

# Engineered Deafness Reveals That Mouse Courtship Vocalizations Do Not Require Auditory Experience

Elena J. Mahrt,<sup>1</sup> David J. Perkel,<sup>2,3,4</sup> Ling Tong,<sup>3,4</sup> Edwin W. Rubel,<sup>3,4,5</sup> and Christine V. Portfors<sup>1</sup>

<sup>1</sup>School of Biological Sciences, Washington State University, Vancouver, Washington 98686, and <sup>2</sup>Department of Biology, <sup>3</sup>Department of Otolaryngology-Head and Neck Surgery, <sup>4</sup>Virginia Merrill Bloedel Hearing Research Center, and <sup>5</sup>Department of Physiology and Biophysics, University of Washington, Seattle, Washington 98195

Auditory experience during development is necessary for normal language acquisition in humans. Although songbirds, some cetaceans, and maybe bats may also be vocal learners, vocal learning has yet to be well established for a laboratory mammal. Mice are potentially an excellent model organism for studying mechanisms underlying vocal communication. Mice vocalize in different social contexts, yet whether they learn their vocalizations remains unresolved. To address this question, we compared ultrasonic courtship vocalizations emitted by chronically deaf and normal hearing adult male mice. We deafened CBA/CaJ male mice, engineered to express diphtheria toxin (DT) receptors in hair cells, by systemic injection of DT at postnatal day 2 (P2). By P9, almost all inner hair cells were absent and by P16 all inner and outer hair cells were absent in *DTR* mice. These mice did not show any auditory brainstem responses as adults. Wild-type littermates, also treated with DT at P2, had normal hair cells and normal auditory brainstem responses. We compared the temporal structure of vocalization bouts, the types of vocalizations, the patterns of syllables, and the acoustic features of each syllable type emitted by hearing and deaf males in the presence of a female. We found that almost all of the vocalization features we examined were similar in hearing and deaf animals. These findings indicate that mice do not need auditory experience during development to produce normal ultrasonic vocalizations in adulthood. We conclude that mouse courtship vocalizations are not acquired through auditory feedback-dependent learning.

## Introduction

In learning to speak, humans copy the sounds made by others through imitative vocal learning. Few other mammals show evidence of vocal learning (cetaceans, Deecke et al., 2000; bats, Knoernschild et al., 2010). In contrast, thousands of songbird species are vocal learners (Konishi, 1965; Price, 1979; Nowicki and Marler, 1988; Tchernichovski et al., 1999), and they have become the model system for mechanistic studies of this process.

However, songbirds have limitations as a model for mammalian vocal learning. Vocal learning likely evolved independently in these taxa, and genetic manipulations are not yet routine in songbirds. An experimentally accessible mammal with vocal learning would allow direct study of mechanisms underlying human vocal learning and associated disorders.

Mice are a candidate organism for studying mechanisms underlying vocal communication. They emit ultrasonic social vocalizations (D'Amato and Moles, 2001; Holy and Guo, 2005;

Portfors, 2007), and the genetic basis of these signals can provide insight into human communication disorders (Enard et al., 2009; Scattoni et al., 2009; Wöhr et al., 2011; Schmeisser et al., 2012). Whether mice learn their vocalizations, however, is unresolved.

Several experimental approaches can test for vocal learning, including rearing in isolation, artificial tutoring, cross-fostering, and deaf-rearing. Only two of these approaches have been applied to mice. Kikusui et al. (2011) cross-fostered two mouse strains and showed that the animal's adult vocalizations resembled those of its own genetic strain. Because inbred mouse strains can have compromised high-frequency hearing (Henry and Lepkowski, 1978; Zheng et al., 1999), negative cross-fostering results may not indicate a lack of vocal learning. Positive results could reflect altered social interactions rather than true imitation. A more sensitive approach to detect learning is deprivation of auditory experience (Konishi, 1965).

Two studies have come to opposite conclusions using gene knock-outs expected to induce hearing loss. *Otoferlin* knock-out mice (Hammerschmidt et al., 2012), with disrupted synaptic transmission by inner hair cells, were found to emit adult vocalizations with normal acoustic features. In contrast, *caspase-3* knock-out mice, with some loss of hair cells by postnatal week 5, emitted abnormal adult vocalizations (Arriaga et al., 2012). Two potentially important limitations of these mouse strains could help explain this discrepancy. First, it is unclear how much auditory experience each mouse strain received. Hearing has not been assessed in *Otoferlin* knock-out animals before postnatal (P) day 30. *Caspase-3* knock-out mice have residual hearing up to 5 weeks

Received Oct. 29, 2012; revised Jan. 4, 2013; accepted Jan. 15, 2013.

Author contributions: E.J.M., D.J.P., E.W.R., and C.V.P. designed research; E.J.M. and L.T. performed research; E.J.M., D.J.P., L.T., and C.V.P. analyzed data; E.J.M., D.J.P., E.W.R., and C.V.P. wrote the paper.

This work was supported by National Science Foundation Grant IOS-0920060 to C.V.P. and National Institute for Deafness and Communication Disorders Grants DC03829 and DC04661 to E.W.R. We thank Richard Palmiter for development of the mouse, Roian Egnor for MUSCat software, Amy Boyle for programming support, Linda Robinson for mouse husbandry, and the two anonymous reviewers for valuable comments that improved this manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Christine V. Portfors, Washington State University, 14204 NE Salmon Creek Avenue, Vancouver, WA 98686. E-mail: portfors@vancouver.wsu.edu.

DOI:10.1523/JNEUROSCI.5054-12.2013

Copyright © 2013 the authors 0270-6474/13/335573-11\$15.00/0

of age. In addition, *caspace-3* knock-out mice have abnormal brain morphology (Kuida et al., 1996; Takahashi et al., 2001); this phenotype alone, independent of any hearing loss, could cause abnormal vocalizations.

To overcome these limitations, we used transgenic mice that allowed us to prevent all auditory experience. These mice were engineered to express diphtheria toxin (DT) receptors (DTRs) in hair cells. Injection of DT at P2 led to complete hair cell death before the onset of hearing (Tong et al., 2011; Golub et al., 2012). We were thus able to compare temporal and acoustic features of vocalizations made by deaf and hearing adult male mice.

## Materials and Methods

### Animals

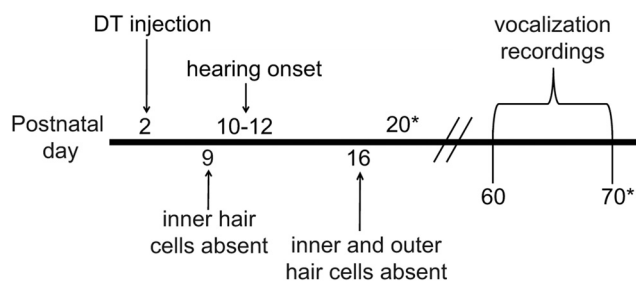
To compare vocalizations emitted by hearing and deaf male mice, we used a mouse line that expresses DTRs in hair cells to induce deafness. CBA/CaJ mice had the human *dtr* gene inserted into the *pou4f3* gene (Golub et al., 2012). *Pou4f3<sup>+DTR</sup>* mice express DTRs on all hair cells, whereas *Pou4f3<sup>+/+</sup>* (WT) mice do not. We used the CBA/CaJ strain because it has good hearing up to at least one year (Willott, 1986; Willott, 1991; Zheng et al., 1999; Zheng and Johnson, 2001). We gave 49 P2 mice a systemic injection of DT (4 ng/g i.m.; List Biological Laboratories) (Fig. 1). Pups were raised with their mothers until weaning at P21. After weaning, pups were group housed in same sex, mixed genotype cages. To identify individual animals throughout the course of the experiment, we placed Sharpie markings on the pups' paws and then cut unique patterns in their fur as adults. Genotyping for *Pou4f3* of all mice was performed at P18 by tail clippings and a DNeasy Blood & Tissue Kit C250 (QIAGEN).

All animal care and experimental procedures followed the guidelines of the National Institutes of Health and were approved by the University of Washington Institutional Animal Care and Use Committee (protocol 2048-02).

### Assessing the effectiveness of DT injections

**Cochlear whole mounts.** We examined the extent of inner and outer hair cell loss in *Pou4f3<sup>+DTR</sup>* and *Pou4f3<sup>+/+</sup>* mice at P9 (7 d after DT injection) and P16 (14 d after DT injection). Mice were killed, and the temporal bones were dissected free. After removal of the bulla, the stapes was lifted from the oval window, the membrane was removed from the round window, and a small opening was made in the apical turn. Cold 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4, was perfused slowly through the cochlea from the opening of the apex turn, after which the temporal bones were kept in the same fixative for 2 h at room temperature. After fixation, the temporal bones were washed three times (10 min each) in PBS, pH 7.4. The tissue was prepared as a whole-mount preparation. The cochlea segments of the organ of Corti were carefully dissected free from the cochlea. The stria vascularis was removed or trimmed, and the tectorial membrane was removed with forceps.

We used two antibodies to label hair cells: a mouse monoclonal anti-parvalbumin antibody (catalog #MAB 1572, Millipore, 1:1000) and a rabbit anti-Myosin6 (catalog #25-6791, Proteus Bioscience, 1:500). To label supporting cells, we used a goat anti-Sox2 (catalog #SC-17320, Santa Cruz Biotechnology, 1:500).



**Figure 1.** Experimental timeline. Injection of DT was given at P2. The majority of inner hair cells were gone by P9, and all cochlear hair cells were eliminated by P16 in *Pou4f3<sup>+DTR</sup>* mice. \*Auditory brainstem responses obtained.

The tissue was permeabilized for 30 min with 0.1% saponin/0.1% Tween 20 in PBS. To prevent nonspecific binding of the primary antibody, we incubated the tissue for 1 h in a blocking solution consisting of 5% normal serum/0.1% Triton X-100 in PBS. Primary antibody incubations were performed for 1 d at 4 deg C in PBS, 5% serum and 0.1% Triton X-100. We used fluorescent-labeled secondary antibodies (Alexa-488, Alexa-568, Invitrogen) at a dilution of 1:400 in the same buffer for 2 h at room temperature. For mouse antibodies against parvalbumin, we used the Mouse-on-Mouse kit (catalog #BMK2202) as specified by the manufacturer (Vector Laboratories). Sections were washed after each antibody incubation (three times, 10–15 min each) in 0.1% Tween 20 in PBS. After counterstaining nuclei with DAPI (catalog #D9542, Sigma-Aldrich, 1 μg/ml), the specimens were mounted in Vectashield (Vector Laboratories), coverslipped, and examined with confocal fluorescence microscopy.

