Science
AAAS

Supplementary Materials for

Cultural transmission of vocal dialect in the naked mole-rat

Alison J. Barker[1]*, Grigorii Veviurko[1], Nigel C. Bennett[2], Daniel W. Hart[2], Lina Mograby[1], Gary R. Lewin[1]*.

*Correspondence to: grlewin@mdc-berlin.de; Alison.Barker@mdc-berlin.de

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Animal care and maintenance

Naked mole-rats (*Heterocephalus glaber*) were maintained in a humidity (50-70%) and temperature (30-32 °C) controlled environment, under low illumination levels. A diet of tubers (primarily sweet potatoes, celery root and carrots) was provided daily (*ad libitum*) and supplemented weekly with ProNutro (Bokomo). Animals were housed by colony in a series of custom designed interconnected plastic chambers (Fräntzel Kunstsstoffe, Rangsdorf, Germany). All experimental procedures and husbandry protocols were approved by the local governmental authorities in Berlin (Landesamt für Gesundheit und Soziales, Berlin, License: G0196/17). Colonies in South Africa were housed in tunnel systems with several plastic chambers serving as food storage, toilet and sleeping areas and connected by acrylic glass tunnels. They were fed a variety of chopped vegetables (primarily sweet potatoes, cucumbers and carrots) and supplemented weekly with ProNutro (Bokomo). Nesting material consisted of wood shavings. The ambient room temperature ranged from between 29 and 32 °C and humidity from 50-70%. The experimental protocol was approved by the Animal Ethics Committee of the University of Pretoria (ECO73-17). All the naked mole-rats used in this study are descended from multiple colonies captured by Prof. Jenny Jarvis primarily in Mtito Andei, and Lerata Kenya and constitute a mixed parentage (*1*). Detailed information related to animals from each colony can be found in fig.S3.

Audio recordings

Audio recordings were made from a total of 166 animals from seven colonies; six housed at the Max Delbrück Center for Molecular Medicine in Berlin, Germany and one at the University of Pretoria, South Africa. Audio recordings were acquired using Sennheiser MKH8020 microphones and a Behringer U-Phoria UMC1820 audio interface with Avisoft-Recorder Software (Avisoft Bioacoustics, Glienicke/Nordbahn) (sampling rate 32 kHz, 24 bit resolution). Of recorded soft chirps, 2,526 soft chirps did not pass our quality control checks and were removed as outliers, the remaining 36,190 soft chirps were used for analysis. Raw sound recordings were visualized as spectrograms for further analysis using Avisoft-SASLab Pro software (Avisoft Bioacoustics, Glienicke/Nordbahn) or with custom python scripts.

Code availability

All custom-written scripts are archived here: DOI 10.5281/zenodo.4104396.

Audio data processing and feature extraction

Raw audio recordings were digitized as wav files which were then segmented into individual sounds using a custom developed “SoundSplitter” program. Following the segmentation of putative individual sounds, wav file segments relating to single soft chirps were converted into spectrograms as described in detail below and manually confirmed by the experimenter. Soft chirp spectrograms were then traced using a custom script to identify and generate a new spectrogram file containing only the fundamental (lowest) harmonic frequency. Feature extraction was

performed on both raw sound files and spectrograms and combined parameters used for training machine learning classifiers. A schematic of feature extraction can be found in fig.S1.

SoundSplitter

Audio recordings (wav files) were processed with a custom python "SoundSplitter" script to segment individual vocalizations from a longer recording often containing multiple vocalizations, and some background noise. In brief, the algorithm parses the entire audio recording into bins of 23 ms, extracts forty Mel-frequency cepstral coefficients to use as a features vector, and designates each bin as sound or noise using a Random Forest Classifier adapted from the Python3 sklearn module (with default parameters and 256 trees) (*24*). All splits were manually validated by experimenters with a built-in validation step.

Spectrogram generation

Audio recordings (wav files) were converted into spectrograms using the short time Fourier transform (STFT) function from the Python librosa library using default parameters and bin size set to 512.

SoftTrace

We developed a "SoftTrace" algorithm for extracting the features of individual soft chirps from a spectrogram image, using a convolutional auto encoder (CAE) trained on manually traced soft chirps. CAE is a neural network composed of three convolution layers with max-pooling followed by symmetrical deconvolution and up-sampling layers (*25*). All soft chirp spectrograms were traced using our CAE script to generate a new image file of the first harmonic of the soft chirp, which was used for feature extraction. Soft traces were passed through an automated filter to remove outliers (i.e. other non-soft chirp sound types that occasionally occurred during the recording) and all soft chirp traces were manually validated for accuracy before proceeding to feature extraction. (see fig. S1).

Feature extraction

For comparative analysis of soft chirp sound parameters we extracted eight features: (**1**) *pitch*, (**2**) *wiener entropy*, and (**3**) *zero crossing rate* using the librosa library (*4*) and (**4**) *asymmetry*, (**5**) *duration*, (**6**) *height*, (**7**) *peak frequency* and (**8**) *slope* using a custom script designed to work with soft chirp traces made from spectrogram images. Features (**1**)– (**3**) were extracted from the original soundwave files (down sampled from 32,000 Hz to 22,050 Hz) using the librosa library for audio data processing (*26*). (**1**) *Pitch* was computed using the piptrack function from the librosa library. The frequency range was set to [2kHz, 8kHz] and window size 10ms. Other parameters were set at default values. Determination of pitch differed from the extraction of the fundamental frequency in that the pitch determination weighted the contribution of all harmonics. (**2**) *Wiener entropy* measures how close the soundwave signal is to white noise. It was computed using implementation from the librosa library with bin size set to 10ms. (**3**) *Zero crossing rate* is a commonly used feature for audio data analysis counting how many times the soundwave changes in sign. The built-in

librosa zero_crossings function was used to compute this feature. Parameters from the soft chirp trace were computed directly from the spectrogram image and determined as follows: (**4**) *Asymmetry* describes the vertical (frequency) axis difference between the left- most and rightmost points of the soft chirp harmonic, ($Start_{frequency} - End_{frequency}$). (**5**) *Duration* is defined as the length of the soft chirp on the horizontal (time) axis and calculated by subtracting the start point from the end point of the soft chirp in (s) ($End_{time} - Start_{time}$). (**6**) *Height* describes the difference in the vertical (frequency) axis between the lowest and highest points of the soft chirp calculated by subtracting the (start or end frequency, whichever has the lowest value) from the peak frequency ($Peak_{frequency} - (End_{frequency}/Start_{frequency})$). (**7**) *Peak frequency* and (**8**) *slope* were computed from a parabolic fit to the soft chirp spectrogram image. Using the Python3 numpy.linalg module, a parabola $y = -(a(x-h)^2 + k)$ was fit to each soft chirp such that the mean square error was minimized. The frequency of the vertex (k) was taken as the peak frequency and slope as the coefficient, a preceding x^2.

Machine learning classifiers

Using the Random Forest Classifier module from the librosa library classifiers were trained on the eight extracted soft chirp features (described above, fig. S1). The classifier used 256 trees with additional parameters set to default settings (*24,26*). Confusion matrices were obtained by averaging cross validation results and display the prediction strength of the classifier as a fraction (i.e. the fraction of times the predicted data label matches the actual data label, where a value of 1 denotes a 100% correct prediction rate). Data splits for cross-validation testing colony dialects were performed on individual animals, i.e. all sounds from each animal were either in the training or testing dataset. This step was necessary in order for cross validation results to test the ability of the classifier to learn intercolonial differences rather than memorize individuals. For prediction of other features (i.e. age, rank, sex, body mass and body length) cross-validation was also split by individual animal id. For cross-validation testing of individual soft chirp recognition, soft chirps were split by recording date. To classify colony dialects of foster pups, the trained colony classifier was used to test soft chirps recorded from foster pups and output a classification prediction rate (as a percentage) for each colony. For each foster pup the colony classifier was trained to distinguish between 5 colony dialects (including foster and birth colonies). In addition to confusion matrices, we provide a list of feature importance which show the impact each feature has on the soft chirp classification (Fig. 2C). Importance calculation was implemented in the scikit-learn library. Prediction accuracy for Colony M was on average 76.18% $\pm$ 6.7, SEM).

