbio.0030386 Medline

19. K. M. Seagraves, B. J. Arthur, S. E. R. Egnor, Evidence for an audience effect in mice: Male social partners alter the male vocal response to female cues. *J. Exp. Biol.* **219**, 1437–1448 (2016). doi:10.1242/jeb.129361 Medline

20. D. Y. Takahashi, D. Z. Narayanan, A. A. Ghazanfar, Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr. Biol.* **23**, 2162–2168 (2013). doi:10.1016/j.cub.2013.09.005 Medline

21. C. T. Miller, A. Wren Thomas, Individual recognition during bouts of antiphonal calling in common marmosets. *J. Comp. Physiol. A* **198**, 337–346 (2012). doi:10.1007/s00359-012-0712-7 Medline

22. N. Uchida, A. Kepecs, Z. F. Mainen, Seeing at a glance, smelling in a whiff: Rapid forms of perceptual decision making. *Neuroscience* **7**, 485–491 (2006). Medline

23. J. R. Miller, M. D. Engstrom, Vocal Stereotypy and Singing Behavior in Baiomyine Mice. *J. Mammal.* **88**, 1447–1465 (2007). doi:10.1644/06-MAMM-A-386R.1

24. P. Campbell, B. Pasch, J. L. Pino, O. L. Crino, M. Phillips, S. M. Phelps, Geographic variation in the songs of neotropical singing mice: Testing the relative importance of drift and local adaptation. *Evolution* **64**, 1955–1972 (2010). Medline

25. P. Campbell, B. Pasch, A. L. Warren, S. M. Phelps, Vocal ontogeny in neotropical singing mice (Scotinomys). *PLOS ONE* **9**, e113628 (2014). doi:10.1371/journal.pone.0113628 Medline

26. B. Pasch, B. M. Bolker, S. M. Phelps, Interspecific dominance via vocal interactions mediates altitudinal zonation in neotropical singing mice. *Am. Nat.* **182**, E161–E173 (2013). doi:10.1086/673263 Medline

27. S. J. Steppan, J. J. Schenk, Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. *PLOS ONE* **12**, e0183070 (2017). doi:10.1371/journal.pone.0183070 Medline

28. M. Konishi, Birdsong: From behavior to neuron. *Annu. Rev. Neurosci.* **8**, 125–170 (1985). doi:10.1146/annurev.ne.08.030185.001013 Medline

29. G. Pavan, T. J. Hayward, J. F. Borsani, M. Priano, M. Manghi, C. Fossati, J. Gordon, Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985-1996. *J. Acoust. Soc. Am.* **107**, 3487–3495 (2000). doi:10.1121/1.429419 Medline

30. G. Arriaga, E. P. Zhou, E. D. Jarvis, Of mice, birds, and men: The mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLOS ONE* **7**, e46610 (2012). doi:10.1371/journal.pone.0046610 Medline

31. G. A. Castellucci, M. J. McGinley, D. A. McCormick, Knockout of Foxp2 disrupts vocal development in mice. *Sci. Rep.* **6**, 23305 (2016). doi:10.1038/srep23305 Medline

32. N. A. Hessler, A. J. Doupe, Social context modulates singing-related neural activity in the songbird forebrain. *Nat. Neurosci.* **2**, 209–211 (1999). doi:10.1038/6306 Medline

33. V. Demartsev, A. Strandburg-Peshkin, M. Ruffner, M. Manser, Vocal Turn-Taking in Meerkat Group Calling Sessions. *Curr. Biol.* **28**, 3661–3666.e3 (2018). doi:10.1016/j.cub.2018.09.065 Medline

34. D. A. Liao, Y. S. Zhang, L. X. Cai, A. A. Ghazanfar, Internal states and extrinsic factors both determine monkey vocal production. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3978–3983 (2018). doi:10.1073/pnas.1722426115 Medline

35. Y. B. Sirotin, M. E. Costa, D. A. Laplagne, Rodent ultrasonic vocalizations are bound to active sniffing behavior. *Front. Behav. Neurosci.* **8**, 399 (2014). doi:10.3389/fnbeh.2014.00399 Medline

36. T. Riede, in *Handbook of Behavioral Neuroscience*, S. M. Brudzynski, Ed. (Academic Press, 2018), vol. 25, pp. 45–60.

37. U. Jürgens, On the elicitability of vocalization from the cortical larynx area. *Brain Res.* **81**, 564–566 (1974). doi:10.1016/0006-8993(74)90853-1 Medline

38. T. Komiyama, T. R. Sato, D. H. O'Connor, Y.-X. Zhang, D. Huber, B. M. Hooks, M. Gabitto, K. Svoboda, Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464**, 1182–1186 (2010). doi:10.1038/nature08897 Medline

39. U. Jürgens, The neural control of vocalization in mammals: A review. *J. Voice* **23**, 1–10 (2009). doi:10.1016/j.jvoice.2007.07.005 Medline

40. K. Hammerschmidt, G. Whelan, G. Eichele, J. Fischer, Mice lacking the cerebral cortex develop normal song: Insights into the foundations of vocal learning. *Sci. Rep.* **5**, 8808 (2015). doi:10.1038/srep08808 Medline

41. M. A. Long, M. S. Fee, Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456**, 189–194 (2008). doi:10.1038/nature07448 Medline

42. M. A. Long, K. A. Katlowitz, M. A. Svirsky, R. C. Clary, T. M. A. Byun, N. Majaj, H. Oya, M. A. Howard 3rd, J. D. W. Greenlee, Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation. *Neuron* **89**, 1187–1193 (2016). doi:10.1016/j.neuron.2016.01.032 Medline

