

## CORRECTION

# Correction: Mouse vocal emission and acoustic complexity do not scale linearly with the size of a social group

Megan R. Warren, Morgan S. Spurrier, Daniel T. Sangiamo, Rachel S. Clein and Joshua P. Neunuebel

There were several errors in *J. Exp. Biol.* (2021) **224**, jeb239814 (doi:10.1242/jeb.239814).

In Fig. 4A, the label for the light blue section for the Continuous, 1-Mouse pie chart was missing and now reads 2%; the dark blue section of the 1-Jump, 4-Mouse chart is now labelled 13% rather than 12%.

In Fig. 5A, the label for the green section of the Continuous, 4-Mouse chart was missing and now reads 49%.

In Fig. 5D the *y*-axis label is now labelled 0.03 rather than 0.3.

All versions of the article have been updated.

We apologise to the authors and readers for these errors, which were introduced during figure preparation, and for any inconvenience caused.

## RESEARCH ARTICLE

# Mouse vocal emission and acoustic complexity do not scale linearly with the size of a social group

Megan R. Warren<sup>1,2</sup>, Morgan S. Spurrier<sup>1</sup>, Daniel T. Sangiamo<sup>1</sup>, Rachel S. Clein<sup>1</sup> and Joshua P. Neunuebel<sup>1,\*</sup>

## ABSTRACT

Adult mice emit ultrasonic vocalizations (USVs), sounds above the range of human hearing, during social encounters. While mice alter their vocal emissions between isolated and social contexts, technological impediments have hampered our ability to assess how individual mice vocalize in group social settings. We overcame this challenge by implementing an 8-channel microphone array system, allowing us to determine which mouse emitted individual vocalizations across multiple social contexts. This technology, in conjunction with a new approach for extracting and categorizing a complex, full repertoire of vocalizations, facilitated our ability to directly compare how mice modulate their vocal emissions between isolated, dyadic and group social environments. When comparing vocal emission during isolated and social settings, we found that socializing male mice increase the proportion of vocalizations with turning points in frequency modulation and instantaneous jumps in frequency. Moreover, males change the types of vocalizations emitted between social and isolated contexts. In contrast, there was no difference in male vocal emission between dyadic and group social contexts. Female vocal emission, while predominantly absent in isolation, was also similar during dyadic and group interactions. In particular, there were no differences in the proportion of vocalizations with frequency jumps or turning points. Taken together, the findings lay the groundwork necessary for elucidating the stimuli underlying specific features of vocal emission in mice.

**KEY WORDS:** Animal communication, Mouse ultrasonic vocalizations, Social behavior, Sound source localization

## INTRODUCTION

Vocalizations play an important role across the animal kingdom, communicating personal information including physical attributes (Gamba et al., 2012; Ji et al., 2013; Stoeger and Baotic, 2016) and physiological state (Knutson et al., 2002; Sehrsweeney et al., 2019), as well as survival-related information such as the presence of predators (da Silva et al., 2002; Seyfarth et al., 1980). Animals emit vocalizations with varying levels of acoustic complexity. The acoustic complexity of vocalizations ranges from simple continuous sounds, with frequencies that increase, decrease or remain constant, to complex sounds with elements such as instantaneous jumps in frequency or reversals in the direction of frequency modulation (Behr and von Helversen, 2004; Bradbury and Vehrencamp, 1998;

White and White, 1970; Zuberbühler et al., 1997). By varying acoustic complexity, signaling animals may alter the information content of the emitted auditory cues and potentially change the behavior of conspecific receivers (Kershenbaum et al., 2016). Consequently, animals are thought to modulate vocal activity and acoustic complexity as a function of their social context. Male zebra finches (*Taeniopygia guttata*), for instance, sing differently when a female is present than when alone (Chen et al., 2016; Kao and Brainard, 2006). The golden rocket frog (*Anomaloglossus beebei*) emits vocalizations with greater acoustic complexity during courtship behaviors compared with those produced during aggressive behaviors (Pettitt et al., 2012). The yellow mongoose (*Cynictic penicillate*) produces calls to indicate the presence of a predator when group members are present, but not in isolation (le Roux et al., 2008). Thus, across numerous species, animals regulate their vocal emissions based upon their social environment.

The house mouse (*Mus musculus*) also modulates ultrasonic vocal emission across different social contexts, emitting distinct types of vocalizations during opposite-sex interactions and non-social conditions (Hanson and Hurley, 2012; Yang et al., 2013). Mouse ultrasonic vocalizations (USVs), signals ranging in frequency from 30 to 110 kHz (Gourbal et al., 2004; Holy and Guo, 2005), are predominantly emitted during opposite-sex encounters (Wysocki et al., 1982), and these signals are typically assumed to be produced by males (Warburton et al., 1989; Whitney et al., 1973). The features of mouse USVs are context dependent, differing in acoustic complexity (as measured by the percent of vocalizations with instantaneous jumps in frequency), vocalization type and spectral-temporal features across discrete behavioral contexts (Gaub et al., 2016; Grimsley et al., 2016; Hammerschmidt et al., 2012). Moreover, mice produce complex, multi-syllabic vocalizations containing frequency jumps when socially interacting (Matsumoto and Okanoya, 2018; Miller and Engstrom, 2007; Scattoni et al., 2008; Weiner et al., 2016) and USV emission changes over the course of interaction (Matsumoto and Okanoya, 2018). However, direct knowledge of how individual mice alter their vocal emissions across different social contexts is lacking. One study found that males produce more complex vocalizations in the presence of a male listener than when alone (Seagraves et al., 2016). Other work has found that vocalizations in a mixed-sex dyad differ from vocalizations emitted when a male is in isolation (Chabout et al., 2015; Hanson and Hurley, 2012). These studies, however, typically employed a dyadic social encounter, thus not determining the role of group size in vocal emission, and also assumed that all vocalizations were emitted by the male mouse. Recent work, however, revealed that female mice emit USVs while interacting with males (Heckman et al., 2017; Neunuebel et al., 2015; Sangiamo et al., 2020; Warren et al., 2018b), indicating that all socializing mice are capable of emitting USVs. As such, to fully elucidate the role social engagement plays in altering mouse vocal emissions, the vocal behavior of each mouse needs to be individually tracked across different social contexts.

<sup>1</sup>Department of Psychological and Brain Sciences, University of Delaware, Newark, DE 19716, USA. <sup>2</sup>Department of Biology, Emory University, Atlanta, GA 30322, USA.

\*Author for correspondence (jneun@udel.edu)

 J.P.N., 0000-0002-4894-4364

In the current study, we combined a novel approach to extract a full repertoire of USVs with our previously established sound-source localization system (Warren et al., 2018a). Together, these tools allowed us to pinpoint where individual vocalizations originated and accurately assign vocalizations to their respective emitter. We tracked the vocal behavior of individual mice in three social contexts: (1) an isolated context involving a single male or female, (2) dyadic social interactions consisting of a male and a female, and (3) group social interactions comprising 2 males and 2 females. We directly showed that male mice emitted more acoustically complex vocalizations (i.e. the proportion of vocalizations with frequency jumps and turning points in frequency modulation) and altered the types of vocalizations they emitted in the presence of a social partner compared with during periods of isolation. Male vocal emission was unaltered by increasing the number of social partners from 1 to 3, as vocal emission was indistinguishable between dyadic and group social interactions. Similarly, female mice showed no vocal differences between dyadic and group encounters. Together, our results indicate that while male and female mice increase the complexity of their vocal repertoire between isolated and social contexts, complexity does not scale linearly with the size of a mouse's current social group.

## MATERIALS AND METHODS

### Subjects

Adult mice (age 2–5 months) on a B6 background were used to examine acoustic complexity across multiple social contexts. In the first context, isolated males were exposed to female cues (previously described in Warren et al., 2018a;  $n=19$  males) or isolated females were exposed to male cues ( $n=35$  females). The second context consisted of a male and female pairing (previously described in Warren et al., 2018b;  $n=10$  males and 8 females). The third context consisted of groups of mice, two males and two females (previously described in Sangiamo et al., 2020;  $n=22$  males and 22 females). All mice were raised in a colony in the Life Sciences Research Facility at the University of Delaware. Colony founders were purchased from Jackson Laboratory (Bar Harbor, ME). At 3 weeks old, mice were weaned, group-housed by sex (maximum of four per cage) and implanted with light-activated microtransponders (*P*-Chip injector, PharmaSeq Inc., Monmouth Jct, NJ) for identification purposes. Mice were housed in cages containing ALPHA-dri bedding (Animal Specialties and Provisions, LLC, Watertown, TN) and environmental enrichment, and allowed *ad libitum* access to food and water.

Individually recorded males were housed in isolation ( $n=2$ ), with same-sex littermates ( $n=14$ ), or in a breeding pair ( $n=3$ ). Females recorded in isolation and all mice used in a social encounter were isolate-housed for at least 2 weeks before recordings to minimize group-housing effects on social behavior (Hilakivi-Clarke and Lister, 1992; Jones and Nowell, 1989; Konig, 1994). All mice were maintained on a 12 h:12 h light:dark cycle and experiments were conducted during the dark phase of the light cycle.

All experiments were conducted at the University of Delaware in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The University of Delaware Animal Care and Use Committee approved all experimental protocols (protocol number: 1275).

### Experimental setup

To make mice visually distinct, the fur of each mouse recorded in a group social encounter was bleached in a unique pattern with hair dye (Clairol Born Blonde, Proctor and Gamble, Cincinnati, OH;

Ohayon et al., 2013). Patterns were two vertical stripes, two horizontal stripes, one slash or five dots. If the dye faded over time, mice were repainted with the same pattern. For dyadic recordings, all males were painted with a five-dot pattern; females remained unpainted. For individually recorded males, as mice did not need to be visually distinguishable, hair dyeing depended on whether that male was being prepared for use in another experiment. All individually recorded females were first used in dyadic or group recordings; therefore, patterns depended on the experimental condition. No mice were used in both dyadic and group recordings.

USV emission is enhanced by previous experience with the opposite sex (Arriaga et al., 2012). Therefore, all mice recorded with a social partner were exposed to a single animal of the opposite sex for 10 min 1 day after males were first marked with hair dye. Opposite-sex exposures were conducted in 11 singly recorded males and every singly recorded female. Opposite-sex exposures occurred in a clean cage with no bedding, and sessions were terminated before successful copulation or after 10 min. Animals used as opposite-sex stimuli were never used in behavioral recordings.

Males are typically non-vocal in isolation unless female scent cues are present (Guo and Holy, 2007; Musolf et al., 2010). Thus, for males recorded without a social partner, female scent cues were introduced into the arena by allowing a female to explore the environment for 3–5 min and/or by introducing freshly soiled bedding from the cage of a female prior to recording (Warren et al., 2018a). For isolated female recordings, male scent cues were provided by having the male in the arena prior to the recording. Recordings of individual animals lasted for 10 min.

Females were only recorded in the estrus stage of the estrous cycle, as determined by non-invasive lavage (Sangiomo et al., 2020; Warren et al., 2018b). Lavages were performed 30 and 120 min before a potential recording for dyadic and group interactions, respectively. The vaginal cavity was flushed with 30  $\mu$ l of saline, the resultant solution was placed on a microscope slide, and the cells were stained with crystal violet. Pictures of the cells were taken using a camera (World Precision Instruments, cat. #USBCAM50) attached to a microscope (VWR, cat. #89404-890) via a coupler (World Precision Instruments, cat. #501381). Estrous stage was determined by assessing the observed cell types, with estrus defined as the majority of cells being anucleated, cornified squamous epithelial cells (Cora et al., 2015). Mice were recorded with each social partner only once. Mice in dyadic recordings participated in multiple interactions (Warren et al., 2018b).

All experiments were conducted in an anechoic chamber while audio and video data were recorded. Audio data was sampled by an 8-channel microphone array (microphones from Avisoft-Bioacoustics, Glienicke, Germany; cat. #CM16/CMPA40-5V) using equipment from National Instruments (Austin, TX; cat. #PXIe-1073, PXIe-6356, BNC-2110). Each microphone (model 3384, Krohn-Hite, Brockton, MA) sampled at 250,000 Hz and was low pass filtered at 200 kHz. Microphones were surrounded by rings of LED lights to visually determine microphone position prior to recordings. Video data was recorded via a single overhead camera (FLIR, Richmond, BC, Canada; cat. #GS3-U3-41C6M-C) using BIAS software (<https://bitbucket.org/iorodeo/bias/downloads/>) developed by Kristin Branson, Alice Robie, Michael Reiser and Will Dickson. The camera was triggered at 30 Hz via a counter pulse sent from the PXIe-6356 hardware. The triggering pulse was sent to both the camera and the National Instruments equipment concurrently through a BNC splitter to facilitate alignment of the audio and video data. Custom-written MATLAB v. 2014b software (MathWorks, Natick, MA) was used to control all recording

devices. All data were stored on a PC (Hewlett-Packard, Palo Alto, CA).