**Table 1. Description of syllable categories**

Syllable type	Spectrographic representation	Criteria
One jump		One frequency step of at least 2 kHz with <5 ms separation
Two jump		Two frequency steps of at least 2 kHz with <5 ms separation
Three jump		Three frequency steps of at least 2 kHz with <5 ms separation
Four jump		Four frequency steps of at least 2 kHz with <5 ms separation
Five jump		Five frequency steps of at least 2 kHz with <5 ms separation
FM upsweep		Upward change in frequency >6 kHz
FM downsweep		Downward change in frequency >6 kHz
Reverse chevron		Downsweep followed by an upsweep; total bandwidth >6 kHz
Chevron		Upsweep followed by a downsweep; total bandwidth >6 kHz
Complex		Three or more phases with frequency modulation of >6 kHz
Constant frequency		Frequency did not change by >6 kHz
Short		Duration <5 ms

**Table 2. All measured acoustic parameters in 12 syllable categories for hearing mice<sup>a</sup>**

Parameter	Location	One jump	Two jump	Three jump	Four jump	Five jump	FM upsweep	FM downsweep	Reverse chevron	Chevron	Complex	Constant frequency	Short
Duration (ms)	Line	40.1 (7.6)	69.7 (12.9)	66.1 (22.9)	92.4 (29.1)	69.9 (20.3)	28.8 (5.8)	39.8 (12.9)	33.9 (4.2)	49.6 (6.8)	58.5 (11.6)	30.7 (8.2)	4.1 (0.4)
Segment 1 duration (ms)	Line	19.4 (3.7)	20.4 (5.0)	10.9 (3.5)	19.5 (10.2)	7.0 (1.6)			17.0 (4.8)	22.9 (3.5)			
Segment 2 duration (ms)	Line	19.8 (7.7)	21.3 (6.6)	17.3 (5.1)	14.2 (5.7)	10.5 (6.2)			16.9 (3.6)	26.7 (4.9)			
Segment 3 duration (ms)	Line		25.4 (7.7)	18.1 (15.6)	11.0 (0.9)	10.6 (3.5)							
Segment 4 duration (ms)	Line			15.6 (8.5)		14.1 (5.6)							
Segment 5 duration (ms)	Line				29.1 (26.7)	15.0 (9.5)							
Segment 6 duration (ms)	Line					11.8 (6.4)							
Gap 1 (ms)	Line	0.9 (0.4)	1.5 (1.0)	1.5 (1.0)	1.4 (0.8)	0.6 (0.2)							
Gap 2 (ms)	Line		1.2 (0.9)	1.8 (1.2)	1.3 (1.0)	1.5 (1.3)							
Gap 3 (ms)	Line			0.8 (0.3)	0.9 (0.7)	1.9 (1.8)							
Gap 4 (ms)	Line				0.9 (0.6)	0.6 (0.2)							
Gap 5 (ms)	Line					0.5 (0.0)							
Peak time (ms)	Fcontour								17.0 (4.8)	22.9 (3.5)			
Start f1 (kHz)	Line	81.6 (7.4)	72.4 (7.2)	87.1 (6.3)	75.3 (7.1)	83.8 (6.9)	70.1 (5.7)	87.4 (8.5)	83.0 (8.2)	74.1 (8.5)	76.1 (8.4)	79.3 (6.1)	79.4 (5.0)
Start f2 (kHz)	Line	79.9 (6.7)	90.8 (9.7)	77.4 (9.8)	89.5 (7.5)	79.7 (14.2)			74.3 (6.6)	91.6 (7.8)			
Start f3 (kHz)	Line		76.5 (6.2)	97.2 (15.1)	87.5 (14.6)	100.8 (12.9)							
Start f4 (kHz)	Line			78.6 (5.8)	83.6 (6.6)	76.9 (9.5)							
Start f5 (kHz)	Line				76.9 (8.7)	96.0 (9.6)							
Start f6 (kHz)	Line					68.4 (8.0)							
End f (kHz)	Line						89.9 (5.8)	73.6 (6.3)			79.1 (6.0)	80.1 (5.3)	
End f1 (kHz)	Line	88.5 (8.4)	84.0 (9.3)	95.5 (7.5)	89.1 (5.8)	89.1 (4.2)			74.3 (6.6)	91.6 (7.8)			79.4 (4.9)
End f2 (kHz)	Line	77.8 (7.0)	91.5 (9.9)	82.5 (8.6)	87.6 (3.8)	84.7 (9.2)			83.6 (6.3)	74.0 (5.8)			
End f3 (kHz)	Line		77.3 (6.5)	91.3 (16.5)	86.2 (14.0)	98.1 (16.1)							
End f4 (kHz)	Line			72.5 (6.7)	79.9 (8.2)	78.3 (8.3)							
End f5 (kHz)	Line				71.0 (7.8)	86.3 (10.8)							
End f6 (kHz)	Line					70.0 (10.7)							
Minimum f (kHz)	Fcontour	69.5 (5.1)	60.2 (7.3)	64.3 (10.0)	60.6 (7.7)	63.3 (5.7)	69.5 (5.7)	72.1 (7.3)	74.4 (6.6)	69.4 (6.1)	70.3 (6.9)	76.4 (5.6)	78.4 (4.8)
Maximum f (kHz)	Fcontour	95.2 (7.9)	101.9 (3.7)	108.4 (7.4)	107.9 (4.9)	103.3 (13.2)	89.2 (6.2)	87.7 (8.2)	87.2 (8.0)	91.1 (7.4)	90.1 (7.9)	82.6 (5.8)	82.8 (4.7)
Peak f (kHz)	Line								74.3 (6.6)	91.6 (7.8)			
Mean f (kHz)	Fcontour	81.7 (5.9)	77.8 (6.6)	82.4 (9.5)	80.5 (5.2)	85.0 (7.6)	80.1 (5.5)	80.4 (7.3)	79.3 (6.5)	82.0 (7.0)	80.3 (7.4)	79.7 (5.6)	80.0 (4.8)
Bandwidth (kHz)	Fcontour	25.7 (6.2)	41.6 (5.2)	44.2 (8.1)	47.3 (5.9)	40.0 (10.0)	19.7 (2.6)	15.6 (2.6)	12.8 (3.1)	21.7 (3.5)	19.8 (2.9)	6.2 (1.1)	
Δ f1 (kHz)	Fcontour	17.7 (4.5)	28.0 (7.4)	29.5 (7.5)	29.6 (7.1)	24.6 (4.3)							
Δ f2 (kHz)	Fcontour		20.5 (5.0)	25.8 (7.4)	25.1 (6.0)	20.4 (5.6)							
Δ f3 (kHz)	Fcontour			17.6 (1.6)	19.7 (2.8)	17.3 (4.6)							
Δ f4 (kHz)	Fcontour				15.7 (1.2)	15.3 (5.3)							
Δ f5 (kHz)	Fcontour					11.5 (5.7)							
FM amplitude (kHz)	Line										10.6 (3.9)		
F offset (kHz)	Line										73.6 (6.1)		
FM frequency (Hz)	Line										12.2 (3.1)		
F initial phase (°)	Line										−5.4 (15.6)		
Segment slope 1 (kHz/s)	Line	0.4 (0.4)	0.7 (0.3)	0.9 (0.7)	0.9 (0.4)	1.0 (0.8)	0.8 (0.1)	−0.4 (0.2)	−0.6 (0.2)	0.8 (0.2)		0.0 (0.1)	
Segment slope 2 (kHz/s)	Line	−0.1 (0.2)	−0.3 (0.4)	0.5 (0.4)	−0.4 (0.7)	0.4 (0.5)			0.6 (0.2)	−0.8 (0.2)			
Segment slope 3 (kHz/s)	Line		−0.7 (0.5)	−0.4 (0.7)	−0.2 (0.3)	−0.2 (0.4)							
Segment slope 4 (kHz/s)	Line			−1.3 (1.8)	−0.4 (0.3)	0.1 (0.3)							
Segment slope 5 (kHz/s)	Line				−0.5 (0.3)	−0.8 (0.7)							
Segment slope 6 (kHz/s)	Line					−0.2 (0.7)							
Peak time shift ABS (ms)	Both								0.9 (0.7)	0.3 (0.3)			
Peak time shift (ms)	Both								0.6 (1.1)	0.2 (0.4)			
Residual (kHz)	Line	1.2 (0.4)	1.4 (0.2)	1.3 (0.3)	1.1 (0.2)	1.0 (0.5)	1.4 (0.2)	1.5 (0.4)	0.9 (0.2)	1.3 (0.2)	1.7 (0.2)	0.9 (0.2)	

<sup>a</sup>Data are mean (SD). ABS, Absolute value.

We viewed whole-mount preparations on an IX-81 inverted microscope (Olympus) integrated into an FV-1000 laser scanning confocal microscope (Olympus). We collected the images with a 10×/1.40 NA UPLSAPO objective or a 100×/1.40 NA UPLSAPO oil-immersion objective. The fluorescent labels were excited with a 405 nm laser diode, 488 nm argon ion laser, and a 561 nm diode pumped solid state laser. We collected the images with a 4-channel dichroic mirror (blue/green/red/far red), a 490 nm long pass dichroic, and a 425–475 nm diffraction filter setting on channel 1, a 560 nm long pass dichroic mirror and 500–550 nm diffracting setting on channel 2 with a 585–655 nm emission filter on channel 3. Sequential image acquisition was performed to avoid bleed-through using Fluoview software, version 1.3a. We imported images into ImageJ 1.42a (National Institutes of Health) to create maximum intensity projections from z-series stacks, which were saved as 24-bit RGB TIFFs. The figures we created with Adobe Photoshop CS version 8 (Adobe) were subjected to histogram stretch and  $\gamma$  adjustment to fill the dynamic range and compensate for printing.

