Genes, Brain and Behavior
Official publication of the International Behavioural and Neural Genetics Society

Genes, Brain and Behavior (2016) ***15:*** *243–259* doi: 10.1111/gbb.12274

Ultrasonic vocalizations of adult male *Foxp2*-mutant mice: behavioral contexts of arousal and emotion

S. Gaub[†], S. E. Fisher[‡,§] and G. Ehret[†]*

[†]Institute of Neurobiology, University of Ulm, Germany, [‡]Department of Language and Genetics, Max Planck Institute for Psycholinguistics, and [§]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
**Corresponding author: G. Ehret, Institute of Neurobiology, University of Ulm, D-89069 Ulm, Germany. E-mail: guenter.ehret@uni-ulm.de*

Adult mouse ultrasonic vocalizations (USVs) occur in multiple behavioral and stimulus contexts associated with various levels of arousal, emotion and social interaction. Here, in three experiments of increasing stimulus intensity (water; female urine; male interacting with adult female), we tested the hypothesis that USVs of adult males express the strength of arousal and emotion via different USV parameters (18 parameters analyzed). Furthermore, we analyzed two mouse lines with heterozygous *Foxp2* mutations (*R552H* missense, *S321X* nonsense), known to produce severe speech and language disorders in humans. These experiments allowed us to test whether intact *Foxp2* function is necessary for developing full adult USV repertoires, and whether mutations of this gene influence instinctive vocal expressions based on arousal and emotion. The results suggest that USV calling rate characterizes the arousal level, while sound pressure and spectrotemporal call complexity (overtones/harmonics, type of frequency jumps) may provide indices of levels of positive emotion. The presence of *Foxp2* mutations did not qualitatively affect the USVs; all USV types that were found in wild-type animals also occurred in heterozygous mutants. However, mice with *Foxp2* mutations displayed quantitative differences in USVs as compared to wild-types, and these changes were context dependent. Compared to wild-type animals, heterozygous mutants emitted mainly longer and louder USVs at higher minimum frequencies with a higher occurrence rate of overtones/harmonics and complex frequency jump types. We discuss possible hypotheses about *Foxp2* influence on emotional vocal expressions, which can be investigated in future experiments using selective knockdown of *Foxp2* in specific brain circuits.

Keywords: Adult mice, arousal, emotional vocalization, *Foxp2* mutation, positive emotion, speech disorder, ultrasonic vocalization

Received 9 April 2015, revised 1 July 2015, 13 September 2015, and 9 November 2015, accepted for publication 10 November 2015

House mouse (*Mus musculus*) ultrasonic vocalizations (USVs) allow the study of function and dysfunction of neural and genetic control of vocal production, emotional expression and regulation of social behavior (Fischer & Hammerschmidt 2011; Portfors & Perkel 2014; Scattoni 2011). In contrast to humans, house mice (laboratory strains) seem to lack vocal learning abilities (Hammerschmidt *et al.* 2012; Kikusui *et al.* 2011; Mahrt *et al.* 2013). Nonetheless, mouse USVs have been used to characterize arousal states (Bell 1974); sexual emotion and motivation (Guo & Holy 2007; Hanson & Hurley 2012; Nunez & Tan 1984; Nyby *et al.* 1977; Sales 1972; White *et al.* 1998); development and regulation of social interactions (Arriaga & Jarvis 2013; Chabout *et al.* 2012; Grimsley *et al.* 2011; Hammerschmidt *et al.* 2012; Moles *et al.* 2007); etiology of neurodevelopmental disorders such as autism (Jamain *et al.* 2008; Rotschafer *et al.* 2012; Scattoni *et al.* 2011; Woehr *et al.* 2010) and Down syndrome (Zampieri *et al.* 2014); muscarinic influence on the dopaminergic regulation of reward in the brain (Wang *et al.* 2008). Against this background, the present study pursued two primary objectives.

First, we hypothesized that acoustic parameters of USVs provide indices of the strength of arousal and/or emotion, as suggested in general models of affective states (Barrett & Russell 1999; Russell 1980). In this context, we defined 18 acoustic parameters of adult male USVs and compared USV properties under three conditions: (1) low arousal while presenting a cotton pad with water, (2) increased arousal with positive emotion while presenting female urine, and (3) high arousal with positive emotion when males interacted with a live female. These experiments addressed which features of USVs (if any) can provide robust indices of affective states in mice (Lahvis *et al.* 2011).

Second, we assessed USVs of adult male mice with *Foxp2* mutations known to cause human speech and language deficits (Lai *et al.* 2001; MacDermot *et al.* 2005; Watkins *et al.* 2002). This gene also plays important roles in auditory-guided vocal learning in songbirds (Haesler *et al.* 2007; Murugan *et al.* 2013; Wohlgemuth *et al.* 2014). Mouse studies with heterozygous *Foxp2* mutations (matching the state of the related human disorder) revealed generally normal motor function but deficits in acquisition of motor skills (French & Fisher 2014; French *et al.* 2012; Groszer *et al.* 2008) and learning of auditory-motor associations (Kurt *et al.* 2012). Innate vocalizations of *Foxp2* heterozygous pups had largely normal acoustic properties (Gaub *et al.* 2010). Consequences of *Foxp2* mutation for vocalization abilities of adult mice are unknown. Therefore, we analyzed the USVs of two mouse lines harboring different etiological *Foxp2* mutations (Groszer *et al.* 2008) and compared USVs of heterozygous mutants (het) with those of wild-type (wt) littermates. The mouse

lines carried either a *R552H* missense mutation leading to defective protein with disturbed transcription factor function (Vernes & Fisher 2009; Vernes *et al.* 2006), or a *S321X* mutation with a null allele and reduced levels of functional Foxp2 protein (Groszer *et al.* 2008; MacDermot *et al.* 2005; Vernes *et al.* 2006).

Materials and methods

Animals

Mutants of the *Foxp2-R552H* and of the *Foxp2-S321X* mouse lines were originally generated by a gene-driven *N-ethyl-N-nitrosourea* (ENU) mutagenesis strategy as described in Groszer *et al.* (2008). Marker-assisted backcrossing of founder males into the C3H/HenNHsd background strain was used to accelerate the homogenization of genomic background and to eliminate non-relevant mutations generated by ENU mutagenesis, prior to behavioral analyses. The parental generation of the test mice of both studied *Foxp2* lines was crossed with the background strain C3H/HenNHsd for at least five generations. Following this, animals heterozygous for one of the *Foxp2* mutations (*R552H* or *S321X*) were paired either with a heterozygous animal of the same *Foxp2* mouse line or with a wild-type animal (C3H/HenNHsd) in order to obtain the 24 wild-type (12 wt littermates each of *R552H* and *S321X* heterozygotes) and the 24 heterozygous adult males (12 het of each *R552H* or *S321X* mutation) assessed in this study. The behavioral tests were carried out at the age of nine to ten weeks. Animals were housed under standard conditions in plastic cages with wood shavings as nest material under a 12 h-light–dark cycle (light on at 0800 h) and an average temperature of 23°C. Genotyping of the mice was conducted as previously described in Gaub *et al.* (2010).

Experimental situations, recording of ultrasounds and behavior

Experiments were performed between 0800 h and 1800 h under dim red light (although all the animals, just like other C3H lines, had retinal degeneration due to the presence of a naturally occurring mutation (rd1) in the phosphodiesterase 6B gene on this strain background) in a sound-proof and anechoic room at an average temperature of 23°C. One week before the experiments started, the males were housed alone in a plastic cage (26.5 × 20 × 14 cm). Three types of experiments were performed, each with the same 12 males of each genotype (wt, het) of the *R552H* mouse line, experiments 1 and 2 also with the same 12 males of each genotype (wt, het) of the *S321X* line. The duration of the measurements was always 30 min in which the male's USVs were recorded synchronously with its behavior.

Experiment 1: Five minutes before the recordings started, the metal grid of the cage (holding food and water bottle) was removed and the cage walls were heightened by a 14-cm plastic top allowing free sight and access into the cage without males being able to escape. A cotton pad with a 20-µl drop of demineralised water was placed in the male's cage.

Experiment 2: After preparation of the cage as in experiment 1, a cotton pad with a 20 µl drop of female urine was placed in the male's cage. The urine sample consisted of 10 µl fresh (maximally 3 min old) yellow female urine diluted in 10 µl demineralised water. Six males of each genotype were presented with urine from a female in estrus, the other six males from a female in metestrus/diestrus. Estrus phase was determined by microscopical evaluation of vaginal smears (Becker *et al.* 2005; Van der Gulden *et al.* 1975). The females were at least 61 days old and of the same genotype and *Foxp2* mouse line, but from a different litter as the tested male.

Experiment 3: This took place in a large plastic cage (42.5 × 26.5 × 15 cm) with fresh wood shavings which was divided in two compartments of equal size by a plastic wall with a small closable passage. A male (marked with a dot of opaque white at his head) and an adult (at least 63 days old) female of the same genotype and *Foxp2* mouse line, but from a different litter were placed in each compartment, respectively. They could acclimatize for two hours with food and water available and for another half hour with the grid covers removed and the cage walls heightened by a 15-cm plastic top, for the same purpose as described for experiments 1 and 2. To identify the actual start point of the 30 min interaction, the recordings of USVs and behavior started one minute before the closure between the compartments was opened. The two compartment cage was indirectly illuminated by two infrared lamps. Six *R552H* heterozygous and wild-type males respectively, interacted either with a female in estrus or with a female in metestrus/diestrus. Estrus phase was determined (as mentioned above) before the female was put into the test cage.

Ultrasonic vocalizations were recorded with a calibrated condenser microphone (Bruel and Kjaer, Model 4135, used without protection grid; Bruel & Kjear, Bremen, Germany) with preamplifier (Bruel and Kjaer, Model 2633; Bruel & Kjaer, Bremen, Germany) positioned centrically about 28 cm above the test cage. The microphone output was high-pass (cutoff at 20 kHz; tests 1 and 2) or band-pass (500–100 kHz; test 3) filtered (Kemo VBF 10M, 132 dB/octave slope; Kemo Limited, Dartford, Kent, UK). The filter output was amplified (Bruel and Kjaer measuring amplifier, Model 2610, 40 dB setting; Bruel & Kjear, Bremen, Germany) and digitized (Toshiba notebook CPU, 500 kHz DAQCard-6062E National Instruments, NIDISK software version 2.0; Engineering Design, Berkeley, CA, USA) with a sampling rate of 357 143 Hz and a gain of 10.0 (experiments 1 and 2). In experiment 3, recordings were made with two microphones centered above the cage. This two-channel recording (gain of 20.0 for each channel) reduced the sampling rate of the equipment to 200 000 Hz restricting the frequency range of analysis to 100 kHz.

For synchronization of sound and video recordings, a digital control signal was sent by the sound recording system, when active, to trigger a red LED positioned in the view field of the video camera (Conrad CCD camera S/W; Conrad Electronic SE, Hirschau, Germany). The video camera output was digitized (Pinnacle Studio 500-PCI and software, version 10) and then displayed on a computer monitor so that the experimenter could observe the animals' behavior. The video files (MPEG2 format) were stored with a resolution of 720 × 576 and a rate of 25 frames/second (Pinnacle Studio software, version 10) for later analysis. In addition to the sound recordings described above, an ultrasonic microphone (SM2 Microphone, Ultra Sound Advice, Wimbledon, London, UK) connected with a bat detector (S-25 Bat Detector, Ultra Sound Advice, Wimbledon, London, UK) was fixed above the test cage in order to receive and transform the animals' vocalizations to the hearing range of the experimenter who was outside the room and thus enabled notes to be taken about occurrences of USVs and types of behavior.