Recordings were conducted in a cage with walls constructed of mesh (McMaster-Carr, Robbinsville, NJ; cat. #9318T25) with a frame of extruded aluminium (cuboid arena: 76.2×76.2×61 cm width×length×height; cylindrical arena: 68.6×91.4 cm, diameter×height; 80/20 Inc., Columbia City, IN) surrounded by Sonex foam (Pinta Acoustic Inc., Minneapolis, MN; cat. #VLW-35). Recordings were conducted in the dark, with infrared lights (GANZ, Cary, NC; cat. # IR-LT30) positioned above the cage. The arena floor was covered with ~1.5 cm of ALPHA-dri bedding to enhance contrast between the mice and the cage floor. A ruler, placed in the center of the arena, was used to convert camera pixels into meters. Two 15 s pre-tests were run before each recording. In the first, the LEDs surrounding each of the microphones were illuminated and used to determine the position of each microphone. In the second, the LEDs were turned off and the overhead infrared lights were turned on to confirm the focus of the camera. Following the two pre-tests, the ruler was removed while the infrared lights remained on for the duration of each recording. Recordings lasted 10, 30 and 300 min for the mice recorded in isolation, pairs and groups, respectively. To control for potential temporal differences in vocal emission, all analyses focused upon the first 10 min of recording.

### Data processing

A data analysis pipeline was created on the University of Delaware's Farber computer cluster. The pipeline was used to determine each mouse's trajectory, to localize the source of individual vocalizations and to extract information about the vocalizations assigned to each animal.

### Video tracking

Machine learning approaches were used to automatically identify and track the position of each mouse (Motr; Ohayon et al., 2013). For every video frame, an ellipse was fitted around each recorded mouse, and the *x*- and *y*-positions of their bodies, their length and width, and their heading direction were calculated. Motr used the dyed patterns to identify individual mice within group recordings. Tracking accuracy was manually inspected in MATLAB.

### Vocal extraction

Extracting and assigning vocalizations to individual mice required four major steps, outlined below.

#### Step 1: audio segmentation

Multi-taper spectral analysis was used to automatically extract continuous stretches of sound (Seagraves et al., 2016; Warren et al., 2018a) from the audio recordings across the 8 microphone channels. All data were bandpass filtered between 30 and 110 kHz. Then, temporally overlapping segments were Fourier transformed via discrete prolate spheroidal sequences as window functions ( $K=5$ ,  $NW=3$ ). Each time-frequency point was compared with noise using an *F*-test (Percival and Walden, 1993;  $P<0.05$ ). This was repeated on multiple segment lengths to capture a range of spectral and temporal scales (non-equispaced fast Fourier transform,  $NFFT=64$ , 128 and 256). Data were combined into a single spectrogram and convolved with a square box (11 pixels in frequency by 15 in time) to fill in small gaps before continuous stretches of sound, containing a minimum of 1500 pixels and lasting a minimum of 5 ms, were extracted. The minimum duration of 5 ms was chosen as localization accuracy plateaued for vocal signals lasting  $\geq 5$  ms (data not shown; replicating work from Neunuebel et al., 2015). Extracted

signals were stored as traces of frequency over time or frequency contours.

#### Step 2: consolidating continuous signals

Mice are known to emit vocalizations containing frequency jumps or instantaneous jumps between two different frequencies. Owing to this discontinuity in frequency, each unique segment would be extracted as a unique signal using our previous methods (Warren et al., 2018a). To more accurately characterize mouse USV complexity, we aimed to extract vocalizations, including complex vocalizations containing one or more frequency jumps. Therefore, in vocal recordings from isolated males, independent viewers manually assessed whether successive vocal signals were part of a single vocalization or from different vocalizations. The independent viewers were four different members of the lab, each with at least a year of experience studying mouse ultrasonic vocalizations, and the inter-rater reliability was  $96.4\pm 1.5\%$  (mean $\pm$ s.d.), with a minimum value of 94.4%. Based on the assessment of the independent viewers, signals separated by less than 13 ms in time (start time of the second signal minus end time of the first signal) were consistently categorized as segments of a single vocalization that had been segmented into multiple parts. These vocalizations contained both continuous and complex components.

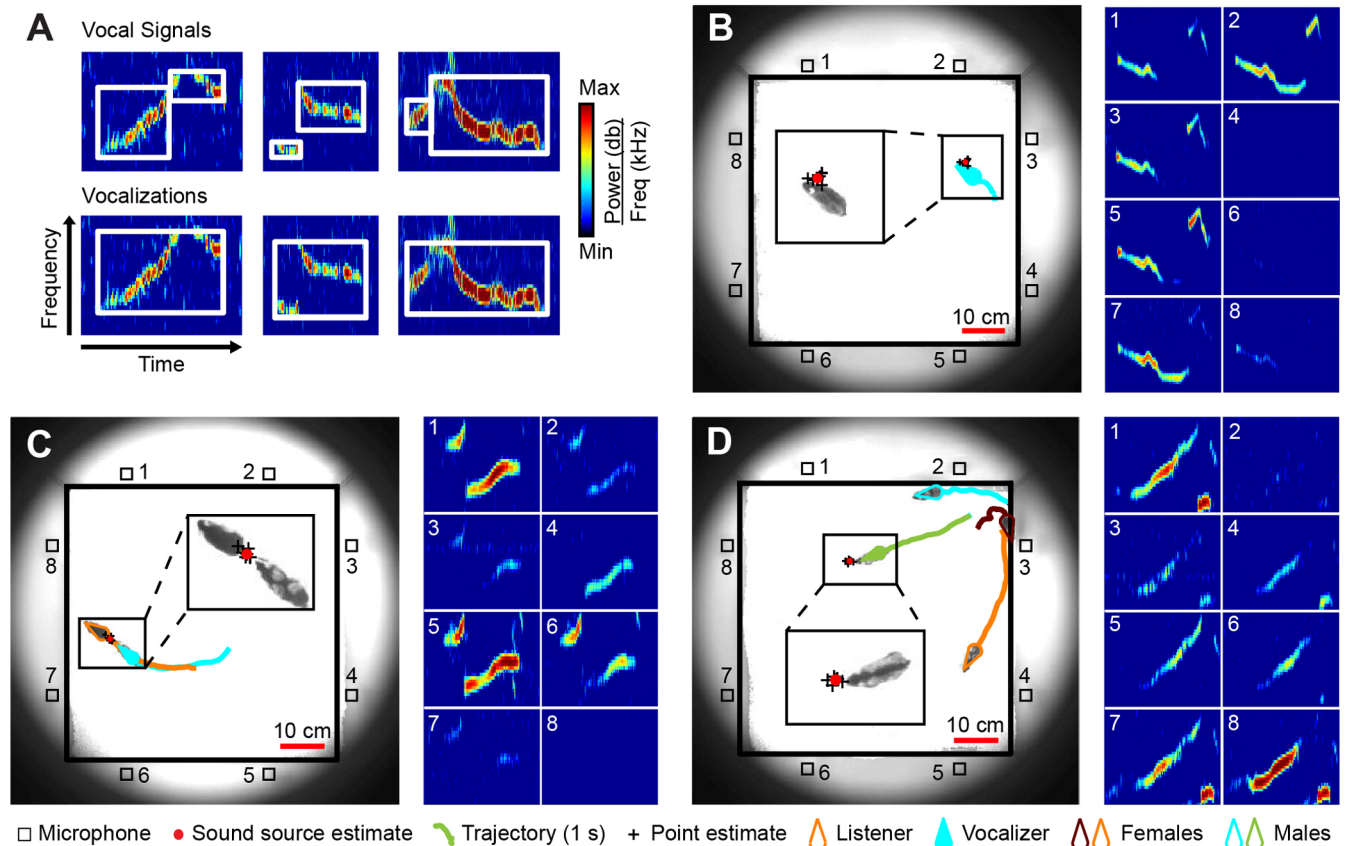
When an animal emitted a continuous vocalization, we assumed, *a priori*, that the whole sound came from a single animal. Thus, we could confidently assign the signal to a single animal prior to estimating the source location of each component. However, since complex signals can consist of multiple discontinuous sounds, it is possible that instead of being a single complex vocalization, the individual pieces of sound were actually unique vocalizations emitted by different animals. As such, consolidating these segments prior to determining which animal emitted each segment has the potential to inaccurately pinpoint the vocalizer and represent the vocal data. To avoid this potential pitfall, we asked trained viewers to assess vocal data from a single male and determine whether or not pairs of successive signals appeared to be single complex vocalizations containing frequency jumps. We found that successive signals separated by greater than 10 kHz (start frequency of the second signal–end frequency of the first signal) were consistently categorized as vocalizations containing a frequency jump.

Based on the findings of our trained viewers, two criteria needed to be met to collapse vocal signals into vocalizations. First, successive signals had to be separated by fewer than 13 ms. Second, successive signals needed to differ in frequency by fewer than 10 kHz. If successive signals met both of these criteria, they were consolidated into vocalizations. Signals falling within the temporal threshold but exceeding the frequency threshold were kept as independent elements until after we could discern which mouse emitted each element.

#### Step 3: sound source localization and signal assignment

Localizing each vocal signal relied upon a jack-knife resampling method (Warren et al., 2018a). Briefly, for each extracted sound, the audio data from one microphone was omitted and the location of emission was estimated by assessing the time differences in when the sound arrived at each of the 7 remaining microphones. This process was repeated 7 additional times, omitting a different microphone each time, to generate a total of 8 estimated locations, or point-estimates (Fig. 1B–D). The point-estimates, in conjunction with the overall estimate (the average of the 8 point-estimates), were then used to calculate a probability density across the floor of the cage, representing the likelihood that the sound originated from any





**Fig. 1. Sound source localization system extracts and assigns vocalizations to individual mice.** (A) Examples of vocal signals (top) and corresponding vocalizations (bottom). (B–D) Images of arena (left) at the time of a vocal emission from (B) one male in isolation, (C) a dyadic interaction between a male and a female, and (D) a group interaction between two males and two females. Point estimate (plus symbol) shows one estimated sound source location when using a unique combination of 7 microphones. For every vocalization, there are 8 estimates generated. The sound source estimate (red dot) denotes the overall estimated position where the vocalization originated. Insets show enlarged images of sound source estimate. Trajectory indicates the position of the mouse over the preceding 1 s. Spectrograms on right correspond to individual microphones in the arena image.

location in the arena. Each animal was assigned the probability density value ( $D$ ) corresponding to the position of their nose, with nose position determined using Motr. These values were used to calculate a mouse probability index (MPI), indicating the probability that the signal was emitted by each animal in the arena, using the following formula:

$$\text{MPI}_n = \frac{D_n}{\sum_{i=1}^M D_i}, \quad (1)$$

where  $n$ =mouse number and  $M$ =the total number of mice. To assign a signal to an individual mouse, the MPI value for that mouse needed to exceed 0.95, indicating a 95% likelihood that the signal was emitted by that specific mouse. For recordings with two or more mice, the MPI values were based upon the locations of the mice within the arena. For isolated recordings, during processing we randomly generated positions for three additional virtual mice. This allowed us to test the system's ability to assign vocalizations to the correct mice when there were four possible mice, while still being certain which mouse emitted each individual vocalization (Fig. S1). This knowledge was used to test the accuracy of our system (described below).

#### Step 4: concatenating complex vocalizations

At this point, all unambiguously continuous vocalizations were assigned to the vocalizing mouse. However, vocal signals meeting

the criteria for complex vocalizations remained separated. Thus, we next determined whether these signals were multiple vocalizations emitted by multiple animals, or instead individual components of a complex vocalization emitted by a single mouse. If successive signals were emitted by a single animal, they were consolidated into a single vocalization. If the two signals were emitted by two different vocalizers, each was extracted as a vocalization. If one of the two signals did not meet the threshold for assignment, meaning we were not 95% sure of the identity of the vocalizer for one of the signals, we assessed the MPI values for the unassigned signal. We temporarily assigned the vocalizer as the mouse with the highest MPI value, or the mouse that was most likely to have emitted that signal. If that mouse matched the vocalizer of the assigned signal, we consolidated the two signals into a single vocalization. If not, we extracted the two signals as two unique vocalizations, one assigned to a mouse and one unassigned. Once this was completed for all pairs of successive signals, all vocalizations lasting <10 ms were removed.