**Auditory brainstem recordings.** To confirm that the *DTR* males were functionally deaf and that the WT males had normal hearing, we measured auditory brainstem responses (ABRs) of all animals after the vocalization recordings. In addition, we measured the ABRs of 11 mice at P20 to ensure that DT injections deafened animals. Mice were anesthetized (ketamine, 100 mg/kg; xylazine, 5 mg/kg, i.p.), placed on a heating pad to maintain body temperature near 37°C, and placed in a sound-attenuating chamber. We recorded ABRs using standard subcutaneous needle electrodes with the positive electrode at the vertex of the skull and the reference electrode in the ipsilateral thigh. Sound stimuli were generated and ABR recordings digitized using custom software. Responses were preamplified (100×; Grass Technologies P15 amplifier), sent through an MA3 amplifier with an additional 20 dB post-preamp gain (Tucker Davis Technologies), bandpass filtered (100–3000 Hz; Krohn-Hite filter model 3550), and digitized at 24.4 kHz. We sampled responses for a 15 ms window (with a 5 ms stimulus onset delay). The threshold was

**Table 3. All measured acoustic parameters in 12 syllable categories for deaf mice<sup>a</sup>**

Parameter	Location	One jump	Two jump	Three jump	Four jump	Five jump	FM upsweep	FM downsweep	Reverse chevron	Chevron	Complex	Constant frequency	Short
Duration (ms)	Line	43.4 (6.7)	54.4 (15.6)	64.7 (8.8)	69.6 (16.5)	77.4 (19.9)	30.7 (3.8)	38.8 (10.9)	34.9 (7.5)	52.6 (6.5)	63.1 (11.8)	33.3 (8.6)	3.7 (0.7)
Segment 1 duration (ms)	Line	18.2 (5.8)	24.2 (8.1)	12.7 (7.3)	23.5 (9.9)	22.4 (11.5)			14.3 (2.5)	26.5 (3.6)			
Segment 2 duration (ms)	Line	24.6 (7.1)	15.7 (6.7)	21.2 (5.4)	7.8 (4.4)	12.6 (6.8)			20.6 (5.4)	26.1 (5.0)			
Segment 3 duration (ms)	Line		13.0 (4.0)	17.8 (5.4)	11.6 (7.4)	10.6 (4.8)							
Segment 4 duration (ms)	Line			10.3 (4.0)	11.5 (6.3)	7.0 (2.2)							
Segment 5 duration (ms)	Line				11.7 (6.2)	14.2 (6.3)							
Segment 6 duration (ms)	Line					6.4 (6.6)							
Gap 1 (ms)	Line	0.6 (0.1)	0.9 (0.5)	0.7 (0.2)	0.9 (0.6)	0.6 (0.2)							
Gap 2 (ms)	Line		0.6 (0.1)	1.2 (0.5)	0.6 (0.3)	1.4 (1.3)							
Gap 3 (ms)	Line			0.8 (0.2)	1.0 (0.3)	0.5 (0.0)							
Gap 4 (ms)	Line				0.8 (0.4)	0.8 (0.5)							
Gap 5 (ms)	Line					0.8 (0.2)							
Peak time (ms)	Fcontour								14.3 (2.5)	26.5 (3.6)			
Start f1 (kHz)	Line	87.7 (9.1)	71.3 (4.5)	87.0 (4.5)	76.3 (11.9)	84.6 (7.8)	68.5 (4.5)	81.4 (3.8)	73.7 (8.3)	68.2 (4.4)	70.9 (4.0)	75.1 (3.6)	77.2 (5.3)
Start f2 (kHz)	Line	77.8 (6.3)	94.5 (15.0)	75.8 (6.9)	95.3 (12.5)	89.9 (19.4)			66.0 (8.5)	84.7 (3.7)			
Start f3 (kHz)	Line		78.7 (6.3)	107.8 (6.6)	84.7 (6.8)	96.6 (10.2)							
Start f4 (kHz)	Line			76.5 (3.0)	92.2 (19.0)	94.9 (15.6)							
Start f5 (kHz)	Line				83.3 (8.0)	99.7 (4.7)							
Start f6 (kHz)	Line					78.6 (3.7)							
End f (kHz)	Line						85.8 (4.9)	69.5 (3.7)			74.8 (5.2)	76.4 (4.3)	
End f1 (kHz)	Line	89.5 (7.9)	84.7 (6.1)	91.5 (6.3)	85.9 (7.3)	89.5 (6.2)			66.0 (8.5)	84.7 (3.7)			77.6 (5.1)
End f2 (kHz)	Line	77.5 (8.8)	94.9 (15.4)	84.8 (4.5)	94.5 (11.0)	93.0 (13.0)			76.5 (9.3)	71.1 (3.4)			
End f3 (kHz)	Line		79.0 (6.3)	100.1 (4.7)	88.7 (5.9)	97.3 (8.8)							
End f4 (kHz)	Line			68.9 (1.4)	88.2 (15.6)	94.4 (16.0)							
End f5 (kHz)	Line				78.8 (13.6)	84.3 (14.5)							
End f6 (kHz)	Line					74.3 (5.3)							
Minimum f (kHz)	Fcontour	67.5 (3.3)	63.8 (4.8)	65.9 (3.0)	63.2 (3.8)	67.0 (2.9)	67.9 (4.7)	68.6 (3.6)	66.3 (8.0)	65.6 (3.9)	65.4 (3.1)	72.4 (4.1)	76.6 (5.2)
Maximum f (kHz)	Fcontour	97.0 (7.5)	103.5 (10.2)	110.7 (5.3)	105.7 (6.1)	111.7 (4.8)	84.7 (5.0)	81.2 (3.6)	78.1 (8.6)	84.3 (3.8)	84.5 (7.8)	78.3 (3.8)	78.9 (5.2)
Peak f (kHz)	Line								66.0 (8.5)	84.7 (3.7)			
Mean f (kHz)	Fcontour	80.3 (4.1)	82.1 (6.2)	86.5 (3.9)	83.4 (4.1)	87.3 (0.7)	77.3 (4.5)	75.5 (3.6)	71.1 (8.5)	76.7 (3.6)	75.4 (4.6)	75.8 (3.9)	77.6 (5.1)
Bandwidth (kHz)	Fcontour	29.5 (6.2)	39.7 (9.4)	44.8 (5.3)	42.5 (6.8)	44.7 (7.8)	16.8 (2.7)	12.6 (2.6)	11.8 (2.0)	18.7 (2.3)	19.1 (7.6)	5.9 (1.2)	
Δ f1 (kHz)	Fcontour	19.1 (7.2)	23.0 (8.6)	26.2 (2.4)	25.9 (6.9)	26.2 (3.4)							
Δ f2 (kHz)	Fcontour		19.8 (7.2)	22.1 (1.9)	23.5 (6.1)	24.3 (2.2)							
Δ f3 (kHz)	Fcontour			16.1 (1.9)	15.5 (6.2)	23.0 (1.8)							
Δ f4 (kHz)	Fcontour				13.5 (5.9)	22.0 (1.3)							
Δ f5 (kHz)	Fcontour					16.1 (3.1)							
FM amplitude (kHz)	Line										11.5 (4.3)		
F offset (kHz)	Line										66.0 (12.2)		
FM frequency (Hz)	Line										12.3 (3.4)		
F initial phase (°)	Line										−0.2 (13.9)		
Segment slope 1 (kHz/s)	Line	0.0 (0.3)	0.6 (0.2)	0.4 (0.4)	0.5 (0.5)	−0.2 (1.0)	0.6 (0.1)	−0.4 (0.1)	−0.6 (0.3)	0.7 (0.1)		0.1 (0.0)	
Segment slope 2 (kHz/s)	Line	−0.1 (0.4)	−0.2 (0.8)	0.4 (0.4)	−0.1 (0.3)	−0.2 (0.9)			0.6 (0.3)	−0.7 (0.1)			
Segment slope 3 (kHz/s)	Line		−0.6 (0.4)	−0.4 (0.6)	0.4 (0.3)	0.2 (0.4)							
Segment slope 4 (kHz/s)	Line			−1.2 (0.4)	−0.4 (0.4)	0.0 (0.1)							
Segment slope 5 (kHz/s)	Line				−0.6 (0.6)	−1.1 (0.3)							
Segment slope 6 (kHz/s)	Line					−1.0 (0.3)							
Peak time shift ABS (ms)	Both								0.5 (0.4)	0.5 (0.4)			
Peak time shift (ms)	Both								0.3 (0.6)	0.3 (0.5)			
Residual (kHz)	Line	1.3 (0.2)	1.3 (0.4)	1.2 (0.2)	1.0 (0.5)	1.1 (0.2)	1.3 (0.2)	1.3 (0.3)	0.8 (0.2)	1.1 (0.2)	1.6 (0.4)	0.8 (0.2)	

<sup>a</sup>Data are mean (SD). ABS, Absolute value.

defined as the lowest sound pressure level (SPL) in which a recognizable waveform was present and repeatable. For WT animals, stimuli were presented 500 times from 80 to 20 dB SPL in steps of 10 and then 1000 repetitions in steps of 5 dB SPL when approaching threshold. Stimuli for *DTR* animals were presented at 1000 repetitions at intensities of 90 and 70 dB SPL. Thresholds were determined at 4, 8, 16, and 32 kHz (stimuli were 5 ms duration, 1 ms rise/fall time, repetition rate 19/s) and for a broad-band click.

#### Vocalization recordings

Vocalizations emitted in the presence of a female were recorded from 21 *DTR* and 12 WT male mice between the ages of P60 and P70. The male was placed in an empty acrylic cage (10 × 19 × 8 in.), located inside a dark, single-walled sound-attenuating chamber lined with anechoic foam and allowed to acclimate for 2 min. An age-matched female was then placed in the cage, and vocalizations were recorded for 15–20 min.

