The breath shape controls intonation of mouse vocalizations

Alastair MacDonald[1], Alina Hebling[2], Xin Paul Wei[1,3], and Kevin Yackle[1*]

[1]Department of Physiology, University of California-San Francisco, San Francisco, CA 94143

[2]Neuroscience Graduate Program, University of California-San Francisco, San Francisco, CA 94143, USA

[3]Biomedical Sciences Graduate Program, University of California-San Francisco, San Francisco, CA 94143, USA

*Lead contact

* Lead author email: kevin.yackle@ucsf.edu

Abstract:

Intonation in speech is the control of vocal pitch to layer expressive meaning to communication, like increasing pitch to indicate a question. Also, stereotyped patterns of pitch are used to create distinct sounds with different denotations, like in tonal languages and, perhaps, the ten sounds in the murine lexicon. A basic tone is created by exhalation through a constricted laryngeal voice box, and it is thought that more complex utterances are produced solely by dynamic changes in laryngeal tension. But perhaps, the shifting pitch also results from altering the swiftness of exhalation. Consistent with the latter model, we describe that intonation in most vocalization types follows deviations in exhalation that appear to be generated by the re-activation of the cardinal breathing muscle for inspiration. We also show that the brainstem vocalization central pattern generator, the iRO, can create this breath pattern. Consequently, ectopic activation of the iRO not only induces phonation, but also the pitch patterns that compose most of the vocalizations in the murine lexicon. These results reveal a novel brainstem mechanism for intonation.

Introduction:

Modulation of the frequency of produced sound, perceived as pitch, creates meaning within words or phrases through intonation (*Prieto et al., 2015*). For example, in English, an increasing pitch is used to indicate a question or stress importance and a decreasing pitch communicates a declaration. Additionally, the concatenation of specialized sounds accented by variations in pitch enriches the composition of spoken language, like how the same sound at divergent pitches can relay different meanings (*Howie et al., 1976*). Two key pieces of the phonation system are the larynx (the "voice-box") and the breathing muscles *(Berke et al., 2009, Laplange et al., 2018*). Succinctly, the breathing muscles drive airflow through a narrowed larynx to produce a basic vocalization (*Finck et al., 2009*). The local speed of the airflow through the larynx dictates the fundamental frequency of the tone, so changes in either the swiftness of the breath exhalation or the extent of laryngeal closure can both, presumably, alter the pitch (*Kelm-Nelson et al., 2018, Herbst C.T., 2016, Mahrt et al., 2016*). While control of the size of the laryngeal opening is well established as a mechanism to regulate the dynamic changes in pitch for human to rat and mouse vocalizations (*Poeppel et al., 2020, Johnson et al., 2010, Riede et al., 2017*), the contribution of exhalation itself remains to be carefully defined. In fact, it is presumed that the velocity of expiration only modulates the vocal amplitude or loudness (*Riede T., 2011, Riede T., 2013*). This perception stems from the airflow of the rodent breath not strongly predicting the pitch. Yet paradoxically, an injection of air below the larynx to enhance flow increases pitch (*Riede T., 2011*). This incongruity even extends to songbirds, a leading vocalization model system (*Suthers et al., 2002, Schmidt et al., 2014, Plummer et al., 2008, Goller et al., 2004*). Here, we seek to resolve this inconsistency by taking advantage of the experimental, behavioral, and genetic approaches in the mouse (*Yackle K., 2023*). If two independent variables are used to alter pitch, like the larynx and the breath airflow, then the interplay would enhance the ability to produce a diverse repertoire of sounds and thereby enable a broader lexicon.

The medullary brainstem possesses at least two means that might account for the two control points proposed above, laryngeal diameter and exhalation speed. First, modulation of some laryngeal premotor neurons in the retroambiguus (RAm) and perhaps intermingled or nearby motor neurons modulates the size of the laryngeal opening (*Kelm-Nelson et al., 2018, Hage S.R., 2009a, Park et al*., 2024) and is even sufficient to evoke vocalizations across species, albeit mostly abnormal (*Zhang et al.,* 1995, *Hartmann & Brecht*, 2020, *Veerakumar et al.,* 2023, *Park et al*., 2024). Second, the vocalization central pattern generator we recently described, called the intermediate Reticular Oscillator (iRO), induces coordinated changes in the expiratory airflow and laryngeal closure (*Wei et al., 2022*). For example, during neonatal cries, the iRO oscillates exhalation speed and larynx activity to time the syllable sounds. Thus, the RAm provides a mechanism to modulate pitch by controlling laryngeal diameter independently from the iRO altering the tone by dictating the speed of the breath expiratory airflow. While the contribution of Ram in adult phonation has recently been established in mice (*Veerakumar et al., 2023, Park et al., 2024*), the role of the iRO remains undefined.

Here, we describe the coordinated changes in breath airflow and pitch in the ten vocalizations of the adult murine lexicon (*Grimsley et al., 2011*). We describe that the modulation of pitch for the different vocalizations either correlates or anticorrelates with the changes in exhalation velocity. These results support a model in which two independent mechanisms involving changes in laryngeal opening airflow control tone. Consistently, we found that the pattern of breathing muscle activity is different for each mechanism. For example, the muscle that drives the inspiration phase of the breath is ectopically re-engaged to break expiratory airflow and decrease pitch. This mirrors how the iRO rhythmically patterns the breathing muscles to time syllables during neonatal cries. Using anatomical, molecular, and functional approaches, we demonstrate that the iRO vocal central pattern generator drives changes in breath expiratory airflow to pattern pitch and produce seven of the ten vocalizations in the endogenous lexicon. These data resolve the prior paradoxical role of exhalation speed in sound production and show

it can directly control pitch. Additionally, we establish the iRO as a mechanistic basis for intonation. And lastly, these results generalize the crucial role of the iRO in phonation across developmental stages and, we presume, across species.

Results:

Vocalizations are produced by a program coupled to breathing.

It is possible that the ten murine ultrasonic vocalizations (USVs) defined by unique pitch patterns ('syllable types'; *Grimsley et al., 2011*) are formed by distinct breaths or as substructures nested in a common breath. Prior work has suggested the latter (*Sirotin et al., 2014*). To expand upon this, we simultaneously measured breathing and USVs by customizing the lid of a whole-body plethysmography chamber to accommodate a microphone. Male mice in the chamber were exposed to fresh female urine and robustly sniffed and vocalized for the first 5-10 minutes of the recording at a peak rate of about 4 events per second (n = 6) (Fig. 1A-B). A vocalization was classified as a narrow-band sound in the 40-120 kHz ultrasonic frequency range during a single breath (Fig. 1A). The instantaneous frequency of vocalization breaths was typically between 5-10 Hz (mean: 7.5 Hz) (Fig. 1C) and mostly occurred during episodes of rapid sniffs (8.5 - 10 Hz) (Fig. 1A and B), as previously reported (*Sirotin et al., 2014, Castellucci et al., 2018*). When compared to neighboring breaths, vocalization breaths were slightly slower overall (Fig. 1C) with subtly larger inspiratory and perhaps expiratory airflow despite similar durations of each phase (Fig. 1D and E). These data reveal that a vocalization breath appears mostly like a normal breath but with the addition of a nested sound pattern. This led us to hypothesize that a distinct sub-program is activated within a breath to generate a vocalization.

The adult murine lexicon is composed of at least 10 USV syllable types that are defined by different, but stereotyped, patterns of pitch. Most breaths contain a single syllable (88%), which we define as a continuous USV event (Fig. 1F and S2-3), however, on some occasions, we observed 2 (8%) or 3 (4%) syllables separated by >20 ms within a single breath (Fig. S3). This

structure mirrors that described in neonatal cries (*Wei et al*., 2022). A pre-trained convolutional neural network (CNN) was used to classify USVs into different types based on changes in pitch (*Fonseca et al., 2021*), and the on- and offset of each vocalization was overlayed upon the corresponding breath airflow (Fig. S1-3). Vocalizations began and ended throughout expiration (Fig. S1), and the most common tended to start near the onset of exhalation and ended shortly there-after (like the up frequency modulated, step down, flat, and short types) (Fig. S2). Vocalizations with more intricate changes in pitch had more variable times of on- and offset (like complex, chevron, two step, multi, step up, down frequency modulated) (Fig. S3). And lastly, when the vocalizations occurred late in expiration, the duration of this breath phase was prolonged (Fig. S4). The bias of USV timing by the breath, combined with the USV modulation of breath length demonstrate these programs are independent but reciprocally coupled.

Two mechanisms create the changes in pitch pattern.

Fluctuations in airflow speed through the larynx produce changes in the sound's pitch. For example, augmenting airflow through the explanted rodent larynx increases pitch (*Mahrt et al., 2016*). We proposed two potential mechanisms that could contribute to how the laryngeal airflow is modulated to form the distinct USV types in the murine lexicon: one based on the swiftness of exhalation pushing air through the larynx (model 1), and the other based on the diameter of the laryngeal opening (model 2). According to the first model, changes in pitch positively correlate with the breath expiratory airflow measured by plethysmography, which we term positive intonation. On the other hand (model 2), a narrowed larynx increases pitch by speeding local airflow while simultaneously impeding the overall expiratory airflow measured by plethysmography; we call this negative intonation. Note, these models can form similar expiratory airflow patterns, but predict opposite relationships to pitch.

We assessed for evidence of each model by calculating the correlation coefficient (r) between instantaneous expiratory airflow and the corresponding USV fundamental frequency.

Down or up frequency modulated (fm) USVs served as simple USV examples and we found that these were positively or negatively correlated, respectively (median r = 0.62 and -0.46, Fig. 2A and C). These two were also distinguished by the sound onset and offset, whereby down fm started and ended later during the expiration (Fig. S2-3). These distinguishing features are consistent with the sounds being produced by separate mechanisms to alter pitch, positive and negative intonation.

Between these mechanisms, most syllable types displayed positive intonation. Five of the other eight USV syllable types had positively shifted intonation, the complex, step down, chevron, two step, and multi (median r = 0.31, 0.28, 0.32, 0.19, and 0.24 respectively) (Fig. 2B, D and Fig. S5). In particular, the mirrored oscillations in the breath expiratory airflow and pitch during a complex vocalization best illustrated the positively coupled relationship (model 1, Fig. 2C). In contrast, the up fm was the only USV with negative intonation. The r values for positively and negatively correlated USVs were not random, as they fell outside r values computed in simulated datasets composed of USVs with shuffled expiration airflow or expirations with shuffled pitch frequencies (Fig. 2 and Fig. S5).

The two step and step up USVs (median r = 0.19 and -0.03) appeared to have a portion of the pitch pattern correlated with the expiratory airflow, while the other part(s) were un- or anti-correlated (e.g., the two step, Fig. 2E). This suggests that the pitch is produced by switching between positive and negative intonation mechanisms within the breath. The remaining two USV types (flat and short) occurred with various breath shapes which resulted in a wide range of r values (Fig. S5). Some USV syllable types are defined by the presence of large, instantaneous jumps in frequency (like the step up, step down, two- and multi-step). A jump was not associated with a corresponding change in airflow. Also, the jumps did not predict the subsequent intonation pattern. For example, step up syllables could be negative to negative (20%), negative to positive (20%), or positive to positive (60%) intonations.

Across all USV types, we found that the relationship between airflow and pitch is relative rather than absolute which is expected since the relationship is determined by at least two independent variables (laryngeal tone and exhalation speed). In summary, these results establish an important positive connection between the breath expiratory airflow to modulate pitch (model 1). This supports the hypothesis that a vocalization pattern generator must integrate with and even control the breath airflow as a key mechanism to produce various USV types in the murine lexicon.