#### Quantifying complexity

Once we had extracted and determined which mouse emitted each vocalization, we next aimed to determine whether acoustic complexity differed across the three social conditions. Only vocalizations assigned to specific animals were included in analyses. For the 1-mouse context, vocalizations assigned to any of the virtual mice were excluded. We therefore quantified

complexity using two different metrics. First, we compared the proportion of vocalizations emitted by individual mice in each context that had zero (i.e. continuous), one, or multiple frequency jumps (Fig. 3). As a secondary approach, we determined the number of vocalizations containing a discrete change in the direction of frequency modulation (i.e. any point in the vocalization where the frequency was increasing and switched to decreasing, or vice versa; Seagraves et al., 2016). If this pattern existed, we considered the vocalization to be complex. Our second method for quantifying complexity required smoothing the vocalizations. Therefore, we first normalized the data such that every vocalization spanned 100 unique time points and had an average frequency of zero. We then applied an envelope function to each vocalization to mark the upper and lower bounds. We averaged the upper and lower envelopes to generate a smooth line through the signal, which was used to assess complexity.

### Vocal clustering: *k*-means

As unique components of vocalizations containing frequency jumps could be separated by as much as 13 ms, some of the complex signals contained temporal gaps, or periods of time with no frequency data. However, our clustering algorithm required all input to be continuous. Therefore, we represented temporal gaps in vocalizations with frequencies of zero. Then, we began the clustering program by splitting all vocalizations into two clusters. Across all vocalizations categorized into cluster one, we found the mean and standard deviation of normalized frequency at each of the 100 time points. This process was repeated for all vocalizations categorized into cluster two. If more than 3% of vocalizations within either of the clusters had more than 25 of their 100 data points falling outside of 2.5 standard deviations from the group mean, we considered that there was too much variability in the shapes of the vocalizations within the cluster. Therefore, we reran the clustering program with one additional cluster, continuing to add one cluster at a time until we found the optimal number of clusters to sufficiently encompass the variability across all vocalizations. With each iteration, we saved the cluster identity of each vocalization so that we could track how clusters changed as we further segmented the data.

### Vocal clustering: elbow function

Once we determined the optimal number of clusters to explain the variability across all of our vocalization data, we separated the vocalizations in each of the clusters into three groups: vocalizations with zero, one, or multiple frequency jumps. As we had saved the cluster identities with each iteration, we knew the cluster identity of each vocalization when there were only two clusters, when there were only three clusters, all the way through the optimal number of clusters. Starting with the continuous vocalizations, we computed the average variability in shape within and between vocalizations from each of the first two clusters. Variability within a cluster was calculated using Euclidean distance to compare the shapes of random pairs of vocalizations within the cluster. Variability between the two clusters was calculated using Euclidean distance to compare the shapes of random pairs of one vocalization from within the cluster of interest and one vocalization from the other cluster. We found the average of all within-cluster distances and all between-cluster distances and generated a ratio of within divided by between. This process was repeated with three through our pre-determined optimal number of clusters to provide us with difference ratios for all possible numbers of clusters with continuous vocalizations.

We fitted a line to the difference ratios across each possible number of clusters (2–23) and applied an elbow function to the result. The elbow function moves along the curve one bisection point (cluster) at a time and fits two lines; one to all points to the left of the bisection and one to all points to the right. The bisection point that minimizes the sum of error for the two fits, or the point of optimal trade-off between within- and between-cluster variability, is then determined. This indicates the optimal number of clusters. This analysis was also applied to both the 1-jump and multi-jump data to determine the optimal number of shapes needed to characterize continuous, 1-jump and multi-jump vocalizations. This methodology allowed us to determine the optimal number of vocal groups at each level of acoustic complexity.

Previous work using our progressive *k*-means clustering method has shown that male mice emit behavior-associated vocal signals, and this pattern holds when applying other clustering methods (Sangiomo et al., 2020). However, a vital aspect of our clustering procedure is the fact that we incorporate information about the vocalizing animal. Here, we added the quantification of structurally complex, multi-syllabic vocalizations, which allows us to attribute entire multi-syllabic vocalizations to individual animals. Thus, with this novel capacity to represent the full range of mouse USVs, we used the term vocalization type as a categorical label for the final clustering output.

### Quantifying cluster similarity

To confirm that our clustering method was not being driven by a specific social context, we next quantified the within- and between-type similarity for each context separately (Fig. S3A–C). Thus, within a context we used a Euclidean distance measure to assess the similarity of vocalizations within a type. We randomly split the vocalizations from a type into two equally sized groups, excluding a single vocalization in the case of an uneven number of vocalizations, and found the distance between each pair of vocalizations (the first vocalization from each half, the second vocalization from each half, etc.).

To assess similarity between types, we used the same Euclidean distance measure, but compared vocalizations from different types. Thus, we first extracted all vocalizations of the type of interest. We then randomly selected that many vocalizations from the remaining vocalizations (i.e. vocalizations from all other types), and found the distance between each pair of vocalizations (the first vocalization from the type of interest versus the first vocalization from the remaining data, the second vocalization from both categories, etc.). This was repeated for all types of vocalizations to determine whether vocalizations of the type of interest were more similar to each other than they were to vocalizations of other types from the same context. This was repeated separately for all three social contexts.

To quantify vocal similarity across social contexts, we employed a similar Euclidean distance measure, but employed the measure across all three contexts (Fig. S3D). To quantify within-type similarity, we found all vocalizations from the type of interest across all three social contexts. We determined which context contained the fewest vocalizations of the type of interest, and randomly extracted that many vocalizations from the type of interest in each context. For example, 297, 598 and 723 type 1 vocalizations were emitted in the 1-mouse, 2-mouse and 4-mouse, respectively, so we randomly selected 297 type 1 vocalizations from each context. We next made pairwise comparisons, comparing the first vocalization of the type of interest between all possible pairs of contexts: 1-mouse versus 2-mouse, 2-mouse versus 4-mouse, and 1-mouse versus 4-mouse. This was repeated for the second vocalization

through the final vocalization, and all resultant distances were averaged to generate a summary value for the type of interest. To generate a between-type comparison, we also randomly selected 297 vocalizations from across each of the other types. Then we compared vocalizations from type 1 emitted in the 1-mouse condition only with type 2 vocalizations from the same context. This was replicated for all 3 contexts and averaged to get a summary value for the difference between type 1 and type 2 vocalizations within a social context. This analysis was repeated for each possible pair of types to determine how similar vocalizations within a type were across contexts compared to how similar vocalizations were within a context but across type.

### Quantifying proportional vocal emission

To determine how mice vocalized across social contexts, we assessed how frequently every mouse emitted each type of vocalization. To control for differences in the number of signals emitted by individual animals, we generated a vocal proportion. Thus, for each mouse, we determined the total number of vocalizations emitted. Then, we found the proportion of each type of vocalization out of the total number of vocalizations emitted (i.e. we calculated the total number of signals that the mouse emitted of each type and divided those values by the total number of signals emitted by that mouse). This was repeated for all mice in all conditions.

### Statistics

All statistical tests were nonparametric. Mann–Whitney tests were used for pairwise comparisons. For group-wise comparisons, a Kruskal–Wallis test was used followed by Dunn–Sidak *post hoc* tests to correct for multiple comparisons. All alpha values were set to 0.05.

### RESULTS

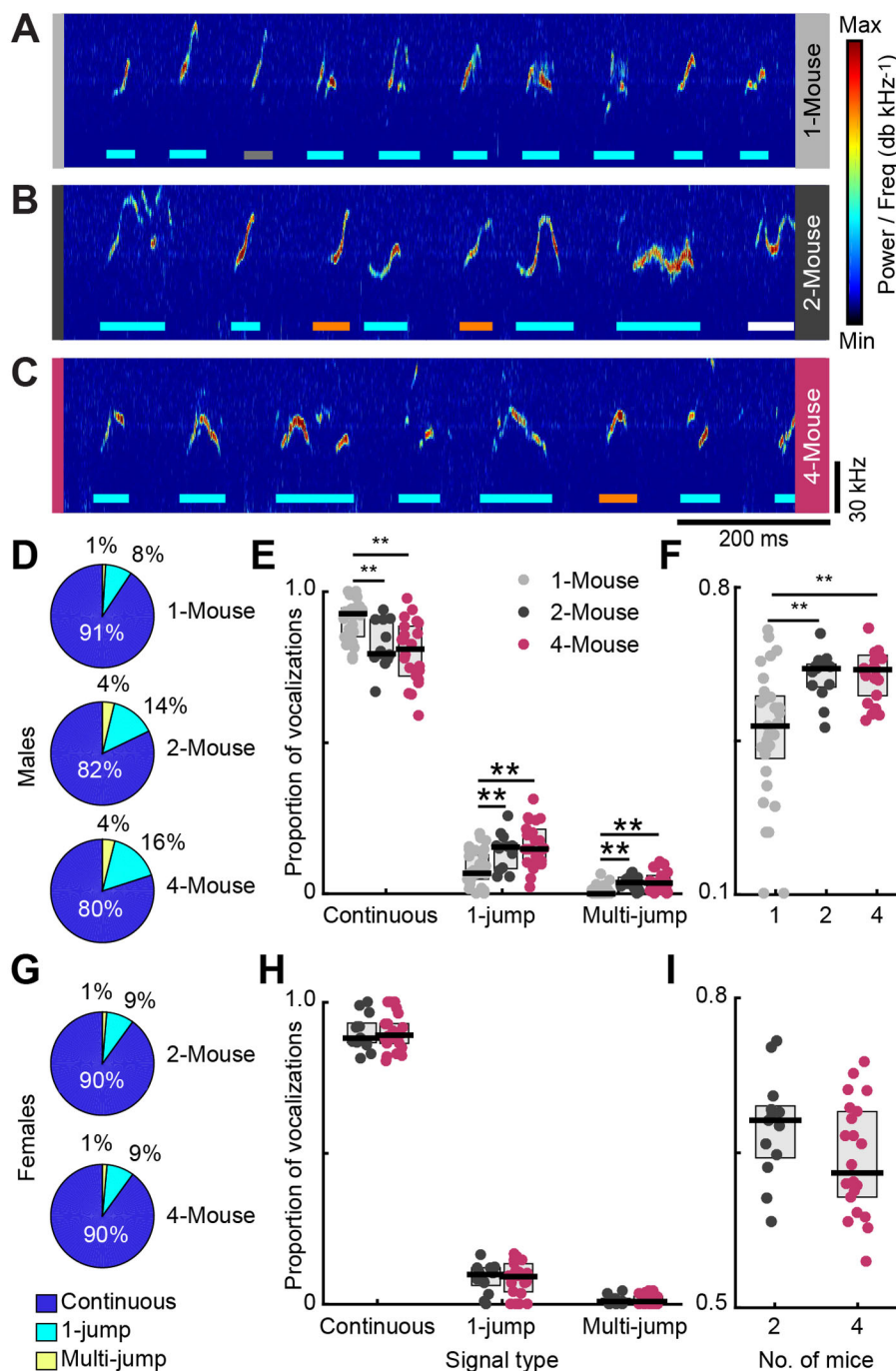
Using an 8-channel microphone array to assign vocalizations (Fig. 1A; Fig. S1) to individual mice, we compared the complexity of male- or female-emitted vocalizations in three different contexts: an isolated context (when mice were exposed to urine and/or other bodily scent cues from the opposite sex), a dyadic context (when one male and one female were present) and a group context (when two males and two females were present). As shown in Fig. 1B–D, male mice vocalized in all three contexts (isolated:  $n=29$  recordings, 3905 assigned vocalizations out of 4641 detected vocalizations; dyadic:  $n=13$  recordings, 12,153 assigned male-emitted vocalizations out of 18,721 detected vocalizations; group:  $n=11$  recordings, 22 males, 15,616 assigned male-emitted vocalizations out of 23,287 detected vocalizations). We therefore aimed to quantify how acoustic complexity differed across social contexts (Fig. 2A–C). Similar to previous work (Arriaga et al., 2012; Chabout et al., 2015, 2016; Holy and Guo, 2005; Matsumoto and Okanoya, 2018; Scattoni et al., 2008; Weiner et al., 2016), we split vocalizations into three groups based on complexity: signals without frequency jumps (continuous), signals with a single jump in frequency (1-jump) or signals with two or more jumps in frequency (multi-jump). We then found the proportion of vocalizations emitted by each male that was classified into each group (for visualization purposes, means averaged across individuals are represented in Fig. 2D). In the isolated condition, 92.73% (IQR=85.12–94.55%) of vocalizations emitted by an individual male were continuous (Fig. 2E), with 6.78% (IQR=4.80–12.86%) having one jump and 0% (IQR=0–1.40%) having multiple jumps. In the dyadic condition, in contrast, only

79.41% (IQR=78.09–90.73%) of vocalizations were continuous, while 15.44% (IQR=8.30–16.65) had one jump and 3.69% (IQR=1.82–5.41%) had multiple jumps. In a group setting, acoustic complexity was similar to the two-mouse condition. For group contexts, 81.00% (IQR=72.07–88.59%) of vocalizations were continuous, 14.82% (IQR=10.40–21.33%) contained a single frequency jump, and 3.53% (IQR=1.01–6.01%) contained multiple jumps. Moreover, the proportion of vocalizations that were continuous was significantly different across social contexts (Kruskal–Wallis with Dunn *post hoc*:  $\chi^2=18.34$ ,  $P=1.04e-4$ ), with isolated males emitting a significantly greater proportion of continuous signals than males in a dyad ( $P=0.0084$ ) or a group ( $P=0.002$ ). The proportion of continuous vocalizations emitted by males did not differ between dyads and groups ( $P=0.98$ ). Isolated males emitted a significantly lower proportion of vocalizations with one jump than males in a dyad ( $P=0.022$ ) or a group ( $P=0.0007$ ), with no difference between dyads and groups ( $P=0.97$ ). Lastly, isolated males also emitted a significantly lower proportion of vocalizations with multiple jumps than either dyads ( $P=0.0006$ ) or groups ( $P=0.0006$ ), with no difference detected between dyads and groups ( $P=0.93$ ). Thus, our findings suggest that males, regardless of the number of partners, increase the complexity of their vocal emissions in social contexts compared with non-social contexts.

