Mice produce ultrasonic vocalizations by intra-laryngeal planar impinging jets

Elena Mahrt, Anurag Agarwal, David Perkel, Christine Portfors, Coen P.H. Elemans

Supplemental Figures

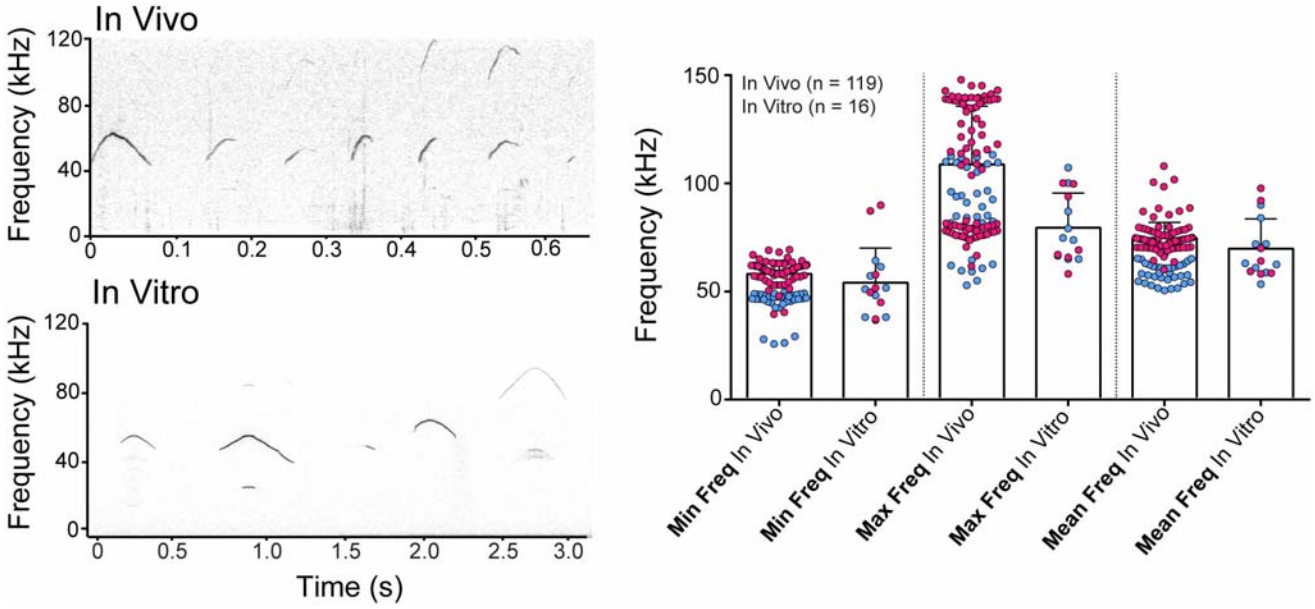

Figure S1. USVs and UVs produced by the excised larynx *in vitro* correspond well to *in vivo* USVs.

A) Exemplary spectrograms of *in vivo* (top) and *in vitro* USVs (bottom) (Power spectral density spectrogram; FFT length: 1024, overlap: 75%, Hamming window, dynamic range: 50 dB in units^2/Hz)

B) Minimum and mean peak frequency of *in vivo* USVs and *in vitro* USVs were not significantly different (Student's T-Test; P > 0.01), while max frequency was greater *in vivo* (Student's T-Test; P< 0.01). Blue colored circles represent sound repetitions from one animal and pink from another animal from *in vivo* and *in vitro* conditions.

