# iScience

# Article

# Vocal recognition of partners by female prairie voles

## **Graphical abstract**



# **Highlights**

Check for

- Muroid rodents can display individual vocal recognition
- Adult prairie vole vocalizations vary more across individuals than social experience
- Voles emit behavior-dependent vocalizations
- Female prairie voles behaviorally recognize their mate's vocalizations

## **Authors**

Megan R. Warren, Jenny Zha, Larry J. Young, Robert C. Liu

Correspondence robert.liu@emory.edu

# In brief

Rodent behavior; Bioacoustics; Evolutionary biology



# **iScience**



## Article

# Vocal recognition of partners by female prairie voles

Megan R. Warren,<sup>1,2</sup> Jenny Zha,<sup>1</sup> Larry J. Young,<sup>2,3,4</sup> and Robert C. Liu<sup>1,2,5,\*</sup>

<sup>1</sup>Department of Biology, Emory University, Atlanta, GA, USA

<sup>2</sup>Center for Translational Social Neuroscience, Emory National Primate Research Center, Atlanta, GA, USA <sup>3</sup>Department of Psychiatry and Behavioral Science, Emory University School of Medicine, Atlanta, GA, USA <sup>4</sup>Deceased

Deceased

<sup>5</sup>Lead contact \*Correspondence: robert.liu@emory.edu https://doi.org/10.1016/j.isci.2025.111796

#### **SUMMARY**

Recognizing conspecifics is vital for differentiating mates, offspring, and social threats. Individual recognition is often reliant upon chemical or visual cues but can also be facilitated by vocal signatures in some species. In common laboratory rodents, playback studies have uncovered communicative functions of vocalizations, but scant behavioral evidence exists for individual vocal recognition. Here, we find that the socially monogamous prairie vole (*Microtus ochrogaster*) emits behavior-dependent vocalizations that can communicate individual identity. Vocalizations of individual males change after bonding with a female; however, acoustic variation across individuals is greater than within-individual variation. Critically, females behaviorally discriminate their partner's vocalizations from a stranger's, even if emitted to another stimulus female. These results establish the acoustic and behavioral foundation for individual vocal recognition in prairie voles, where neurobiological tools enable future studies revealing its causal neural mechanisms.

#### **INTRODUCTION**

Navigating complex social environments often requires an ability to recognize individuals, as differentiating between them is vital for handling dominance hierarchies,<sup>1</sup> identifying threats from neighboring territories,<sup>2</sup> and even identifying one's own offspring for parental care.<sup>3</sup> While multimodal cues for individual identification are common across the animal kingdom, identification by a single modality is often necessary, particularly when the environment restricts access to other cues. Indeed, many species, including various mammals such as humans,<sup>5</sup> can recognize conspecifics based solely upon the unique acoustic signatures in vocalizations.<sup>1,6,7</sup> For instance, agile frogs (Rana dalmatina) call more in response to stranger calls than to familiar calls<sup>8</sup>; many birds prefer mate-emitted calls over those from a stranger<sup>9,10</sup>; and Mexican free-tailed bat (Tadarida braseliensis mexicana) mothers prefer their own pups' vocalizations over those of stranger pups.<sup>11</sup> Surprisingly though, among the large Muroidea superfamily of rodents that encompasses many laboratory model species like mice and rats, which do emit communicative vocalizations,<sup>12,13</sup> there is scant behavioral evidence that they are used to recognize individuals. Given the utility of rodents for implementing causal manipulations to study neural mechanisms, this absence has occluded progress in understanding the neurobiology of individual vocal recognition.

Common laboratory rodents emit vocalizations that exhibit individual acoustic variation<sup>12,14–18</sup> and have been suggested

to carry individual information<sup>14-16</sup> (though see Mahrt et al.<sup>19</sup>). A lack of behavioral demonstrations of individual vocal recognition may be due to numerous factors. For one, social recognition is thought to be largely chemically mediated in laboratory rodents, and thus researchers have not focused upon individual recognition through audition.<sup>20,21</sup> Another is that rodent ultrasonic vocalizations (USVs) are commonly thought to solely convey a vocalizer's arousal state,<sup>22,23</sup> so studies<sup>24-27</sup> have typically looked only for a communicative function rather than testing for individual recognition. Additionally, habituation to sounds is commonly observed in playback studies, making it difficult to characterize any sustained preference for one individual's vocalizations over another's.<sup>25,28</sup> These issues together may explain why vocal discrimination studies have been relatively rare in laboratory rodents.

Nevertheless, one rodent model for which individual vocal recognition may be highly beneficial is the prairie vole (*Microtus ochrogaster*), which forms socially monogamous pair bonds between mated adults.<sup>29</sup> These enduring bonds require partner recognition across many timescales and distances, likely involving multiple sensory modalities. Adult prairie voles do in fact emit vocalizations in both the audible and ultrasonic frequency range.<sup>30,31</sup> Therefore, we investigated the possibility that these vocalizations communicate individual identity, both acoustically and behaviorally, by designing a paradigm centered around the social experience to establish a lifelong pair bond.

1



#### Figure 1. Experimental protocol

(A) Timeline of experiments, created with BioRender.com.

(B) Spectrogram of 1 s of vocal activity from a Day 1 recording.

(C) Schematic of vocal feature extraction. Top shows vocalization spectrogram, with warmer colors indicating louder sound. Black line traces the actual fundamental frequency, which is reproduced in the Bottom panel along with the frequency fitted (red) according to a sinFM model.

(D) Distributions of duration (left) and onset frequency (right) parameters from the sinFM fits for all vocal segments from each animal pair recorded on Day 1.

#### RESULTS

# Acoustic analysis uncovers greater individual-specific rather than experience-dependent vocal variation

We focused on recording and testing USVs from adult prairie voles. We allowed pairs of male and female voles to freely interact within an arena for 30 min while recording audio and video data (Figures 1A and 1B). Individual males (n = 7) were first placed with an unfamiliar stimulus female (Day 0) and then a different unfamiliar female (Day 1) who was to become his mated partner. Following this "pre-cohabitation" (cohab) condition, the males were co-housed with their soon-to-be partners for seven days in the colony room to solidify their pair bond. Pairs were then separated for 24 h and brought back for both a playback study (Day 9, see in the following) and a "post-cohab" social interaction (Day 9).

We extracted 141,788 total vocal segments across 21 recording sessions (6,752  $\pm$  3,618 segments/recording,

mean ± standard deviation). We fit each vocal segment to a combination of linear and sinusoidal frequency modulated (sinFM) tones whose parameters (Figure 1C) are known to modulate auditory responses in rodents.<sup>32</sup> The distributions of each of the fitted parameters (Figure 1D) were coarsely similar, yet often statistically distinct from individual to individual (Pairwise Kolmogorov-Smirnov tests, with Bonferroni correction. n = 7 recordings. Duration: 19/21 significant comparisons; onset frequency: 21/21 significant comparisons; uncorrected p < 0.0024).

To visualize the multidimensional acoustic structure, we applied dimensionality reduction techniques<sup>33</sup> to embed each segment's six-dimensional parameters into a two-dimensional vocal space (see STAR Methods; Figure 2A). Three clear peaks emerged, suggesting highly used vocal subtypes. To determine whether these stereotyped vocalizations were used preferentially during particular behaviors, we manually demarcated three distinct behaviors throughout the videos: male follows female (MFF; an affiliative behavior), fight (an aggressive behavior),



#### Figure 2. Emission of vocal subtypes is linked to behavior

(A) (Top) 2D heatmap of sinFM features across all emitted vocal segments after dimensionality reduction (t-SNE, see STAR Methods). Black dashed lines show peak outlines, with all sounds within a peak being one subtype. (Bottom) Example vocalizations that fall into respective subtypes.

(B) Bar graph of number of behavioral engagements for both pre- and postcohabitation recordings. Gray lines connect the same pairs across experience. n.s., non significant.

(C) Vocal rate for vocalizations from each of the three subtypes shown in (A). Bars show mean +standard deviation.

\*, p < 0.05.



