



Title	Optimization of spectral-domain optical coherence tomography with a supercontinuum source for <i>in vivo</i> motion detection of low reflective outer hair cells in guinea pig cochleae
Author(s)	Nin, Fumiaki; Choi, Samuel; Ota, Takeru et al.
Citation	Optical Review. 2021, 28(2), p. 239-254
Version Type	VoR
URL	https://hdl.handle.net/11094/93413
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka



Optimization of spectral-domain optical coherence tomography with a supercontinuum source for in vivo motion detection of low reflective outer hair cells in guinea pig cochleae

Fumiaki Nin¹ · Samuel Choi^{2,5} · Takeru Ota³ · Zhang Qi⁴ · Hiroshi Hibino^{3,5}

Received: 6 August 2020 / Accepted: 15 March 2021 / Published online: 29 March 2021
© The Author(s) 2021

Abstract

Sound evokes sub-nanoscale vibration within the sensory epithelium. The epithelium contains not only immotile cells but also contractile outer hair cells (OHCs) that actively shrink and elongate synchronously with the sound. However, the in vivo motion of OHCs has remained undetermined. The aim of this work is to perform high-resolution and -accuracy vibrometry in live guinea pigs with an SC-introduced spectral-domain optical coherence tomography system (SD-OCT). In this study, to reveal the effective contribution of SC source in the recording of the low reflective materials with the short total acquisition time, we compare the performances of the SC-introduced SD-OCT (SCSD-OCT) to that of the conventional SD-OCT. As inanimate comparison objects, we record a mirror, a piezo actuator, and glass windows. For the measurements in biological materials, we use in/ex vivo guinea pig cochleae. Our study achieved the optimization of a SD-OCT system for high-resolution in vivo vibrometry in the cochlear sensory epithelium, termed the organ of Corti, in mammalian cochlea. By introducing a supercontinuum (SC) light source and reducing the total acquisition time, we improve the axial resolution and overcome the difficulty in recording the low reflective material in the presence of biological noise. The high power of the SC source enables the system to achieve a spatial resolution of $1.72 \pm 0.00 \mu\text{m}$ on a mirror and reducing the total acquisition time contributes to the high spatial accuracy of sub-nanoscale vibrometry. Our findings reveal the vibrations at the apical/basal region of OHCs and the extracellular matrix, basilar membrane.

Keywords In vivo vibrometry · Organ of Corti · Outer hair cells · Cochlear partition · Supercontinuum · Optical coherence tomography

1 Introduction

Fumiaki Nin, Samuel Choi contributed equally to this work.

✉ Fumiaki Nin
nin@gifu-u.ac.jp

¹ Division of Biological Principles, Department of Physiology, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

² Department of Electrical and Electronics Engineering, Niigata University, 8050 Ikarashi-2, Niigata 950-2181, Japan

³ Division of Glocal Pharmacology, Department of Pharmacology, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita 565-0871, Japan

⁴ Department of Otolaryngology Head and Neck Surgery, Graduate School of Medicine and Dental Sciences, Niigata University, 757 Ichibancho, Asahimachi, Niigata 951-8510, Japan

⁵ AMED-CREST, AMED, Suita, Osaka, Japan

In the field of cochlear mechanics, quantitative observation of the vibration within the sensory epithelium is important since hair cells in the epithelium transduce sound vibration into electrical signals that are essential for audition [1–4]. In mammals, the epithelium contains various cells and structures, including inner hair cells (IHCs), outer hair cells (OHCs), basilar membrane (BM), and the reticular lamina (RL). Sound elicits vibration in all of these constituents. Among them, only OHCs cell bodies actively and periodically change length in association with the vibration [5, 6]. The unique mechanical activity of OHCs modifies the vibration in the epithelium with a feedback mechanism [7–9]. The phenomenon, in conjunction with the nonlinear mechanical properties of sensory IHCs, is considered crucial to establish the high-sensitivity and sharp-frequency tuning associated with normal hearing

[9–11]. Nevertheless, because the activity of OHCs has only been measured in isolated cells, there is still uncertainty regarding the real motion in situ [5, 6].

To examine the nano-vibrations within the epithelium, optical coherence tomography (OCT)-based vibrometry systems have been recently developed [12–25]. These OCT systems successfully analyze the tissues or cells in the epithelium [15–19, 22–26]. Though their theoretical resolutions in inanimate materials are cellular level, the difference in vibrations between the apical and basal regions of OHCs in animals *in vivo* has not yet been fully addressed. To precisely resolve these regions during the *in vivo* vibrometry, we consider two factors, axial resolution and spatial accuracy, in this study.

The reflectance of the object affects the effective optical bandwidth that is the key parameter for the axial resolution of the OCT system. The reflectance of the sensory epithelium is 0.04–0.004%, which is 200–2000 times less than that of the cornea in the human eye (~8%) [27]. To accumulate the backscattered light sufficiently with the conventional OCT vibrometry systems, the total acquisition time for one session is typically more than 1 s [13, 19, 22, 24, 25]. However, the background fluctuation arising from the physiological motion such as the heartbeat and respiration often exceeds some dozen micrometers in a cycle of more than 1 s. Because this fluctuation is significantly larger than the pixel size in an A-scan image, the long total acquisition time can smear out the image and degrade the spatial accuracy. To avoid the effect of the physiological motion and accurately detect the vibrations inside the epithelium, we reduced the total acquisition time. The reduced total acquisition time and the low reflectance of the sensory epithelium deteriorate the power of the accumulated light in the sensor. In this situation, the above conventional spectral-domain OCT (SD-OCT) system is insufficient because of the power and spectrum of the light source; the power of a super-luminescent diode (SLD) typically used in SD-OCT systems is less than 5 mW [16, 19, 22, 24, 25, 28]. Because of the Gaussian distribution of the light source spectrum, the power at both ends of the spectrum is often weaker than around the central wavelength. The effective optical bandwidth of the system is therefore narrowed, leading to the spatial resolution being degraded. We overcome this impediment by introducing a supercontinuum (SC) source into the SD-OCT system. Owing to the nonlinear optical phenomenon, it provides more power than the SLD and a smoother spectral continuum within the sensitive optical bandwidth of the sensor. These modifications achieve high axial resolution and high spatial accuracy in imaging and vibrometry in transparent biological material even with the physiological movements. Consequently, we demonstrate the ability of our system to measure simultaneously and accurately

sub-nanometric motion in the apical and basal areas of OHCs in reference to that occurring in the BM region.

In this study, we first evaluate the performances of the SC-introduced SD-OCT system (SCSD-OCT), and apply it to the biological measurement in *ex/in vivo* guinea pigs. Lastly, we analyze the vibration data in *in vivo* guinea pigs in comparison with post-mortem condition, and discuss its physiological significance.

2 Methods

The details on the basic performances of the SCSD-OCT, sound frequencies applied to the animals, cochlear electrophysiological measurements, the laser interferometer used in this study, and *in vivo* vibrometric data are provided in Supplemental materials.

2.1 Animal preparation

All animal experiments, which were designed in accordance with the Japanese Animal Protection and Management Law, were carried out in compliance with the protocol approved by the Institutional Animal Care Committee and the President of Niigata University (Permission Number: Niigata Univ. Res. SA00407). Hartley guinea pigs (200–300 g, 2–4 weeks of age; SLC Inc., Hamamatsu, Japan) were housed at the animal facility and kept in a 12-h light and 12-h dark cycle. Food and water were provided *ad libitum*. All animal handling and reporting complied with the animal research: reporting of *in vivo* experiments (ARRIVE) guidelines [29]. For *ex vivo* preparation, all animals were anesthetized by intraperitoneal injection of sodium pentobarbital (1 ml per guinea pig; Somnopentyl; Kyoritsu, Japan). Once deep anesthesia was confirmed, the left cochlea was surgically exposed, and then dissected from the temporal bone. It was also washed in the chilled standard phosphate-buffered saline (PBS) consisting of NaCl, 137 mM; Na₂HPO₄, 8.1 mM; KCl, 2.68 mM; KH₂PO₄, 1.47 mM at pH 7.4. For *in vivo* preparation, we first intramuscularly injected atropine sulfate (0.05 mg/kg) and Midazolam (4 mg/kg) into the animals as a premedication. Animals were then deeply anaesthetized with an intra-peritoneal injection of urethane (1.5 g/kg). Supplemental doses of anesthesia were administered to ensure areflexia to toe pinches. After tracheotomy, the animal was artificially ventilated with room air using a respirator (SN-408-7; Shinano Manufacturing, Japan). Surgical access to the cochlea was performed using a similar method as previously described for biophysical measurements [19, 22, 30]. The pinna was removed along with the muscles covering the bulla. The ventral and anterior portions of the bulla were surgically opened using a ventrolateral approach to expose the round window in the basal turn of

the cochlea. The guinea pig's head was fixed on the acrylic plate, which was tightly connected to a base stage (SL20/M; Thorlabs, USA). The animal's body was positioned on a heating pad atop a movable platform under the OCT system. The platform could move planarly with a housed speaker (T250D, Fostex, Japan), calibrated with an ultrasound microphone (Sokolich ultrasonic probe microphone system, US). The tip of the speaker was inserted into the left ear canal of the guinea pig. To monitor the auditory function of the live animal, we measured the auditory brainstem response (ABR) thresholds before and after the recording at four different frequencies of 7000, 14,000, 21,000, and 28,000 Hz. The experiment including surgery was completed within 4 h.