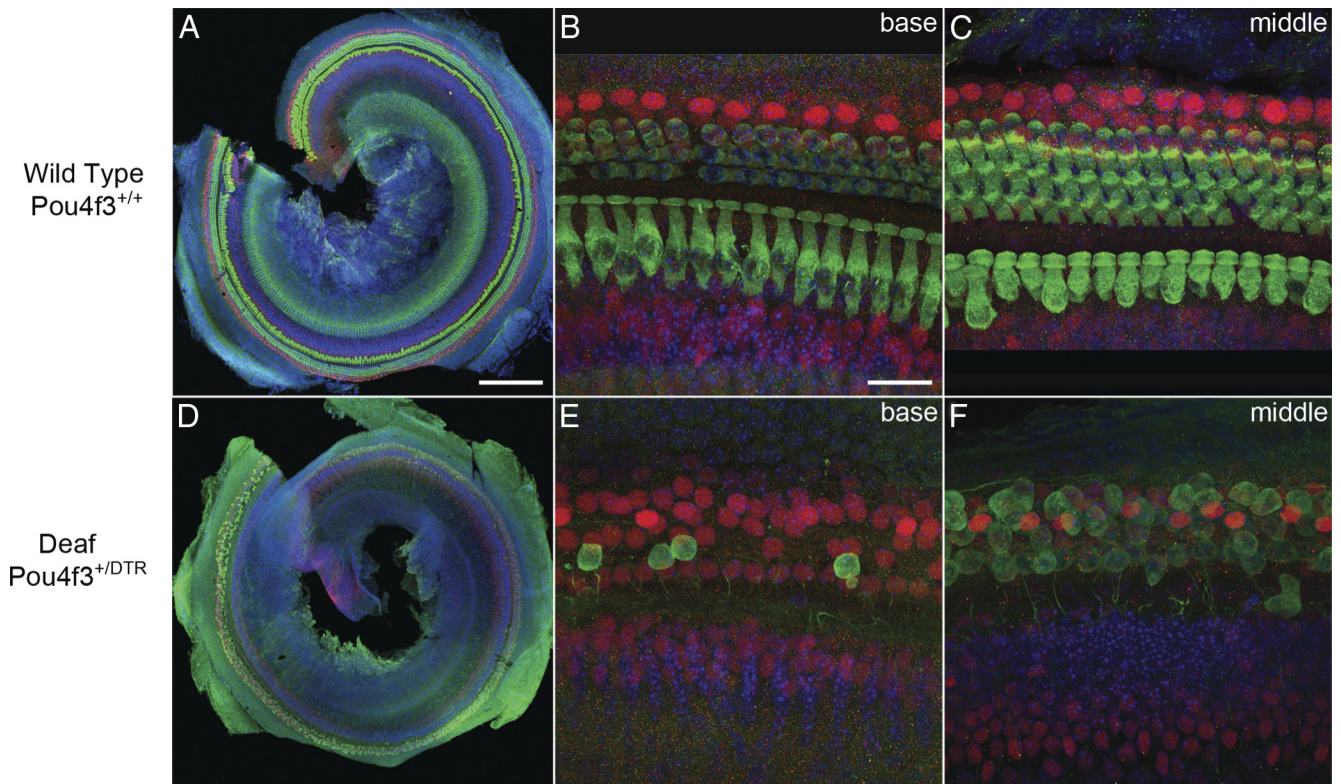
Females were paired once with each male but were never paired with more than three males in a day, and never consecutively. Recordings were always done during their active (dark) period and at approximately the same time across days. We recorded up to five different sessions from each male, each on different days.

Vocalizations were recorded with an UltraSoundGate CM16 microphone (Avisoft Bioacoustics) positioned 20 cm above the cage floor. The microphone was connected to an Avisoft UltraSoundGate 416H preamplifier (Avisoft Bioacoustics), and the acoustic signals were amplified and digitized at a sampling rate of 375 kHz with 16-bit resolution. Gain was manually adjusted during recordings to optimize acoustic sampling while preventing saturation.

#### Data analysis

We used Avisoft SASLab Pro software (Avisoft Bioacoustics) for initial analysis of the same two recording sessions for all animals. We calculated



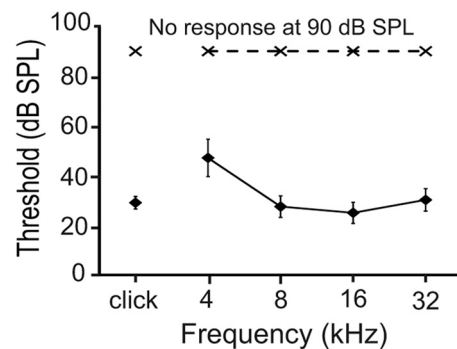


**Figure 2.** Low- and high-power confocal images of whole-mount preparations from representative P9 mice. **A–C**, WT (*Pou4f3*<sup>+/+</sup>) mouse. **D–F**, *Pou4f3*<sup>+/DTR</sup> mouse. Both mice were injected with DT on P2. Red cells show antigenicity to SOX2, indicating an organ of Corti support cell phenotype; hair cells are indicated by antigenicity to a mixture of antibodies against myosin6 and parvalbumin, and blue represents DAPI-stained nuclei. **A**, In WT mice, the low-power image shows that the full complement of hair cells was present throughout the basal to apical turns. **B**, High-power image shows the characteristic single row of inner hair cells, a space for the tunnel of Corti, and then three rows of outer hair cell somata in the basal region. **C**, Middle turn region. **D**, In the *Pou4f3*<sup>+/DTR</sup> mice, however, the low-power image reveals a dramatic loss of hair cells by 7 d after the DT injection. **E**, High-power images reveal complete loss of inner hair cells and almost complete loss of outer hair cells in the basal region, coupled with what appears to be complete survival of supporting cells. **F**, In the middle region, an occasional inner hair cell remained, but it usually appeared swollen and degenerative, whereas a large complement of outer hair cells had not yet degenerated. Again, supporting cell complement appeared intact. By P16, all hair cells had been lost (data not shown). Scale bars: (in **A**, **D**) 100  $\mu$ m; (in **B**) **B**, **C**, **E**, **F**, 10  $\mu$ m.

spectrograms (Hamming window; FFT length 1024; 100% frame size; 75% temporal overlap) and adjusted the element detection threshold parameters (minimum duration 1 ms; maximum entropy of 1; hold time 15–20 ms) manually for each recording session to maximize detection of vocalizations and minimize detection of nonvocal sounds. The software automatically detected syllables, defined as continuous sounds above a certain power threshold bounded by silence of at least 2 ms, and provided syllable beginning and end times. We calculated the number of syllables emitted per minute for each recording session.

**Temporal analysis of syllables.** We defined a syllable bout as a group of three or more syllables bounded by at least 130–180 ms of silence. The general range for this threshold was determined by examining the valley in the distribution of the duration of silent periods (corresponding to intersyllable and interbout intervals) for all animals. The specific value for each individual animal was determined by an experimenter who was blind to the condition of the animal, and based on the highly reliable intersyllable interval within a bout. Avisoft automatic detection software (Avisoft Bioacoustics) automatically determined the number of bouts. For each recording session, we calculated the fraction of syllables in a bout, number of bouts/min, mean number of syllables/bout, and mean intersyllable interval (from start of one syllable to start of the next syllable). We averaged values across recording sessions to determine mean values for each animal.

**Syllable categorization.** We assigned each syllable to one of 12 categories by visually inspecting the spectrogram, generally following the scheme of Scattoni et al. (2008). Table 1 illustrates the categorization criteria. For each recording session, we calculated the fraction of syllables in each category and averaged these values across recording sessions for each animal.



**Figure 3.** *Pou4f3*<sup>+/DTR</sup> mice were deaf. ABRs from all male mice were obtained at P70. Thresholds are mean  $\pm$  SD for hearing ( $\blacklozenge$ ) and deaf ( $\times$ ) animals. The highest level tested with the *DTR* mice was 90 dB SPL, and none of the animals had any response at this intensity.

**Syllable-sequence analysis.** For each recording session, we calculated a matrix of syllable-transition events, including each of the 12 syllable categories and silent periods between bouts. We normalized this matrix in two ways. First, each element was divided by the total number of transitions in that session, such that the value at element *i, j* in the matrix represents the probability during that session of observing a transition from syllable category *i* to category *j*. The second type of normalization gave us conditional transition probabilities. We normalized across rows of the matrix to calculate the probability that, given that the previous syllable was of category *i*, the next one would be of type *j*. Each of these matrices was averaged across recording sessions for each animal.

**Acoustic measurements.** For a subset of syllables with the highest signal-to-noise ratio, we calculated detailed acoustic parameters. For this analysis, we used the Matlab software package Mouse Vocalization Categorizer (MUSCat), originally developed by Dr. S. E. Roian Egnor (HHMI, Janelia Farm), with our own modifications. MUSCat facilitated semi-automated syllable detection and categorization and calculated a spectrogram (Hamming window, no overlap, 0.5 ms time bins, 122 Hz frequency resolution) for each syllable. For each time bin with acoustic power above a given threshold, MUSCat measured the frequency with highest power. The set of these frequency values within the spectrogram determined the frequency contour of the syllable (Fcontour). To ensure precise contouring, an experimenter manually viewed each Fcontour overlaid on the spectrogram of the syllable. If necessary, the experimenter manually corrected the contour to match the actual syllable. Syllable frequency contours allowed us to examine many more acoustic parameters than more traditional methods.

To reduce the dimensionality of the data, we then fit the contour points with simple functions. Frequency-jump syllables were fitted with 2–6 line segments, one for each continuous acoustic element. FM up-sweeps, FM down-sweeps, constant frequency, and short syllables were fitted with a single line segment. Chevrons and reverse chevrons were fitted with two intersecting line segments. Complex syllables were fitted with a sinusoid, varying center frequency, amplitude of frequency modulation, frequency of frequency modulation, initial phase, and duration. Although most parameters were measured from the fitted line segments, some were measured directly from the Fcontours for greater precision. The fitted parameter values for each syllable type and their measurement locations are listed in Tables 2 and 3. We averaged values for each measurement across recording sessions for each animal.

