Rats emit aversive 44-kHz ultrasonic vocalizations during prolonged Pavlovian fear conditioning

Krzysztof H. Olszyński[1]*, Rafał Polowy[1]*, Agnieszka D. Wardak[1], Izabela A. Łaska[1], Aneta W. Grymanowska[1], Wojciech Puławski[2], Olga Gawryś[3], Michał Koliński[2], Robert K. Filipkowski[1,4]

[1] Behavior and Metabolism Research Laboratory, Mossakowski Medical Research Institute, Polish Academy of Sciences, Warsaw, Poland
[2] Bioinformatics Laboratory, Mossakowski Medical Research Institute, Polish Academy of Sciences, Warsaw, Poland
[3] Department of Renal and Body Fluid Physiology, Mossakowski Medical Research Institute, Polish Academy of Sciences, Warsaw, Poland
[4] corresponding author
* equal contribution

Abstract
Rats are believed to communicate their emotional state by emitting two distinct types of ultrasonic vocalizations. The first is long "22-kHz" vocalizations (>300 ms, <32 kHz) with constant frequency, signaling aversive states and the second, is short "50-kHz" calls (<100 ms, >32 kHz), often frequency-modulated, in appetitive situations. Here we are first to describe a new-type of aversive vocalizations emitted at a higher pitch by Wistar and spontaneously hypertensive rats (SHR) in an intensified aversive state – prolonged fear conditioning. These calls, which we named "44-kHz vocalizations", are long (>150 ms), generally at a constant frequency (usually within 35-50 kHz range) and have an overall spectrographic image similar to 22-kHz calls. Some 44-kHz vocalizations are comprised of both 22-kHz-like and 44-kHz-like elements. Furthermore, two separate clustering methods confirmed that these 44-kHz calls are distinct from other vocalizations. We also show that some of rats' responses to the playback of 44-kHz calls were more akin to that of aversive calls, e.g. heart rate changes, whereas other responses were at an intermediate level between aversive and appetitive calls. Our results suggest that rats have a wider vocal repertoire than previously believed, and current definitions of major call types may require reevaluation.

Charles Darwin wrote: "That the pitch of the voice bears some relation to certain states of feeling is tolerably clear"[1]. This has also been tolerably clearly observed and widely described for ultrasonic vocalizations of rats[2-4] which emit low-pitched aversive calls and high-pitched appetitive calls. The former are "22-kHz" vocalizations (Figs 1A, 2A), with 18 to 32 kHz frequency range, monotonous and long, usually >300 ms, and are uttered in distress[2-5]. The latter are "50-kHz" vocalizations (Fig. 1C), are relatively short (10-150 ms), frequency-modulated, usually within 35-80 kHz, and they signal appetitive and rewarding states[2-5]. Therefore, these two types of calls communicate the animal's emotional state to their social group[5]. Low-pitch (<32 kHz), short (<300 ms; Fig. 1B) calls, assumed to also express a negative aversive state, have been described but their role is not clearly established[5]. Notably, high-pitch (>32 kHz), long and monotonous ultrasonic vocalizations have not yet been described. Here we show these unmodulated rat vocalizations with peak frequency[a] at about 44 kHz (Figs 1B, 1E, 2B), emitted in aversive experimental situations, especially in prolonged fear conditioning.

[a] *peak frequency* of a given vocalization is defined here as the highest power peak in the averaged spectrum of the entire element.

New calls are high, long, unmodulated

In three separate experiments, two with delay-fear-conditioning, one of which has already been described[6,7], and the third with trace-fear-conditioning (see Methods), 46 of 68 conditioned Wistar rats Figs 1B, 1E, S1BC displayed a new-type vocalization that was high-pitched, i.e. in the range of 50-kHz calls, but long and monotonous (Fig. 2B). These vocalizations, e.g. top-right group in Figs 1B, S1C, were outside the defined range[2-4] for both 50-kHz (bottom-right group in Figs 1C, S1BC) and 22-kHz calls (top-left group in Figs 1A, 1B, S1).

This new-type of vocalization was also observed in a different rat strain acquired from a different breeding colony, i.e. spontaneously hypertensive rats (SHR)[8], also trained in delay fear conditioning[6]. Six of the 32 SHR tested displayed high-pitch, long, monotonous vocalizations (e.g. Fig. S2AG).

Overall, we analyzed 140,149 vocalizations from all fear conditioning experiments and through trial-and-error, we set **new criteria,** namely **peak frequency of >32 kHz and >150 ms duration** to define the new-type calls. We manually verified the results on the spectrogram using these parameters and only 308 calls (0.2%) were incorrectly assigned (i.e. exceptionally long 50-kHz vocalizations misplaced in the new-type group or borderline-short vocalizations of the new-type misplaced to 50-kHz calls). Hence the new parameters correctly assigned 99.8% of cases and are thus effective to distinguish the new-type calls in an automated fashion. Finally, 10,445 new-type calls were identified, which constituted 7.5% of the total calls during fear conditioning experiments (comp. Fig. 1G). These vocalizations have a peak frequency range from 32.2 to 51.5 kHz (95% of cases) with an average peak frequency of 42.1 kHz, and they exhibited 43.8 kHz peak frequency at the cluster center in a DBSCAN analysis (Fig. 3A). In line with the accepted nomenclature convention, underlining the relationship with 22-kHz vocalizations, we christened this new-type of ultrasonic calls as "**44-kHz vocalizations**".

New calls in long aversive stimulation

We found 44-kHz vocalizations especially in rats which received multiple electric shocks. These vocalizations were less frequent following the first trial (1.7 ± 0.9% of all long aversive calls), and increased in subsequent trials, particularly during the 5th (12.6 ± 6.5%) to 10th (24.0 ± 8.4%) trials, where 44-kHz calls gradually replaced 22-kHz vocalizations in some rats (Fig. 1F, Supplementary Video 1; comp Fig. 1D vs. 1E). Please note, in the Wistar rats that undergone 10 trials of delay fear conditioning, majority of the 22-kHz calls were emitted after the 3rd shock, i.e. during the 3rd ITI (inter-trial-interval), while 44-kHz vocalizations were emitted in the second part of the training, i.e. 5th to 10th ITI (Fig. 1F). We also observed the frequencies of 22-kHz calls to gradually rise throughout fear conditioning training, i.e. during subsequent ITI (Fig. S1DE). The frequency levels of 44-kHz vocalizations also appeared to rise (Fig. S1DE) but we were unable to statistically demonstrate it. There were more 44-kHz vocalizations during fear conditioning training than testing (4.42 ± 1.27 vs. 0.30 ± 0.17 calls/min; all three experiments; $p < 0.0001$, Wilcoxon).

44-kHz sorted into five subtypes

While the majority of 44-kHz vocalizations were of continuous unmodulated frequency (Fig. 2B), some comprised additional elements. Based on the composition of individual call elements and their relation to each other, we manually sorted the calls into five categories (Fig. 2B-F). If the start (*prefix*) or end (*suffix*) portion of a call was less than 1/5th the length of the following or previous element, this portion of the call was not considered in its categorization into the five subtypes. The names and descriptions of the five subtypes are: **flat** – single element with near constant frequency with little to no interruptions to the sound continuity on the spectrogram; **step up** – two elements with an instantaneous frequency jump, where the first element is of lower frequency; **step down** – two elements with an instantaneous frequency jump, where the first element is of higher frequency; **insert** – three elements with an instantaneous frequency change, where the middle element is of different frequency; **complex** – more than three elements with instantaneous frequency changes.

44-kHz and 22-kHz calls closely related

44-kHz were emitted in aversive behavioral situations – as 22-kHz calls are observed[9-11]. Both types of calls are long (usually >300 ms) and frequency-unmodulated. Some of the elements constituting such as step up; step down; insert and complex 44-kHz vocalizations (Fig. 2C-F) were at a lower frequency – typical for 22-kHz vocalizations. *Vice versa* we also observed 22-kHz calls with 44-kHz-like elements. Therefore, we propose that these long 22-kHz and 44-kHz vocalizations constitute a *supertype* group of long unmodulated aversive calls ("22/44-kHz vocalizations").

We observed a stable, approximately 1.5 ratio in peak frequency levels between 22-kHz and 44-kHz vocalizations within individual rats. Specifically, in fourteen rats (13 Wistar and 1 SHR) at their transition from 22-kHz to 44-kHz calls during the fear conditioning session, the proportion between the frequencies of the long 22-kHz vocalizations and the long 44-kHz calls was 1.48 ± 0.02. Similar results were obtained for 70 step up (1.53 ± 0.03) and 65 step down (1.59 ± 0.02) 44-kHz calls – altogether suggesting a 1.5-times or 3:2 frequency ratio. This ratio and its relevance has been observed in invertebrates and vertebrates including human speech and music[12]. In music theory, 3:2 frequency ratio is referred to as a perfect fifth and is often featured, e.g. the first two notes of the *Star Wars* 1977 movie (ascending, i.e. step up; comp. Fig. 2C, Supplementary Audio 1) and *Game of Thrones* 2011 television series (descending, i.e. step down; comp. Fig. 2D, Supplementary Audio 2) theme songs. All of which may point to a common evolutionary basis for this sound interval and its prevalence (also discussed in[13]).