The total number of animals used in this study was 87. Among them, four animals were used for ex vivo experiments, four for evaluation of the effectiveness of the power of the SC light source for in vivo measurements, twelve for the irradiation effect on ABR, and one for recording the vibration tuning curve with a laser interferometer. The other 66 animals underwent the surgery described above for in vivo recordings by means of SD-OCT vibrometry. In 34 animals, the preparation was successful and data were acquired prior to and 30 min after anoxia. Eleven out of 34 preparations showed minimal ABR threshold changes during the experiment; if the ABR thresholds changed more than 30 dB in any frequencies, we eliminated the data from the analysis. Of the 11 sensitive cochleae, the vibration amplitude of four exhibited a clear nonlinear response to the sound stimulation at BM or RL, a hallmark for physiological reaction of the sensory epithelium. Four samples with a clear nonlinear response were selected for this study, because the cochleae without the nonlinear response were likely to be damaged by surgery or anesthesia. Accordingly, we analyzed the limited number of samples that strictly met the ABR threshold criteria as stated above and the active vibration of the epithelium as the other groups adopted [16, 18, 20, 31, 32]. Although some groups have worked on this type of in vivo measurement in various animals, the output has been very low [16, 20, 31–34]. The invasive surgery, low reflectance of the epithelia, vulnerability of sensitive hearing, and time-consuming data collection are known to be the main reasons for this outcome. In our study, the number of samples is almost equivalent to that in all the previous studies [16, 18, 20, 31, 32].

2.2 The basic performance of the SC-introduced SD-OCT

Figure 1a details the SCSD-OCT system that we constructed for this study. For this setup, we used the spectrometer and scanning system in a commercial SD-OCT (Ganymede GAN520; Thorlabs, US) as well as an SC source (SuperK EXT-20; NKT Photonics, Denmark). The light source was

connected to an optical bandpass filter (SuperK SPLIT; NKT Photonics, Denmark), a fiber delivery system (SuperK CONNECT; NKT Photonics, Denmark), and a single-mode broadband fiber (FD7; NKT Photonics, Denmark). We used a custom objective lens (a modified M Plan Apo NIR 10 \times , Mitutoyo, Japan), with improved transmission at the wavelength employed, a long focal length of 30.5 mm, and a depth of focus of 4.1 μ m. It had improved transmission at the wavelengths employed in this study. The average light power onto the sample was 117 mW. The effective bandwidth and central wavelength in the spectrometer were approximately 200 nm and 900 nm, respectively (Table 1).

3 Results

3.1 The performance of the SC-introduced SD-OCT in low reflective objects

To analyze the interference signal acquired in the spectrometer and evaluate the axial resolution, we tested the image of a fused silica glass of $50 \pm 0.01 \mu\text{m}$ thick (Atock, Japan) (Fig. 1b). In a B-scan image of the glass, a line was selected for the A-scan. The interference signal intensity was temporally averaged over 10,000 A-scans, and the result is shown in Fig. 1c. The axial pixel size calculated from the number of pixels between the interference peaks was $1.25 \pm 0.00 \mu\text{m}$ in the glass. Because the refractive index of glass and air is 1.45 and 1, respectively, the pixel size in air was $0.86 \pm 0.00 \mu\text{m}$. In the recording of the sensory epithelium, we estimated the refractive index to be 1.35; therefore, the pixel size is $1.16 \pm 0.00 \mu\text{m}$. The axial resolutions, which were calculated from the full width at half maximum (FWHM), of the SCSD-OCT were $1.72 \pm 0.00 \mu\text{m}$ and $2.73 \pm 0.00 \mu\text{m}$ in air and glass, respectively (see Supplementary text 1.1, and Fig. S2 (c)). We also assessed the lateral resolution by discriminating separate lines in an OCT image on a resolution test target (Edmund optics: USAF 1951 target) (Fig. 1d). Because the peak intensity of the OCT signal was over 5 dB higher than the background in the region of 8-1 (Fig. 1e), the lateral resolution was $1.95 \mu\text{m}$. The performance of the SCSD-OCT is summarized in Table 1, and its detailed description and comparison of the performance to a SLD-introduced SD-OCT systems are described in Supplemental materials.

3.2 Doppler vibrometry

To analyze the vibration inside the sensory epithelium, we collected three parameters, that is, intensity signal of the epithelium, vibration amplitude, and phase. Before collecting such information, we first selected a line of interest from a B-scan image. Then, we performed a spectral

Fig. 1 Depiction and performances of the supercontinuum-introduced spectral-domain optical coherence tomography (SCSD-OCT) system. **a** A supercontinuum (SC) source was used in the recording of the inanimate and biological materials. **b** 2D cross-sectional tomographic image of air/glass/air layers. From the top, the layers are air, glass of $50 \pm 0.01 \mu\text{m}$ thickness, and air. The line recorded from this is highlighted by the red line. **c** Signal intensity (arbitrary unit) on linear scale. **d** Upper panel shows the CCD view of a test target. The selected line in lower panel is highlighted by yellow. The dotted yellow line in lower panel indicates the line shown in **d**. **e** OCT signal intensity on logarithmic scale. 8-1 and 9-1 ~ 3 show the group number and element of the test target (color figure online)

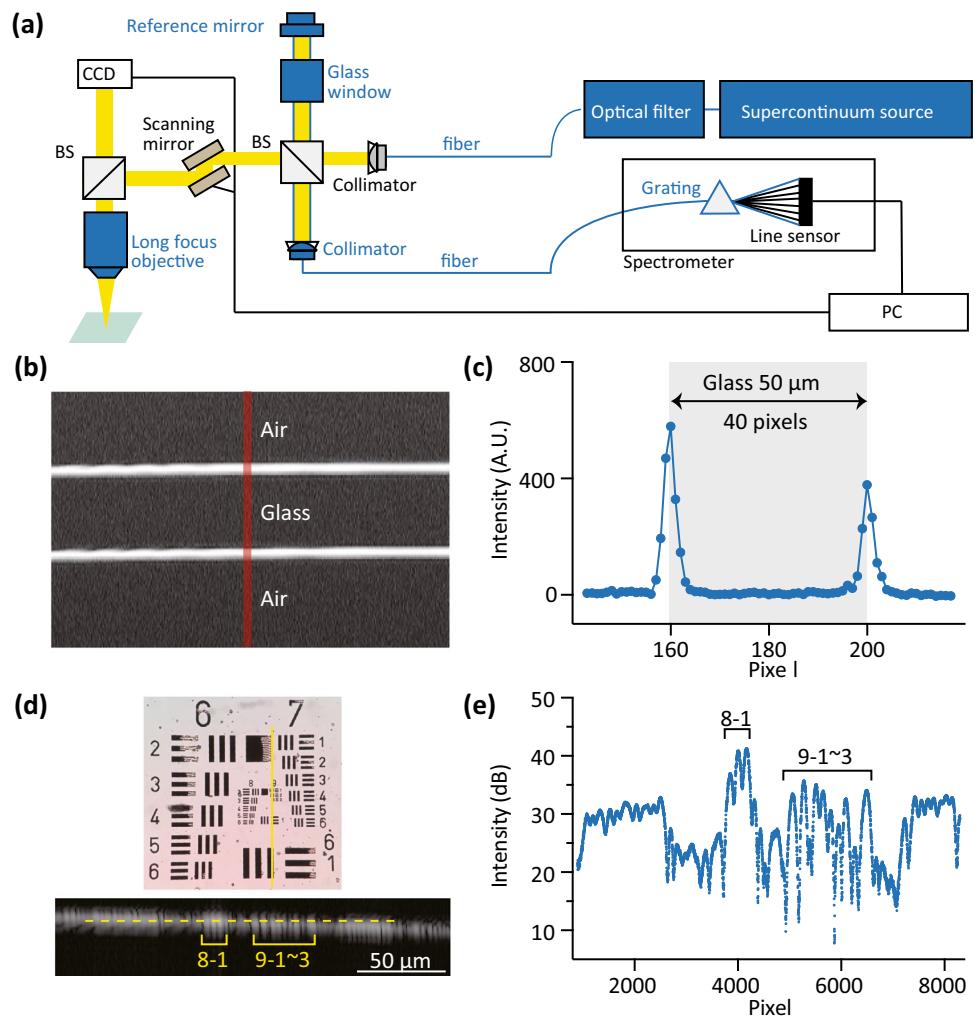


Table 1 Performance on various reflective objects of the supercontinuum-introduced SD-OCT (SCSD-OCT) system