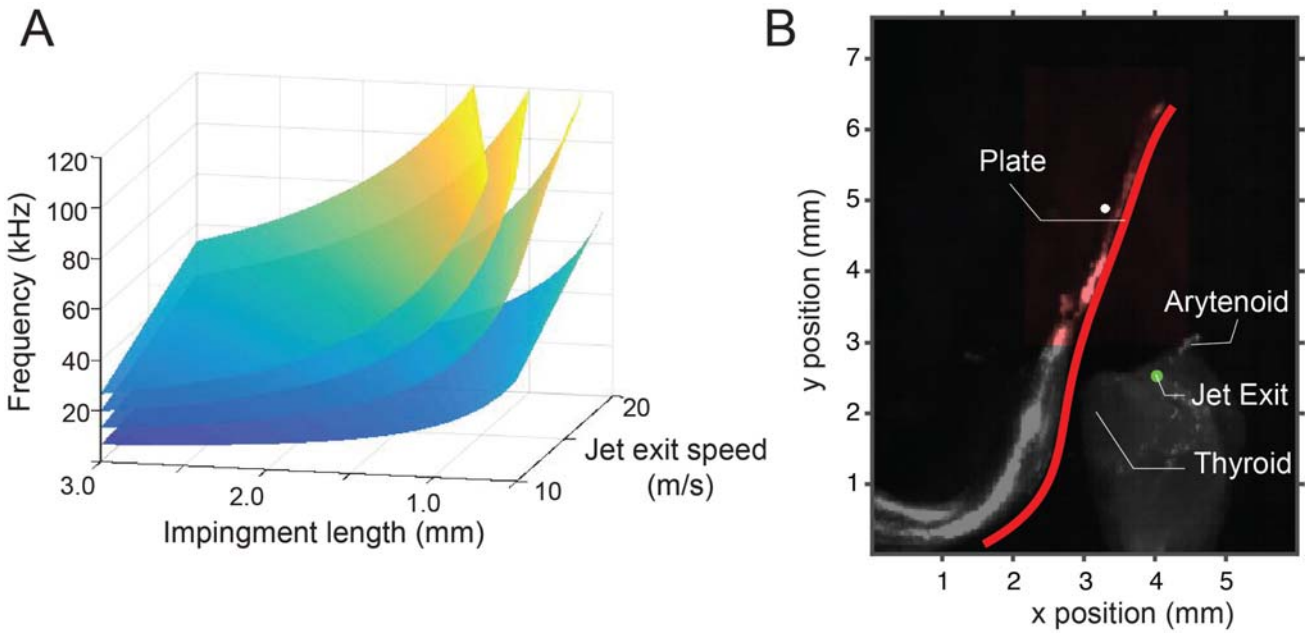

Figure S2. Testing frequency mode predictions of the planar wall impinging jet model. A) Surface plot of predicted whistle frequencies as function of jet flow and impingement length. Only modes n=2, 4, 6, and 8 are shown for clarity. B) Side view (movie still) of thyroid replacement experiment showing the extracted boundary of plate (red line) that was positioned by cross-correlating the plate region of interest (red shading) with the entire image. The white dot indicates the subpixel position of the center of the plate sub-image used for cross-correlation. The green dot is the estimated exit position of the glottal jet. The upper part of the thyroid is cut away. Pixel resolution in the original image was 25.4 µm.

Supplemental Experimental procedures

Subjects

Because *in vivo* recordings have shown that female mice emit ultrasonic vocalizations with similar acoustic features as male mice [S1], we used both males and females in our experiments. For *in vitro* experiments, 15 out of 17 larynges tested made ultrasonic sound. Of these 17 larynges tested, 10 were adult BALB (male n = 4, female n = 6), 5 were adult C57 (male n = 4, female n = 1), and 2 were adult NMRI (female n = 2) mice from Taconic Europe, Denmark. Detailed video and acoustic analysis was conducted on 8 adult BALB (male n = 4, female n = 4) and 1 adult NMRI (male n = 1) mice. Of these larynges, *in vivo* recordings were recorded from 2 BALB males. All experiments were conducted at the University of Southern Denmark in accordance with the Danish Animal Experiments Inspectorate (Copenhagen, Denmark).

Excised larynx preparation

Mice were euthanized with isoflurane and the larynx, trachea, lower jaw and tongue were dissected, flash frozen in liquid nitrogen, and then maintained at -80°C until just prior to the experiment. At the start of each experiment, tissue was thawed in refrigerated ringer's solution [S2] and the lower jaw, tongue, esophagus, hyoid, fat, and oropharynx were removed. The exterior, ventral, and dorsal views of each larynx were photographed with a Leica DC425 mounted on a stereomicroscope (M165-FC, Leica Microsystems). The trachea and attached larynx were fixed on a blunted 19G needle using 10-0 monofilament suture. Care was taken that the end of the needle was greater than 0.5 mm from the caudal end of cricothyroid muscles. The arytenoids were held in place by two micromanipulators, but also without these USVs were readily elicited and all recordings were done with the larynx in resting

position, i.e. without additional adduction of the arytenoids. Tissue was kept moist during the experiment with Ringer's solution.