New calls form separate, distinct group

Next, we showed that 44-kHz calls indeed constitute a distinct, separate type of ultrasonic vocalizations as it was sorted into isolated clusters by two different methods. First, using the DBSCAN algorithm method based on calls' peak frequency and duration, we were able to divide all vocalizations recorded during all training sessions into 44-kHz vocalizations vs. all other vocalizations as two separate clusters (Fig. 3A). Secondly, a clustering algorithm that includes call contours, i.e. k-means with UMAP projection done via DeepSqueak, sorted 44-kHz vocalizations of different subtypes including unusual ones (Fig. S2A-F), into topologically-separate groups. Notably, flat 44-kHz calls were consistently in a separate cluster from 22-kHz calls Figs 3C, S3B).**Specific response to 44-kHz playback**

To describe the behavioral and physiological impact of 44-kHz vocalizations, we performed playback experiments in two separate groups of naive rats (Methods, Figs 4, S4). Overall, the responses to 44-kHz aversive calls presented from the speaker were either similar to 22-kHz vocalizations or in-between responses to 22-kHz and 50-kHz playbacks. For example, the heart rate of rats exposed to 22-kHz and 44-kHz vocalizations decreased, and increased to 50-kHz calls (Figs 4A, comp.[14]). Whereas the number of vocalizations emitted by rats was highest during and after the playback of 50-kHz, intermediate to 44-kHz and lowest to 22-kHz playbacks (Figs 4BC, S4EF). Additionally, the duration of 50-kHz vocalizations emitted in response to 44-kHz playback was also intermediate, i.e. longer than following 22-kHz playback (Fig. 4D) and shorter than following 50-kHz playback (Figs 4D, 4SG). Finally, similar tendencies were observed in the distance travelled and time spent in the half of the cage adjacent to the speaker (Fig. 4SA-D).

Discussion

As Charles Darwin noted above[1] and other researchers have confirmed[15], the frequency level of animal calls is a vocal parameter that changes in accordance with its arousal state (intensity) or emotional valence (positive/negative state). The frequency shifts towards both higher and lower levels, i.e. alterations were observed during both positive (appetitive) and negative (agonistic/aversive) situations, however, as a general rule, frequency increases with an increase in arousal. It could be argued that our prolonged fear conditioning increased the arousal of the rats with no change in the valence of the aversive stimuli.

The question is, why have the 44-kHz vocalizations been overlooked until now? On one hand, long (or not that long[16]), frequency-stable high-pitch vocalizations have been reported before (e.g.[17,18]), notably as caused by intense cholinergic stimulation[19] or higher shock-dose fear

conditioning[20]. However, they have not been systematically defined, described or demonstrated to be a separate type of vocalization. On the other hand, 44-kHz calls were likely omitted as the analyses were restricted to canonical groups (flat 22-kHz and short 50-kHz calls), moreover – many older bat-detectors had limited frequency-range detection when stress-evoked types of ultrasonic calls were being established. Finally, 44-kHz vocalizations are emitted much fewer than 22-kHz calls (Fig. 1F). Here we established that these 44-kHz vocalizations are a separate and behaviorally-relevant group of rat ultrasonic calls. Our results bring to awareness that rats employ these previously unrecognized, long, high-pitched and flat aversive calls in their vocal repertoire. Researchers investigating rat ultrasonic vocalizations should be aware of their potential presence and to not rely fully on automated detection of high vs. low-pitch calls.

Literature

1. Darwin, C. (1872). "Chapter 4: Means of Expression in Animals", The Expression of the Emotions in Man and Animals (D. Appleton & Company).
2. Brudzynski, S.M. (2019). Emission of 22 kHz vocalizations in rats as an evolutionary equivalent of human crying: Relationship to depression. Behav Brain Res *363*, 1-12. 10.1016/j.bbr.2019.01.033.
3. Brudzynski, S.M. (2021). Biological Functions of Rat Ultrasonic Vocalizations, Arousal Mechanisms, and Call Initiation. Brain Sci *11*. 10.3390/brainsci11050605.
4. Simola, N., and Granon, S. (2019). Ultrasonic vocalizations as a tool in studying emotional states in rodent models of social behavior and brain disease. Neuropharmacology *159*, 107420. 10.1016/j.neuropharm.2018.11.008.
5. Brudzynski, S.M. (2013). Ethotransmission: communication of emotional states through ultrasonic vocalization in rats. Curr Opin Neurobiol *23*, 310-317. 10.1016/j.conb.2013.01.014.
6. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., Zielinski, J., and Filipkowski, R.K. (2022). Spontaneously hypertensive rats manifest deficits in emotional response to 22-kHz and 50-kHz ultrasonic playback. Prog Neuropsychopharmacol Biol Psychiatry *120*, 110615. 10.1016/j.pnpbp.2022.110615.
7. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., and Filipkowski, R.K. (2021). Increased Vocalization of Rats in Response to Ultrasonic Playback as a Sign of Hypervigilance Following Fear Conditioning. Brain Sci *11*. 10.3390/brainsci11080970.
8. Okamoto, K., and Aoki, K. (1963). Development of a strain of spontaneously hypertensive rats. Jpn Circ J *27*, 282-293. 10.1253/jcj.27.282.
9. Antoniadis, E.A., and McDonald, R.J. (1999). Discriminative fear conditioning to context expressed by multiple measures of fear in the rat. Behav Brain Res *101*, 1-13. 10.1016/s0166-4328(98)00056-4.
10. Dupin, M., Garcia, S., Boulanger-Bertolus, J., Buonviso, N., and Mouly, A.M. (2019). New Insights from 22-kHz Ultrasonic Vocalizations to Characterize Fear Responses: Relationship with Respiration and Brain Oscillatory Dynamics. eNeuro *6*. 10.1523/ENEURO.0065-19.2019.
11. Taylor, J.O., Urbano, C.M., and Cooper, B.G. (2017). Differential patterns of constant frequency 50 and 22 khz usv production are related to intensity of negative affective state. Behav Neurosci *131*, 115-126. 10.1037/bne0000184.
12. Hoeschele, M. (2017). Animal Pitch Perception: Melodies and Harmonies. Comp Cogn Behav Rev *12*, 5-18. 10.3819/CCBR.2017.120002.
13. Bowling, D.L., and Purves, D. (2015). A biological rationale for musical consonance. Proc Natl Acad Sci U S A *112*, 11155-11160. 10.1073/pnas.1505768112.
14. Olszynski, K.H., Polowy, R., Malz, M., Boguszewski, P.M., and Filipkowski, R.K. (2020). Playback of Alarm and Appetitive Calls Differentially Impacts Vocal, Heart-Rate, and Motor Response in Rats. iScience *23*, 101577. 10.1016/j.isci.2020.101577.
15. Briefer, E.F., Padilla de la Torre, M., and McElligott, A.G. (2012). Mother goats do not forget their kids' calls. Proc Biol Sci *279*, 3749-3755. 10.1098/rspb.2012.0986.

16. Bialy, M., Podobinska, M., Barski, J., Bogacki-Rychlik, W., and Sajdel-Sulkowska, E.M. (2019). Distinct classes of low frequency ultrasonic vocalizations in rats during sexual interactions relate to different emotional states. Acta Neurobiol Exp (Wars) *79*, 1-12.
17. Sales, G.D. (1979). Strain Differences in the Ultrasonic Behavior of Rats (Rattus norvegicus) Am Zool *19*, 513-527.
18. Shimoju, R., Shibata, H., Hori, M., and Kurosawa, M. (2020). Stroking stimulation of the skin elicits 50-kHz ultrasonic vocalizations in young adult rats. J Physiol Sci *70*. 10.1186/s12576-020-00770-1.
19. Brudzynski, S.M., and Bihari, F. (1990). Ultrasonic vocalization in rats produced by cholinergic stimulation of the brain. Neurosci Lett *109*, 222-226. 10.1016/0304-3940(90)90567-s.
20. Wohr, M., Borta, A., and Schwarting, R.K. (2005). Overt behavior and ultrasonic vocalization in a fear conditioning paradigm: a dose-response study in the rat. Neurobiol Learn Mem *84*, 228-240. 10.1016/j.nlm.2005.07.004.