Statistics

All statistics were performed using GraphPad Prism 8 or custom python scripts. Data were tested for normality (D'Agostino-Pearson test) and for homoscedasticity (Bartlett's test.) For multiple comparisons normally distributed datasets were tested with a one-way ANOVA and post-hoc tests performed with Tukey's multiple comparisons test. For nonparametric data Kruskal-Wallis test with Dunn's multiple comparisons test were used. For single comparisons of normally distributed data unpaired t tests were used. Confusion matrices were generated using standard python libraries. For data shown in fig.S7 (A-D) linear regressions were plotted in Prism. Residuals for all features are also plotted in fig. S7 (E-H). Significance values are reported as: * P value $\leq$ 0.05; ** P value $\leq$ 0.005; *** P value $\leq$ 0.0005. All error bars are standard error of the mean (SEM).

Generation of playback and artificial stimuli

Playback stimuli: Playback soft chirps were selected to be representative of each colony. A 10 s excerpt of soft chirps were compiled from a rank and weight matched individual from each colony and the temporal sequence of soft chirps was maintained from the original recording. A single trial consisted as presentation of the entire playback stimuli (between 10-18 soft chirps). Soft chirp sequences were followed by 10s of silence before repetition. Artificial stimuli: Artificial stimuli were generated using the graphical synthesizer function in SASLab Pro software (Avisoft Bioacoustics, Glienicke/Nordbahn). Duration was kept constant at 120 ms for all stimuli. Artificial stimuli were designed for both Colony B and Colony T using empirically determined mean frequencies and mean asymmetries for each colony. We also generated a pure tone of 4.5 kHz (Colony B, mean peak frequency) and peak frequency-doubled stimuli (Colony B peak frequency, 9.0 kHz). For frequency – doubled artificial stimuli, the mean asymmetry values for the colony were maintained. For all artificial stimuli, a sequence of 8 soft chirps constituted a single trial. Individual soft chirps were presented at intervals ranging from 0.5- 3s. Soft chirp sequences were followed by 10s of silence before repetition.

Behavioral experiments

Place Preference Assay

Four subordinates from Colony B and five from Colony T were used for all behavioral experiments. Animals were habituated to the behavioral chamber (as shown in Fig. 3A top) consisting of three interconnected chambers for ~ 10 minutes prior to the start of the experiments. Each chamber (left, right) contained a microphone (Sennheiser MKH8020) and loudspeaker (UKHONK Mini USB Speaker, HK-5002) connected to a laptop for simultaneous audio recordings and audio playback presentation. All individuals tested were in the worker class (rank 3-5). For each animal a minimum of 36 trials for each audio stimulus was performed with each trial containing between 8-18 soft chirp presentations. During the behavioral trial the animal was allowed to move freely throughout all chambers, and sound presentation begun only when the animal was in the central chamber. The chamber with sound presentation (left or right speakers) was pseudo-randomly alternated. Each playback stimulus was presented both in the right and left chambers with equal frequency and trials from all chambers averaged to control for any inherent chamber preferences. Time spent in each chamber was logged manually by a partially-blinded experimenter (via a handheld timer) as the animal transitioned into the chamber during times of audio stimulus presentation. Soft chirp responses made by experimental animals were analyzed offline from audio recordings made during the behavioral trial. All experiments were performed on two separate experimental days to ensure repeatability. A Place Preference Index was calculated as the (amount of time spent in sound chamber – the amount of time spent in non-sound chamber) divided by (the total amount of time spent in either sound presentation chamber, left or right). Thus, maximal avoidance of sound stimuli would be scored with a value of -1 and maximal preference for the sound stimulus would be scored with a value of 1. A soft chirp occurring within 0.5 s of a presented playback stimulus was counted as a response. Response rate was calculated per animal per day (averaging all trials for each stimulus) and normalized to the maximal response

per animal. Response rates were then averaged across animals and across days for final statistical analysis. As part of the habituation process, bedding from the home colony was evenly distributed throughout the behavioral apparatus. To test that bedding was not providing an olfactory cue that may influence soft chirp responses, we performed control behavioral experiments with the six animals (four from Colony B and two from Colony T) with bedding from a foreign colony (Fig.3F). All playback stimuli were presented at the same volume for all trials and measured at 81.9 dB (+/- 1.5dB, SEM) in the sound presentation chamber.

Forced Choice Assay

A modified version of the Place Preference Assay was also performed in which audio playback from two colonies was presented simultaneously. We quantified the place preference in each chamber using the place preference index as described above (see fig. S8).

Cross fostering experiments

Two separate cross fostering experiments were performed involving three animals. For all cross-fostering experiments foster pups were washed with warm water and coated in a slurry of fecal matter removed from the foster colony before being placed in the new colony, and identifying marks made on their forepaw digits which could be continually tracked until they were old to chip with RFID chips at ~ 6 months of age. In the first experiment two orphaned pups (pups Jo and pup Da) born into Colony S, were cross fostered into two separate colonies that both had litters within one week of the birth of pups J and D (Colonies T and M respectively). Audio recordings were made at periodic intervals and pups were frequently monitored to ensure survival in the colonies. In a second cross foster experiment a pup (pup Mi) that was abandoned from Colony T was fostered into Colony M within the first postnatal week of life. Several foster siblings from Colony M also survived, namely pups Ob and Ny, and were tracked along with pup Mi with periodic audio recordings. After six months of fostering, colony dialects were tested.

Queen transitions

In Colony S, the first queen, who had reigned stably, producing several litters died due a pregnancy related complication, which was confirmed upon autopsy. The ascending queen was considered established once she gave birth to her first litter. She was attacked following the birth of the first litter by several males within her colony and had to be euthanized due to her severe injuries (~1 month after the birth of the pups). Two surviving pups from this queen that were cross-fostered (pups Jo and Da in Fig. 4). In Colony B, the queen was also attacked and overthrown by colony members (fig.S12). Additional soft chirp recordings were made in the three months following her overthrow and constitute the anarchy phase described in fig.S12.

Hierarchy assessment

To reliably assess the rank of individuals within colonies, we modified an assay for dominance in rodents (*28*, *29*). When two naked mole-rats approach each other head on in a tunnel, the more dominant individual is often observed to climb over the more subordinate individual (*29*). Using this principle, we established a behavioral ranking test, which was reliably able to predict the

queen, who is morphologically distinct and therefore easily visually confirmed. In brief, two plastic chambers were connected via a transparent plastic tube. Two naked mole-rats were placed in each chamber and allowed to move freely between the two chambers, when the animals entered the tube simultaneously the identity of the mole-rat that climbed over the other was recorded. Hierarchy was assessed in a winner take all, single elimination strategy (with a minimum of three trials for each pairing). Animals were pseudo-randomly selected for pairing and hierarchies were tested over multiple months to ensure animals were tested with equal frequency. In some cases, the colony was split into lower and higher-ranking individuals based on previous hierarchy tests to perform a more fine-grained ranking of individuals. A Ranking index (R.I.) was defined as follows: (number of wins) divided by (the total number of behavioral trials). Ranking indices were normalized to maximum values for each colony to allow for comparisons across colonies. Thus, the winner, the individual with the highest fraction of wins would have ranking index of 1. Ranks were assigned as rank 1, R.I. ≥ 0.8, rank 2, R.I. ≥ 0.6, rank 3, R.I. ≥ 0.4, rank 4, R.I. ≥ 0.2, rank 5, R.I. < 0.2 (as shown in fig. S6). The queen was always assigned a rank of 1 when testing rank in our dialect classifiers.

Supplementary Text

Body size and its influence on soft chirp sound features

Previous reports have suggested that the frequency of the soft chirp may be highly dependent on body mass (*16*, *20*) and inverse correlations between vocal pitch and body length have been reported in numerous mammalian species (*29*). Surprisingly, we observed a significant positive correlation between soft chirp vocal pitch and both body length and body mass (fig.S7A,B; R^2= 0.189, P = 0.0006, body length and R^2= 0.295, P = 0.0007, body mass). In some cases, the queen, who is consistently the largest individual (in both body mass and length), displayed the highest frequency soft chirps (fig. S7K,L).