Inspiratory and laryngeal muscles have coordinated activity that represents positive and negative intonation.

To explore the mechanisms underlying the two intonation models, we simultaneously recorded the electrical activity of the primary muscles for breathing, the diaphragm, and laryngeal control (thyroarytenoid and cricothyroid) during basal breathing and bouts of vocalizations in male mice. Electromyography (EMG) electrodes were permanently placed along the diaphragm and inserted into the larynx, while breathing and USVs were measured in parallel by whole-body plethysmography and a microphone (as in Fig. 1). Electrocardiogram signals were annotated and removed from the diaphragm EMG recordings post-hoc. According to our models and the timing of the syllable types within the breath (Fig. S2-3), we anticipated that USVs with positive intonation would have a coordinated re-activation of the diaphragm and laryngeal muscles later in expiration, while the up fm would only have laryngeal activity at expiration onset.

EMG activity during basal breathing displayed the three phases of the breath cycle (*Yackle, 2023)*. The diaphragm was active during inspiration and larynx in the subsequent post-inspiration period, and these were followed by a late expiration phase where neither muscle was active (Fig. 3A). The airflow measured by plethysmography had a ~10ms lag when compared to the EMG activities (blue and red arrows, Fig. 3A). Upon exposure to female odors, the male mice produced bouts of USVs containing the entire lexicon for several minutes, demonstrating that

electrodes interfered with neither breathing nor sound production (Fig. 3B). Breaths containing USVs were distinguished from adjacent breaths lacking sound by an increase in laryngeal muscle activity that prolonged into expiration. Like airflow, the sound followed EMG activity by ~10ms. More than 10 mice were studied and, unsurprisingly, given the invasiveness of the EMGs only three produced robust signals and the number of acquired vocalizations varied from tens to thousands between these mice (n=70, 1482, 2819). The analysis below is from a single animal that had the clearest EMG signals, but the same results and relationships were observed in all three animals.

To study positive intonation, we analyzed the down fm and complex USVs. Down fms started later during expiration and began at a drop in expiratory airflow (Fig. 3C). The diaphragm activity occurred just before the decrease in expiratory airflow (blue arrow, Fig. 3C) and the following laryngeal activity persisted during the sound. This pattern reflects the activity of the muscles during a normal breath, inspiration (diaphragm) then post-inspiration (larynx), but this pattern was ectopically embedded within an expiration. Similarly, complex vocalizations had the diaphragm then laryngeal activity pattern during expiration, but this cycled multiple times concurrent with the increase and decrease in expiratory airflow and pitch (Fig. 3C). Complex USVs had an average of 1.5 cycles per expiration, the interval between diaphragm bursts was 43 $\pm$ 14 ms, and the laryngeal activity occurred at 69 $\pm$ 12% through this diaphragm-to-diaphragm interval. Also, in ~19% of the complex USVs, sound, laryngeal, and diaphragm activity co-occurred, suggesting that other mechanisms contribute to the diversity of sounds.

Up fm represents negative intonation, and correspondingly, the activity of the muscles distinct from the positively intonated types. At the peak of the post-inspiratory period of the breath, the vocalization began and persisted throughout the laryngeal muscle activity (Fig. 3C). The sound ended at the onset of a burst in the diaphragm EMG (Fig. 3C).

In summary, these EMG studies of the key inspiratory muscle and larynx serve to reflect the core components of the breathing central pattern generator (CPG) that produce inspiration

and post-inspiration. Ectopic activation of these antiphase patterns during the expiration of a vocal breath appears to result in a USV with positive intonation and the cyclic engagement of this leads to an oscillating pitch. Additionally, the termination of the USV with negative intonation corresponds to the re-activation of inspiratory muscles (Fig. 3D). The novel finding that the endogenous pattern of the breathing CPG is re-engaged within an adult vocal breath, a 'mini-breath', mimics the rhythmic syllables of the neonatal vocalizations patterned by the intermediate Reticular Oscillator (iRO;*Wei et al., 2022*). This suggests that the iRO also plays a central role in the production of adult vocalizations.

The iRO resides within the adult brainstem phonation circuit.

The iRO has yet to be identified in adult mice. The iRO is molecularly defined in the neonate by the co-expression of *Preproenkephalin* (*Penk*) and *Vesicular glutamate transporter 2* (*Vglut2*) and is anatomically localized to the medullary ventral intermediate Reticular Formation (iRT) directly medial to the compact nucleus ambiguus (*Wei et al., 2022*). We determined that the iRO molecular and anatomical features exist in adults in two ways. First, we generated triple transgenic mice that label *Penk*$^{+}$*Vglut2*$^{+}$ neurons and the derived lineages with tdTomato (*Penk*-Cre; *Vglut2*-Flp; Ai65) (Fig. 4A). And second, we stereotaxically injected the iRO region of *Penk*-Cre; *Vglut2*-Flp mice with a Cre and Flp dependent reporter adeno-associated virus (AAV Cre^{ON}Flp^{ON}-ChR2::YFP) (Fig. 4B-C). Consistent with the definition of the iRO in neonatal mice, tdTomato+ and YFP+ *Penk*$^{+}$*Vglut2*$^{+}$ neurons were found in the iRT adjacent to the compact nucleus ambiguus (Fig. 4A-C). These results demonstrate that the ventrolateral medulla of adult mice contains neurons with the molecular and anatomical identity of the iRO.

Neonatal iRO neurons are presynaptic to the kernel of breathing, the pacemaker for inspiration (preBötzinger Complex, preBötC) (*Smith et al., 1991*), and premotor to multiple laryngeal and tongue muscles. We traced the YFP+ axons of *Penk*$^{+}$*Vglut2*$^{+}$ neurons (*Penk*-Cre; *Vglut2*-Flp and AAV Cre^{ON}Flp^{ON}-ChR2::YFP) and found they elaborated within the nucleus

ambiguus (NA) and retroambiguus (RAm) where laryngeal premotor and motor neurons localize (Fig. 4A, D), the breathing pacemaker (Fig. 4E), and the hypoglossal (tongue) motor nucleus (Fig. 4E). The projection patterns of these *Penk*$^{+}$*Vglut2*$^{+}$ neurons provide additional evidence that these adult neurons maintain the same connectivity properties as the neonatal iRO neurons, indicating they can control the key elements for vocalization: the breath airflow and larynx.

In adult mice, vocalizations have been triggered by activation of the midbrain periaqueductal gray (PAG), namely glutamatergic neurons in the lateral to ventrolateral subregion (*Michael et al., 2020, Chen et al., 2021, Tschida et al., 2019*). Note, an exact definition of PAG-USV stimulating neurons remains to be satisfactorily described. To assess if the iRO region is positioned downstream of the ventrolateral PAG, we unilaterally injected *Vglut2*-Cre mice with a Cre^{ON}-ChR2::YFP expressing retrograde traveling AAV (AAVrg) (Fig. 4F). Among the labeled brain regions, we found YFP+ neurons in a region of the midbrain PAG overlapping with areas that contain PAG-USV neurons. To our surprise, neurons from the ipsi- and contralateral PAG projected to the iRO region in nearly equal numbers (Fig. 4F). These molecular, anatomical, and neural morphology characterizations reveal that the iRO exists in adults and is embedded within the brainstem phonation network (PAG → iRO → the preBötC, NA, RAm, hypoglossal) (Fig. 4G).

Ectopic activation of the putative iRO induced vocalization.

If these labeled *Penk*$^{+}$*Vglut2*$^{+}$ neurons are indeed the iRO, we anticipated that ectopic activation would induce vocalization. We tested this in two ways. First, we generated *Penk*-Cre;*Vglut2*-Flp;Cre^{ON}Flp^{ON}-ReaChR triple transgenic mice which express the red-shifted Channel Rhodopsin in *Penk*$^{+}$;*Vglut2*$^{+}$ neurons and the derived lineage (ReaChR mice) and second, we stereotaxically injected the AAV Cre^{ON}Flp^{ON}-Channel Rhodopsin2::YFP (ChR2) into the iRO region of *Penk*-Cre;*Vglut2*-Flp mice. In both instances we implanted optic fibers above the iRO bilaterally to further localize neural activation (Fig. 5A and S6A). In both experimental regimes, ectopic light activation of the *Penk*$^{+}$*Vglut2*$^{+}$ neurons induced bouts of vocalizations where the breathing rate

was entrained by the frequency of stimulation (Fig. 5A, S6A, S6I). Most bouts and the breaths within contained vocalizations (Fig. 5B and S6B), and the amplitudes of all elicited breaths were significantly increased (Fig. S6J). Some AAV-ChR2 mice showed previously described broad band harmonic vocalizations (like *Grimsley et al., 2011)*, while others did not vocalize (n=5/9), likely due to incomplete labeling. Note, the audible sound during inspirations in these animals reflects orofacial movement artifacts that result in the fiberoptic colliding with the plethysmography chamber walls. Additionally, the ReaChR animals without vocalizations were found to have "off target" optic fiber implants (n=2/6). Taken together, these data are consistent with the notion that the iRO is sufficient to induce phonation via control of both breath airflow and laryngeal opening, just as it does in neonatal cries.

To demonstrate the specialization of the iRO neurons for vocalization and the inability of modulated breathing alone to elicit USVs, we performed several additional control experiments. First, to ensure that just stimulation of breathing is insufficient to elicit vocalization, we optogenetically excited the glutamatergic preBötC neurons (*Vglut2*-Cre with AAV Cre^{ON}-ChR2). Indeed, we found that, although breathing sped up, optogenetic stimulation never elicited vocalizations (Fig. S6C, H, I). And second, to determine if the ability to elicit vocalizations was generalizable to other neural types in the iRO anatomical region, we activated *Penk*$^{+}$, *μ-opioid receptor*$^{+}$*Vglut2*$^{+}$, *Tachykinin 1*$^{+}$, and *Vesicular GABA transporter*$^{+}$ neurons and found that vocalizations were never induced upon light stimulation, although breathing was altered in various ways (Fig. S6). In summary, these data functionally demonstrate the existence of *Penk*$^{+}$*Vglut2*$^{+}$ iRO neurons in adult mice and their ability to create vocalizations by modulating both breathing and presumably the larynx.

Excitation of the iRO evoked nearly the entire murine lexicon.

Above, we described that one mechanism for generating the different patterns of vocalizations was via the modulation of the breath airflow (positive intonation). Once again, this was defined as

a positive correlation between expiratory airflow and pitch (Fig. 2). We hypothesized that this property stems from the iRO's capacity to control breathing, and so we made the following predictions: 1) that the USVs evoked after stimulation would be biased to those with an endogenous positive correlation between airflow and pitch (like the down fm and step down), and 2) that the elicited USVs would be transformed to become more positively correlated.

We classified the evoked iRO vocalizations (*Penk*-Cre;*Vglut2*-Flp;Cre^{ON}Flp^{ON}-ReaChR) with the CNN, and to our surprise, seven of the ten types of endogenous USVs were induced upon activation of the iRO (Fig. 5). The most abundant elicited USV was the down fm which, in the endogenous dataset, had the strongest positive intonation (Fig. 5G and Fig. 2). Conversely, the USV with the strongest negative intonation was rarely found, up fm. These results are striking since the down fm is the least common endogenous USV and up fm is the most common (Fig. 2). These results are consistent with the first prediction where the optically evoked USV types were biased towards those with endogenous positive intonation. Beyond this, all the ectopic USVs combined had a more positive association between airflow and pitch compared to all endogenous USVs (Fig. 5D). This was not explained purely by the abundance of down fm since, when these were omitted from the analysis the positive bias was unchanged. This aligns with the second prediction. These data demonstrate that the iRO is sufficient to pattern nearly all USV types, and that the pitch of the induced vocalizations tightly follows the breathing airflow.