Analysis of ultrasounds and behavior

The recorded USVs were analyzed with SIGNAL software versions 4.0 and 4.1 (Engineering Design, Berkeley, CA, USA) using a 1024-point Fourier fast transformation and a Hanning window. Vocalizations were displayed as waveforms and spectrograms. The following measures of the four parameter classes USV-type occurrence, rate of USV production, USV intensity and spectrotemporal parameters of the USVs were taken and the respective data (points a, b, c, … below) analyzed. The measures concern levels of arousal and emotion and the motor ability to produce simple and complex sounds at various rates, as follows.

USV-type occurrence: (a) Latency to first USV occurrence after placement of the cotton pad with water/urine in experiments 1 and 2 (in experiment 3 this latency could not be specified because the male's vocalizations depended on the behavior of the interacting female); (b) division of USVs in simple calls without frequency jumps and calls with frequency jumps using the terminology of jump types according to Holy and Guo (2005) as well as two further jump types seen in our experiments; (c) percentage of USVs with at least one frequency jump; and (d) percentage of USVs with more than one frequency jump.

Rate of USV production: (e) Number of USVs per second during the first minute of stimulus presentation in experiments 1 and 2

or during a specified (see below) type of behavior in experiment 3; (f) number of USV series during the first minute of stimulus presentation in experiments 1 and 2; (g) number of single USVs and number of USVs belonging to a series during the first minute of stimulus presentation in experiments 1 and 2; (h) number of USVs within a series during the first minute of stimulus presentation in experiments 1 and 2; (i) USV duration; (j) duration of inter-USV intervals in series of USVs. A USV series was defined as USVs separated by inter-USV intervals of 600 milliseconds or less. On average, the number of inter-USV intervals in the vocalizations had a peak near 100 milliseconds and then decreased with an exponential function of the type $y = y_0 + A\ e^{-ax}$ to reach a constant low value by 600 milliseconds (see Gaub & Ehret 2005).

Intensity and spectrotemporal parameters: (k) Peak sound pressure level (SPL); (l) percentage of USVs with overtones or harmonics; (m) peak SPL of USVs either with or without overtones/harmonics from the USVs recorded in experiment 2; (n) minimum frequency of USVs; (o) minimum frequency of USVs in relation to the number of jumps in experiment 2; (p) maximum frequency and frequency bandwidth (maximum minus minimum frequency) of USVs recorded in experiments 1 and 2. (These last two parameters were not available for many USVs in experiment 3 because of the 100 kHz upper frequency limit of recordings and, therefore, were not evaluated in experiment 3.)

For analyzing the peak SPLs of the USVs, their peak amplitudes were measured as voltages in the waveform displays. These voltages were calibrated by the voltage produced by a 25-kHz pure tone of 60 dB peak SPL recorded in exactly the same experimental conditions (without mice present) and equipment settings as the USVs in the experiments. SPLs were determined only when the USVs were free of superimposed noise, which could occur by movements of the mice in the cage (all three experiments), and free of superimposed broadband defensive calls of the female in experiment 3.

Video recordings were played back with the software ULEAD VIDEOSTUDIO (version 7.0). They could be synchronized in time with the recorded sound files with an accuracy of 40 milliseconds for detecting the starting point (cotton pad viewable at the top edge of the cage) of the sound analysis time window (experiments 1 and 2) and for allocating the sequences of the sound files to the corresponding behavior sequences of the video (experiment 3). We specified the following types of male behavior and analyzed the USVs occurring in synchrony with these behavior types in experiment 3:

- first sniffing of a body region of the female (*first sniffing*): first sniffing activity of the male during the 30-min phase of interactions at any body region of the female (head, genital region, flank, ventral or dorsal region);
- sniffing at the female's head (*head sniffing*): male sniffed at the female's head from the nose up to the ears inclusive;
- sniffing at the female's genital region (*genital sniffing*);
- sniffing each other (*mutual sniffing*): male and female sniffed at the same time at any body region of the partner.
- aimed approach to the female (*aimed approach*): male moved directly toward the female (from a distance of six centimeters at least) and touched any part of the female's body with its snout or its vibrissae. During *aimed approach*, the female was either passive or moved in any direction except the direction toward the male;
- successful mounting or attempt to mount the female (*successful mounting*, *attempt to mount*): mounting the female was successful, when the male was with copulation movements on the back of the female tightly clasping her flanks with his forepaws for at least three seconds. All attempts to mount the back of the female or, by mistake, the female's head for less than three seconds were counted as *attempt to mount*. A sniffing action immediately before mounting and sniffing actions in between a sequence of mounting actions that were not more than 2 seconds apart counted to the mounting actions, and were not noted separately as sniffing. Sniffing and mounting actions of the above-mentioned types were separately noted when there was a pause of more than 2 seconds between them.

Except for the behavior of *first sniffing*, we analyzed all USVs from five interactions of each of the above-mentioned behavior types arbitrarily taken from all interactions that occurred over the whole 30 min observation period. If possible, at least one interaction of a given type was taken from each 10 min period of the whole 30 min observation time. If a certain behavior type occurred less than five times, USVs from all the interactions of this type were analyzed.

Statistical analyses

The data samples recorded for each animal were used to calculate an average number, percentage or value for the respective measure (points a – p, see above) that characterizes the animal. These individual averages were then used to calculate the average values of experimental and/or genetic groups. These group data were the basis for statistical tests. There were several cases in which an experimental animal did not emit any USVs or USV series or did not show a certain interaction type in experiment 3. Such cases reduced the individuals from initially 12 to a smaller number contributing to the group data shown in the figures. In addition, there were a few cases in which extreme values of individual averages were identified with $P < 0.05$ as outliers of the group (Dixon 1953; see Sachs 1999). These values were not considered in the group averages and in the statistical tests based on them.

Statistical analyses of data were performed with SigmaStat software, version 3.1. In general, statistical differences between genotypes and stimuli (experiments 1, 2) or genotypes and interaction types (experiment 3) were detected with a two-way ANOVA in which one factor was the genotype and the other factor the stimulus or type of interaction. In case of significant effects of the factors, pairwise comparisons were automatically done with the Holm-Sidak method. When ANOVA conditions were not given and data could not be normalized by transformations, statistical differences were analyzed depending on the type of experiment: (1) Comparisons between the stimuli water and female urine (experiments 1, 2) were done separately for each genotype (t- or U-test) and then, if means or medians of the genotypes in a given experiment suggested differences, one or two further *t*- or *U*-tests were done to establish significance of the differences. In case of multiple comparisons (up to 4 comparisons), the α-value of 0.05 was adjusted with the Bonferroni procedure leaving differences of $P < 0.01$ significant. (2) Comparisons between the interaction types (experiment 3) were done separately for each genotype (ANOVA or ANOVA on ranks) and then, if means of the genotypes in a given interaction suggested differences, one further *t*- or *U*-test were done to establish significance of the differences. The α-value of 0.05 was again adjusted with the Bonferroni procedure. In the course of the data analysis for each parameter, a maximum of 24 comparisons turned out to be meaningful reducing the α-value to 0.002135 so that differences of $P < 0.001$ were considered significant. In the legends to the figures it is indicated which significant differences have been adjusted by the Bonferroni procedure. In order to reduce the probability of false positive results also in the cases in which the two-way ANOVA could be applied, significant differences in these cases were noted only if $P < 0.01$. All statistical tests were two-tailed. Tables S1 and S2 (Supporting Information) provide details of the statistical methods used for comparing given data sets in the figures of the main part of the study (Table S1) and the figures in the supplement (Table S2).

Statistically significant differences will be mentioned in the text and plotted as ** for $P < 0.01$ and *** for $P < 0.001$. Since the great majority of group data in the main body of the paper was normally distributed, all figures show the mean ± SD in order to facilitate comparisons, although in the statistical comparisons non-normally distributed data were properly treated. In the supplement to the paper, data from experiments 1 and 2 are plotted separately for both mouse lines (*R552H*, *S321X*) using boxplots with median, 25%-quartile, 75%-quartile and range for non-normally distributed data.

Results

General

The acoustic parameters of male USVs did not differ with respect to estrus state of the females used in the experiments. Therefore, we pooled data from experiments using

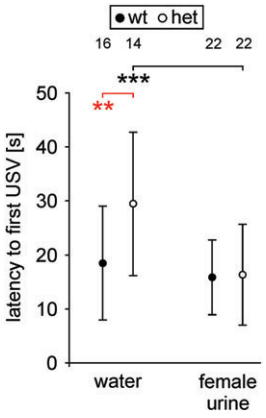

Figure 1: Latency to first USV occurrence. The latencies to the first USV occurrence in experiment 1 (water) and experiment 2 (female urine) are shown as means ± SD for the groups of wild-type (wt) and heterozygous (het) animals. Significant differences are indicated by ** ($P < 0.01$) or *** ($P < 0.001$). Significant differences marked in red color characterize differences between wild-types (wt) and heterozygotes (het). Animals of the het group responded significantly faster to female urine compared to water. Compared to the wt group the animals of the het group vocalized later the first USV in response to water. The numbers in the top row of the figure indicate the number of animals included in the respective measures.

females in estrus and in metestrus/diestrus. Further, as expected by the nearly identical genetic background of wt animals (due to extensive backcrossing prior to behavioral testing) there were no significant differences in any of the measures taken among the wild-types of the two mouse lines. Interestingly, we also did not detect differences in any of the measures as a consequence of the specific type of *Foxp2* mutation: *R552H* hets were not significantly different from *S321X* hets (data shown in Supporting Information). Prior studies have shown that these distinct mutations, which disturb *Foxp2* in different ways, display consistent outcomes for certain behavioral features (e.g. Groszer *et al.* 2008), while showing genotype-specific differences for others (e.g. Kurt *et al.* 2012). As consequence of the absence of statistically significant differences among the wt animals on the one hand and among the het animals on the other hand, we merged the data from all wt animals to form the wt group and from all heterozygous animals to form the het group in order to evaluate possible effects of arousal/emotion and *Foxp2* mutation on the vocalization of USVs. The results of these evaluations are shown in the main body of the paper.

Analyses of USV occurrence and acoustic parameters

Latency to first USV occurrence

In experiments 1 and 2 it took, on average, between 15–30 s for the first USV to occur after placement of the stimulus (Fig. 1). Animals of the het group were significantly faster to vocalize to urine compared to water. They were slower than the wts to vocalize to water. Wt animals showed very similar latencies to the first USV after both stimuli.

Types of USVs

USVs could have none, one or more than one frequency jump. First, we classified the USVs in those without frequency jumps (Fig. 2a) and in others with frequency jumps (Fig. 2b). We used the classification scheme of Holy and Guo (2005) to differentiate between the jump types h, d, u, hd, hdu and du (Fig. 2b) and added two further types. The 'hh' type equals the h-type (downward jump from a high frequency), but starts (110–135 kHz) and ends (70–100 kHz) in an even higher frequency range. The 'uh' type equals the u-type (upward jump), but starts (65–85 kHz) and ends (95 – 115 kHz) in a higher frequency range (Fig. 2b). As in Holy and Guo (2005), some USVs could not be classified in the mentioned jump types, partly because their frequency ranges were outside the indicated ones. Their occurrence rates are indicated by white areas in the pie charts of Fig. 2b.

A quantitative analysis showed that, in general, ~20–40% of the USVs of all genotypes in experiments 1 and 2 had at least one frequency jump (Fig. 3). In the female interactions (experiment 3), all mice emitted USVs with the highest rate of frequency jump(s) during the attempt to mount/mounting and head sniffing (Fig. 3). Often more than 40% of the USVs in these interactions had at least one frequency jump, which is a significantly higher rate compared to the rate from several other interaction types, which was often below 20% (Fig. 3). Figure 3 includes also the percentage of USVs without frequency jumps which is just the reverse of the plotted measure, on average, 100% minus the shown means. This view on Fig. 3 clearly indicates that most of the USVs emitted in the experiments had no frequency jumps and that the rate of USVs with frequency jumps, mainly with more than one frequency jump (see Fig. 4), is increased especially when the males sniffed the head of a female or tried to mount.