Once secured onto the needle, the trachea and larynx were set upright onto a custom made apparatus with computer controlled air delivery system. Air driven through the larynx was pressure controlled, humidified, and temperature monitored (temperatures ranged from 22 – 30°C) as described in detail in [S3]. We subjected the larynx to an increasing and decreasing pressure ramp from 0-2 kPa in 4 s. Tracheal mass flow was measured with MEMS flow sensors (PMF series, Posifa Microsystems, San Jose, USA) and a response time of 1 ms. To our knowledge no calibrated subglottal pressure or tracheal mass flow measurements during *in vivo* USV production are available for mice, but we consider the applied 0-2 kPa range physiologically relevant, because (1) in rats subglottal pressure ranges from 0-2 kPa during *in vivo* USV production [S4,S5], (2) in similarly sized birds subsyringeal pressures during song and induced sound production range from 0-3 kPa [S3,S6-S8], and (3) increasing subglottal pressure above 3 kPa *in vitro* lead to a substantial increase of acoustic noise suggesting that the upper limit for stable mode production was reached. The applied subglottal pressure ramps resulted in tracheal mass flow of 0-4 ml/s (Fig 1C). During quiet respiration in rats, tracheal mass flow is 15-20 ml/s/kg [S9]. Extrapolating these values to a mouse of 20 gr would suggest a tracheal mass flow during quiet respiration of 0.3-0.4 ml/s, which is below, and thus consistent with, the observed phonation threshold flow of 1 ml/s in Fig 1C.

Sound was recorded with a 1/4-inch pressure microphone-pre-amplifier assembly (model 46BD, G.R.A.S., Denmark), amplified and high-pass filtered (10 Hz, 3-pole Butterworth filter, model 12AQ, G.R.A.S., Denmark). The microphone sensitivity was measured before each experiment (sound calibrator model 42AB, G.R.A.S., Denmark). The microphone was placed at 22-24 mm away from the mounted larynx in the acoustic near field, and on a 90° angle to avoid the air jet from the tracheal outlet.

Microphone, pressure and flow signals were low-pass filtered at 100, 10 and 10 kHz, respectively (custom-built filter) and digitized at 250 kHz (USB 6259, 16 bit, National Instruments, Austin, Texas). All control and analysis software was written in Matlab.

The laryngeal glottal opening was imaged with a light-sensitive 16 bit high-speed camera (Fastcam SA1, Photron, San Diego, CA, USA; 100,000 frames/s) mounted on a stereomicroscope (M165-FC, Leica Microsystems) and illuminated by a plasma light source (HPLS200, Thorlabs, Germany) through liquid light guides. The glottis consisted of the cartilaginous glottis between the arytenoids, and never reached into the membranous vocal fold portion. Due to the large amounts of data collected when imaging at 100,000 fr/sec we typically recorded 1-2 seconds.

We developed an algorithm for automated parameter extraction of projected glottal area and width per frame. After finding the glottal shape in each frame by Canny edge detection, the midline of the glottis was defined as the major axes orientation of a fitted ellipsoid to the extracted shape. Glottal width (i.e. arytenoid abduction times two) was calculated as the maximum of the perpendicular distance between the midline and all points along the glottal shape. Due to the downward ~45° angle of the glottis with the stereomicroscope's objective (see Fig 1A), the actual glottal area was about [1/cos(45°)=] 1.4 larger than the projected glottal area, and ranged from 0.1 to 0.45 mm^2 between preparations, resulting in jet speeds around 10-15 m/s. Glottal width was parallel to the objective and thus was correctly projected, and varied little within - but more between - preparations ranging from 36 to 80 pixels (0.33 to 0.73 mm). With Canny edge detection we could in principle not detect vibration below a single pixel of 9.2 μm. To determine if we could detect smaller vibrations masked by noise, we also inspected high dynamic range spectrograms of the glottal width signal. Careful visual inspection of our raw image data, played-back at various relevant speeds, and the spectrograms of glottal width did not reveal any motion at the frequencies of the USVs. Peak frequency in the glottal width signal was

finally determined by calculating power spectral density estimates of the glottal width of 500 consecutive images using the periodogram method (zero padded to 512 points). The peak frequency always occurred at 0 Hz (Fig 1C).

We removed the epiglottis with a horizontal cut using micro-scissors. We subsequently cut slices of the cranial part of the thyroid until just above the vocal folds.