43. A. Yamaguchi, D. Gooler, A. Herrold, S. Patel, W. W. Pong, Temperature-dependent regulation of vocal pattern generator. *J. Neurophysiol.* **100**, 3134–3143 (2008). doi:10.1152/jn.01309.2007 Medline

44. L. S. Tang, M. L. Goeritz, J. S. Caplan, A. L. Taylor, M. Fisek, E. Marder, Precise temperature compensation of phase in a rhythmic motor pattern. *PLOS Biol.* **8**, e1000469 (2010). doi:10.1371/journal.pbio.1000469 Medline

45. A. Pires, R. R. Hoy, Temperature coupling in cricket acoustic communication. II. Localization of temperature effects on song production and recognition networks in Gryllus firmus. *J. Comp. Physiol. A* **171**, 79–92 (1992). doi:10.1007/BF00195963 Medline

46. A. Miri, C. L. Warriner, J. S. Seely, G. F. Elsayed, J. P. Cunningham, M. M. Churchland, T. M. Jessell, Behaviorally Selective Engagement of Short-Latency Effector Pathways by Motor Cortex. *Neuron* **95**, 683–696.e11 (2017). doi:10.1016/j.neuron.2017.06.042 Medline

47. G. Lopes, J. Nogueira, G. Dimitriadis, J. A. Menendez, J. J. Paton, A. R. Kampff, A robust role for motor cortex. bioRxiv 058917 [Preprint]. 18 May 2017. https://doi.org/10.1101/058917.

48. G. H. Otazu, H. Chae, M. B. Davis, D. F. Albeanu, Cortical Feedback Decorrelates Olfactory Bulb Output in Awake Mice. *Neuron* **86**, 1461–1477 (2015). doi:10.1016/j.neuron.2015.05.023 Medline

49. M. J. Siniscalchi, V. Phoumthipphavong, F. Ali, M. Lozano, A. C. Kwan, Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nat. Neurosci.* **19**, 1234–1242 (2016). doi:10.1038/nn.4342 Medline

50. S. J. Eliades, X. Wang, Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* **453**, 1102–1106 (2008). doi:10.1038/nature06910 Medline

51. S. Roy, L. Zhao, X. Wang, Distinct Neural Activities in Premotor Cortex during Natural Vocal Behaviors in a New World Primate, the Common Marmoset (Callithrix jacchus). *J. Neurosci.* **36**, 12168–12179 (2016). doi:10.1523/JNEUROSCI.1646-16.2016 Medline

52. C. T. Miller, A. W. Thomas, S. U. Nummela, L. A. de la Mothe, Responses of primate frontal cortex neurons during natural vocal communication. *J. Neurophysiol.* **114**, 1158–1171 (2015). doi:10.1152/jn.01003.2014 Medline

53. S. R. Hage, A. Nieder, Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* **4**, 2409 (2013). doi:10.1038/ncomms3409 Medline

54. C. S. Simões, P. V. Vianney, M. M. de Moura, M. A. Freire, L. E. Mello, K. Sameshima, J. F. Araújo, M. A. Nicolelis, C. V. Mello, S. Ribeiro, Activation of frontal neocortical areas by vocal production in marmosets. *Front. Integr. Neurosci.* **4**, 123 (2010). doi:10.3389/fnint.2010.00123 Medline

55. C. L. Ebbesen, M. Brecht, Motor cortex - to act or not to act? *Nat. Rev. Neurosci.* **18**, 694–705 (2017). doi:10.1038/nrn.2017.119 Medline

56. S. R. Hage, A. Nieder, Dual Neural Network Model for the Evolution of Speech and Language. *Trends Neurosci.* **39**, 813–829 (2016). doi:10.1016/j.tins.2016.10.006 Medline

57. T. T. Burkhard, R. R. Westwick, S. M. Phelps, Adiposity signals predict vocal effort in Alston's singing mice. *Proc. R Soc. B* **285**, 20180090 (2018). doi:10.1098/rspb.2018.0090 Medline

58. S. Schöneich, B. Hedwig, Neural basis of singing in crickets: Central pattern generation in abdominal ganglia. *Naturwissenschaften* **98**, 1069–1073 (2011). doi:10.1007/s00114-011-0857-1 Medline

59. J. M. Kittelberger, B. R. Land, A. H. Bass, Midbrain periaqueductal gray and vocal patterning in a teleost fish. *J. Neurophysiol.* **96**, 71–85 (2006). doi:10.1152/jn.00067.2006 Medline

60. B. Hedwig, Control of cricket stridulation by a command neuron: Efficacy depends on the behavioral state. *J. Neurophysiol.* **83**, 712–722 (2000). doi:10.1152/jn.2000.83.2.712 Medline

61. J. I. Benichov, S. E. Benezra, D. Vallentin, E. Globerson, M. A. Long, O. Tchernichovski, The forebrain song system mediates predictive call timing in female and Male Zebra finches. *Curr. Biol.* **26**, 309–318 (2016). doi:10.1016/j.cub.2015.12.037 Medline

62. A. Gomez-Marin, J. J. Paton, A. R. Kampff, R. M. Costa, Z. F. Mainen, Big behavioral data: Psychology, ethology and the foundations of neuroscience. *Nat. Neurosci.* **17**, 1455–1462 (2014). doi:10.1038/nn.3812 Medline

63. J. W. Krakauer, A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, D. Poeppel, Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron* **93**, 480–490 (2017). doi:10.1016/j.neuron.2016.12.041 Medline