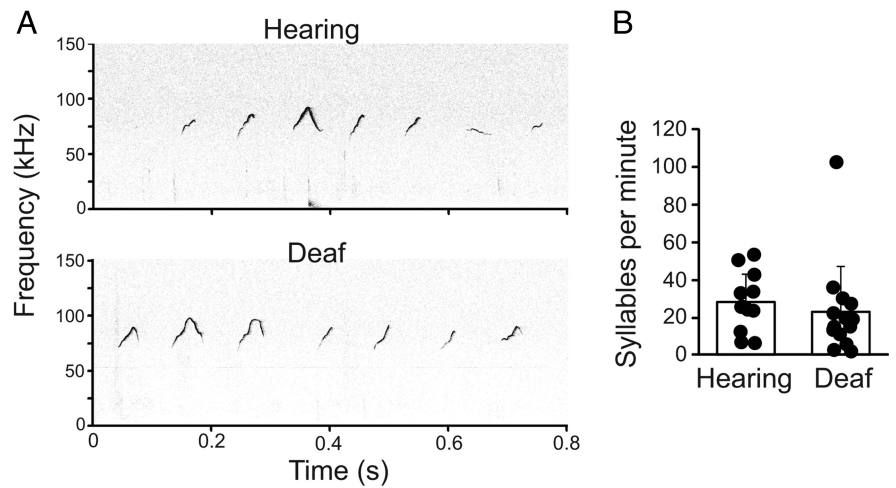
To assess the quality of fits to Fcontour values, we calculated a residual value, defined as the mean absolute frequency error per contour point across each syllable. A perfect fit had a residual value of 0 Hz. We empirically determined that syllables with poor fits had residual values  $>3$  kHz and thus excluded syllables with residual values larger than this threshold, resulting in exclusion of 5.6% of syllables from the acoustic analysis.

**Statistical analysis.** We tested for normal distribution of all vocalization parameters using Kolmogorov–Smirnov tests. All parameters had non-normal distributions so we used only nonparametric statistical tests. All results presented in the text and tables are mean  $\pm$  SD. To compare group means for bout structure, syllable transition pattern, and acoustic parameters, we used Mann–Whitney U tests with correction for multiple testing (Simes correction,  $\alpha = 0.05$ ) (Simes, 1986). For all of these tests, the sample size was 12 hearing animals and 14 deaf animals. To compare distribution of syllable types emitted by WT and *DTR* groups, we used a Kruskal–Wallis test. To compare similarity of vocalization parameters of siblings and nonsiblings, we calculated coefficients of variation (SD/mean) and compared these using Mann–Whitney U tests with Simes correction for multiple testing. We set our significance level at  $p < 0.05$  for all statistical tests.

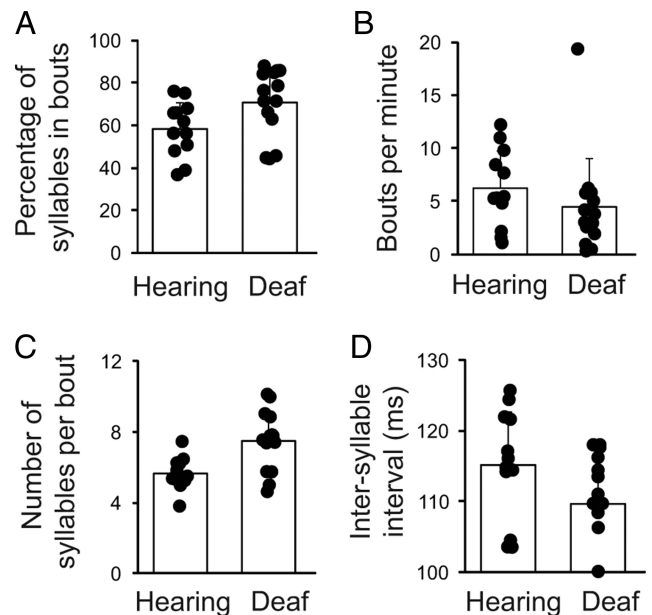
## Results

Of the 49 mice that were given injections of DT, 21 were male *Pou4f3*<sup>+/*DTR*</sup> (*DTR*), and 12 were male *Pou4f3*<sup>+/*+*</sup> (WT). Of these males, 14 *DTR* and 12 WT animals emitted  $>30$  syllables across recording sessions and were included in the data analysis.

To assess the efficacy of DT injections given at P2 to eliminate hair cells, we examined whole-mounts 7 (P9) and 14 (P16) days after injection. Figure 2 illustrates that, at 7 d after DT injection,



**Figure 4.** Vocalizations of hearing and deaf male mice were qualitatively similar. **A**, Example spectrograms of ultrasonic vocalizations emitted by one hearing and one deaf male mouse in the presence of a female. **B**, Mean number of syllables emitted per minute  $>15$  min recording sessions of hearing and deaf mice. The circles represent the mean values for each animal, and the bars represent the mean values for the group. Error bars indicate the SDs of the group means.

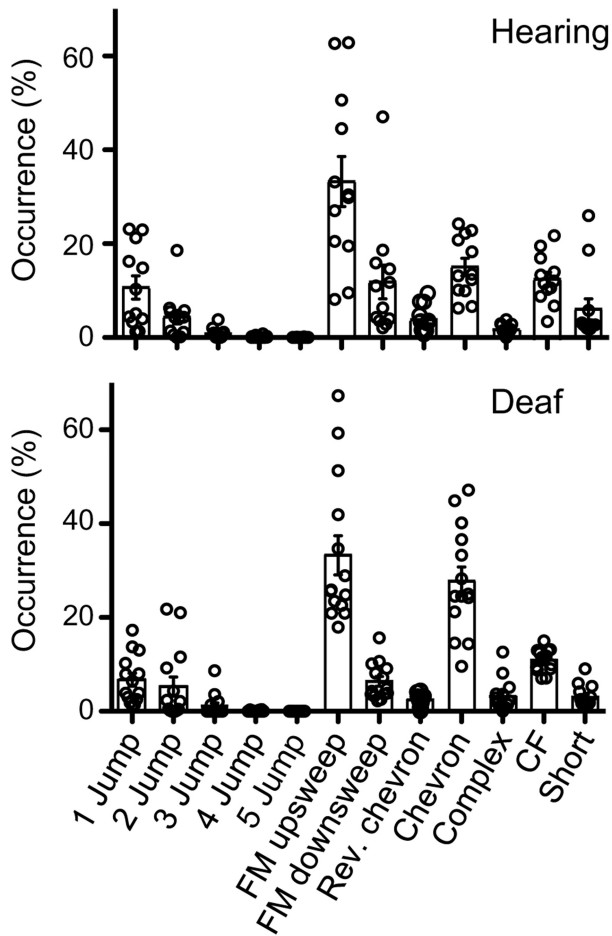


**Figure 5.** The temporal organization of vocalizations emitted by hearing and deaf mice was similar. **A**, Fraction of syllables contained in bouts. **B**, Number of bouts emitted per minute. **C**, Number of syllables per bout. **D**, Intersyllable interval. In all plots, the circles represent the mean values for each animal, and the bars represent the mean values for the group. Error bars indicate the SDs of the group means.

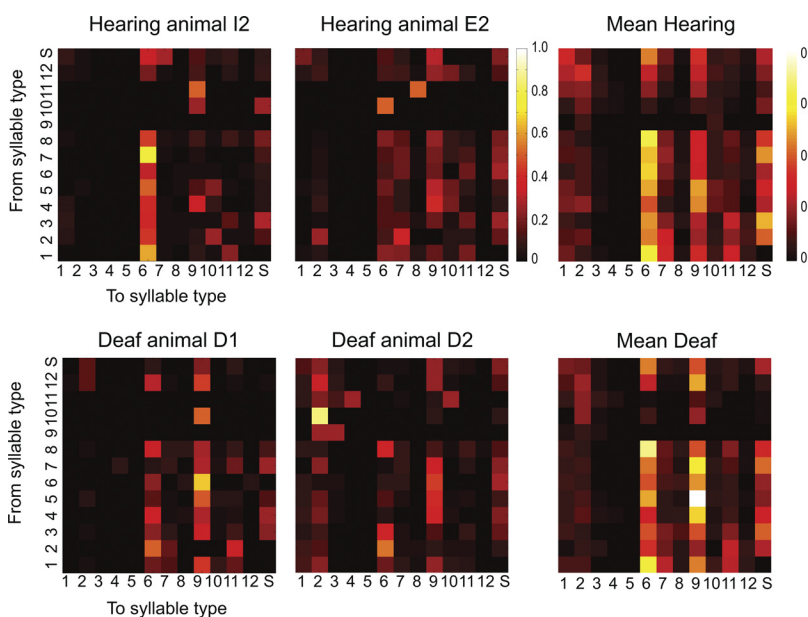
the majority of inner hair cells were gone (base, 100% gone; middle,  $>95\%$  gone; apex, 95% gone) and  $>50\%$  of outer hair cells were gone (base,  $>90\%$  gone; middle, 60–70% gone; apex, 50–70% gone). In contrast, all inner and outer hair cells were normal throughout the cochlea of WT mice. At 14 d after DT injection (P16), all inner hair cells were gone throughout the cochlea and  $<5\%$  of outer hair cells remained in the *DTR* mice. WT mice had a normal complement of inner and outer hair cell loss. Thus, we confirmed that *DTR* mice had no hearing ability during development and that the DT injections in WT mice had no effect.

To test auditory function, we obtained ABR thresholds for clicks and pure tones (4–32 kHz). In contrast to the WT mice that had normal ABR thresholds (Fig. 3), none of the 14 *DTR* mice





**Figure 6.** Hearing and deaf mice emitted the same types of syllables. The relative occurrence of each syllable type for hearing and deaf animals. For each syllable, the circles represent the mean values for each animal, and the bars represent the mean values for the group. Error bars indicate the SDs of the group means.



**Figure 7.** The pattern of syllable emissions was not different between hearing and deaf males. Transition probability matrices for two hearing males, two deaf males, and the hearing and deaf group means. Syllable categories 1–12 are as follows: one jump, two jump, three jump, four jump, five jump, FM upsweep, FM downsweep, reverse chevron, chevron, constant frequency, complex, and short, respectively. S, Silence.

showed an ABR at 90 dB SPL (the maximum intensity available) for all sound stimuli presented (Fig. 3). For the remainder of this report, we refer to WT mice as hearing and *DTR* mice as deaf.