OCT system	SCSD-OCT
Central wavelength	900 nm
Effective bandwidth	200 nm
Sampling frequency	~ 100,000 Hz
z-axis pixel size (in air)	$0.86 \pm 0.00 \mu\text{m}$
z-axis pixel size (in glass)	$1.25 \pm 0.00 \mu\text{m}$
z-axis pixel size (in tissue)	$1.16 \pm 0.00 \mu\text{m}$
z-axis resolution (in air)	$1.72 \pm 0.00 \mu\text{m}$
z-axis resolution (in glass)	$2.73 \pm 0.00 \mu\text{m}$
z-axis resolution (in tissue)	~ 3 μm (estimated)
xy-axis resolution	1.95 μm
Power on sample	117 mW
Power on round window	88.5 mW
Power on epithelium	47.6 mW

Fourier transform over the wavenumber on the interference signal using ThorImageOCT (Thorlabs, US) software. The software applied a Hamming window in the processing, and captured the time-domain interference signal of 10,000 scans in one recording session. In terms of the intensity signal, which reflects the structural image of the epithelium, the custom-made LabVIEW program temporally averaged the assembled intensity of the time-domain interference signal. Moreover, the program also performed a spectral Fourier transform over the captured phase of the time-domain interference signal. The magnitude of the spectral Fourier transform was translated into the vibration amplitude, whereas the phase of the transform corresponded to the vibration phase. When a sample vibrates, backscattered light from a position z' within the sample alters the optical path in the sample arm over time t . This alteration induces a change in the phase of the time-domain interference signal, $\varphi(z', t)$. Assuming the sample vibration is composed of a single frequency f , with an initial phase θ , $\varphi(z', t)$ becomes:

$$\varphi(z', t) = A(z') \sin(2\pi ft + \theta) + \xi(t), \quad (1)$$

where $A(z')$ is the amplitude of a sinusoidal vibration sample, and $\xi(t)$ is the phase noise of the interference signal mainly due to physiological motion such as heartbeat, respiration, and muscle contractions. The amplitude $A(z')$ was converted into units of nanometers by multiplying by $\lambda/(4\pi n)$; the actual amplitude of the vibration is calculated as $A(z') \times \lambda/(4\pi n)$. For air and intracochlear soft tissue, we assumed n to be 1.0 and 1.35, respectively [35]. To smooth the distribution of the amplitude and phase along the depth, a triangular moving average filter with a half-window of 3 pixels was applied in the spatial domain.

To control sample vibration and its recording, we prepared two sinusoidal voltages using a function generator (WF1948, NF Corporation, Japan) with specific amplitudes, frequencies, and delay. One voltage drove a piezo or speaker for the stimulation, and the other voltage served as an external trigger input of the OCT device to control the sampling rate. To detect an amplitude and phase in steady-state vibrations, we started the latter voltage with a 200 ms delay to the former voltage. Accordingly, no recording data included the information in a transient response of vibration. Notably,

the phase in vibration was referenced to the phase of the sinusoidal voltage for the sample. To avoid the micrometer-scale physiological motion during in vivo vibrometry, we reduced the total acquisition time for one session. Therefore, in this study, all the total acquisition times were 100 ms except for M-scan imaging of 1 s (see Sect. 3, 4). During the 100 ms, the software collected 10,000 A-scans in one session because the sampling rate of the line sensor was 100,000 Hz.

The frequency response of SCSD-OCT was assessed using a piezo-electric actuator (PZT). Figure 2a shows a B-scan image of PZT. This target was sinusoidally moved by providing voltages of 0.5, 1.0, and 10 V at stimulus frequencies between 5000 and 30,000 Hz with 1000 Hz increments (Fig. 2b). To capture many cycles in a recording of low-frequency vibration for accuracy, the sampling frequency for stimuli of less than 15,000 Hz was 40,000 Hz, whereas it was 100,000 Hz for stimuli of more than 16,000 Hz. We, therefore, used a 100 ms total acquisition time for the sample's vibration frequencies from 16,000 to 30,000 Hz, and a 250 ms total acquisition time for those from 5000 to 15,000 Hz. In both situations, the number of scans collected in one session was 10,000. The amplitude was collected

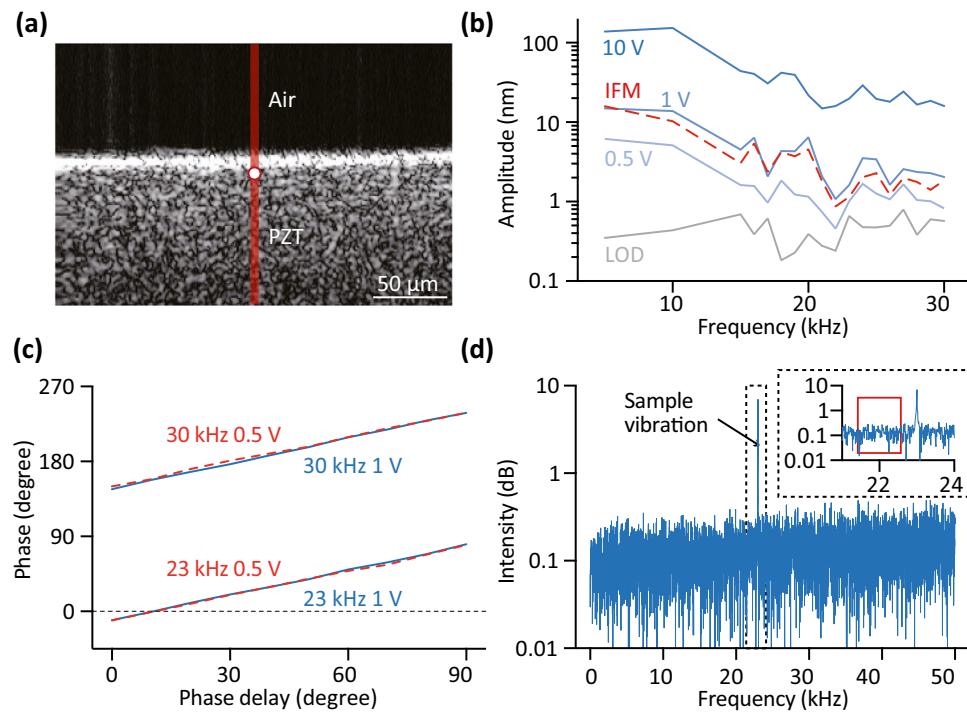


Fig. 2 Performance of the supercontinuum spectral-domain optical coherence tomography (SCSD-OCT) system during vibrometry. **a** 2D tomographic image of the surface of a piezo-electric actuator (PZT). The selected line is highlighted in red. The red open circle indicates the recording point on the PZT. **b** Acquired amplitudes of a vibrating PZT with a SCSD-OCT (blue) and a laser interferometer (IFM) (dotted red). The input voltage was varied from 0.5 to 10 V. **c** Phases

of the vibrating PZT at 23,000 and 30,000 Hz stimuli, and 0.5 V (dotted red) and 1 V input voltage (blue). **d** The vibration amplitude of PZT using a 23,000 Hz stimulus, 100 ms duration, and 10 V input voltage. The limit of detection (LOD) was calculated from the signals from the red squared region in the inset and was plotted with a gray line in **b** (color figure online)

from the location of the PZT surface (Fig. 2a, open red circle). The detected amplitude increased linearly with the strength of the applied voltage, and the values in 1 V stimuli were consistent with those acquired by laser interferometry. The detailed system of the laser interferometer is described in the Supplemental materials [30]. The average measurement error between the two methods was less than ~ 0.72 nm. To ensure that our SCSD-OCT system permitted the phase measurement, we drove the PZT with sinusoidal stimuli at both 23,000 Hz and 30,000 Hz frequencies with different phase delays. The delays were varied from 0° to 90° in 10° increments. We then compared the vibration phase to the delay. Figure 2c displays the delay versus the measured phase of the vibrating PZT. The measured phase grew in parallel with the shift of the delay. Figure 2d indicates the fast Fourier transform (FFT) vibration amplitude measured from the surface of the vibrating PZT at 23,000 Hz with 10 V stimuli. We evaluated a limit of detection (LOD) value of vibration amplitude in the frequency domain. The vibration amplitude in all the frequency was derived from the change in phase of the interference signal, $\varphi(z', t)$ as we mentioned above (Eq. 1). LOD was defined by the noise floor (NF) and standard deviation (SD) as $\text{LOD} = \text{NF} + 5\text{SD}$. To calculate NF and SD, 10,000 scans of the vibration amplitudes in the frequencies from $f=1500$ Hz to $f=500$ Hz were used (Fig. 2d, red square). NF is the noise floor, i.e., averaged value, and

SD is the standard deviation. f is the frequency of the vibration of the target sample. We adopted this criterion for all recordings. Under the PZT vibration in a 24,000 Hz stimulus with 0.5 V voltage, the LOD was 477 pm (Fig. 2b, gray line).