Wall impinging jet model

We derived our model from [S10]. Resonance occurs when the feedback is in phase with the instability waves in the jet. The phase of the feedback disturbance at frequency *f* is given by $2\pi \cdot \left(f\frac{x}{u} + f\frac{x}{c}\right)$, where *x* is the distance between the jet exit (glottis) and the planar wall, i.e. the impingement length, *u* is the mean convection speed of the coherent structures, approximated as jet exit speed, and c is sound speed. This feedback phase will be in phase with the initial disturbance if equal to $2\pi \cdot n$, where *n* is an integer, and thus follows: $n = f_n\left(\frac{x}{c} + \frac{x}{u}\right)$. Because $u/c \ll 1$, the whistling frequency is given by: $f_n = n * u/x$.

Jumps between modes thus need not be harmonically related. If the n=1 mode is observed the steps would be harmonics. In Figure 1F for example, n=1 is not excited, and thus the whistling frequencies are not harmonically related.

Our model predicts that the resonance frequencies are independent of the speed of sound if the jet speed remains constant. However, generating a whistle with the same pressure difference (from sub- to supraglottal) in a gas with lower density ρ results in a higher jet speed *u* that is inversely proportional to the square root of air density ($u \propto 1/\sqrt{\rho}$). Thus, from our model ($f_n = n * u/x$) follows that the predicted frequency of the whistle is inversely proportional to the square root of density ($f_n \propto 1/\sqrt{\rho}$). However, a lower density gas also has a different speed of sound, which is *also* inversely proportional to

the square root of density, because $c = \sqrt{K/\rho} \propto 1/\sqrt{\rho}$, where K is an adiabatic constant. Thus, because both parameters are proportional to $1/\sqrt{\rho}$, jet speed u and the speed of sound c are directly proportional to each other, i.e. $f_n \propto c$. This relation explains Roberts' observations that rodent USV frequencies shifted upwards in heliox gas with an increased speed of sound compared to normal air [S11].

Varying jet speed u between 0-20 m/s at a fixed impingent length x of 2 mm results already in a wide range of predicted frequencies up to 80 kHz for modes 1-8 (Figure S2A), and up to 160 kHz for x=1 mm. Recently it was shown that the USVs produced *in vivo* can have a wider range (20-100 kHz) [S12] than the previously accepted range of 30-100 kHz [S13] and our model thus covers the frequency range found both in our *in vitro* as well as *in vivo* observations.

Modulating impingement length

Mice probably only have a limited range to modulate the distance between glottis and thyroid wall, i.e. impingement length x. Most of the frequency modulation will therefore likely be driven by changes in jet speed u, by changing mass flow and glottis shape. Therefore we reasoned that changing laryngeal impingement length experimentally would provide further support for our model.

To modulate jet impingement length after thyroid removal, a 4x2 mm metal plate made out of a flattened 19G needle was placed at the original position of the thyroid (Figure S2B). The precise position and motion of this plate was controlled with an ergometer (Model 300C, Aurora Scientific, Ontario, Canada), which measured displacement at the tip of the lever arm (displacement resolution 1 µm). The larynges were placed in the setup as described above with a constant jet speed (10-30 ml/s) in combination with no movement, or sinusoidal movement of the plate.

To determine the impingement length *x*, we quantified the exact position of the plate and glottal opening by filming the larynx with a high-speed camera (MotionPro-X4, 12 bit CMOS sensor, IDT; 500 frames/s) simultaneously from the top and side by placing a 5x5x5 mm 45˚ angled aluminium surface-coated prism (Thorlabs, Newton, New Jersey, US) next to the larynx. Because the plate often obscured the glottal opening only the side views were used for extraction of plate position and impingement length. To calculate the exact position of the plate (wall) and glottal opening (jet exit) we cross-correlated regions of interest containing only the plate or larynx with the entire image. We used linear interpolation to position the peak in the cross-correlation at a precision of 2.54 µm.

The mean convection speed of the coherent structures is approximated as jet exit speed *u*, which we calculated as the measured tracheal mass flow divided by glottal area. Because we could not directly measure glottal area in this experiment, we needed to fit the area to obtain the measured sound frequency traces. In all four larynges the area was 0.2 – 0.4 mm^2, values that corresponded well with values observed in the intact excised larynx experiments (0.1 – 0.45 mm^2).

In vivo vocalization recordings

USVs emitted in the presence of a familiar female mouse were recorded from two adult (>100 days) male mice prior to *in vitro* experiments. The females were devocalized by unilateral nerve cut [S14]. Mice were placed into a standard acrylic cage located within a single walled attenuating chamber. USVs were recorded with a ½-inch (model 26AC, G.R.A.S.) microphone placed 10 cm above the cage floor. The acoustic signals were amplified and digitized by an Avisoft UltraSoundGate (model 416H, Avisoft Bioacoustics) at a sampling rate of 250 kHz and 16-bit resolution.