Queen influence on colony dialect

We were able to examine a second colony, Colony B which experienced an overthrow of the queen. In this case we could confirm that the queen was killed by other colony members and we observed a similar loss of soft chirp colony dialect cohesion in the three months following her death (fig.S12A,B) as also observed in Colony S (Fig.4I-K). This new result provides further evidence that the queen's presence is necessary for the maintenance of colony dialects. Comparisons of individual soft chirp variability between queens, breeding males and representative subordinates from five colonies revealed the highest variability in individual soft chirps (comparing peak frequency and asymmetry, fig.S5) from the breeding males, which except for the queen are the only reproductively active members of the colony (*1, 28, 30*). Increased vocal variability in the breeding males suggests a link between vocal plasticity and hormonal state. We hypothesize that a similar mechanism may be at work when the queen is lost, as colony-wide reproductive suppression is also lost (*28,30*) and thus accompanying physiological changes may contribute to overall variability in individual soft chirps.

Developmental acquisition of soft chirps

Young animals (from 1-3 months of age) produce at least three distinct sounds (pup combo, pup squawk, pup cheveron), not observed in adults. Additionally, several juvenile versions of adult sound types namely the upsweep, downsweep, phee and soft chirp begin to appear around ~1 month of life but do not recapitulate adult sound features until ~ 6 months of age (*6*, *15;* fig.S9-S10). Developmental transitions in sound type usage may have a purely anatomical basis, due to an immature vocal tract or may represent a type of vocal learning: either as an active sensory motor learning process where practice of vocalizations is required for refinement (i.e. vocal production learning), or from context learning (i.e. usage learning, where pup sounds may still be vocally possible for adults but animals learn to use only adult sounds as they mature) (*23*). While the underlying mechanisms are not yet known, it is clear that during early life the vocal repertoire of the naked mole-rat differs from the adult, suggesting the possibility of a developmental window for acquisition of colony dialects.

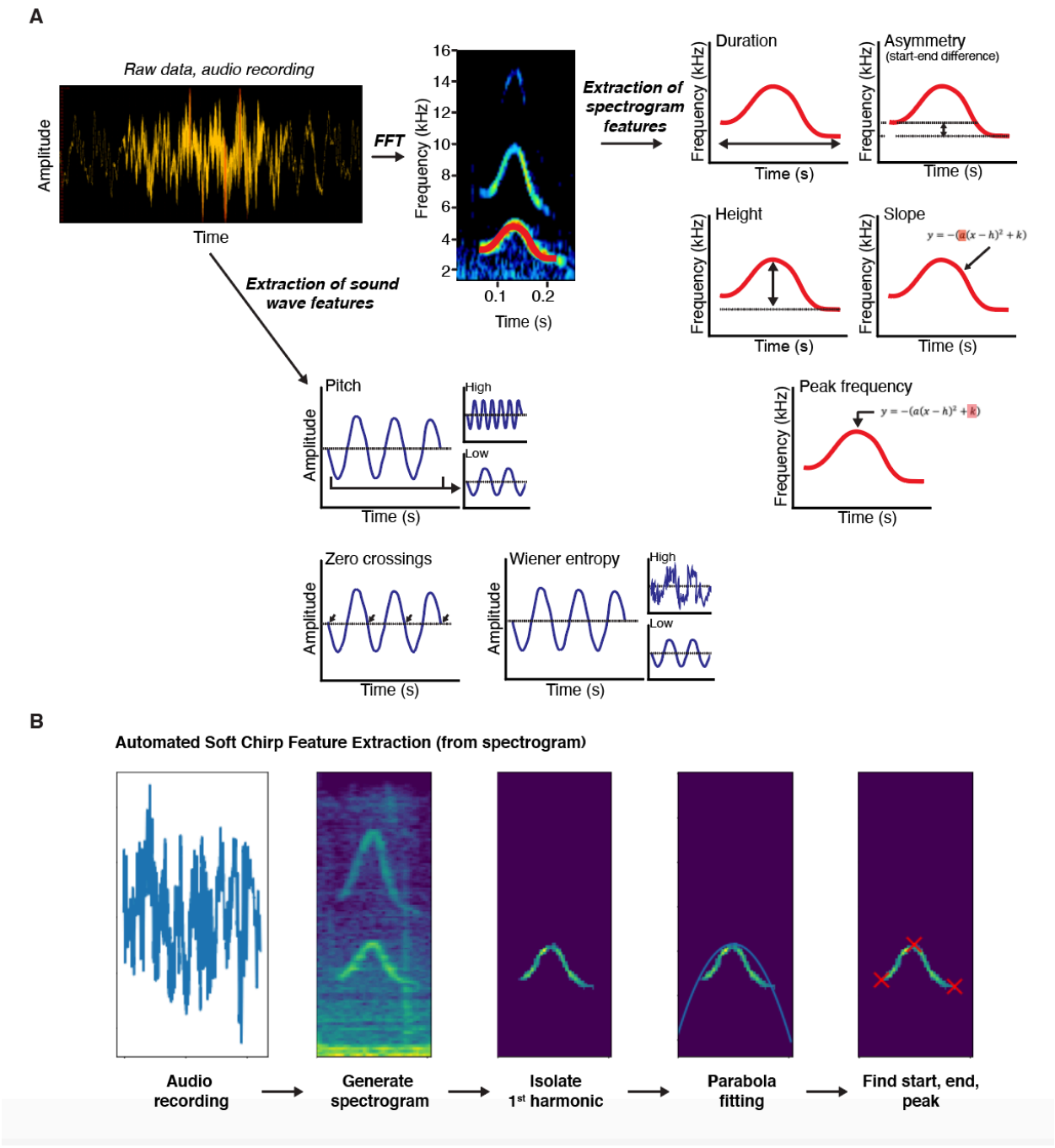

Fig. S1.
Workflow for soft chirp feature extraction. (**A**) Features from both the soundwave and spectrogram are used. (**B**) Example of automated soft chirp spectrogram feature detection from custom -written scripts.

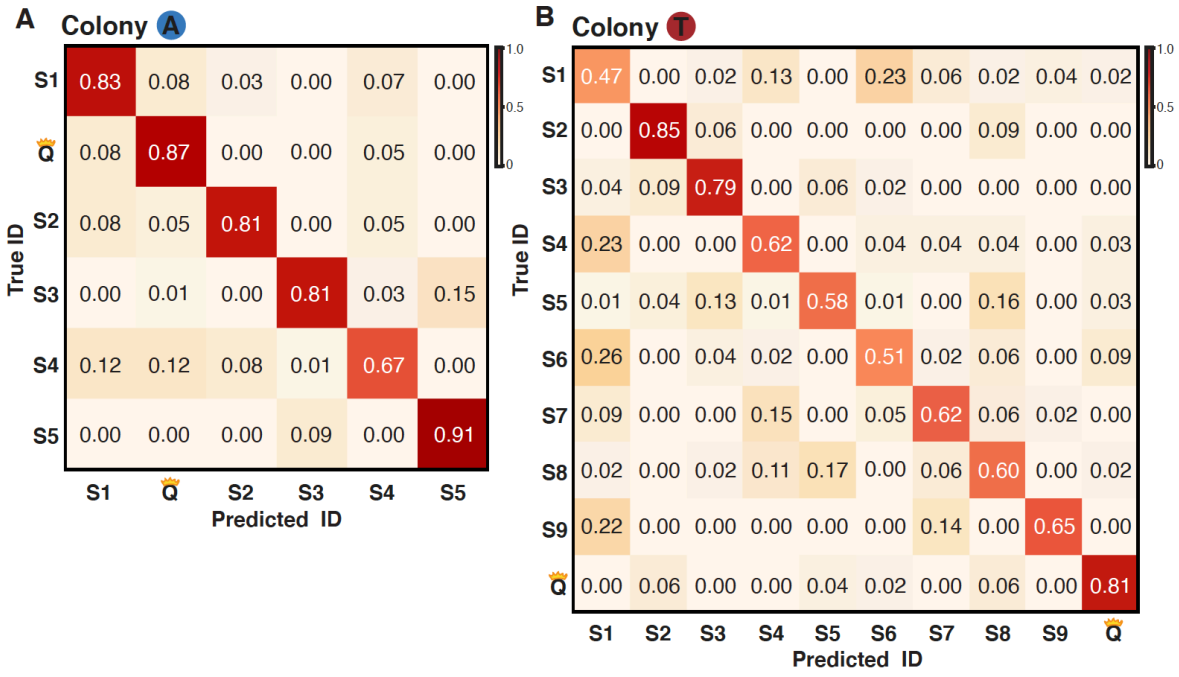

Fig. S2.