3.3 Ex vivo imaging of isolated cochlear partition

During in/ex vivo recordings, we limited the power of the SC source to 429.0 mW. In this situation, the power of the beam onto the sample was 117 mW, almost 10 times higher than that of 3–16 mW in previous OCT systems [12–18, 20–25, 36–39]. In addition to the object reflectance and short total acquisition time, we had two reasons for the necessity of the high power irradiation in the in vivo recordings of guinea pig cochleae from the anatomical perspective: mechanical vignetting and reflection of the beam. These factors were likely to be initiated by two characteristic structures inside the temporal bone. First, compared to the cochleae of the mouse and gerbil, the organ of guinea pigs is located deeper inside the bone (Fig. 3a) [19, 24, 31]. The beam was likely to be mechanically vignetted because of the narrow path in mastoid air cells (Fig. 3b). Second, the angle of the round window membrane in guinea pig cochleae is steeper toward the surface of the cochlear partition (Fig. 3c). To demonstrate these effects on the power of the beam, we monitored the light strengths on the pathway through the temporal bone

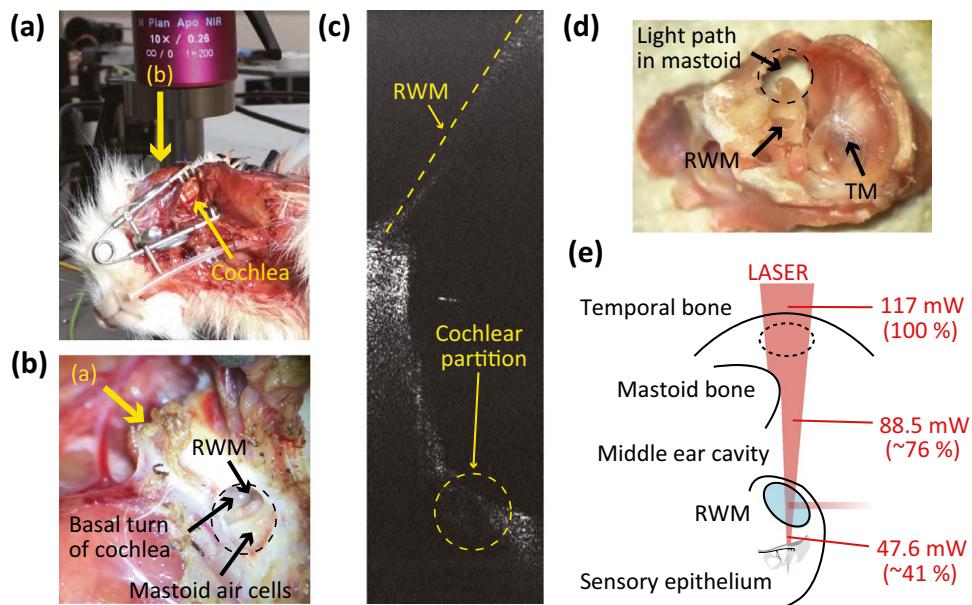


Fig. 3 Effect of the mechanical vignetting and reflection on the beam by the anatomical structures in a guinea pig. **a** A lateral view of the guinea pig cochlea during an in vivo vibration recording. The thick yellow arrow indicates the direction of the view shown in **b**. **b** A view of the basal turn and round window membrane of a guinea pig cochlea through the mastoid air cells. The thick yellow arrow indicates the direction of the view shown in **a**. **c** Optical coherence tomography

(OCT) image of the round window membrane (RWM) and cochlear partition in a live animal. **d** An inside view of the middle ear cavity in a dissected temporal bone. Cochlea bony wall except for around the RWM is removed, and a light path is made on the bulla. **e** Schematic of the beam power between the surface of the temporal bone and the sensory epithelium

and round window to the sensory epithelium in a dissected temporal bone (Fig. 3d). Almost 24% of the total power on the animal's temporal bone was vignetted by the disturbance of the mastoid bone, whereas ~35% was scattered by the round window membrane (Fig. 3e). These results indicate that the light power on the epithelium by SLDSD-OCT can be estimated to be ~0.29 mW (Table S1). Accordingly, compared to that in other species [19, 24, 25], averaging 4,000,000 scans with a total acquisition time of ~27 s was likely to be required in guinea pigs [22].

Prior to in vivo analysis of the sensory epithelium, we next recorded an image of the ex vivo tissue using SCSD-OCT. After dissection of the cochlea, we positioned the sample upside down on a phosphate-buffered saline (PBS) filled petri dish and exposed the epithelium by removing a round window membrane. In the schematic of the epithelium, the outer tunnel and tunnel of Corti are highlighted with a dotted pink circle and a triangle, respectively (Fig. 4a). These are major landmarks in the epithelium. The apical and basal regions of OHCs are highlighted in blue and purple, respectively. Figure 4b shows a camera image of the view including the sensory epithelium. A region of approximately 300 μm width in the sensory epithelium (red line) was subjected to a

B-scan, and the unaveraged image is displayed in Fig. 4c. In this image, we identified the outer tunnel, the tunnel of Corti, and OHCs based on the schematic in Fig. 4a and literature [12, 15, 17, 20, 22, 38]. The OHC diameter and length at the basal turn of the guinea pig cochlea are ~7 μm and ~40 μm , respectively [16], whereas the ideal lateral resolution of our system was 1.95 μm . Nevertheless, the system distinguished not the three rows but the apical and basal regions of OHCs, and displayed OHCs as a cluster (Fig. 4c yellow dotted rectangle) [12, 17, 22]. This inconsistency may be due to the strong speckle initiated in the OCT image by SC source. To further identify the apical and basal regions of OHCs, we held the position of the beam on the center of the OHCs image (Fig. 4c, red arrowheads). Figure 4d shows the intensity distribution averaged over 10,000 A-scans. The x-axis is the axial location, upper to the left and lower to the right. The cochlear partition was visualized above the background signal in lymph fluids with a signal-to-noise ratio (SNR) of more than 5 dB (Fig. 4d, bidirectional arrow). The BM directly faces the perilymph, as such in Fig. 4d the contrast neighboring the weak signal to this fluid (pixels 211–220), mirrored the border zone of the BM [13–20, 22–24, 36, 37]. In this context, the peak observed around pixel 220