As a secondary quantification of acoustic complexity, we determined whether each male-emitted vocalization contained a turning point in frequency (switching from increasing in pitch to decreasing in pitch or vice versa (see Mahrt et al., 2013; Panksepp et al., 2007; Scattoni et al., 2008; Seagraves et al., 2016). Using this metric, male mice emitted a lower proportion of complex signals during one-mouse recordings compared to 2-mouse recordings (Fig. 2F; Kruskal–Wallis with Dunn *post hoc*:  $\chi^2=15.79$ ,  $P=0.011$ ; 1-mouse: median=48.39%, IQR=40.98–55.22%; 2-mouse: median=61.49%, IQR=57.26–62.49%) or 4-mouse recordings ( $P=0.0010$ ; 4-mouse: median=61.29%, IQR=55.32–64.53). No differences were observed between the 2- and 4-mouse contexts ( $P=1.0$ ). Therefore, these results provide further evidence that only a single social partner is needed for male acoustic complexity to reach an upper bound.

Although our results suggest that acoustic complexity differs based on the presence or absence of social partners, other factors may be at play. One possible explanation is that different experiences might contribute to the differences in acoustic complexity observed between singly and socially recorded mice. For the isolate-housed recordings of male mice, multiple prior experiences were uncontrolled. Some of the males used in the isolated recordings were marked with individual dye patterns, while others were unmarked. Some mice were previously exposed to the opposite sex, while others were not. Some animals lived with other mice, while others were isolate housed. In contrast, all males used in recordings with social partners were marked with dye patterns, were previously exposed to a female and were singly housed prior to recordings. To directly control for the potential impact that prior experience might play in our study, we ran two additional analyses using recordings of isolated males that had the same experiences as males in a group context (i.e. prior opposite-sex exposure and having their fur dyed;  $n=11$  isolated males). When comparing the acoustic complexity of males with similar experiences, we found that there were still significant differences in the proportion of vocalizations containing a change in the direction of frequency modulation in the 1-mouse and 2-mouse contexts (Kruskal–Wallis with Dunn:  $\chi^2=12.75$ ,  $P=0.01$ ; 1-mouse: median=41.94%, IQR=31.28–51.99%; 2-mouse: median=61.49%, IQR=57.26–62.49%) as well as the





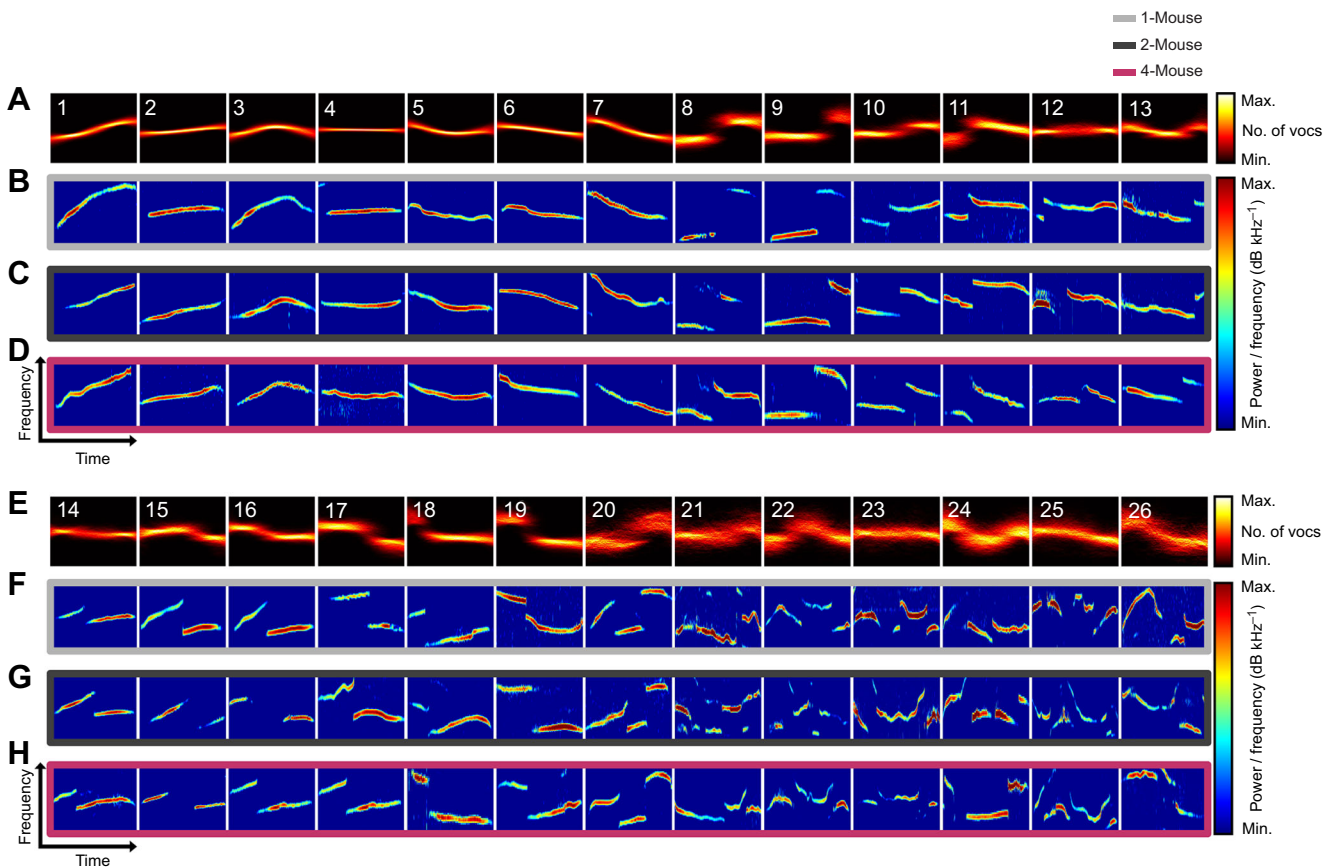
1-mouse and 4-mouse contexts ( $P=0.002$ , 4-mouse: median=61.27%, IQR=55.32–64.53%). These results were replicated when comparing the proportion of vocalizations with frequency jumps (Kruskal–Wallis with Dunn *post hoc*:  $\chi^2=9.98$ ; 1-mouse versus 2-mouse:  $P=0.047$ ; 1-mouse, median=6.45%, IQR=4.91–14.78%; 2-mouse, median=15.44%, IQR=8.30–16.65%; 1-mouse versus 4-mouse:  $P=0.006$ ; 4-mouse, median=14.82%, IQR=10.40–21.33%). Therefore, these results align with evidence from prior studies (Hanson and Hurley, 2012; Seagraves et al., 2016) and suggest that the presence of a social partner played a substantial role in increasing acoustic complexity.

While females were long thought to be silent during mixed-sex interactions, recent work has shown that female mice vocalize while engaged with males, albeit less frequently than males (Heckman

et al., 2017; Neunuebel et al., 2015; Sangiamo et al., 2020; Warren et al., 2020,b). Therefore, we sought to determine whether female mice modulate their acoustic complexity across social contexts. We found that females were vocally active in both dyadic ( $n=13$  females, 1614 female-emitted vocalizations) and group ( $n=11$  recordings, 1570 vocalizations from 22 females) settings. However, females were typically silent in isolation (non-social:  $n=35$  females, 39 female-emitted vocalizations). The lack of female vocalizations in isolation is similar to prior findings (Maggio and Whitney, 1985) and prevents a powered comparison to the two social contexts. Thus, we only compared female vocal emission between dyadic and group contexts. Fig. 2G shows the mean proportion of vocalizations categorized as continuous, 1-jump or multi-jump across individual females. Continuous vocalizations constituted

**Fig. 2. Quantification of acoustic complexity across social contexts.** (A–C) Representative images of 1 s of audio data recorded when (A) one male was in isolation (1-mouse), (B) a male and a female interacted (2-mouse), and (C) two males and two females interacted (4-mouse). Colored lines below vocalizations denote vocalizing animal with cyan representing males, orange showing females, and white indicating unassigned vocalizations. Gray in 1-mouse context shows vocalization assigned to virtual mouse (5.7% of all vocalizations in 1-mouse context). (D) Mean percentage of vocalizations emitted by male mice in each of the three social contexts (1-mouse=3905 vocalizations; 2-mouse=12,153 vocalizations; 4-mouse=15,616 vocalizations) that were continuous, had one jump or had multiple jumps. (E) Proportion of vocalizations per male mouse in each context with zero, one or multiple jumps (1-mouse:  $n=29$  males; 2-mouse:  $n=13$  males; 4-mouse:  $n=22$  males). (F) Proportion of male-emitted vocalizations with at least one change in the direction of frequency modulation. (G) Mean percentage of female-emitted vocalizations during the dyadic and group contexts (2-mouse=1614 vocalizations; 4-mouse=1570 vocalizations). Females were predominantly silent in isolation ( $n=35$  females, 39 vocalizations), therefore, the 1-mouse condition is not shown. (H) Proportion of vocalizations per female in each social context with zero, one or multiple jumps (2-mouse:  $n=13$  females; 4-mouse:  $n=22$  females). (I) Proportion of female vocalizations with at least one change in the direction of frequency modulation. In E,F,H,I, each dot represents a single mouse; gray boxes represent IQR; thick horizontal black lines represent group medians; Kruskal–Wallis with Dunn's *post hoc*:  $*P<0.01$ .