Soft chirps contain information that is sufficient to distinguish individuals within a colony. Confusion matrices show prediction success rates for Random Forest classifier trained on individual identity with eight soft chirp features. Results from two additional colonies are shown here. (**A**) Colony A. (**B**) Colony T. S denotes subordinates from each colony.

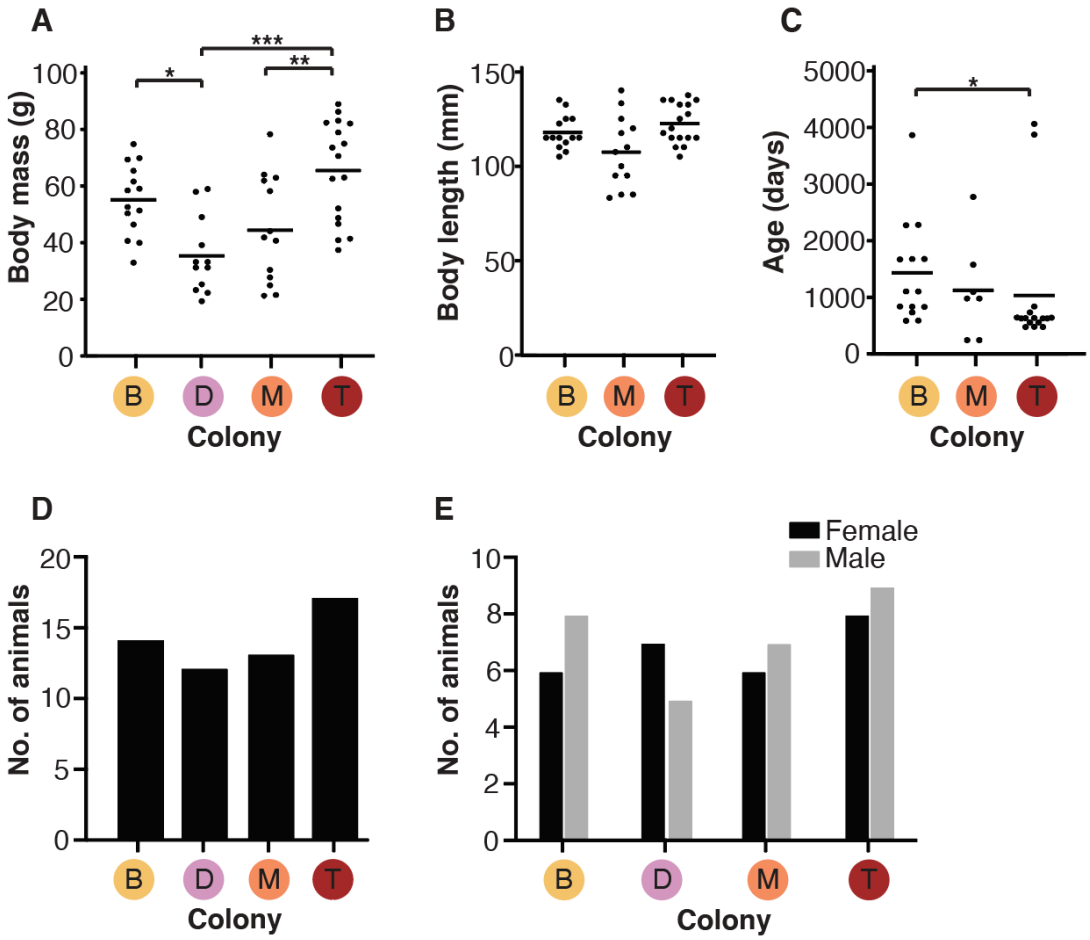

Fig. S3.

Colony biometrics. (**A-C**) Distribution of body mass (A), body length (B) and age (C) across colonies. (D) Number of animals recorded from each colony. (**E**) Sex distribution by colony (one-way ANOVA, Tukey's multiple comparisons test in (A, B), Kruskal-Wallis test with Dunn's multiple comparisons test in (C); * P <0.05, ** P < 0.005, *** P,< 0.0005 . Error bars, SEM.

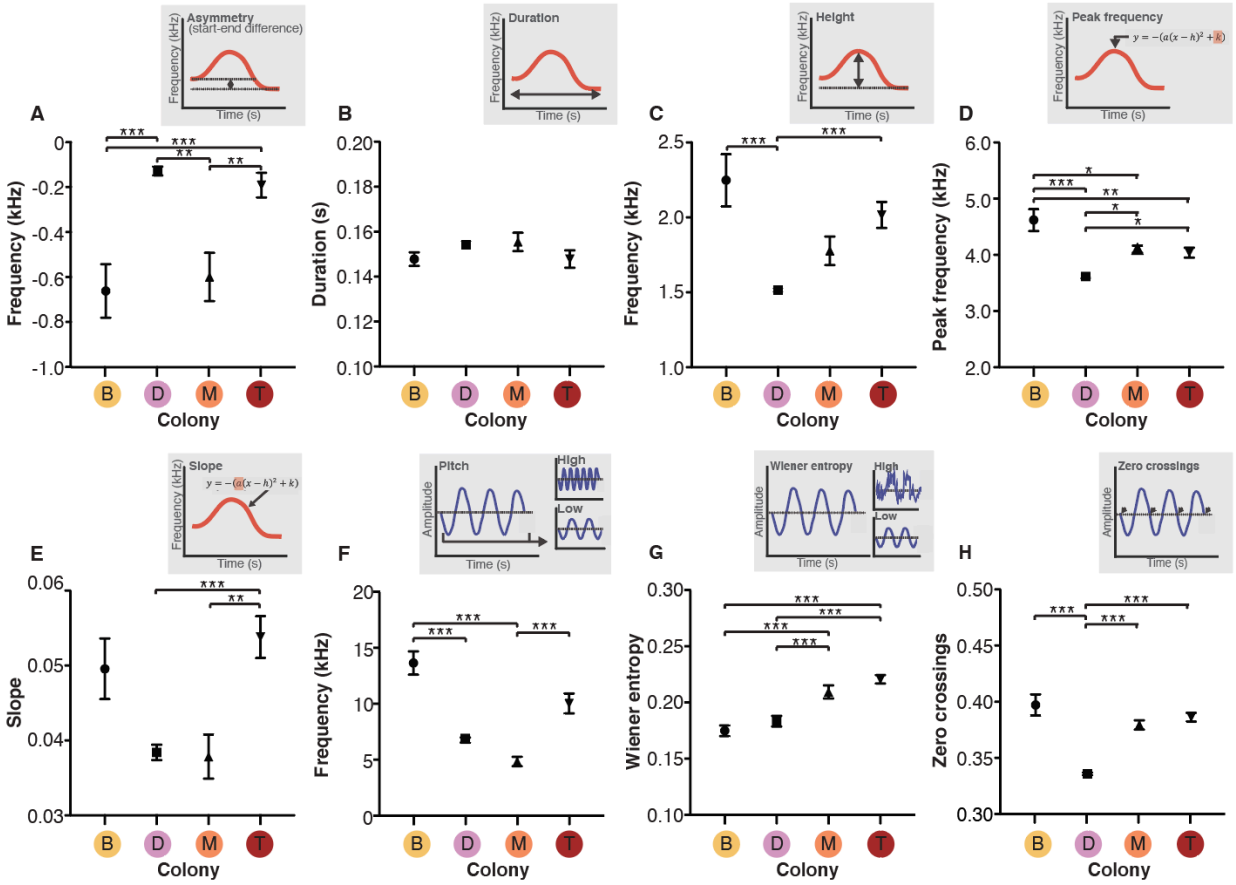

Fig. S4.