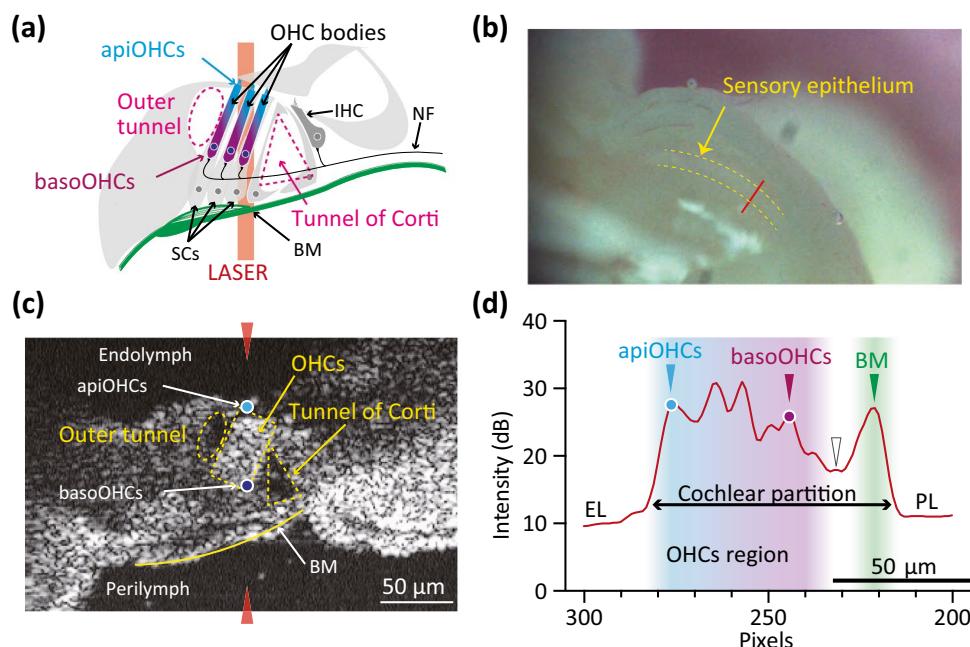


Fig. 4 Ex vivo imaging of the cochlea partition isolated from a guinea pig. **a** Schematics of a cochlear partition. Red lines show the direction of the laser. **b** A view of the guinea pig epithelium in the recording. Red line indicates the scanned line. **c** Optical coherence tomography (OCT) image of the epithelium in the basal turn of the cochlea with indicative landmarks. Outer tunnel, tunnel of Corti, and a cluster of outer hair cell (OHCs) are highlighted in a dotted yellow circle, triangle, and rectangle. Red arrowheads indicate the selected line shown in **d**. **d** Intensity collected in the line was plotted on a

logarithmic scale. Top of the epithelium lies to the left. In this and subsequent illustrations, blue and purple shading indicates presumed OHCs region, and green shading corresponds to BM. Bidirectional arrow, blue, purple, green, and open arrowheads exhibit cochlear partition, apiOHCs, basoOHCs, BM, and supporting cells region. *RW* round window, *IHC* inner hair cell, *OHC* outer hair cell, *apiOHCs* apical region of OHCs, *basoOHCs* basal region of OHCs, *SCs* supporting cells, *BM* basilar membrane, *NF* neuronal fiber, *EL* endolymph, *PL* perilymph (color figure online)

corresponded to the inside of the BM (Fig. 4d, green shading). Similarly, the other contrast around pixels 275–283 represented the border zone of the sensory epithelium, which is exposed to the endolymph. The peak at pixel 275 represented the RL containing the apical edge of OHCs, similar to observations in the literature [13–20, 22–24, 36, 37]. In addition, we observed a few contrasts inside the epithelium. Recent studies using OCT showed that OHCs provide continuous high intensity [12, 17, 22] and supporting cells elicit relatively low intensity [12, 17]. Therefore, it seems probable that, in the cochlear partition described in Fig. 4d, the area eliciting minimal intensity (pixel ~232) is the location of the supporting cells (open arrowhead). From this perspective, the region with relatively high intensities ranging from pixels 275 to 243 was likely to represent the inside of OHCs. In this regard, the moderate contrast between pixels 232 and 243 correspond to a border zone of OHCs. In summary, the region from pixels 232 to 283 was presumed to be described as the OHCs region (Fig. 4d, blue and purple shading). Note that the width of this region was 53.6 μm . This distance should result from a diagonal cross of the cluster of OHCs (one OHC: ~7 μm in diameter and ~40 μm in length) by the beam (see Fig. 4a, c). Similar results were recorded in two different dissected cochleae.

On the basis of these observations, in this and following measurements, we denoted that the peak region neighboring the contrast to the endolymph as the apical region of OHCs (i.e., apiOHCs) (Fig. 4d, blue arrowhead). The peak next to the contrast bordering the supporting cells was denoted the basal region of OHCs (i.e., basoOHCs) (Fig. 4d, purple arrowhead). From the histological viewpoint, the distance between these peaks should be roughly <50 μm [16]; in our data of Fig. 4d, this parameter was 53.6 μm . In most of the

recordings, the number of pixels representing this region exceeded 30. This feature was due to the high axial resolution of our system, and the strong advantage in distinguishing the apical and basal regions of OHCs.

3.4 In vivo measurements of cochlear partition motion

Motion effect of the live animal on the OCT signal has not yet been fully examined in the previous studies [13–22, 24–26, 28, 32, 33, 36, 37, 40]. Therefore, we evaluate the effect of the total acquisition time on the spatial accuracy during vibrometry by M-scan. The M-scan is the time-motion display of the OCT image along a chosen line. Figure 5a displays the OCT images of the cochlear partition including the sensory epithelium. We selected a line on the OCT image and then recorded the M-scan image for 1 s without sound stimuli as shown in Fig. 5b. The sampling rate was 10,000 Hz. Owing, not to the sound, but to the background fluctuation arising from the physiological motion, recording position seemed to be shifted laterally. This fluctuation seemed to decrease the thickness of the cochlear partition by more than 10 pixels (Fig. 5b). Figure 5c–e show the intensity distributions of the partition along the depth axis. The distance between two peaks at the edge of the partition at 30 ms was similar to those at 130 ms (Fig. 5c, d). By contrast, the distance at 700 ms significantly shrank (Fig. 5e) this deformation seems to arise from the micrometer-scale physiological motion in lateral and/or axial directions. In vibrometry with conventional OCT, the total acquisition time was more than 1 s. Therefore, the physiological motion smears out the averaged intensity distribution, resulting in loss of spatial

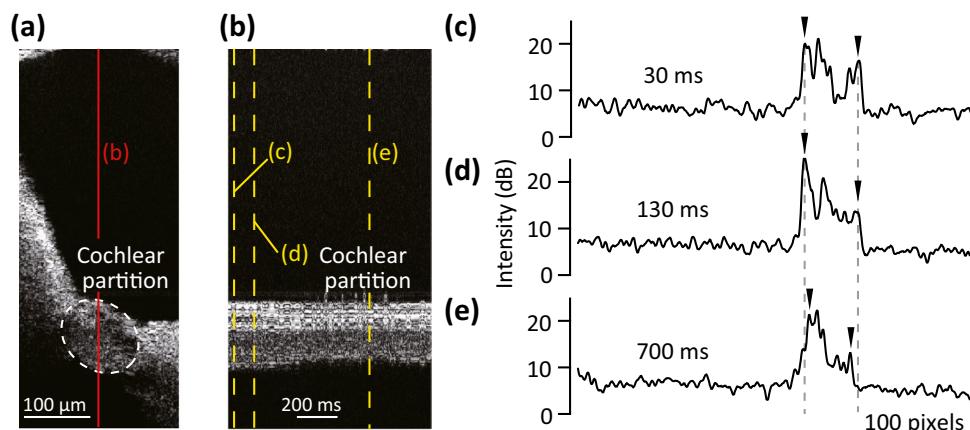


Fig. 5 In vivo OCT and M-scan imaging of the cochlea partition without sound stimuli in a live guinea pig. **a** OCT image of the partition. Red line indicates the selected line for **b**. **b** M-scan image of the line in **a**. The yellow dotted lines are the times of the A-scan shown

in **c–e**. **c–e** Intensity distribution in OCT image at different times. **c–e** show the intensities in the time of 30 ms, 130 ms, and 700 ms, respectively. Filled arrowheads indicate the peaks at the edge of the partition (color figure online)

accuracy. To avoid this effect, we set the total acquisition time to 100 ms in this study. Of note, in this series of experiments, the image of the cochlear cross section was obtained through the intact round window in a live guinea pig. The data were taken from a location similar to that of Fig. 4.