Even though pre- and post-cohab social behaviors remained similar on average, vocalization usage might still be modulated by pair bonding experience. We therefore generated vocal maps (Figure 3A) separately for each individual pair's pre- and post-cohab sessions and embedded them into the common space (Figure 2A) to quantify the differences between all possible pairs of vocal spaces. Using Jensen-Shannon divergence,<sup>34</sup> we found that vocal maps differed by 0.307–0.584 units (0.447  $\pm$ 0.063). To put this in perspective, we also pooled across all individuals to generate overall vocal spaces for all pre- and all postcohab recordings (Figure 3A, bottom right panels). Those two maps were largely similar, with a divergence of 0.326 that fell near the lowest end of all empirically measured comparisons across individuals (Figure 3B, star, n = 84. z-test, z = 1.93, p =0.03). Hence, even though individual male-female pairs vocalized differently after their own cohabitation experience, which vocalizations they emitted in those contexts were not, on average, simply dictated by pair bonding.

The lack of a discernible systematic effect of pair bonding on vocalizations could imply that each male-female encounter simply produces an independent collection of USVs, which are as variable across individual pairs as they are across cohabitation experience. If true, such randomness would not be conducive to individual vocal recognition. To test for this, we considered whether divergences derived from comparing the pre- and post-cohab vocal spaces within the same individuals were any different than those calculated from all possible comparisons across different individuals (Figure 3B). We found that these two distributions were significantly different (Kolmogorov-Smirnov [KS] test, n = 7 within,  $n = 7 \times 6 \times 2$  between. KS stat = 0.62, p = 0.01), with the latter shifted to larger divergences than the former. In fact, the residual across-individual divergence for subjects once their own within-individual pre-to-post divergence was subtracted out was significantly different from zero (signed rank, z = 3.86, p < 0.001; Figure 3C) and positive. Hence, the vocalizations of different prairie voles tended to be more distinct from each other than those emitted by the same prairie voles between their initial and pair-bonded social interactions. In fact, when using more traditional measures of vocal features in rodents (e.g., high and low frequency), we found no significant differences in pre- and post-cohab vocal features (Figure S1; paired t tests, all p > 0.05). Thus, vocal emissions, while largely









Figure 3. Vocalizations differ more between individuals than across experiences

(A) Individual t-SNE maps for each of the seven unique male-female pairing for pre-cohabitation (columns 1 and 3) and post-cohabitation (columns 2 and 4) recordings. Colored circles indicate different pairs (consistent across figures). (Bottom right) Pre- and post-cohabitation t-SNE maps pooled across all pairs. (B) Cumulative distribution of Jensen-Shannon divergences computed for pairwise comparisons of the t-SNE maps within all seven individuals from pre- to post-cohabitation (red line,  $n = 7 \times 6 \times 2$ ), and across pre- to post-cohabitation for vocal segments pooled over all individuals (star, n = 1).

(C) Cumulative distribution of voles' (n = 7) across-individual Jensen-Shannon divergences (included in black line in B) normalized by its own within-individual divergence (included in red line in C). Median (brown dot with interquartile range in gray) of this distribution is significantly higher than 0, indicating that individuals are more similar vocally to themselves than to others.

\*, *p* < 0.05.

consistent across social experiences, exhibit differences between individuals.

#### Individual identity can be decoded from sinFM features

Greater variability in the vocalizations of different male-female pairs could provide the basis for individuals to be recognized from their vocal emissions. We investigated this at the acoustic level by training a classifier (Figure 4A) to predict individual identity from vocalizations. Given the variability seen across contexts, we first limited ourselves to the vocalizations emitted from the same social context. For post-cohabitation, our classifier successfully predicted the identity of individual males with an accuracy of 0.74 ± 0.15, which was significantly above chance (n = 1,000 samples, t = 128.7, p < 0.001, chance = 0.14, Figure 4B). Hence, vocalizations from the same context were emitted with sufficiently individualized features. These were presumably dominated by the male's emissions, 35-37 but, to control for the fact that different females were present during each of the post-cohab recordings with the partner, we separately trained a classifier on the vocal segments from Day 0, when all males were exposed to the same stimulus female. We were still able to predict identity with an accuracy of 0.85 ± 0.02, which was again well above chance (n = 1,000 samples, t = 986.4, p < 0.001, Figure 4C). Thus, any female vocal contributions were unlikely to provide the only identifying information in the recording. Male prairie voles must instead emit vocalizations with enough acoustic individuality that the emitter can in principle be accurately recognized, regardless of who he vocalizes to.

While these results show that the six sinFM features can be used to identify individuals, they do not indicate which features are important. To address that, we trained and tested a series of classifiers on data wherein each feature was successively removed. We found that excluding duration, phi, or onset frequency still allowed classifier accuracy significantly above chance (n = 1,000 samples, all  $p < 10^{-3}$ , z-test with Bonferroni correction; Figure 4D), indicating that these parameters are not essential. However, removing the frequency trajectory's amplitude or frequency of frequency modulation (AFM or FFM) or its slope led to chance-level accuracy, pointing to their discriminative function. Together, these results indicate that vocal differentiation is likely reliant upon these three extracted sinFM features.

If prairie vole vocalizations truly carry systematic identifying information, we next reasoned that this should allow individual identification across days. We therefore tested whether vocalizations emitted on one day of our recordings could be used to predict individual identities on other recording days (Figure 4E).





#### Figure 4. Male identity can be determined from sinFM features

(A) Schematic of how a single-vector multi-class classifier works to predict the class of unlabeled test data (black dots) after supervised training (red, blue, and cyan dots). Colored regions depict corresponding hyperplanes.

(B) Classifier accuracy in determining which male emitted individual vocal segments based on their sinFM acoustic features while males are socializing to their tobe partners (training and testing data from Day 1). Boxes on the identity line show correct identity predictions. Values off the identity line show incorrect predictions.

(C) Same as (B), but vocal signals were from times when all males interacted with the same stimulus female (training and test data from Day 0). Red line indicates chance level.

(D) Violin plot of classifier accuracy with all 6 features ("None") or with each of the six sinFM features removed individually. Horizontal green line shows median, and vertical shows standard deviation. \*, significantly greater than chance accuracy using a z-test with a Bonferroni correction.

(legend continued on next page)



We first established the chance accuracy of individual classification by shuffling Day 1 vocal data 1,000 times (black line; 14%  $\pm$ 2% accuracy). To determine how consistently animals vocalized across days, we trained a series of classifiers using Day 0 data to see if we could predict individual identities when testing on Day 1 data. We could accurately predict individual identities on Day 1 at levels significantly above chance (n = 1,000 samples, 23% ± 1% accuracy; KS test, KS stat = 0.98,  $p < 10^{-3}$ ). Next, we combined Day 0 and Day 1 data (i.e., all data before the animals were pair bonded) to see whether our classifier could generalize when given data from multiple recording sessions and two separate females. We again found that identity could be accurately predicted at levels above chance  $(47\% \pm 3\%)$ . n = 1,000 samples, KS test, KS stat = 1.0,  $p < 10^{-3}$ ). Thus, non-pair bonded voles emit distinguishable vocalizations across both days and different social partners.

To test whether distinguishability would be retained even after pair bonding, we trained another series of classifiers on data collapsed across Day 0 and Day 1 (AllPre), but this time we assessed how accurately we could predict individual identity using Day 9 data (post). We still found that identity could be predicted at levels significantly above chance (n = 1,000 samples,  $31\% \pm 2\%$  accuracy; KS test, KS stat =  $1.0, p < 10^{-3}$ ), indicating a consistency in an individual's vocalizations that extends over longer periods of time and across social experiences. Taken together, these results demonstrate that prairie voles emit vocalizations with individual specificity over time, across experience levels, and even across social partners.