(**A-I**) Analysis of individual soft chirp features for each colony display significant differences across multiple spectrogram and soundwave extracted features , * P <0.05, ** P < 0.005, *** P,< 0.0005 . Error bars, SEM. The eight features displayed here were used to train Random Forest classifiers in Fig. 2.

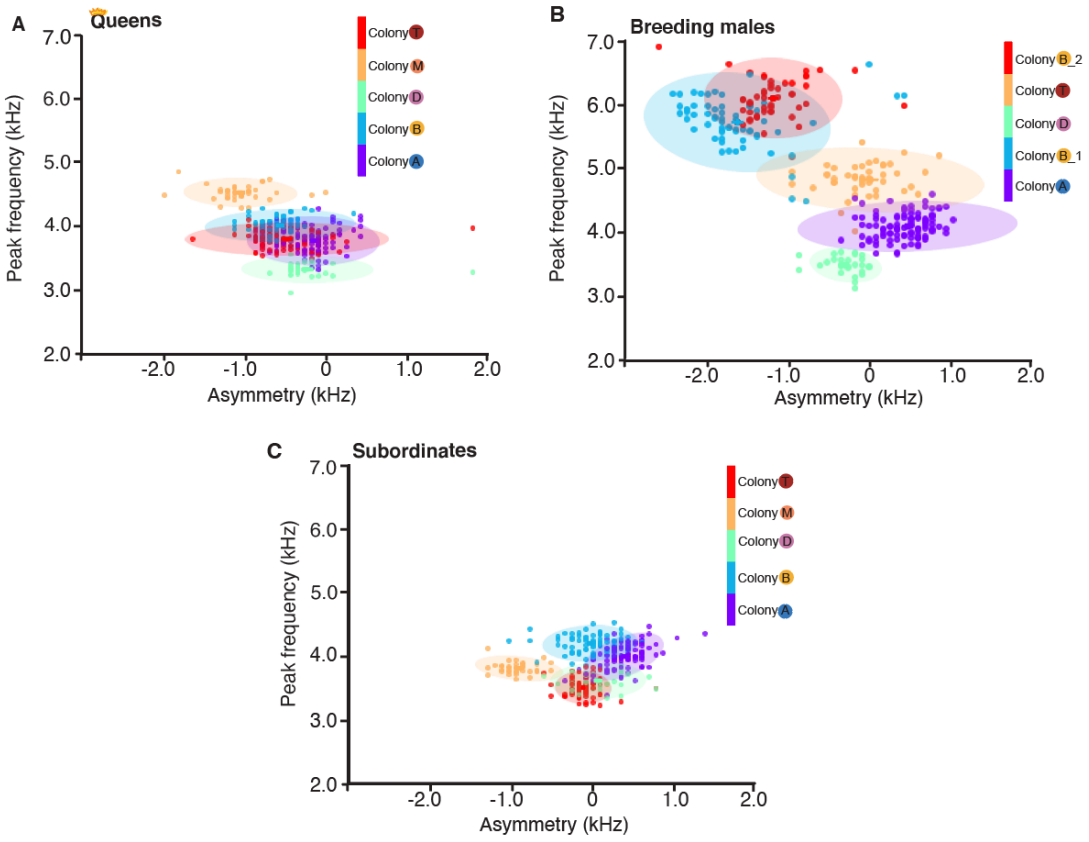

Fig. S5.

Breeding males produce the most variable soft chirps within colonies. (**A-C**) Soft chirp peak frequency vs. soft chirp asymmetry graphs provide a snapshot of variability across individuals for queens (A), breeding males (B) and subordinates (C). Colony indicated by color bar. Ellipses display 95% confidence intervals.

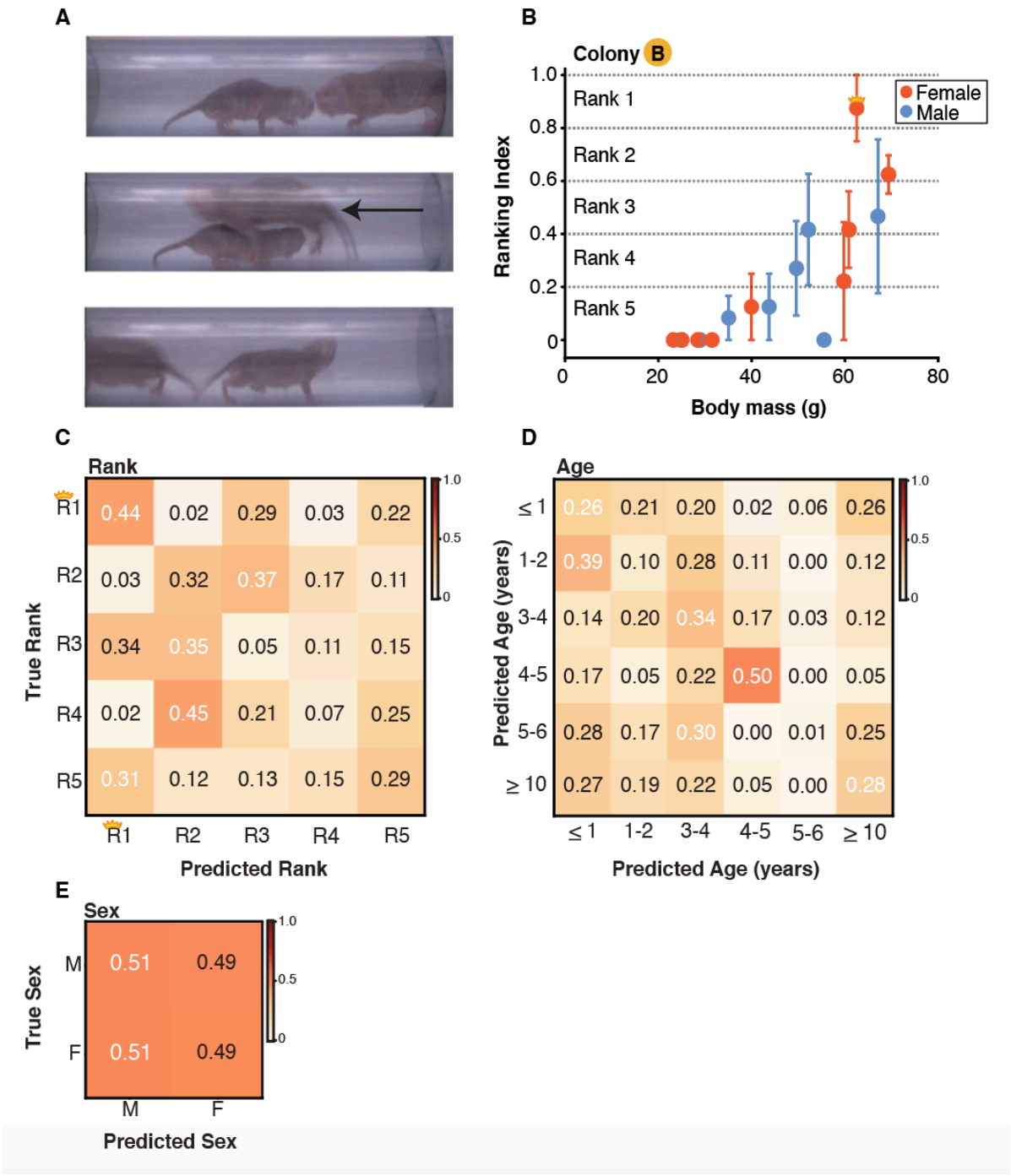

Fig. S6.

Rank, age and sex are not strongly predicted by soft chirp features. (**A-B**) Example of hierarchy test for assessing individual rank in a colony, data shown from Colony B and previously described in (*28*). More dominant individuals will climb over less dominant individuals (A, and see Movie S2) an assay which successfully predicts the queen (B). (**C-E**) Rank (C), age (D), and sex (E) are not strongly encoded in soft chirp features. Confusion matrices for Random Forest classifier training with rank, age and sex respectively.