Discussion:

Here we propose that the intonation that establishes the diversity of the adult murine lexicon is explained by two mechanisms, the modulation of the breath waveform and presumably the size of the laryngeal opening. We describe that unique vocalization types have characteristic fluctuations in the expiratory airflow, whereby some changes in pitch are strongly correlated with airflow (positive intonation) while others are anti-correlated (negative intonation). These two mechanisms can even be used in the same breath to produce complex changes in pitch. To our

surprise, six of the ten USV types primarily used the positive intonation mechanism. These data support a novel and key role for the breathing system in the production of various types of vocalizations. In support of this hypothesis, we found re-activation of the breathing central pattern generator during the expiration phase of the vocal breath. This ectopic activity appeared as a 'mini-breath' (inspiratory diaphragm then post-inspiratory laryngeal muscle activities) nested within a normal expiration and corresponded with modulated airflow and pitch. This resembles how the vocalization central pattern generator, the iRO, patterns the rhythmic syllable structure of neonatal cry vocalizations. We show that the iRO is sufficient to induce most of the endogenous USV syllable types via the modulation of the breath airflow. In contrast to the natural lexicon, the pitch of the evoked USVs is primarily explained by positive intonation. These data imply that the iRO can produce the mechanism to pattern positive intonation, thereby suggesting that negative intonation derives from a separate neuronal component of the phonatory system. We propose that these two mechanisms can be used independently or in conjunction to generate the diverse repertoire of vocalizations (Fig. 5H).

The iRO likely patterns intonation for endogenous phonation.

The description of the iRO within the adult neural circuit for phonation suggests a key role in patterning the endogenous adult vocalizations. In this case, we propose that the upstream periaqueductal gray input would "turn-on" the iRO which then co-opts the breathing pacemaker and coordinates its anti-phase activity with the laryngeal muscles to produce and pattern the changes in breath airflow and vocal pitch (positive intonation). The iRO can do this since it is presynaptic to both the breathing pacemaker and the laryngeal motor neurons. In this case, the brief re-activation of inspiratory muscles we observed would slow ongoing expiration, enabling a decrease in airflow speed, and thus pitch. After, relaxation of these muscles results in an increase in expiration airflow and pitch. This type of oscillatory modulation we observed has also been seen in neonatal cries generated by the iRO (*Wei et al., 2022*). An important next step will be to validate

the necessity of the iRO in adult phonation, as anticipated from its necessary role in neonatal vocalization. None-the-less, the presence of the iRO across developmental stages implies a conserved role in innate vocalizations within the mouse and perhaps across the animal kingdom, where vocalization central pattern generators have been hypothesized and even identified in species from fish to birds to primates (*Zhang et al., 2020, Chagnaud et al., 2011, Hage S.R., 2009b, Kelley et al., 2020*).

The iRO can autonomously produce multiple vocalization patterns.

A surprising finding is that ectopic activation of the iRO produces seven of the ten syllable types within the murine lexicon. How might this occur? One possibility is that the iRO has multiple modes which can each produce a different pattern of activity. Such a phenomenon has been demonstrated in other central pattern generating systems like the crustacean stomatogastric ganglia (*Marder et al., 2001, Marder E., 2012*). A more likely option is that additional mechanisms of vocal modulation are layered upon a basic pattern produced by the iRO. For example, other regions with direct control of the laryngeal motor neurons within RAm would add complexity to the vocalization induced by the iRO, akin to how vocal control by the human laryngeal motor cortex is conceptualized (Fig. 5H) (*Dichter et al., 2018, Silva et al., 2022)*. Here we propose that perhaps just two mechanisms (breath airflow and laryngeal opening) account for the intricacy of the murine sounds produced, and the layering of these enables a basic pitch structure within a breath to become sophisticated. Of note, these resulting models do not simplistically explain the origin of the pitch jumps that are present in some USV types and these may instead arise from mechanisms like an active process to alter the conformation of the larynx or upper airway.

Recently, glutamatergic laryngeal premotor neurons in RAm were identified that are sufficient to elicit vocalizations, albeit somewhat abnormal, and also necessary to produce sound (*Park et al., 2024, Veerakumar et al., 2023)*. This raises the possibility that these RAm neurons compose the cellular basis for the laryngeal control mechanism we propose produces negative

intonation (Fig. 5H). However, the fact that ectopic excitation of these neurons or iRO elicits vocalizations suggests that reciprocal connections are engaged to create the dynamics of vocalizations. Consistently, iRO neurons project to RAm and vice versa (Fig. 3 and *Park et al., 2024, Veerakumar et al., 2023)*. Interesting future studies will be to assess the necessity of either site to produce USVs upon ectopic excitation of the complementary region.

The control of breathing airflow is a novel biomechanical mechanism for intonation.

Intonation is a key aspect of communication, whereby the same word or phrase could be used as a question or a statement simply by different fluctuations in pitch. Additionally, in tonal languages, the same sound with differences in pitch can have completely different meanings. Our findings describe a novel biophysical mechanism for intonation and a cellular basis. Now, the iRO or direct modulation of breathing can serve as a starting point to map higher level components of brain-wide vocalization circuits that structure additional subliminal layers of perception in speech.

Acknowledgments: We thank Dr. YoonJeung Chang and Beatriz Cuevas for assistance with microscopy. We thank Dr. David Julius, and members of the Yackle lab for their input and revision of the manuscript. Funding: This work was supported by the Brain Initiative R34 NS127104, NINDS R01 NS126400, and the Simon's Foundation.

Author Contributions: A.M. performed stereotaxic brain injections, optogenetic stimulations, and histological studies. A.H. and X.P.W. developed and performed EMG studies. A.M. collected and analyzed vocalization, breathing, and EMG data. A.M. and K.Y. designed experiments and wrote the manuscript.

Competing interests: Authors declare no competing interests. Data and materials availability: All data collected in this study and code use for analysis are available upon request from the corresponding author.

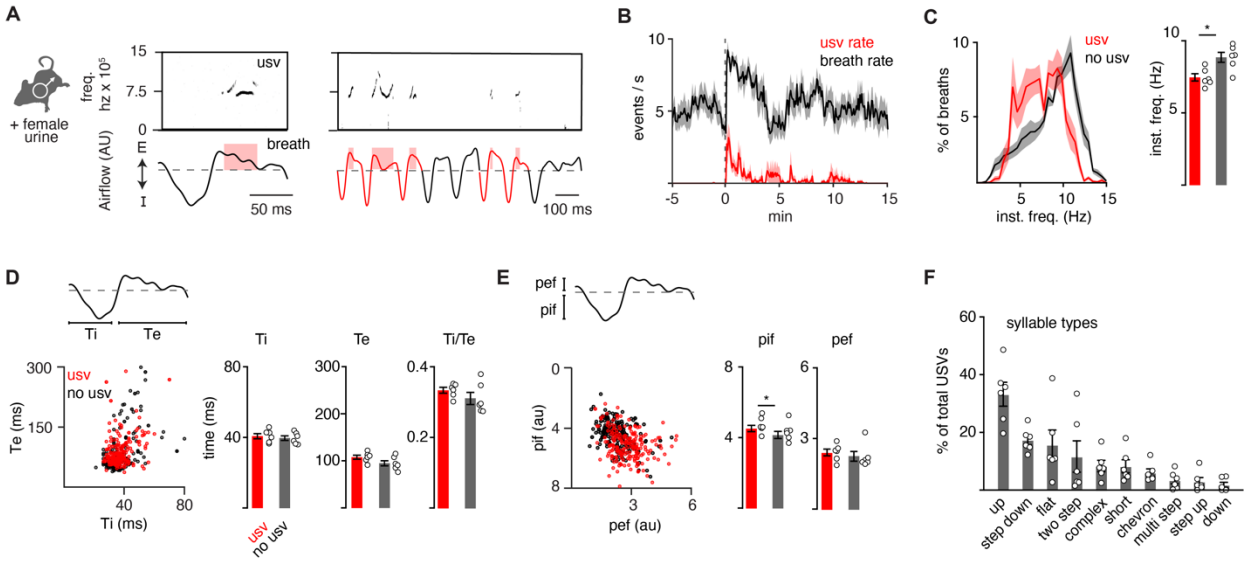

Figure 1. The full repertoire of vocalizations occurs within a normal appearing breath. A, Male mice exposed to female urine produce ultrasonic vocalizations (USV) at about 75 kilohertz (top) that coincide with the expiratory airflow (E, arbitrary units) of the breath cycle (bottom). Red box indicates the length of the USV. A bout of vocalizations contains breaths with USVs (red) interspersed among sniff breaths (black). **B**, Rates of breathing (black) and USV production (red). Exposure to female urine at time 0, n=6 mice. **C**, Left, histogram of the instantaneous frequency of breaths with and without USVs from n=6 animals. Right, average instantaneous frequency for each mouse (mean ±SEM). Each dot is the mean from each animal. p-value 0.03; paired t-test. **D**, Scatter plot of the inspiratory (Ti) and expiratory time (Te) for USV (red) and non-USV (black) breaths from a representative animal. Right, bar graph of mean ± SEM of Ti, Te, and the ratio for n=6. p-values 0.40, 0.18, and 0.25; paired t-test. **E**, The breath peak inspiratory (pif) and expiratory (pef) airflow represented as in **D**. p-values 0.01, 0.27; paired t-test. **F**, Bar graph (mean ±SEM) of the percent of total USVs for each type from n=6 mice.

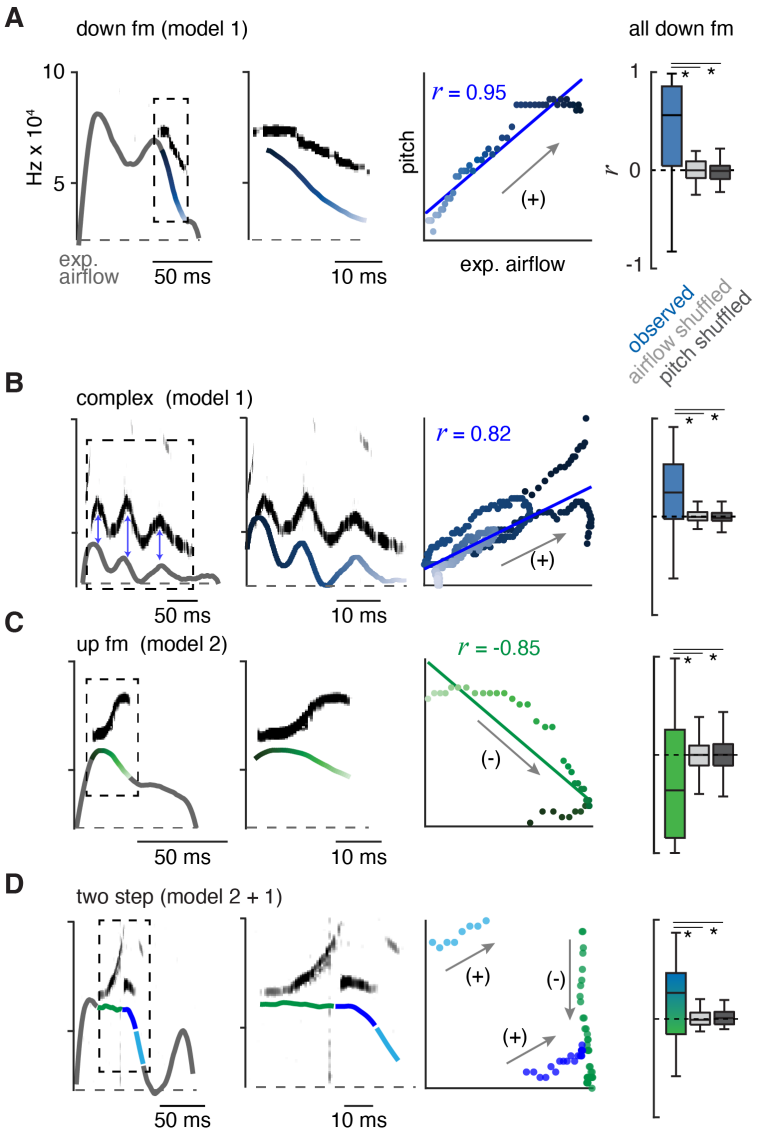

Figure 2. The ten types of USVs are produced by at least two mechanisms that modulate airflow. A, Left, example of the expiratory airflow and pitch for a down frequency modulated (fm) USV. Middle, magnification of airflow and sound. The scale of airflow is not displayed. The time of breath airflow from expiration onset during the USV is color coded blue to white. Scatter plot of instantaneous expiratory airflow and pitch for the single USV and the correlation (line, r). Note, the change in pitch mirrors airflow (annotated as "+"), consistent with the 'breath modulation' called model 1. Box and whisker plot of n=40 down fm correlation coefficients (r). Controls, normal