Both hearing and deaf mice readily investigated the female and emitted qualitatively similar vocalizations (Fig. 4A). Although syllable production rates were quite variable across individuals, this parameter did not differ significantly between the hearing and deaf groups (Mann–Whitney U test;  $U = 55.0$ ;  $p = 0.15$ ) (Fig. 4B). Over the 15 min recording sessions, hearing animals emitted an average of  $27.9 \pm 14.7$  syllables/min and deaf animals emitted an average of  $22.7 \pm 24$  syllables/min. Because our recording sessions were much longer than other studies (Hammerschmidt et al., 2012), we examined emission rates for the first 3 min after the female was introduced into the cage. Hearing animals emitted an average of  $93 \pm 46.8$  syllables/min, and deaf animals emitted an average of  $120 \pm 89$  syllables/min. These values were not statistically different (Mann–Whitney U test;  $U = 74$ ,  $p = 0.60$ ).

We analyzed the temporal organization of 21,251 and 18,662 syllables emitted by hearing and deaf males, respectively. Hearing and deaf males emitted the majority of syllables in bouts, with no significant difference between hearing and deaf animals ( $58 \pm 12.3\%$ , hearing;  $70.4 \pm 15.3\%$ , deaf; Mann–Whitney U test;  $U = 42.0$ ;  $p = 0.08$ ; Fig. 5A). The number of bouts emitted per minute did not differ ( $6.2 \pm 3.5$ , hearing;  $4.4 \pm 4.5$ , deaf; Mann–Whitney U test;  $U = 54.0$ ;  $p = 0.16$ ; Fig. 5B). Deaf animals emitted a slightly higher number of syllables/bout ( $5.6 \pm 0.8$ , hearing;  $7.5 \pm 1.6$ , deaf; Mann–Whitney U test;  $U = 33.0$ ;  $p = 0.04$ ; Fig. 5C). Finally, the intersyllable intervals did not differ ( $115 \pm 7.4$  ms, hearing;  $109.5 \pm 7.3$  ms, deaf; Mann–Whitney U test;  $U = 53.0$ ;  $p = 0.2$ ; Fig. 5D).

We categorized the syllables into 12 types. Although there was high variability across individual animals, both hearing and deaf mice emitted all 12 types of syllables (Fig. 6). FM upsweeps and chevrons were the most commonly emitted syllable types in both hearing and deaf animals. Frequency-jump vocalizations with 4 and 5 jumps were emitted only rarely. Except for the deaf mice emitting significantly more chevrons (15% hearing; 27% deaf; Kruskal–Wallis;  $H_{(2)} = 9.5$ ;  $p = 0.002$ ), the hearing and deaf mice emitted the syllable types in similar proportions.

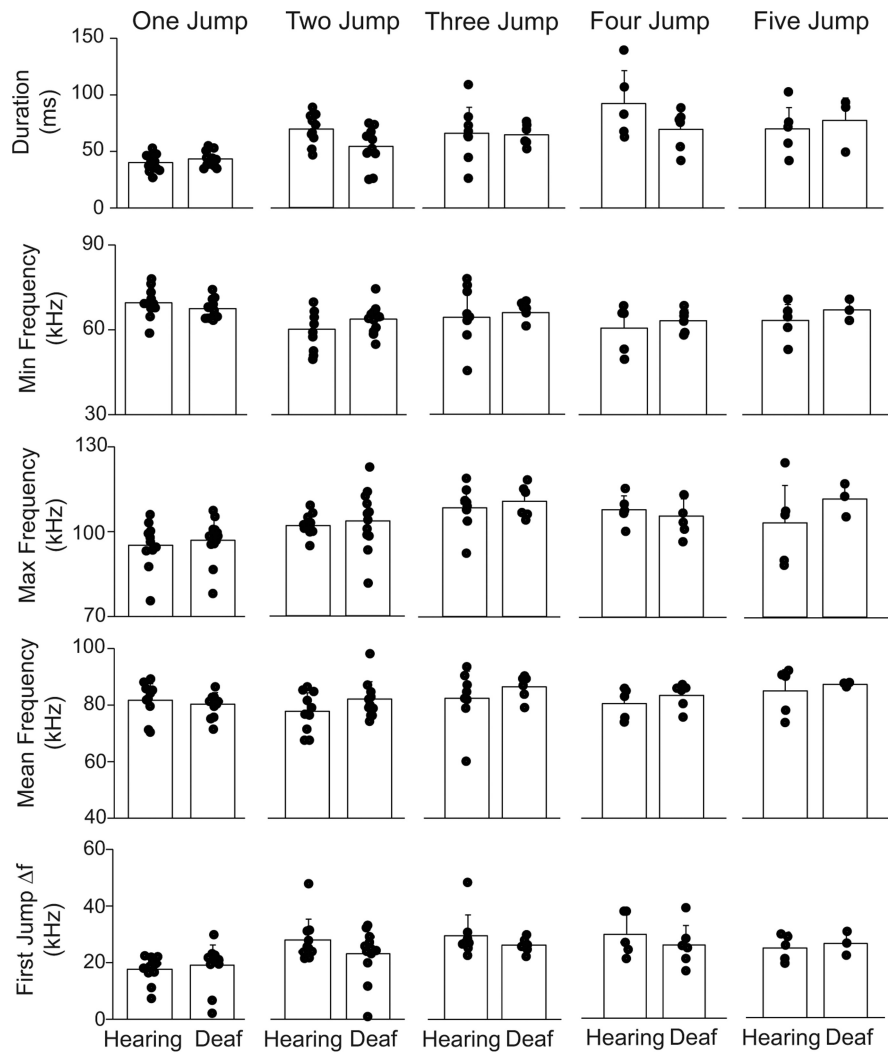
We calculated transition-probability matrices to describe the likelihood of transitions from any given syllable type to any other syllable type. When we analyzed the conditional syllable transitions of 12 different syllable categories, there were no differences between hearing and deaf male mice (Mann–Whitney U tests;  $p > 0.169$  for all tests). Males within each group exhibited greater variability in the ordering of their syllables than did males across the hearing and deaf groups. This is illustrated by the heat maps in Figure 7, where no obvious patterns appear across individual males within either hearing or deaf groups. The average transition probabilities of the hearing and deaf groups were not significantly different. When we analyzed the normalized syllable transitions of 12 different syllable categories,

there were no differences between hearing and deaf male mice, except for chevron-to-chevron transitions (Mann–Whitney *U* test;  $U = 13.0$ ;  $p = 0.04$ ).

We found no significant differences in any of the acoustic parameters between hearing and deaf animals (Mann–Whitney *U* tests;  $p > 0.3$  for all tests). Figures 8 and 9 illustrate a subset of acoustic parameters that had the lowest *p* values for syllables with and without frequency jumps, respectively. Tables 2 and 3 show the mean  $\pm$  SD for all acoustic parameters for the hearing and deaf groups, respectively.

A striking aspect of both hearing and deaf mice vocalizations was the amount of variability in almost all features. Much of this variability was not accounted for in our tests for statistical differences between hearing and deaf vocalization parameters because of the way we pooled the data. In all measurements, we first calculated the mean for each animal and then calculated the mean for the group. Thus, the variance was determined based on the means of the animals within a group, and variability within an animal was ignored. As can be seen in Figure 10, the variability within an animal was often greater than the variability within a group. Figure 10A shows how the number of syllables per bout was highly variable within an individual animal. Figure 10B shows how the mean  $\pm$  SD of individual animals also varied within a group. Moreover, the within animal variability for this parameter was clearly greater than the within group variability. Figure 10 also shows that the within animal variability was greater in the deaf animals than in the hearing animals. Although we show this variability analysis for only the number of syllables/bout, similar results held for the majority of parameters we measured. One notable exception was intersyllable interval. The amount of between animal variability was low (Fig. 5D), and the within animal variability was also low, suggesting tighter control over this particular parameter of vocalization behavior.

To test whether nonauditory feedback, perhaps via social signals from the mother or cage-mates, could shape mouse vocalizations, we compared the similarities of vocalizations emitted by siblings with those emitted by nonsiblings. We reasoned that, if social interactions (or genetic similarity) consistently shaped vocalization parameters, siblings should exhibit more similarity in their vocalization parameters than nonsiblings. We calculated coefficients of variation (SD/mean) for temporal and acoustic parameters of all pairs of siblings and pairs of nonsiblings. We found that only one parameter was significantly more similar in sibling pairs than nonsibling pairs: intersyllable interval (Mann–Whitney *U* test;  $N = 24$  sibling pairs, 301 nonsibling pairs;  $U = 2119$ ;  $p = 0.03$ , all other parameters,  $p > 0.4$ ). However, because all animals, hearing or deaf, had very similar values of intersyllable intervals (Fig. 5D), the actual differences in CVs between



