

RESEARCH ARTICLE

Evidence for an audience effect in mice: male social partners alter the male vocal response to female cues

Kelly M. Seagraves^{1,2}, Ben J. Arthur² and S. E. Roian Egnor^{2,*}

ABSTRACT

Mice (*Mus musculus*) form large and dynamic social groups and emit ultrasonic vocalizations in a variety of social contexts. Surprisingly, these vocalizations have been studied almost exclusively in the context of cues from only one social partner, despite the observation that in many social species the presence of additional listeners changes the structure of communication signals. Here, we show that male vocal behavior elicited by female odor is affected by the presence of a male audience – with changes in vocalization count, acoustic structure and syllable complexity. We further show that single sensory cues are not sufficient to elicit this audience effect, indicating that multiple cues may be necessary for an audience to be apparent. Together, these experiments reveal that some features of mouse vocal behavior are only expressed in more complex social situations, and introduce a powerful new assay for measuring detection of the presence of social partners in mice.

KEY WORDS: Social network, Mouse, Social behavior, Ultrasonic vocalization

INTRODUCTION

Interactions with conspecifics can have profound effects on health and reproductive success in social species (Cameron et al., 2009; Fernald and Hirata, 1977; Marler, 1955; Schjelderup-Ebbe, 1922). One avenue for exploring the mechanisms underlying social behavior is to focus on the communicative signals that transfer information between social partners (Bradbury and Vehrencamp, 1998). In mice (*Mus musculus*), ultrasonic vocalizations (USVs) are produced in many contexts, including mother–pup, female–female and male–female interactions, and these vocalizations are known to carry social information (reviewed in Portfors, 2007; Asaba et al., 2014; von Merten et al., 2014; Musolf et al., 2010).

Although mice often live in dense colonies, mouse vocal behavior has traditionally been studied in the context of cues from only one social partner, despite the observation that in many animals the structure of communication signals can be modified by the presence of additional listeners, a phenomenon known as the ‘audience effect’ (reviewed in Matos and Schlupp, 2005). For instance, many anurans increase signaling duration, complexity or rate in response to a competitive audience (Gerhardt and Huber, 2002; Ryan, 1985). Similar findings have been shown in birds – the male zebra finch (*Taeniopygia guttata*), for example, increases the rate of vocal response to his mate’s calls when in the presence of another mated pair (Vignal et al., 2004). Finally, in humans,

typically developing children show improved performance in a task administered in a social context over a non-social version – an improvement that is absent in children with autism spectrum disorder (Chevallier et al., 2014). These studies show that complex social environments can influence signaling behaviors, and lead to the hypothesis that additional social partners could also have an effect on mouse vocal behavior. If true, such a finding would expand our ability to assess the genetic and neural underpinnings of this phenomenon.

To test this hypothesis, we investigated whether the presence of a male social partner affects the male vocal response to female chemosensory cues, because male audiences have been shown to affect conspecific courtship displays in other animals (Ryan, 1985; Gerhardt and Huber, 2002; Fisher and Rosenthal, 2007). Additionally, as it is difficult to distinguish between male and female vocalizations (Hammerschmidt et al., 2012), this configuration ensures that any observed effects are solely the result of a change in male vocal behavior. Our results demonstrate an audience effect in mice, including changes in vocalization rate and structure. We further show that single sensory cues (odor or vocalizations) indicating the presence of a male audience are not sufficient to elicit an effect, suggesting that multiple cues may be necessary for an audience to be apparent. Together, these findings not only provide insight into the roles that mouse USVs play in social interactions but also introduce a novel paradigm that could be used to measure the behavioral influence of social cues, as well as abnormalities in social behavior in mouse models of human social-deficit disorders.

MATERIALS AND METHODS

Experimental animals

SWR/J mice from the Janelia Research Campus breeding facility (founders from The Jackson Laboratory, ME, USA) were housed in 28.5×18×15.2 cm cages (Allentown, Inc., NJ, USA) with ALPHA-dri bedding (Shepherd Specialty Papers, MI, USA), a plastic tube (amber polycarbonate Mouse Tunnels, Bio-Serv, NJ, USA), and a cotton nesting block (Nestlets™, Ancare Corp., NY, USA) and/or paper-strip nesting material (Bed-r’Nest®, The Andersons, Inc., OH, USA). Food (LabDiet 5053, MO, USA) and water were available *ad libitum*. Animals were maintained at 21.4±0.3°C on a 12 h light:dark cycle (lights off at 11:00 h). All experiments were performed at the Janelia Research Campus in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All procedures were carried out according to a protocol approved by the Janelia Research Campus Institutional Animal Care and Use Committee (IACUC; protocol no. 11-70). The Janelia Research Campus Vivarium maintains full AAALAC accreditation.

Male subjects were isolate housed at least 1 week before experiments, and had been exposed to a female (≥9 min, no intromission) to ensure familiarity with female odor (Dizinno et al.,

¹Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. ²Howard Hughes Medical Institute, Janelia Research Campus, 19700 Helix Drive, Ashburn, VA 20147, USA.

*Author for correspondence (egnorrr@janelia.hhmi.org)

1978). Three subject groups were used: those in experiments 1–3, those in experiment 4 and those in experiments 5–6 (see experiment descriptions below). Subjects were 14–50 weeks old at the time of the experiments.

Experimental design

Experiments were performed in sterile 48.3×15.6×26.7 cm plastic cages (Ancare Corp., NY, USA) with ALPHA-dri bedding, unless otherwise stated (Fig. S1A). Two microphones were positioned 36 cm above the floor. The arena was surrounded by acoustically attenuating foam (SONEX™ Valueline Acoustical Foam, 1.5 in thick; Pinta Acoustic, Inc., MN, USA) to dampen echoes and background noise. When required, 20 µl of urine was presented on a sterile nesting block, unless otherwise stated.

Experiment 1: solo males versus paired males exposed to a sterile environment

Pairs of males ($N=18$ animals) were exposed individually (Solo-no-odor condition) or as a pair (Paired-no-odor condition) to a sterile experimental arena and nesting block. At least 2 h prior to testing, subjects were anesthetized with isoflurane and marked for individual identification (black Sharpie®, Newell Rubbermaid Office Products, IL, USA). Animals were given 20 min to acclimate to the arena, and then were tested for 6 min in either the Solo-no-odor or Paired-no-odor condition. After 4 h, animals were tested in the other condition. The order of experimental sessions was counterbalanced across pairs. Each male was tested with two different social partners (separated by ≥10 days) from the pool of 17 possible partners (resulting in $N=18$ pairs), and paired males had not been exposed to one another since weaning (postnatal day 21), if ever.

As it is not possible to distinguish which animal in a pair produced each vocalization, we also created a null model for the vocal behavior of two independently vocalizing animals in this condition by pooling together the vocalizations produced by pairs of males when tested individually (Summed-solo-no-odor condition). Vocalization frequency contours were pooled together (not overlaid in time) after automatic identification and post-processing (see details below).

Experiment 2: solo males versus paired males exposed to female odor

The protocol from experiment 1 was replicated for experiment 2, except that in all experimental sessions female urine was present on a sterile nesting block ($N=20$ animals; referred to as Solo-female-odor and Paired-female-odor conditions). Again, each male was tested with two different social partners from the subject pool ($N=20$ pairs).

Experiment 3: solo males exposed to female odor only or to male urine and female urine simultaneously

Experiment 3 was nearly identical to experiment 2, except that in paired sessions the social partner was replaced with urine from an unfamiliar male presented on an additional nesting block (Female-and-male-urine). All males ($N=18$) were tested once each in the Female-urine-only and Female-and-male-urine conditions.

Experiment 4: solo males exposed to female urine only or to male body odor and female urine simultaneously

Experiment 4 was identical to experiment 2, except that in paired sessions the social partner was replaced with body odor from an unfamiliar male presented on an additional nesting block (Male-body-odor condition). Body odor consisted of fur clippings, saliva and tears from the inner corners of the donor

male's eyes. All subjects were tested once each in the Male-body-odor and Male-body-odor-control conditions ($N=20$).

Experiment 5: solo males exposed to female urine and playbacks of either male USVs or silence

Individual males were exposed to female urine and playbacks of either male USVs or silence in a custom-built acrylic arena ($N=20$), consisting of three chambers (15.25×15.25×15.9 cm) separated by solid dividers with 2.54 cm diameter holes for subjects to pass through (Fig. S1B). The outer walls were made of plastic mesh and the top of the arena was infrared-transparent/visible light-opaque acrylic. Two speakers (Ultrasonic Electrostatic Speaker ESS16 and UltraSoundGate 116 amplifier, Avisoft Bioacoustics, Glienicke, Germany) and four microphones were arranged outside the arena such that a microphone and speaker faced one another through each outer chamber, and two microphones faced one another through the middle chamber.

The recording of silence was several concatenated segments of a Solo-female-odor audio recording that did not contain USVs (30 s total). The USV recording was a 30 s audio segment recorded during a Paired-female-odor experiment, containing 143 USVs in natural bouts. The loudest vocalizations had maximum sound pressure levels (SPL) between 74 and 78 dB SPL at the location of the microphone (37–41 dB SPL background; 41–49 dB SPL USV maximums in center chamber; 37–39 dB SPL USV maximums at far chamber). Both recordings were filtered between 35 and 110 kHz.

In all playback sessions, 10 µl of female urine was located in each playback chamber (no bedding) and the silence recording was played from one speaker. The other speaker played either the silence recording (Silent-playback condition) or the USV recording (USV-playback condition). Subjects were tested once in each condition (3 days between trials), with order counterbalanced between subjects. Playbacks were looped continuously throughout the 6 min session.

Experiment 6: solo males exposed to female odor and an anesthetized male

Experiment 6 was identical to experiment 2, except that data were only collected in paired sessions with an anesthetized social partner (Anesthetized-male condition; $N=19$). Stimulus animals were anesthetized with 0.01 ml g⁻¹ body mass of 10 mg ml⁻¹ ketamine and 0.8 mg ml⁻¹ xylazine.

Data acquisition

Acoustic signals were recorded using ultrasound microphones (CM16/CMPA40-5V; Avisoft Bioacoustics), amplified by 20 dB (40 dB adjustable preamplifier; Avisoft Bioacoustics) and low-pass filtered at 200 kHz (Krohn-Hite Model 3384 Four Channel Filter; Krohn-Hite Corp., MA, USA) before being sampled at 450,450 Hz (NI PXIe-6356 DAQ device; National Instruments Corp., TX, USA) and stored on a hard disk. Video was recorded from above at 29 frames s⁻¹ (camera: Basler CMOS Firewire.A camera with IR filter; Edmund Optics, Inc., NJ, USA; software: StreamPix, NorPix, Inc., Montreal, QC, Canada). All recordings were performed under infrared light (GANZ IR-LT30 Outdoor IR Illuminator, 850 nm; Reytec Imaging, Inc., NY, USA).