#### Prairie voles behaviorally prefer partner vocalizations

Lastly, we wanted to know whether acoustic distinguishability in principle translates in practice to behavioral recognition of the vocalizing male by his female partner. We devised a playback study to present each female prairie vole with the Day 0 USVs emitted *to the stimulus female* by either their own partner or by a stranger male. We used sound files in which every 30 s the stimulus emitted from a given speaker was either USVs from the appropriate male (partner or stranger, speaker side randomized and counterbalanced per playback session) or background noise (Figures 5A and 5B). Then, to gauge sound-motivated behavior, we assessed two types of behavioral engagement: times when the females wall-touched on the two walls with active speakers (Figure 5C, top) and times when the females actively nose poked toward speakers emitting sounds (Figure 5C, bottom).

Because rodents are known to quickly acclimate to sound playback,<sup>25,28</sup> we first looked at whether our paradigm was sufficient to elicit prolonged interest in sound playback. We found no significant decrement in wall touching (n = 6 recordings,  $F_{5,90} = 1.4$ , p = 0.22) or nose poking (n = 6 recordings,  $F_{5,90} = 1.13$ , p = 0.35; Figure 5D) over time, indicating that our design was sufficient to encourage consistent behavioral engagement

# iScience Article

over a 10-min period. Given our ability to characterize interest over an extended period of time, we next separated these responses into USV-directed behaviors (Figure 5C, left) or background noise-directed behaviors (Figure 5C, right). We found no significant preference in the proportion of time females walltouched toward USVs compared to background noise (n = 6 recordings, ranksum = 48, p = 0.14; Figure 5E). In contrast, females spent significantly more time nose poking toward USVs than toward background noise (n = 6 recordings, ranksum = 54, p =0.02; Figure 5F). Thus, female prairie voles exhibit a behavioral preference for USVs over background noise, as assessed by nose poking behavior.

We next aimed to determine whether females preferred to nose poke toward partner-emitted versus stranger-emitted sounds. Although we found no significant difference in the total *number* of nose pokes toward the partner or stranger USVs (n = 6 recordings, Wilcoxon signed rank, signrank = 10, p =0.13; Figure 6A), the total *duration* of nose poking did differ significantly, with females spending more time nose poking toward her partner's sounds (n = 6 recordings, signrank = 21, p = 0.02; Figure 6B). Thus, female prairie voles show a behavioral preference for actively trying to reach the source of her partner's USVs over stranger USVs.

In creating a paradigm that avoided behavioral habituation to playback, we scheduled blocks of time during the 10 min when only USVs from one male were playing from just one speaker (non-competing), times when USVs from both males were playing simultaneously from different speakers (competing), and times when no USVs were playing from either speaker (only background sounds). This did not appear to affect female responses to USVs, since the total time they spent nose poking toward USVs was not different between the competing and non-competing conditions (n = 6 recordings, signrank = 15, p = 0.44; Figure 6C). When it came to preference for partneremitted USVs though, females exhibited a clear preference in the non-competing condition (n = 6 recordings, ranksum = 54.p = 0.02; Figure 6D) but were only trending toward a preference in the competing condition (n = 6 recordings, ranksum = 51, p =0.06; Figure 6E), hinting that simultaneous playback schemes may partly obscure vocal preferences in rodents. Regardless, when collapsed across both playback conditions, a preference based on the proportion of time nose poking recovered a significant preference for partner-emitted compared to stranger-emitted USVs (n = 6 recordings, ranksum = 54, p =0.02; Figure 6F). Thus, female prairie voles can show their partner recognition through a behavioral preference for the sounds of their mate over a stranger male.

#### DISCUSSION

Rodents are not generally known for an ability to recognize individuals based on their vocalizations, yet by leveraging an ethologically grounded playback paradigm in prairie voles, we

<sup>(</sup>E) Cumulative density plot of classifier accuracies across different training/testing contexts. All distributions were significantly greater than chance distribution (two-sample Kolmogorov-Smirnov tests with Bonferroni correction). Horizontal gray line indicates median. In legend, text before the arrow indicates training data and text after the arrow indicates testing data. AllPre means data collapsed across Day 0 and Day 1. Post, Day 9.

Dur., duration; FFM, frequency of frequency modulation; AFM, amplitude of frequency modulation; Onset, onset frequency. \*, p < 0.05.





#### Figure 5. Female prairie voles nose poke more toward USVs than background noise

(A) Representative spectrogram showing 2 s of audio data recorded during free social interaction between a male and a female. Playback context switches at the vertical gray line (30 s mark), switching from USV playback on right speaker (top) to USV playback on left speaker (bottom). Thick horizontal white lines show when USVs were being played. Dashed white line shows 20 kHz or dividing line between audible and ultrasonic sounds.

(B) Schematic of playback organization, created with BioRender.com. During playback, sounds were concurrently played back from two speakers on opposite ends of the arena; one file containing USVs emitted by the experimental female's partner and one containing USVs from an unfamiliar male. Speaker side (left or right) and playback ordering (1 vs. 2) were randomized.

(C) Schematics of wall touch (top) and nose poke (bottom) behavior toward USVs (left) and background noise (right).

(D) Bar plot of total duration of wall touch (top) and nose poke (bottom) time in each 30 s playback bin. Bars show mean + standard deviation.

(E) Percent of wall touch time spent touching sides playing USVs (green) versus sides playing background noise (gray). Rectangle to far right denotes interquartile range (IQR) and median (horizontal black line) for "toward USVs."

(F) Same as (C), but looking at nose poke times.

\*, *p* < 0.05.

demonstrated that males emit behavior-dependent and acoustically distinct USVs that females are able to use behaviorally to identify their partner. Our discovery brings monogamous rodents into the diverse collection of species in which individual vocal recognition has been found, spanning amphibians to birds and mammals,<sup>6</sup> but no *Muroid* rodents. Importantly, as a laboratory rodent model,<sup>38–41</sup> the prairie vole adds the opportunity to uncover the causal neural mechanisms by which the auditory processing of identifying vocal features elicits individual-directed behavioral responses. We found that prairie voles emit behavior-dependent vocalizations. During aggressive behaviors (fight), prairie voles rely more upon vocalization types that are less common during non-aggressive behaviors (MFF and groom), and vice versa. Work in mice has similarly shown that male mice emit different vocalizations when being an aggressor versus when being aggressed.<sup>42</sup> One likely explanation for this is that an animal's motivation state dictates their emission of communicative sounds.<sup>43</sup> While prior work has posited that vocalizations are solely a readout of motivational or arousal state,<sup>22,23</sup> our







Figure 6. Females show a behavioral preference for partner USVs

(A) Total duration of nose pokes toward partner USVs that are during periods when speakers are simultaneously playing USVs from both the partner and stranger males (purple) and also times when only partner USVs are playing (with background noise on the second speaker). Rectangles to right show IQR for each group. Upper and lower bounds of rectangles denote 75<sup>th</sup> and 25<sup>th</sup> percentile, respectively, for both partner (blue) and stranger (orange). Horizontal black line represents median

(B) Proportion of times nose poking toward partner and stranger while both speakers are simultaneously playing USVs.

(C) Proportion of times nose poking toward partner and stranger while only one speaker at a time is playing USVs.

(D-F) (D) Number of nose pokes toward partner and stranger USV playback. (E) Total duration of nose pokes toward partner and stranger USV playback. IQRs for (D)-(F) show only partner-directed behavior. (F) Percentage of nose poking time directed toward partner and stranger USVs.

\*, p < 0.05. #, p = 0.06. (A)–(C) are Wilcoxon sign ranks. (D)–(F) are ranksums comparing partner proportions to 0.5 (chance).

work indicates that vocal emissions also exhibit acoustic individuality that is sufficient for behavioral identification of conspecifics.