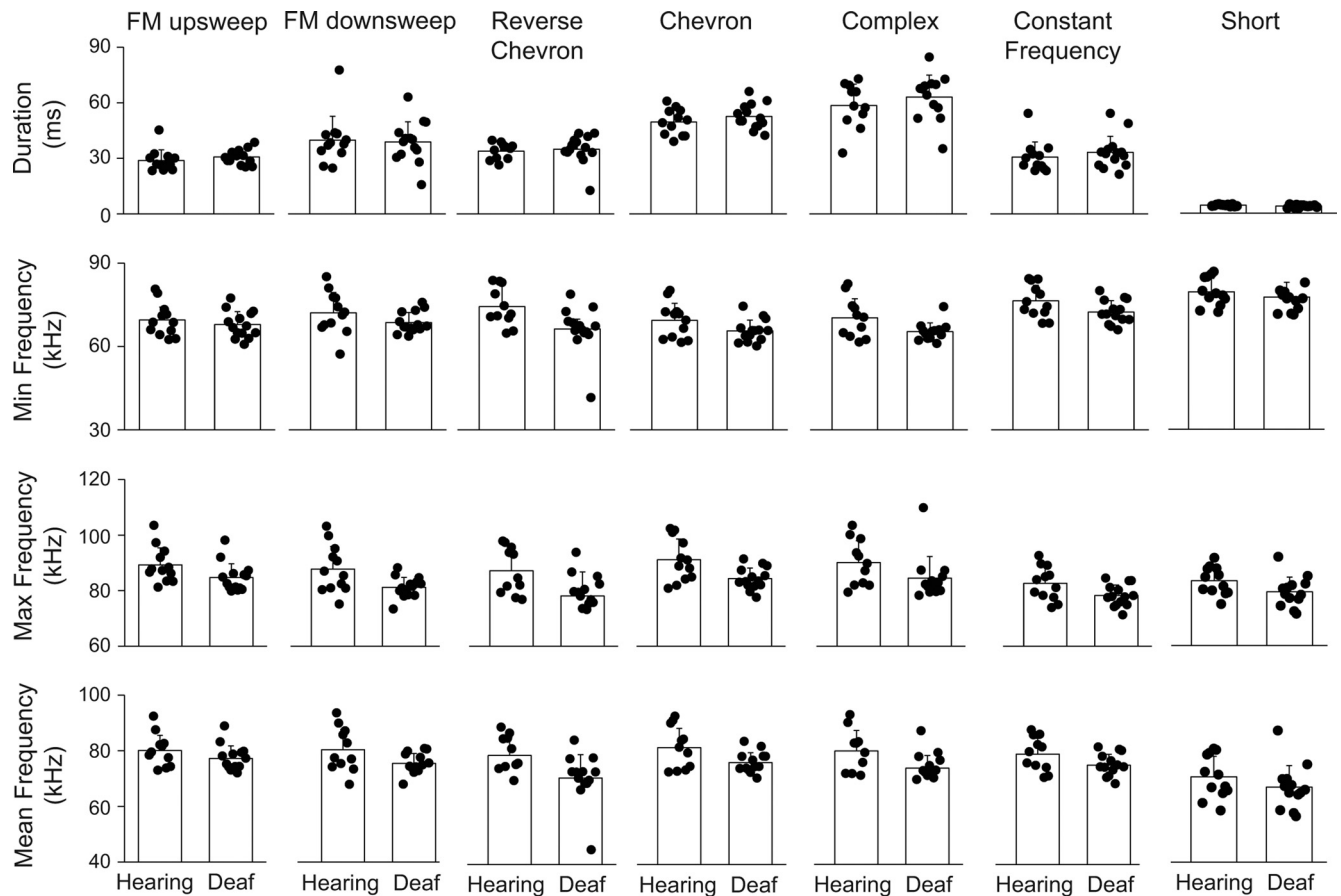
**Figure 8.** Acoustic parameters in jump syllables were not different for hearing and deaf animals. In all plots, the circles represent the mean values for each animal, and the bars represent the mean values for the group. Error bars indicate the SDs of the group means.

sibling and nonsibling pairs were very small. Because deaf mice emitted a higher relative percentage of chevron syllables, we asked whether siblings were more likely to emit a similar fraction of chevrons than were nonsiblings. The CVs of the fraction of chevrons emitted by siblings and nonsiblings were not significantly different (Mann–Whitney *U* test;  $N = 24$  sibling pairs, 301 nonsibling pairs;  $U = 2614$ ;  $p = 0.24$ ).

## Discussion

We compared vocalizations emitted by normal hearing and deaf male mice. To induce deafness, we used mice engineered to express DTRs in hair cells so that, by injecting DT at P2, we were able to eliminate all inner and most outer hair cells by P9, before mice normally begin to hear. Thus, we were able to prevent auditory experience during development. We used this approach because in songbirds, which exhibit vocal learning, juvenile deafening is the manipulation that most dramatically disrupts adult song (Konishi, 1965). We found that deafness failed to disrupt the production of normal adult courtship vocalizations in mice. Our detailed, quantitative analysis of the temporal structure of vocalization bouts, the types of vocalizations, the patterns of syllables, and the acoustic features of each syllable type showed that vocalizations emitted by hear-





**Figure 9.** Acoustic parameters in nonjump syllables were not different for hearing and deaf animals. In all plots, the circles represent the mean values for each animal, and the bars represent the mean values for the group. Error bars indicate the SDs of the group means.

ing and deaf mice were nearly indistinguishable. Specifically, deaf mice emitted all the same syllable types, with the same acoustics, as hearing mice. These findings indicate that mice, in contrast to songbirds and humans, have little, if any, need for auditory experience during development to produce normal adult vocalizations. Further, our finding that vocalizations of siblings are not more similar than those of nonsiblings suggests that social interactions do not shape vocalizations in mice.

### Strengths of the *Pou4f3*<sup>+DTR</sup> mouse

The *DTR* mouse has a number of advantages over other mice with engineered deafness or spontaneous mutations of genes required for inner ear development. First, because the hair cells are only killed once the DT is injected, the timing of deafness can be controlled and all hearing ability eliminated quickly at any age. Here, we injected DT at P2 and found that all inner and most outer hair cells are eliminated by P9. This is in contrast to mice engineered with gene knock-outs where the timing and extent of hearing loss are less certain. For example, *caspase-3* knock-outs have residual hearing up to 5 weeks of age (Takahashi et al., 2001). Thus, it is unclear the extent to which these mice have auditory experience during development. Second, because *pou4f3* is not widely expressed throughout the brain, the DT injections selectively targeted hair cells. In contrast, knocking out genes can have global effects. For example, knock-out of *caspase-3* leads to loss of inner hair cells (Takahashi et al., 2001) and also causes abnormal morphology throughout the brain (Kuida et al., 1996).

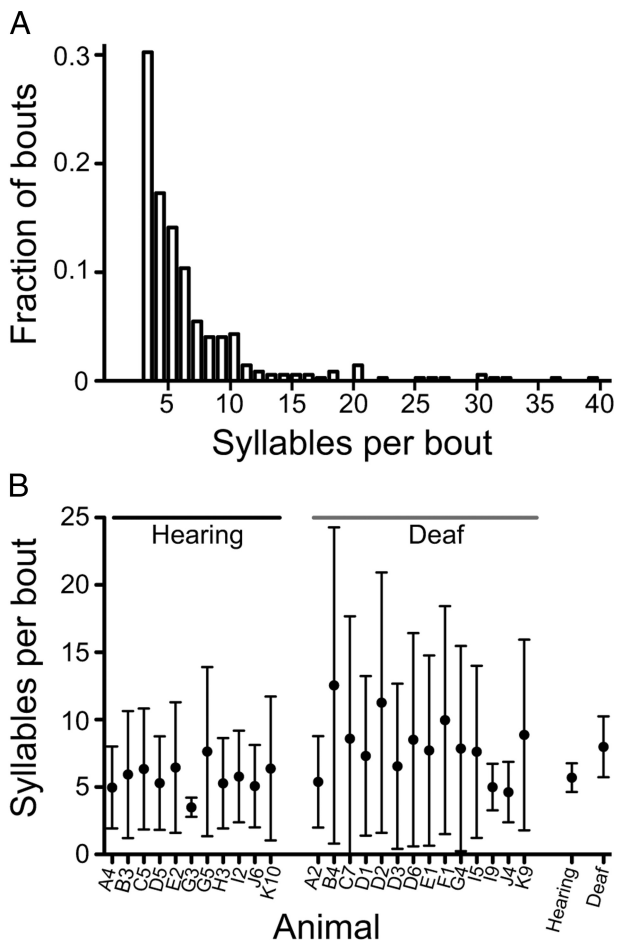
The degraded vocalizations produced by these mice (Arriaga et al., 2012) may well be independent of any hearing loss in these animals.

Finally, our *DTR* mice were of the CBA/CaJ strain, which has normal hearing (Zheng et al., 1999). In contrast, many mice with gene knock-outs are bred on strains with compromised hearing. For example, the C57Bl/6 strain, which is often used in vocalization studies, starts losing its high-frequency hearing at 3 months of age (Henry and Lepkowski, 1978; Henry and Chole, 1980; Zheng et al., 1999). Thus, even “control” animals may have abnormal auditory experience, potentially compromising the interpretation of any negative results.

Considering the strengths of the *DTR* mouse model, our finding that deaf mice emit vocalizations that are very similar to those emitted by mice with normal hearing indicates that the fundamental features of mouse courtship vocalizations are not altered by a lack of hearing during development.

### Deaf-rearing does not degrade mouse vocalizations

Rearing deaf animals has the potential to detect any of several possible roles of auditory experience in the development of normal vocalizations. As in humans and songbirds, the normal developmental process could include a stage of acquiring a template for later vocal production and a stage of practicing to learn to produce an acquired or even an innate template (Marler, 1970; Doupe and Kuhl, 1999; Kuhl, 2003). Loss of hearing would be expected to disrupt all of these processes, resulting in abnormal vocalizations. However, secondary effects of deafening, such as



**Figure 10.** The variability of vocalization parameters within an animal was high. **A**, Histogram of syllables per bout for all of one animal's vocalizations. **B**, Syllables per bout for each animal (mean  $\pm$  SD) and group mean  $\pm$  SD.

decreased social interaction, could also contribute to degraded vocalizations, complicating the interpretations of results. Our finding that deaf mice do not have degraded vocalizations leads to the conclusion that mice do not require auditory feedback to develop essentially normal vocalizations.

We examined a number of parameters of bout structure and found that almost all features were similar between hearing and deaf animals. In particular, the intersyllable interval was remarkably similar. Both groups emitted syllables with intervals (start-to-start) of 92–120 ms. This  $\sim$ 9 Hz periodicity is in line with that reported previously in CBA mice (Liu et al., 2003) and suggests that a central pattern generator underlies this innate rhythmicity. The one statistically significant difference in bout structure is that deaf animals emitted slightly more syllables per bout. Considering that several possible mechanisms could underlie the control of this parameter (e.g., efference copy signals, sensory or social feedback), it is unlikely that the slight difference is the result of imitative vocal learning.

It is well documented that, in the presence of a female or her urine, male mice produce different syllable types (Holy and Guo, 2005; Portfors, 2007; Hammerschmidt et al., 2012; Hanson and Hurley, 2012). Different strains have been reported to emit approximately the same types of syllables but at different relative probabilities of occurrence (Panksepp et al., 2007; Choi et al., 2011). Here, the overall category distributions were similar between groups, although deaf animals emitted significantly more

chevron syllables. The finding that deaf mice emitted all the same syllable types as hearing mice indicates that the production of specific syllable types is not learned. It is unclear why deaf males produce more chevrons than hearing males, but because chevron production is not correlated with courtship behavior (Hanson and Hurley, 2012), it seems unlikely that these affect female mating decisions.