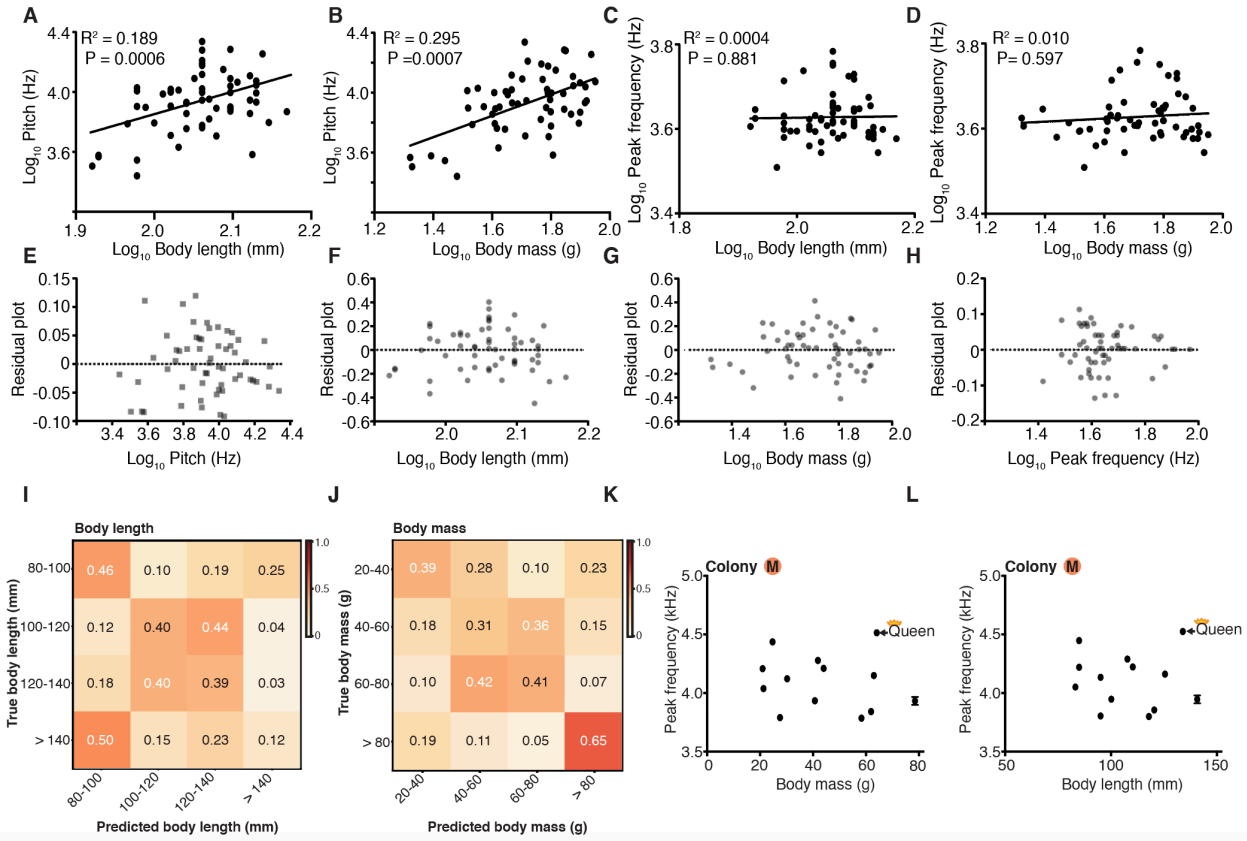

Fig. S7.

Naked mole-rat soft chirp pitch but not peak frequency shows a positive correlation with body size. (**A-B**) Pitch vs. body length (A) and body mass (B). (**C-D**) Peak soft chirp frequency vs. body length (A) and body mass (B). (**E-H**) Residuals for pitch, body length, body mass and peak frequency. (**I,J**) Confusion matrices for Random Forest classifier trained on body length (I) and body mass (J). (**K, L**) Despite having one of the longest body lengths and largest body masses relative to the rest of the colony, the queen does not produce the lowest frequency vocalizations when compared across the entire colony, as predicted by an inverse correlation between pitch or frequency and body size, data plotted for Colony M.

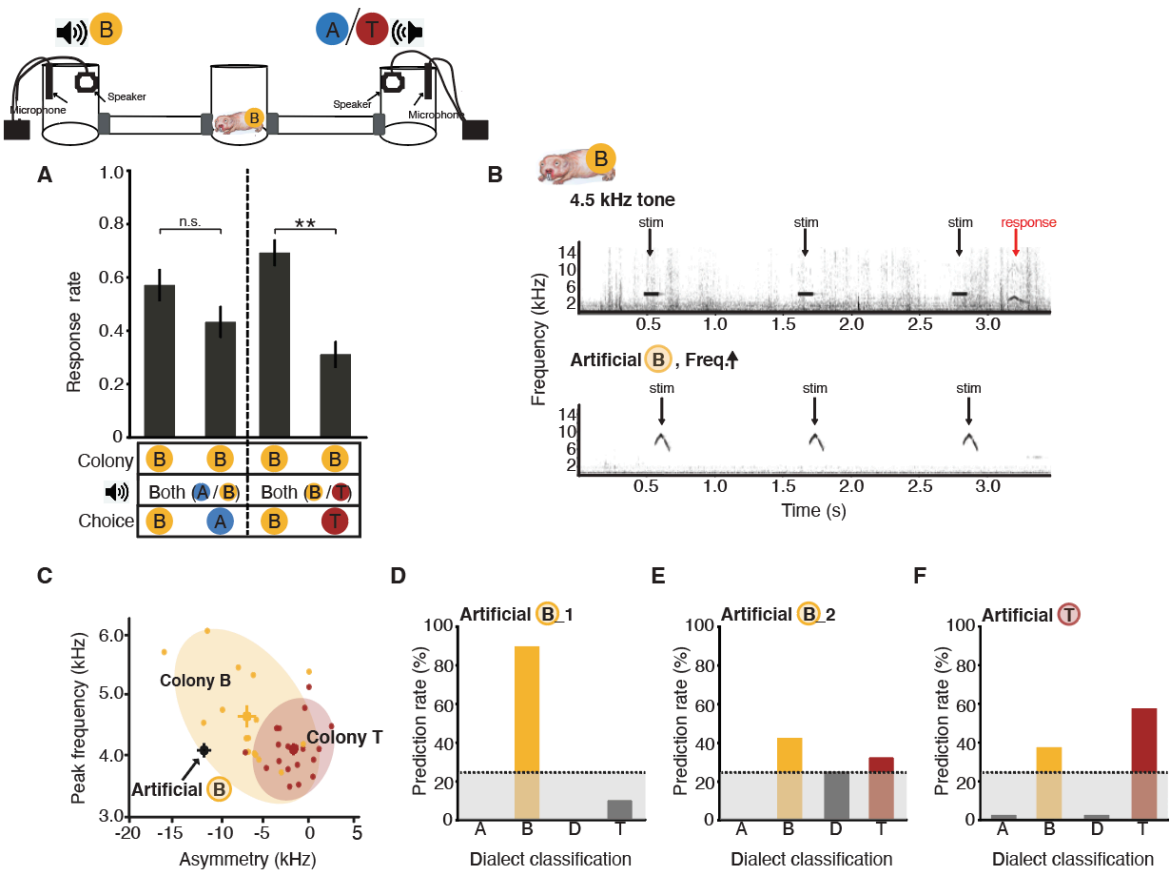

Fig. S8.

Recognition of colony dialects drives preferential behavioral responses. (**A**) Schematic of forced choice behavioral assay (top). When soft chirp audio playbacks from two colonies are presented simultaneously animals prefer to spend time in the chamber presenting soft chirps from their own colony * P <0.05, ** P < 0.005, *** P,< 0.0005 . Error bars, SEM (**B**) Representative examples of responses to additional audio stimuli: pure tone (4.5 kHz) and peak frequency-doubled stimuli from Colony B. (**C**) Artificial playback stimuli generated for Colony B overlap with Colony B features but not with any individual Colony B member (peak frequency is plotted vs. asymmetry). (**D-F**) Artificial stimuli generated for Colony B and Colony T are correctly assigned to Colony B and T respectively when tested with our colony classifier. Error bars are SEM, ellipses display 95% confidence intervals.