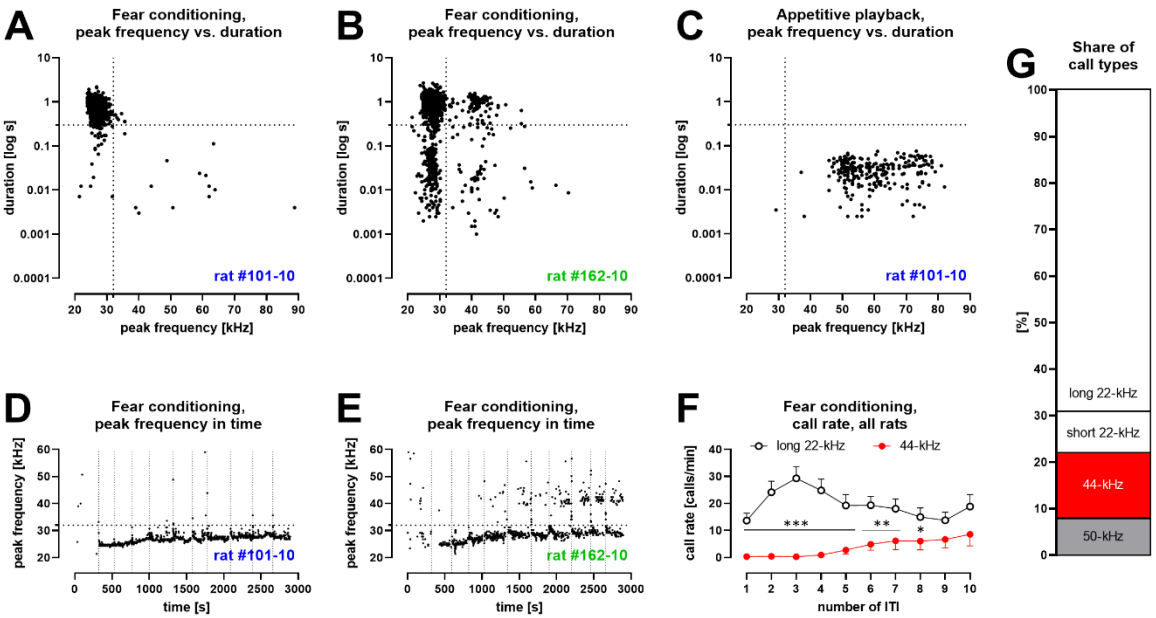

Fig. 1. Characteristics of vocalizations emitted by Wistar rats during delay fear conditioning with ten aversive foot-shocks. A – some rats produced aversive 22-kHz vocalizations with typical features i.e. constant-frequency of <32 kHz, >300 ms duration – both values marked as dotted lines); example emission from one rat. **B** – some rats produced 44-kHz vocalizations with constant frequency of >32 kHz and long duration (>150 ms); example emission from one rat. **C** – rats which emitted aversive vocalizations during fear session, produced 50-kHz vocalizations during appetitive playback session the following day (full data published in[1]); representative data from same rat in **A**. **D** – the onset of long 22-kHz alarm calls typically occurred after first shock stimulus (vertical dotted lines mark time of shock deliveries in **DE**); note the gradual rise in peak frequency[a], not exceeding 32 kHz (horizontal dotted line in **DE**); data from same rat as **AC**. **E** – in rats that emitted 44-kHz calls, the onset was usually delayed to after several foot-shocks; note the gradual rise in peak frequency of both long 22-kHz and 44-kHz vocalizations throughout training (comp. Fig. S1DE); data from same rat in **B**). **F** – call rate of long 22-kHz calls was higher than 44-kHz calls (*p < 0.05, **p < 0.01, ***p < 0.001; Mann-Whitney) and with different time-course – maximum number of 22-kHz calls at ITI-3 (higher than ITI-1, 2, 5-10; <0.0001–0.0265 p levels); and higher number of 44-kHz calls at ITI-5-10, i.e. 5.8 ± 2.9 vs. ITI-1-4, i.e. 0.5 ± 0.2; p < 0.0001; all Wilcoxon); numbers of ITI (inter-trial-intervals) correspond to the numbers of previous foot-shocks, values are means ± SEM. **G** – long 44-kHz vocalizations had a higher incidence rate (14.1%) than short 22-kHz (8.9%) and 50-kHz calls (7.9%); values are calculated for sum of all vocalizations obtained during entire training sessions (there were fewer 50-kHz calls, i.e. 4.9%, when vocalizations prior to the first shock were not included). **A-E**: dots reflect specified single rat values. **FG**: n = 29; other results from these rats are previously published[1,2].

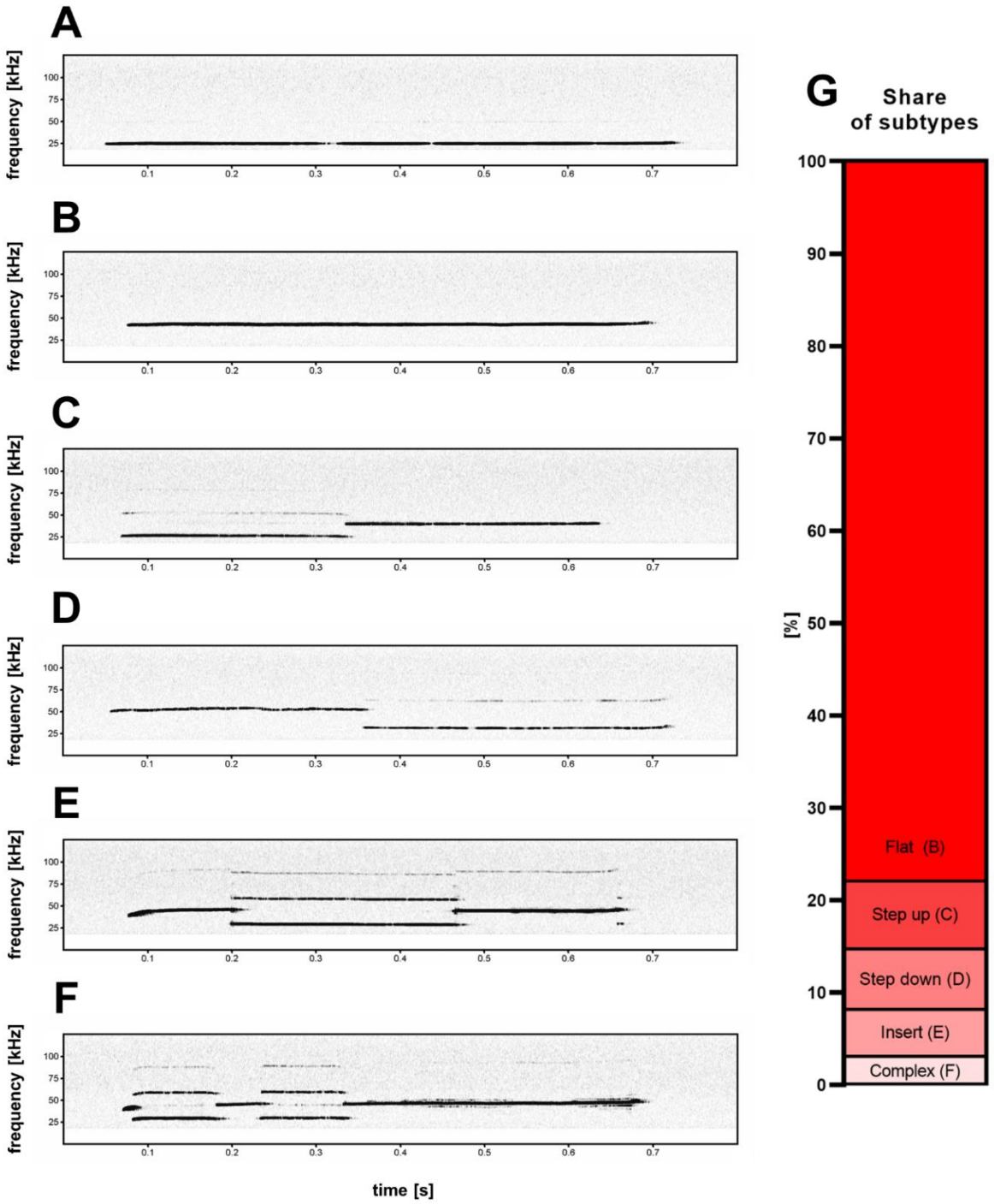

Fig. 2. Five subtypes (B-F) of high frequency 44-kHz aversive vocalizations. A – standard aversive 22-kHz vocalization with peak frequency <32 kHz (peak frequency = 24.4 kHz). 44-kHz aversive vocalization subtypes: **B** – flat (constant frequency call; peak frequency = 42.4 kHz), **C** – step up (peak frequency = 39.5 kHz), **D** – step down (peak frequency = 52.2 kHz), **E** – insert (peak frequency = 38.5 kHz), **F** – complex (peak frequency = 46.3 kHz). **G** – percentage share of 44-kHz call-subtypes in all cases of detected 44-kHz vocalizations.