Unique acoustic signatures in the vocalizations of common laboratory rodents have been debated previously,<sup>15,16,18,19</sup> but readouts in earlier behavioral studies were not able to confirm individual identification. One complication has been a lack of robust effects of a sender's vocalizations on a receiver's behavior. Playing back calls (often versus silence) elicits initial investigation of the speaker that habituates rapidly,<sup>24,27</sup> although by properly controlling motivation with operant training over many months, auditory discrimination of natural vocalizations is possible in mice.<sup>44</sup> Wild female mice can distinguish non-kin vocalizations from those of kin they had not actually heard vocalize, suggesting a heritable group-level recognition.<sup>28</sup> Similarly, behavioral studies of pup vocalizations suggest that dams can at least discriminate calls from pups of their own litter<sup>45</sup> or of different sexes.<sup>46</sup> In fact, while prairie vole pups vocalize at higher rates than pups of similar species,<sup>47,48</sup> their calls become more stereotyped between early infancy and later stages (e.g., post-natal day [P]6 to P16),49 making acoustic individuality seemingly less likely in adults. Nevertheless, our acoustic analysis and decoding results indicate that adult prairie voles endogenously exhibit individual vocal recognition, which to our knowl-

edge is among the first evidence of individual distinctiveness and recognizability in lab rodents.

Individual vocal recognition presumably evolved to facilitate differential responses to distant conspecifics with whom interactions would have divergent costs or benefits. For animals that form lifelong pair bonds, recognizing a mate's vocalizations before they are seen and discriminating them from that of a stranger could mean the difference between welcoming home a co-parent or defending their offspring from an intruder. The adaptive benefit of multimodal partner recognition may be one reason why we were able to observe a robust sign of a female prairie vole's high interest in the playback of her partner's vocalizations in the absence of seeing or smelling him-particularly after introducing a random block playback design and nose poke readout to sustain and gauge motivation, respectively.

Our block design included periods when only one speaker played USVs, while the other played background sounds, as well as periods when USVs from both speakers competed for the female's attention. Interestingly, the strongest preference for partner-emitted sounds arose when these were playing from only one speaker at a time, confirming that, similar to other rodents,<sup>27,50</sup> female prairie voles prefer USVs over background noise. However, rodent playback experiments often rely upon a competitive playback model where rodents can choose one



A further innovation allowed us to circumvent another common limitation in rodent USV studies: emissions are difficult to localize to specific animals.<sup>51</sup> We made sure that our playback used recordings from males who were stimulated with the same female so that any female-emitted vocalizations would be acoustically similar on both speakers and thus should not be the basis for a preferential response. Furthermore, playbacks were counterbalanced so that, if one female heard audio files from her partner and a stranger male, then that stranger's partner heard the same two files in her own playback session. With this, we still found a consistent nose poke preference toward each female's own partner, making it unlikely that the stimulus female's vocalizations drove behavioral discrimination. We therefore concluded that female prairie voles can recognize their partners' vocalizations, thereby establishing a behavioral foundation for uncovering the neural mechanism of individual vocal recognition.

An open question concerns whether the individual recognition we found is based on familiarity or pair bonding. To attempt to address this, we also recorded from a set of unfamiliar female-female pairs to mimic our Day 0 male-female pairings. However, without a male present, females were not highly vocal (n = 5pairs, averaging 58 vocalizations in 30 min. Data not shown), and thus we could not generate appropriate playback files. This finding is similar to work in other rodents showing that female mice, for instance, emit vocalizations at much lower rates than males.<sup>37,51,52</sup> Thus, we cannot yet exclude the possibility that a female's vocal recognition of her partner is based on social familiarity rather than social affiliation.

In either case, our study builds a foundation for studying the neurobiology of vocal recognition in rodents. Rodents are already widely studied for elucidating neural underpinnings of social behavior and recognition more broadly. For instance, "nepotopy"-wherein cells responsive to non-kin versus kin-related stimuli are spatially organized-was discovered in the rat lateral septum.<sup>53</sup> Neurons with causal roles in social recognition memory were found in mice by manipulating circuits between the hippocampus and the lateral septum or nucleus accumbens.54-56 Furthermore, neuromodulatory mechanisms to establish longterm preferences for specific individuals were uncovered by studies in prairie voles, which revealed the importance of the oxytocin and vasopressin neuropeptide systems<sup>57-59</sup> and how they change with experience.<sup>39,60</sup> Despite this broad neurobiological foundation, no work has explored the neurobiological basis for individual vocal recognition in rodents because the behavioral evidence for it has been lacking. This is likely because vocal differences are instead typically attributed to species, strain, sex, arousal, or social context-not individual identity.<sup>61,62</sup> Our paradigm in prairie voles creates an opportunity to find neural correlates of individual voices, as has been found in primates, <sup>63,64</sup> and track their integration into the social recognition circuitry that drives behavioral responses.<sup>65</sup> Intriguingly, prairie voles have an unusually large auditory cortex compared to other rodent



species,<sup>66</sup> which could reflect an evolutionary adaptation to enhance acoustic cues in the vocalizations from one's monogamous partner.

Finally, elucidating the basic science of individual vocal recognition offers translational potential. A relatively understudied clinical deficit known as phonagnosia manifests as an impairment in recognizing people by their voice.<sup>67</sup> Patient studies suggest a neural origin starting within the human temporal lobe and extending beyond it depending on whether there are dysfunctions in processing basic vocal features or the sense of familiarity generated by a vocal percept.<sup>68</sup> While species differences would need to be factored in, establishing an ethologically grounded behavioral paradigm for individual vocal recognition in rodents will make future studies of the underlying voice perception-toaction circuits possible.

#### Limitations of the study

One limitation of this work is our inability to localize vocalizations to an individual. This is a common issue in the field, with groups attempting to solve this problem by relying on solutions wherein physical location in an arena matter<sup>51</sup> or wherein the animals are constrained to a very specific physical arrangement.<sup>37</sup> To alleviate any potential impact of this limitation on our results, we used Day 0 audio data for our playback recordings such that, if playback files contained female vocalizations, every playback file would contain vocalizations from the same female, thus making male vocalizations the only distinguishing factor between files. Another limitation of this study of a nontraditional animal model is the relatively slim sample size for our playback experiment after one pair's video recording was corrupted. Lastly, the study only assessed responses to vocalizations in females, as we were unable to elicit female vocalizations in a female-female context to use in testing male behavioral responses. Nevertheless, we still uncovered significant preferences of female voles for the vocalizations of their partners, indicating that this is a robust, ethologically relevant effect.

#### **RESOURCE AVAILABILITY**

#### Lead contact

Inquiries should be addressed to the lead contact, Robert C. Liu (robert.liu@ emory.edu).

#### Materials availability

This study did not generate new unique materials.

#### Data and code availability

- Data have been deposited at GitHub and are publicly available at Github: https://doi.org/10.5281/zenodo.14617947.
- All original code has been deposited at and is publicly available at Github: https://doi.org/10.5281/zenodo.14617947.
- Any additional information is available from the lead contact upon request.

#### ACKNOWLEDGMENTS

We thank Lorra Julian and the Emory National Primate Research Center veterinary and animal care staff for vole care. We thank Danial Arslan and Jim Kwon for setting up the ultrasound recording and playback system. This work was supported by NIH grants P50MH100023 (L.J.Y. and R.C.L.), R01MH115831 (R.C.L. and L.J.Y.), 5R01DC008343 (R.C.L.), and P510D11132 (EPC).

### AUTHOR CONTRIBUTIONS

CelPress

Conceptualization, M.R.W. and R.C.L.; methodology, M.R.W. and R.C.L.; investigation, M.R.W.; data curation, M.R.W. and J.Z.; visualization, M.R.W. and R.C.L.; supervision, R.C.L. and L.J.Y.; writing – original draft, M.R.W.; writing – review and editing, M.R.W. and R.C.L.; funding acquisition, R.C.L. and L.J.Y.; resources, R.C.L. and L.J.Y.

#### **DECLARATION OF INTERESTS**

The authors declare no competing interests.

#### STAR \* METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS • Subjects
- METHOD DETAILS
  - Data collection
  - Vocal Extraction
  - Vocal plavback
  - Behavior scoring
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantifying acoustic features of sounds
  - Characterizing vocal space
  - Characterizing vocalizations during behavior
  - Characterizing acoustic discriminability
  - Characterizing female behavior
  - Statistics

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci. 2025.111796.