3.5 In vivo measurements of the motion in apical/basal region of OHCs and BM

We also examined the effectiveness of the SC light source power on the OCT images and vibration measurements in vivo. The data were recorded by varying the estimated

light intensities on the epithelium (2.0–47.6 mW) (Fig. 6). Figure 6a shows the OCT images of the sensory epithelium. Under 2.0 mW intensity, the epithelium was insufficiently visualized and when the intensity increased to 4.1 mW, an outline of the epithelium appeared. The cluster of OHCs, however, was not clearly distinguishable. At an intensity greater than 14.2 mW, structures inside the epithelium were observed. To further test the vibrometry data, we applied 24,000 Hz sound with 90 dB sound pressure level (SPL) to the animal. Figure 6b–e presents data on SNR in the signal intensity, the noise floor (NF), amplitude, and phase of the vibration under sound stimulation of 24,000 Hz at 90 dB. These data were averaged over 10,000 A-scans acquired in

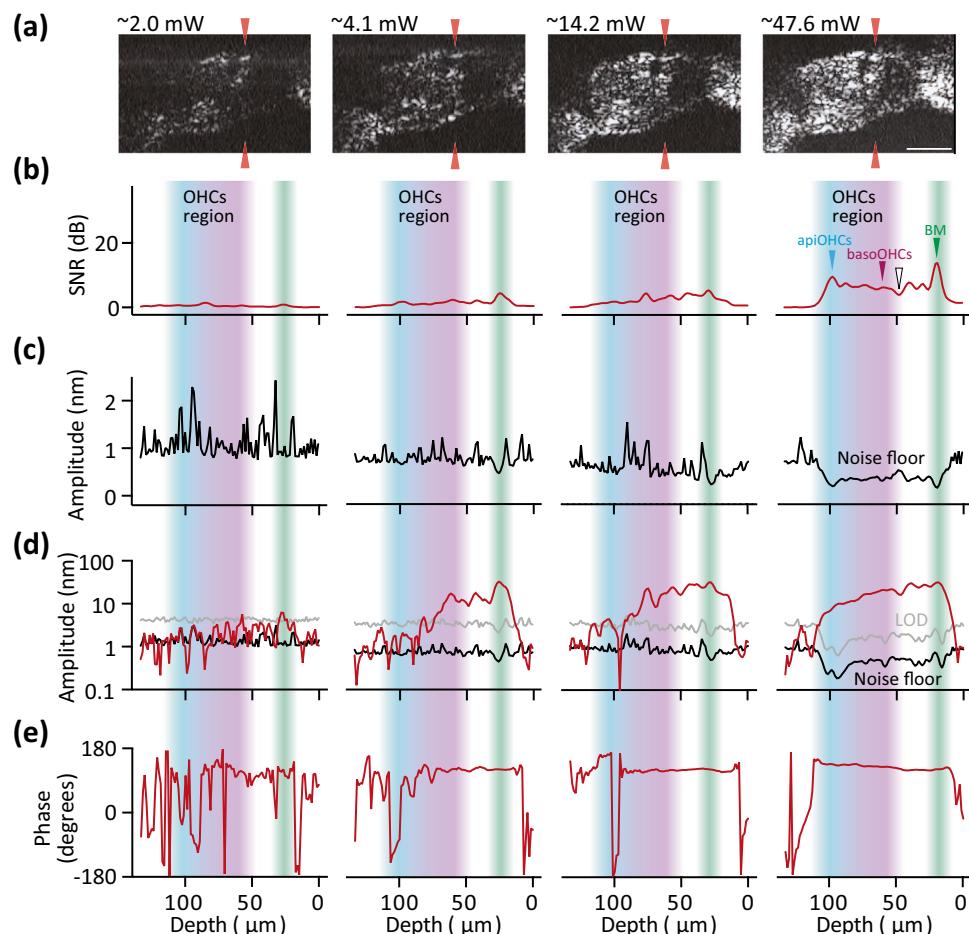


Fig. 6 Analyses of the sensory epithelium at varying light powers with the optical coherence tomography (OCT) system. **a** OCT images of the epithelium in the basal turn of the cochlea through the round window membrane. Red arrowheads show the selected line in **b–e**. Scale bar in the right panel is 50 μ m. **b** Signal-to-noise ratio (SNR) in the intensities of the averaged A-scan image during the sound stimulation. Blue, purple, green, and open arrowheads exhibit cochlear partition, apiOHCs, basoOHCs, the basilar membrane (BM), and supporting cells region. Blue and purple shadings represent OHCs region, which is defined in the data under the light power of 47.6 mW. The intensity is referenced to the data from inside the fluid

space above the epithelium, the endolymph. **c** Noise floor (NF) of the vibration amplitude on a linear scale. As power increases, the NF decreases. **d** NF, the limit of detection (LOD), and vibration amplitudes in response to pure-tone stimulation at 24,000 Hz with 90 dB on a logarithmic scale. Black and gray lines indicate NF and LOD ($LOD = NF + 5SD$). **e** Responses of the vibration phase in the epithelium. As the power is gained, phase fluctuation is suppressed due to the improvement in SNR. Blue, purple, green, and open arrowheads exhibit apiOHCs, basoOHCs, basilar membrane, and supporting cell regions (color figure online)

one recording session. Based on the criteria described in Fig. 4d, at the 47.6 mW measurement, we detected the two SNR peaks in the region of OHCs (Fig. 6b, light blue and purple arrowheads). One was neighboring the endolymph, and another was next to supporting cells. Because these landmarks corresponded to the border zones of the OHCs, these SNR peaks in the zone were denoted as the apiOHCs and basoOHCs. Note that SNR in the OHCs region was significantly enhanced by ~ 10 dB at 47.6 mW compared to the corresponding region at 4.1 mW (Fig. 6b). A similar improvement was also observed in NF; in the majority of the region inside the OHCs, the value at 47.6 mW was half of that at 4.1 mW (Fig. 6c). Because of this low NF at 47.6 mW, all the amplitude's data points between apiOHCs and basoOHCs exceeded LOD ($LOD = NF + 5SD$) (Fig. 6d, red line). The spatial distribution pattern of these data points was relatively smooth. In contrast, the amplitude data was unstable along the depth axis at the light intensity of 14.2 mW, and significantly attenuated at intensities less than 4.1 mW. These impairments may be attributed to high NF arising from the low SNR (Fig. 6b, c), which increased the LOD values to approximately 4 and 5 nm at 14.2 and 2.0 mW, respectively (Fig. 6d, gray line). In terms of the vibration phase, the values inside the region where the vibration amplitude satisfied the LOD were almost flat whereas those in the region where the amplitude was less than the LOD had fluctuated (Fig. 6e). Compared to the LOD, vibration amplitudes were spatially significant at 47.6 mW. Therefore, we chose this intensity and LOD in subsequent *in vivo* assays.

A different guinea pig was tested further, and the OCT image and vibration data obtained under the control and post-mortem conditions are shown in Fig. 7. Figure 7a summarizes the schematics of the beam, applied sound, and evoked vibration. The range scanned in the epithelium was 300 μ m (Fig. 7b). The focus of the objective lens was positioned at the depth of the OHCs. On the basis of the structural information from *ex vivo* experiments, we identified common landmarks mentioned in Fig. 4 (Fig. S4 (a) and (b)). We also detected the tectorial membrane, and collagen-fiber layer of the BM (Fig. 7b green and yellow arrowheads). The additional peak inside the endolymphatic region and the dip in the bimodal peak of the BM signal reflect the tectorial membrane and collagen-fiber layer of BM, respectively (Fig. 7c, green and yellow arrowheads). These observations indicate that the spatial resolution of our system was higher than that of conventional systems [15, 17–19, 21, 22, 25]. For vibrometry, the scanning beam was held at the location on the OHCs (Fig. 7b, red arrowheads). The signal intensity, amplitude, and phase of the vibration were then acquired while stimulating the ear canal with a pure-tone sound. The sound frequency was 23,000 Hz, and SPLs ranged from 50 to 90 dB. In mammalian cochleae, biomechanical properties of the organ of Corti tune each region to a specific frequency

of sound; the base of the cochlea resonates at higher frequencies, and the apex at lower frequencies. Because our recording point observed through the round window was located in the base, the high-frequency sound elicited a larger vibration in the epithelium than low-frequency sound. The sound frequency that maximally vibrates the epithelium with small SPLs at the recording point is known as the best frequency (BF) [33]. In BF and near-BF, the amplitude response to different SPLs exhibits compressive growth [8]. According to the frequency response of the epithelial vibration recorded with an interferometer, we noted the frequencies from 22,000 to 24,000 Hz as near-BF and those of $< 22,000$ Hz as sub-BF throughout this study as previously reported [30]. In this trial, we chose 23,000 Hz as a near-BF stimulation.