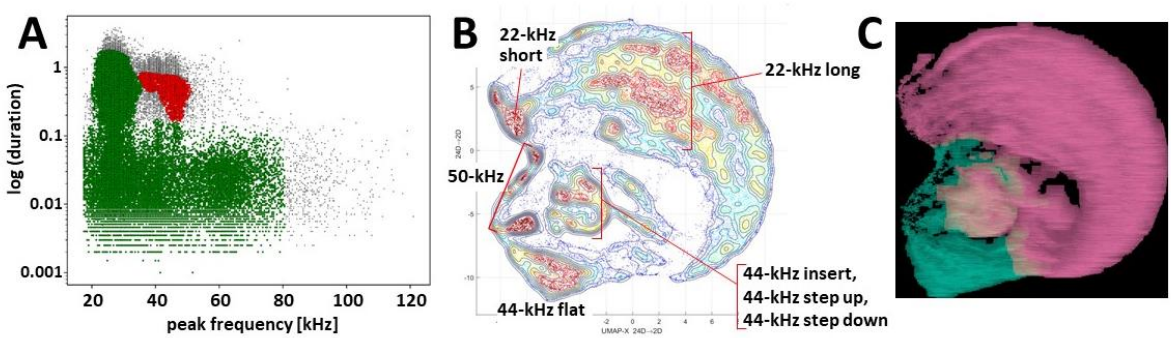

Fig. 3. Clustering of ultrasonic vocalizations from fear conditioning sessions using two independent methods. A – DBSCAN algorithm (ε = 0.14) clustering of vocalizations from all fear conditioning experiments (Wistar rats, n = 138; SHR, n = 80), silhouette coefficient = 0.198, two clusters emerge, cluster of green dots n = 77,243 (due to high generality of cluster average peak frequency and duration deemed redundant), cluster of red dots n = 5,646 (average peak frequency = 43,826.6 Hz, average duration = 0.524 s), some calls were not assigned to any cluster, i.e. outlier vocalizations, black dots, n = 4,139. **BC** – clustering by k-means algorithm and visualization of calls emitted by rats (n = 26) during trace and delay fear conditioning training, total number of calls n = 44,859; see also Fig. S3. **B** – topological plot of ultrasonic calls using UMAP embedding, particular agglomerations of calls labeled with their type or subtype. **C** – spectrogram images from DeepSqueak software superposed over plot B, colors denote clusters from unsupervised clustering, number of clusters set using elbow optimization (max number = 4), two clusters emerge.

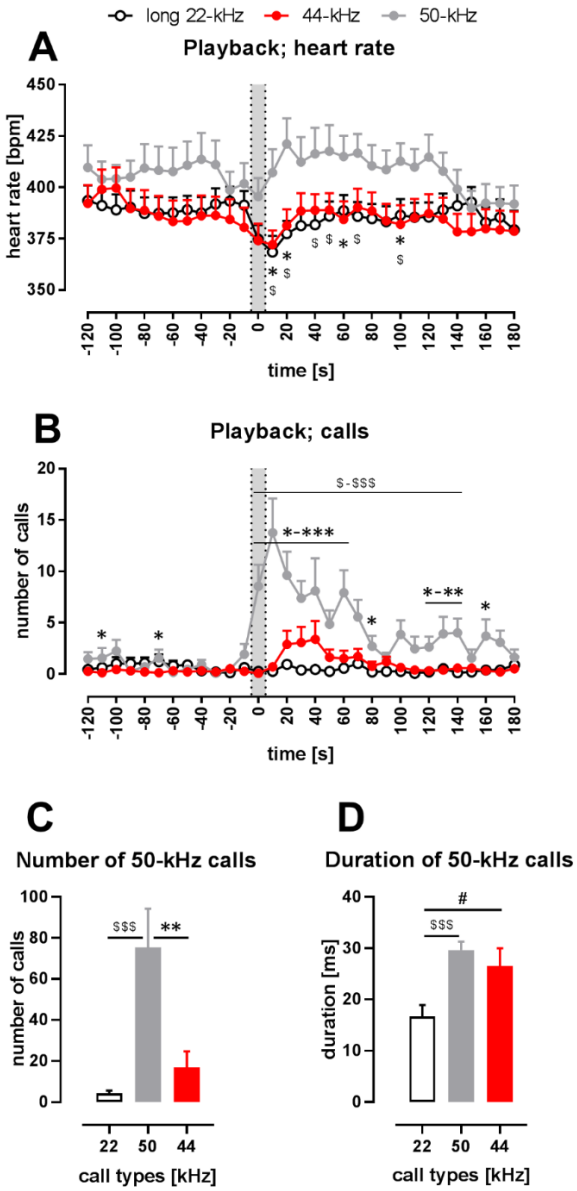

Fig. 4. Physiological and behavioral response to playback of 44-kHz calls (vs. 50-kHz and 22-kHz calls) presented from a speaker to naïve Wistar rats. A – heart rate (HR); **B** – the number of emitted vocalizations. **AB** – gray sections correspond to the 10-s-long ultrasonic playback. Each point is a mean for a 10-s-long time-interval with SEM. **CD** – properties of 50-kHz vocalizations emitted in response to ultrasonic playback, i.e. number of calls (**C**) and duration (**D**) calculated from the 0-120 s range. **A** – 50-kHz playback resulted in HR increase (playback time-interval vs. 10-30 s time-interval, $p = 0.0007$), while the presentation of the aversive playbacks resulted in HR decrease, both in case of 22-kHz ($p < 0.0001$) and 44-kHz ($p = 0.0014$, average from -30 to -10 time-intervals (i.e. "*before*") vs. playback interval, all Wilcoxon), which resulted in different HR values following different playbacks, especially at +10 s ($p = 0.0097$ for 50 kHz vs. 22-kHz playback; $p = 0.0275$ for 50 kHz vs. 44-kHz playback) and +20 s time-intervals ($p = 0.0068$, $p = 0.0097$, respectively, all Mann-Whitney). **B** – 50-kHz playback resulted also in a rise of evoked vocalizations (*before* vs. 10-30 s time-interval, $p = 0.0002$, Wilcoxon) as was the case with 44-kHz playback ($p = 0.0176$ in respective comparison), while no rise was observed following 22-kHz playback ($p = 0.1777$). However, since the increase in vocalization was robust in case of 50-kHz playback, the number of emitted vocalizations was higher than after 22-kHz playback (e.g. $p < 0.0001$

during 0-30 time-intervals) as well as after 44-kHz playback (e.g. $p < 0.0001$ during 0-10 time-intervals, both Mann-Whitney). Finally, when the increases in the number of emitted ultrasonic calls in comparison with *before* intervals were analyzed, there was a difference following 44-kHz vs. 22-kHz playbacks during 30 s and 40 s time intervals ($p = 0.0420$ and 0.0430, respectively, Wilcoxon). **C** – During the 2 min following the onset of the playbacks, rats emitted more ultrasonic calls during and after 50-kHz playback in comparison with 22-kHz ($p < 0.0001$) and 44-kHz ($p = 0.0011$) playbacks. The difference between the effects of 22-kHz and 44-kHz playbacks was not significant ($p = 0.2725$, comp. Fig. S4F; all Mann-Whitney). **D** – Ultrasonic 50-kHz calls emitted in response to playback differed in their duration, i.e. they were longer to 50-kHz ($p = 0.0004$) and 44-kHz ($p = 0.0273$, both Mann-Whitney) playbacks than to 22-kHz playback. * 50-kHz vs. 44-kHz, $ 50-kHz vs. 22-kHz, # 22-kHz vs. 44-kHz; one character (*, $ or #), $p < 0.05$; two, $p < 0.01$; three, $p < 0.001$; Mann-Whitney (**AB**) or Wilcoxon (**CD**). Values are means ± SEM, n = 13-16.

Literature for Figures

1. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., and Filipkowski, R.K. (2021). Increased Vocalization of Rats in Response to Ultrasonic Playback as a Sign of Hypervigilance Following Fear Conditioning. Brain Sci *11*. 10.3390/brainsci11080970.
2. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., Zielinski, J., and Filipkowski, R.K. (2022). Spontaneously hypertensive rats manifest deficits in emotional response to 22-kHz and 50-kHz ultrasonic playback. Prog Neuropsychopharmacol Biol Psychiatry *120*, 110615. 10.1016/j.pnpbp.2022.110615.
3. Brudzynski, S.M., Bihari, F., Ociepa, D., and Fu, X.W. (1993). Analysis of 22 kHz ultrasonic vocalization in laboratory rats: long and short calls. Physiol Behav *54*, 215-221. 10.1016/0031-9384(93)90102-l.

Methods

Animals

Wistar rats (n = 167) were obtained from The Center for Experimental Medicine of the Medical University of Bialystok, Poland; spontaneously hypertensive rats (SHR, n = 80) and Sprague-Dawley rats (n = 16) were from Mossakowski Medical Research Institute, Polish Academy of Sciences, Poland. All rats were males, 7 weeks of age on arrival, randomly assigned into groups and cage pairs where appropriate; housed with a 12 h light-dark cycle, ambient temperature (22–25 °C) with standard chow and water provided *ad libitum.* The animals were left undisturbed for at least one week before any procedures, then handled at least four times for 2 min by each experimenter directly involved for one to two weeks. All procedures were approved by Local Ethical Committees for Animal Experimentation in Warsaw.