Next, we took the USVs with any number of frequency jumps (now 100%) and separated the USVs with one frequency jump from those with more than one frequency jump. Figure 4 shows that few (less than average 10%) USVs of wt and het animals had more than one frequency jump in response to water. Significantly more USVs of all genotypes had more than one frequency jump in response to urine compared with water. In addition, female interactions *sniffing/aimed approach* and *attempt to mount/mounting* produced significantly higher percentages of USVs with more than one frequency jump than water presentation, *attempt to mount/mounting* in all genotypes also more than female urine (Fig. 4). Altogether, Fig. 4 shows for both wt and het animals an increase in the occurrence rate of highly complex USVs, i.e. those with more than one frequency jump, with changing stimulus from water to urine to sniffing a female to mount a female.

The pie charts of Fig. 2b show a quantitative analysis of the occurrence rates of the jump types of USVs from experiments 1 and 2. Because of the very low occurrence rate of USVs with frequency jump(s) in the interaction types *first sniffing, genital sniffing, mutual sniffing*, *aimed approach*, especially in the heterozygous animals (Fig. 3), we did not classify the USVs from experiment 3 in jump types. When comparing jump types in response to water and urine separately in wt and het animals (ANOVA on ranks followed by U-test for verifying the main effects between vocalizations to

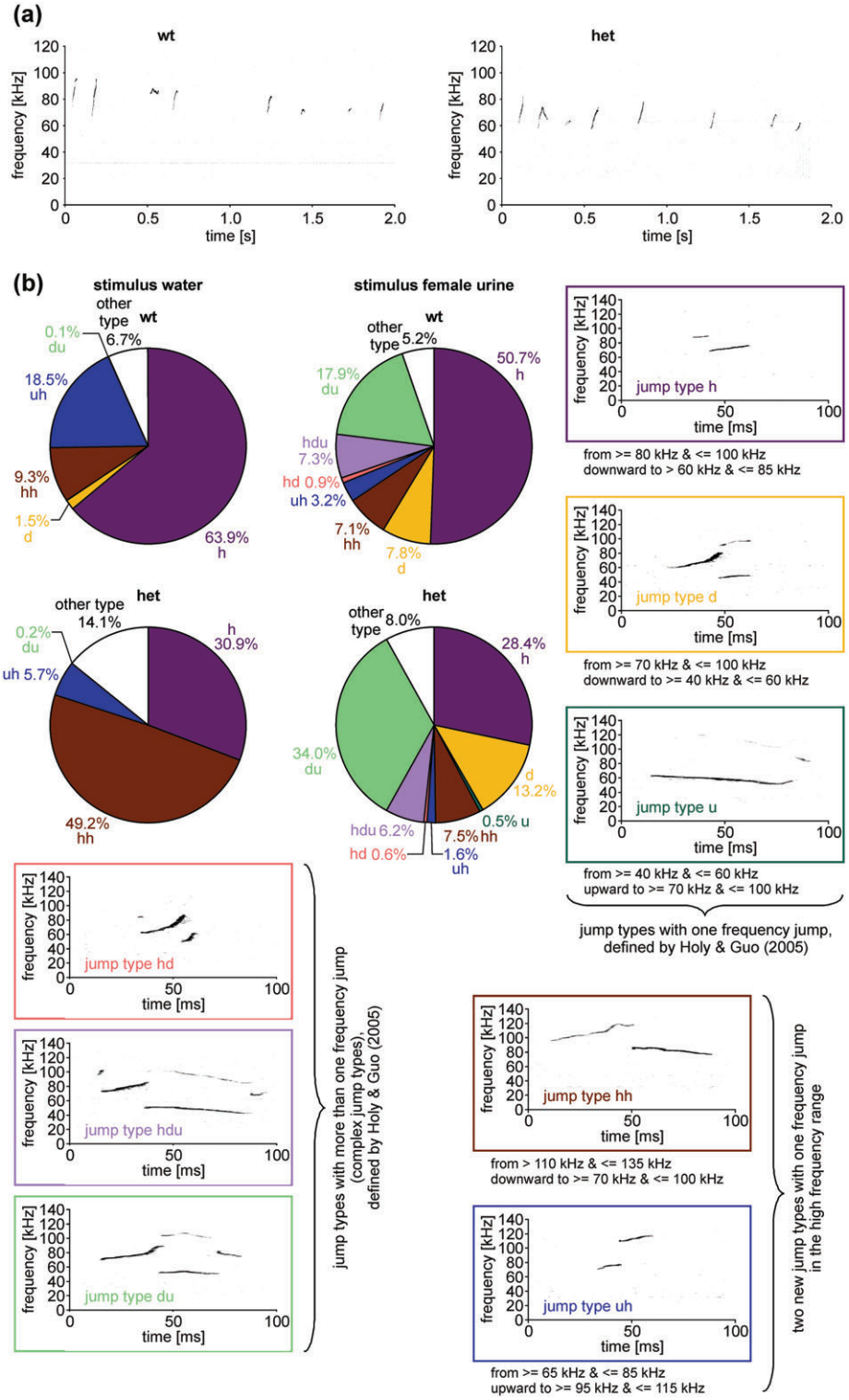

Figure 2: Types of USVs. (a) Example spectrograms of series of USVs without frequency jumps (simple type of USVs) from a wild-type (wt) and a heterozygous mutant (het). (b) Division of USVs with frequency jumps (complex types of USVs). The types of frequency jumps were defined according to Holy and Guo (2005) as jump types h, d, u, hd, hdu, du, other types, and expanded by the two further jump types hh and uh. The average occurrence rates of these jump types in experiments 1 (stimulus water) and 2 (stimulus female urine) are expressed in % of the USVs that had frequency jumps at all (compare Fig. 3). The colors in the pie graphs are the same as the colors of the frames of the USV jump types, the 'white' section representing 'other types' most of which did not fit with their frequency ranges to the indicated ones. USVs with upward jumps (u, uh) are generally rare. USVs with downward jumps (h, hh) are most abundant in response to water, highly complex types with more than one frequency jump (hd, hdu, du) are virtually absent in response to water but present at higher rates in response to urine. For further explanations, see text.

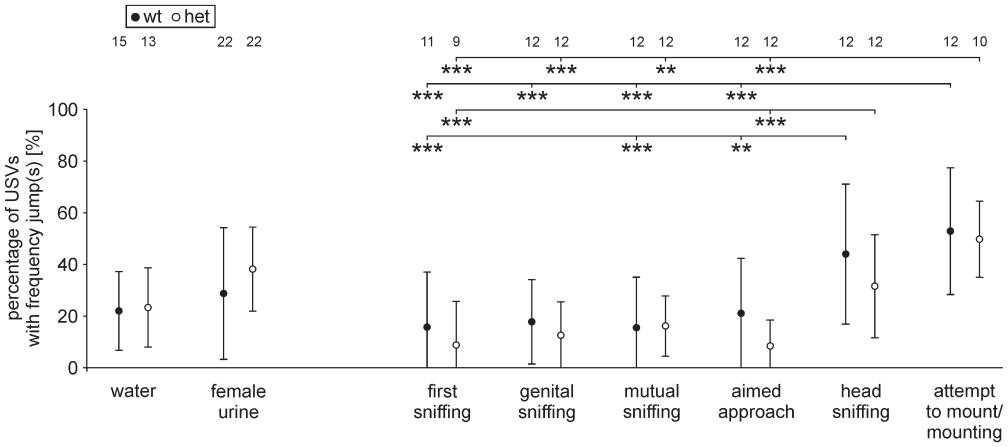

Figure 3: Percentage of USVs with frequency jump(s). The average percentages of USVs (mean ± SD) with frequency jump(s), i.e. USVs with at least one frequency jump, are shown for both groups (wt, het) in experiment 1 (water), experiment 2 (urine) and in the different female interactions in experiment 3. The means vary between about 10 – 50% indicating that most USVs were without frequency jump(s). Only during the female interaction of head sniffing and attempt to mount/mounting, nearly 50% of the USVs had frequency jumps, which are significantly more jumps than in some of the other female interactions. For further explanations see Fig. 1.

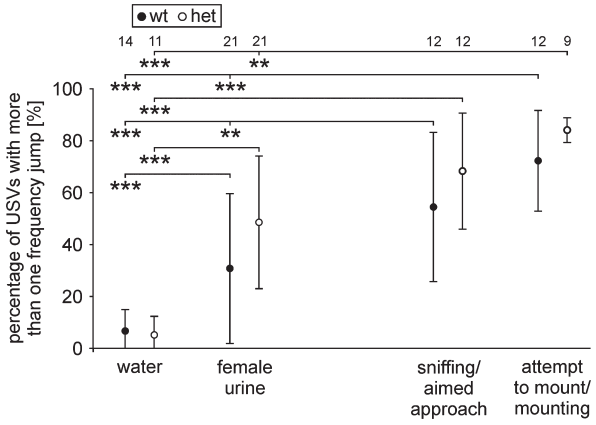

Figure 4: Percentage of USVs with more than one frequency jump. The average percentages of USVs with more than one frequency jump are shown for both groups (wt, het) and the indicated experimental situations. In this figure the data from all situations of female interaction in experiment 3 except attempt to mount/mounting have been combined to sniffing/aimed approach. There is a significant increase of the rates of USVs with more than one frequency jump from experiment 1 (water) to experiment 2 (female urine) to attempt to mount/mounting (experiment 3) with rates in the other female interactions being intermediate between the urine and mounting condition. For further explanations see Fig. 1.

water and urine and between genotypes), the following statistically significant differences were noted: Both wt and het animals vocalized in response to urine more jump types d, du and hdu than in response to water (each $P < 0.01$) (compare Fig. 2b). With regard to the genotypes we found that het compared to wt males vocalized significantly ($P < 0.01$) more jump type hh in response to water and more jump type du in response to urine.

Number of USVs

Heterozygotes of both mouse lines and their wild-type littermates all produced significantly more USVs per second to female urine (2–3/second) compared to water (less than 1/s) (Fig. 5). During interaction with a female, wt males produced significantly more USVs at *first sniffing* and *head sniffing* (near 5/s) compared with the other interaction types (2–3/second). Such significant differences occurred in the het group only between *head sniffing* and *aimed approach*, i.e. on average, heterozygotes were less variable in the number of emitted USVs over all interaction types than the wild-types (Fig. 5).

Most of the mounting actions were attempts to mount the female. Altogether, only four successful mounting sequences occurred, one by a wild-type and three by heterozygous males. The average number of USVs from the three successful mountings of the heterozygotes is shown separately from the average number of USVs from the attempt to mount (Fig. 5). Successful mounting did not elicit more USVs than the attempt to mount.

USV series, single USVs, number of USVs in a USV series

Without differences between the genotypes, most USVs (average 87.5%) were emitted in series, not as single USVs. Female urine elicited in the het group an average of 10 USV series during the 1 min observation time, which is significantly more USV series than to the stimulus water (2 series; Fig. 6a). This difference did not occur in the wt animals which produced average 8 series to urine and 6 series to

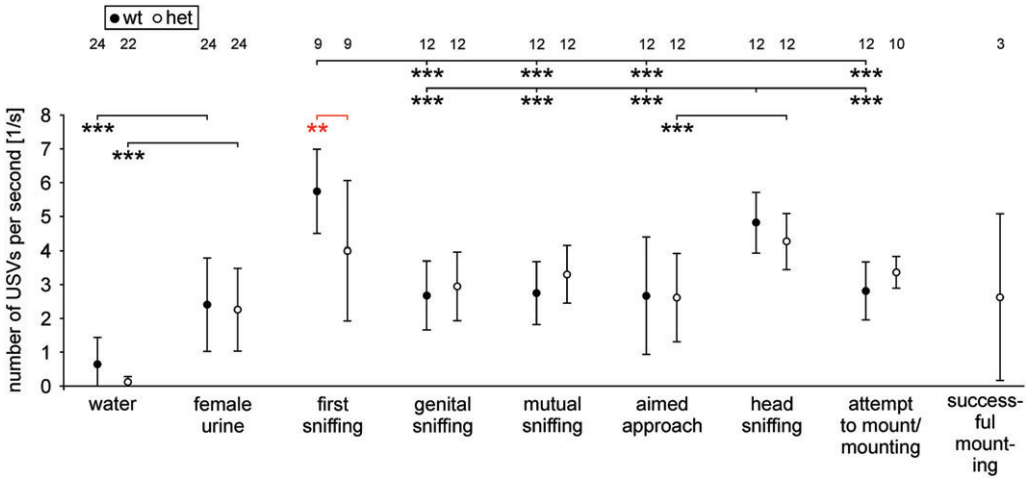

Figure 5: Number of USVs per second. The average numbers of USVs per second (mean ± SD) are shown for both groups (wt, het) and the indicated experimental situations. Female urine released significantly more USVs than water. The highest rates of USVs were detected during the first sniffing female interaction and during head sniffing. The average value for the het animals in response to water was 0.12 with a SD of 0.16. Statistical comparisons in the experiments with water and female urine had Bonferroni adjustment. For further explanations see Fig. 1.