Received: July 31, 2024 Revised: November 11, 2024 Accepted: January 9, 2025 Published: January 10, 2025

#### REFERENCES

- Tibbetts, E.A., and Dale, J. (2007). Individual recognition: it is good to be different. Trends Ecol. Evol. 22, 529–537. https://doi.org/10.1016/j.tree. 2007.09.001.
- Jaeger, R.G. (1981). Dear Enemy Recognition and the Costs of Aggression between Salamanders. Am. Nat. 117, 962–974. https://doi.org/10.1086/ 283780.
- Jouventin, P., Aubin, T., and Lengagne, T. (1999). Finding a parent in a king penguin colony: the acoustic system of individual recognition. Anim. Behav. 57, 1175–1183. https://doi.org/10.1006/anbe.1999.1086.
- Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: animal communication in a world in flux. Trends Ecol. Evol. 25, 292–300. https:// doi.org/10.1016/j.tree.2009.11.003.
- Gustafsson, E., Levréro, F., Reby, D., and Mathevon, N. (2013). Fathers are just as good as mothers at recognizing the cries of their baby. Nat. Commun. 4, 1698. https://doi.org/10.1038/ncomms2713.
- Carlson, N.V., Kelly, E.M., and Couzin, I. (2020). Individual vocal recognition across taxa: a review of the literature and a look into the future. Philos. Trans. R Soc. B Biol. Sci. 375, 20190479. https://doi.org/10.1098/rstb. 2019.0479.



- Bradbury, J.W., and Vehrencamp, S.L. (2011). Principles of Animal Communication (Oxford University Press).
- Lesbarreres, D., and Lode, T. (2002). Variations in male calls and responses to an unfamiliar advertisement call in a territorial breeding anuran, Rana dalmatina: evidence for a "dear enemy" effect. Ethol. Ecol. Evol. 14, 287–295. https://doi.org/10.1080/08927014.2002.9522731.
- Wanker, R., Apcin, J., Jennerjahn, B., and Waibel, B. (1998). Discrimination of different social companions in spectacled parrotlets (Forpus conspicillatus): evidence for individual vocal recognition. Behav. Ecol. Sociobiol. 43, 197–202. https://doi.org/10.1007/s002650050481.
- Buhrman-Deever, S.C., Hobson, E.A., and Hobson, A.D. (2008). Individual recognition and selective response to contact calls in foraging brownthroated conures, *Aratinga pertinax*. Anim. Behav. 76, 1715–1725. https://doi.org/10.1016/j.anbehav.2008.08.007.
- Balcombe, J.P. (1990). Vocal recognition of pups by mother Mexican freetailed bats, *Tadarida brasiliensis mexicana*. Anim. Behav. 39, 960–966. https://doi.org/10.1016/S0003-3472(05)80961-3.
- Liu, R.C., Miller, K.D., Merzenich, M.M., and Schreiner, C.E. (2003). Acoustic variability and distinguishability among mouse ultrasound vocalizations. J. Acoust. Soc. Am. *114*, 3412–3422. https://doi.org/10.1121/1. 1623787.
- Sewell, G.D. (1970). Ultrasonic Communication in Rodents. Nature 227, 410. https://doi.org/10.1038/227410a0.
- Hoffmann, F., Musolf, K., and Penn, D.J. (2012). Spectrographic analyses reveal signals of individuality and kinship in the ultrasonic courtship vocalizations of wild house mice. Physiol. Behav. 105, 766–771. https://doi.org/ 10.1016/j.physbeh.2011.10.011.
- Holy, T.E., and Guo, Z. (2005). Ultrasonic Songs of Male Mice. PLoS Biol. 3, e386. https://doi.org/10.1371/journal.pbio.0030386.
- Marconi, M.A., Nicolakis, D., Abbasi, R., Penn, D.J., and Zala, S.M. (2020). Ultrasonic courtship vocalizations of male house mice contain distinct individual signatures. Anim. Behav. *169*, 169–197. https://doi.org/10.1016/j. anbehav.2020.09.006.
- Fernández-Vargas, M., Riede, T., and Pasch, B. (2022). Mechanisms and constraints underlying acoustic variation in rodents. Anim. Behav. 184, 135–147. https://doi.org/10.1016/j.anbehav.2021.07.011.
- Schwarting, R.K.W., Jegan, N., and Wöhr, M. (2007). Situational factors, conditions and individual variables which can determine ultrasonic vocalizations in male adult Wistar rats. Behav. Brain Res. *182*, 208–222. https:// doi.org/10.1016/j.bbr.2007.01.029.
- 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. https://doi.org/ 10.1523/JNEUROSCI.5054-12.2013.
- Hurst, J.L., Payne, C.E., Nevison, C.M., Marie, A.D., Humphries, R.E., Robertson, D.H., Cavaggioni, A., and Beynon, R.J. (2001). Individual recognition in mice mediated by major urinary proteins. Nature 414, 631–634. https://doi.org/10.1038/414631a.
- Keller, M., Baum, M.J., Brock, O., Brennan, P.A., and Bakker, J. (2009). The main and the accessory olfactory systems interact in the control of mate recognition and sexual behavior. Behav. Brain Res. 200, 268–276. https://doi.org/10.1016/j.bbr.2009.01.020.
- 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. https://doi.org/10.1037/a0012889.
- Ehret, G. (2005). Infant Rodent Ultrasounds A Gate to the Understanding of Sound Communication. Behav. Genet. 35, 19–29. https://doi.org/10. 1007/s10519-004-0853-8.
- 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. https://doi.org/10.1098/rsbl.2009.0317.



- Perrodin, C., Verzat, C., and Bendor, D. (2023). Courtship behaviour reveals temporal regularity is a critical social cue in mouse communication. Elife *12*, RP86464. https://doi.org/10.7554/eLife.86464.
- Saito, Y., Tachibana, R.O., and Okanoya, K. (2019). Acoustical cues for perception of emotional vocalizations in rats. Sci. Rep. 9, 10539. https:// doi.org/10.1038/s41598-019-46907-0.
- Shepard, K.N., and Liu, R.C. (2011). Experience restores innate female preference for male ultrasonic vocalizations. Genes Brain Behav. 10, 28–34. https://doi.org/10.1111/j.1601-183X.2010.00580.x.
- Musolf, K., Hoffmann, F., and Penn, D.J. (2010). Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. Anim. Behav. 79, 757–764. https://doi.org/10.1016/j.anbehav.2009.12.034.
- Thomas, J.A., and Birney, E.C. (1979). Parental Care and Mating System of the Prairie Vole, Microtus ochrogaster. Behav. Ecol. Sociobiol. 5, 171–186. https://doi.org/10.1007/BF00293304.
- Colvin, M.A. (1973). Analysis of Acoustic Structure and Function in Ultrasounds of Neonatal Microtus. Behaviour 44, 234–263.
- Terleph, T.A. (2011). A comparison of prairie vole audible and ultrasonic pup calls and attraction to them by adults of each sex. Beyond Behav. 148, 1275–1294. https://doi.org/10.1163/000579511X600727.
- Chong, K.K., Anandakumar, D.B., Dunlap, A.G., Kacsoh, D.B., and Liu, R.C. (2020). Experience-Dependent Coding of Time-Dependent Frequency Trajectories by Off Responses in Secondary Auditory Cortex. J. Neurosci. 40, 4469–4482. https://doi.org/10.1523/JNEUROSCI.2665-19.2020.
- Tabler, J.M., Rigney, M.M., Berman, G.J., Gopalakrishnan, S., Heude, E., Al-Lami, H.A., Yannakoudakis, B.Z., Fitch, R.D., Carter, C., Vokes, S., et al. (2017). Cilia-mediated Hedgehog signaling controls form and function in the mammalian larynx. D.Y. Stainier, ed. 6, e19153. https://doi.org/10. 7554/eLife.19153.
- 34. Lin, J. (2006). Divergence measures based on the Shannon entropy. IEEE Trans. Inf. Theor. *37*, 145–151. https://doi.org/10.1109/18.61115.
- Warren, M.R., Spurrier, M.S., Roth, E.D., and Neunuebel, J.P. (2018). Sex differences in vocal communication of freely interacting adult mice depend upon behavioral context. PLoS One *13*, e0204527. https://doi.org/10. 1371/journal.pone.0204527.
- 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. https://doi.org/10.1016/ 0031-9384(89)90164-9.
- Heckman, J.J., Proville, 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. https://doi.org/10. 1038/s41598-017-02954-z.
- Amadei, E.A., Johnson, Z.V., Jun Kwon, Y., Shpiner, A.C., Saravanan, V., Mays, W.D., Ryan, S.J., Walum, H., Rainnie, D.G., Young, L.J., and Liu, R.C. (2017). Dynamic corticostriatal activity biases social bonding in monogamous female prairie voles. Nature 546, 297–301. https://doi.org/ 10.1038/nature22381.
- Borie, A.M., Agezo, S., Lunsford, P., Boender, A.J., Guo, J.D., Zhu, H., Berman, G.J., Young, L.J., and Liu, R.C. (2022). Social experience alters oxytocinergic modulation in the nucleus accumbens of female prairie voles. Curr. Biol. *32*, 1026–1037.e4. https://doi.org/10.1016/j.cub.2022. 01.014.
- Forero, S.A., Sailer, L.L., Girčytė, A., Madrid, J.E., Sullivan, N., and Ophir, A.G. (2023). Motherhood and DREADD manipulation of the nucleus accumbens weaken established pair bonds in female prairie voles. Horm. Behav. *151*, 105351. https://doi.org/10.1016/j.yhbeh.2023.105351.
- Pierce, A.F., Protter, D.S.W., Watanabe, Y.L., Chapel, G.D., Cameron, R.T., and Donaldson, Z.R. (2024). Nucleus accumbens dopamine release reflects the selective nature of pair bonds. Curr. Biol. 34, 519–530.e5. https://doi.org/10.1016/j.cub.2023.12.041.