Animal details: groups of animals used

Trace fear conditioning experiment. Wistar rats, n = 34, both single-housed (n = 14) and pair-housed (n = 20), were implanted with radiotelemetric transmitters for measuring heart rate in an ultrasonic vocalization playback experiment previously described by us[1] after which, at 13 weeks of age, half of them (n = 17) were fear-conditioned (10 shocks), while the other half (n = 17) served as controls.

Delay fear conditioning experiment, rats with transmitters. Wistar rats (n = 94) and SHR (n = 80) were implanted with a radiotelemetric transmitter one week before fear conditioning during which they received 0, 1, 6 or 10 shocks at 12 weeks of age. All the details are described in[2,3].

Delay fear conditioning experiment, rats without transmitters. Wistar rats (n = 10) were housed in pairs; were not implanted with radiotelemetric transmitters to eliminate the potential effect of surgical intervention on vocalization; they received 10 conditioning stimuli at 12 weeks of age as described in[2,3].

Playback experiment, rats with transmitters. Wistar rats (n = 29) were housed in pairs; all were implanted with a radiotelemetric transmitter one week before the playback experiment. At 12 weeks of age, one group (n = 13) heard 50-kHz appetitive vocalization playback while the other (n = 16) 22-kHz and 44-kHz aversive calls.

Playback experiment, rats without transmitters. Sprague Dawley rats (n = 16) were housed in pairs, were not implanted with the transmitters, and received appetitive and aversive ultrasonic vocalization playback at 8 weeks of age.

Surgery, transmitter implantation, heart-rate registration

Radiotelemetric transmitters (HD-S10, Data Sciences International, St. Paul, MN, USA) were implanted into the abdominal aorta of rats in specified groups as previously described[1,3]. An illustrative image with the surgery details can be found elsewhere (Figure 5 in[4]; please note, tissue glue was used instead of cellulose patches and silk sutures). The signal was collected by receivers (RSC-1, Data Sciences International, St. Paul, MN, USA) as previously described[1-3]. Readings were processed using Dataquest ART (version 4.36 Data Sciences International) for trace fear conditioning and Ponemah (version 6.32, Data Sciences International) software for other experiments.

Fear conditioning

All conditioning procedures were conducted in a chamber (VFC-008-LP, Med Associates, Fairfax, VT, USA) located in an outer cubicle (MED-VFC2-USB-R, Med Associates) equipped with an ultrasound CM16/CMPA condenser microphone (Avisoft Bioacoustics, Berlin, Germany). Ultrasonic vocalizations were recorded via Avisoft USGH Recorder (Avisoft Bioacoustics), and rat behavior was recorded via NIR monochrome camera (VID-CAM-MONO-6, Med Associates). All procedures were described in detail before[2,3].

Trace fear conditioning was performed similarly to some previous reports (e.g.[5]). Rats were individually placed in the fear conditioning apparatus in one of two different contexts: A (safe) or B

(unsafe). Context A was in an illuminated room with the cage interior with white light, the cage floor was made of solid plastic, and the cage was scented with lemon odor, cleaned with a 10% ethanol solution; the experimenter was male wearing white gloves. Context B was a different, dark room, with the cage interior with green light, the floor was made of metal bars, and the cage was scented with mint odor, cleaned with 1% acetic acid; the experimenter was female with violet gloves. The procedure: on day -2, each rat was habituated to context A for 20 min; on day -1, habituated to context B for 20 min; on day 0, each rat was placed for 52 min in context A; on day 1, after 10 min in context B, the rat received 10 conditioning stimuli (15-s-long sine wave tone, 5 kHz, 85 dB) followed by a 30 s trace period and a foot-shock (1 s, 1 mA) and 210 s inter-trial interval (ITI) (total session duration: 52 min). The animals were tested with the same protocol without shocks in context A (day 2) and context B (day 3). During the test session, control animals showed a lower level of freezing than conditioned animals (1.32 ± 0.80% vs. 19.72 ± 4.29% during the first 5 min in unsafe context B and 0.39 ± 0.28% vs. 9.94 ± 1.85% during 10 s following the time of expected shock in context B, results averaged from the first three out of ten trials; $p = 0.0003$ and $p = 0.0001$, respectively, Mann-Whitney); none of the control animals emitted 44-kHz calls, neither the fear conditioning day nor the test days.

Delay fear conditioning. The procedure and its results were described before[2,3]; rats received 1, 6 or 10 conditioning stimuli (20-s-long white light co-terminating with an electric foot-shock, 1 s, 1 mA). Control animals showed a lower level of freezing than conditioned animals. There were only 4 ultrasonic calls we classified as 44-kHz vocalizations among 4,126 vocalizations emitted by the control rats during training and testing. We did not observe any difference in the number of 44-kHz vocalizations between Wistar rats with transmitters vs. without transmitters during delay conditioning training ($p = 0.8642$, Mann-Whitney). These two groups were therefore reported together.

Freezing behavior was scored automatically using Video Freeze software (Med Associates). To avoid including brief moments of the animal's stillness, freezing was measured only if the animal did not move for at least 1 s.

Ultrasonic playback

It was performed as described in[1-3] in individual experimental cages with acoustic stimuli presented through a Vifa ultrasonic speaker (Avisoft Bioacoustics, Berlin, Germany) connected to an UltraSoundGate Player 116 (Avisoft Bioacoustics). Ultrasonic vocalizations emitted by the rat were recorded by a CM16/CMPA condenser microphone (Avisoft Bioacoustics). Both playback and recording of calls were performed using Avisoft Recorder USGH software (version 4.2.28, Avisoft Bioacoustics). The locomotor activity was recorded with an acA1300-60gc camera (Basler AG, Ahrensburg, Germany). Sets of ultrasonic vocalizations presented:

- 22-kHz long calls, 8 calls in 1 repeat, typical 22-kHz band constant frequency (max-min frequency difference 1.9 ± 0.9 kHz), 24.5 ± 0.2 kHz peak frequency, 1066.4 ± 90.2 ms duration with 195.6 ± 15.5 ms sound intervals;
- 44-kHz long calls, 8 calls in 1 repeat, constant frequency (max-min frequency difference 2.7 ± 0.1 kHz), 42.1 ± 0.2 kHz peak frequency, 1064.3 ± 89.6 ms duration with 199.0 ± 14.7 ms sound intervals;
- 50-kHz modulated calls, 23 calls in 2 repeats, moderately modulated calls (max-min frequency difference 8.6 ± 0.3 kHz), 61.0 ± 0.8 kHz peak frequency, 37.6 ± 1.5 ms duration with 183.7 ± 4.5 ms sound intervals.

Playback procedure, rats with transmitters; as previously described in[1-3]. In short, after 10 min of silence, Wistar rats were exposed to four 10-s-long call sets with 5-min-long ITI in-between; the order of the presented sets was randomized with alternated appetitive or aversive playbacks.

Playback procedure, rats without transmitters; after 5 min of initial silence, Sprague-Dawley rats were presented with two 10-s-long playback sets of either 22-kHz or 44-kHz calls, followed by one 50-kHz modulated call 10-s set and another two playback sets of either 44-kHz or 22-kHz calls not previously heard. The playback presentations were separated by 3 min ITI.

Locomotor activity in playback. An automated video tracking system (Ethovision XT 10, Noldus, Wageningen, The Netherlands) was used to measure the total distance travelled (cm). Proximity to the

speaker was expressed as the percentage of time spent in the half of the cage closer to the ultrasonic speaker. Center-point of each animal's shape was used as a reference point for measurements of locomotor activity thus registering only full-body movements.

Analysis of ultrasonic vocalizations

Audio recordings were analyzed manually using SASLab Pro (version 5.2.xx, Avisoft Bioacoustics) as described[1-3] to measure key features of calls and categorize them into subtypes.

22-kHz vs. 44-kHz frequency ratio. A clear transition point between 22-kHz and 44-kHz long calls was observed in n = 13 Wistar rats and n = 1 SHR. In each case, ten 22-kHz calls followed by ten 44-kHz calls were analyzed.

Step up and step down frequency ratio. Rats which emitted at least five vocalizations of the specific subtype were analyzed (step up, n = 14; step down, n = 13; 5 calls of the two subtypes from each rat were chosen randomly and the frequencies of their elements were measured.

Ultrasonic vocalizations clustering (two independent methods)

Calls of conditioned and control animals were taken from all fear conditioning training sessions (n = 218 rats). We used **DBSCAN algorithm**[6]; a density based method, from the scikit-learn (sklearn) Python package, because of its ability to detect a desired number of clusters of arbitrary shape; with two main input parameters: MinPts (minimal number of points forming the core of the cluster) and ε (the maximum distance two points can be from one another while still belonging to the same cluster). To avoid detecting small clusters, we limited MinPts to 150 samples. The heuristic method described by Ester et al.[6] was implemented to find the initial range of ε. All the input data were standardized. The silhouette coefficient[7] was used to control the quality of the clustering. Maximizing ε among different ranges helped to select the most relevant number of identified clusters. Clustering with ε in the range of 0.14–0.2 resulted in a silhouette coefficient around 0.2–0.5.