We examined in detail the first-order syllable-transition probabilities. Because we examined all possible transitions, not only those between common syllable types (Choi et al., 2011) or just between jump and nonjump syllables (Holy and Guo, 2005; Kikusui et al., 2011), our assay had the sensitivity to detect subtle syllable sequencing differences between hearing and deaf animals. Despite this sensitivity, the only difference was that deaf animals exhibited more chevron-to-chevron transitions. This result is likely an unavoidable consequence of the deaf animals emitting more chevron syllables, and is not evidence of imitative vocal learning.

Last, we examined whether the acoustics of syllables emitted by hearing and deaf males were different. We divided syllables into more categories than a previous study using *Otoferlin* knock-out mice (Hammerschmidt et al., 2012) to reduce the chance of missing an effect of learning on subtle syllable characteristics. Because deafening did not cause significant changes in any of the 99 independent acoustic parameters we measured, our findings confirm and extend the conclusion that acoustic features are not learned.

In summary, we compared  $>$ 250 parameters describing the features of mouse vocalizations emitted by hearing and deaf male mice and found only three statistically significant differences. However, the statistical significance of these differences may have depended on our use of mean values from each individual, ignoring individual variability. In all parameters measured, the variability within individuals was always greater than the variability of the mean parameter values across individuals within the group. The origin of this high within-individual variability is unknown. It could reflect intrinsic variations in neural and/or muscular signals, or be driven by changing environmental conditions not addressed here.

It is also unclear whether these differences are biologically important. For example, it is unknown whether the number of syllables in a bout or the distribution and sequencing of syllable types emitted by males alter female mating decisions. In addition, because of the large within-animal variability in vocalization parameters, a female would need to hear many repetitions to determine whether the emitter was hearing or deaf.

However, even if the changes we observed in deaf mice are statistically and biologically significant, they do not necessarily provide evidence for imitative vocal learning. For example, these differences could be the result of operant conditioning, a form of learning that could shape vocalizations through external feedback, such as social interactions. On the other hand, we did not find evidence for such effects when we compared vocalizations of siblings and nonsiblings. It will be important in future studies to identify the environmental and genetic factors that cause individual and strain differences in mouse vocalizations.

#### Mice are useful models for genetic disorders of vocal communication

Our finding that mice do not learn their vocalizations indicates that mice are not a good model for studying mechanisms of vocal learning. They are, however, useful for understanding the genetic and neural mechanisms of normal vocal communication and

associated disorders. For example, manipulation of particular genes has already shed light on the genetic basis of human communication disorders (Enard et al., 2009; Scattoni et al., 2009; Wohr et al., 2011; Fujita et al., 2012; Schmeisser et al., 2012; Srivastava et al., 2012). It seems likely that the genetic mechanisms underlying vocal communication are conserved across mammals. Thus, understanding these mechanisms in mice will provide a basis for teasing apart innate and learned aspects of vocal communication in humans.

## References

- Arriaga G, Zhou EP, Jarvis ED (2012) Of mice, birds and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS One* 10:e46610. [CrossRef Medline](#)
- Choi H, Park S, Kim D (2011) Two genetic loci control syllable sequences of ultrasonic courtship vocalizations in inbred mice. *BMC Neuroscience* 12:104. [CrossRef Medline](#)
- D'Amato FR, Moles A (2001) Ultrasonic vocalizations as an index of social memory in female mice. *Behav Neurosci* 115:834–840. [CrossRef Medline](#)
- Deecke VB, Ford JK, Spong P (2000) Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Anim Behav* 60:629–638. [CrossRef Medline](#)
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* 22:567–631. [CrossRef Medline](#)
- Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J (2009) A humanized version of FOXP2 affects cortico-basal ganglia circuits in mice. *Cell* 137:961–971. [CrossRef Medline](#)
- Fujita E, Tanabe Y, Imhof BA, Momoi MY, Momoi T (2012) CADM1-expressing synapses on Purkinje cell dendrites are involved in mouse ultrasonic vocalization activity. *PLoS One* 7:e30151. [CrossRef Medline](#)
- Golub JS, Tong L, Ngyuen TB, Hume CR, Palmiter RD, Rubel EW, Stone JS (2012) Hair cell replacement in adult mouse utricles after targeted ablation of hair cells with diphtheria toxin. *J Neurosci* 32:15093–15105. [CrossRef Medline](#)
- Hammerschmidt K, Reisinger E, Westekemper K, Ehrenreich L, Strenze N, Fischer J (2012) Mice do not require auditory input for the normal development of their ultrasonic vocalizations. *BMC Neurosci* 13:40. [CrossRef Medline](#)
- Hanson JL, Hurley LM (2012) Female presence and estrous state influence mouse ultrasonic courtship vocalizations. *PLoS One* 7:e40782. [CrossRef Medline](#)
- Henry KR, Chole RA (1980) Genotypic differences in behavioral, physiological and anatomical expressions of age-related hearing loss in the laboratory mouse. *Audiology* 19:369–383. [CrossRef Medline](#)
- Henry KR, Lepkowski CM (1978) Evoked potential correlates of genetic progressive hearing loss: age-related changes from the ear to the inferior colliculus of C57BL/6 and CBA/CaJ mice. *Acta Otolaryngol* 86:366–374. [CrossRef Medline](#)
- Holy TE, Guo Z (2005) Ultrasonic songs of male mice. *PLoS Biol* 3:e386. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Knörnschild M, Nagy M, Metz M, Mayer F, von Helversen O (2010) Complex vocal imitation during ontogeny in a bat. *Biol Lett* 6:156–159. [CrossRef Medline](#)
- Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* 22:770–783. [Medline](#)
- Kuhl PK (2003) Human speech and birdsong: communication and the social brain. *Proc Natl Acad Sci U S A* 100:9645–9646. [CrossRef Medline](#)
- Kuida K, Zheng TS, Na S, Kuan C, Yang D, Karasuyama H, Rakic P, Flavell RA (1996) Decreased apoptosis in the brain and premature lethality in CPP32-deficient mice. *Nature* 384:368–372. [CrossRef Medline](#)
- Liu RC, Miller KD, Merzenich MM, Schreiner CE (2003) Acoustic variability and distinguishability among mouse ultrasound vocalizations. *J Acoust Soc Am* 114:3412–3422. [CrossRef Medline](#)
- Marler P (1970) Birdsong and speech development: could there be parallels? *Am Sci* 58:669–673. [Medline](#)
- Nowicki S, Marler P (1988) How do birds sing? *Music Perception* 5:391–426. [CrossRef](#)
- Panksepp JB, Jochman KA, Kim JU, Koy JJ, Wilson ED, Chen Q, Wilson CR, Lahvis GP (2007) Affiliative behavior, ultrasonic communication and social reward are influenced by genetic variation in adolescent mice. *PLoS One* 2:e351. [CrossRef Medline](#)
- Portfors CV (2007) Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J Am Assoc Lab Anim Sci* 46:28–34. [Medline](#)
- Price PH (1979) Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93:260–277. [CrossRef](#)
- Scattoni ML, Gandhi SU, Ricceri L, Crawley JN (2008) Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism. *PLoS One* 3:e3067. [CrossRef Medline](#)
- Schmeisser MJ, Ey E, Wegener S, Bockmann J, Stempel AV, Kuebler A, Jansen AL, Udvardi PT, Shibani E, Spilker C, Balschun D, Skryabin BV, Dieck St, Smalla KH, Montag D, Leblond CS, Faure P, Torquet N, Le Sourd AM, Toro R, et al. (2012) Autistic-like behaviours and hyperactivity in mice lacking *prosp1/Shank2*. *Nature* 486:256–260. [CrossRef Medline](#)
- Simes RJ (1986) An improved bonferroni procedure for multiple tests of significance. *Biometrika* 73:751–754. [CrossRef](#)
- Srivastava DP, Jones KA, Woolfrey KM, Burgdorf J, Russell TA, Kalmbach A, Lee H, Yang C, Bradberry MM, Wokosin D, Moskal JR, Casanova MF, Waters J, Penzes P (2012) Social, communication, and cortical structural impairments in *epac2*-deficient mice. *J Neurosci* 32:11864–11878. [CrossRef Medline](#)
- Takahashi K, Kamiya K, Uruse K, Suga M, Takizawa T, Mori H, Yoshikawa Y, Ichimura K, Kuida K, Momoi T (2001) Caspase-3-deficiency induces hyperplasia of supporting cells and degeneration of sensory cells resulting in the hearing loss. *Brain Res* 894:359–367. [CrossRef Medline](#)
- Tchernichovski O, Lints T, Mitra PP, Nottebohm F (1999) Vocal imitation in zebra finches is inversely related to model abundance. *Proc Natl Acad Sci U S A* 96:12901–12904. [CrossRef Medline](#)
- Tong L, Hume C, Palmiter R, Rubel EW (2011) Ablation of mouse cochlea hair cells by activating the human diphtheria toxin receptor (*DTR*) gene targeted to the *pou4f3* locus. *Proc Assoc Res Otolaryngol*. Abstract 836.
- Willott JF (1986) Effects of aging, hearing loss, and anatomical location on thresholds of inferior colliculus neurons in C57BL/6 and CBA mice. *J Neurophysiol* 56:391–408. [Medline](#)
- Willott JF (1991) Central physiological correlates of ageing and presbycusis in mice. *Acta Otolaryngol Suppl* 476:153–156. [Medline](#)
- Wöhr M, Rouillet FI, Hung AY, Sheng M, Crawley JN (2011) Communication impairments in mice lacking Shank1: reduced levels of ultrasonic vocalizations and scent marking behavior. *PLoS One* 6:e20631. [CrossRef Medline](#)
- Zheng QY, Johnson KR (2001) Hearing loss associated with the modifier of deaf waddler (*mdfw*) locus corresponds with age-related hearing loss in 12 inbred strains of mice. *Hearing Res* 154:45–53. [CrossRef Medline](#)
- Zheng QY, Johnson KR, Erway LC (1999) Assessment of hearing in 80 inbred strains of mice by ABR threshold analyses. *Hearing Res* 130:94–107. [CrossRef Medline](#)