Estrous cycle confirmation and female urine collection

Male bedding was added to the home cages of isolate-housed stimulus females to induce estrous cycling (Dalal et al., 2001; Whitten, 1956). Estrous state was determined daily using vaginal cytology, between 08:00 h and 12:00 h. Vaginal cell samples were collected via lavage (Caligioni, 2009; McLean et al., 2012), and estrous stage was

determined by the ratio of cell types observed (Karim et al., 2003). See Fig. S1C for example smears and stage definitions.

Stimulus females were selected based on estrus state, with preference given to estrus day 1, followed by proestrus and then estrus day 2. Naturally voided urine collected from a single female in a sterile environment was stored at $-15\pm 1^\circ\text{C}$ until minutes before experimental use (<10 h).

Vocalization processing

Vocalization identification

Vocalization segments (USV components continuous in time and frequency) were detected and extracted from audio recordings using custom-written software based on multitaper spectral analysis (Ax, available at <https://github.com/JaneliaSciComp/Ax>). Only data between 20 and 120 kHz were analyzed. Overlapping segments in time were Fourier transformed using multiple discrete prolate spheroidal sequences as windowing functions ($NW=22$, $K=43$). An *F*-test was used to infer whether the independent estimates of intensity at each time–frequency point were significantly above noise ($P<0.01$; Thomson, 1982). The frequency resolution of points containing signal was enhanced using the time-derivative of the phase (Charpentier, 1986; Brown and Puckette, 1993). This procedure was performed for multiple segment lengths on each audio channel to capture data at different temporal scales and spatial positions (nonequispaced fast Fourier transform, $NFFT=128$, 256, 512). The data were combined in a single sonogram whose pixel size corresponded to the time resolution of the shortest segment and frequency resolution of the longest. This image was then convolved with a rectangle (1300 Hz tall, 1 ms wide) to fill in small gaps. Locations of contiguous pixels that did not exceed a minimum area of 18.75 Hz s were discarded from further characterization. This procedure was performed on data from each microphone separately, and then the independent lists of vocalizations were combined for further analysis.

The output of the automated vocalization identification process was a series of data points in time and frequency describing each vocal segment (referred to as frequency contours). We then used post-processing heuristics to create the list of whole vocalizations for analysis. Vocal segments were discarded if they had no power above 30 kHz, and then segments with <15 ms of silence between them were combined. Next, putative whole vocalizations shorter than 5 ms or with characteristics similar to the sound of the door closing (>40 ms duration and no segments with frequency modulation $>|0.02|$ kHz ms^{-1}) were excluded from further analysis. Finally, overlapping vocalizations (defined below) and vocalizations from the USV playback were identified and excluded.

Quantifying automatic vocalization processing accuracy

We quantified the accuracy of the automatic vocal segment extraction and post-processing steps by having a human observer visually inspect 10 s of audio recordings from the Solo-no-odor, Paired-no-odor, Solo-female-odor, Paired-female-odor, Female-and-male-urine and Anesthetized-male conditions. The observer marked the start and end times and the highest and lowest frequencies of vocalization segments, and then the segments were processed with the same post-processing heuristics as the experimental data; 99.62% of the human-labeled vocalizations were identified by Ax (262 of 263; human-labeled vocalizations correctly identified if an Ax vocalization overlapped them by any amount in time and frequency), 0.38% were missed (1 of 263) and 1.12% of the automatically identified vocalizations were false positives (3 of 268; Ax-labeled vocalizations not overlapping a human-labeled vocalization). We also measured the accuracy of the human

observer relative to another human by comparing vocalization labels for the same 55 s of data from two files; 95.37% of the vocalizations were correctly identified (268 of 281), 4.63% were missed (13 of 281) and 0% were false alarms (0 out of 264).

Automatically labeled vocalizations more accurately overlapped human-labeled vocalizations in the time domain than in the frequency domain (time: 94.66% overlapped by $\geq 70\%$; frequency: 42.75% overlapped by $\geq 70\%$). The poor accuracy in the frequency domain was driven by low-amplitude harmonic components, which were identified by the humans but not by Ax (time: 98.51% of human-labeled vocalizations overlapped by $\geq 70\%$; frequency: 89.55% overlapped by $\geq 70\%$).

Calculating false alarm rates

To generate a more accurate estimate of the false alarm rate (FA; automatically identified vocalizations that did not contain vocal signal) for recordings of one or two vocally competent animals, we manually identified false alarms in all audio recordings from Solo-no-odor and Paired-no-odor experiments. Using these data, we defined an FA threshold for each condition as three standard deviations above the FA average. A Summed-solo threshold was also calculated from the FA counts from the Solo-no-odor condition in each pair. Only experiments with more automatically labeled vocalizations than the FA threshold were considered to contain a significant number of vocalizations: Solo mean=5.72 FAs (s.d.=7.84), Solo FA threshold=29.24 FAs; Summed-solo mean=11.44 FAs (s.d.=9.88), Summed-solo FA threshold=41.08 FAs; Paired mean=3.72 FAs (s.d.=2.78), Paired FA threshold=12.07 FAs. These thresholds were used to exclude conditions from vocal structure analysis in cases where there were too few vocalizations.

Removal of overlapping vocalizations

Because of concerns about how simultaneous vocalizations could affect our results, we visually identified overlapping vocalizations in spectrograms of the audio data from each of the Paired-female-odor ($N=20$) sessions and excluded them from analyses. We defined overlapping vocalizations as vocalizations that had at least two intersecting vocal segments (see Fig. S1D for example). We could not rule out that some simultaneous vocalizations may look continuous, or that some near-simultaneous but non-overlapping vocalizations may have been combined based on their proximity in time. The average percentage of overlapping vocalizations in the Paired-female-odor condition was 2.1% (s.d.=2.9%, overlaps in 16 of 20 sessions).

Data analysis and statistics

All *P*-values reported in the main text were corrected using False Discovery Rate procedures (Benjamini and Hochberg, 1995) to control for multiple comparisons, and tests were considered significant if adjusted *P*-values (P_{adj}) were <0.05 .

Vocalization counts

Vocalization count is the number of vocalizations recorded in an experimental session. As vocalization counts were not always normally distributed (Lilliefors goodness-of-fit test, $\alpha=0.05$), non-parametric tests were used to make the following vocalization count comparisons: the Solo-no-odor and Paired-no-odor conditions were compared with the matching Solo-female-odor and Paired-female-odor conditions (unpaired, Wilcoxon rank sum tests), while the Summed-solo-no-odor condition was compared with the Paired-no-odor condition, and the Summed-solo-female-odor condition was compared with the Paired-female-odor condition (paired, Wilcoxon

signed rank tests). The Female-urine-only condition was compared with the Female-and-male-urine condition, the Male-body-odor condition was compared with the Male-body-odor-control condition, and the Silent-playback condition was compared with the USV-playback condition (paired, Wilcoxon signed rank tests). Finally, vocalization counts when the male social partner was anesthetized (Anesthetized-male condition) were compared with the Solo-female-odor condition (unpaired, Wilcoxon rank sum test). The effects of female odor, the presence of a male audience, and their interaction were also checked using a generalized linear mixed model (GLMM; response variable modeled with Poisson distribution, log link function; performed using the MATLAB 2015b `fitglm` function, MathWorks, Natick, MA, USA).

Acoustic structure

Duration, frequency bandwidth, minimum frequency, maximum frequency, mean frequency, start frequency, end frequency and the number of frequency-modulated chunks (see detailed explanation below) were measured using standard MATLAB functions from the frequency contours of each vocalization. As minimum, maximum, mean, start and end frequency were correlated (Fig. S1E), we limited further analysis of frequency characteristics to frequency bandwidth and mean frequency. To minimize the effect of FAs on our results, we only included experiments with more vocalizations than the FA threshold in analyses of acoustic structure and vocal complexity. Furthermore, we only compared measurements of acoustic structure and vocal complexity between conditions when data were available for ≥ 10 experiments in each condition. If a paired comparison was performed (e.g. a Wilcoxon signed rank test), both experiments in a pair were required to meet the threshold.

As some acoustic structure measurements were not normally distributed, we used the shapes of the distributions to select the percentile to compare between the Summed-solo-female-odor ($N=16$) and Paired-female-odor ($N=16$) conditions. We pooled data across experiments within the two conditions and then visually compared the resulting distributions by plotting their empirical cumulative distribution functions against one another in a plot called a quantile–quantile ($Q-Q$) plot (Wilk and Gnanadesikan, 1968; only center 99.5% is plotted to exclude extreme outliers). We then statistically tested across individual experiments (Wilcoxon signed rank tests) at a single percentile near where the distributions were maximally different, if the $Q-Q$ plot deviated from the line $y=x$ (see Fig. 1D–F). Duration, frequency bandwidth and mean frequency were also compared (Wilcoxon rank sum tests) at the same percentiles between individual experiments from the Male-body-odor ($N=18$) and Male-body-odor-control ($N=18$) conditions, the Male-body-odor ($N=18$) and Paired-female-odor ($N=20$) conditions, the Anesthetized-male ($N=15$) and Solo-female-odor ($N=22$) conditions, and the Anesthetized-male ($N=15$) and Paired-female-odor ($N=20$) conditions.

Vocal complexity and simple syllable types

The frequency modulation slope often changes within individual mouse vocalizations. To quantify these modulation patterns, we broke vocalizations into frequency-modulated segments (sections of the vocalization with the same frequency modulation trend). To do this, we interpolated missing data points in the frequency contour, and then low-pass filtered the frequency contour to remove high-frequency jitter. Next, we located where the slope of the frequency modulation changed sign, which produced an initial set of frequency-modulated segments. Then, we combined adjacent frequency-modulated segments that collectively had a frequency bandwidth < 1 kHz,

based on the limits of frequency discrimination known for mice (Ehret, 1975). Finally, frequency-modulated segments at the beginning or end of the vocalizations were discarded if they were < 5 ms, and frequency-modulated segments were combined in the middle of the vocalizations if there was less than a 0.1 kHz s^{-1} difference in their frequency modulations and they were separated by a frequency-modulated segment < 5 ms (see Fig. S1F for examples).

We categorized vocalizations as simple (1 frequency-modulated segment) or complex (> 1 frequency-modulated segment), and then further divided simple vocalizations into upsweep (> 5 kHz change), downsweep (< -5 kHz change) and flat (≤ 5 kHz change) syllable types based on their frequency bandwidth and modulation slope (Mahrt et al., 2013; Panksepp et al., 2007; Scattoni et al., 2008). The proportions of complex vocalizations, upsweeps, downsweeps and flats were compared between individual experiments from the Summed-solo-female-odor ($N=16$) and Paired-female-odor ($N=16$) conditions (Wilcoxon signed rank test), the Male-body-odor ($N=18$) and Male-body-odor-control ($N=18$) conditions (Wilcoxon signed rank test), the Male-body-odor ($N=18$) and Paired-female-odor ($N=20$) conditions (Wilcoxon rank sum test), the Anesthetized-male ($N=15$) and Solo-female-odor ($N=22$) conditions (Wilcoxon rank sum test), and the Anesthetized-male ($N=15$) and Paired-female-odor ($N=20$) conditions (Wilcoxon rank sum test).