The axial M-scan data recorded from the aforementioned regions were analyzed with different stimulus intensities and the results are displayed in Fig. 7c–e. Figure 7c shows the signal intensity on the axial M-scan that included OHC signal in the recording. The amplitude and phase of vibrations are presented in Fig. 7d, e, respectively. The signal intensity was obtained from a temporal average of 10,000 A-scans over the total acquisition time, as opposed to a single A-scan data selected from a B-scan. The intensities of apiOHCs, basoOHCs, and BM can be visualized above the background signal inside the lymph fluids by over 20 dB, whereas the intensity was minimum at the location of the supporting cells (Fig. 7c, open arrowheads). These observations were consistent with those in the dissected cochlea (see Fig. 4d). In addition, under both control and post-mortem conditions, there was little detected variation in signal intensity throughout the recording despite changes in the sound pressure. This result implies that there was a significant spatial accuracy in vibrometry. In other words, the beam was negligibly moved from the initial target onto the different cells or structures during measurement. The behavior of the epithelial vibrations is shown in Fig. 7d, e. Of note, in conventional OCT systems, although the apiOHCs and basoOHCs are separately visualized, vibrations in these regions have not yet been fully analyzed [15–19, 22, 23, 36, 37]. Under the acoustic stimuli of ≤ 70 dB SPL, the vibration amplitudes of apiOHCs and BM were larger than those of basoOHCs. For intense stimuli of ≥ 80 dB SPL, the amplitudes of apiOHCs, basoOHCs, and BM were all similar (Fig. 7d). Figure 7e illustrates the phases in the epithelium. Below 70 dB SPL, the phase in apiOHCs and BM remarkably led the basoOHCs phase. Under intense sound, the phases were almost constant throughout the cochlear partition. Under post-mortem conditions the partition was characterized by completely different behavior. The vibration amplitudes at the basoOHCs exceeded those at the apiOHCs amplitudes, whereas those at BM were almost equal to the apiOHCs amplitudes at all stimulus intensities (Fig. 7d). In

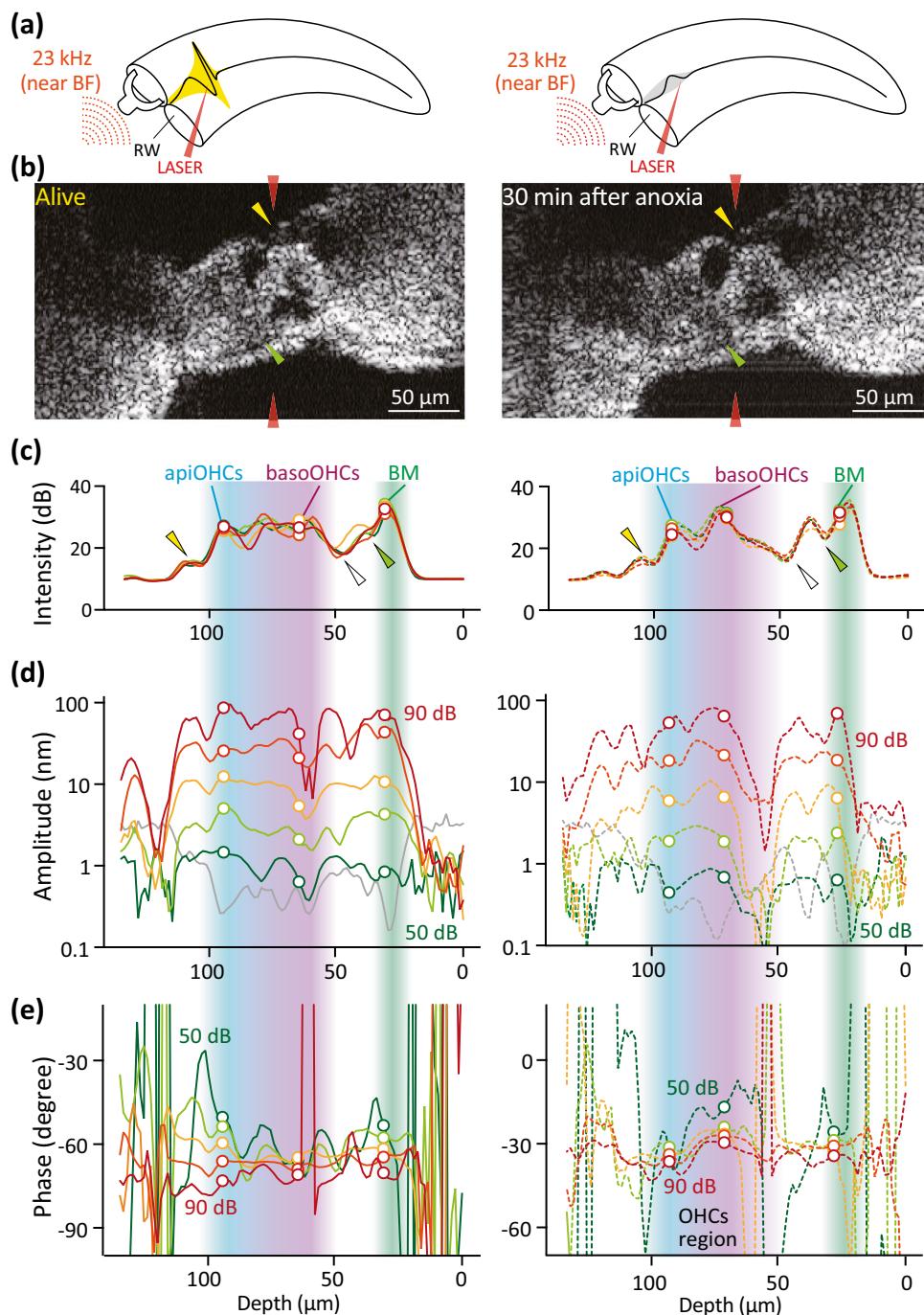


Fig. 7 Optical coherence tomography (OCT) analyses of the sensory epithelium under near-best frequency (BF) stimulation. Right and left panels show the data under physiological and post-mortem conditions in an animal, respectively. **a** Depiction of the location of recording and sound-elicited vibration. **b** OCT images of a guinea pig's epithelium in the basal turn of the cochlea. The red, green, and yellow arrowheads show the selected line for **c–e**, tectorial membrane, and collagen-fiber layer of the basilar membrane (BM), respectively. **c** The intensity of the OCT image on the selected line. The stimulation results are plotted in 10 dB increments from 50 to 90 dB SPL under physiological and post-mortem conditions. In this and subsequent illustrations, levels are arranged in a chromatic sequence. The data used for the analysis in Fig. 8 are open circles. The green, yellow and open arrowheads indicate the tectorial membrane, collagen-fiber layer of the BM and supporting cells region. **d** Vibration amplitudes inside the epithelium in response to pure-tone stimulation at 23,000 Hz. The mild convergence of the peaks toward that under 90 dB stimulation indicates that the vibrations under physiological conditions are compressive (left panel). Amplitudes measured 30 min after anoxia scale linearly with sound pressure (right panel). Gray lines indicate LOD. **e** Responses of the vibration phase inside the epithelium. The divergence of the response from positive value implies that phase leads that of the other part in the epithelium. In the phase responses under post-mortem condition, phases overlapped in various sound pressures. Phases are plotted in 10 dB increments in stimulation from 50 to 90 dB SPL under physiological and post-mortem conditions (color figure online)

cate the tectorial membrane, collagen-fiber layer of the BM and supporting cells region. **d** Vibration amplitudes inside the epithelium in response to pure-tone stimulation at 23,000 Hz. The mild convergence of the peaks toward that under 90 dB stimulation indicates that the vibrations under physiological conditions are compressive (left panel). Amplitudes measured 30 min after anoxia scale linearly with sound pressure (right panel). Gray lines indicate LOD. **e** Responses of the vibration phase inside the epithelium. The divergence of the response from positive value implies that phase leads that of the other part in the epithelium. In the phase responses under post-mortem condition, phases overlapped in various sound pressures. Phases are plotted in 10 dB increments in stimulation from 50 to 90 dB SPL under physiological and post-mortem conditions (color figure online)