K-means algorithm. Vocalizations of all fear-conditioned rats with 6-10 shocks and >30 of 44-kHz calls (n = 26) were detected using a built-in neural network for long rat calls (Long Rat Detector YOLO R1) on DeepSqueak[8] software (version 3.0.4) running under MATLAB (version 2021b, MathWorks, Natick, MA, USA) and manually revised for missed and mismatched calls. Unsupervised k-means clustering was based on call contour, frequency and duration variables, with equal weights assigned, and several descending elbow optimization parameters were used to obtain different maximum numbers of clusters together with Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP)[9] for superimposing and visualization of clusters.

Quantification and statistical analysis

Data were analyzed using non-parametric Friedman, Wilcoxon, Mann-Whitney tests with GraphPad Prism 8.4.3 (GraphPad Software, San Diego, CA, USA); the p values are given, $p < 0.05$ as the minimal level of significance. Figures were prepared using the same software and depict average values with a standard error of the mean (SEM).

Acknowledgements

We thank Iryna Artemieva for her help with DeepSqueek analysis. This research was funded by the National Science Centre, Poland, grant OPUS no. 2015/19/B/NZ4/03393 (R.K.F.) and by Mossakowski Medical Research Institute, PAS, Poland, Internal Research Fund no. FBW-17 (R.K.F.).

Author contributions

K.H.O, and R.P, and R.K.F. designed the study and wrote the manuscript. K.H.O., R.P., A.D.W., A.W.G, and O.G. performed the experiments. W.P. and M.K. performed DBSCAN analysis. R.P. performed k-means analysis. K.H.O., R.P., I.A.Ł., and A.D.W. analyzed the data. R.K.F. acquired the funding and supervised the project. All authors reviewed and approved the final version of the manuscript.

Literature for Methods

1. Olszynski, K.H., Polowy, R., Malz, M., Boguszewski, P.M., and Filipkowski, R.K. (2020). Playback of Alarm and Appetitive Calls Differentially Impacts Vocal, Heart-Rate, and Motor Response in Rats. iScience *23*, 101577. 10.1016/j.isci.2020.101577.
2. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., Zielinski, J., and Filipkowski, R.K. (2022). Spontaneously hypertensive rats manifest deficits in emotional response to 22-kHz and 50-kHz ultrasonic playback. Prog Neuropsychopharmacol Biol Psychiatry *120*, 110615. 10.1016/j.pnpbp.2022.110615.
3. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., and Filipkowski, R.K. (2021). Increased Vocalization of Rats in Response to Ultrasonic Playback as a Sign of Hypervigilance Following Fear Conditioning. Brain Sci *11*. 10.3390/brainsci11080970.
4. Pestana-Oliveira, N., Nahey, D.B., Johnson, T., and Collister, J.P. (2020). Development of the Deoxycorticosterone Acetate (DOCA)-salt Hypertensive Rat Model. Bio Protoc *10*, e3708. 10.21769/BioProtoc.3708.
5. Jaholkowski, P., Kiryk, A., Jedynak, P., Ben Abdallah, N.M., Knapska, E., Kowalczyk, A., Piechal, A., Blecharz-Klin, K., Figiel, I., Lioudyno, V., et al. (2009). New hippocampal neurons are not obligatory for memory formation; cyclin D2 knockout mice with no adult brain neurogenesis show learning. Learn Mem *16*, 439-451. 10.1101/lm.1459709.
6. Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In 34. pp. 226-231.
7. Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics *20*, 53-65.
8. Coffey, K.R., Marx, R.G., and Neumaier, J.F. (2019). DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. Neuropsychopharmacology *44*, 859-868. 10.1038/s41386-018-0303-6.
9. McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

Supplementary Figures

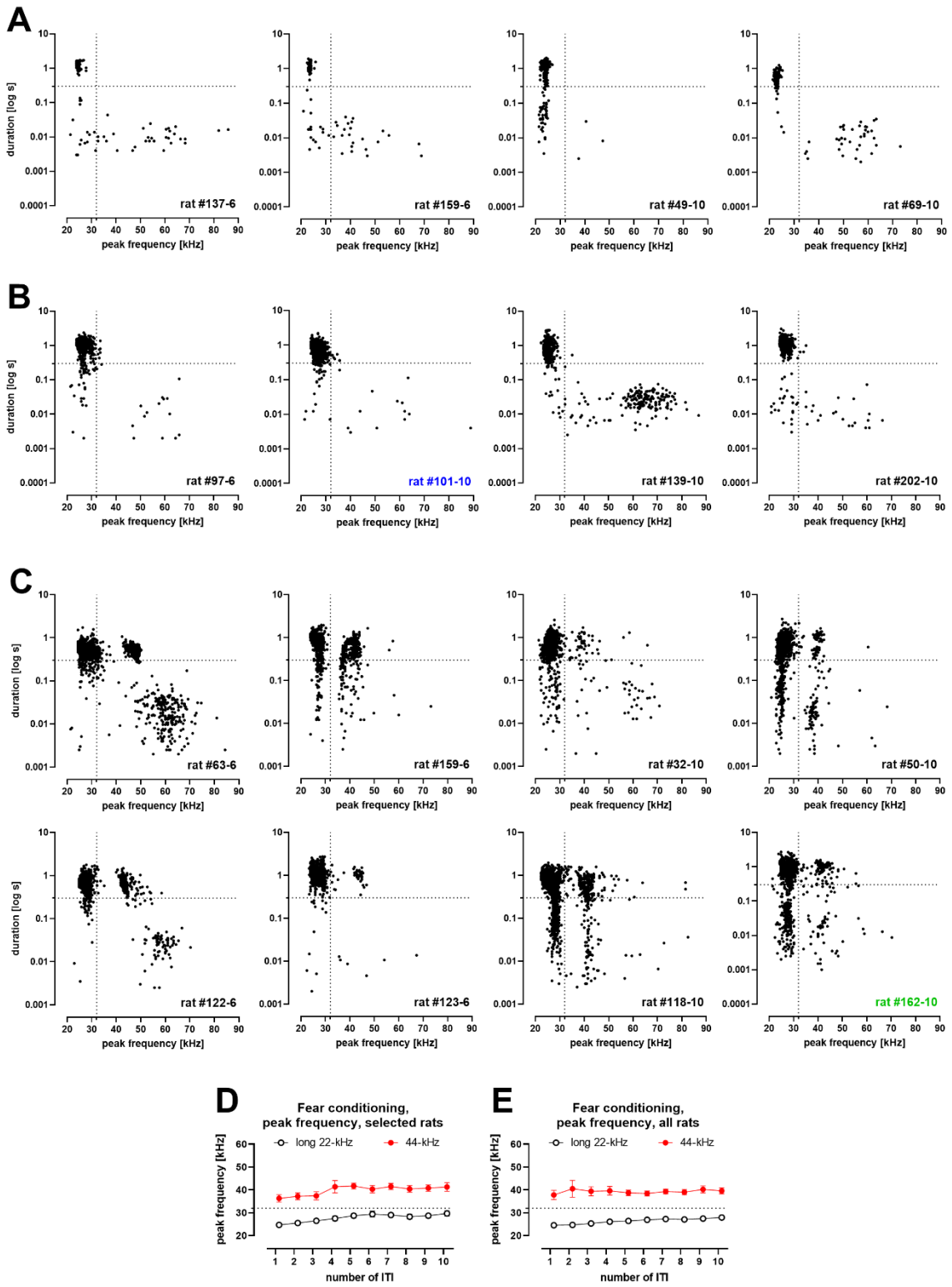

Fig. S1. Variations of call frequency; shown in relation to call duration (ABC) and over ten subsequent aversive trials (DE) in Wistar rats. ABC – Vocalizations plotted in relation to peak frequency (x axis) and duration (y axis). Each point corresponds to one vocalization. Vertical dotted line marks threshold

value (32 kHz) between 22-kHz and 50-kHz calls. Horizontal dotted line marks threshold value (300 ms) between short and long 22-kHz calls[3]. Rat identifier is given in lower right corner; the number after dash indicates the number of conditioning trials. **A** – examples from four rats which emitted typical long 22-kHz calls (no 44 kHz calls). **B** – four typical long 22-kHz vocalizations with few long 22-kHz calls crossing the 32 kHz threshold. **C** – eight sample rats which emitted typical long 22-kHz vocalizations and atypical high-frequency aversive calls forming a distinct 44-kHz group. **DE** – frequencies of 22-kHz and 44-kHz vocalizations in Wistar rats over ten aversive trials: only rats that emitted 44-kHz calls in at least seven ITI are plotted (**D**); all Wistar rats which received 10 conditioning trials (**E**). Horizontal dotted line marks threshold value (32 kHz) between 22-kHz and 50-kHz calls. The numbers of inter-trial intervals (ITI) correspond to the numbers of the previous stimuli. **D** – peak frequency in subsequent ITI rose gradually (for long 22-kHz calls: $p < 0.0001$, Friedman, $p = 0.0039$, Wilcoxon; for 44-kHz calls: $p = 0.0155$, Friedman, $p = 0.0977$, Wilcoxon). **E** – peak frequency of long 22-kHz calls in subsequent ITI rose gradually (for long 22-kHz calls, $p < 0.0001$, Friedman, $p = 0.0005$, Wilcoxon); unable to determine for 44-kHz calls due to low n number. Values are means ± SEM, **D**: n = 9, **E**: n = 46.