Principal component analysis of acoustic structure, vocal complexity and simple syllable types

As a secondary confirmation of the acoustic structure, vocal complexity and simple syllable type patterns observed across the Paired-female-odor ($N=20$), Summed-solo-female-odor ($N=16$), Male-body-odor ($N=19$), Male-body-odor-control ($N=19$) and Anesthetized-male ($N=15$) conditions, we performed a principal component analysis on these features. We then tested whether there was an overall difference in principal components 1 and 2 across these social contexts using a GLMM (response variable modeled with normal distribution; performed using the MATLAB 2015b `fitglm` function), and tested for significant differences between the same datasets as in the previous analysis.

Male proximity analysis

To examine the behavior of the males in the Paired-female-odor condition, we manually determined the positions of each male and the nesting block (on which female odor was presented) in 100 video frames in which a vocalization occurred and 100 video frames in which a vocalization did not occur from each of the 20 experiments (4000 frames total, randomly selected). Non-vocalization frames were not within 1 s of a vocalization-associated frame, to avoid selecting non-vocalization frames in the middle of vocal bouts. The position was determined by fitting an ellipse to the body of the mice or the nesting block, which provided the location of the object's center and the nose of the animals. We then defined the males as being close if their centers were < 1.5 body lengths apart, and we defined the males as being close to the female odor if their nose was < 1 nesting block diameter from the center of the nesting block. We compared the percentage of frames that the males spent close to one another between the vocalization-associated and non-vocalization frames (Wilcoxon signed rank test), and we compared the percentage of frames in which at least one of the males was close to the female odor between the vocalization-associated and non-vocalization frames (Wilcoxon signed rank test).

To determine whether fighting contributed to the males spending time close together in the Paired-female-odor condition, the latency to the first fight was measured in Paired-no-odor and Paired-female-

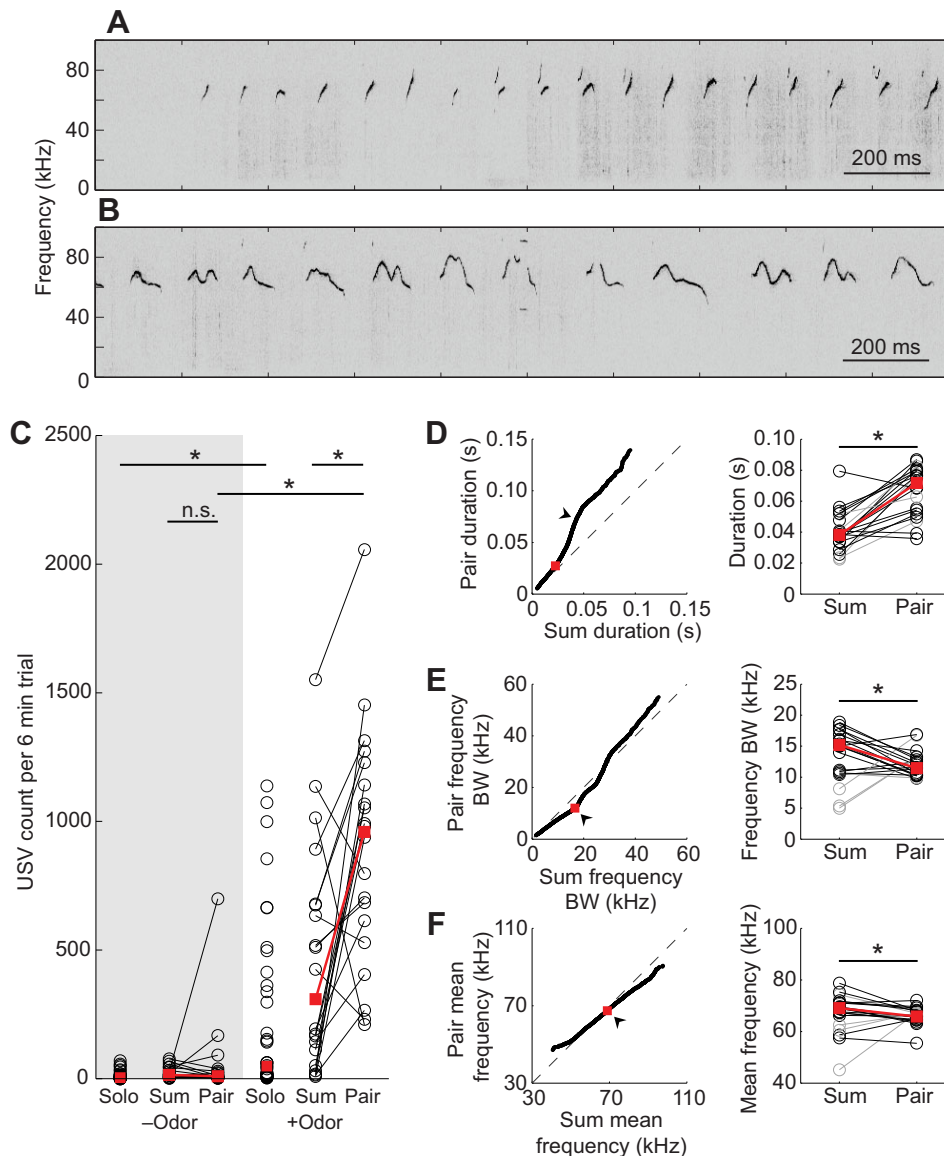


Fig. 1. Vocalizations recorded in solo and paired conditions. Example vocalizations from (A) Solo-female-odor and (B) Paired-female-odor experiments. (C) Ultrasonic vocalization (USV) counts for solo, summed solo (Sum) and paired (Pair) conditions with (+odor) and without (-odor) female odor. Open circles indicate individual experiments. (D–F) Acoustic structure comparisons between the Summed-solo-female-odor and Paired-female-odor conditions. (Left) Q–Q plot comparing pooled acoustic structure distributions (arrowheads indicate percentile at which statistical comparisons were made). (Right) Comparisons of individual experiments at specified percentiles. Gray circles and lines indicate experiments excluded from analysis because of low vocalization counts (see ‘Calculating false alarm rates’ in Materials and methods for details). (D) Vocalization duration, compared at 92nd percentile. (E) Frequency bandwidth (BW), compared at 50th percentile. (F) Mean frequency, compared at 50th percentile. Red squares indicate medians and asterisks indicate statistical significance.

odor sessions by *post hoc* visual inspection of the video recordings. A fight was defined as interlocked wrestling behavior (Van Oortmerssen, 1971). If there was no fighting during the 6 min session, then a score of 360 s was assigned. The latency distributions were not normally distributed (Lilliefors goodness-of-fit test), so a Wilcoxon rank sum test was used to compare the latency distributions between conditions.

RESULTS

A male audience alters the male vocal response to female odor

USVs were recorded from individuals and pairs of males in the absence of other social cues to establish baseline vocalization rates. USVs were also recorded from individuals and pairs when female urine was present, to test the effect of a male audience on the

male vocal response to female odor (example vocalizations in Fig. 1A,B).

Effects of female odor and a male audience on vocalization count

Individuals and pairs of males produced few vocalizations when exposed to a sterile environment (Solo-no-odor and Paired-no-odor conditions; Fig. 1C), demonstrating that neither exposure to a novel environment nor male–male social contact elicits robust vocal behavior. As expected from previous observations (e.g. Nyby et al., 1979), introducing female odor increased the vocal rate of individual males (Solo-no-odor versus Solo-female-odor: $P_{\text{adj}}=0.0066$, $W=240.50$; Fig. 1C). Likewise, female odor increased the number of vocalizations produced by male pairs (Paired-no-odor versus Paired-female-odor: $P_{\text{adj}}=0.000021$, $W=178$; Fig. 1C), indicating that paired males also respond vocally to female

chemosensory cues (these results were corroborated using a GLMM; see Table S1).

By definition, an audience effect is a change in behavior caused by the presence of another animal. We therefore tested whether the presence of a male audience significantly changes the number of female odor-elicited vocalizations. To compare the behavior of two animals with the behavior of individuals, we generated a null model for the expected vocal behavior of two independently vocalizing males (Summed-solo-no-odor and Summed-solo-female-odor) by pooling the vocalizations produced individually by each member of a pair in the Solo-no-odor or Solo-female-odor conditions, respectively. The presence of a male social partner did not affect vocal count when female odor was not present (Paired-no-odor versus Summed-solo-no-odor: $P_{\text{adj}}=0.73$, $T=76$; Fig. 1C), but vocalization counts were higher than expected when female urine and a male audience were present (Paired-female-odor versus Summed-solo-female-odor: $P_{\text{adj}}=0.0092$, $T=25$; Fig. 1C; corroborated using a GLMM; Table S1), indicating that male audiences do affect the male vocal response to female odor, and demonstrating an audience effect in mice.

Effect of a male audience on acoustic structure

In addition to the effect on vocalization count, the presence of a male audience also modified the acoustic structure of female odor-elicited USVs. We visualized the differences between the distributions of vocalization durations in the Summed-solo-female-odor and Paired-female-odor conditions using a quantile–quantile ($Q-Q$) plot (Fig. 1D, left), which shows the empirical cumulative distribution functions of each pooled dataset plotted against one another. If the distributions have similar shapes, then the data will lie along the line $y=x$. The $Q-Q$ plot comparing durations in the Summed-solo-female-odor and Paired-female-odor conditions deviated from the line $y=x$, such that the Paired-female-odor vocalizations were longer than the Summed-solo-female-odor vocalizations above the 60th percentile, with a maximal difference near the 92nd percentile. We tested whether this difference in the pooled data was consistent across individuals. A comparison of the duration at the 92nd percentile in individual experiments showed that vocalizations produced with a male audience were significantly longer ($P_{\text{adj}}=0.0062$, $T=130$; Fig. 1D, right). In contrast, the $Q-Q$ plots for frequency bandwidth (Fig. 1E, left) and mean frequency (Fig. 1F, left) each lay approximately on the line $y=x$, and therefore these measurements were compared at the 50th percentile. Frequency bandwidths at the 50th percentile were larger in individual Summed-solo-female-odor experiments ($P_{\text{adj}}=0.0188$, $T=121$; Fig. 1E, right), while mean frequencies at the 50th percentile were higher in the Summed-solo-female-odor condition ($P_{\text{adj}}=0.0206$, $T=120$; Fig. 1F, right). These analyses show that a male audience modifies features of vocal structure, including

vocalization duration, frequency bandwidth and mean frequency, and thus there is an audience effect on each of these features.

Effect of a male audience on vocal complexity and simple syllable types

We further investigated whether a male audience affects frequency modulation in individual vocalizations. We categorized USVs as simple (no change in modulation direction) or complex (≥ 1 change in modulation direction), and found that the proportion of complex vocalizations was higher in the Paired-female-odor condition than in the Summed-solo-female-odor condition ($P_{\text{adj}}=0.0289$, $T=117$; Fig. 2A).