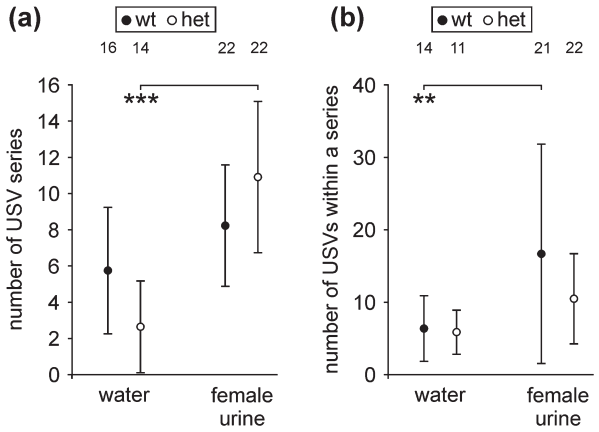

Figure 6: Number of USV series and number of USVs in a series. The average numbers of USV series **(a)** and the average numbers of USVs in a USV series **(b)** (mean ± SD) emitted during the first minute of stimulus presence in experiment 1 (water) and experiment 2 (urine) are shown for both groups (wt, het), respectively. In the het group significantly more USV series occurred in response to urine than in response to water (a), and in the wt group the number of USVs in a series was larger in response to female urine than to water (b). For further explanations see Fig. 1.

water (Fig. 6a). In the wt group, the number of USVs within a series was larger in response to urine compared to water (Fig. 6b). Taking the data of Fig. 6a,b together indicates that the increase in the number of USVs in response to urine compared to water (Fig. 5) may derive from both an increase of USV series and USVs within a series in the responses of all genotypes.

USV duration

The change from water to urine as stimulus significantly increased the USV duration in both wt and het males (Fig. 7a). In addition to this stimulus effect, we observed a genotype effect, i.e. the het group vocalized significantly longer USVs than the wt group in response to water and urine (Fig. 7a). Interaction with females led to significantly longer calls during *head sniffing* and *attempt to mount/mounting* (40–60 milliseconds) compared to other interaction types (20–40 milliseconds) both in wt and het males (Fig. 7a).

Duration of inter-USV intervals in series of USVs

USV series of het males elicited by female urine had significantly shorter inter-USV intervals (average 100 milliseconds) than those in response to water (average 200 milliseconds; Fig. 7b). This difference was not observed in the wt group. The inter-USV interval durations in series of USVs emitted in interactions of *head sniffing* and *attempt to mount/mounting* were shorter (less than 100 milliseconds) than in other sniffing interactions and aimed approach, significant in several cases for both wt and het males (Fig. 7b). In addition, wts had shorter inter-USV intervals than hets in the interaction types *first sniffing* and *aimed approach*.

Taking the data of USV durations (Fig. 7a) and inter-USV interval durations (Fig. 7b) together, it becomes evident that the sound density of vocalizations (presence of sound in a given time interval) increases, especially in het animals, when the stimulus changes from water to urine and from *first/genital/mutual sniffing/aimed approach* to *head sniffing and attempt to mount/mounting*.

Peak sound pressure level (SPL)

The USVs of all mouse groups were significantly louder to the stimulus urine than to water (Fig. 8). During interaction

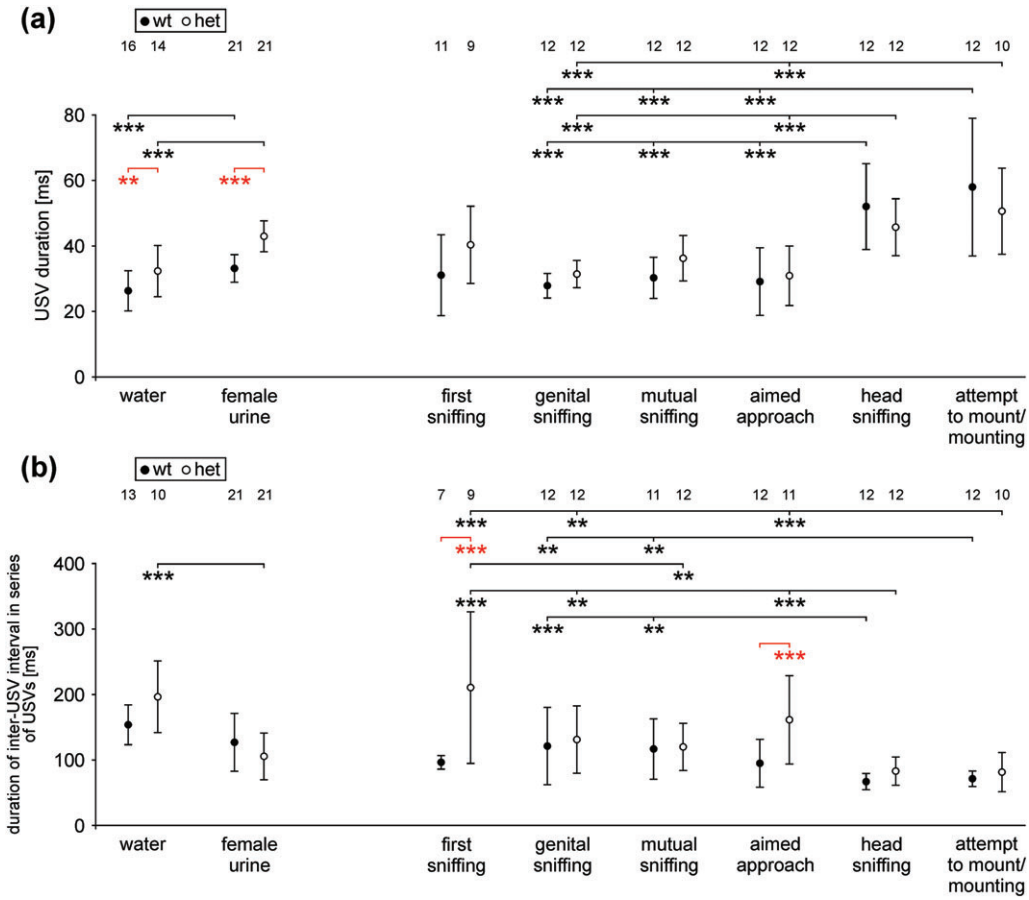

Figure 7: Duration of USVs and of inter-USV intervals in series of USVs. (a) The average durations of USVs (mean ± SD) and (b) the average durations of inter-USV intervals in USV series (mean ± SD) are shown for both groups (wt, het) and the indicated experimental situations, respectively. Female urine led to significantly longer USVs than water (a) and the longest USVs were recorded during head sniffing and attempt to mount/mounting interactions. Het group animals had USVs with longer durations than wt group animals in response to both water and female urine. Female urine led to significantly shorter inter-USV intervals than water in the het group (b) and the shortest inter-USV intervals were recorded during head sniffing and attempt to mount/mounting interactions. Het group animals had longer inter-USV interval durations than wts in the interactions of first sniffing and aimed approach (b). Statistical comparisons of USV duration in the interaction experiment (a) had Bonferroni adjustment. For further explanations see Fig. 1.

with a female, wt and het mice emitted the loudest USVs in *attempts to mount/mounting* and *head sniffing*, resulting in statistically significant differences with USVs from the other interaction types (Fig. 8). Further, a group difference was noted in the USVs to urine where the het mice produced significantly louder USVs than the wt littermates (Fig. 8).

Overtones or harmonics

All genotypes emitted significantly more USVs with overtones or harmonics in response to urine as compared to water (Fig. 9). In the female interactions, both wt and het mice emitted the highest rates (mostly more than 30%) of USVs with overtones or harmonics in the *attempt to mount/mounting* and during *head sniffing* (Fig. 9). These rates were often significantly higher than the rates (mostly below average 20%) emitted in other types of behavioral interactions. Further, a group difference was noted in the USVs to urine where the het mice produced significantly more USVs with overtones or harmonics than their wt littermates (Fig. 9).

Relationship between peak SPL and overtone/harmonics

The profiles both of peak SPL (Fig. 8) and percentage of USVs with overtones or harmonics (Fig. 9) in dependence on the stimulus situation were very similar. Louder USVs consistently had a higher probability of presence of overtone/harmonics compared to rather soft USVs. In order to confirm this relationship for all genotypes, we separated the USVs in response to female urine into those with or without overtone/harmonics and calculated the peak SPL separately for these groups. We did this only for the urine stimulus because USVs with overtone/harmonics to the water

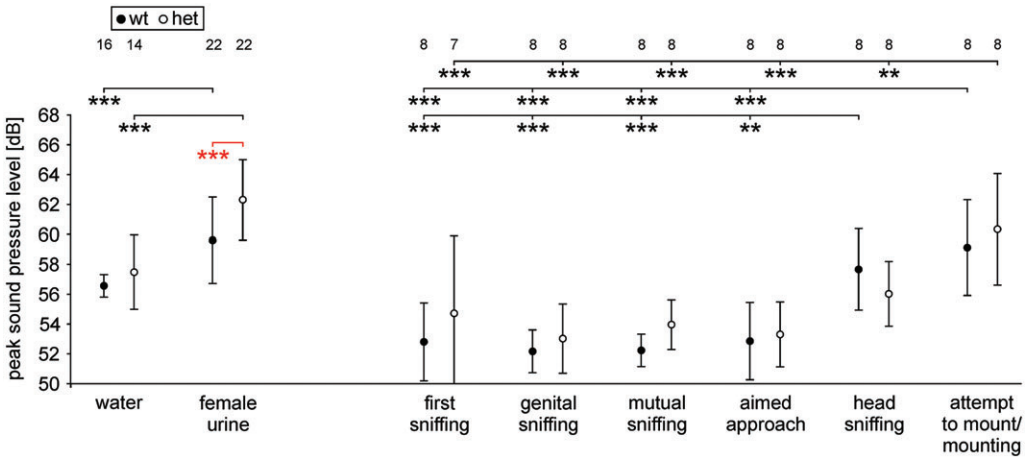

Figure 8: Peak sound pressure level (dB SPL) of USVs. The average peak SPLs of USVs (mean ± SD) are shown for both groups (wt, het) and the indicated experimental situations. Female urine led to significantly louder USVs than water. USVs recorded during head sniffing and attempt to mount/mounting interactions were often significantly louder than USVs from the other female interactions. Het group animals had louder USVs than wt group animals in response to female urine. For further explanations see Fig. 1.