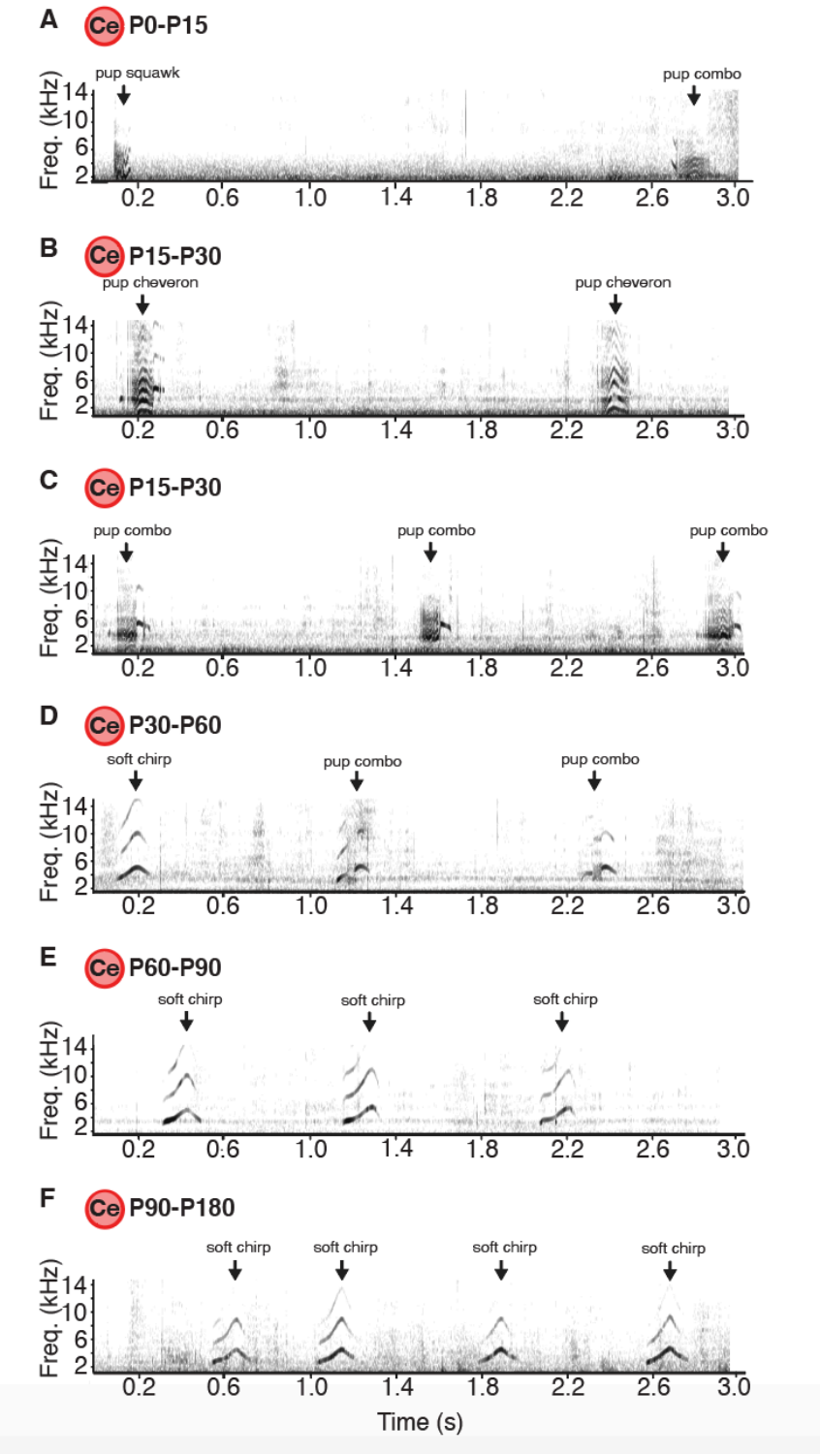

Fig. S9.

(**A-F**) Representative examples of developmental sound type usage in a single pup, Ce, from Colony L. Note transition from exclusively pup sounds to soft chirps begins around P60-P90.

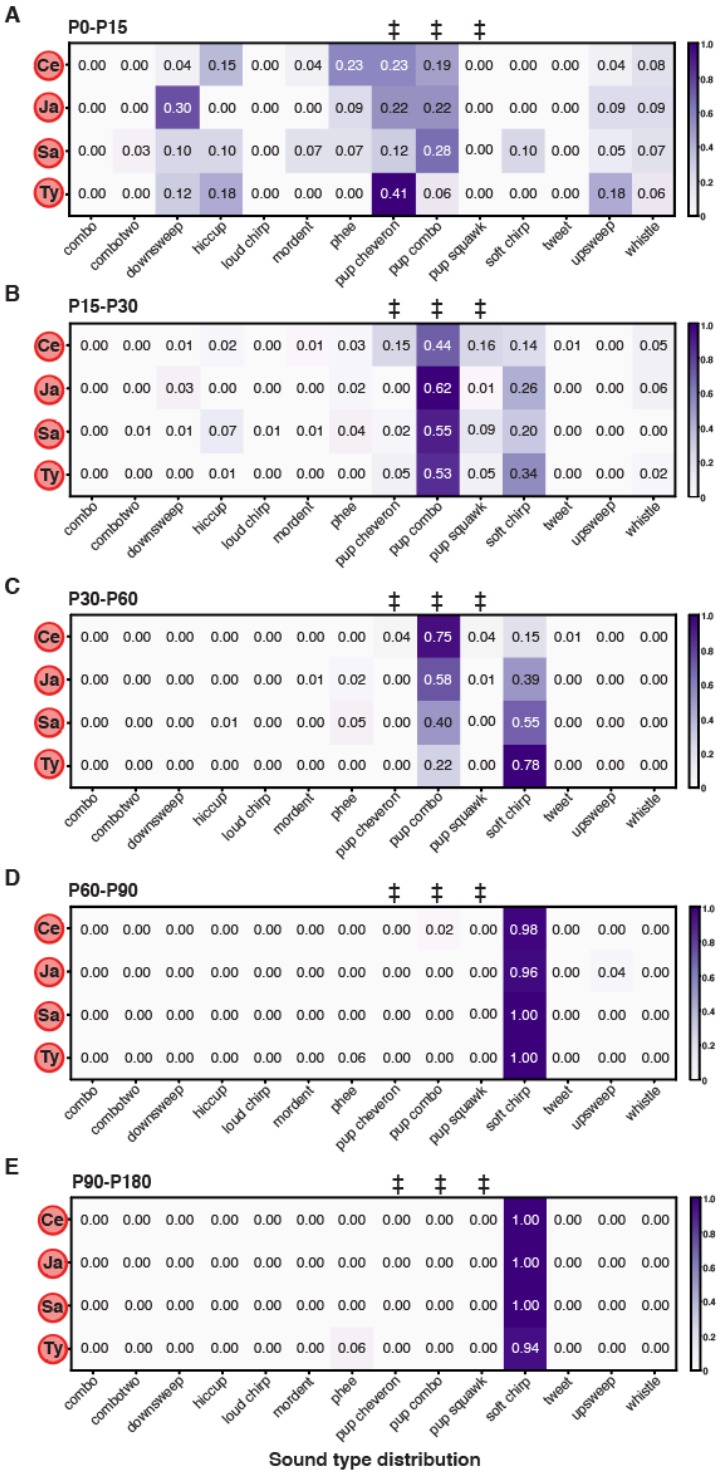

Fig.S10.

(**A-E**) Distribution of sound type usage across early development for four pups in Colony L (pups Ce, Ja, Sa, and Ty. Note transition to soft chirp primary soft chirp usage over time. ‡, denotes exclusively pup produced sound types.