Another way to address differences in frequency modulation between vocalizations is to group vocalizations into categories based on their modulation pattern. We divided the simple vocalizations into upsweep, downsweep and flat categories, and found that the proportion of upsweeps decreased when an audience was present (Summed-solo-female-odor versus Paired-female-odor: $P_{\text{adj}}=0.0057$, $T=7$; Fig. 2B), while the proportion of downsweeps and flats remained the same (downsweeps: $P_{\text{adj}}=0.085$, $T=108$; flats: $P_{\text{adj}}=0.65$, $T=57$; Fig. 2C,D). These results show that there are significant differences in the proportions of some syllable types produced between the with- and without-audience conditions, suggesting that syllable proportions could convey information about the specific social context to the audience or other receivers.

Together, these results show that the presence of a male audience affects many aspects of vocal communication – vocalization count, acoustic structure, vocal complexity and the proportional use of syllable types.

Paired males vocalize while investigating one another

To begin to assess possible functions of the observed audience effects, we examined the behavior of males in the Paired-female-odor condition. We found that males spent more time close to one another (centroids within 1.5 body lengths) while vocalizing than when not vocalizing ($P_{\text{adj}}=0.00081$, $T=210$), and that even when not vocalizing they still spent considerable time close to one another (Fig. 3A, left). In contrast, males spent relatively little time vocalizing near the female odor, and there was no difference in the amount of time that they spent near the female odor when they were and were not vocalizing ($P_{\text{adj}}=0.070$, $T=47$; Fig. 3A, right). As unfamiliar males fight when female odor is not present (Fig. 3B), we looked to see whether fighting drove the large amount of time that the males spent close to one another when female odor was present. On the contrary, the latency to the first fight was significantly longer when female odor was present ($P_{\text{adj}}=0.00028$, $W=214$; Fig. 3B), such that there were no fights between the

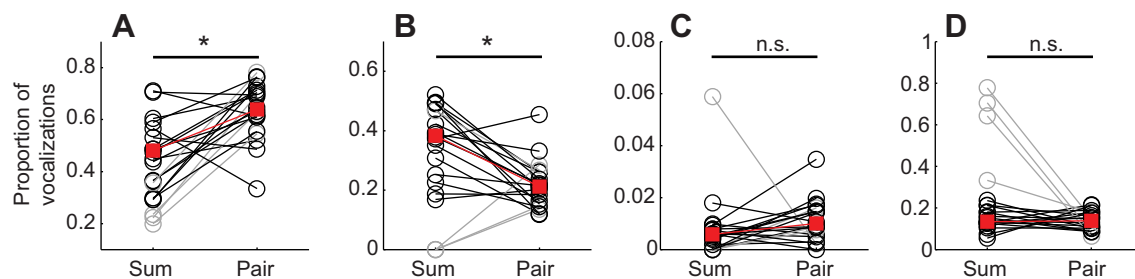


Fig. 2. Comparing vocalization complexity and simple syllable types between Summed-solo-female-odor and Paired-female-odor conditions.

Proportions of (A) complex, (B) upsweep, (C) downsweep and (D) flat vocalizations. Open circles indicate individual experiments, red squares indicate the median and asterisks indicate statistical significance. Gray lines and circles indicate experiments excluded from analysis because of low vocalization counts.

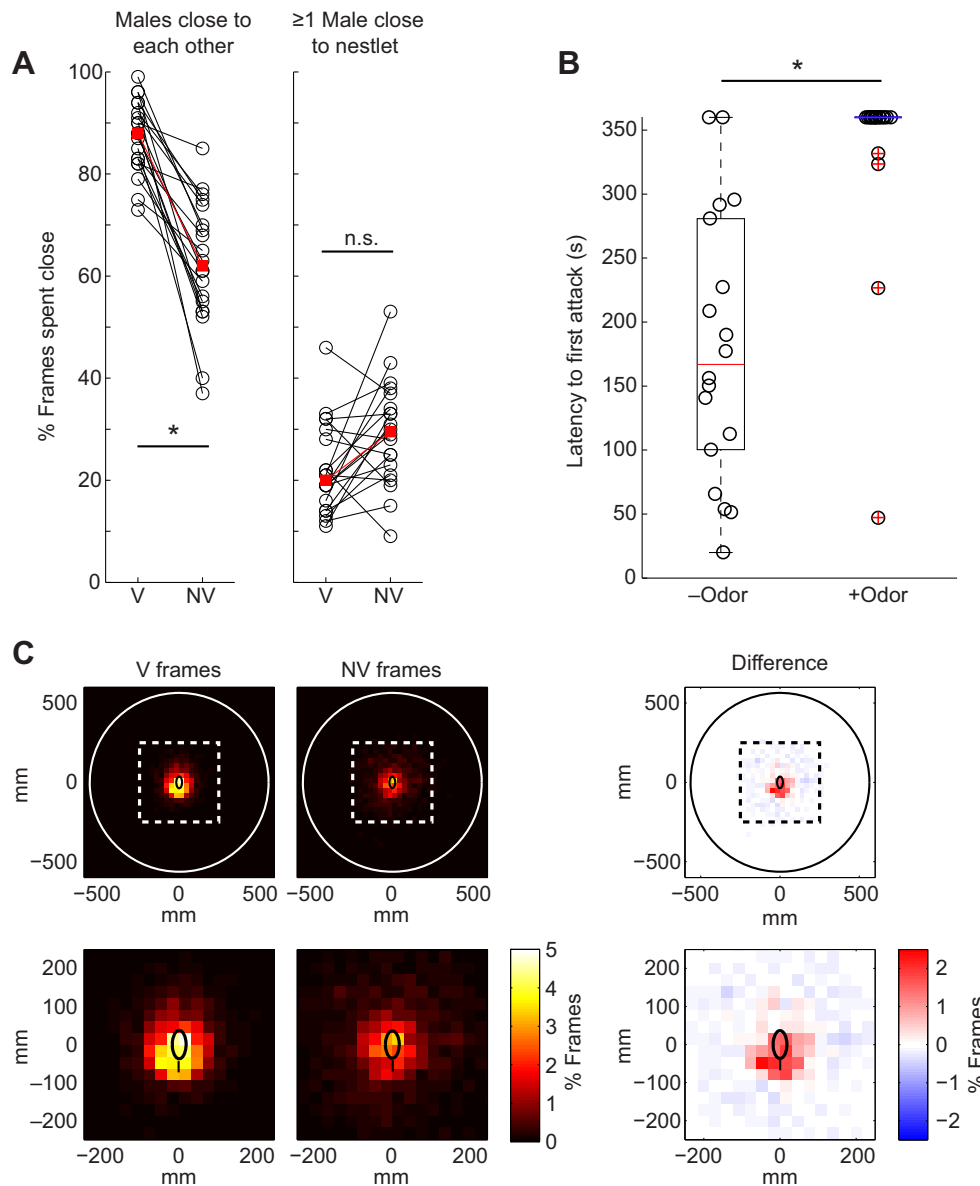


Fig. 3. Male behavior during interactions with a male audience. (A, left) Percentage of vocalization-associated (V) and non-vocalization (NV) frames in which the males were close to one another, where close is defined as their centroids being less than 1.5 mouse body lengths apart. (A, right) Percentage of vocalization-associated and non-vocalization frames in which at least one male was close to the female odor, where close is defined as the mouse's nose being within 1 nesting block diameter of the nesting block's center. Open circles indicate individual experiments, red squares indicate the median and asterisks indicate statistical significance. (B) Latency to the first attack between paired males, with and without female odor present. Experiments with no fights were assigned a value of 360 s. Red lines indicate the median, upper and low edges of the black boxes indicate the interquartile range, red crosses indicate outliers and asterisks indicate statistical significance. (C) Density plots showing the position of one male's nose relative to the center of the other male, in vocalization-associated frames and non-vocalization frames. The difference between the vocalization-associated density plot and the non-vocalization density plot is shown on the right; red and blue indicate more time spent at that location in the vocalization-associated or non-vocalization frames, respectively. The top row shows all of the possible locations of the noses, while the bottom row shows an enlarged view of the center region around the reference mouse (indicated in the top plots by a dashed white or black line). The white or black outer circle indicates the maximum distance that one male's nose could be relative to the other male's center. A black ellipse is plotted in the center to represent the reference mouse, with a vertical line showing the position of the tail.

majority of pairs. Instead, the males spent considerable time with their noses near the anogenital region of the other male, particularly while producing vocalizations (Fig. 3C).

Only some male cues elicit changes in male vocal behavior

Although a male audience affects the male vocal response to female odor in several ways, it is not clear what features of the audience are necessary to elicit these changes. We therefore investigated whether individual male cues are sufficient to elicit an audience effect. Male

subjects were exposed to female odor and one of the following male cues: male urine, male body odor (fur, saliva and tears), USVs from a pair of males, or an anesthetized male. We did not test whether visual cues from the male audience were necessary because all experiments were conducted in the dark.

Effect of male cues on vocalization count

None of the male cues had an effect on vocalization count (Female-urine-only versus Female-and-male-urine: $P_{\text{adj}}=0.327$, $T=56.5$;

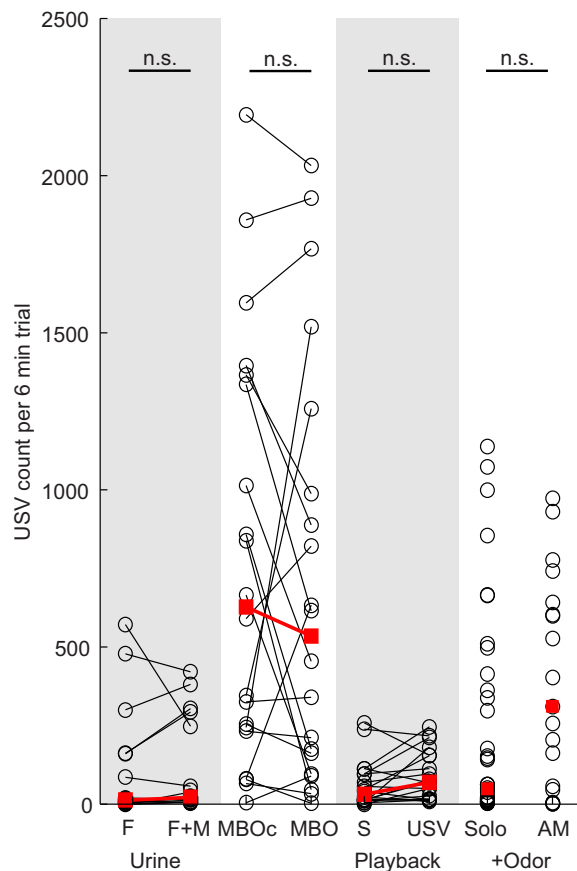


Fig. 4. Influence of male cues on male vocal counts. USV counts for Female-urine-only (F), Female-and-male-urine (F+M), Male-body-odor (MBO), Male-body-odor-control (MBOc), Silent-playback (S), USV-playback (USV), Solo-female-odor (Solo) and Anesthetized-male (AM) conditions. Open circles indicate individual experiments and red squares indicate the median.