- 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. https://doi.org/10.1038/s41593-020-0584-z.
- Morton, E.S. (1977). On the Occurrence and Significance of Motivation-Structural Rules in Some Bird and Mammal Sounds. Am. Nat. 111, 855–869. https://doi.org/10.1086/283219.
- Neilans, E.G., Holfoth, D.P., Radziwon, K.E., Portfors, C.V., and Dent, M.L. (2014). Discrimination of Ultrasonic Vocalizations by CBA/CaJ Mice (Mus musculus) Is Related to Spectrotemporal Dissimilarity of Vocalizations. PLoS One 9, e85405. https://doi.org/10.1371/journal.pone.0085405.
- Mogi, K., Takakuda, A., Tsukamoto, C., Ooyama, R., Okabe, S., Koshida, N., Nagasawa, M., and Kikusui, T. (2017). Mutual mother-infant recognition in mice: The role of pup ultrasonic vocalizations. Behav. Brain Res. 325, 138–146. https://doi.org/10.1016/j.bbr.2016.08.044.
- Bowers, J.M., Perez-Pouchoulen, M., Edwards, N.S., and McCarthy, M.M. (2013). Foxp2 mediates sex differences in ultrasonic vocalization by rat pups and directs order of maternal retrieval. J. Neurosci. 33, 3276–3283. https://doi.org/10.1523/JNEUROSCI.0425-12.2013.
- Blake, B.H. (2002). Ultrasonic Calling in Isolated Infant Prairie Voles (Microtus Ochrogaster) and Montane Voles (M. Montanus). J. Mammal. 83, 536–545. https://doi.org/10.1644/1545-1542(2002)083<0536:UCIIIP>2. 0.CO;2.
- Rabon Jr, D.R., Webster, W.D., and Sawrey, D.K. (2001). Infant ultrasonic vocalizations and parental responses in two species of voles (Microtus). Can. J. Zool. 79, 830–837. https://doi.org/10.1139/z01-043.
- Warren, M.R., Campbell, D., Borie, A.M., Ford, C.L., 4th, Dharani, A.M., Young, L.J., and Liu, R.C. (2021). Maturation of Social-Vocal Communication in Prairie Vole (Microtus ochrogaster) Pups. Front. Behav. Neurosci. 15, 814200. https://doi.org/10.3389/fnbeh.2021.814200.
- Pomerantz, S.M., Nunez, A.A., and Bean, N.J. (1983). Female behavior is affected by male ultrasonic vocalizations in house mice. Physiol. Behav. 31, 91–96. https://doi.org/10.1016/0031-9384(83)90101-4.
- Warren, M.R., Sangiamo, D.T., and Neunuebel, J.P. (2018). High channel count microphone array accurately and precisely localizes ultrasonic signals from freely-moving mice. J. Neurosci. Methods 297, 44–60. https:// doi.org/10.1016/j.jneumeth.2017.12.013.
- Neunuebel, J.P., Taylor, A.L., Arthur, B.J., and Egnor, S.E.R. (2015). Female mice ultrasonically interact with males during courtship displays. P. Mason, ed. 4, e06203. https://doi.org/10.7554/eLife.06203.
- Clemens, A.M., Wang, H., and Brecht, M. (2020). The lateral septum mediates kinship behavior in the rat. Nat. Commun. *11*, 3161. https://doi. org/10.1038/s41467-020-16489-x.
- Hitti, F.L., and Siegelbaum, S.A. (2014). The hippocampal CA2 region is essential for social memory. Nature 508, 88–92. https://doi.org/10.1038/ nature13028.
- Okuyama, T., Kitamura, T., Roy, D.S., Itohara, S., and Tonegawa, S. (2016). Ventral CA1 neurons store social memory. Science 353, 1536– 1541. https://doi.org/10.1126/science.aaf7003.
- Rashid, M., Thomas, S., Isaac, J., Karkare, S., Klein, H., and Murugan, M. (2024). A ventral hippocampal-lateral septum pathway regulates social novelty preference. Elife 13. https://doi.org/10.7554/eLife.97259.1.
- Cho, M.M., DeVries, A.C., Williams, J.R., and Carter, C.S. (1999). The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (Microtus ochrogaster). Behav. Neurosci. *113*, 1071–1079. https://doi.org/10.1037/0735-7044.113.5.1071.
- Lim, M.M., Wang, Z., Olazábal, D.E., Ren, X., Terwilliger, E.F., and Young, L.J. (2004). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. Nature 429, 754–757. https://doi.org/10.1038/nature02539.
- Young, L.J., Lim, M.M., Gingrich, B., and Insel, T.R. (2001). Cellular Mechanisms of Social Attachment. Horm. Behav. 40, 133–138. https://doi.org/ 10.1006/hbeh.2001.1691.





- Borie, A.M., Young, L.J., and Liu, R.C. (2022). Sex-specific and social experience-dependent oxytocin–endocannabinoid interactions in the nucleus accumbens: implications for social behaviour. Philos. Trans. R Soc. B Biol. Sci. 377, 20210057. https://doi.org/10.1098/rstb.2021.0057.
- Hanson, J.L., and Hurley, L.M. (2012). Female Presence and Estrous State Influence Mouse Ultrasonic Courtship Vocalizations. PLoS One 7, e40782. https://doi.org/10.1371/journal.pone.0040782.
- Heckman, J., McGuinness, B., Celikel, T., and Englitz, B. (2016). Determinants of the mouse ultrasonic vocal structure and repertoire. Neurosci. Biobehav. Rev. 65, 313–325. https://doi.org/10.1016/j.neubiorev.2016.03.029.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. Nature 403, 309–312. https://doi. org/10.1038/35002078.
- Perrodin, C., Kayser, C., Logothetis, N.K., and Petkov, C.I. (2011). Voice Cells in the Primate Temporal Lobe. Curr. Biol. 21, 1408–1415. https:// doi.org/10.1016/j.cub.2011.07.028.
- Scribner, J.L., Vance, E.A., Protter, D.S.W., Sheeran, W.M., Saslow, E., Cameron, R.T., Klein, E.M., Jimenez, J.C., Kheirbek, M.A., and Donaldson, Z.R. (2020). A neuronal signature for monogamous reunion. Proc. Natl. Acad. Sci. USA *117*, 11076–11084. https://doi.org/10.1073/pnas.1917287117.