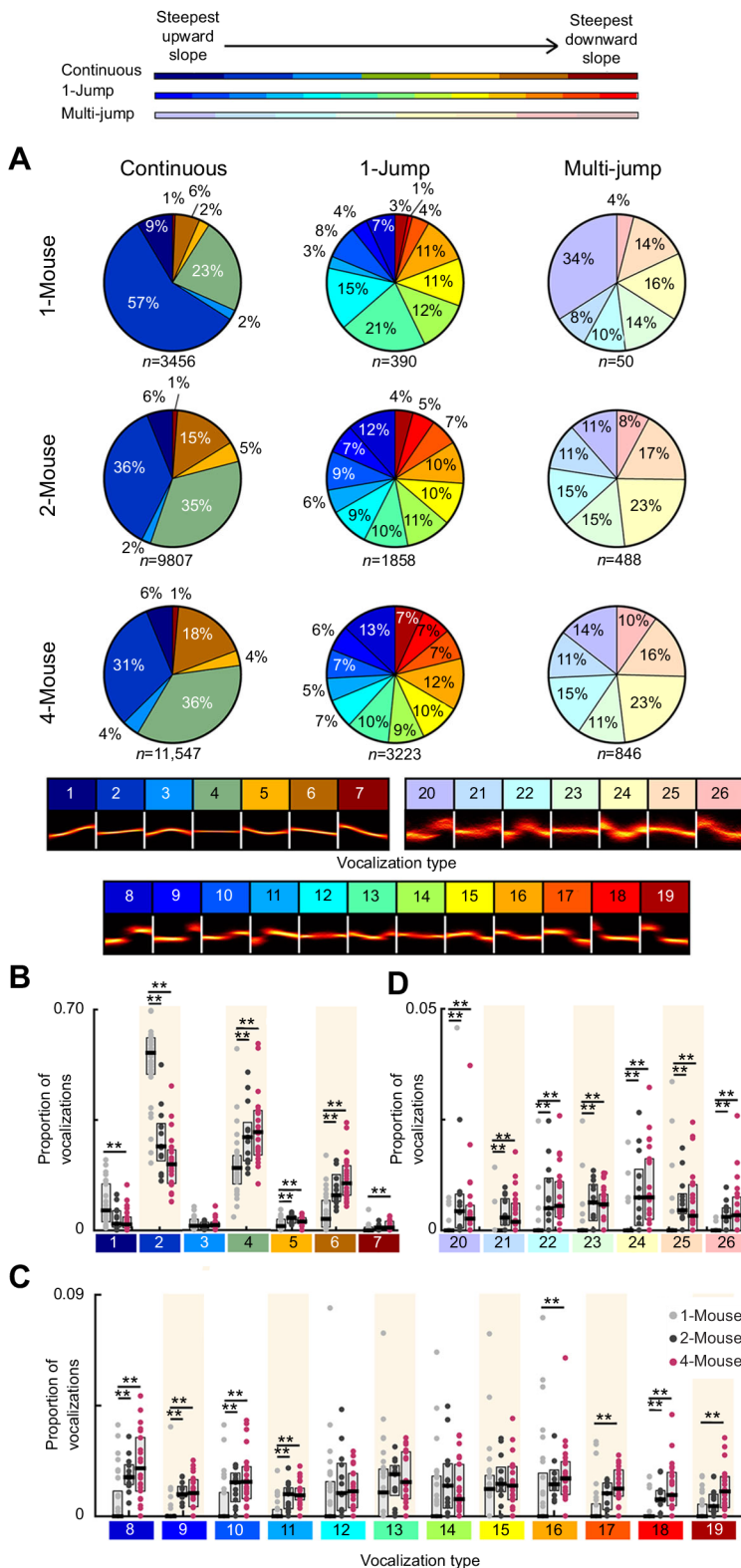


**Fig. 3. Automated vocal clustering determined that mice emit 26 types of vocalizations.** (A,E) Images with a black background show cumulative plots where all vocalizations of the indicated type are overlaid. Lighter colors represent more vocalizations, whereas darker colors denote fewer vocalizations. (B–D,F–H) Images with a blue background show individual examples of each vocalization type (column) emitted from an isolated male (1-Mouse; B,F; light gray), dyadic interaction (2-Mouse; C,G; dark gray) or group interaction (4-Mouse; D,H; pink).  $y$ -axis for each individual vocalization=60 kHz.

88.12% (IQR=86.63–93.14%) of vocalizations emitted by females in a dyadic condition compared with 89.09% (IQR=86.36–92.94%) in group settings (Fig. 2H; Mann–Whitney: ranksum=232,  $P=0.96$ ). Vocalizations containing a single frequency jump constituted 9.90% (IQR=6.26–12.14%) of female-emitted signals in dyads compared to 9.21% (IQR=4.12–13.46%) in group settings (Fig. 2H; Mann–Whitney: ranksum=238,  $P=0.91$ ). Finally, vocalizations with two or more jumps constituted 0.94% (IQR=0–1.93%) of the signals in dyadic settings versus 0.91% (IQR=0–2.67%) in group settings (Fig. 2H; Mann–Whitney: ranksum=263.5,  $P=0.94$ ). When quantifying acoustic complexity using our secondary metric, whereby a signal was considered complex if it contained at least one inflection in the direction of frequency modulation (Fig. 2I), the proportion of complex signals emitted by females was again similar between dyadic and group contexts (Mann–Whitney: ranksum=275.5;  $P=0.16$ ). During dyadic interactions, 68.15% (IQR=64.51–69.55%) of female-emitted vocalizations switched direction in frequency modulation. Similarly, 63.06% (IQR=60.71–69%) of female vocalizations during group interactions changed direction. Together, these results suggest that the complexity of female-emitted vocalizations, like male-emitted vocalizations, is unaltered by increasing the number of social partners.

Previous work has shown that mice emit different types of vocalizations in different behavioral contexts (Matsumoto and Okanoya, 2018). Therefore, we assessed whether mice produced different types of vocalizations across the three social contexts. Using an automated  $k$ -means clustering algorithm, vocalizations

emitted across each of the three contexts were clustered together based on shape. The clustering algorithm defined 23 unique vocalization types (Fig. S2). Based on evidence that mice emit vocalizations containing frequency jumps in specific behavioral contexts (Matsumoto and Okanoya, 2018), we reasoned that vocalizations containing frequency jumps may convey different meaning than continuous signals. Consequently, we sorted the vocalizations within each of the 23 types into groups based on complexity. This produced 23 types of continuous vocalizations, 23 types of vocalizations with a single frequency jump and 23 types of vocalizations with multiple frequency jumps. Then, to determine the optimal number of types necessary to accurately represent the vocalizations from each complexity level, we applied an elbow function to the vocalization clusters from each group. We found that 26 vocalization types were optimal to encompass the variability in both vocalization shape and complexity (Fig. 3; Fig. S2): 7 continuous types (vocalization types 1–7), 12 types with one jump (types 8–19) and 7 types with multiple jumps (types 20–26). Unique to this clustering approach, the method incorporates information about the vocalizing animal to ensure structurally complex, multi-syllabic vocalizations are produced by a single animal. To confirm that clustering was equally effective across all three social conditions, we used a Euclidian distance measure to compare within-type variability in shape (how similarly shaped are signals of the same type) and between-type variability (how similar in shape are vocalizations of different types). Within-type variability was consistently less than between-type variability (Fig. S3A–C).



**Fig. 4. Male mice alter the types of vocalizations they emit across social contexts.** (A) Pie charts show the mean percentage of all vocalizations, averaged across individual male mice, that were continuous (left column, vocalization types 1–7), had one jump (middle column, types 8–19) or had multiple jumps (right column, types 20–26). Continuous, 1-jump and multi-jump represent levels of complexity. Contexts include a single male (1-Mouse; top row; 29 males), a male and a female (2-Mouse; middle row; 13 males), or two males and two females (4-Mouse; bottom row; 22 males). The *n* below each pie chart denotes the number of vocalizations emitted during the specific context and with the indicated level of complexity. (B–D) The proportion of vocalizations that were emitted by each male and classified as continuous (B), 1-jump (C) and multi-jump (D). Each dot represents a single mouse. Gray boxes represent IQR; thick horizontal black lines represent group medians. Kruskal–Wallis with Dunn's *post hoc*. \*\**P*<0.01.

Furthermore, vocalizations within a type, but emitted in different contexts, were consistently more similar to each other than to different types of vocalizations emitted within the same context (Fig. S3D). These results demonstrate that, even with the addition of new vocalization types containing frequency jumps, our clustering method accurately grouped vocalizations with other similarly shaped signals across all social contexts.

With the new-found knowledge of the vocalization types that mice emit, we next determined whether the emission of specific types of vocalizations differed across social contexts (Fig. 4; statistics in Table 1). Males in dyadic conditions produced fewer type 1 and 2 vocalizations (continuous) than males in isolation. In contrast, males emitted more of types 4–6 (continuous), types 8–11 and 18 (1-jump), and types 20–26 (multi-jump) during interactions

**Table 1. Proportional vocal emission of male mice across three social contexts**

Vocalization type	1-Mouse		2-Mouse		4-Mouse		$\chi^2$	P	Pairwise P-values		
	Median	IQR	Median	IQR	Median	IQR			1 vs. 2	1 vs. 4	2 vs. 4
1	0.06	0.03–0.15	0.02	0.01–0.06	0.02	0.01–0.04	10.08	<0.01*	0.13	<0.01*	0.93
2	0.56	0.50–0.61	0.27	0.22–0.34	0.21	0.15–0.25	38.48	<10 <sup>-8</sup> *	<0.01*	<10 <sup>-8</sup> *	0.37
3	0.01	0–0.03	0.01	0.01–0.02	0.02	0.01–0.03	0.74	0.69	1.00	0.84	0.85
4	0.20	0.15–0.24	0.29	0.22–0.34	0.31	0.24–0.38	18.53	<10 <sup>-5</sup> *	0.01*	<10 <sup>-4</sup> *	0.96
5	0.01	0–0.04	0.04	0.02–0.05	0.03	0.02–0.04	10.75	<0.01*	<0.01*	0.10	0.51
6	0.04	0.01–0.09	0.11	0.09–0.18	0.15	0.11–0.20	26.81	<10 <sup>-6</sup> *	<0.01*	<10 <sup>-6</sup> *	0.68
7	0	0–0.01	0.01	0.01	0.01	0–0.03	12.13	<0.01*	0.09	<0.01*	0.89
8	0	0–0.01	0.02	0.01–0.02	0.02	0.01–0.03	19.12	<10 <sup>-5</sup> *	<0.01*	<10 <sup>-4</sup> *	0.97
9	0	0	0.01	0.01–0.02	0.01	0–0.01	22.65	<10 <sup>-5</sup> *	<10 <sup>-4</sup> *	<10 <sup>-4</sup> *	0.96
10	0	0–0.01	0.01	0.01–0.02	0.01	0.01–0.02	18.18	<10 <sup>-4</sup> *	0.01*	<10 <sup>-4</sup> *	0.92
11	0	0	0.01	0–0.01	0.01	0.01	23.12	<10 <sup>-5</sup> *	<0.01*	<10 <sup>-5</sup> *	0.98
12	0	0–0.01	0.01	0–0.02	0.01	0–0.02	3.92	0.14	0.27	0.28	0.99
13	0.01	0–0.02	0.02	0.01–0.02	0.01	0.01–0.03	3.89	0.14	0.70	0.15	0.89
14	0	0–0.02	0.01	0–0.02	0.01	0–0.02	2.85	0.24	0.29	0.65	0.87
15	0.01	0–0.02	0.01	0.01–0.02	0.01	0.01–0.02	1.69	0.43	0.54	0.75	0.96
16	0	0–0.02	0.01	0.01–0.02	0.02	0.01–0.02	9.39	<0.01*	0.09	0.01*	0.99
17	0	0–0.01	0.01	0–0.01	0.01	0.01–0.02	17.95	<10 <sup>-4</sup> *	<0.05*	<10 <sup>-4</sup> *	0.66
18	0	0	0.01	0.01	0.01	0–0.02	31.24	<10 <sup>-8</sup> *	<10 <sup>-4</sup> *	<10 <sup>-6</sup> *	0.88
19	0	0	0	0–0.01	0.01	0–0.02	15.97	<0.01*	0.13	<10 <sup>-4</sup> *	0.47
20	0	0	0	0–0.01	0	0–0.01	13.28	<0.01*	<0.01*	0.02*	0.68
21	0	0	0	0–0.01	0	0–0.01	19.83	<10 <sup>-5</sup> *	<10 <sup>-4</sup> *	<0.01*	0.80
22	0	0	0	0–0.01	0	0–0.01	22.41	<10 <sup>-5</sup> *	<0.01*	<10 <sup>-5</sup> *	0.93
23	0	0	0.01	0–0.01	0.01	0–0.01	16.51	<10 <sup>-4</sup> *	<0.01*	<0.01*	0.94
24	0	0	0.01	0–0.01	0.01	0–0.01	21.34	<10 <sup>-5</sup> *	<0.01*	<10 <sup>-5</sup> *	0.98
25	0	0	0	0–0.01	0	0–0.01	16.14	<10 <sup>-4</sup> *	<0.01*	<0.01*	0.74
26	0	0	0	0	0	0–0.01	21.83	<10 <sup>-5</sup> *	<10 <sup>-4</sup> *	<10 <sup>-4</sup> *	1.00

1 vs. 2=1-mouse versus 2-mouse context; 1 vs. 4=1-mouse versus 4-mouse context; 2 vs. 4=2-mouse versus 4-mouse context. Vocalization type: 1–7, continuous; 8–19, 1-jump; 20–26, multi-jump. \*P<0.05, Kruskal–Wallis with a Dunn *post hoc* test.

with a female than in the absence of a social partner. During group interactions, males emitted significantly higher proportions of types 4–7 (continuous), types 8–11 and 16–19 (1-jump), and types 20–26 (multi-jump) than in non-social conditions. Most notably, however, no significant differences were observed in male vocal emission between dyadic and group contexts. Interestingly, females significantly modulated the emission of only two vocalization types between dyadic and group contexts (Fig. 5; statistics in Table 2). These findings provide further evidence that male mice regulate vocal emission based on the presence (or absence) of a social partner. Moreover, male and female mouse vocal emission is generally unaffected by the presence of additional social partners.