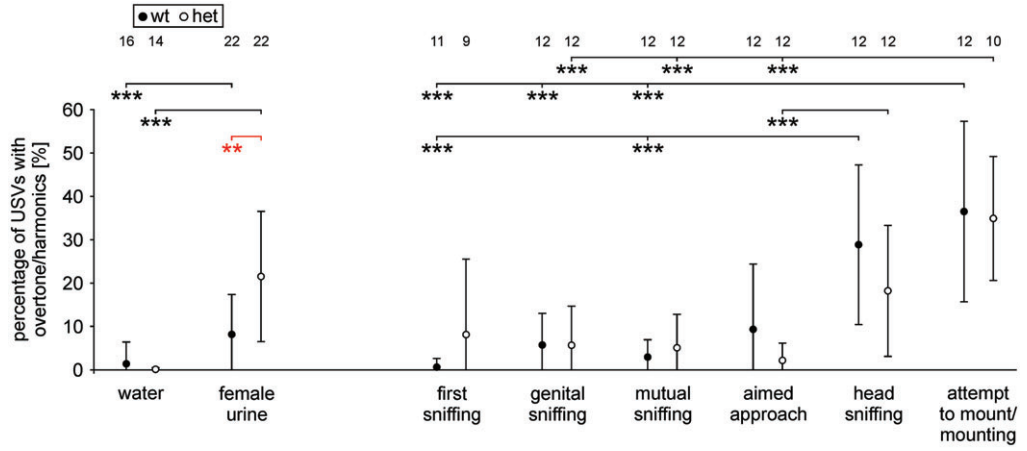

Figure 9: Percentage of USVs with overtones or harmonics. The average percentages of USVs with overtones or harmonics (mean ± SD) are shown for both groups (wt, het) and the indicated experimental situations. Female urine led to significantly more USVs with overtone/harmonics than water. USVs recorded during head sniffing and attempt to mount/mounting interactions had often significantly more overtone/harmonics than USVs from the other female interactions. Het group animals had higher percentages of USVs with overtones/harmonics than wts in response to female urine. Statistical comparisons in the experiments with water and female urine and in the interaction experiment had Bonferroni adjustment. For further explanations see Fig. 1.

stimulus were very rare (Fig. 9). As shown in the supplementary Fig. S11, the USVs with overtone/harmonics of all genotypes were, in fact, significantly (average 5dB) louder than the USVs without overtone/harmonics. This relationship must be expected because the louder sounds are, the easier overtones/harmonics can be detected by a given measuring device.

Minimum frequency

All genotypes produced USVs with significantly lower minimum frequency in response to urine compared to water (Fig. 10). In the female interactions, USVs emitted by both wt and het mice in attempts to mount/mounting had the lowest minimum frequency in comparison with other female interaction types, leading to several statistically significant differences (Fig. 10). Statistically significant differences between the groups were also found. In the experiments with water as stimulus and in the female interaction *mutual sniffing* USVs of the het group had higher minimum frequencies than those of the wt group.

Context of minimum frequency and frequency jumps

Minimum frequency of USVs in relation to their content of frequency jumps was analyzed in all genotypes for the urine

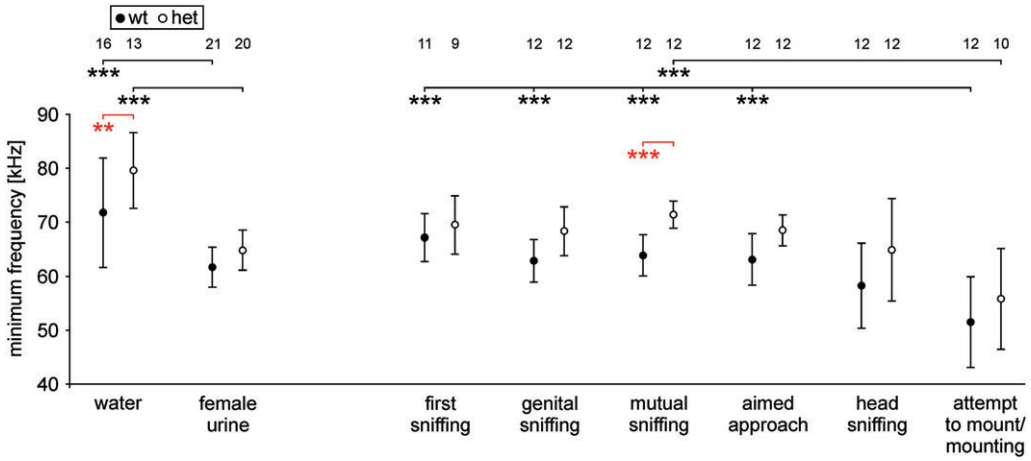

Figure 10: Minimum frequency of USVs. The average minimum frequencies of USVs (mean ± SD) are shown for both groups (wt, het) and the indicated experimental situations. Female urine led to significantly lower minimum frequencies than water. USVs recorded during attempt to mount/mounting interactions had often significantly lower minimum frequencies than USVs from other female interactions. Het group animals had USVs with higher minimum frequencies than wts in response to water and in the mutual sniffing interaction. Statistical comparisons in the interaction experiment had Bonferroni adjustment. For further explanations see Fig. 1.

stimulation (experiment 2), in which enough USVs without a jump, with one jump or with more than one jump were recorded to allow a meaningful analysis. Clearly, in both groups, USVs with more than one jump had the significantly lowest minimum frequencies (Fig. 11a). Further, animals of the het group had significantly higher minimum frequencies in the USVs without a jump than their respective wt littermates. This difference indicates that the small difference in minimum frequency between wt and het groups seen for urine stimulation in Fig. 10 originated mainly from USVs without jumps.

Maximum frequency and frequency bandwidth

Because the upper frequency limit of analysis was restricted to 100 kHz in experiment 3, we determined the maximum frequency and the frequency bandwidth (maximum minus minimum frequency) only for USVs from experiments 1 and 2. As Fig. 11b shows, all genotypes emitted USVs with significantly higher maximum frequencies to water compared with urine (Fig. 11b).

The comparison of the frequency bandwidths of the USVs (Fig. 11c) indicates similar bandwidths of wt USVs to water and urine while the het USVs had significantly larger frequency bandwidths to urine compared with water.

Summary of results concerning arousal and emotion

There are five USV parameters that change continuously with increasing stimulus intensity from experiment 1 (water) to experiment 2 (female urine) and to *head sniffing* and the *attempt to mount/mounting* interaction with the female in experiment 3. The parameters are USV duration (Fig. 7a), duration of inter-USV intervals in series of USVs (Fig. 7b), USV sound pressure level (Fig. 8) and with that the percentage of USVs with overtone/harmonics (Fig. 9), the percentage of USVs with more than one frequency jump (Fig. 4), and the USV minimum frequency (Fig. 10). These parameters are summarized in Table 1 and can be discussed (see below) in the context of increasing arousal/emotion of the animals as consequence of increasing stimulus intensity. Parameters without such a systematic change with the stimuli in the experiments are the number of emitted USVs (Fig. 5) and the percentage of USVs with at least one frequency jump and the reverse, of USVs without frequency jumps (Fig. 3).

Summary of results concerning animals without (wt) and with Foxp2 mutations (het)

A summary of the differences between wt and het animals with regard to the analyzed parameters in experiments 1–3 is shown in Table 2. The 11 listed measures all indicate quantitative, not qualitative differences, with varying degrees of statistical support. They concern the latency to the first USV, USV number, duration of USVs and inter-USV intervals, USV intensity with the dependent parameter of overtone/harmonics, USV type defined by the pattern of frequency jumps, and USV minimum frequency.

In addition to the group differences indicated in Table 2, a comparison of the measures taken in experiments 1 and 2 shows that wt animals did not differ in their responses to water and urine in (a) the latency to the first USV, (b) the number of USV series, (c) the duration of inter-USV intervals in series of USVs, and (d) the frequency bandwidth. In these four cases, animals of the het group showed significant differences with $P < 0.001$ (Figs. 1,6a,7b,11c), i.e. with urine as stimulus compared to water the latencies to the first USV were shorter, the number of USV series larger, the duration of inter-USV intervals in series of USVs shorter, and the frequency bandwidths of USVs larger.

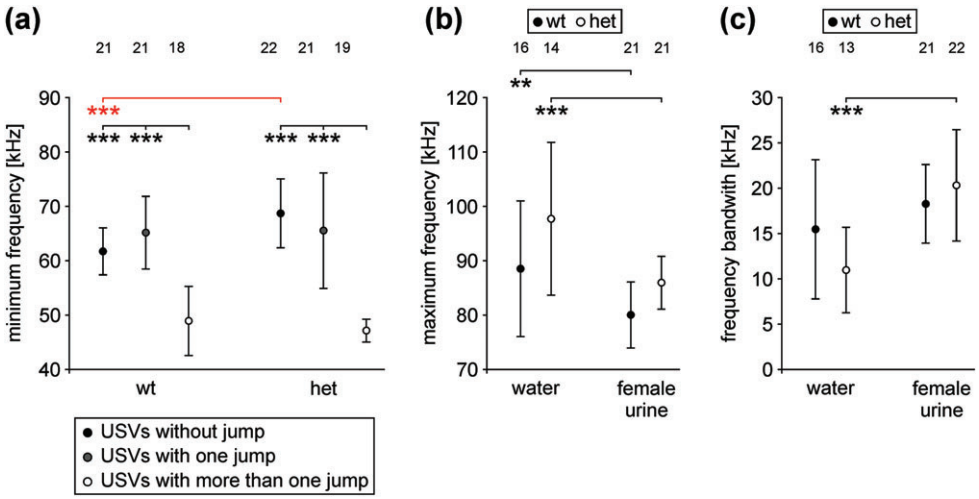

Figure 11: Minimum frequency of USVs in relation to their jump content, maximum frequency and frequency bandwidth of USVs. (a) The average minimum frequencies of the USVs (mean +/− standard deviation) in experiment 2 (female urine) are shown for both groups (wt, het) separately for USVs without frequency jump, USVs with one frequency jump, and USVs with more than one jump (compare Fig. 2). USVs with more than one frequency jump had the lowest minimum frequencies, and males of the het group emitted USVs without jump with significantly higher minimum frequencies than the animals of the wt group (a). (b) The average maximum frequencies of USVs and (c) the average frequency bandwidths (maximum minus minimum frequency) of USVs (mean +/− standard deviation) are shown for both groups and the experiments 1 (water) and 2 (female urine), respectively. Female urine led to significantly lower maximum frequencies than water (b). In the het group, female urine released significantly larger frequency bandwidths than water (c). The statistical tests in panel (a) had the Bonferroni adjustment. For further explanations see figures 1, 2.

Discussion

USV parameters expressing arousal and/or positive emotion in male mice

Forty years ago, Bell (1974) hypothesized that USVs in small rodents are arousal-produced and arousal-producing. The arousal hypothesis is supported by the fact that males emit USVs related to the *intensity of change* of their current situation, especially with regard to olfactory stimuli. Small changes, while exploring a new environment (Chabout *et al.* 2012; Hoffmann *et al.* 2012) or receiving a cotton pad with water or male urine as new stimulus, led to low rates of USVs (Guo & Holy 2007; Musolf *et al.* 2010; Nyby *et al.* 1979; Whitney & Nyby 1979; present experiment 1). USV rates were increased with more significant changes while getting in contact with female odor cues Guo & Holy 2007; Musolf *et al.* 2010; (Nyby *et al.* 1977; Sipos *et al.* 1995; present experiment 2) or being involved in sniffing interactions with an adult female (Hammerschmidt *et al.* 2012; Nyby 1983; Sales 1972; Scattoni *et al.* 2011; Zampieri *et al.* 2014; present experiment 3). Remarkably, the highest USV rates in experiment 3 of the current study were noted, especially in wts, during the first sniffing of the male after getting in touch with a female. These findings suggest that a high arousal level and, thus, a high rate of USVs is produced by the large change in the stimulus situation, i.e. the sudden presence of a female mouse. High USV rates were also reached when sniffing at the female's forehead suggesting enforced perception of stimulating pheromones of the female's facial glands (Kimoto *et al.* 2005; Luo *et al.* 2003). USV rates while attempting to mount or actually mounting the female for copulation were not increased compared to sniffing actions at female urine or body parts (except face) of the female. This again is evidence that the USV rate does not reflect the level of general excitation while doing something, but rather the strength of a *stimulus change* that is currently perceived. In accordance with this view, increasing USV rates have been observed to increasing concentration of female urine (Guo & Holy 2007) and to the change induced by *removal* of a female with which the male interacted before (Hanson & Hurley 2012). Therefore, our experiments add further evidence to state that the rate of USVs emitted and USV production at all is not necessarily related to the context of courtship behavior. The term 'courtship calls/vocalizations' for a general characterization of USVs of adult male house mice (e.g. Hanson & Hurley 2012; Nunez & Tan 1984; Nyby *et al.* 1977) is inappropriate.