USV pitch vs shuffled airflow or shuffled pitch to normal USV expiratory airflow. **B-D**, Representative expiratory airflow and pitch, box and whisker plot of all r values, and onset / offset for complex (n=165), up fm (n = 589), and two step (n=61) vocalizations represented as in **A**. The airflow for each unique USV element is uniquely color coded as green, blue, or purple. Note, the change in pitch for two components correlates and one anticorrelates. This is consistent with both mechanisms being sequentially used. Annotated as mixed blue and green box and whisker plot. * = $p < 0.05$, one-way ANOVA with Sidak's post-hoc test.

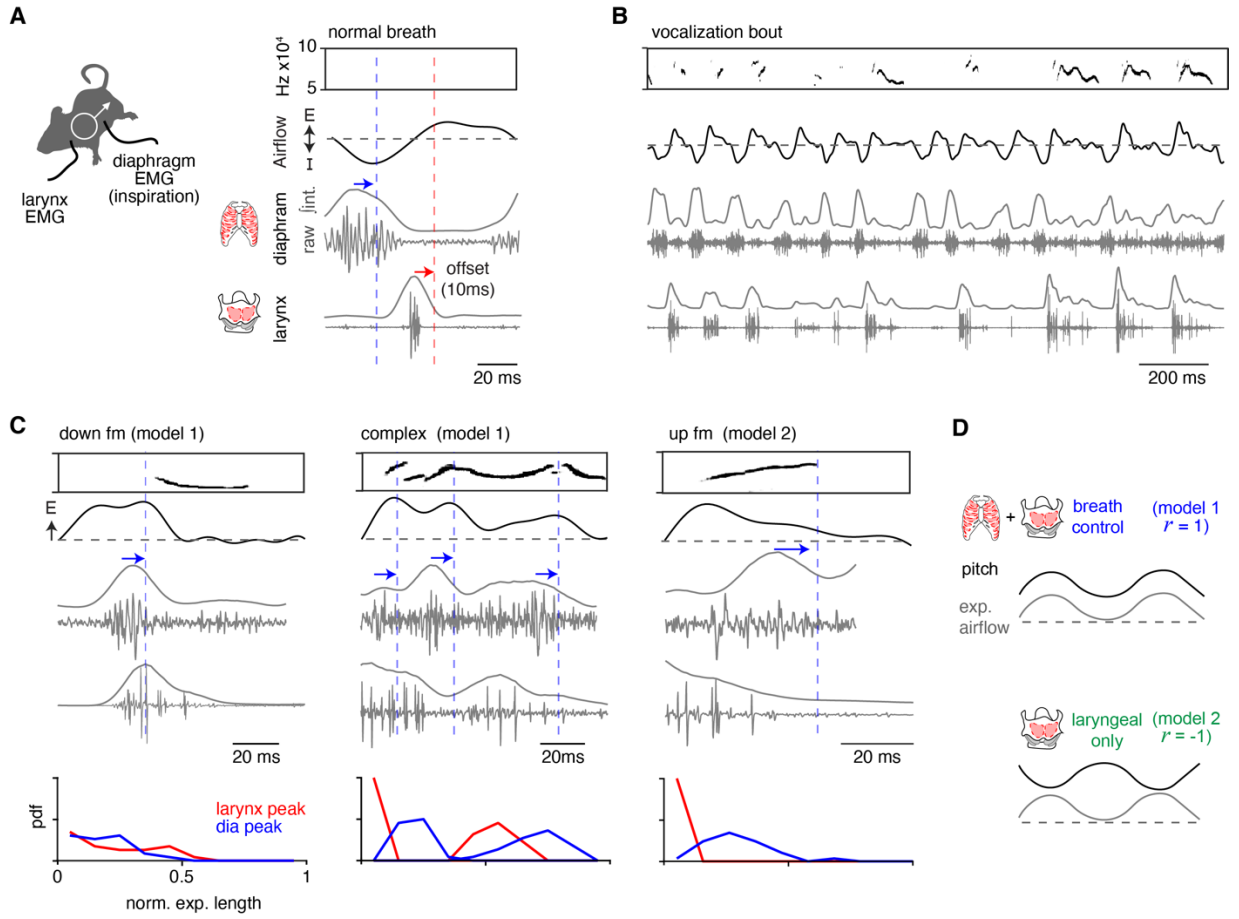

Figure 3. Ectopic activation of inspiratory and laryngeal muscles corresponds to changes in vocalization pitch. **A**, Activity of the diaphragm (inspiratory) and laryngeal (thyroarytenoid and cricothyroid) muscles were recorded *in vivo* by electromyography (EMG) simultaneously with breathing and USVs. Right, example of sound, breath airflow, and muscle activities during a basal breath. The diaphragm shows restricted activity during inspiration and laryngeal muscles are active during the post-inspiration period. Note, blue and red arrows / lines indicate the ~10ms offset of the peak EMG activity and airflow / sound measurements. **B**, Representative vocalization bout shows robust vocalizations and breathing in mice with implanted EMGs. **C**, Representative sound, airflow, diaphragm and laryngeal EMGs during the expiration of a down fm (n = 23 annotated), complex (n = 43 annotated), and up fm USV (n = 29 annotated). Blue arrow and dashed line indicate the airflow and sound offset from the diaphragm peak muscle activity. Bottom,

probability density function (PDF) for the peak of the integrated EMG activity for the diaphragm (blue) and larynx (red) during the normalized expiration. Y – axis is from 0 to 1. Note, data in A-D is from one animal with clear EMG signals that represents the findings in all three animals studied. **D**, Schematics for each model. Model 1 – breath control: inspiratory and laryngeal muscles have alternating activity throughout the sound / expiration and a $r > 0$ for pitch vs expiratory airflow. Muscle activities correspond to an increase (laryngeal) and a decrease (diaphragm) in pitch. Model 2 – laryngeal only: laryngeal but not diaphragm activity occurs during the sound and produces a $r < 0$ for pitch vs expiratory airflow.

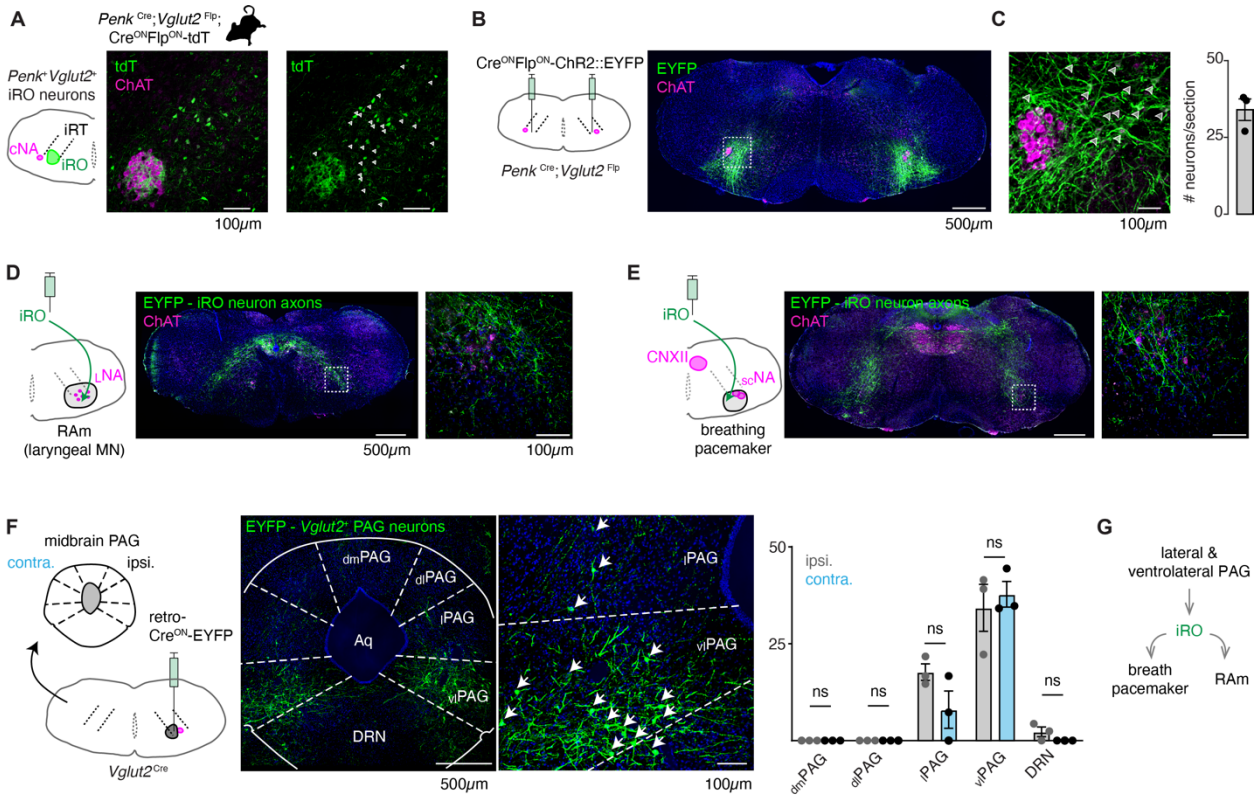

Figure 4. Anatomically and molecularly defined iRO neurons form a brainstem phonatory circuit. **A**, Labeling of *Penk*+*Vglut2*+ neurons in the iRO anatomical region in adult *Penk*-Cre;*Vglut2*-Flp;Ai65 mice (Cre^{ON}Flp^{ON}-tdTomato) (observed in n=5 mice). The iRO region is defined as medial to the compact nucleus ambiguus (cNA, ChAT +) in the ventral intermediate Reticular Tract (iRT). Note, the cNA is filled with tdTomato labeled axons. Cell bodies marked with arrowhead. **B**, Bilateral stereotaxic injection of AAV Cre^{ON}Flp^{ON}-ChR2::EYFP into the iRO anatomical region of *Penk*-Cre;*Vglut2*-Flp adult mice. **C**, Magnified boxed region in **B**. Arrowheads label neuron soma quantified right (n=3). **D**, Axons of EYFP expressing iRO neurons from **B** in the retroambiguus (RAm) anatomical region where laryngeal premotor and motor neurons are located. **E**, Axons of EYFP expressing iRO neurons from **B** in the breathing pacemaker. **F**, Unilateral retrograde AAV Cre^{ON}-EYP (AAVrg) stereotaxic injection into the iRO region in *Vglut2*-Cre adults (n=3). Glutamatergic neurons were identified in the contralateral (contra.) and

ipsilateral (ipsi.) midbrain periaqueductal gray (PAG). Anatomical regions of the PAG: dorsomedial (dm), dorsolateral (dl), lateral (l), ventrolateral (vl) nearby to the dorsal raphe nucleus (DRN) and surrounding the cerebral aqueduct (Aq). Quantification of glutamatergic PAG neurons in each region demarcated, ns = not statistically significant; two-way ANOVA with Sidka's post-hoc test. **G**, Model schematic of the iRO as a central component of the brainstem phonation circuit to convert a vocalization "go" cue from the PAG into a motor pattern.