## DISCUSSION

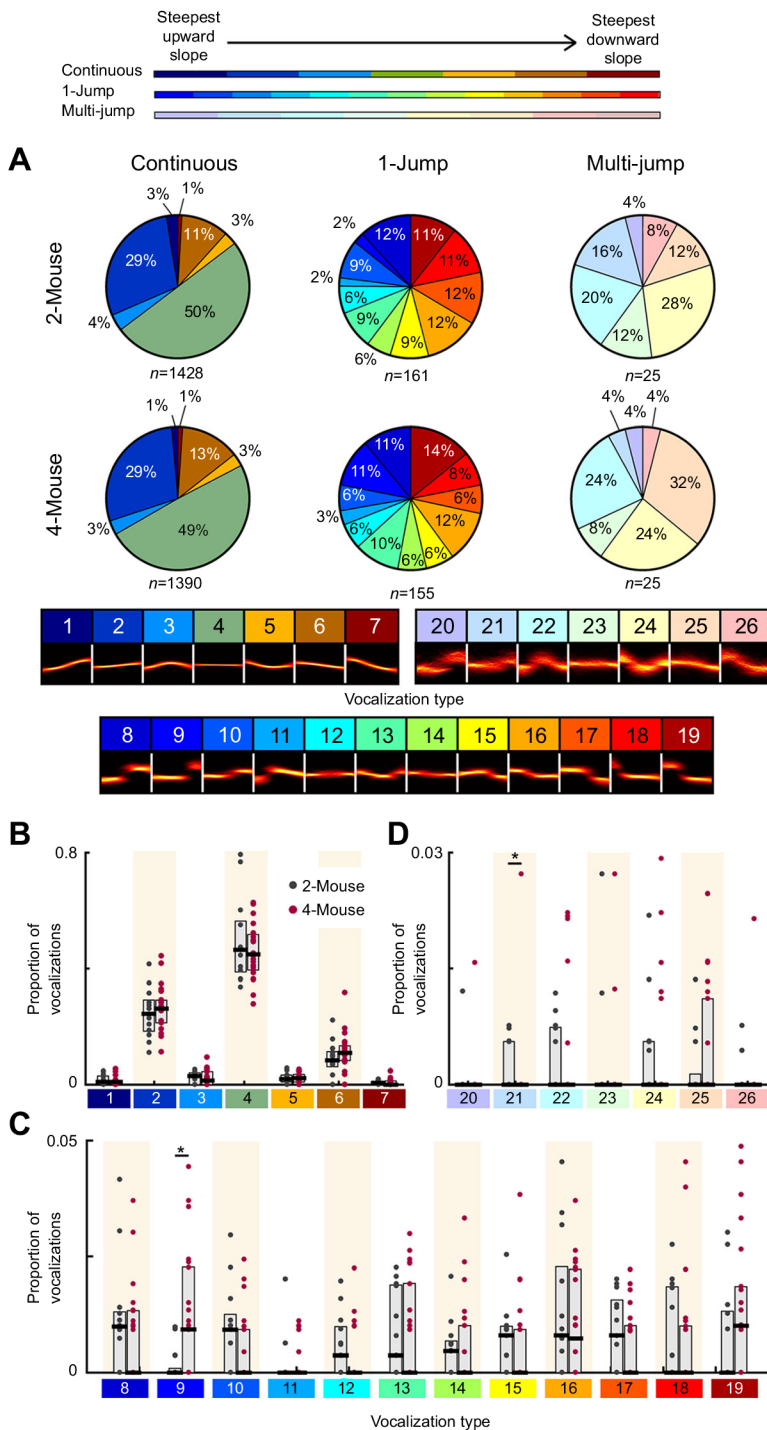
With the ability to assign a full repertoire of ultrasonic vocalizations to individual mice interacting with conspecifics, we examined the role that social context plays in mouse acoustic complexity. Our results indicate that male mice alter both the complexity of their vocalizations and the relative proportion of the vocalization types produced in isolated and social contexts. The most interesting finding, however, is that the complexity of both male and female vocal emissions, as well as the types of vocalizations emitted, is preserved as the number of social partners increases from 1 to 3.

Context-specific alterations in acoustic complexity have been shown in numerous species. One of the most studied models is birds, wherein males often produce complex courtship songs to attract females (Byers and Kroodsmas, 2009). Females are generally more attracted to males with larger repertoires of sound, altered phrasing or different song types (Buchanan and Catchpole, 1997; Collins, 1999; Mountjoy and Lemon, 1996; Reid et al., 2004). However, this phenomenon is not specific to birds. In *geladas*

(*Theropithecus gelada*), females prefer to spend time near speakers playing complex calls compared with simple calls (Gustison and Bergman, 2016). In greater sac-winged bats (*Saccopteryx bilineata*), females consistently roost in the harem of males exhibiting greater variation in their vocal emissions than males producing fewer syllable types (Davidson and Wilkinson, 2004). In Túngara frogs (*Engystomops pustulosus*), females are more attracted to complex calls (Rand and Ryan, 1981). Thus, across the animal kingdom, enhancing acoustic complexity may be reproductively advantageous for males.

Mice also show context-specific alterations in acoustic complexity, specifically between isolated and social conditions (Matsumoto and Okanoya, 2018). C57/B6 mice emit more diverse signals while socially engaged than when exploring a novel environment or when isolated and under stress (Chabout et al., 2012). When a female is removed from a male–female dyad, fewer vocalizations containing frequency jumps are emitted and there are also fewer vocalizations that decrease in pitch (Yang et al., 2013). However, once the female is returned to the arena, vocal activity resembles that seen in the original dyadic conditions. In contrast, for the CBA/CaJ mouse, more vocalizations are emitted after a female is removed than while a female is present (Hanson and Hurley, 2012). Additional work showed that vocalizations are more complex in response to female scent cues than during social interactions (Chabout et al., 2015), which contrasts with our findings. One possible explanation for the differences is that each study used a different strain of mice. The previous study employed B5D2F1/J mice and we used C57BL/6J mice. Using C57BL/6J mice, Chabout et al. (2016) showed that more complex vocal syntax was produced during dyadic interactions than in isolation. Thus, these studies provide additional evidence that USV acoustic complexity may be regulated by context in different strains of mice.





**Fig. 5. Female mice produce similar types of vocalizations during dyadic and group interactions.** (A) Pie charts show the mean percentage of all vocalizations, averaged across individual female mice, that were continuous (left column, types 1–7), had one jump (middle column, types 8–19), or had multiple jumps (right column, types 20–26). Contexts include a male and a female (2-Mouse; top row; 13 females) or two males and two females (4-Mouse; bottom row; 22 females). The *n* below each pie chart denotes the number of vocalizations emitted during the specific context and with the indicated level of complexity. (B–D) The proportion of vocalizations that were emitted by each female and classified as continuous (B), 1-jump (C) and multi-jump (D). Each dot represents one female. Thick horizontal black lines indicate group medians. Boxes represent IQR. Mann–Whitney: \**P*<0.05.

Another example of the influence of social context on vocal emission in mice is the audience effect, whereby males upregulate the complexity of vocalizations in the presence of another male (Seagraves et al., 2016). However, males in the Seagraves et al. (2016) study were unable to access a live female. Instead, male vocal activity was recorded either solely in the presence of female scent cues or in the presence of female scent cues plus a live male. Our results extend this finding and directly show that male acoustic complexity also increases in the presence of a female listener. Taken together, these results suggest that male mice increase the complexity of vocal emissions in the presence of any social partner. Additionally, our results uncover a potential upper bound

for mouse acoustic complexity that is reached with a single social partner, as the acoustic complexity of both males and females is maintained between contexts containing one or three partners. The findings are surprising because increasing the number of audience members is believed to increase acoustic complexity in many other species (Matos and Schlupp, 2005).

The internal states of an animal strongly influence vocal emission (Morton, 1977). In many species, motivational and emotional states may underlie context-dependent changes in the frequency [e.g. rats (Brudzynski, 2007; Burgdorf et al., 2008) and humans (Bachorowski and Owren, 1995)] or temporal structure [e.g. bats (Bastian and Schmidt, 2008) and tree shrews (Schehka et al., 2007)]

**Table 2. Proportional vocal emission of female mice across two social contexts**

Vocalization type	2-Mouse		4-Mouse		P-value
	Median	IQR	Median	IQR	
1	0.01	0–0.03	0.01	0–0.01	0.72
2	0.24	0.18–0.29	0.26	0.21–0.29	0.61
3	0.03	0–0.04	0.01	0–0.05	0.50
4	0.46	0.39–0.56	0.45	0.39–0.52	0.82
5	0.02	0.01–0.03	0.02	0.01–0.04	0.84
6	0.08	0.06–0.11	0.11	0.08–0.13	0.15
7	0.01	0–0.01	0	0–0.01	0.97
8	0.01	0–0.01	0	0–0.01	0.37
9	0	0	0.01	0–0.02	<0.03*
10	0.01	0–0.01	0	0–0.01	0.11
11	0	0	0	0	0.64
12	0	0–0.01	0	0	0.13
13	0	0–0.02	0	0–0.02	0.94
14	0	0–0.01	0	0–0.01	0.46
15	0.01	0–0.01	0	0–0.01	0.26
16	0.01	0–0.02	0.01	0–0.02	0.65
17	0.01	0–0.02	0	0–0.01	0.28
18	0	0–0.02	0	0–0.01	0.54
19	0	0–0.01	0.01	0–0.02	0.22
20	0	0	0	0	0.73
21	0	0	0	0	<0.05*
22	0	0–0.01	0	0	0.52
23	0	0	0	0	0.60
24	0	0	0	0	0.52
25	0	0	0	0–0.01	0.47
26	0	0	0	0	0.31

Vocalization type: 1–7, continuous; 8–19, 1-jump; 20–26, multi-jump. \* $P < 0.05$ , Kruskal–Wallis with a Dunn *post hoc* test.

of vocalizations. Moreover, vocal emission may be affected by arousal (Bell, 1974) as seen in prairie vole pups, where increases in heart rate co-occur with decreases in USV duration and complexity (Stewart et al., 2015). While these examples strongly argue that motivation, emotion and arousal play important roles in vocal emission, it is less clear how internal states influence mouse acoustic complexity and emission. Because we found that the complexity and types of mouse vocalizations mice are similar across different social contexts, one possible explanation is that the underlying emotional, motivational or arousal states are similar in the two social contexts. Alternatively, internal states may not largely impact mouse vocal emission, such that internal states only push acoustic complexity towards an upper limit. The exact behavioral repertoire of individual animals may also be a driving factor that underlies mouse vocal emissions. For instance, mice emit vocalizations containing frequency jumps prior to and during mounting behavior (Hanson and Hurley, 2012; Matsumoto and Okanoya, 2018). Additionally, more vocalizations are emitted when two males are investigating each other than when males are separated from each other (Seagraves et al., 2016). More directly, male mice have been shown to emit distinct types of signals during specific behaviors (Sangiameo et al., 2020). Therefore, a potential explanation for the lack of vocal differences between social contexts is that vocalizations are closely linked to behavior, and the behavioral repertoire of mice in the two social conditions may have been similar. In all likelihood, both internal state and behavioral actions influence vocal emission, but determining the specific contributions of each variable requires further investigation.

Female vocal emission is highly regulated and influences dynamic social interactions. For instance, calls by female Alaskan moose (*Alces alces gigas*) are thought to discourage advances from subordinate males and ensure mating opportunities with more

dominant males (Bowyer et al., 2011). Female dunnocks (*Prunella modularis*), while consistently vocal during territorial conflicts with rival females, are even more vocal when competing for male attention (Langmore and Davies, 1997). Female mice have also been shown to regulate their vocal emissions. For example, the sex of a female's social partner influences her latency to vocalize, with females taking longer to vocalize in same-sex dyads than mixed-sex interactions (Warren et al., 2020). Furthermore, female vocalizations appear to change the behavior of males in specific behavioral contexts (Warren et al., 2020). Familiarity with a social partner also impacts vocal emission, with females vocalizing more in the presence of a novel than a familiar animal (D'Amato and Moles, 2001). Interestingly, our findings indicate that females do not alter the complexity or types of vocalizations emitted based on the size of their social group, with vocalizations being indistinguishable in the presence of one or multiple listeners. However, female vocal rate increased whenever a listener was present, as few vocalizations were detected when recording females in isolation. These results could also be interpreted as an increase in complexity in the presence of a social partner, as any type of vocalization is more complex than silence. Thus, because the vocal behavior of females is strongly influenced by multiple environmental and social conditions and, importantly, these signals directly modulate social interactions, our working models of mouse social communication need to account for female vocalizations.