Male USV production depends on the presence of male sex hormones (Nunez & Tan 1984; Warburton *et al.* 1989) and the functioning of M5 muscarinic receptors in the brain. The latter are necessary for stimulating dopaminergic neurotransmission in the nucleus accumbens for the perception of reward (Basile *et al.* 2002; Forster *et al.* 2001; Wang *et al.* 2008; Yeomans *et al.* 2000, 2001). USV emission can also indicate the anticipation of a female mouse. When female mice are odorized with various chemicals, males learn to associate the smell of the chemicals with the presence of a female and vocalize ultrasounds just to the presentation of the chemicals (review by Nyby 2010). Thus, USV production does not only characterize arousal as induced by a change of the stimulus context but also implicates a positive background state

Table 1: Summary of USV parameters that differed in relation to changes in stimulus context. Possible differences between genotypes are not considered

Increasing arousal and/or emotion lead to	increasing stimulus intensity	*P* value
.... longer duration (Fig. 7a)	from water to female urine	$P < 0.001$
	from sniffing/aimed app. to head sniffing and attempt to mount/mounting	$P < 0.001$
.... shorter inter-USV interval durations in USV series (Fig. 7b)	from water to female urine	$P < 0.001$
	from first sniffing to head sniffing and attempt to mount/mounting	$P < 0.01$ or $P < 0.001$
.... higher peak SPL (Fig. 8)	from water to female urine	$P < 0.001$
	from first sniffing to head sniffing and attempt to mount/mounting	$P < 0.01$ or $P < 0.001$
.... higher percentage of USVs with overtone/harmonics (Fig. 9)	from water to female urine	$P < 0.001$
	from first sniffing to head sniffing and attempt to mount/mounting	$P < 0.001$
.... higher percentage of USVs with more than one frequency jump (Fig. 4)	from water to female urine to sniffing interactions and attempt to mount/mounting	$P < 0.01$ or $P < 0.001$
.... lower minimum frequency (Fig. 10)	from water to female urine	$P < 0.001$
	from sniffing interactions to attempt to mount/mounting	$P < 0.001$

(i.e. being attracted by the stimulus context), which may be named 'positive emotion', when perceiving or anticipating rewarding stimuli.

Having concluded that the USV rate reflects the arousal level, we ask next which USV parameters may provide the most reliable indices of the level of positive emotion. The possible parameters are, besides USV rate, as listed in Table 1, USV duration, duration of inter-USV intervals in USV series, sound pressure level and presence of overtones/harmonics, percentage of highly complex calls (USVs with more than one frequency jump), and minimum frequency. Other studies comparing the acoustic structure of calls in low and high arousal situations of kittens (Scheumann *et al.* 2012), Weddel seal pups in interaction with their mothers (Collins *et al.* 2011) and social interactions of marmoset monkeys (Yamaguchi *et al.* 2010) found longer call durations and shorter inter-call intervals (increases of call rates) with increasing level of arousal. These changes were unrelated to a positive or negative background state of the caller assumed from the behavioral context. Hence, it seems that temporal parameters of vocalization series such as call duration and inter-call intervals determining the call rate, express levels of arousal rather than emotion. If we adopt this view for the present study, then sound pressure level and parameters expressing variability and complexity of vocalizations remain as potential indicators of positive emotion. Actually, in human speech, as in joyful laughter, emotions of joy and happiness are expressed by increased sound intensity, and increased pitch and variability and range of the voice's fundamental frequency (Johnstone & Scherer 2000; Juslin 2013; Juslin & Laukka 2001, 2003; Scherer 1981, 1989; Scherer *et al.* 1991; Sobin & Alpert 1999). Similarly in our experiments, the USVs vocalized in the contexts of head sniffing and attempt to mount/mounting may have expressed an increased level of positive emotion by increased sound pressure level and increased call complexity in the spectrotemporal domain (overtones/harmonics, frequency jumps). Compared to other situations of male–female interactions, an increased rate of complex USVs with harmonics and often with frequency jumps was also observed in males of the CBA/J mouse strain shortly before mounting a female (Hanson & Hurley 2012). Recently, Chabout *et al.* (2015) tested male USVs in response to female urine and in male–female interactions. They found amounts of USVs with frequency jumps to urine (35%) and in male–female interactions (20%) that were very similar to our values (Fig. 3), if we do not consider head sniffing and attempt to mount/mounting. Since these authors did not differentiate sniffing/mounting actions of the males during the short (5 min) interaction time in that study, they did not find the increase in the amount of USVs with more than one jump (complex calls) related to different situations as shown in Fig. 4 of the current study. Therefore, our data do not, in general, support their hypothesis that complex USVs are used to attract females at a distance and simple USVs are emitted while interacting with a female.

In summary, our present data together with comparative evidence from mice, other mammals and humans suggest that USVs of male house mice express via different parameters levels both of arousal and positive emotion. Arousal can be scaled with call rate (including the parameters call duration and inter-call intervals), positive emotion with a combination of call intensity (and the depending measure of presence of overtone/harmonics) and spectrotemporal call complexity. The latter can be expressed in the case of USVs by a scale rating the presence of simple calls, calls with one frequency jump and complex calls with two or three frequency jumps. These scales can be plotted on the arousal and valence axes of the circumplex model describing human affect (Juslin 2013; Russell 1980) suggesting that acoustic parameters of USVs of male house mice can provide insights into affective states of these animals within a given behavioral context. Since largely normal USVs are vocalized by deaf

Table 2: Summary of differences between wild-types and *Foxp2* mutants with regard to parameters of the USVs

Compared to wild-types, USVs of heterozygous mutants are, on average, different by	Stimulus	p value
... longer latency to first USV (Fig. 1)	Water	$P < 0.01$
... smaller number of USVs/s (Fig. 5)	First sniffing	$P < 0.01$
.... longer duration (Fig. 7a)	Water and female urine	$P < 0.01$ or $P < 0.001$
... longer duration of inter-USV interval (Fig. 7b)	First sniffing and aimed approach	$P < 0.01$
.... higher peak SPL (Fig. 8)	Female urine	$P < 0.001$
.... higher percentage of USVs with overtone/harmonics (Fig. 9)	Female urine	$P < 0.01$
.... higher percentage of jump type du (Fig. 2b)	Female urine	$P < 0.01$
.... higher percentage of jump type hh (Fig. 2b)	Water	$P < 0.01$
.... higher minimum frequency (Fig. 10)	Water	$P < 0.01$
... higher minimum frequency in female interactions (Fig. 10)	Mutual sniffing	$P < 0.001$
.... higher minimum frequency of USVs without jump (Fig. 11a)	Female urine	$P < 0.001$

males (Hammerschmidt *et al.* 2012; Mahrt *et al.* 2013) and males lacking the hippocampus and cerebral cortex (Hammerschmidt *et al.* 2015), USVs of mice seem to be suitable to measure arousal and positive emotion without possible influences by immediate auditory feedback and cognition.

USVs of Foxp2 heterozygous males show changes in the expression of arousal and emotion

Heterozygotes of both mouse lines emitted, as did their wild-type littermates, all defined USV types (Fig. 2) as single USVs and in USV series in the given stimulus contexts. This absence of qualitative differences in USV production between the wt and het groups leads to the conclusion that one intact *Foxp2* gene is sufficient to support the production of wt-like vocalizations that are generated by innate, arousal- and emotion-based processes. Hence, our present findings build on the analyses of vocal production in *R552H* and *S321X* mutant mouse lines and the derived conclusions from pups (Gaub *et al.* 2010) and extend them to adults. *Foxp2* has been implicated in learning motor skills and auditory-motor associations in mice (French *et al.* 2012; Groszer *et al.* 2008; Kurt *et al.* 2012), as well as song learning in songbirds (Haesler *et al.* 2007; Heston & White 2015; Teramitsu *et al.* 2010; Wohlgemuth *et al.* 2014) and language acquisition in humans (Fisher & Scharff 2009; Lai *et al.* 2001; MacDermot *et al.* 2005). Thus, functions of this gene appear to be important in the context of learning, i.e. formation and/or retrieval of memory for the control of certain motor acts. According to our data, motor acts such as vocalizations elicited by arousal and emotions in contexts of instinctive behavior appear qualitatively intact in mice carrying heterozygous *Foxp2* mutations.

Animals of the wt and het groups did not differ in the ability to produce all USV types in the appropriate situations, as single USVs and/or in series of USVs. However, it is evident from Table 2 that the het group differed from the wt group in a quantitative way with regard to several USV parameters. Compared to the wt group, heterozygous mutants started later with the first USV to water, emitted significantly smaller numbers of USVs with longer inter-USV interval to *first sniffing* interactions, significantly longer and louder USVs with a higher percentage of overtone/harmonics, USVs with higher percentages of certain types of frequency jumps, and USVs with higher minimum frequencies. In addition, USVs of heterozygous mutants expressed, on average, larger differences in response to increasing stimulus intensity from water in experiment 1 to urine in experiment 2 with regard to the parameters latency to first USV (Fig. 1), number of USV series (Fig. 6a), duration of inter-USV intervals in series of USVs (Fig. 7b), and frequency bandwidth (Fig. 11c). The USVs of wt animals to urine or water did not significantly differ in these parameters.

Given the conceptual framework of USVs as indices of arousal and positive emotion (as set out above), one potential interpretation of the quantitative differences in USV expression between the wt and het group is as follows. A stimulus context that arouses without the component of positive emotion (like the sudden presence of a cotton pad with water or the sudden presence of a female in the first sniffing interaction) seems to arouse heterozygous *Foxp2* mutants less than wts, or hets express arousal less intensely than wts; the heterozygotes take longer to start vocalizing (water) or produce fewer USVs and longer inter-USV intervals (Table 2). In a stimulus context with positive emotion, like the presence of female urine, hets express more intense changes in USVs than wt animals, i.e. the higher SPL and the higher percentage of overtone/harmonics of the USVs may indicate a higher positive emotional response to female urine in hets compared to the wts. Significant SPL differences between het and wt animals were not recorded in experiment 3, perhaps due to the fact that the USVs of the moving animals were recorded from various angles by the fixed microphones. As consequence of the high directionality of the microphone characteristics in the high ultrasonic range, SPLs were measured with a large jitter that may have masked small but systematic SPL differences of USVs. In experiment 2, the microphone and the sniffing male were in a rather fixed spatial relationship. Analyses of the complexity of USVs provide additional support for the suggestion that het animals either were in a state of higher positive emotion and/or expressed a given emotional state more intensely than the wt group when emitting USVs to female urine. The complex USV type 'du' occurred in the het group more frequently than in the wts (Table 2) and USVs with more than one frequency jump occurred in the hets of the *R552H*-line more frequently than

in the wts of this line (supplement Fig. S3). In summary, our data suggest that heterozygous *Foxp2* mutants are, compared to wts, less aroused (or express arousal less strongly) but more emotional and/or with stronger emotional expression. To follow up this hypothesis further, it would be interesting to compare our data with analyses of expressions of arousal (e.g. exclamations of surprise) and emotion (e.g. laughing, crying) in humans having the same type of *FOXP2* mutations. Such data are, however, not available. Studies reporting about emotional vocalizations in persons carrying heterozygous deletions of the *FOXP2* gene together with various numbers of other surrounding genes (review in Zilina *et al.* 2012, eTable 1) found that crys were not affected by small deletions, but laughing could be suppressed (Lai *et al.* 2000), produced with unusual effort (Zeesman *et al.* 2006), or changed to squeals and shrieks (Rice *et al.* 2012). Further, Zilina *et al.* (2012) reported about two persons with increased aggressiveness in their behavior. Therefore, it seems that heterozygous mutations in the *FOXP2* (*Foxp2*) gene can affect emotions and emotional vocal expressions in humans.