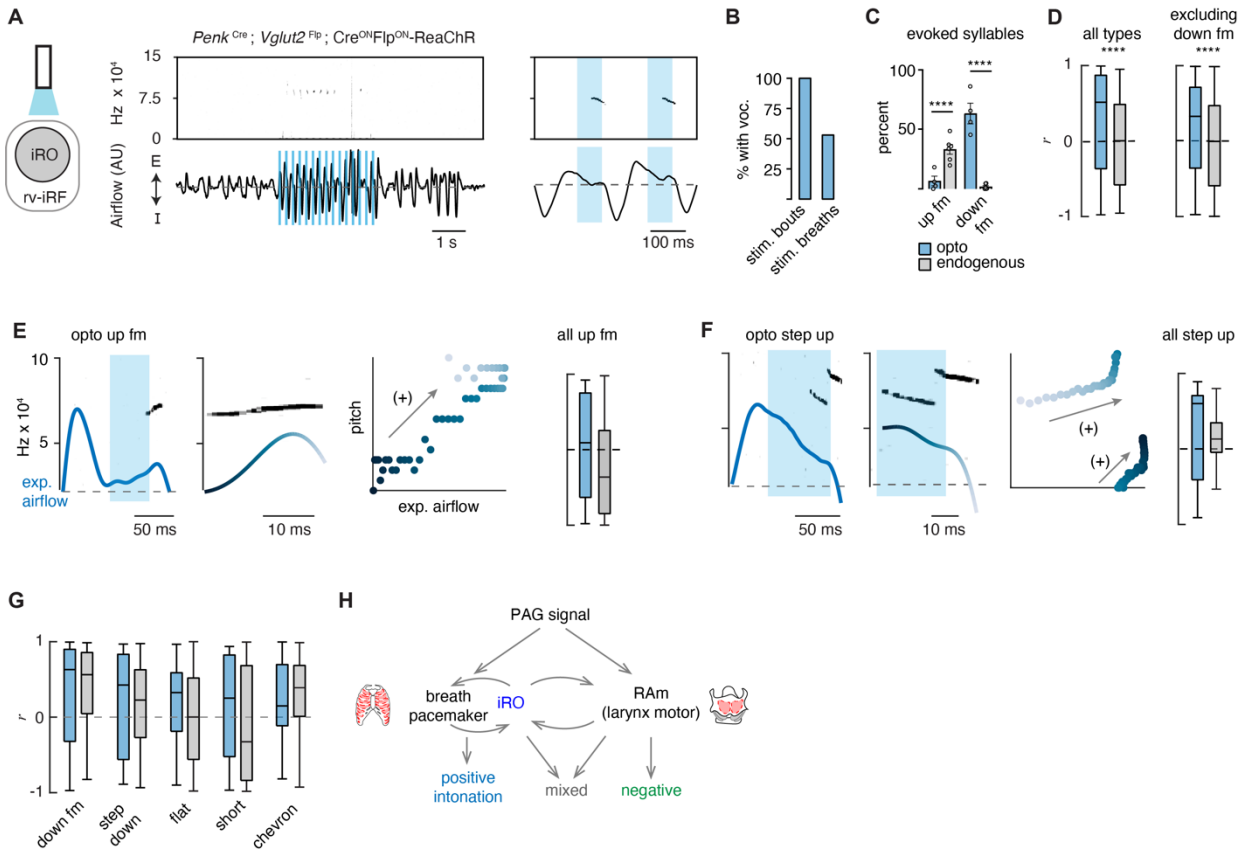

Figure 5. Ectopic activation of the iRO evokes airflow correlated USV types and switches the relationship of the anti-correlated types. **A,** Optogenetic activation of the iRO region in *Penk*-Cre;*Vglut2*-Flp;Cre^{ON}Flp^{ON}-ReaChR mice evokes USVs (blue box, 5 Hz stimulation). USVs occur during, or shortly after laser onset. **B,** Percentage of stimulation bouts containing at least one USV and the percentage of breaths within the stimulation window containing a USV. **C**, Percentage of optogenetically evoked (blue) or endogenously occurring (gray) syllables that are up fm of down fm. **** p-value< 0.001; two-way ANOVA with Sidak's post-hoc test, p > 0.05 for all other types. **D**, Box and whisker plot of the correlation coefficient between breathing airflow and pitch (r) for all opto evoked (n=395) and endogenous (n=1850) USVs and all opto evoked and endogenous vocalizations without down fm (n=143 and 1810) from n=4 opto and n=6 endogenous mice. **** p-value < 0.001; Mann-Whitney test. **E,** Left, example of the expiratory airflow and pitch for an optogenetically evoked up fm USV. Middle, magnification of airflow and sound. Time of

breath airflow during the USV is color coded blue to white. Scatter plot of instantaneous expiratory airflow and pitch for the single USV. Compare to endogenous up fm USV in Fig. 2D. Right, box and whisker plot of correlation coefficients (r) for each optogenetically evoked and endogenous up fm USV (n=15 vs. 589). **F**, Step up USV as in **E**. box and whisker plot of correlation coefficients (r) for each optogenetically evoked and endogenous step up USV (n=15 vs. 40). **G**, r value box and whisker plots for the remaining optogenetically evoked USV types (down fm n=242 vs. 40, step down n=38 vs. 293, flat n=34 vs. 337, short n=27 vs. 168, and chevron n=10 vs. 99 from n=4 opto and 6 endogenous mice. **E-G**, Two-way ANOVA with Sidak's post-hoc test for two way comparisons was used; all p-values >0.05. **H**, Schematic illustrating the two mechanisms to pattern the USV pitch. Left, the reciprocal connection between the iRO and breathing pacemaker patterns the USVs with a positive correlation between pitch and airflow. Right, the retroambiguus (RAm) control of the larynx dictates the anti-correlated USV types. Middle, the combination of these two mechanisms within a single breath create additional USV patterns.

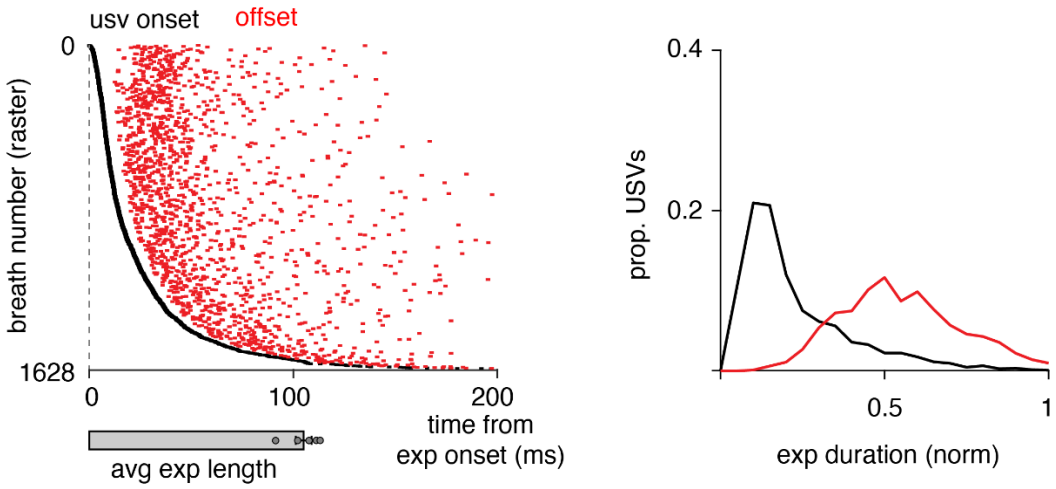

Figure Supplemental 1. USV onset and offset during expiration. Left, raster plot of USV onset and offset times (ms) aligned to the beginning of expiration (onset, black and offset, red) for 1850 events. Below, the average expiratory length for n=6 animals. Right, histogram of the onset for each vocalization during a normalized expiratory duration. Note, while onset is biased to early expiration, vocalizations can begin throughout and even in late expiration.

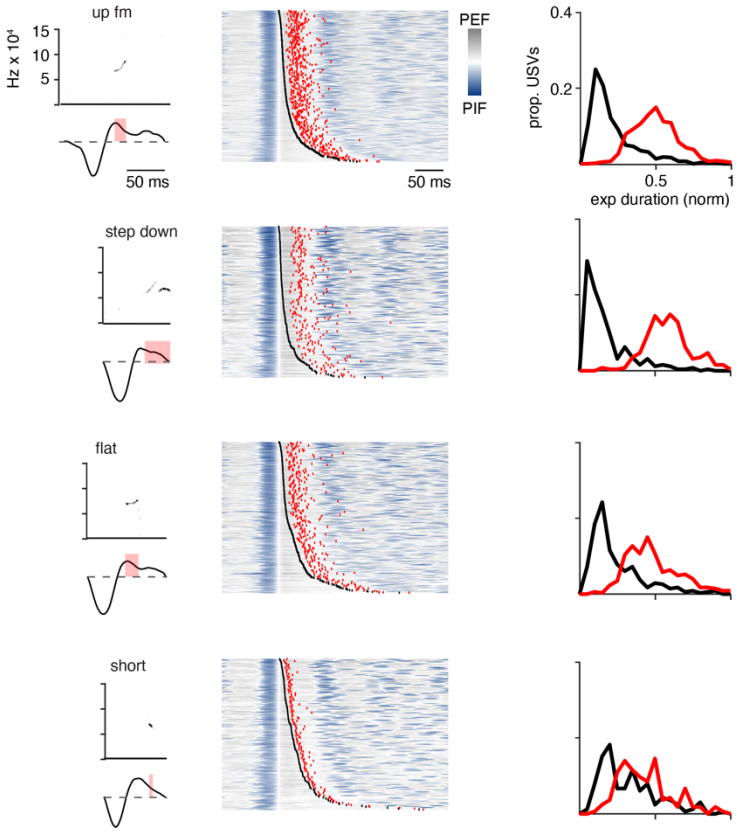

Figure Supplemental 2. Representative example of the most common USV types and the onset and offset times during expiration. Representative examples for each if the most common USV types and the representation of the onset and offset as Fig. S1.