Male-body-odor versus Male-body-odor-control: $P_{\text{adj}}=0.567$, $T=77$; Silent-playback versus USV-playback: $P_{\text{adj}}=0.152$, $T=59.5$; Anesthetized-male versus Solo-female-odor: $P_{\text{adj}}=0.129$, $W=335$; Fig. 4).

Effect of male body odor and an anesthetized male audience on acoustic structure

Males produced too few vocalizations in the Female-and-male-urine and USV-playback conditions to compare acoustic structure characteristics between these conditions and their controls. However, we did compare acoustic structure characteristics between the Male-body-odor and Male-body-odor-control conditions (example vocalizations in Fig. 5A,B), the Male-body-odor and Paired-female-odor conditions, the Solo-female-odor and Anesthetized-male conditions (example vocalizations in Fig. 6A,B), and the Anesthetized-male and Paired-female-odor conditions.

We found that duration (92nd percentile) was not different between the Male-body-odor and Male-body-odor-control conditions ($P_{\text{adj}}=0.418$, $T=108$; Fig. 5C), but that it was different between the Male-body-odor and Paired-female-odor conditions ($P_{\text{adj}}=0.000072$, $W=550$; Fig. 5C), indicating that male body odor is not sufficient to induce the same changes in vocalization duration as an awake audience. Likewise, male body odor was not sufficient to induce changes in frequency bandwidth or mean frequency at the 50th percentile (Male-body-odor versus Male-body-odor-control frequency bandwidth: $P_{\text{adj}}=0.463$, $T=65$; Male-body-odor versus

Paired-female-odor frequency bandwidth: $P_{\text{adj}}=0.0022$, $W=268$; Male-body-odor versus Male-body-odor-control mean frequency: $P_{\text{adj}}=0.557$, $T=69$; Male-body-odor versus Paired-female-odor mean frequency: $P_{\text{adj}}=0.0353$, $W=306$; Fig. 5D,E).

In contrast to male body odor, we found that the effect of an anesthetized audience on vocalization duration is indistinguishable from that of an awake audience. Anesthetized-male vocalization durations were longer than Solo-female-odor durations ($P_{\text{adj}}=0.0023$, $W=302$; Fig. 6C), but not different from Paired-female-odor durations ($P_{\text{adj}}=0.199$, $W=407$; Fig. 6C). Meanwhile, the anesthetized audience only partially reproduced the effects that an awake audience had on frequency bandwidth and mean frequency. Frequency bandwidths at the 50th percentile were not different between the Solo-female-odor and Anesthetized-male conditions ($P_{\text{adj}}=0.111$, $W=481$; Fig. 6D) or between the Anesthetized-male and Paired-female-odor conditions ($P_{\text{adj}}=0.90$, $W=365$), suggesting that the Anesthetized-male frequency bandwidths are intermediate between the Solo-female-odor and Paired-female-odor bandwidths. Likewise, Anesthetized-male mean frequencies at the 50th percentile were not different from either the Solo-female-odor or Paired-female-odor mean frequencies (Solo-female-odor versus Anesthetized-male: $P_{\text{adj}}=0.152$, $W=474$; Paired-female-odor versus Anesthetized-male: $P_{\text{adj}}=0.363$, $W=326$; Fig. 6E). Thus, the anesthetized audience fully recapitulates the effect of an awake audience on vocalization duration, while having an intermediate effect on frequency bandwidth and mean frequency.

Effect of male body odor and an anesthetized male audience on vocalization complexity and simple syllable types

Male body odor was not sufficient to induce changes in vocalization complexity or the usage of simple syllable types (Male-body-odor versus Male-body-odor-control complex vocalizations: $P_{\text{adj}}=0.983$, $T=86$; Male-body-odor versus Paired-female-odor complex vocalizations: $P_{\text{adj}}=0.0015$, $W=518$; Male-body-odor versus Male-body-odor-control upsweeps: $P_{\text{adj}}=0.744$, $T=77$; Male-body-odor versus Paired-female-odor upsweeps: $P_{\text{adj}}=0.0001$, $W=235$; Male-body-odor versus Male-body-odor-control downsweeps: $P_{\text{adj}}=0.743$, $T=52$; Male-body-odor versus Paired-female-odor downsweeps: $P_{\text{adj}}=0.0060$, $W=498.5$; Male-body-odor versus Male-body-odor-control flats: $P_{\text{adj}}=0.332$, $T=113$; Male-body-odor versus Paired-female-odor flats: $P_{\text{adj}}=0.334$, $W=432$; Fig. 5F–I), while an anesthetized social partner had the same effect as an awake social partner. Compared with the Solo-female-odor condition, the proportion of complex vocalizations increased when an anesthetized social partner was present ($P_{\text{adj}}=0.0242$, $W=370$; Fig. 6F), but there was no difference in the proportion of complex vocalizations between the Anesthetized-male and Paired-female-odor conditions ($P_{\text{adj}}=0.119$, $W=213$; Fig. 6F). Additionally, there were fewer upsweeps produced in the Anesthetized-male condition than in the Solo-female-odor condition ($P_{\text{adj}}=0.0057$, $W=181$; Fig. 6G), but the proportion of downsweeps and flats remained constant (downsweeps: $P_{\text{adj}}=0.124$, $W=345$; flats: $P_{\text{adj}}=0.358$, $W=321$; Fig. 6H and I, respectively). Meanwhile, the proportions of upsweeps, downsweeps and flats were not different between the Anesthetized-male and Paired-female-odor conditions (upsweeps: $P_{\text{adj}}=0.363$, $W=304$; downsweeps: $P_{\text{adj}}=0.666$, $W=286.5$; flats: $P_{\text{adj}}=0.324$, $W=308$; Fig. 6G–I, respectively).

Thus, an anesthetized male elicits similar changes in acoustic structure, syllable complexity and relative syllable proportions as an awake male, while male body odor does not produce an effect (corroborated using a GLMM analysis of the first two components in a principal component analysis; Fig. S2, Tables S2 and S3).

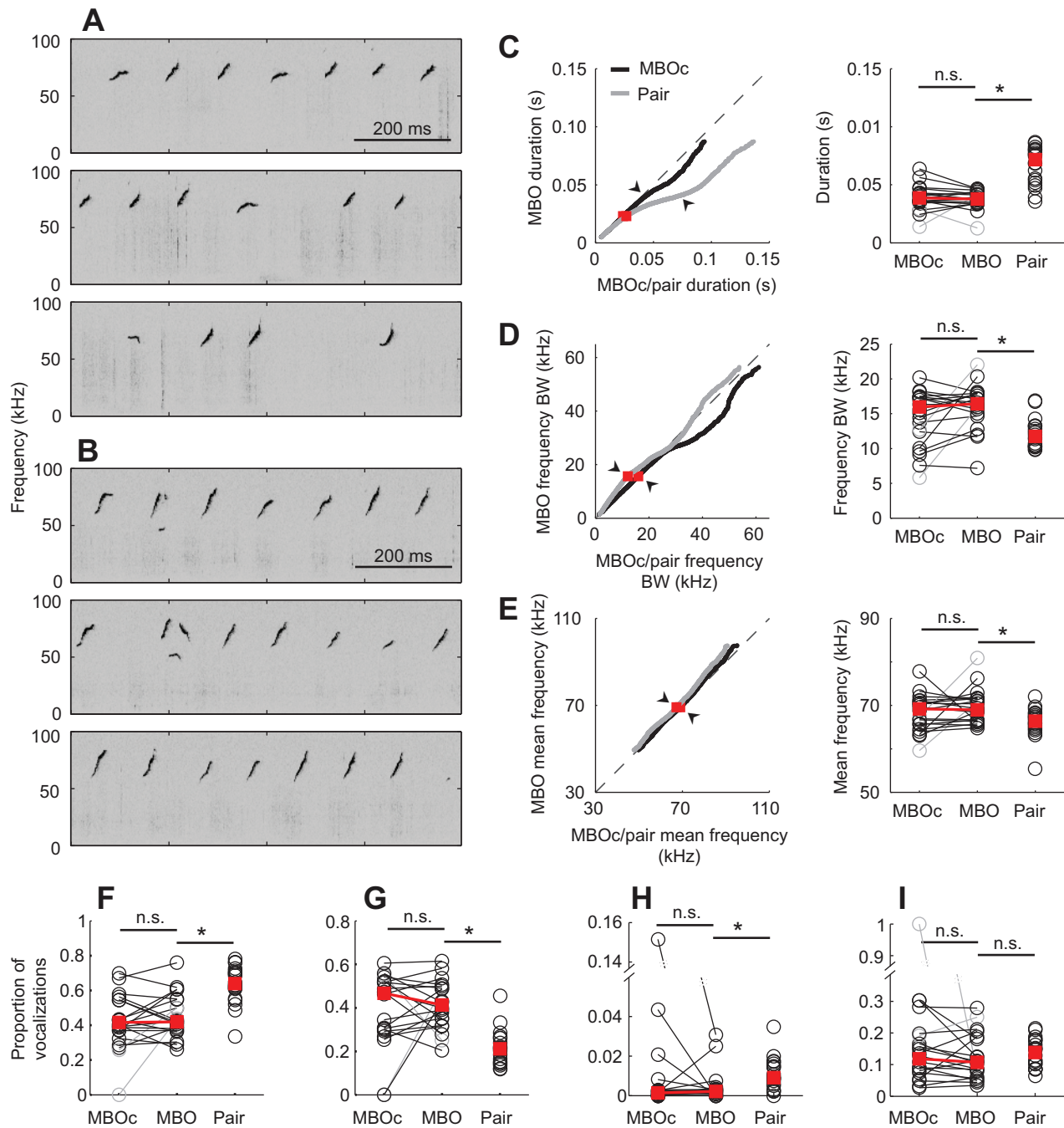


Fig. 5. Influence of male body odor on vocalization features. Example vocalizations from the (A) Male-body-odor-control and (B) Male-body-odor conditions. (C–E) Comparisons of vocalization duration, frequency bandwidth and mean frequency, respectively, between the Male-body-odor and Male-body-odor-control conditions, and between the Male-body-odor and Paired-female-odor conditions. (F–I) Proportions of complex vocalizations, upsweeps, downsweeps and flats in the Male-body-odor-control, Male-body-odor and Paired-female-odor conditions. Red squares indicate medians and asterisks indicate statistical significance.

DISCUSSION

An audience effect is any change in signaling behavior driven by the identity of target and/or non-target receivers (Gyger et al., 1986; Marler et al., 1986; Matos and Schlupp, 2005). Although audience effects have been shown in other animals (Matos and McGregor, 2002; Ryan, 1985; Vignal et al., 2004), there have been few studies on this effect in rodents (Blanchard et al., 1991; Wöhr and Schwarting, 2008), limiting our ability to assess the genetic and neural underpinnings of this phenomenon.

Our findings show that a male audience modifies male vocal responses to female odor, thus demonstrating an audience effect in

mice. All aspects of male vocal behavior examined – vocalization count, acoustic structure and syllable complexity – were modified by the presence of a male audience. In addition, we show that while single sensory cues indicating the presence of a male audience (urine, body odor or vocalization playback) were not sufficient to elicit changes in vocal behavior, the presence of an anesthetized male produced changes that were similar to those observed with an awake audience.