In conclusion, our ability to localize sounds as animals freely interact enabled us to directly show that the complexity of ultrasonic vocalizations emitted by individual mice increases in the presence of a social partner. Moreover, acoustic complexity appears to reach an upper bound with the addition of a single listener, as vocal emission is similar in dyadic and group interactions. Our findings provide the foundation for understanding the relationship between complex social communication and natural behavior, as well as

disentangling the roles that internal states and external cues play in regulating the vocal emissions of both male and female mice.

#### Acknowledgements

We thank the staff from the Life Science Research Facility for assistance as well as James Farmer and Jaime Quesenberry for their help in building the acoustic chamber. We thank Anita Schwarz and other members of the University of Delaware Information Technologies for facilitating our ability to process the data. Finally, we thank the reviewers for all their helpful feedback.

#### Competing interests

The authors declare no competing or financial interests.

#### Author contributions

Conceptualization: M.R.W., J.P.N.; Methodology: M.R.W., J.P.N.; Software: M.R.W., M.S.S., D.T.S., R.S.C.; Validation: M.R.W., J.P.N.; Formal analysis: M.R.W.; Investigation: M.R.W., M.S.S., D.T.S., R.S.C.; Resources: J.P.N.; Data curation: J.P.N.; Writing - original draft: M.R.W.; Writing - review & editing: M.R.W., M.S.S., D.T.S., R.S.C., J.P.N.; Visualization: M.R.W.; Supervision: J.P.N.; Project administration: J.P.N.; Funding acquisition: J.P.N.

#### Funding

This work was funded by the University of Delaware Research Foundation, General University Research Program, and the National Institutes of Health (2P20GM103653 and R01MH122752). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Deposited in PMC for release after 12 months.

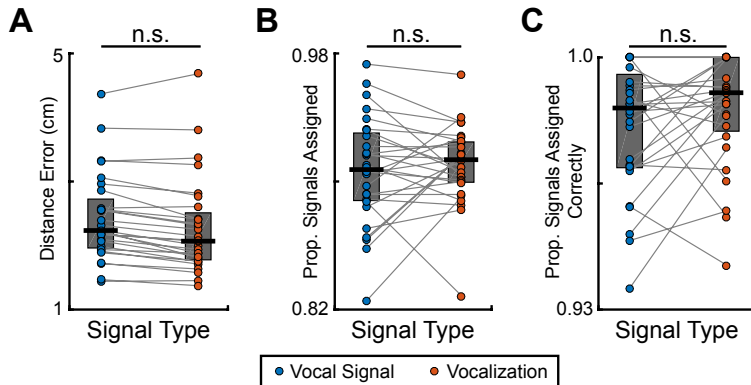
#### References

- Arriaga, G., Zhou, E. P. and Jarvis, E. D. (2012). Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS ONE* 7, e46610. doi:10.1371/journal.pone.0046610
- Bachorowski, J.-A. and Owren, M. J. (1995). Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. *Psychol. Sci.* 6, 219-224. doi:10.1111/j.1467-9280.1995.tb00596.x
- Bastian, A. and Schmidt, S. (2008). Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am.* 124, 598-608. doi:10.1121/1.2924123
- Behr, O. and von Helversen, O. (2004). Bat serenades—complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behav. Ecol. Sociobiol.* 56, 106-115. doi:10.1007/s00265-004-0768-7
- Bell, R. W. (1974). Ultrasounds in small rodents: Arousal-produced and arousal-producing. *Dev. Psychobiol.* 7, 39-42. doi:10.1002/dev.420070107
- Bowyer, R. T., Rachlow, J. L., Stewart, K. M. and Van Ballenberghe, V. (2011). Vocalizations by Alaskan moose: female incitation of male aggression. *Behav. Ecol. Sociobiol.* 65, 2251-2260. doi:10.1007/s00265-011-1234-y
- Bradbury, J. W. and Vehrencamp, S. L. (1998). *Principles of Animal Communication*. Sunderland, Massachusetts: Sinauer Associates, Inc.
- Brudzynski, S. M. (2007). Ultrasonic calls of rats as indicator variables of negative or positive states: Acetylcholine-dopamine interaction and acoustic coding. *Behav. Brain Res.* 182, 261-273. doi:10.1016/j.bbr.2007.03.004
- Buchanan, K. L. and Catchpole, C. K. (1997). Female choice in the sedge warbler *Acrocephalus schoenobaenus* multiple cues from song and territory quality. *Proc. R. Soc. Lond. B Biol. Sci.* 264, 521-526. doi:10.1098/rspb.1997.0074
- Burgdorf, J., Kroes, R. A., Moskal, J. R., Pfaus, J. G., Brudzynski, S. M. and Panksepp, J. (2008). Ultrasonic vocalizations of rats (*Rattus norvegicus*) during mating, play, and aggression: Behavioral concomitants, relationship to reward, and self-administration of playback. *J. Comp. Psychol.* 122, 357-367. doi:10.1037/a0012889
- Byers, B. E. and Kroodsmas, D. E. (2009). Female mate choice and songbird song repertoires. *Anim. Behav.* 77, 13-22. doi:10.1016/j.anbehav.2008.10.003
- Chabout, J., Sarkar, A., Dunson, D. B. and Jarvis, E. D. (2015). Male mice song syntax depends on social contexts and influences female preferences. *Front. Behav. Neurosci.* 9, 76. doi:10.3389/fnbeh.2015.00076
- Chabout, J., Sarkar, A., Patel, S. R., Radden, T., Dunson, D. B., Fisher, S. E. and Jarvis, E. D. (2016). A Foxp2 mutation implicated in human speech deficits alters sequencing of ultrasonic vocalizations in adult male mice. *Front. Behav. Neurosci.* 10, 197. doi:10.3389/fnbeh.2016.00197
- Chabout, J., Serreau, P., Ey, E., Bellier, L., Aubin, T., Bourgeron, T. and Granon, S. (2012). Adult male mice emit context-specific ultrasonic vocalizations that are modulated by prior isolation or group rearing environment. *PLoS ONE* 7, e29401. doi:10.1371/journal.pone.0029401
- Chen, Y., Matheson, L. E. and Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci. USA* 113, 6641-6646. doi:10.1073/pnas.1522306113
- Collins, S. A. (1999). Is female preference for male repertoires due to sensory bias? *Proc. R. Soc. B* 266, 2309-2309. doi:10.1098/rspb.1999.0924
- Cora, M. C., Kooistra, L. and Travlos, G. (2015). Vaginal cytology of the laboratory rat and mouse: review and criteria for the staging of the estrous cycle using stained vaginal smears. *Toxicol. Pathol.* 43, 776-793. doi:10.1177/0192623315570339
- D'Amato, F. R. and Moles, A. (2001). Ultrasonic vocalizations as an index of social memory in female mice. *Behav. Neurosci.* 115, 834-840. doi:10.1037/0735-7044.115.4.834
- da Silva, K. B., Mahan, C. and da Silva, J. (2002). The trill of the chase: eastern chipmunks call to warn kin. *J. Mammal.* 83, 546-552. doi:10.1093/mamm/83.3.546
- Davidson, S. M. and Wilkinson, G. S. (2004). Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Anim. Behav.* 67, 883-891. doi:10.1016/j.anbehav.2003.06.016
- Gamba, M., Colombo, C. and Giacoma, C. (2012). Acoustic cues to caller identity in lemurs: a case study. *J. Ethol.* 30, 191-196. doi:10.1007/s10164-011-0291-z
- Gaub, S., Fisher, S. E. and Ehret, G. (2016). Ultrasonic vocalizations of adult male Foxp2-mutant mice: behavioral contexts of arousal and emotion. *Genes Brain Behav.* 15, 243-259. doi:10.1111/gbb.12274
- Gourbal, B. E., Barthelemy, M., Petit, G. and Gabrion, C. (2004). Spectrographic analysis of the ultrasonic vocalisations of adult male and female BALB/c mice. *Naturwissenschaften* 91, 381-385. doi:10.1007/s00114-004-0543-7
- Grimsley, J. M., Sheth, S., Vallabh, N., Grimsley, C. A., Bhattal, J., Latsko, M., Jasnow, A. 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, 38. doi:10.3389/fnbeh.2016.00038
- Guo, Z. and Holy, T. E. (2007). Sex selectivity of mouse ultrasonic songs. *Chem. Senses* 32, 463-473. doi:10.1093/chemse/bjm015
- Gustison, M. L. and Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Sci. Rep.* 6, 19680. doi:10.1038/srep19680
- 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. doi:10.1371/journal.pone.0041133
- Hanson, J. L. and Hurley, L. M. (2012). Female presence and estrous state influence mouse ultrasonic courtship vocalizations. *PLoS ONE* 7, e40782. doi:10.1371/journal.pone.0040782
- Heckman, J. J., Provillo, R., Heckman, G. J., Azarfar, A., Celikel, T. and Englitz, B. (2017). High-precision spatial localization of mouse vocalizations during social interaction. *Sci. Rep.* 7, 3017. doi:10.1038/s41598-017-02954-z
- Hilakivi-Clarke, L. A. and Lister, R. G. (1992). The role of body weight in resident-intruder aggression. *Aggr. Behav.* 18, 281-287.
- Holy, T. E. and Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biol.* 3, e386. doi:10.1371/journal.pbio.0030386
- Ji, A., Johnson, M. T., Walsh, E. J., McGee, J. and Armstrong, D. L. (2013). Discrimination of individual tigers (*Panthera tigris*) from long distance roars. *J. Acoust. Soc. Am.* 133, 1762-1769. doi:10.1121/1.4789936
- Jones, R. B. and Nowell, N. W. (1989). Aversive potency of urine from dominant and subordinate male laboratory mice (*Mus musculus*): resolution of a conflict. *Aggr. Behav.* 15, 291-296.
- Kao, M. H. and Brainard, M. S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J. Neurophysiol.* 96, 1441-1455. doi:10.1152/jn.01138.2005
- Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akcay, C., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Casar, C. et al. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol. Rev. Camb. Philos. Soc.* 91, 13-52. doi:10.1111/brv.12160
- Knutson, B., Burgdorf, J., Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychol. Bull.* 128, 961-977. doi:10.1037/0033-2909.128.6.961
- Konig, B. (1994). Fitness effects of communal rearing in house mice: the role of relatedness versus familiarity. *Animal Behav.* 48, 1449-1457
- Langmore, N. E. and Davies, N. B. (1997). Female dunlocks use vocalizations to compete for mates. *Anim. Behav.* 53, 881-890. doi:10.1006/anbe.1996.0306
- le Roux, A., Cherry, M. I. and Manser, M. B. (2008). The audience effect in a facultatively social mammal, the yellow mongoose, *Cynictis penicillata*. *Anim. Behav.* 75, 943-949. doi:10.1016/j.anbehav.2007.07.014
- Maggio, J. C. and Whitney, G. (1985). Ultrasonic vocalizing by adult female mice (*Mus musculus*). *J. Comp. Psychol.* 99, 420-436. doi:10.1037/0735-7036.99.4.420
- 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, 5573-5583. doi:10.1523/JNEUROSCI.5054-12.2013
- 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 University Press. doi:10.1017/CBO9780511610363.006
- Matsumoto, Y. K. and Okanoya, K. (2018). Mice modulate ultrasonic calling bouts according to sociosexual context. *R Soc. Open Sci.* 5, 180378. doi:10.1098/rsos.180378
- Miller, J. R. and Engstrom, M. D. (2007). Vocal stereotypy and singing behavior in baiomyine mice. *J. Mammal.* 88, 1447-1465. doi:10.1644/06-MAMM-A-386R.1