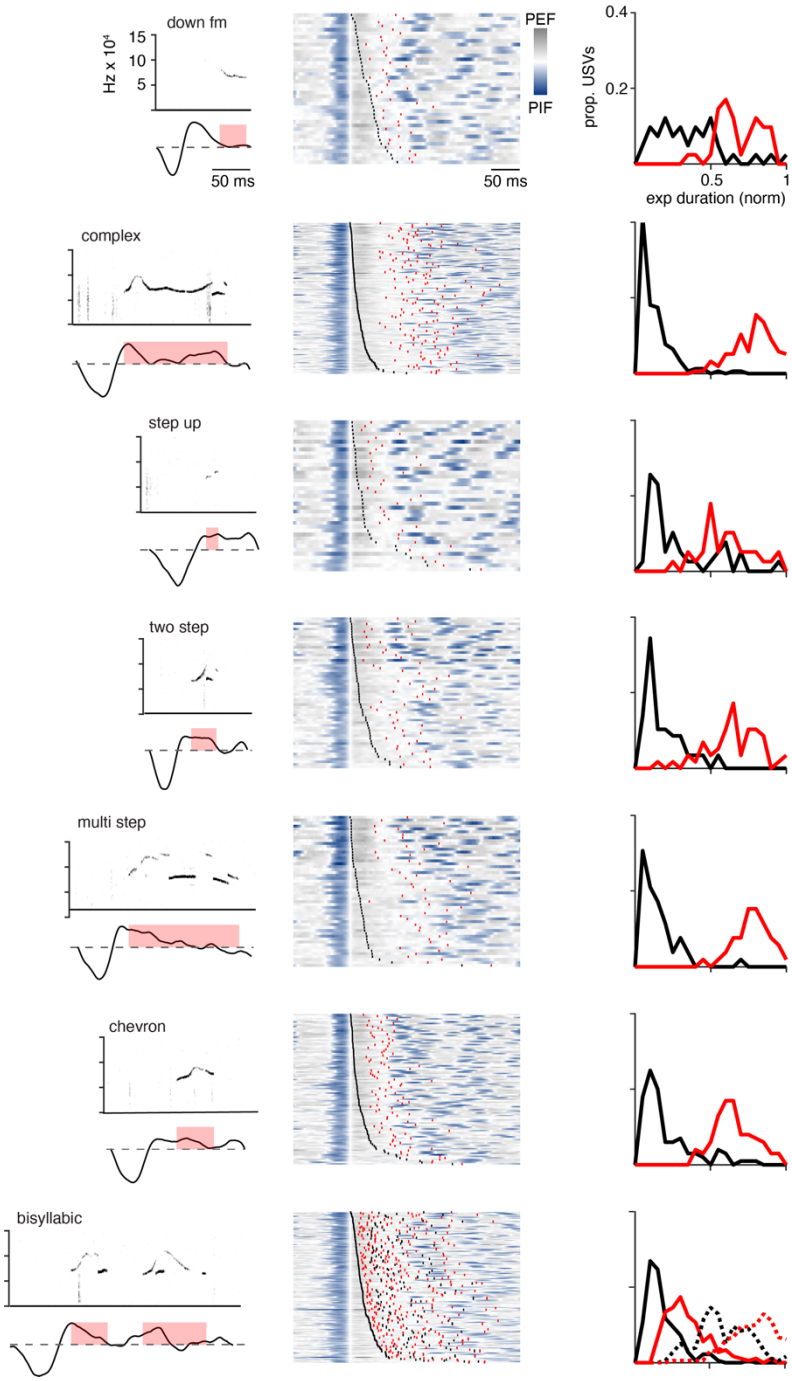

Figure Supplemental 3. Representative example of the many USV types and the onset and offset times during expiration. Representative examples for the remaining USV types and the representation of the onset and offset as Fig. S1. Note, more complex vocalizations have onset and offset times that occur later in expiration.

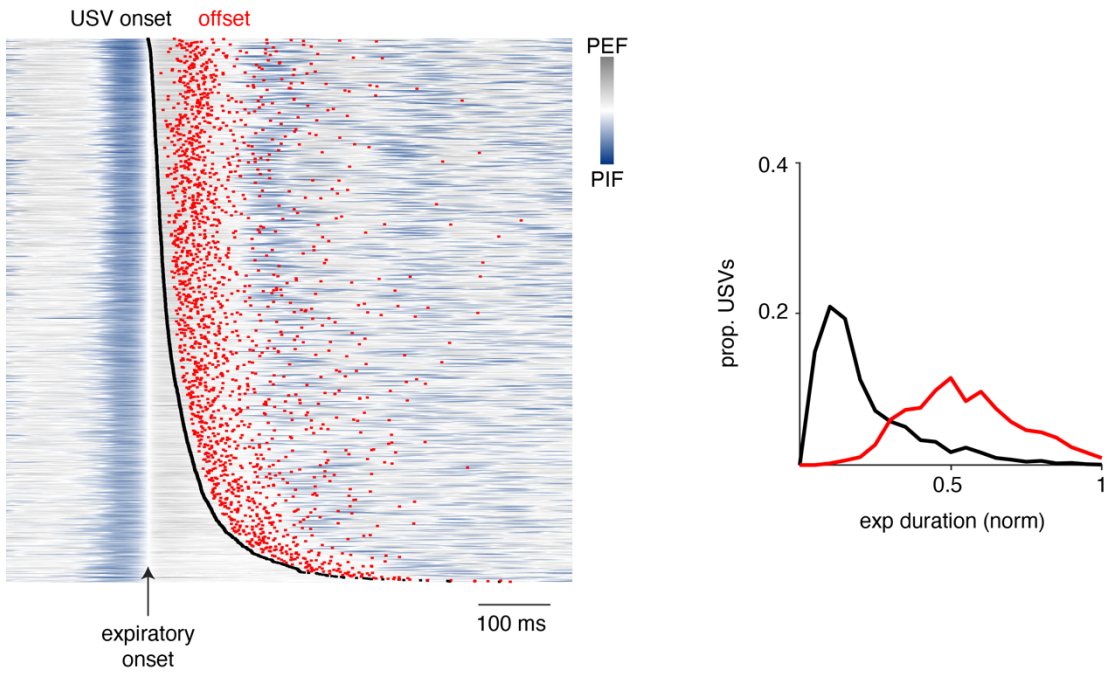

Figure Supplemental 4. Raster plot of USV on- and offset plotted upon the breathing rhythm. Raster plot of 1850 USVs aligned by the beginning of expiration with the sound onset and offset annotated by dots. The breath airflow is represented the gradient from blue to gray, where inspiration is blue, and expiration is gray. Note that breaths after ~1200 have late onset during expiration and delay the onset of the subsequent inspiration.

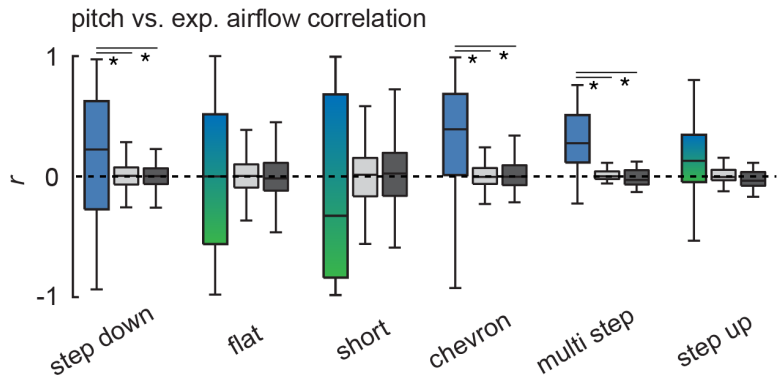

Figure Supplemental 5. Correlation coefficient and onset / offset time for six USVs. Box and whisker plot of correlation coefficients (r) for step down (n=293), flat (n=337), short (n=168), chevron (n=99), multi (n=58), and step up (n=40) USVs. * = $p < 0.05$, one-way ANOVA with Sidak's post-hoc test.

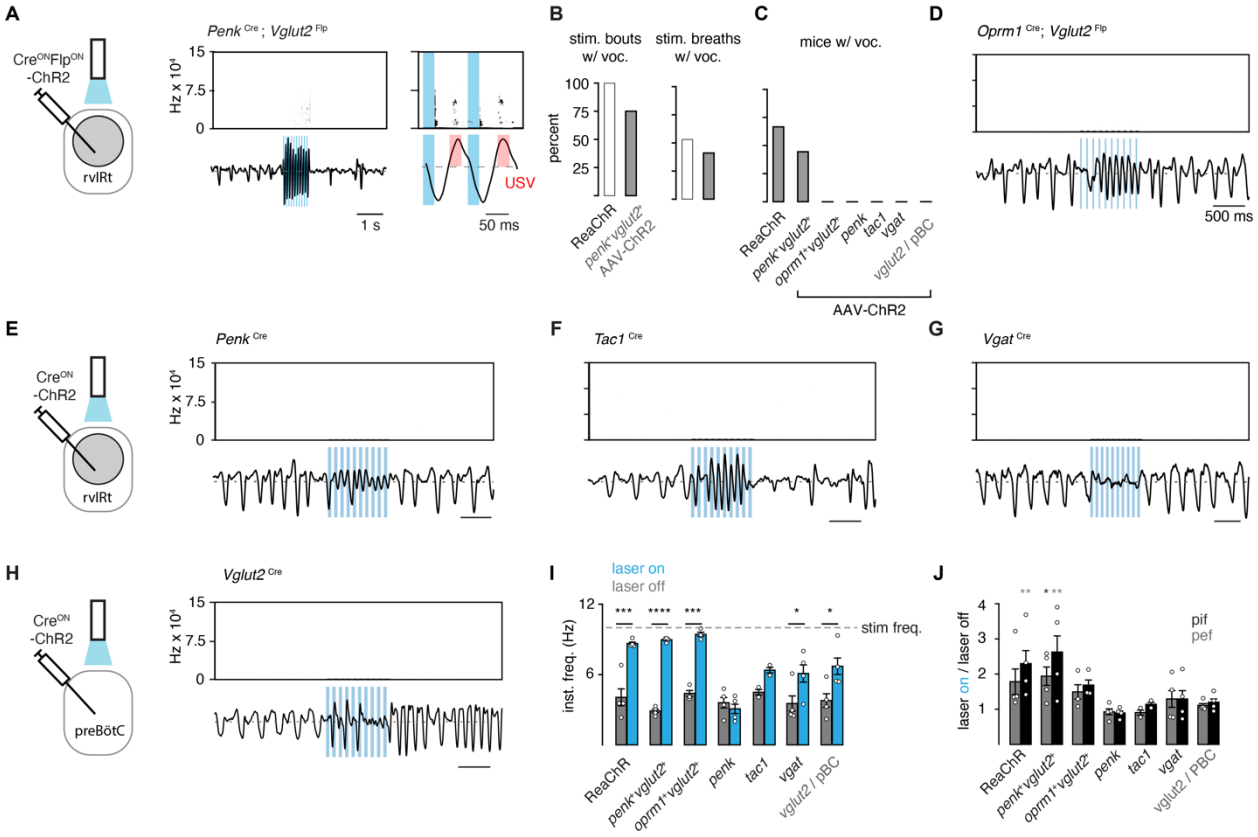

Figure Supplemental 6. Optogenetic modulation of breathing and USVs for different molecularly defined cell types in the iRO anatomical region. **A**, Representative example of the change in breathing and ultrasonic vocalizations during a single light stimulation bout (blue box, 10 Hz) in *Penk*-Cre;*Vglut2*-Flp mice stereotaxically injected with AAV Cre^{ON}Flp^{ON}-Channel Rhodopsin2::YFP (ChR2) in the iRO (gray circle). Breathing rate is entrained by light and the amplitude is increased. USVs occur at the peak of expiration. rvlRt, rostral ventral Intermediate Reticular tract. **B**, Percent of stimulation bouts and breaths within each bout that contain USVs or broad band vocalizations in *Penk*-Cre;*Vglut2*-Flp;ReaChR and *Penk*-Cre;*Vglut2*-Flp virally injected mice. **C**, Percent of mice with vocalizations for each tested genotype and injection site.