USV Analysis

Acoustic parameters of all *in vitro* and *in vivo* USVs were analyzed using Avisoft SASLab Pro software (Avisoft Bioacoustics), as described in detail in [S15]. USVs were automatically detected by the software when energy was above a certain threshold, bounded by >2 ms of silence, minimum duration of 1 ms, and hold time of 20 ms. For each USV, the minimum, maximum, and mean peak frequencies were extracted from spectrograms of each USV (FFT length: 1024, overlap: 75%, Hamming window). Because maximum peak frequencies were measured from the highest frequency bin of the detected signal, USVs with multiple modes had higher maximum frequencies than those with only one mode.

Supplemental References

[S1] Hammerschmidt K., Radyushkin K., Ehrenreich H., Fischer J. (2012). The structure and usage of female and male mouse ultrasonic vocalizations reveal only minor differences. PLoS One. 7(7): e41133.

[S2] Elemans, C. P. H., Spierts, I. L. Y., Hendriks, M., Schipper, H., Müller, U. K., and Van Leeuwen, J. L. (2006). Syringeal muscles fit the trill in ring doves (Streptopelia risoria L.). J. Exp. Biol. 209 (5), 965-977.

[S3] Elemans C.P.H. *et al*. (2015) Universal mechanisms of sound production and control in birds and mammals. Nat. Commun. 6:8978

[S4] Riede, T. (2011). Subglottal pressure, tracheal airflow, and intrinsic laryngeal muscle activity during rat ultrasound vocalization. J. Neurophysiol. 106 (5), 2580–2592.

[S5] Riede, T. (2013). Stereotypic laryngeal and respiratory motor patterns generate different call types in rat ultrasound vocalization. J. Exp. Zool. A. Ecol. Genet. Physiol. 319 (4), 213-224.

[S6] Elemans C.P.H., Zaccarelli R., Herzel H. (2008). The biomechanics and neuromuscular control of vocalisation in a non-songbird. J. R. Soc. Interface. 5, 691-703.

[S7] Elemans CPH, Laje R, Mindlin GB, Goller F (2010). Smooth operator: Avoidance of subharmonic bifurcations through mechanical mechanisms simplifies song motor control in adult zebra finches. J. Neurosci. 30, 13246 –13253.

[S8] Riede, T. and Goller, F. (2010). Peripheral mechanisms for vocal production in birds – differences and similarities to human speech and singing. Brain. Lang. 115, 69–80.

[S9] Mortola, J. P., and Noworaj, A. (1983). Two-sidearm tracheal cannula for respiratory airflow measurements in small animals. J. Appl. Physiol. Respir. Environ. Exerc. Physiol. 55(1 Pt 1), 250–253.

[S10] Ho, C. M., and Nosseir, N. S. (1981). Dynamics of an impinging jet. Part 1. The feedback phenomenon. J. Fluid Mech. 105, 119-142.

[S11] Roberts, L. H. (1975). The rodent ultrasound production mechanism. Ultrasonics. *13*(2), 83-88.

[S12] Grimsley, J. M., Sheth, S., Vallabh, N., Grimsley, C. A., Bhattal, J., Latsko, M., and Wenstrup, J. J. (2016). Contextual Modulation of Vocal Behavior in Mouse: Newly Identified 12 kHz “Mid-Frequency” Vocalization Emitted during Restraint. Front. Behav. Neurosci. 10.

[S13] Sales, G. D. (1972). Ultrasound and mating behaviour in rodents with some observations on other behavioural situations. J Zool, 168, 149-164.

[S14] Nunez, A. A., Pomerantz, S. M., Bean, N. J., and Youngstrom, T. G. (1985). Effects of laryngeal denervation on ultrasound production and male sexual behavior in rodents. Physiol. Behav. 34 (6), 901-905

[S15] Mahrt, E. J., Perkel, D. J., Tong, L., Rubel, E. W., and Portfors, C. V. (2013). Engineered deafness reveals that mouse courtship vocalizations do not require auditory experience. J. Neurosci. 33 (13), 5573-5583.

Author Contributions

Conceptualization, AA, DP, CP, CPHE; Methodology, EM, AA and CPHE; Investigation, EM, CPHE; Writing – Original Draft, EM, AA and CPHE; Writing – Review & Editing, EM, AA, DP, CP, CPHE; Funding Acquisition, CP and CPHE; Resources, CPHE; Supervision, CPHE.