The male competitor may be the intended receiver

In the absence of other social partners, female odor is required to elicit vocalizations from male mice (Nyby et al., 1979). This has

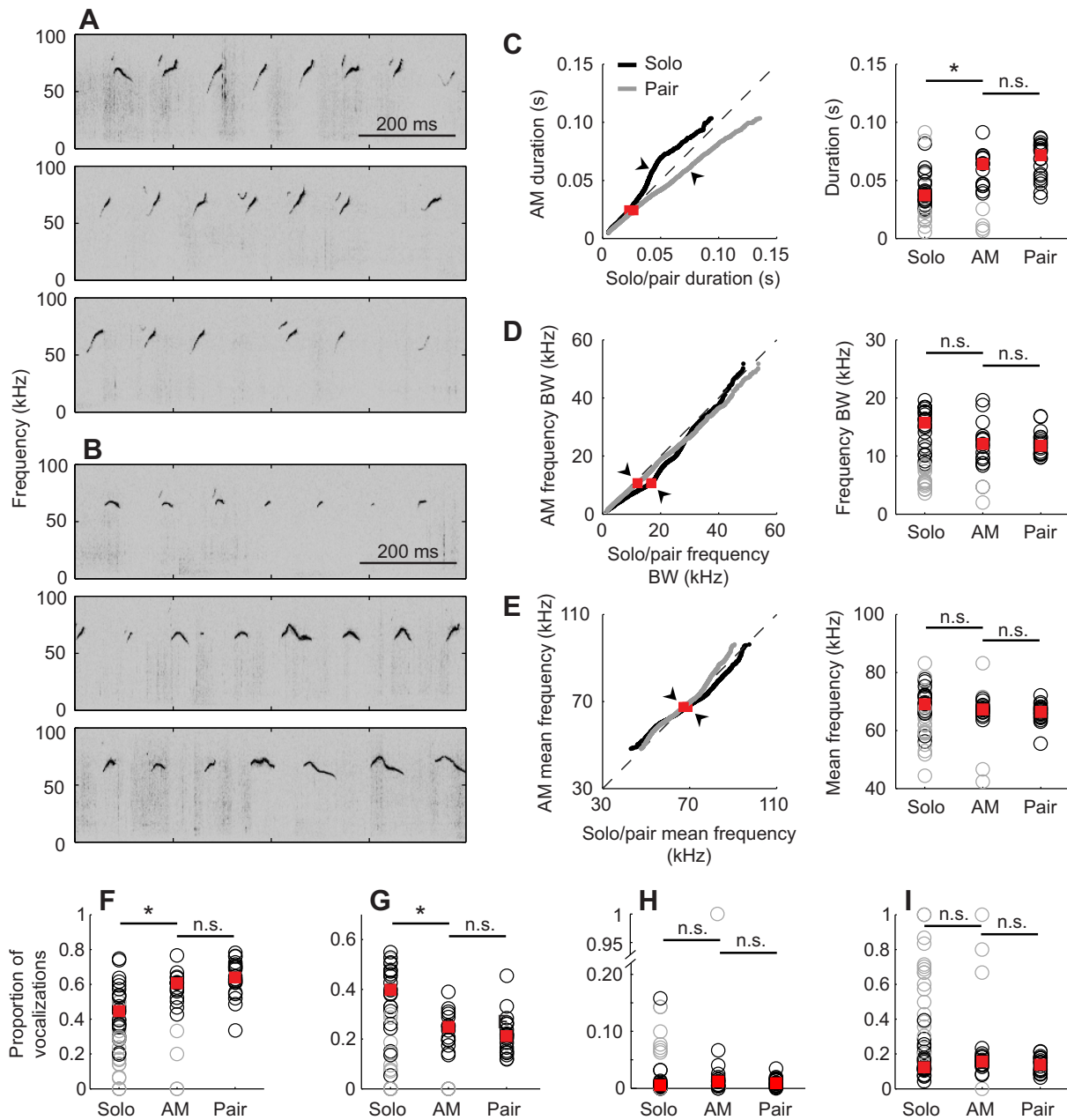


Fig. 6. Influence of an anesthetized male on vocalization features. Example vocalizations from the (A) Solo-female-odor and (B) Anesthetized-male conditions. (C–E) Comparisons of vocalization duration, frequency bandwidth and mean frequency, respectively, between the Anesthetized-male and Solo-female-odor conditions, and between the Anesthetized-male and Paired-female-odor conditions. (F–I) Proportions of complex vocalizations, upsweeps, downsweeps and flats in the Solo-female-odor, Anesthetized-male and Paired-female-odor conditions. Red squares indicate medians and asterisks indicate statistical significance.

led to the belief that these vocalizations are directed at the female and that they play a role in male–female interactions. While female odor is also necessary to reliably elicit vocalizations from pairs of male mice (current study; Nyby et al., 1976; Sales, 1972; Stowers et al., 2002; Whitney et al., 1973), we found that in this context, males primarily vocalize when they are close to one another, rather than when they are near the female odor or when they are independently exploring the environment. This suggests that when a male audience is present, males may be directing their vocalizations at the other male, instead of, or in addition to, the potential female partner. Additionally, the fact that males alter their vocal behavior when a male audience is present suggests that the

function of the vocal display is context dependent. In other animals, increases in vocalization count and duration, similar to those observed in this study, are associated with higher levels of energy expenditure (Prestwich, 1994; Taigen and Wells, 1985) and provide a measure of fitness (Andersson, 1994; Klump and Gerhardt, 1987). Together, these observations are consistent with the males using their vocal behavior as a competitive display. Thus, while one interpretation of our results is that the male audience influences the other male's female-directed vocal behavior, it is also possible to interpret our findings as the potential female audience acting as a releaser for vocal behavior between males.

The audience effect as a readout of the behavioral influence of social cues

For the presence of an audience to change the behavior of a signaler, the signaler must be aware of the audience. However, direct contact with the audience may not be necessary for the audience to be apparent. For example, chemical cues from another male induce changes in the courtship displays of male newts (*Lissotriton boscai*; Aragón, 2009), birds robustly respond to playbacks of conspecific vocalizations (Stoddard et al., 1991, 1988), and humans behave more generously in an economic game when shown stylized visual cues of an observer (Haley and Fessler, 2005). In our experiments we replaced a male audience with male odor (urine or body odor), playbacks of male USVs or an anesthetized male (mixture of tactile and body odor cues) to test whether any of these cues are sufficient to produce an audience effect.

We found no difference in vocalization number when males were exposed to female odor in the presence of male urine, male body odor or playbacks of male USVs compared with when they were exposed to female odor alone. Furthermore, we observed no change in the acoustic features of vocalizations when males were exposed to female urine and male body odor. These results suggest that male odor and playbacks of male USVs are not sufficient to elicit a measurable audience effect. This is surprising, because both male odor and male USVs are thought to affect conspecific behavior in mice. Odors of unfamiliar males increase aggressive behavior between males (Archer, 1968; Mackintosh and Grant, 1966), females use male scent marks to assess male fitness (Rich and Hurst, 1998), and body odors have been shown to affect how mice respond to other social cues (Asaba et al., 2014; Ferrero et al., 2013). Similarly, female mice approach playbacks of male USVs (Hammerschmidt et al., 2009; Musolf et al., 2010; Shepard and Liu, 2011). However, male odors and vocalizations may not be good predictors of the presence of another male in the wild, or the presentation of these cues in our experiments may have lacked unknown naturalistic elements. It is also worth noting that the minimal vocal response to female odor in the presence of male USV playbacks suggests that male mice, unlike some long-distance vocalizers like birds and frogs (McGregor, 1992), may not vocally respond to USVs in the absence of other social cues.

In contrast, an anesthetized male reproduced many of the audience effects elicited by an awake male – vocalizations produced in this condition were longer and more complex, and there were fewer upsweep vocalizations. Thus, tactile contact with the male audience and exposure to male body odor together are sufficient to elicit changes in female odor-elicited vocal behavior. That male body odor alone was not sufficient to elicit an effect suggests that tactile contact plays a key role. This experiment also shows that active behavior (motion or vocalization) on the part of the audience is not required to induce an effect. However, that the changes in vocal behavior elicited by an anesthetized male did not entirely match those elicited by an awake male suggests that the effect of an audience is not all-or-none and may be driven by specific features of the audience. If true, this finding advances a suite of possible future studies that could use the audience effect as a tool to examine how the features of male social partners affect vocal behavior.

Conclusions

Our finding that male mice respond to the presence of an audience indicates that mice are sensitive to their social environment, and that they may be adapted to signal in communication networks (McGregor, 2005). This observation expands the social contexts

in which adult USVs are produced, and provides a foundation for future studies that examine USV behavior in complex social groups. Furthermore, it opens new avenues of inquiry about the factors that influence the audience effect (e.g. experience, strain variation) and the mechanisms by which sensory information about potential social partners modifies the structure of vocal signals. Finally, progressing beyond the social dyad is likely to uncover novel opportunities for understanding the details of mouse vocal and social behavior, and may drive the development of improved assays relevant to human models of socio-cognitive and communication disorders.

Acknowledgements

We thank Megan Atkins for assistance with Ax ground-truthing and Sara Hänni for running pilot experiments that inspired this work. We also thank Juan Nunez-Iglesias, Jeremy Freeman and Jason Wittenbach for statistical advice. Finally, we thank Gül Dölen, Berthold Hedwig, Anthony Leonardo, Albert Lee and Vivek Jayaraman for their helpful remarks on the manuscript.

Competing interests

The authors declare no competing or financial interests.

Author contributions

K.M.S. and S.E.R.E. designed the experiments, B.J.A. wrote the acoustic segmentation software, K.M.S. collected and analyzed the data, K.M.S. and S.E.R.E. interpreted the data and wrote the article.

Funding

This work was supported by the Howard Hughes Medical Institute (HHMI) and HHMI's Janelia Graduate Scholars Program. Deposited in PMC for release after 6 months.