- Campi, K.L., Karlen, S.J., Bales, K.L., and Krubitzer, L. (2007). Organization of sensory neocortex in prairie voles (Microtus ochrogaster). J. Comp. Neurol. 502, 414–426. https://doi.org/10.1002/cne.21314.
- 67. Roswandowitz, C., Maguinness, C., and Kriegstein, K. (2018). Deficits in Voice-Identity Processing: Acquired and Developmental Phonagnosia. https://doi.org/10.20944/preprints201806.0280.v2.
- Van Lancker, D.R., and Canter, G.J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. Brain Cogn. 1, 185–195. https://doi.org/10.1016/0278-2626(82)90016-1.
- Tachibana, R.O., Kanno, K., Okabe, S., Kobayasi, K.I., and Okanoya, K. (2020). USVSEG: A robust method for segmentation of ultrasonic vocalizations in rodents. PLoS One 15, e0228907. https://doi.org/10.1371/journal. pone.0228907.
- Friard, O., and Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol. Evol. 7, 1325–1330. https://doi.org/10.1111/2041-210X. 12584.
- Razavi, N. (2024). Jensen-Shannon divergence. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/20689jensen-shannon-divergence.



### **STAR**\***METHODS**

#### **KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Estradiol	Fisher Scientific	17-β-Estradiol-3-Benzoate
Experimental models: Organisms/strains		
Wildtype prairie voles	Emory University	N/A
Software and algorithms		
Matlab	Mathworks	https://www.mathworks.com/
Behavioral Observation Research Interactive Software	Boris	https://www.boris.unito.it/
USVSeg	Tachibana et al. <sup>69</sup>	https://github.com/rtachi-lab/usvseg
Other		
Source code and data	This paper	https://doi.org/10.5281/zenodo.14617947
Schematic generation	Biorender	https://www.biorender.com/

#### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All experiments were conducted in strict accordance with the guidelines established by the National Institutes of Health and approved by Emory University's Institutional Animal Care and Use Committee.

#### **Subjects**

We used experimentally naïve wildtype prairie voles (*Microtus ochrogaster*) to assess the role of ultrasonic vocalizations (USVs) as a means of partner identification. All 15 voles (7 male, 8 female) were adults (postnatal day P60+ days old). Animals originated from a laboratory breeding colony derived from field-captured voles in Champaign, Illinois. Animals were housed with a 14/10 h light/dark cycle at 68-72°F with *ad libitum* access to water and food (Laboratory Rabbit Diet HF # 5326, LabDiet, St. Louis, MO, USA). Cages contained Bedo'cobbs Laboratory Animal Bedding (The Andersons; Maumee, Ohio) and environmental enrichment, which included cotton pieces to facilitate nest building. Animals were weaned at 20-23 days of age then group housed (2-3 per cage) with age- and sex-matched pups. Experiments occurred during the light cycle (between 9 a.m. and 5 p.m.).

All females were ovariectomized prior to experiments. Females were then primed with subcutaneous administration of estradiol benzoate (17- $\beta$ -Estradiol-3-Benzoate, Fisher Scientific, 2  $\mu$ g dissolved in sesame oil) for the 3 days preceding any days on which they were recorded.

#### **METHOD DETAILS**

#### **Data collection**

All recordings were conducted in a designated behavioral-recording room separate from the animal colony. To record sociallyinduced USVs, males were first removed from their home cage and placed into a plexiglass recording chamber (24.5 x 20.3 mm) lined with clean Alpha-DRI bedding. Males were recorded in three social settings. First on Day 0, a stimulus female (the same stimulus female was used with all males) was placed into the arena and free interaction with the male was audio and video recorded for 30 minutes. The following day, Day 1, males were recorded for 30 minutes with the randomly selected female that he would be pair-housed with. The male and female were then removed from the recording arena and placed into a shared home-cage, where they remained for 7 days. On Day 8, the male and female were separated. On Day 9, the female was brought back for a playback experiment (outlined below). Later on Day 9, the pair-housed male and female were reintroduced in the recording arena and audio and video data was recorded for another 30 minutes. During all recordings, animals had *ad lib* access to water gel (Clear H2O Scarborough; Scarborough, ME) and food (Laboratory Rabbit Diet HF # 5326).

A microphone (Avisoft CM16/CMPA microphone) was placed above the chamber to record audio data. Audio data were sampled at 300 kHz (Avisoft-Bioacoustics; Glienicke, Germany; CM15/CMPA40-5V), and an UltraSoundGate (116H; Avisoft-Bioacoustics; Glienicke, Germany) data acquisition system was used and integrated with Avisoft-RECORDER software to store the data. A video camera (Canon Vixia HF R800) recorded a top-down view of the chamber at 30 frames-per-second.



Given the use of a single microphone, vocalizations could not be localized to a single individual. As such, vocalizations were attributed to the relevant recording used, not an individual vocalizer.

#### **Vocal Extraction**

To extract vocalizations and vocal segments (continuous units of sound), audio files were processed with USVSEG, an open-source MATLAB-based program for detecting and extracting USVs.<sup>69</sup> Files were bandpass filtered between 15 and 125 kHz, then characterized with a timestep of 0.5 ms. Sounds with fewer than 6 samples (i.e., shorter than 3 ms) were excluded. USVSEG was modified as reported previously<sup>49</sup> to generate frequency contours – traces of the time, sound frequency, and sound amplitude at each sample within all extracted vocalizations. The contours were then further refined using custom-written MATLAB scripts.

To delineate time blocks where vocalizations occurred in a recording, we ran files through USVSEG and a postprocessing script to generate a structure indicating the time of each USV emission. The files were then broken up into two second intervals and labelled as 'background noise' (no vocalizations present) or 'USV-containing' (one or more USVs present). These two second intervals were used to generate our playback files below.

#### **Vocal playback**

To characterize female interest in USVs, females were placed into a plexiglass arena (20 cm x 24.6 cm) which had a speaker behind an opaque barrier on the left and right sides. The barrier contained holes 1 cm in diameter which the females could poke their nose through.

For each playback session, females were acclimated in the arena while 10 minutes of background noise (see below) played. The acclimation file was generated by finding all two-second intervals without vocal emissions in a control audio file (a male-female interaction not used for our experiment) and combining a random subset of the intervals into a 10-minute acclimation file solely containing background noise. This file was then played back on both speakers simultaneously during the acclimation period prior to each recording.

We used audio data from the male-stimulus pairings on Day 0 to generate playback files for preference testing. Day 0 was chosen because all males were interacting with the same stimulus female on Day 0, and thus any contamination from female vocalizations would be from the same female across all playback files.

During the test period, females heard 10 minute audio files, consisting of blocks of background noise and blocks containing USVs. To generate blocks of background noise, 15 unique 'background-noise' intervals were randomly selected and combined to generate one 30-second file. Background noise periods include rustling and other noise generated by the voles that did not include USVs. This process was repeated 10 times to generate 10 unique background-noise blocks for a single recording. The same process was followed to generate blocks of 'USV-containing' files, using two second intervals with vocalizations. These 20 blocks were then consolidated into a single 10-minute playback file using the ordering seen in Figure 4. Whether Order A corresponded to the partner file or the stranger file was randomly assigned. Audio files were manually checked by an experienced observer to ensure accurate file generation.

Which speaker the stranger and partner sounds came from was randomized, but sides were kept consistent within an individual playback session to mimic a more realistic scenario wherein the males were not inexplicably teleporting across the arena. Playbacks were counterbalanced such that when possible, two different females heard the same two playback files. E.g., the playback files for the female from pair 1 were from male 1 (partner) and male 3 (stranger). The same files were played back to the female from pair 3, such that the sounds from male 1 became the stranger sounds, and the sounds from male 3 became the partner sounds. One planned pairing did not occur (pair 8), because the male needed to be removed from the study after the Day 0 recordings. Thus, the female from pair 6 heard playback sounds from partner male 6 and stranger male 8, which was not counterbalanced with an eighth partnered female.