ReaChR with iRO optic fiber implantation, n=6. iRO stereotaxic viral injection: *Penk*-Cre;*Vglut2*-Flp, n=9; *Oprm1*-Cre;*Vglut2*-Flp, n=4; *Penk*-Cre, n=4; *Tac1*-Cre, n=5, *Vgat*-Cre, n=4. PreBötC stereotaxic viral injection: *Vglut2*-Cre, n=4. **D-H**, Representative examples of stimulation bouts for each genotype with rvIRt or preBötC viral injection. **I**, Bar graph of average ± standard deviation and average for each animal (circle) for the instantaneous breathing frequency before and during the optogenetic laser pulse (10 Hz). * $p<0.05$; *** $p<0.001$; **** $p<0.0001$; two-way ANOVA with Sidak's post-hoc test. Genotypes and injection sites as in **C-H**. **J**, Bar graph of average ± standard deviation and average for each animal (circle) for the ratio of the peak inspirator flow (pif, black) and peak expiratory flow (pef, gray) for optogenetically stimulated breaths versus nearby unstimulated breaths for each genotype. * $p<0.05$; ** $p<0.01$; two-way ANOVA with Sidak's post-hoc test. Genotypes and injection sites as in **C-H.**

Resource Availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Kevin Yackle (kevin.yackle@ucsf.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All reported data collected in this study will be shared by the lead author upon request.
- All original code has been deposited at Github. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Resource Table

Abbreviation	Full Name	Identifier	Supplier	Reference
Mice				
Vglut2^{FlpO}	Slc17a6-IRES2-FlpO-D knock-in	030212	The Jackson Laboratory	Daigle *et al* 2018
Penk^{Cre}	Penk-IRES2-Cre	025112	The Jackson Laboratory	Tasic *et al* 2018
LSL-FSF-ReaChR	R26 LSL FSF ReaChR-mCitrine	024846	The Jackson Laboratory	Hooks *et al* 2015
Tac1^{Cre}	Tac1-IRES2-Cre-D	021877	The Jackson Laboratory	Harris *et al* 2014
Oprm1^{Cre}	Oprm1$^{Cre:GFP}$ KIKO	035574	The Jackson Laboratory	Liu *et al* 2022
Vgat^{Cre}	Vgat-ires-cre knock-in	028862	The Jackson Laboratory	Vong *et al* 2011
Ai65	Ai65(RCFL-tdT)-D	021875	The Jackson Laboratory	Madisen *et al* 2015
Virues				
AAV-$C_{ON}F_{ON}$-ChR2-Eyfp	AAV5-hSyn-Con/Fon-hChR2(H134R)-EYFP	55645-AAV5	Addgene	Fenno *et al* 2014
AAV-DIO-ChR2	AAV5-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-HGHpA	20298-AAV5	Addgene	
AAVrg-DIO-ChR2-EYFP	AAVrg-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-HGHpA	20298-AAVrg	Addgene	
Antibodies				
Chicken anti-GFP		GFP-1020	Aves	
Goat anti-ChAT		AB144p	Millipore	
Donkey anti-Chicken 488		A78948	Invitrogen	
Donkey anti-Goat 546		A11056	Invitrogen	
Donkey anti-Goat 647		A21447	Invitrogen	
Software and algorithms				
Matlab		Matlab 2022b	Mathworks	
VocalMat		https://github.com/ahof1704/VocalMat		Fonseca *et al* 2021
USVseg		https://github.com/rtachi-lab/usvseg		Tachibana *et al* 2020
Bespoke code		https://github.com/YackleLab		This paper, Bachmutsky *et al* 2020, Wei *et al* 2022
Prism 9			GraphPad	

Experimental model and subject details

Vglut2^{FlpO}, *Penk*^{Cre}, *Tac1*^{Cre}, *Oprm1*^{Cre}, *Vgat*^{Cre}, Ai65 and LSL-FSF-ReaChR have been described. Mice were obtained from Jackson Laboratories and bred in-house at the UCSF Laboratory Animal Research Center. Mice were housed in groups of 2-5 unless otherwise stated under a 12:12 light-dark cycle with *ad libitum* access to chow and water. All animal experiments were performed in accordance with national and Institutional Animal Care and Use Committee - University of California San Francisco guidelines with standard precautions to minimize animal stress and the number of animals used in each experiment.

Recombinant viruses

All viral procedures followed the Biosafety Guidelines approved by the University of California, San Francisco (UCSF) Institutional Animal Care and Use Program (IACUC) and Institutional Biosafety Committee (IBC). The viruses used in experiments were AAV5-hSyn-Con/Fon-hChR2(H134R)-EYFP (55645-AAV5, Addgene, 1.8x1013 vg/ml), AAV5-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-HGHpA (20298-AAV5, Addgene, 1x1013 vg/ml), AAVrg-EF1a-DIO-hChR2(H134R)-EYFP-WPRE-HGHpA (20298-AAVrg, Addgene, 2.1x1013 vg/ml).

Methods details

Endogenous USV and breathing recording

Male *Vglut2*^{FlpO};*Penk*^{Cre} mice (aged 8-16 weeks) were individually housed and habituated to experimenter handling and a plethysmography chamber for >4 days. On the test day, the mice were placed in a clean cage base with a female mouse for 5 minutes and then moved to a plethysmography chamber. The chamber was modified to accommodate a microphone to record vocalizations (CM16/CMPA, Avisoft Bioacoustics) and the airflow in the chamber was measured by a spirometer (FE141, AD instruments). Both data streams were acquired through a DAQ board (PCI-6251, National Instruments) and written to disk for offline analysis. Sound was acquired at

400 kHz and airflow at 1 kHz. After a 20-minute habituation period, mice had airflow and sound recorded for 5 minutes before a cotton bud soaked in fresh urine was placed in the chamber, and sound and breathing were recorded for a further 15 minutes. Urine was collected the day of the experiment from a group of 5 female mice temporarily housed in a custom-made wire-bottom cage.

The recordings were run through VocalMat (Fonseca et al., 2021) for USV detection and only mice that produced >50 USVs in response to the stimulus were included for further analysis (5/13 mice). Airflow recordings were imported to MATLAB, high pass filtered (2Hz) and smoothed. Breaths were taken from the first 200s following urine presentation and features (Ti, Te, Pif, Pef, instantaneous frequency) were computed from segmented breaths as previously described (Bachmutsky et al., 2021). USV start and end times from VocalMat were used to identify which breaths contained USVs and calculate timing metrics (relative onset and offset from expiration onset and the same values normalized to expiratory duration). VocalMat was also used to identify the types of USV which were manually checked and corrected if necessary. For analysis of the relationship between airflow and frequency, a multitaper spectrogram was computed using code modified from USVseg (Tachibana et al., 2020) and then the frequency bin with the greatest power was taken from each time bin to create a vector of the peak frequency. The correlation coefficient of this peak frequency vector and the expiratory airflow at the time stamps identified by VocalMat was then calculated for each identified USV.

For electromyography (EMG) recordings of the diaphragm and larynx we modified the protocol of Hérent et al., 2020. Electrodes were prepared from stainless steel (diaphragm, ground; AM systems #793200) and tungsten (larynx; AM systems #795500) wires and soldered to a 5-pin connector. An incision was made in the skin of the scalp and the fascia overlaying the skull was cleared. The twisted diaphragm electrodes were tunneled under the skin and inserted through the posterior ribs and sternum to cover the diaphragm as described in Hérent et al., 2020. To insert the larynx electrodes an incision was made in the neck and layers of muscle parted to expose the

thyroid cartilage. Wires were then tunneled under the skin to the anterior incision, were inserted through the thyroid cartilage and secured with superglue. A ground wire was inserted under the skin of the neck. The connector was then secured to the skull with dental cement, all incisions were closed with suture and mice were transferred to a heated recovery chamber.

After several days of recovery, mice were placed in the plethysmography chamber and the EMG pins were connected to an amplifier (AM Systems 1800). Audio and airflow signals were acquired as described above. EMG signals were bandpass filtered (300 Hz – 20 kHz) and acquired at 10 kHz. Diaphragm EMG recordings that included an ECG signal were manually annotated and the artefact removed from the trace. EMG traces were rectified and integrated offline with a modified Paynter filter in MATLAB. Three types of USVs (Up fm, down fm, complex) were manually inspected and EMG peaks were annotated to generate histograms (figure 3).

Virus injection, fiber implantation and optogenetics

Surgery was conducted with sterile tools and aseptic technique. Mice were first anesthetized with isoflurane (4%), the hair overlaying the scalp was shaved and mice were placed in the stereotaxic frame where isoflurane (0.9-1.5%) was continuously delivered for the duration of the surgery. Mice were then injected with buprenorphine (0.1 mg/kg, s.c.) and carprofen (5 mg/kg, s.c.) and bupivicane (0.25mg, under the skin of the scalp). The skin was then covered with betadine before an incision was made with a scalpel. The fascia was removed, and the skull dried with ethanol. The bregma and lambda sutures were identified and the skull was leveled using these landmarks. A craniotomy was drilled at the injection coordinates and a pulled glass pipette lowered to the injection site. An injection was made at a speed of 100 nl/min from an injection system (Nanoject III, Drummond). The injection pipette was left in place for 10 minutes following the injection then slowly retracted from the brain. In the case of bilateral injections, this process was then repeated on the other side. The skin was then closed by suture and the mouse transferred to a heated recovery cage.

For optogenetic experiments the virus injections were performed as described above. Once the injection pipette was removed the skull was scored with a scalpel blade and a fiber implant composed of a ferrule (CFLC230, Thorlabs) and an optic fiber (FT200EMT, Thorlabs) held in place with epoxy (F112, Thorlabs) inserted into the brain 200 μm dorsal to the injection site. The first fiber was glued in place while the second fiber was inserted. Once both fibers were in place, the skull was covered with dental cement (C&B Metabond) then a second layer of acrylic (Jet). After the skull cap was dried mice were transferred to a heated recovery cage. Coordinates (in mm) were as follows; iRO: 6.35 posterior to bregma, 5.4 ventral to skull surface, 1.2 lateral to midline; pBC: 6.73 posterior to bregma, 5.77 ventral to skull surface,1.3 lateral to midline. To maximally excite the iRO system we did bilateral implants.

Mice were given 6 weeks between injection/implantation surgery and being used for experiments. ReaChR mice were implanted as described above. For optogenetic experiments bilateral fibers were connected to a split-patch cord (SBP(2)_200/220/900-0.37_m_FCM-2xZF1.25, Doric Lenses) and light was delivered from a laser (MBL-III-473, Opto Engine LLC) controlled by a TTL pulse generator (OTPG_4, Doric Lenses). Mice were placed in the plethysmography chamber with the microphone attached to simultaneously record breathing and sound along with the laser pulse commands. All three data streams were acquired through a DAQ board and written to disk for offline analysis. Sound was acquired at 250 kHz, airflow at 1 kHz and laser pulse commands at 1 kHz. After a 20-minute habituation period, laser pulses were delivered at frequencies of 5, 10, 20, and 50 Hz with pulse widths of 10, 25 or 50 ms for durations of 1 or 3 seconds. Laser power was adjusted to deliver ~20 mW of light at the patch cord tip although attenuation of light by the implanted fiber (determined post-hoc) was variable (12-21 mW). Each frequency/pulse width/duration combination was delivered 5 times with 7-9 seconds between presentations and a 30s delay before the next stimulus was delivered.