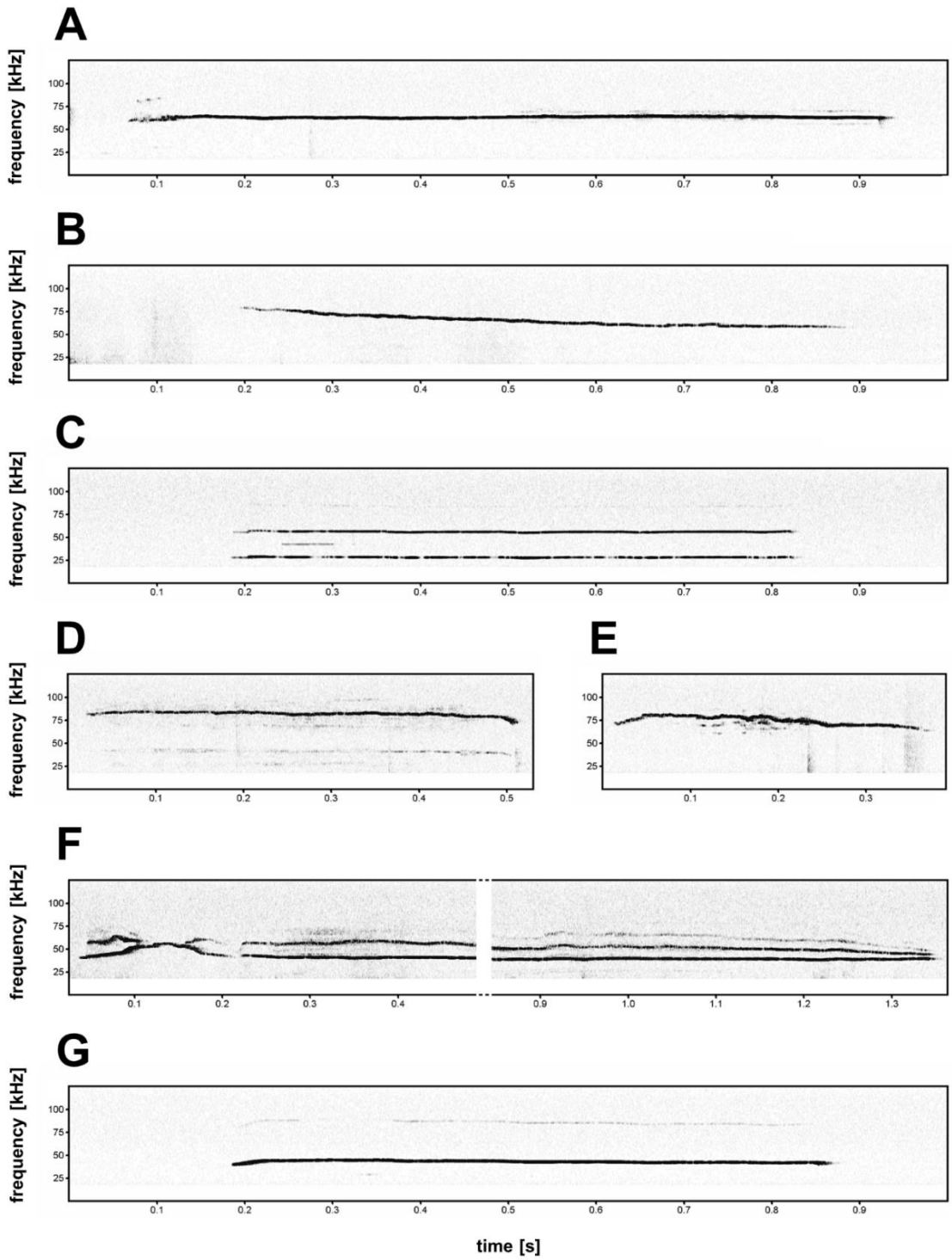

Fig. S2. Non-typical 44-kHz aversive vocalizations. A, **B** – constant frequency calls with very high peak frequency (**A**, peak frequency = 62.9 kHz; **B**, peak frequency = 65.9 kHz, start peak frequency = 78.1 kHz). **C**, **D** – harmonic aversive vocalizations, where element with fundamental frequency (F0, lowest frequency of the vocalization) is not with maximum amplitude, i.e. peak frequency is determined from the higher call component (**C**, F0 = 27.8 kHz, peak frequency = 55.6 kHz; **D**, F0 = 40 kHz, peak frequency = 81.5 kHz). **E**, **F** – vocalizations with prominent duration but with modulated frequency (**E**, peak frequency = 69.3 kHz; **F**, peak frequency = 39.0 kHz). **A**, **G** – constant frequency calls from SHR (**G**, flat 44-kHz call, peak frequency = 42.4 kHz).

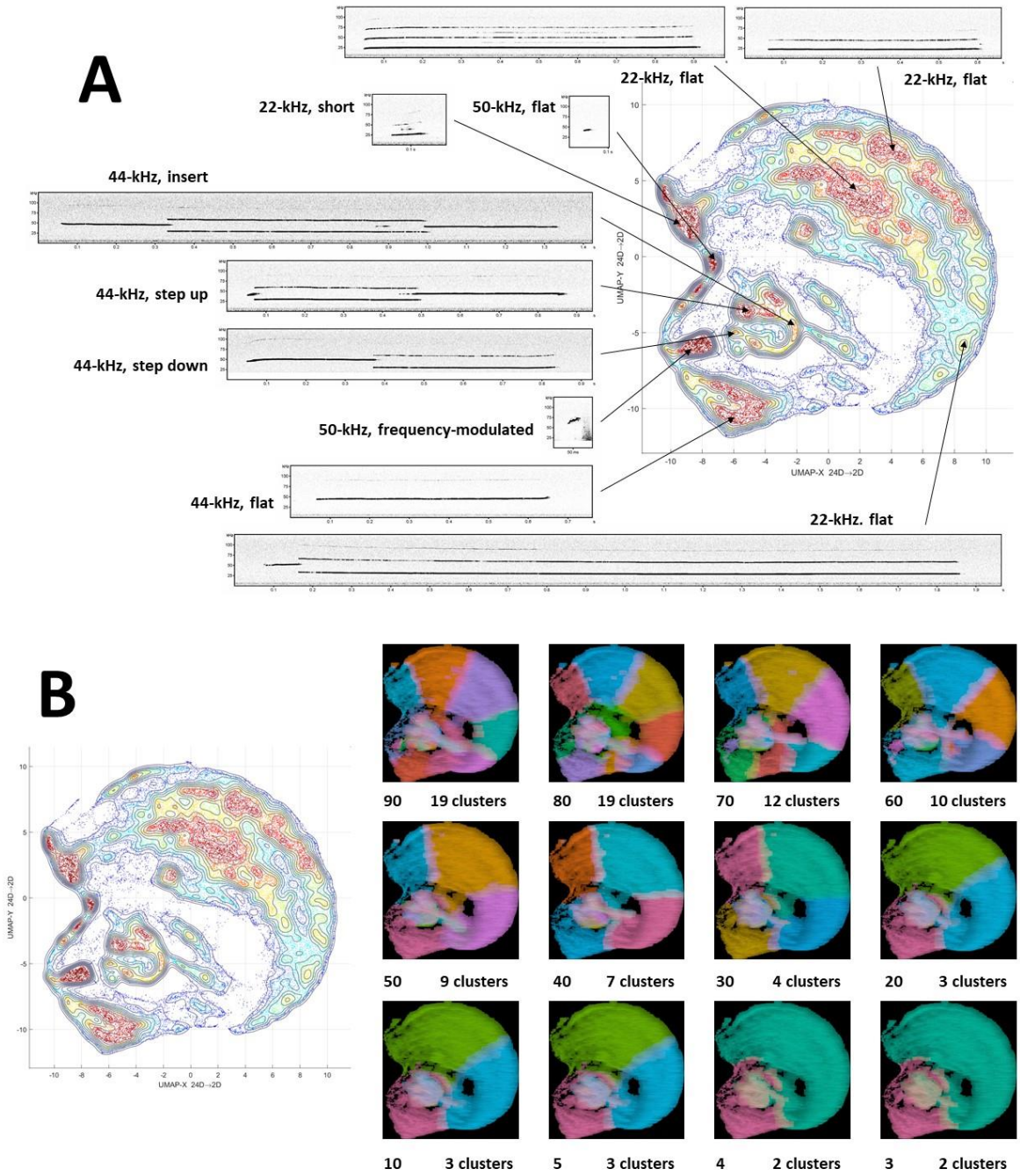

Fig. S3. **Clustering of ultrasonic vocalizations from rats emitting 44-kHz calls using UMAP projection and k-means. A** – topological plot of ultrasonic calls using UMAP embedding from rats emitting 44-kHz vocalizations during trace and delay fear conditioning training (n = 26), total number of calls n = 40,084, with spectrogram miniatures pointing to the general location from which they originated. **B** – comparison of unsupervised k-means clustering with different maximum possible number of clusters using elbow optimization (different clusters denoted by colors) done by DeepSqueak software, superposed over UMAP topological plot, number on the bottom left of the miniature denotes the maximum possible number of clusters set for elbow optimization, number on the bottom right denotes the resulting number of clusters after elbow optimization.