#### **Behavior scoring**

For recordings of free interaction between a male and female, a trained observer used BORIS behavioral annotation software<sup>70</sup> to label the start and stop times of each of the behaviors represented here. Male Follow Female (MFF) was defined as times when the male was behind the female while both voles were traveling in a forward direction. Fight was defined as times when the animals were engaged in a rough and tumble situation. Male self-groom (Groom) was defined as times when the male was stationary but moving his mouth or paws around on his own fur.

For each playback recording, an observer used the video data to score each time a female touched either of the two walls containing a speaker with a paw, as well as all times the female nose-poked towards the speaker (Figure 5C) The time each behavior began and ended was recorded through Boris behavioral annotation software.<sup>70</sup> Unfortunately, one video file was corrupted, so the video from pair 7 was not scored. Observers were blind to vocalizer identity as well as when and which speakers were playing vocalizations during scoring. Matlab scripts were subsequently used to align the recorded behaviors to vocal playback on each speaker.

#### **QUANTIFICATION AND STATISTICAL ANALYSIS**

#### Quantifying acoustic features of sounds

Using custom written code (MATLAB), the audio data corresponding to each individual vocal segment was used to extract a fundamental frequency, which was fit to a linearly and sinusoidally modulated function (sinFM) with six features<sup>32</sup>: onset frequency (f<sub>0</sub>),





amplitude of frequency modulation ( $A_{fm}$ ), frequency of frequency modulation ( $F_{fm}$ ), sine phase at sound onset ( $\phi$ ), linear rate of frequency change (slope), and length of sound (duration) (Figure 1B):

$$f(t) = f_0 + f_{slope}t + A_{fm} \sin(f_{fm}2\pi t + \phi)$$

As a way to examine individual variability in vocal features, we compared the distributions of duration and onset frequency across all individual recordings on Day 1 using pairwise Kolmogrov-Smirnov (KS) tests and a posthoc Bonferroni correction for multiple comparisons (uncorrected p < 0.0024 for significance).

To extract the raw acoustic features of our vocalizations (Figure S1), we used our extracted frequency contours (see: *Vocal Extraction*) to characterize seven unique features for each vocalization. We extracted the duration of each sound, the median, high, and low frequencies, the bandwidth, the slope, and the number of harmonics.

#### **Characterizing vocal space**

To visually depict the vocal space across all six features of each vocal segment, a t-distributed stochastic neighbor embedding (t-SNE) method of dimensionality reduction was used to project the 6-dimensional vocal representations into a 2-dimensional space (Figure 2). t-SNE was chosen to allow a non-linear projection of our 6-dimensional vocal data into a 2-dimensional space. This analysis combined all data from male recordings with his partner (or to-be partner) on Days 9 and 1, respectively. Then, a density function<sup>33</sup> was applied to the data to turn the data points into a probability density function across our 2D vocal space. Thus, a single map (1001x1001 bins, gaussian smoothed with standard deviation of 34 bins) based on the sinFM parameters of all vocalizations across all males was generated<sup>33</sup> (Figure 2A). Then, to compare between contexts or individuals, we plotted either data across all males within the context (Figure 3A, bottom right), or we plotted just the data from a single recording within a single context (Figure 3A). These data were plotted using the original combined tsne mapping locations, rerunning the density function for each set of data. Differences between any two vocal maps were measured by the Jensen-Shannon (JS) divergence,<sup>71</sup> as JS provides a bounded (0-1) metric to calculate the dissimilarity between pairs of probability distributions.

#### **Characterizing vocalizations during behavior**

First, we quantified the total number of behaviors engaged in by each pair. We used the output of Boris to tabulate the total number of times each animal engaged in each behavior, then compared the count data from Day 1 to Day 9 to see whether behavior changed with social familiarity.

To assess the relationship between vocalizations and behavior, we first manually outlined the tsne peaks by outlining the red portions of the plot (Figure 2A, dashed black outlines). Next, we found all vocalizations falling within each of the three visible peaks. For each recording, we then extracted each instance of our three classified behaviors (MFF, Fight, MG) and determined the vocal rate for sounds within each vocal peak that were emitted during the behavioral instance. This was repeated for every vocal instance and every behavior across Day 1 and Day 9 (so all interactions with the partner female). We then averaged this value across all instances to get the average vocal rate for each vocal peak within behavior for each animal.

#### **Characterizing acoustic discriminability**

To determine whether individual animal identities could be determined based solely on the features of emitted USVs, we generated a series of 1000 multi-class single vector machine (mcSVM) classifiers. Classifiers were provided with the six sinFM features of each vocal segment plus the identity of the recording pair and tasked with identifying the recording pair given the features of novel vocal segments. For each of the vocal features listed above, values were normalized between 0 and 1 to put all features on the same scale. Then, to train each classifier, 93 vocalizations from each male were randomly selected as training data, and 31 vocalizations were selected as testing data (75 and 25%, respectively, of the least vocal recording, a stimulus recording containing 125 USVs). Each classifier was trained with data from all seven males, then tested on the ability to predict the identity of the emitter of individual, untrained vocalizations. This was replicated 1000 times per comparison type to generate a range of classifier accuracies.

To determine the generalizability of our classifiers, we ran a series of control analyses. First, we reran our within-day analysis using the pre-cohabitation data (Day 1), but for each series of 1000 classifier runs, we excluded a separate sinFM feature from both the training and testing data (Figure 3D). This allowed us to determine the role of each of the six sinFM features in our classifier accuracy.

Secondly, we modified the training and testing data across a new series of four new types of classifiers (Figure 3E). First, we generated a classifier wherein we used the data from Day1 for both our testing and training data, but we shuffled the Pair IDs corresponding to the training data. Thus, by decoupling the data from the recording, we could determine how often we would accurately predict our pair IDs by chance. Next, we generated a series of 1000 classifiers using combined data from Days 0 and 1 data to predict identity on Day 9 to assess whether vocalizations in the early recordings are predictive of vocal features following pair bonding. Third, we generated a series of 1000 classifiers using data 1 as both our training and testing data to determine how consistent vocal emissions are between days within a male. Lastly, we generated a series of classifiers using Day 0 data to predict identity on Day 1. This allowed us to assess consistency on a short timescale but also with two distinct social partners.





#### **Characterizing female behavior**

Female behavior during playback was manually scored by a trained observer (see above). Proportions of time responding to USVs as they were playing out from each speaker was then calculated as

Total time directed towards speaker playing partner USVs Total time directed towards speakers playing USVs

To characterize changes in female interest over time, the total duration of nose-poking was calculated in each 30-second bin. A repeated-measures ANOVA was used to determine an effect of time of nose-poke duration.

To determine whether either wall touch or nose poking was sufficient to capture a vocal preference, we calculated the total proportion of each behavior that occurred towards a USV-emitting speaker versus the proportion that occurred toward a speaker emitting only background noise (Figures 5E and 5F).

#### **Statistics**

Statistical results are presented in the Results text, including tests used, n values and what they represent. Data are represented as mean  $\pm$  standard deviation unless otherwise indicated. Significance was set at p < 0.05 unless otherwise reported. Comparisons across feature distributions were conducted with a one-way ANOVA with a posthoc Bonferroni correction for multiple comparisons. Comparisons between vocal maps were conducted with Jensen-Shannon divergence. Distributions of within-animal divergences in vocal maps to between-animal divergences in vocal maps were compared via a KS test. The distribution of between-animal divergences was compared to a central value of 0 using a ranksum test. Distributions of classifier accuracies were compared to chance using a z-test with a post-hoc Bonferroni correction. CDFs were compared using KS tests with a Bonferroni correction.

Comparisons between female behavioral preferences and chance preferences (50%) used a t-test. The effect of time on nosepoking behavior was assessed with a repeated-measures ANOVA. Raw acoustic features were compared from pre- to post-cohabitation using paired t-tests. A p-value of 0.05 was used as the threshold for significance.

For Figures 5E, 5F, and 6, we used nonparametric tests to account for small sample sizes. Distributions of proportions were compared to chance (50%) using a Wilcoxon ranksum. For comparing the raw count and duration of nose poking between conditions, a Wilcoxon signed rank test was used.

All statistical analyses were conducted in MATLAB (Mathworks).