Supplementary information

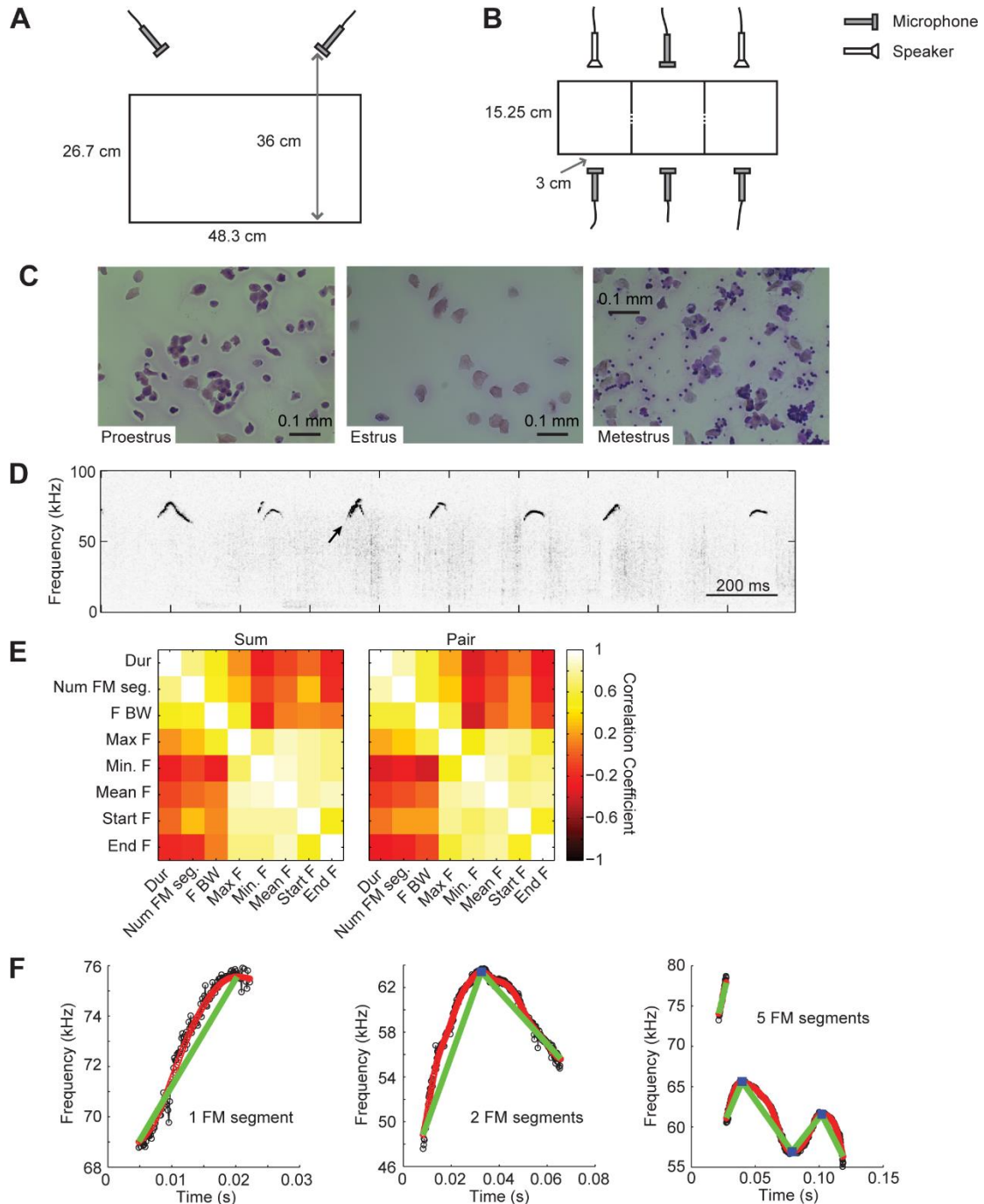
Supplementary information available online at <http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.129361/-/DC1>

References

- Andersson, M. B. (1994). The theory of sexual selection. In *Sexual Selection*, pp. 3–31. Princeton, NJ: Princeton University Press.
- Aragón, P. (2009). Conspecific male chemical cues influence courtship behaviour in the male newt *Lissotriton boscai*. *Behaviour* **146**, 1137–1151.
- Archer, J. (1968). The effect of strange male odor on aggressive behavior in male mice. *J. Mammal.* **49**, 572–575.
- Asaba, A., Okabe, S., Nagasawa, M., Kato, M., Koshida, N., Osakada, T., Mogi, K. and Kikusui, T. (2014). Developmental social environment imprints female preference for male song in mice. *PLoS ONE* **9**, e87186.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 289–300.
- Blanchard, R. J., Blanchard, D. C., Agullana, R. and Weiss, S. M. (1991). Twenty-two kHz alarm cries to presentation of a predator, by laboratory rats living in visible burrow systems. *Physiol. Behav.* **50**, 967–972.
- Bradbury, J. W. and Vehrencamp, S. L. (1998). *Principles of Animal Communication*. Sunderland, MA: Sinauer Associates, Inc.
- Brown, J. C. and Puckette, M. S. (1993). A high resolution fundamental frequency determination based on phase changes of the Fourier transform. *J. Acoust. Soc. Am.* **94**, 662–667.
- Caligioni, C. S. (2009). Assessing reproductive status/stages in mice. *Curr. Protoc. Neurosci.* **A.41.1–A.41.8**.
- Cameron, E. Z., Setsaas, T. H. and Linklater, W. L. (2009). Social bonds between unrelated females increase reproductive success in feral horses. *Proc. Natl. Acad. Sci. USA* **106**, 13850–13853.
- Charpentier, F. (1986). Pitch detection using the short-term phase spectrum. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86. IEEE*, pp. 113–116.
- Chevallier, C., Parish-Morris, J., Tonge, N., Le, L., Miller, J. and Schultz, R. T. (2014). Susceptibility to the audience effect explains performance gap between children with and without autism in a theory of mind task. *J. Exp. Psychol. Gen.* **143**, 972–979.
- Dalal, S. J., Estep, J. S., Valentin-Bon, I. E. and Jerse, A. E. (2001). Standardization of the Whitten Effect to induce susceptibility to *Neisseria gonorrhoeae* in female mice. *Contemp. Top. Lab. Anim. Sci.* **40**, 13–17.

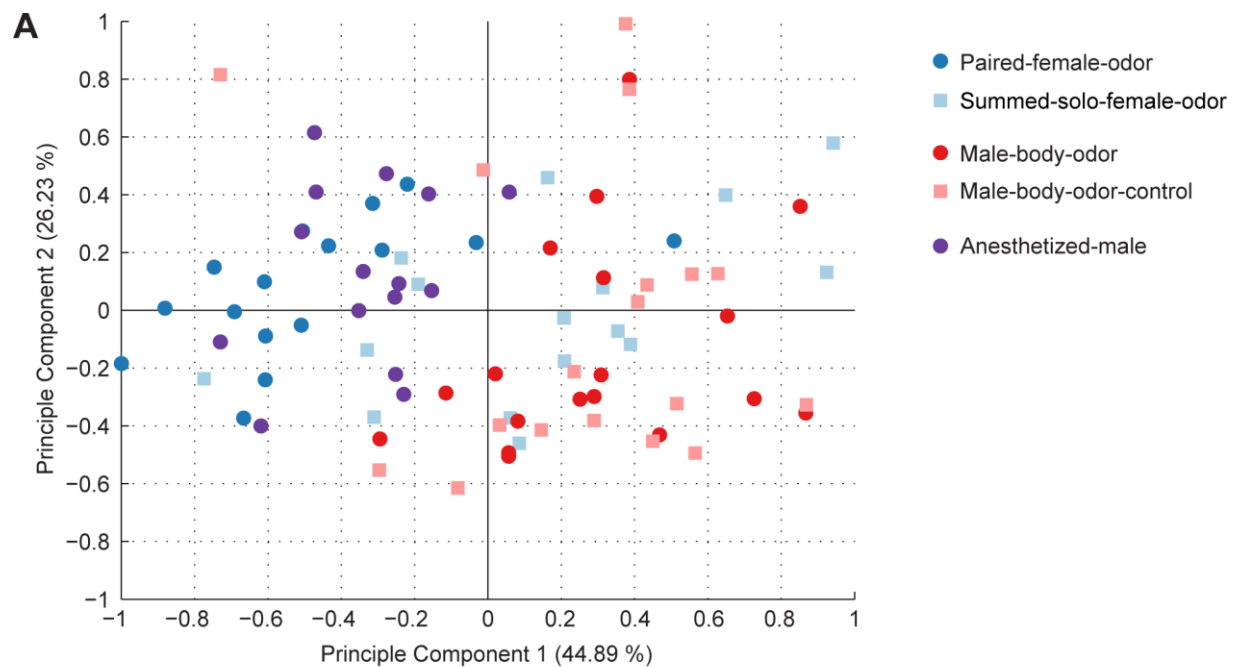
- Dizinno, G., Whitney, G. and Nyby, J.** (1978). Ultrasonic vocalizations by male mice (*Mus musculus*) to female sex pheromone: experiential determinants. *Behav. Biol.* **22**, 104-113.
- Ehret, G.** (1975). Frequency and intensity difference limens and nonlinearities in the ear of the housemouse (*Mus musculus*). *J. Comp. Physiol. A* **102**, 321-336.
- Fernald, R. D. and Hirata, N. R.** (1977). Field study of *Haplochromis burtoni*: quantitative behavioural observations. *Anim. Behav.* **25**, 964-975.
- Ferrero, D. M., Moeller, L. M., Osakada, T., Horio, N., Li, Q., Roy, D. S., Cichy, A., Spehr, M., Touhara, K. and Liberles, S. D.** (2013). A juvenile mouse pheromone inhibits sexual behaviour through the vomeronasal system. *Nature* **502**, 368-371.
- Fisher, H. S. and Rosenthal, G. G.** (2007). Male swordtails court with an audience in mind. *Biol. Lett.* **3**, 5-7.
- Gerhardt, H. C. and Huber, F.** (2002). Acoustic competition and alternative tactics. In *Acoustic Communication in Insects and Anurans* (H. C. Gerhardt and F. Huber), pp. 287-326. Chicago: University of Chicago Press.
- Gyger, M., Karakashian, S. J. and Marler, P.** (1986). Avian alarm calling: is there an audience effect? *Anim. Behav.* **34**, 1570-1572.
- Haley, K. J. and Fessler, D. M. T.** (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* **26**, 245-256.
- Hammerschmidt, K., Radyushkin, K., Ehrenreich, H. and Fischer, J.** (2009). Female mice respond to male ultrasonic "songs" with approach behaviour. *Biol. Lett.* **5**, 589-592.
- Hammerschmidt, K., Radyushkin, K., Ehrenreich, H. and Fischer, J.** (2012). The structure and usage of female and male mouse ultrasonic vocalizations reveal only minor differences. *PLoS ONE* **7**, e41133.
- Karim, B. O., Landolfi, J. A., Christian, A., Ricart-Arbona, R., Qiu, W., McAlonis, M., Eyabi, P. O., Khan, K. A., Dicello, J. F., Mann, J. F. et al.** (2003). Estrous cycle and ovarian changes in a rat mammary carcinogenesis model after irradiation, tamoxifen chemoprevention, and aging. *Comp. Med.* **53**, 532-538.
- Klump, G. M. and Gerhardt, H. C.** (1987). Use of non-arbitrary acoustic criteria in mate choice by female gray tree frogs. *Nature* **326**, 286-288.
- Mackintosh, J. H. and Grant, E. C.** (1966). The effect of olfactory stimuli on the agonistic behaviour of laboratory mice. *Z. Für Tierpsychol.* **23**, 584-587.
- 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. Off. J. Soc. Neurosci.* **33**, 5573-5583.
- Marler, P.** (1955). Studies of fighting in chaffinches (1) behaviour in relation to the social hierarchy. *Br. J. Anim. Behav.* **3**, 111-117.
- Marler, P., Dufty, A. and Pickert, R.** (1986). Vocal communication in the domestic chicken: II. Is a sender sensitive to the presence and nature of a receiver? *Anim. Behav.* **34**, 194-198.
- Matos, R. J. and McGregor, P. K.** (2002). The effect of the sex of an audience on male-male displays of siamese fighting fish (*Betta splendens*). *Behaviour* **139**, 1211-1221.
- Matos, R. J. and Schlupp, I.** (2005). Performing in front of an audience: signalers and the social environment. In *Animal Communication Networks* (ed. P. K. McGregor), pp. 63-83. Cambridge: Cambridge University Press.
- McGregor, P. K.** (1992). *Playback and Studies of Animal Communication*, 1992 edn. New York: Springer.
- McGregor, P. K.** (ed.) (2005). *Animal Communication Networks*. Cambridge University Press. Cambridge: Cambridge University Press.
- McLean, A. C., Valenzuela, N., Fai, S. and Bennett, S. A. L.** (2012). Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. *J. Vis. Exp.* e4389.
- Musolf, K., Hoffmann, F. and Penn, D. J.** (2010). Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. *Anim. Behav.* **79**, 757-764.
- Nyby, J., Dizinno, G. A. and Whitney, G.** (1976). Social status and ultrasonic vocalizations of male mice. *Behav. Biol.* **18**, 285-289.
- Nyby, J., Wysocki, C. J., Whitney, G., Dizinno, G. and Schneider, J.** (1979). Elicitation of male mouse (*Mus musculus*) ultrasonic vocalizations: I. Urinary cues. *J. Comp. Physiol. Psychol.* **93**, 957-975.
- Panksepp, J. B., Jochman, K. A., Kim, J. U., Koy, J. J., Wilson, E. D., Chen, Q., Wilson, C. R. and Lahvis, G. P.** (2007). Affiliative behavior, ultrasonic communication and social reward are influenced by genetic variation in adolescent mice. *PLoS ONE* **2**, e351.
- Portfors, C. V.** (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J. Am. Assoc. Lab. Anim. Sci.* **46**, 28-34.
- Prestwich, K. N.** (1994). The energetics of acoustic signaling in anurans and insects. *Am. Zool.* **34**, 625-643.
- Rich, T. J. and Hurst, J. L.** (1998). Scent marks as reliable signals of the competitive ability of mates. *Anim. Behav.* **56**, 727-735.
- Ryan, M. J.** (1985). *The Tungara Frog: A Study in Sexual Selection and Communication*. Chicago, IL: University of Chicago Press.
- Sales, G. D.** (1972). Ultrasound and aggressive behaviour in rats and other small mammals. *Anim. Behav.* **20**, 88-100.
- Scattoni, M. L., Gandhi, S. U., Ricceri, L. and Crawley, J. N.** (2008). Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PLoS ONE* **3**, e3067.
- Schjelderup-Ebbe, T.** (1922). Beiträge zur sozialpsychologie des haushuhns. *Z. Für Psychol.* **88**, 225-252.
- Shepard, K. N. and Liu, R. C.** (2011). Experience restores innate female preference for male ultrasonic vocalizations. *Genes Brain Behav.* **10**, 28-34.
- Stoddard, P. K., Beecher, M. D. and Willis, M. S.** (1988). Response of territorial male song sparrows to song types and variations. *Behav. Ecol. Sociobiol.* **22**, 125-130.
- Stoddard, P. K., Beecher, M. D., Horning, C. L. and Campbell, S. E.** (1991). Recognition of individual neighbors by song in the song sparrow, a species with song repertoires. *Behav. Ecol. Sociobiol.* **29**, 211-215.
- Stowers, L., Holy, T. E., Meister, M., Dulac, C. and Koentges, G.** (2002). Loss of sex discrimination and male-male aggression in mice deficient for TRP2. *Science* **295**, 1493-1500.
- Taigen, T. L. and Wells, K. D.** (1985). Energetics of vocalization by an anuran amphibian (*Hyla versicolor*). *J. Comp. Physiol. B* **155**, 163-170.
- Thomson, D. J.** (1982). Spectrum estimation and harmonic analysis. *Proc. IEEE* **70**, 1055-1096.
- Van Oortmerssen, G. A.** (1971). Biological significance, genetics and evolutionary origin of variability in behaviour within and between inbred strains of mice (*Mus musculus*). *Behaviour* **38**, 1-91.
- Vignal, C., Mathevon, N. and Mottin, S.** (2004). Audience drives male songbird response to partner's voice. *Nature* **430**, 448-451.
- von Merten, S., Hoier, S., Pfeifle, C. and Tautz, D.** (2014). A role for ultrasonic vocalisation in social communication and divergence of natural populations of the house mouse (*Mus musculus domesticus*). *PLoS ONE* **9**, e97244.
- Whitney, G., Coble, J. R., Stockton, M. D. and Tilson, E. F.** (1973). Ultrasonic emissions: do they facilitate courtship of mice? *J. Comp. Physiol. Psychol.* **84**, 445-452.
- Whitten, W. K.** (1956). Modification of the oestrous cycle of the mouse by external stimuli associated with the male. *J. Endocrinol.* **13**, 399-404.
- Wilk, M. B. and Gnanadesikan, R.** (1968). Probability plotting methods for the analysis of data. *Biometrika* **55**, 1-17.
- Wöhr, M. and Schwarting, R. K. W.** (2008). Ultrasonic calling during fear conditioning in the rat: no evidence for an audience effect. *Anim. Behav.* **76**, 749-760.