Recordings were manually inspected for USVs during the laser epoch and recordings containing USVs were then run through VocalMat to find time stamps and to categorize each USV

by type. Matlab code was then used to quantify the correlation coefficients of optogenetically evoked USVs and the underlying airflow as described above. To analyze the breath statistics of optogenetically evoked breathing, the trial with stimulation parameters: 10 Hz, 25ms pulse width, 3s duration was run through a code to extract breath statistics (Pif, Pef, Instantaneous Frequency) from the 30s period prior to stimulation and from the 5 laser epochs.

Histology

More than 6 weeks following viral injection or the completion of optogenetic testing mice were deeply anaesthetized with isofluorane and transcardially perfused with 0.1M phosphate buffered saline (PBS) then PBS containing 4% paraformaldehyde (PFA). Brains were dissected from the fixed mice and refrigerated in 4% PFA overnight then cryoprotected in 30% sucrose in PBS. Brains were sectioned to 30 μm coronal on a freezing microtome. Sections were washed 3 times for 5 min in PBS before being incubated in blocking solution (PBS, 5% normal donkey serum, 0.3% Triton-X100) for 2 hours. Sections were then incubated overnight in primary antibodies (Chicken anti-GFP, 1:1000, Aves; Goat anti-ChAT, 1:500, Millipore) diluted in a carrier solution (PBS, 1% normal donkey serum, 0.3% Triton-X100). Following incubation sections were washed with PBS 5 times for 5 minutes then incubated in secondary antibodies (Donkey anti-Chicken 488, Donkey anti-Goat 546, Donkey anti-Goat 647) diluted 1:500 in a carrier solution (PBS, 0.3% Triton-X100) for 2 hours at room temperature. After secondary incubation, sections were washed with PBS 5 times for 5 minutes then mounted onto glass slides and cover-slipped with mounting media (Prolong Gold, Invitrogen) and 1 μg/ml DAPI.

Quantification and statistical analysis

Statistics

Data from Matlab was imported to Prism 9 (GraphPad) for statistical analysis. For all statistical analysis except Figure 2 and Figure 5 D-G the mouse was used as the experimental unit. Data were assumed to be normally distributed and of equal variance and parametric tests were used. For data with one discrete variable and measurements made from the same animal (Fig. 1 C-E) paired t-test was used. For data with two variables one or both of which had more than two factors (Fig. 4G, 5C,S2I,J) two-way ANOVA was used with Sidak's post-hoc test for multiple comparisons. To compare pitch-airflow correlations between observed and shuffled datasets (Fig. 2, S5), each USV was considered the experimental unit. For each USV type, we compared the observed values to two null distributions generated by shuffling one of the variables with one-way ANOVA and Sidak's post-hoc test. To compare pitch-airflow correlations of endogenous and optically-evoked USVs (Fig. 5 D-G) each USV was treated as the experimental unit since the vocal repertoire across animals was similar (Fig. 1F) and simply taking a mean from each animal would under-represent the complexity of the data. For the comparison of correlation coefficients between optically evoked and endogenous USVs, two-way ANOVA with Sidak's post-hoc test for two-way comparisons was used. P-values below 0.05 were considered statistically significant.

Bachmutsky, I., Wei, X.P., Kish, E., Yackle, K. (2020) Opioids depress breathing through two small brainstem sites. ELife 9:e52694.

Berke, G.S., and Long, J.L. (2009). Functions of the larynx and production of sounds. Handbook of mammalian vocalization – an integrative neuroscience approach, Chapter 10.1.

Castellucci, G.A., Calbick, D., and McCormick, D. (2018) The temporal organization of mouse ultrasonic vocalizations. PloS One 13(10):e0199929.

Chagnaud, B.P., Baker, R., and Bass, A.H. (2011). Vocalization frequency and duration are coded in separate hindbrain nuclei. Nature communications *2*, 346–11.

Chen, J., Markowitz, J.E., Lilascharoen, V., Taylor, S., Sheurpukdi, P., Keller, J.A., Jensen, J.R., Lim, B.K., Datta, S.R., and Stowers, L. (2021). Flexible scaling and persistence of social vocal communication. Nature *593*, 108–113.

Daigle, T.L. *et al* (2018) A suite of transgenic driver and reporter mouse lines with enhanced brain-cell-type targeting and functionality. Cell *174*, 465-480

Dichter, B.K., Breshears, J.D., Leonard, M.K., and Chang, E.F. (2018). The control of vocal pitch in human laryngeal motor cortex. Cell *174*, 21-31.

Finck, C., and Lejeune, L. (2009). Structure and oscillatory function of the vocal folds. Handbook of mammalian vocalization – an integrative neuroscience approach, Chapter 10.2.

Fenno, L.E. (2014). Targeting cells with single vectors using multiple-feature boolean logic. Nature methods 11(7), 763-772.

Fonseca, A.H., Santana, G.M., Bosque Ortiz, G.M., Bampi, S., and Dietrich, M. (2021) Analysis of ultrasonic vocalizations from mice using computer vision and machine learning. Elife 10:e59161.

Goller, F., and Cooper, B.G. (2004). Peripheral motor dynamics of song production in the zebra finch. Ann. N.Y. Acad. Sci. *1016*, 130–152.

Grimsley, J.M.S., Monaghan, J.J.M., and Wenstrup, J.J. (2011). Development of social vocalizations in mice. Plos One *6*, e17460.

Hage, S.R. (2009a). Neuronal networks involved in the generation of vocalization. Handbook of behavioral neuroscience *19*, 339-349.

Hage, S.R. (2009b). Localization of the central pattern generator for vocalization. Handbook of behavioral neuroscience *19*, 329-337.

Harris, J.A. *et al* (2014). Anatomical characterization of cre driver mice for neural circuit mapping and manipulation. Frontiers in neural circuits 8, 76.

Herbst, C.T. (2016). Vertebrate sound production and acoustic communication. Springer handbook of auditory research, 159–189.

Hooks, B.M., Lin, J.Y., Guo, C., Svoboda, K. (2015). Dual-channel circuit mapping reveals sensorimotor convergence in the primary motor cortex. Journal of neuroscience *35(10)*, 4418-4426.

Howie J. M. (1976) Acoustical studies of mandarin vowels and tones. Cambridge University Press Vol 18.

Johnson, A.M., Ciucci, M.R., Russell, J.A., Hammer, M.J., and Connor, N.P. (2010). Ultrasonic output from the excised rat larynx. The journal of the acoustical society of america *128*, EL75–EL79.

Jürgens, U. (2002). Neural pathways underlying vocal control. Neuroscience and biobehavioral reviews *26*, 235–258.

Kelley, D.B., Ballagh, I.H., Barkan, C.L., Bendesky, A., Elliott, T.M., Evans, B.J., Hall, I.C., Kwon, Y.M., Kwong-Brown, U., Leininger, E.C., et al. (2020). Generation, coordination, and evolution of neural circuits for vocal communication. Journal of neuroscience *40*, 22–36.

Kelm-Nelson, C.A., Lenell, C., Johnson, A.M., and Ciucci, M.R. (2018). Laryngeal activity for production of ultrasonic vocalizations in rats. Handbook of mammalian vocalization – an integrative neuroscience approach, Chapter 4.

Laplagne, D.A. (2018). Interplay between mammalian ultrasonic vocalizations and respiration. Neuronal networks involved in the generation of vocalization, Chapter 6.

Liu, S., Ye, M., Pao, G.M., Song, S.M., Jhang, J., Jiang, H., Kim, J-H., Kang, S.J., Kim, D-I., Han, S. (2022). Divergent brainstem opioidergic pathways that coordinate breathing with pain and emotions. Neuron *110(5),* 857-873.

Mahrt, E., Agarwal, A., Perkel, D., Portfors, C., and Elemans, C.P.H. (2016). Mice produce ultrasonic vocalizations by intra-laryngeal planar impinging jets. Current biology *26*, 1–2.

Madisen, L., *et al.* (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron 85(5), 942-958.

Marder, E., and Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. Current biology *11*, R986–R996.

Marder, E. (2012). Neuromodulation of neuronal circuits: back to the future. Neuron *76*, 1–11.

Michael, V., Goffinet, J., Pearson, J., Wang, F., Tschida, K., and Mooney, R. (2020). Circuit and synaptic organization of forebrain-to-midbrain pathways that promote and suppress vocalization. Elife 9:e63493.

Park, J., Choi, S., Takatoh, J., Zhao, S., Harrahill, A., Han, B.X., Wang, F. (2024). Brainstem control of vocalization and its coordination with respiration. Science 383 (6687).

Plummer, E.M., and Goller, F. (2008). Singing with reduced air sac volume causes uniform decrease in airflow and sound amplitude in the zebra finch. Journal of experimental biology. *211*, 66–78.

Prieto, P. (2015). Intonational meaning. Wiley interdiscip. rev. cognitive sci. *6*, 371–381.

Riede, T. (2011). Subglottal pressure, tracheal airflow, and intrinsic laryngeal muscle activity during rat ultrasound vocalization. Journal of neurophysiology *106*, 2580–2592.

Riede, T. (2013). Stereotypic laryngeal and respiratory motor patterns generate different call types in rat ultrasound vocalization. Journal of experimental zoology part A: ecological genetics and physiology *319*, 213–224.

Riede, T., Borgard, H.L., and Pasch, B. (2017). Laryngeal airway reconstruction indicates that rodent ultrasonic vocalizations are produced by an edge-tone mechanism. Royal society open science *4*, 170976.

Schmidt, M.F., and Wild, J.M. (2014). Chapter 15 The respiratory-vocal system of songbirds Anatomy, physiology, and neural control. Prog. brain res. *212*, 297–335.

Silva, A.B., Liu, J.R., Zhao, L., Levy, D.F., Scott, T.L., and Chang, E.F. (2022). A neurosurgical functional dissection of the middle precentral gyrus during speech production. Journal of Neuroscience *42*, 8416–8426.

Sirotin, Y.B., Costa, M.E., Laplagne, D.A. (2014). Rodent ultrasonic vocalizations are bound to active sniffing behavior. Frontiers in behavior neuroscience. *8*:399.

Smith, J.C., Ellenberger, H.H., Ballanyi, K., Richter, D.W., and Feldman, J.L. (1991). Pre-Bötzinger complex: a brainstem region that may generate respiratory rhythm in mammals. Science *254*, 726–729.

Suthers, R.A., Goller, F., and Wild, J.M. (2002). Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. PNAS *99*, 5680–5685.

Tachibana, R.O., Kanno, K., Okabe, S., Kobayasi, K.I., Okanoya, K. (2020) USVSEG: A robust method for segmentation of ultrasonic vocalizations in rodents. PLoS One

Tasic B. *et al* (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72-78.

Tschida, K., Michael, V., Takato, J., Han, B.-X., Zhao, S., Sakurai, K., Mooney, R., and Wang, F. (2019). A specialized neural circuit gates social vocalizations in the mouse. Neuron, *103*, 459-472.

Vong, L., Ye, C., Yang, Z., Choi, B., Chua Jr, S., Lowell, B.B. (2011). Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. Neuron *71*, 142-54.

Wei, X.P., Collie, M., Dempsey, B., Fortin, G., and Yackle, K. (2022). A novel reticular node in the brainstem synchronizes neonatal mouse crying with breathing. Neuron *110*, 644-657.

Yackle, K. (2023). Transformation of our understanding of breathing control by molecular tools. Annual reviews of physiology *85*, 93-113.

Zhang, Y.S., and Ghazanfar, A.A. (2020). A Hierarchy of autonomous systems for vocal production. Trends in neurosciences *43*, 1–12.