Intriguingly, the *Foxp2* gene has been shown to be expressed in the amygdala (Campbell *et al.* 2009; Ferland *et al.* 2003), which is involved in the control of emotional processes (Pessoa & Adolphs 2010; Phelps & LeDoux 2005) and cry production in mammals (Newman 2007). Little is currently known about potential functions of *Foxp2* in this structure, but one may speculate that the observed quantitative changes in the USVs of heterozygotes could be mediated by effects on the amygdala, a hypothesis that can be tested with future region-specific knockouts using conditional alleles (French *et al.* 2007).

Alternatively, the *Foxp2* mutations may have impacted on vocal variability of adult mouse USVs via effects on the basal ganglia, another region of the brain in which *Foxp2* is highly expressed (e.g. Campbell *et al.* 2009; Ferland *et al.* 2003; Groszer *et al.* 2008; Lai *et al.* 2003; Scharff & Haesler 2005; Teramitsu *et al.* 2004). In relation to this, it has been shown that knockdown of the gene in Area X of adult male zebra finches (corresponding to part of the striatum/pallidum of mammals) does not lead to qualitative changes of their song but significantly increases variability of vocalizations both in the frequency and time domain (Murugan *et al.* 2013). This may correspond to the increased complex type (du) of frequency jumps and the increased rate of overtones or harmonics in the het group of our present study.

References

Arriaga, G. & Jarvis, E.D. (2013) Mouse vocal communication system: are ultrasounds learned or innate? *Brain Lang* **124**, 96–116.

Barrett, L.F. & Russell, J.A. (1999) The structure of current affect: controversies and emerging consensus. *Curr Direct Psychol Sci* **8**, 10–14.

Basile, A.S., Fedorova, I., Zapata, A., Liu, X., Shippenberg, T., Duttaroy, A., Yamada, M. & Wess, J. (2002) Deletion of the M5 muscarinic acetylcholine receptor attenuates morphine reinforcement and withdrawal but not morphine analgesia. *Proc Natl Acad Sci U S A* **99**, 11452–11457.

Becker, J.B., Arnold, A.P., Berkley, K.J., Blaustein, J.D., Eckel, L.A., Hampson, E., Herman, J.P., Marts, S., Sadee, W., Steiner, M., Taylor, J. & Young, E. (2005) Strategies and methods for research on sex differences in brain and behavior. *Endocrinology* **146**, 1650–1673.

Bell (1974) Ultrasounds in small rodents: arousal-produced and arousal-producing. *Dev Psychobiol* **7**, 39–42.

Campbell, P., Reep, R.L., Stoll, M.L., Ophir, A.G. & Phelps, S.M. (2009) Conservation and diversity of Foxp2 expression in muroid rodents: functional implications. *J Comp Neurol* **512**, 84–100.

Chabout, J., Serreau, P., Ey, E., Bellier, L., Aubin, T., Bourgeron, T. & Granon, S. (2012) Adult male mice emit context-specific ultrasonic vocalizations that are modulated by prior isolation or group rearing environment. *PLoS One* **7**, e29401.

Chabout, J., Sarkar, A., Dunson, D.B. & Jarvis, E.D. (2015) Male mice song syntax depends on social contexts and influences female preferences. *Front Behav Neurosci* **9**, 76. DOI:10.3389/fnbeh.2015.00076.

Collins, K.T., McGreevy, P.D., Wheatley, K.E. & Harcourt, R.G. (2011) The influence of behavioural context on Weddell seal (*Leptonychotes weddellii*) airborne mother-pup vocalisation. *Behav Processes* **87**, 286–290.

Dixon, W.J. (1953) Processing data for outliers. *Biometrics* **9**, 74–89.

Ferland, R.J., Cherry, T.J., Preware, P.O., Morrisey, E.E. & Walsh, C.A. (2003) Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *J Comp Neurol* **460**, 266–279.

Fischer, J. & Hammerschmidt, K. (2011) Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication. *Genes Brain Behav* **10**, 17–27.

Fisher, S.E. & Scharff, C. (2009) FOXP2 as a molecular window into speech and language. *Trends Genet* **25**, 166–177.

Forster, G.L., Yeomans, J.S., Takeuchi, J. & Blaha, C.D. (2001) M5 muscarinic receptors are required for prolonged accumbal dopamine release after electrical stimulation of the pons in mice. *J Neurosci* **21** (RC190), 1–6.

French, C.A. & Fisher, S.E. (2014) What can mice tell us about Foxp2 function? *Curr Opin Neurobiol* **28**, 72–79.

French, C.A., Groszer, M., Preece, C., Coupe, A.M., Rajewsky, K. & Fisher, S.E. (2007) Generation of mice with a conditional Foxp2 null allele. *Genesis* **45**, 440–446.

French, C.A., Jin, X., Campbell, T.G., Gerfen, E., Groszer, M., Fisher, S.E. & Costa, R.M. (2012) An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Mol Psychiatry* **17**, 1077–1085.

Gaub, S. & Ehret, G. (2005) Grouping in temporal perception and vocal production is mutually adapted: the case of wriggling calls of mice. *J Comp Physiol A* **191**, 1131–1135.

Gaub, S., Groszer, M., Fisher, S.E. & Ehret, G. (2010) The structure of innate vocalizations in *Foxp2*-deficient mouse pups. *Genes Brain Behav* **9**, 390–401.

Grimsley, J.M.S., Monaghan, J.J.M. & Wenstrup, J.J. (2011) Development of social vocalizations in mice. *PLoS One* **6**, e17460.

Groszer, M., Keays, D.A., Deacon, R.M.J. *et al.* (2008) Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol* **18**, 354–362.

Guo, Z. & Holy, T.E. (2007) Sex selectivity of mouse ultrasonic songs. *Chem Senses* **32**, 463–473.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P. & Scharff, C. (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol* **5**, e321.

Hammerschmidt, K., Reisinger, E., Westekemper, K., Ehrenreich, L., Strenzke, N. & Fischer, J. (2012) Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neurosci* **13**, 40.

Hammerschmidt, K., Whelan, G., Eichele, G. & Fischer, J. (2015) Mice lacking the cerebral cortex develop normal song: insights into the foundations of vocal learning. *Sci Rep* **5**, 8808. DOI:10.1038/srep08808.

Hanson, J.L. & Hurley, L.M. (2012) Female presence and estrous state influence mouse ultrasonic courtship vocalizations. *PLoS One* **7**, e40782.

Heston, J.B. & White, S.A. (2015) Behavior-linked FoxP2 regulation enables zebra finch vocal learning. *J Neurosci* **35**, 2885–2894.
Hoffmann, F., Musolf, K. & Penn, D.J. (2012) Ultrasonic courtship vocalizations in wild house mice: spectrographic analyses. *J Ethol* **30**, 173–180.
Holy & Guo (2005) Ultrasonic songs of male mice. *PLoS Biol* **3**, e386.
Jamain, S., Radyushkin, K., Hammerschmidt, K., Granon, S., Boretius, S., Varoqueaux, F., Ramanantsoa, N., Gallego, J., Ronnenberg, A., Winter, D., Frahm, J., Fischer, J., Bourgeron, T., Ehrenreich, H. & Brose, N. (2008) Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci U S A* **105**, 1710–1715.
Johnstone, T. & Scherer, K.R. (2000) Vocal communication of emotion. In Lewis, M. & Haviland-Jones, J.M. (eds), *Handbook of Emotions*. Guildford Press, New York, pp. 220–235.
Juslin, P.N. (2013) Vocal affect expression: problems and promises. In Altenmüller, E., Schmidt, S. & Zimmermann, E. (eds), *Evolution of Emotional Communication*. Oxford University Press, Oxford, pp. 252–273.
Juslin, P.N. & Laukka, P. (2001) Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* **1**, 381–412.
Juslin, P.N. & Laukka, P. (2003) Communication of emotions in vocal expression and music performance – different channels, same code? *Psychol Bull* **129**, 770–814.
Kikusui, T., Nakanishi, K., Nakagawa, R., Nagasawa, M., Mogi, K. & Okanoya, K. (2011) Cross fostering experiments suggest that mice songs are innate. *PLoS One* **6**, e17721.
Kimoto, H., Haga, S., Sato, K. & Touhara, K. (2005) Sex-specific peptides from exocrine glands stimulate mouse vomeronasal sensory neurons. *Nature* **437**, 898–901.
Kurt, S., Fisher, S.E. & Ehret, G. (2012) *Foxp2* mutations impair auditory-motor association learning. *PLoS One* **7**, e33130.
Lahvis, G.P., Alleva, E. & Scattoni, M.L. (2011) Translating mouse vocalizations: prosody and frequency modulation. *Genes Brain Behav* **10**, 4–16.
Lai, C.S.L., Fisher, S.E., Hurst, J.A., Levy, E.R., Hodgson, S., Fox, M., Jeremiah, S., Povey, S., Jamison, D.C., Green, E.D., Vargha-Khadem, F. & Monaco, A.P. (2000) The *SPCH1* region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am J Hum Genet* **67**, 357–368.
Lai, C.S.L., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F. & Monaco, A.P. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523.
Lai, C.S.L., Gerrelli, D., Monaco, A.P., Fisher, S.E. & Copp, A.J. (2003) FOXP' expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* **126**, 2455–2462.
Luo, M., Fee, M.S. & Katz, L.C. (2003) Encoding pheromonal signals in the accessory olfactory bulb of behaving mice. *Science* **299**, 1196–1201.
MacDermot, K.D., Bonora, E., Sykes, N., Coupe, A.-M., Lai, C.S.L., Vernes, S.C., Vargha-Khadem, F., McKenzie, F., Smith, R.L., Monaco, A.P. & Fisher, S.E. (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet* **76**, 1074–1080.
Mahrt, E.J., Perkel, D.J., Tong, L., Rubel, E.W. & Portfors, C.V. (2013) Engineered deafness reveals that mouse courtship vocalizations do not require auditory experience. *J Neurosci* **33**, 5573–5583.
Moles, A., Costantini, F., Garbugino, L., Zanettini, C. & D'Amato, F.R. (2007) Ultrasonic vocalizations emitted during dyadic interactions in female mice: a possible index of sociability? *Behav Brain Res* **182**, 223–230.
Murugan, M., Harward, S., Scharff, C. & Mooney, R. (2013) Dimished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron* **80**, 1464–1476.
Musolf, K., Hoffmann, F. & Penn, D.J. (2010) Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*. *Anim Behav* **79**, 757–764.
Newman, J.D. (2007) Neural circuits underlying crying and cry responding in mammals. *Behav Brain Res* **182**, 155–165.
Nunez, A.A. & Tan, D.T. (1984) Courtship ultrasonic vocalizations in male Swiss-Webster mice: effects of hormones and sexual experience. *Physiol Behav* **32**, 717–721.
Nyby, J. (1983) Ultrasonic vocalizations during sex behavior of male house mice (*Mus musculus*): a description. *Behav Neural Biol* **39**, 128–134.
Nyby, J.G. (2010) Adult house mouse (Mus musculus) ultrasonic calls: hormonal and pheromonal regulation. In Brudzynski, S.M. (ed), *Handbook of Mammalian Vocalization. An Integrative Neuroscience Approach*. Academic Press, Oxford, pp. 303–310.
Nyby, J., Wysocki, C.J., Whitney, G. & Dizinno, G. (1977) Pheromonal regulation of male mouse ultrasonic courtship (*Mus musculus*). *Anim Behav* **25**, 333–341.
Nyby, J., Wysocki, C.J., Whitney, G., Dizinno, G. & Schneider, J. (1979) Elicitation of male mouse (Mus musculus) ultrasonic vocalizations: I. Urinary cues. *J Comp Physiol Psychol* **93**, 957–975.
Pessoa, L. & Adolphs, R. (2010) Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nat Rev Neurosci* **11**, 773–783.
Phelps, E.A. & LeDoux, J.E. (2005) Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**, 175–187.
Portfors, C.V. & Perkel, D.J. (2014) The role of ultrasonic vocalizations in mouse communication. *Curr Opin Neurobiol* **28**, 115–120.
Rice, G.M., Raca, G., Jakielski, K.J., Laffin, J.J., Iyama-Kurtycz, C.M., Hartley, S.L., Sprague, R.E., Heintzelman, A.T. & Shriberg, L.D. (2012) Phenotype of *FOXP2* haploinsufficiency in a mother and son. *Am J Med Genet* **158A**, 174–181.
Rotschafer, S.E., Trujillo, M.S., Dansie, L.E., Ethell, I.M. & Razak, K.A. (2012) Minocycline treatment reverses ultrasonic vocalization production deficit in a mouse model of Fragile X Syndrome. *Brain Res* **1439**, 7–14.
Russell, J.A. (1980) A circumplex model of affect. *J Personality Social Psychol* **39**, 1161–1178.
Sachs, L. (1999) *Angewandte Statistik*. Springer-Verlag, Berlin.
Sales, G.D. (1972) Ultrasound and mating behaviour in rodents with some observations on other behavioural situations. *J Zool Lond* **168**, 149–164.
Scattoni, M.L. (2011) Special interest section on mouse ultrasonic vocalizations. *Genes Brain Behav* **10**, 1–3.
Scattoni, M.L., Ricceri, L. & Crawley, J.N. (2011) Unusual repertoire of vocalizations in adult BTBR T+tf/J mice during three types of social encounters. *Genes Brain Behav* **10**, 44–56.
Scharff, C. & Haesler, S. (2005) An evolutionary perspective on FoxP2: strictly for the birds? *Curr Opin Neurobiol* **15**, 694–703.
Scherer, K.R. (1981) Speech and emotional states. In Darby, J.K. (ed), *Speech Evaluation in Psychiatry*. Grune and Stratton, New York, pp. 189–220.
Scherer, K.R. (1989) Vocal correlates of emotional arousal and affective disturbance. In Wagner, H. & Manstead, A. (eds), *Handbook of Social Psychophysiology*. Wiley, New York, pp. 165–197.
Scherer, K.R., Banse, R., Wallbott, H.G. & Goldbeck, T. (1991) Vocal cues in emotion encoding and decoding. *Motiv Emotion* **15**, 123–148.
Scheumann, M., Roser, A.-E., Konerding, W., Bleich, E., Hedrich, H.-J. & Zimmermann, E. (2012) Vocal correlates of sender-identity and arousal in the isolation calls of domestic kitten (*Felis silvestris catus*). *Front Zool* **9**, 36.
Sipos, M.L., Wysocki, C.J., Nyby, J.G., Wysocki, L. & Nemura, T.A. (1995) An ephemeral pheromone of female house mice: perception via the main and accessory olfactory systems. *Physiol Behav* **58**, 529–534.
Sobin, C. & Alpert, M. (1999) Emotion in speech: the acoustic attributes of fear, anger, sadness and joy. *J Psycholinguist Res* **28**, 347–365.