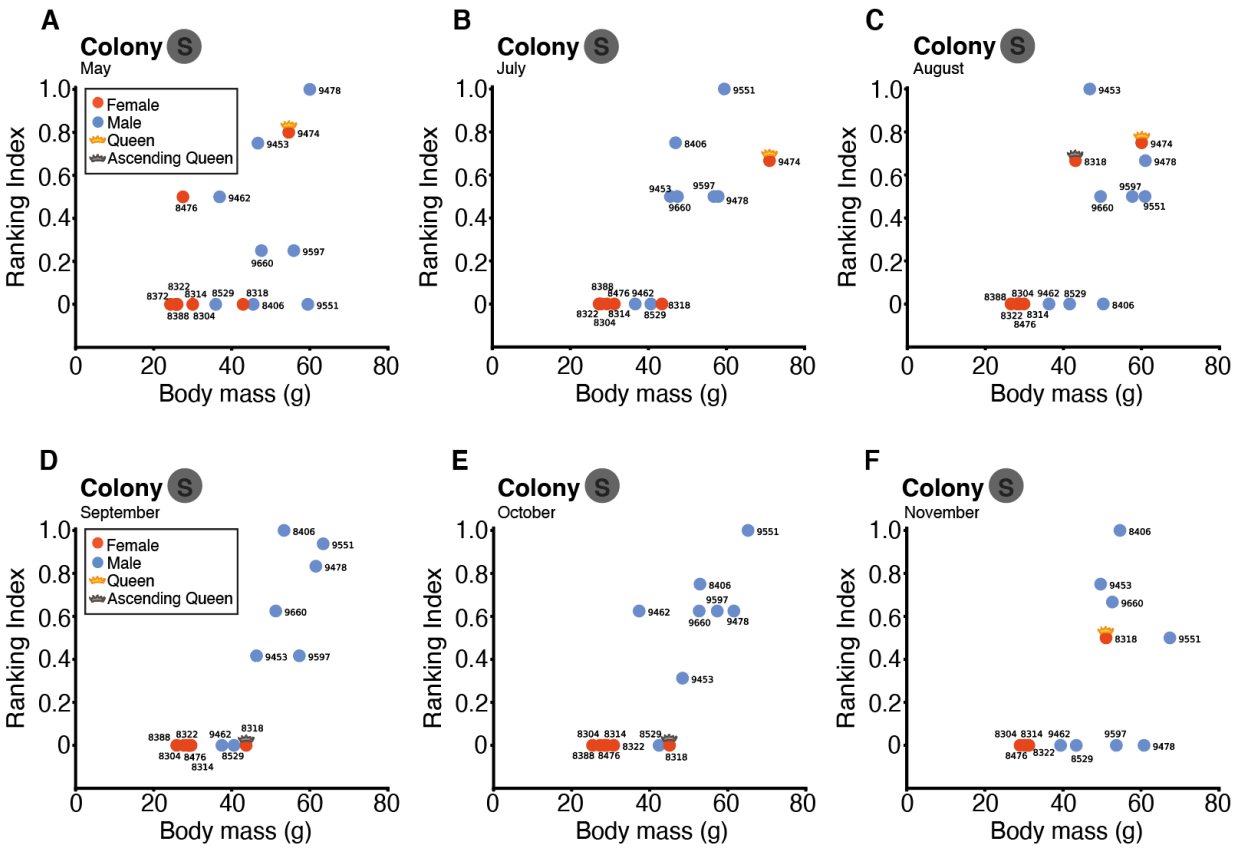

Fig. S11.

(**A-F**) Hierarchy changes in Colony S can be reliably tracked over time. Note ascendance of new queen (ID 8318) which corresponds to Queen Epoch 2 in Fig. 4 I-K.

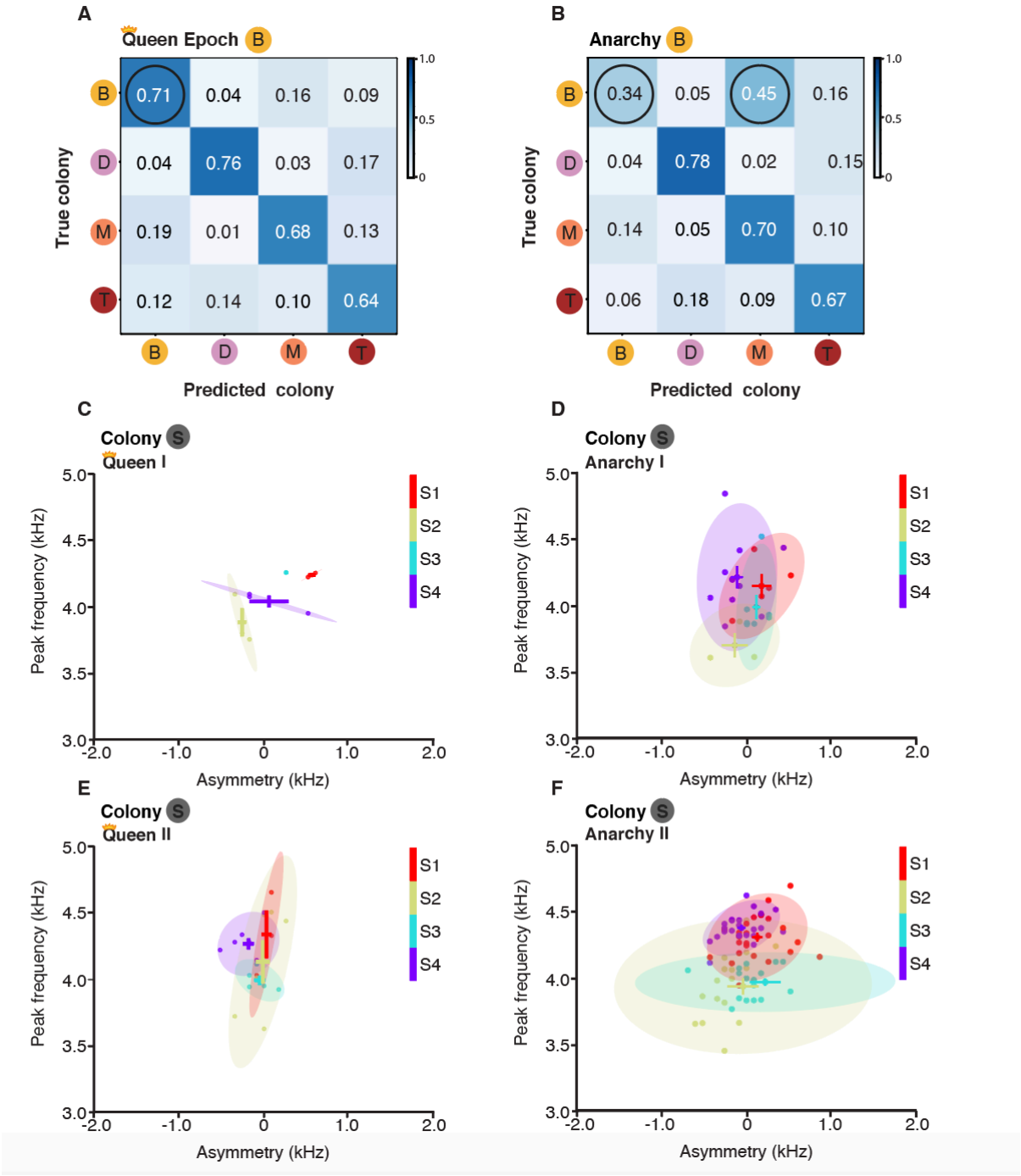

Fig. S12.

Individual variability is increased during periods of anarchy, when no queen is present. (**A**) Colony B prediction accuracy is high when the colony has a stable Queen and is reduced when the Queen is lost (**B**). Black circles indicate changes in prediction accuracy. (**C-F**) Representative examples of soft chirp variability (peak frequency versus asymmetry) are shown in four subordinates (S1-S4) from Colony S during the epoch of Queen I (C), the subsequent phase of Anarchy (I, D), the epoch of Queen II (E) and a second phase of Anarchy (II, F). Ellipses display 95% confidence intervals.

Movie S1. Example of audio playback experiments. Spectrograms with audio stimulus and vocal responses are superimposed on video. Naked mole-rats show preferential responses to home colony playbacks (here Colony B) and artificially generated home colony playbacks but not to playbacks from a foreign colony, here Colony T. Video shown in real time.

Movie S2. Example of dominance interaction assessed with hierarchy test for rank as previously described in (*28*). When two naked mole-rats encounter one another in a tunnel, the dominant individual will climb over the subordinate individual.

Audio S1. Excerpt of a vocal interaction between two naked mole-rats (queen and breeding male) provides a snapshot of the vocal complexity produced by these animals.

Audio S2. Audio excerpt from a whole colony recording demonstrates the near constant volley of soft chirps heard during normal colony interactions.