SUPPLEMENTAL FIGURES



Supplemental Figure 1. Methods. (A) Schematic side view of arena used in Experiments 1-4 and 6. (B) Schematic top view of arena used in Experiment 5. (C) Example slides of vaginal cells collected in proestrus, estrus, and metestrus. Proestrus was characterized as a mixture of basal and squamous cells, with few neutrophils, and estrus was characterized as a large

proportion of squamous cells, with few basal cells or neutrophils. Metestrus and diestrus were combined together due to few observations of diestrus and were characterized as a mixture of basal cells, squamous cells, and neutrophils. (D) Example of overlapping vocalizations from two mice (see arrow). (E) Correlations between measured characteristics of acoustic structure in the Summed-solo-female-odor (Sum) and Paired-female-odor (Pair) conditions (Dur = duration, F = frequency, BW = bandwidth, FM seg. = frequency-modulated segments). (F) Frequency-modulated (FM) segment examples. Black circles are original data points from the frequency contour, and red circles are the interpolated and smoothed contour. Green lines show the frequency modulation of frequency-modulated segments and blue markers show where the frequency modulation changed sign.



B

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
% Variance explained	44.89%	26.23%	13.14%	8.74%	3.80%	3.14%	0.06%
Duration	-0.50	-0.16	-0.06	-0.11	0.41	0.74	0.01
Frequency BW	0.15	-0.65	0.14	0.14	0.63	-0.36	0.04
Mean Frequency	0.24	0.06	0.91	-0.26	-0.02	0.22	-0.01
Complex	-0.51	-0.18	0.13	-0.32	-0.23	-0.31	0.66
Upsweeps	0.52	-0.17	-0.15	0.26	-0.14	0.41	0.66
Downsweeps	-0.37	0.12	0.34	0.85	-0.08	-0.02	0.05
Flats	0.06	0.69	-0.02	-0.03	0.60	-0.15	0.36

Supplemental Figure 2. Principle component analysis of acoustic structure features, complexity, and simple syllable types. Duration and downsweep data were log transformed to make them more normally distributed (a constant of 0.0001 to each downsweep data point to avoid zero values). (A) Principle components 1 and 2 plotted against one another. (B) Component loadings and % variance explained for all seven principle components. Duration = vocalization duration at the 92nd percentile. Frequency BW = frequency bandwidth at the 50th percentile. Mean Frequency = vocalization mean frequency at the 50th percentile. Complex = proportion complex vocalizations. Upsweeps = proportion upsweep vocalizations. Downsweeps = proportion downsweep vocalizations. Flats = proportion flat vocalizations.

Supplemental Table 1. Results of a generalized linear mixed model examining the effects of odor condition and social context on USV count. The response variable was the vocalization counts from the Summed-solo-no-odor, Paired-no-odor, Summed-solo-female-odor, and Paired-female-odor conditions (USV_count), while Odor_condition (no-odor versus female-odor) and Social_context (Summed-solo versus Paired) were fixed factors. Subject ID (n = 20 subjects) was coded as a random factor. The formula given is the model specification formula written in Wilkinson's notation. The comparison groups are also written using Wilkinson's notation.

Formula: USV_count ~ 1 + Odor_condition*Social_context + (1 | Subject1) + (1 | Subject2) + ... + (1 | Subject20)

Comparison groups	Statistical test	α	Test statistic	p-value
Odor_condition	F-test	0.05	F = 3698.6	1.227×10^{-63}
Social_context	F-test	0.05	F = 277.85	2.015×10^{-26}
Odor_condition:Social_context	F-test	0.05	F = 17.27	8.817×10^{-5}

Supplemental Table 2. Results of a generalized linear mixed model investigating the effect of social context on principle component 1 from the acoustic structure PCA. Social context was a fixed factor, Subject ID was a random factor, and Principle Component 1 was the response variable. All five social contexts from the principle component analysis were included (52 unique subjects). The formula given is the model specification formula written in Wilkinson's notation.

Formula: PrincipleComponent1 ~ 1 + Social_context + (1 | Subject1) + (1 | Subject2) + ... + (1 | Subject52)

Comparison groups	Statistical test	α	Test statistic	p -value
Social_context	F -test	0.05	$F = 26.79$	5.391×10^{-14}
Paired-female-odor vs. Summed-solo-female-odor	F -test	0.05	$F = 51.69$	3.427×10^{-10}
Male-body-odor vs. Male-body-odor-control	F -test	0.05	$F = 0.175$	0.677
Anesthetized-male vs. Summed-solo-female-odor	F -test	0.05	$F = 34.94$	8.493×10^{-8}
Anesthetized-male vs. Paired-female-odor	F -test	0.05	$F = 0.122$	0.728

Supplemental Table 3. Results of a generalized linear mixed model investigating the effect of social context on principle component 2 from the acoustic structure PCA. Social context was a fixed factor, Subject ID was a random factor, and Principle Component 2 was the response variable. All five social contexts from the principle component analysis were included (52 unique subjects). The formula given is the model specification formula written in Wilkinson's notation.

Formula: PrincipleComponent2 ~ 1 + Social_context + (1 | Subject1) + (1 | Subject2) + ... + (1 | Subject52)

Comparison groups	Statistical test	α	Test statistic	p -value
Social_context	<i>F</i> -test	0.05	$F = 13.00$	3.686×10^{-8}
Paired-female-odor vs. Summed-solo-female-odor	<i>F</i> -test	0.05	$F = 1.127$	0.292
Male-body-odor vs. Male-body-odor-control	<i>F</i> -test	0.05	$F = 1.489$	0.226
Anesthetized-male vs. Summed-solo-female-odor	<i>F</i> -test	0.05	$F = 2.417$	0.124
Anesthetized-male vs. Paired-female-odor	<i>F</i> -test	0.05	$F = 0.386$	0.536