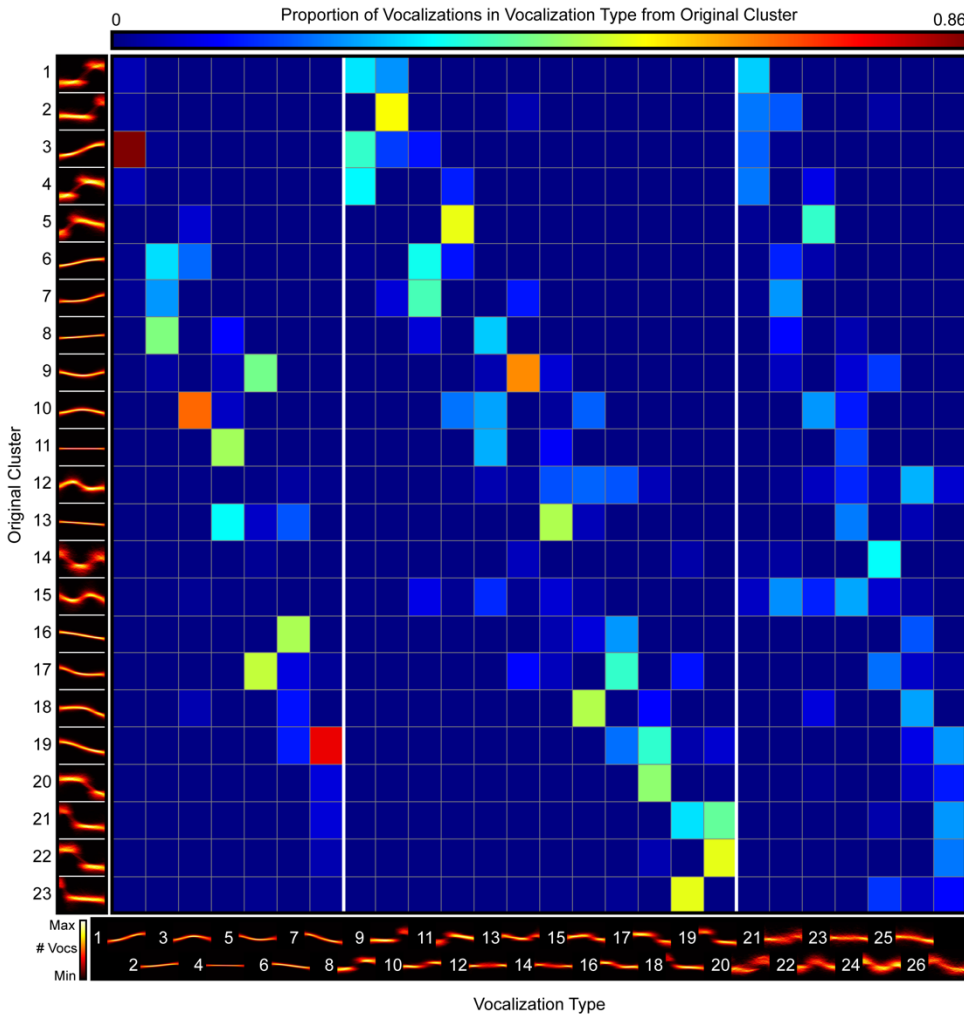
- Morton, E. S.** (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855-869. doi:10.1086/283219
- Mountjoy, D. J. and Lemon, R. E.** (1996). Female choice for complex song in the European starling: a field experiment. *Behav. Ecol. Sociobiol.* **38**, 65-71. doi:10.1007/s002650050218
- Musolf, K., Hoffmann, F. and Penn, D.** (2010). Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. *Anim. Behav.* **79**, 757-764. doi:10.1016/j.anbehav.2009.12.034
- Neunuebel, J. P., Taylor, A. L., Arthur, B. J. and Egnor, S. R.** (2015). Female mice ultrasonically interact with males during courtship displays. *eLife* **4**, e06203. doi:10.7554/eLife.06203
- Ohayon, S., Avni, O., Taylor, A. L., Perona, P. and Roian Egnor, S. E.** (2013). Automated multi-day tracking of marked mice for the analysis of social behaviour. *J. Neurosci. Methods* **219**, 10-19. doi:10.1016/j.jneumeth.2013.05.013
- 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. doi:10.1371/journal.pone.0000351
- Percival, D. B. and Walden, A. T.** (1993). Spectral analysis for physical applications. Cambridge University Press.
- Pettitt, B. A., Bourne, G. R. and Bee, M. A.** (2012). Quantitative acoustic analysis of the vocal repertoire of the golden rocket frog (*Anomaloglossus beebei*). *J. Acoust. Soc. Am.* **131**, 4811-4820. doi:10.1121/1.4714769
- Rand, A. S. and Ryan, M. J.** (1981). The adaptive significance of a complex vocal repertoire in a neotropical frog. *Z. Tierpsychol.* **57**, 209-214. doi:10.1111/j.1439-0310.1981.tb01923.x
- Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Hiebert, S. M., Smith, J. N. M., Stoddard, P. K., Marr, A. B. and Keller, L. F.** (2004). Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Anim. Behav.* **68**, 1055-1063. doi:10.1016/j.anbehav.2004.07.003
- Sangiamo, D. T., Warren, M. R. and Neunuebel, J. P.** (2020). Ultrasonic signals associated with different types of social behavior of mice. *Nat. Neurosci.* **23**, 411-422. doi:10.1038/s41593-020-0584-z
- 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. doi:10.1371/journal.pone.0003067
- Schehka, S., Esser, K.-H. and Zimmermann, E.** (2007). Acoustical expression of arousal in conflict situations in tree shrews (*Tupaia belangeri*). *J. Comp. Physiol. A* **193**, 845-852. doi:10.1007/s00359-007-0236-8
- Seagraves, K. M., Arthur, B. J. and Egnor, S. E.** (2016). Evidence for an audience effect in mice: male social partners alter the male vocal response to female cues. *J. Exp. Biol.* **219**, 1437-1448. doi:10.1242/jeb.129361
- Sehsweneey, M., Wilson, D. R., Bain, M., Boutin, S., Lane, J. E., McAdam, A. G. and Dantzer, B.** (2019). The effects of stress and glucocorticoids on vocalizations: a test in North American red squirrels. *Behav. Ecol.* **30**, 1030-1040. doi:10.1093/behecol/anz044
- Seyfarth, R. M., Cheney, D. L. and Marler, P.** (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801-803. doi:10.1126/science.7433999
- Stewart, A. M., Lewis, G. F., Yee, J. R., Kenkel, W. M., Davila, M. I., Carter, C. S. and Porges, S. W.** (2015). Acoustic features of prairie vole (*Microtus ochrogaster*) ultrasonic vocalizations covary with heart rate. *Physiol. Behav.* **138**, 94-100. doi:10.1016/j.physbeh.2014.10.011
- Stoeger, A. S. and Baotic, A.** (2016). Information content and acoustic structure of male African elephant social rumbles. *Sci. Rep.* **6**, 27585. doi:10.1038/srep27585
- Warburton, V. L., Sales, G. D. and Milligan, S. R.** (1989). The emission and elicitation of mouse ultrasonic vocalizations: the effects of age, sex and gonadal status. *Physiol. Behav.* **45**, 41-47. doi:10.1016/0031-9384(89)90164-9
- Warren, M. R., Sangiamo, D. T. and Neunuebel, J. P.** (2018a). High channel count microphone array accurately and precisely localizes ultrasonic signals from freely-moving mice. *J. Neurosci. Methods* **297**, 44-60. doi:10.1016/j.jneumeth.2017.12.013
- Warren, M. R., Spurrier, M. S., Roth, E. D. and Neunuebel, J. P.** (2018b). Sex differences in vocal communication of freely interacting adult mice depend upon behavioral context. *PLoS ONE* **13**, e0204527. doi:10.1371/journal.pone.0204527
- Warren, M. R., Clein, R. S., Spurrier, M. S., Roth, E. D. and Neunuebel, J. P.** (2020). Ultrashort-range, high-frequency communication by female mice shapes social interactions. *Sci. Rep.* **10**, 2637. doi:10.1038/s41598-020-59418-0
- Weiner, B., Hertz, S., Perets, N. and London, M.** (2016). Social ultrasonic vocalization in awake head-restrained mouse. *Front. Behav. Neurosci.* **10**, 236. doi:10.3389/fnbeh.2016.00236
- White, S. J. and White, R. E. C.** (1970). Individual voice production in gannets. *Behaviour* **37**, 40-54. doi:10.1163/156853970X00222
- 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. doi:10.1037/h0034899
- Wysocki, C. J., Nyby, J., Whitney, G., Beauchamp, G. K. and Katz, Y.** (1982). The vomeronasal organ: primary role in mouse chemosensory gender recognition. *Physiol. Behav.* **29**, 315-327. doi:10.1016/0031-9384(82)90021-X
- Yang, M., Loureiro, D., Kalikhman, D. and Crawley, J. N.** (2013). Male mice emit distinct ultrasonic vocalizations when the female leaves the social interaction arena. *Front. Behav. Neurosci.* **7**, 159. doi:10.3389/fnbeh.2013.00159
- Zuberbühler, K., Noë, R. and Seyfarth, R. M.** (1997). Diana monkey long-distance calls: messages for conspecifics and predators. *Anim. Behav.* **53**, 589-604. doi:10.1006/anbe.1996.0334

## Supplemental Figure 1



**Figure S1: Sound-source localization accuracy is similar for vocal signals and vocalizations.** We evaluated how accurately our sound-source localization system localized vocal signals and vocalizations from singly-recorded males that were vocally stimulated using female scent cues. A) Distance between the nose of the vocalizing mouse and the location where the system estimated the origin of the vocal signals (left; blue) and vocalizations (right; orange). Each dot is the median from a single male. B) Proportion of all extracted signals that were assigned to a mouse. C) Proportion of assigned signals attributed to the real mouse. For B and C, we generated random locations for three additional mice, then quantified the likelihood that each of the four animals (one real and three artificial) emitted each signal. A-C) Each dot is the median value from a single 10-minute recording of an isolated male ( $n = 29$ ). Connected dots are from the same recording. Gray boxes represent IQR; thick horizontal black lines represent group medians. Stats = Mann-Whitney. \*  $p < 0.05$ ; \*\*  $p < 0.01$ . n.s. = non-significant.

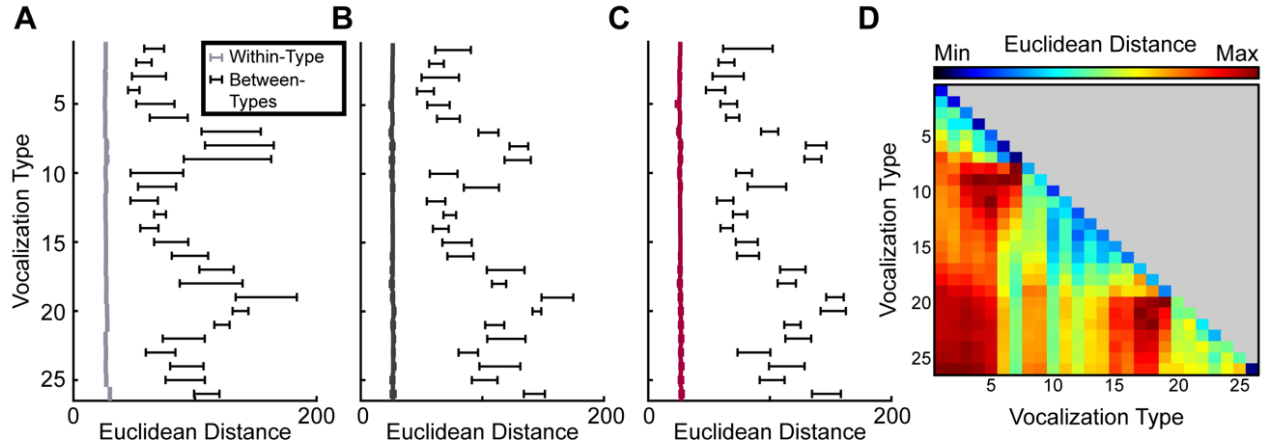
## Supplemental Figure 2



**Figure S2: Vocalizations were organized into 26 unique vocalization types.** Using a k-means clustering program, all vocalizations ( $n = 46,649$ ) were partitioned into 23 clusters (left – cumulative plots). To dissociate vocalizations based on acoustic complexity, we then separated vocalizations within each cluster into three groups: continuous, single jump, and multiple jumps. Then, across all 23 clusters of vocalizations from each group, we applied an elbow function to determine the optimal number of vocalization types to represent the variability within the group. There were 7 types of continuous vocalizations (types 1-7), 12 types of vocalizations with a single jump (types 8-19), and 7 types of vocalizations with multiple jumps (types 20-26). The color of each box indicates the proportion of vocalizations in each of the 26 types that originated from each of the 23 original clusters. Cumulative plots for each of the 26 vocalization types are staggered below the heatmap (odd- and even-numbered vocalization types in the top and bottom rows, respectively). The right edge of each vocalization type is right-aligned to the corresponding column.



### Supplemental Figure 3



**Figure S3: Vocalizations were accurately clustered across all three social contexts.** A-C) Interquartile range of Euclidean distance between vocalizations within a Vocalization Type (colored lines) compared to other vocalization types (black lines) for A) 1-mouse, B) 2-mouse, or C) 4-mouse contexts. Thick colored line represents median distance within a vocalization type. D) Mean Euclidean distance between vocalization types and contexts. The identity line shows comparisons of the same vocalization types emitted in different social contexts. Off-diagonal boxes show comparisons of different vocalization types that were emitted during the same context. See Methods for details about quantification.