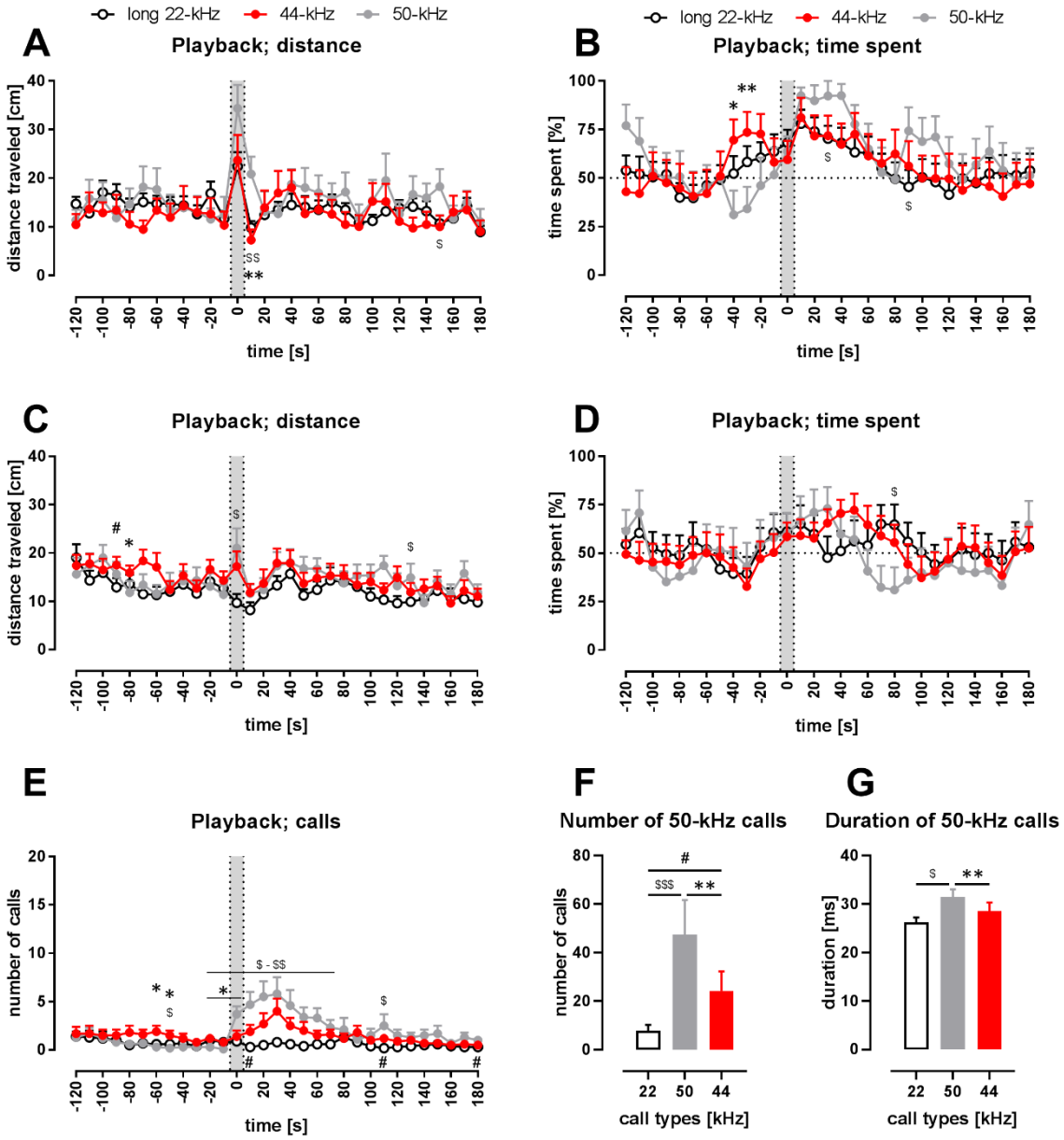

Fig. S4. Behavioral response to playback of 44-kHz calls (vs. 50-kHz and 22-kHz calls). AB – rats with implanted heart-rate transmitters (comp. Fig. 4), Wistar, n = 13-16; **C-G** – rats without transmitters, Sprague-Dawley, n = 15; **AC** – distance traveled; **BD** – time spent in the speaker's half of the cage; the dotted horizontal line marks a 50% chance value for time in a side of the cage; **E** – number of emitted vocalizations; **A-E** – gray sections correspond to the 10-s-long ultrasonic presentation, each point is a mean for a 10-s-long time-interval with SEM. **FG** – properties of 50-kHz vocalizations emitted in response to ultrasonic playback, i.e. number of calls (**F**) and duration (**G**) in 0-120 s range. **A-D** – playback presentation resulted in increased motor activity in case of, especially, 50-kHz playback and 44-kHz playback. Also, all kinds of playback resulted in increased time spent in the half of the cage next to the speaker. **E** – 50-kHz playback resulted in a rise of the number of evoked vocalizations (average from -30 to -10 time-intervals aka *before* vs. 10-30 s time-interval, p = 0.0010) as was the case with 44-kHz playback (p = 0.0142), respectively, while no rise was observed following 22-kHz playback (p = 0.2271, all Wilcoxon). However, since the increase in vocalization was robust in case of 50-kHz playback, the number of emitted vocalizations was higher than both after 22-kHz playback (e.g. p <

0.01 during 0-20 time-intervals) and after 44-kHz playback ($p = 0.0172$, 0 s time-interval, all Mann-Whitney). Finally, when the increases in the number of emitted ultrasonic calls in comparison with *before* intervals were analyzed, there was a difference following 44-kHz vs. 22-kHz playbacks during the 40 s time interval ($p = 0.0017$, Wilcoxon, comp. Fig. 4B). **F** – During the 2 min following the onset of the playbacks, the rats emitted more ultrasonic calls during and after 50-kHz playback in comparison with 22-kHz ($p = 0.0002$) and 44-kHz ($p = 0.0067$) playbacks; also, the rats emitted more ultrasonic calls during and after 44-kHz playback in comparison with 22-kHz playback ($p = 0.0369$), comp. Fig. 4C; all Wilcoxon). **G** – Ultrasonic 50-kHz calls emitted in response differed also in their duration, i.e. they were shorter to 22-kHz ($p = 0.0195$) and 44-kHz ($p = 0.0039$) playbacks than to 50-kHz playback. The difference between the effects of 22-kHz and 44-kHz playbacks was not significant ($p = 0.5469$, comp. Fig. 4D; all Wilcoxon). * 50-kHz vs. 44-kHz, $ 50-kHz vs. 22-kHz, # 22-kHz vs. 44-kHz; one character (*, $ or #), $p < 0.05$; two, $p < 0.01$; three, $p < 0.001$; Mann-Whitney (**AB**) or Wilcoxon (**CD**). Values are means ± SEM.

Literature for Figures

1. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., and Filipkowski, R.K. (2021). Increased Vocalization of Rats in Response to Ultrasonic Playback as a Sign of Hypervigilance Following Fear Conditioning. Brain Sci *11*. 10.3390/brainsci11080970.
2. Olszynski, K.H., Polowy, R., Wardak, A.D., Grymanowska, A.W., Zielinski, J., and Filipkowski, R.K. (2022). Spontaneously hypertensive rats manifest deficits in emotional response to 22-kHz and 50-kHz ultrasonic playback. Prog Neuropsychopharmacol Biol Psychiatry *120*, 110615. 10.1016/j.pnpbp.2022.110615.
3. Brudzynski, S.M., Bihari, F., Ociepa, D., and Fu, X.W. (1993). Analysis of 22 kHz ultrasonic vocalization in laboratory rats: long and short calls. Physiol Behav *54*, 215-221. 10.1016/0031-9384(93)90102-l.

Supplementary Video

Supplementary Video 1. Rat emitting 22-kHz calls followed by 44-kHz calls. The vocalizations are visible on the spectrogram. Ultrasonic vocalizations were modified to be audible to humans.

Supplementary Audios

Supplementary Audio 1. Example of a step up subtype of 44-kHz vocalization modified to be audible to humans.

Supplementary Audio 2. Example of a step down subtype of 44-kHz vocalization modified to be audible to humans.