Teramitsu, I., Kudo, L.C., London, S.E., Geschwind, D.H. & White, S.A. (2004) Parallel *FoxP1* and *FoxP2* expression in songbird and human brain predicts functional interaction. *J Neurosci* **24**, 3152–3163.
Teramitsu, I., Poopatanapong, A., Torrisi, S. & White, S.A. (2010) Striatal *FoxP2* is actively regulated during songbird sensorimotor learning. *PLoS One* **5**, e8548.
Van der Gulden, W.J.I., van Hooijdonk, C.L., de Jong, P. & Kremer, A.K. (1975) *Versuchstiere und Versuchstierkunde*. Vol. **2**. Ges. f. Versuchstierkunde, Basel.
Vernes, S.C. & Fisher, S.E. (2009) Unravelling neurogenetic networks implicated in developmental language disorders. *Biochem Soc Trans* **37**, 1263–1269.
Vernes, S.C., Nicod, J., Elahi, F.M., Coventry, J.A., Kenny, N., Coupe, A.-M., Bird, L.E., Davies, K.E. & Fisher, S.E. (2006) Functional genetic analysis of mutations implicated in a human speech and language disorder. *Hum Mol Genet* **15**, 3154–3167.
Wang, H., Liang, S., Burgdorf, J., Wess, J. & Yeomans, J. (2008) Ultrasonic vocalizations induced by sex and amphetamine in M2, M4, M5 muscarinic and D2 dopamine receptor knockout mice. *PLoS One* **3**, e1893.
Warburton, V.L., Sales, G.D. & Milligan, S.R. (1989) The emission and elicitation of mouse ultrasonic vocalizations: effects of age, sex and gonadal status. *Physiol Behav* **45**, 41–47.
Watkins, K.E., Dronkers, N.F. & Vargha-Khadem, F. (2002) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* **125**, 452–464.
White, N.R., Prasad, M., Barfield, R.J. & Nyby, J.G. (1998) 40- and 70-kHz vocalizations of mice (*Mus musculus*) during copulation. *Physiol Behav* **63**, 467–473.
Whitney, G. & Nyby, J. (1979) Cues that elicit ultrasounds from adult male mice. *Am Zool* **19**, 457–463.
Woehr, M., Roullet, F.I. & Crawley, J.N. (2010) Reduced scent marking and ultrasonic vocalizations in the BTBR T+tf/J mouse model of autism. *Genes Brain Behav* **10**, 35–43.
Wohlgemuth, S., Adam, I. & Scharff, C. (2014) FoxP2 in songbirds. *Curr Opin Neurobiol* **28**, 86–93.
Yamaguchi, C., Izumi, A. & Nakamura, K. (2010) Time course of vocal modulation during isolation in common marmosets (*Callithrix jacchus*). *Am J Primatol* **72**, 681–688.
Yeomans, J.S., Takeuchi, J., Baptista, M., Flynn, D.D., Lepik, K., Nobrega, J., Fulton, J. & Ralph, M.R. (2000) Brain-stimulation reward thresholds raised by an antisense oligonucleotide for the M5 muscarinic receptor infused near dopamine cells. *J Neurosci* **20**, 8861–8867.
Yeomans, J., Forster, G. & Blaha, C. (2001) M5 muscarinic receptors are needed for slow activation of dopamine neurons and for rewarding brain stimulation. *Life Sci* **68**, 2449–2456.
Zampieri, B.L., Fernandez, F., Pearson, J.N., Stasko, M.R. & Costa, A.C. (2014) Ultrasonic vocalizations during male–female interaction in the mouse model of Down syndrome Ts65Dn. *Physiol Behav* **128**, 119–125.
Zeesman, S., Nowaczyk, M.J.M., Teshima, I., Roberts, W., Cardy, J.O., Brian, J., Senman, L., Feuk, L., Osborne, L.R. & Scherer, S.W. (2006) Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves *FOXP2*. *Am J Med Genet* **140A**, 509–514.
Zilina, O., Reimand, T., Zjablovskaja, P., Männik, K., Männamaa, M., Traat, A., Puusepp-Benazzouz, H., Kurg, A. & Õunap, K. (2012) Maternally and paternally inherited deletion of 7q31 involving the *FOXP2* gene in two families. *Am J Med Genet* **158A**, 254–256.

Acknowledgments

We thank Rathi Puliydai for assistance with genotyping, Sabine Schmidt for analyzing part of the vocalizations, and Matthias Groszer for help with initiating the study. This study was supported by the Deutsche Forschungsgemeinschaft (Eh 53/19-1,2).

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Figure S1: Latency to first USV occurrence. The latencies to the first USV occurrence in experiment 1 (water) and experiment 2 (female urine) are shown as means ± SD or as boxplots (median, quartiles, range; in cases when data were not normally distributed) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line. Heterozygous animals of the *R552H*-line responded significantly faster to female urine compared to water. For further explanations see Fig. 1.

Figure S2: Percentage of USVs with frequency jump(s). The average percentages of USVs (mean ± SD) with frequency jump(s), i.e. USVs with at least one frequency jump, are shown for the wt and het groups of each *Foxp2* mouse line in experiment 1 (water) and experiment 2 (urine). In the het group of the *R552H*-line urine released significantly more USVs with frequency jump(s) than water. For further explanations see Fig. 1.

Figure S3: Percentage of USVs with more than one frequency jump. The average percentages of USVs with more than one frequency jump are shown as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to higher rates of USVs with more than one frequency jump than water. For further explanations see Figs. 1, S1.

Figure S4: Number of USVs per second. The average numbers of USVs per second are shown as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine released significantly more USVs than water. For further explanations see Figs. 1,S1.

Figure S5: Number of USV series. The average numbers of USV series (mean ± SD) emitted during the first minute of stimulus presence in experiment 1 (water) and experiment 2 (urine) are shown for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line. In both het groups significantly more USV series occurred in response to urine than in response to water. For further explanations see Fig. 1.

Figure S6: Number of USVs in a USV series. The average numbers of USVs in a USV series are shown as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. For further explanations see Figs. 1, S1.

Figure S7: Duration of USVs. The average durations of USVs (mean ± SD) are shown for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to longer USVs than water in the *R552H* het group. The hets of this line produced significantly longer USVs in response to female urine than the wts of the same line. For further explanations see Fig. 1.

Figure S8: Duration of inter-USV intervals in series of USVs. The average durations of inter-USV intervals in USV series are shown as means ± SD or as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations The *R552H* het group and both genotypes of

the *S321X*-line produced to urine USV series with shorter inter-USV intervals than in response to water. For further explanations see Figs. 1,S1.

Figure S9: Peak sound pressure level (dB SPL) of USVs. The average peak SPLs of USVs are shown as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to significantly louder USVs than water in each genotype group. Statistical comparisons had Bonferroni adjustment. For further explanations see Figs. 1,S1.

Figure S10: Percentage of USVs with overtones or harmonics. The average percentages of USVs with overtones or harmonics are shown as boxplots (median, quartiles, range) for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to significantly more USVs with overtone/harmonics than water in both het groups of the *Foxp2* mouse lines. Statistical comparisons had Bonferroni adjustment. For further explanations see Figs. 1,S1.

Figure S11: Peak sound pressure level (dB SPL) of USVs without and with overtone/harmonics. The average peak SPLs of USVs (mean ± SD) in experiment 2 (female urine) are shown for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line separately for USVs without or with overtone/harmonic(s). In all genotypes, USVs with overtones or harmonics were about 5 dB louder than USVs without overtone or harmonics. For further explanations see Fig. 1.

Figure S12: Minimum frequency of USVs. The average minimum frequencies of USVs (mean ± SD) are shown for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to significantly lower minimum frequencies than water. For further explanations see Fig. 1.

Figure S13: Maximum frequency of USVs. The average maximum frequencies of USVs are shown as means ± SD or as boxplots (median, quartiles, range) wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the indicated experimental situations. Female urine led to lower maximum frequencies than water in the wt group of *R552H*-line and in both groups of the *S321X*-line. For further explanations see Figs. 1,S1.

Figure S14: Frequency bandwidth of USVs. The average frequency bandwidths (maximum minus minimum frequency) of USVs (mean ± SD) are shown for wild-types (wt) and heterozygotes (het) of the *R552H*- and *S321X*-line and the experiments 1 (water) and 2 (female urine). Female urine led to larger frequency bandwidths than water in the het group of *R552H*-line and in both groups of the *S321X*-line. For further explanations see Fig. 1.

Table S1: Statistical tests used to analyze data shown in Figs. 1–11.

Table S2: Statistical tests used to analyze data shown in Figs. S1–S14.