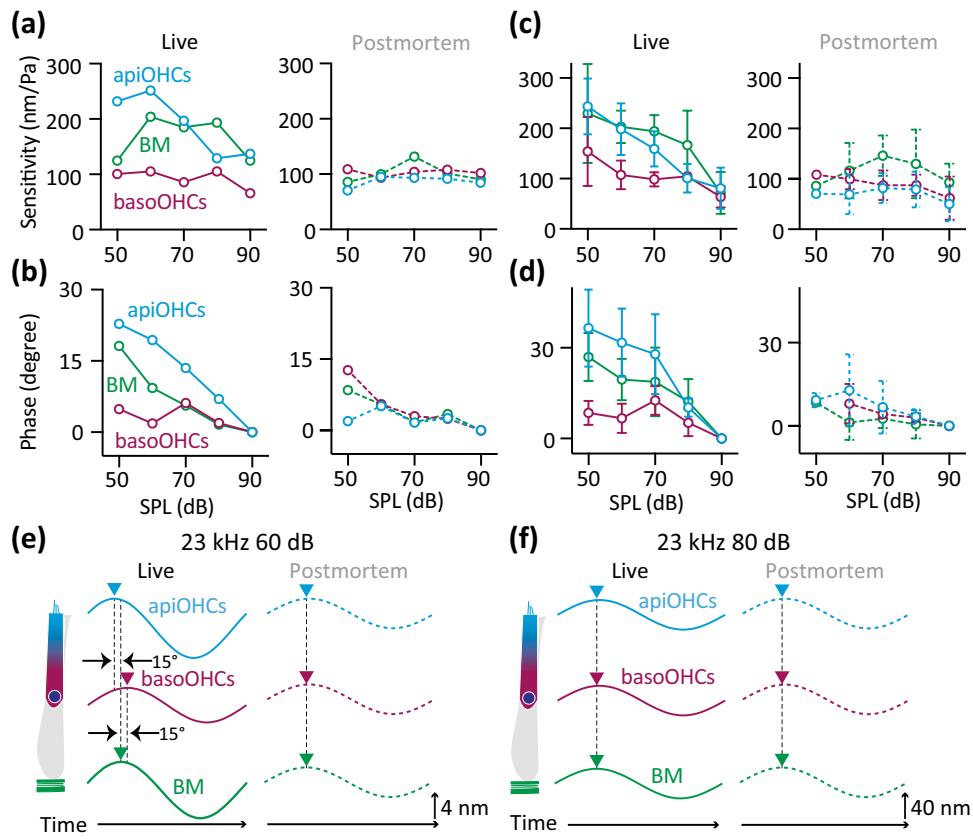


Fig. 8 Summary profile of the vibrations at apical region of outer hair cells (apiOHCs), basal region of OHCs (basoOHCs), and basilar membrane (BM) in one and different guinea pigs ($n=4$). Solid and dotted lines indicate the data under physiological and post-mortem conditions, respectively. Level functions in a live animal relate the active motion of OHCs to the strength of stimulation. In the phase plots, the phase is referenced to the data at 90 dB SPL. **a** The sensitivities under physiological and post-mortem conditions in apiOHCs, basoOHCs, and BM. Sensitivity is calculated from the amplitudes and pressures of sound. The response after anoxia (dotted lines) reveals the behavior of the passive cochlea. At the location of apiOHCs and BM, the physiological data (solid lines) demonstrate higher sensitivity with small-to-moderate stimuli than those under the post-mortem condition. **b** The phases under physiological and post-

mortem conditions in apiOHCs, basoOHCs, and BM. In the high sensitive behavior at apiOHCs and BM during the small-to-moderate stimuli, these phases lead those under 90 dB SPL stimuli. At the location of basoOHCs, neither sensitivity nor phase show any remarkable difference between the physiological and post-mortem conditions. **c, d** Grouped sensitivities and phases of apiOHCs, basoOHCs, and BM and their comparison. **e, f** Schematics of the apiOHCs, basoOHCs, and BM vibrations at the sound level of 60 and 80 dB SPL under physiological and post-mortem conditions. In 60 dB stimulation, vibration amplitudes at apiOHCs and BM are approximately twice greater than those at basoOHCs. The phase at apiOHCs leads that at BM by $\sim 15^\circ$, whereas the phase at basoOHCs lags that at BM by $\sim 15^\circ$. At 80 dB or post-mortem condition, apiOHCs, basoOHCs, and BM vibrate in phase with similar amplitudes

addition, at stimuli of ≤ 70 dB SPL, the phase at apiOHCs and BM never led that at basoOHCs (Fig. 7e).

Figure 8 summarizes the foregoing measurements at apiOHCs, basoOHCs, and BM as a function of the stimulation level (for the plotted data, see Fig. 7c–e open circles). Under physiological conditions, the apiOHCs displayed low sensitivity to the intense sound of more than 80 dB and high sensitivity to small-to-moderate sound (Fig. 8a, solid lines). This result is consistent with the compressive amplification for stimuli reported in the previous studies [8, 15, 17, 20]. The BM sensitivity was also high to moderate sound stimuli. In contrast, the basoOHCs sensitivities little changed and remained almost equivalent to those under the post-mortem conditions, indicating

that basoOHCs motion was not amplified. Under post-mortem conditions, the sensitivities were stably low for various sound pressures; the values at any data point were less than those under control conditions (Fig. 8a, dotted lines). Figure 8b shows the phase plots with reference to data at 90 dB SPL. Under physiological conditions, the phases of both apiOHCs and BM gradually shifted as the sound decreased, and at 50 dB, 22.8° and 15.3° of the phase lead were measured, respectively. Whereas the basoOHCs' phase minimally changed throughout all stimulus intensities. Post-mortem, the phases of apiOHCs and BM were similar to those of basoOHCs and showed no intensity dependence indicating a linear or passive system (Fig. 8b, dotted lines). The trend in the

above vibration measurement was repeatedly confirmed in different animals with similar results (Fig. 8c, d). These results indicated that the basoOHCs vibrations did not occur in synchrony with the mechanical activity of OHCs. To summarize the relative amplitude and phase of vibration in apiOHCs, basoOHCs, and BM under physiological and post-mortem conditions, we displayed the schematics of the vibration at different SPLs in Fig. 8e, f. Under physiological conditions at 60 dB SPL stimuli, the vibration amplitude of apiOHCs and BM exceeded that of basoOHCs. In addition, the phase of apiOHCs' vibration led that of BM by $\sim 15^\circ$. By contrast, the phase of basoOHCs' vibration lagged that of BM by $\sim 15^\circ$. These properties were not remarkable at 80 dB SPL stimuli or under post-mortem conditions. Taken together, the difference in the amplitudes and phases of apiOHCs and BM between control and post-mortem conditions indicates that the OHCs' force contributed to the amplification of the cellular vibration under physiological conditions. Note that we additionally evaluated the vibration of apiOHCs, basoOHCs, and BM in sub-BF stimulation in the same animals used in Figs. 7 and 8 (Fig. S7 and S8). In sub-BF, the amplitude and phase responses to different SPLs exhibited linear growth, and the OHCs activity did not contribute to the epithelial vibration.

4 Discussion and conclusion

The achievements of our study are the optimization of a SD-OCT system for high-resolution *in vivo* imaging and vibrometry in biological material, specifically the sensory epithelium. We described the optical design of an SCSD-OCT system, and demonstrated that the system can image inside of the extremely low reflective material and measure nanometer-scale vibrations in both the apical and basal regions of OHCs as well as BM in live guinea pigs. To focus on the performance of our system in this study, we showed the sound intensity dependence of the vibrations in three regions without detailed frequency dependence in a limited number of guinea pigs. Frequency dependence is likely to reveal the biophysical background of the vibration, as such, it should be undertaken in future experiments.

Conventional OCT vibrometry systems visualize the structure inside the organ of Corti involving apiOHCs and basoOHCs, the tunnel of Corti, and the outer and inner tunnels in the B-scan. However, vibration data in the two regions of OHCs *in vivo* have not yet been fully analyzed owing to the low spatial resolution [13, 15–22, 24, 25]. One of the features of our system is its high axial resolution of $\sim 1.72 \mu\text{m}$ in air. In most conventional OCT vibrometry systems subjected to the recordings of cochlear vibrations,

the spatial resolutions exceed $3.4 \mu\text{m}$ [15–22, 24, 25]. As the length of OHCs is approximately $40\text{--}50 \mu\text{m}$, the pixels inside the OHCs are likely to be less than 12 in those recordings. Comparatively, there was almost twice the number of pixels occupied in this study (Fig. 4d). This axial resolution is equivalent to a high resolution of 'cellular-level' OCT systems described in the literature [41–43].

OCT performance is dependent on the sensitivity in the detector (S), which can be described by a relation [39]:

$$S \propto \frac{\rho A R_s}{2} \Delta t, \quad (2)$$

where ρ is the detector responsivity, A is the power of the light source, R_s is the reflectivity of sample, and Δt is the camera integration time [44–46]. For a transparent sensory epithelium, R_s is only less than 0.04% [27]. To obtain sufficient sensitivity, an increase in Δt as well as A was necessary. However, one of the most serious considerations for *in vivo* vibrometry is to control the noise arising from the physiological motions in the physiological preparation. The amplitudes and frequencies of motions including heartbeat and respiration are mostly high and low, respectively. Thus, a long total acquisition time can smear out the intensity distribution along the depth and degrade the spatial accuracy in vibrometry as we show in Fig. 5. To overcome this shortcoming, we set the total acquisition time to be as short as possible while strengthening light power.

Because of the features and adjustments above, the system that we constructed successfully captured high-resolution images and sub-nanoscale vibrations in the apical and basal regions of OHCs *in vivo*. An ultra-high-resolution OCT with an SC source was previously applied to record the image for guinea pig cochlea *ex vivo* [38]. However, its sampling rate was 40 Hz, which was insufficient for high-frequency vibrometry. This limitation was overcome by the sampling rate by the performance of the SCSD-OCT system described in the present study. Therefore, the vibrometer was capable of visualizing previously unidentified mechanical properties of live OHCs.

The power of the beam onto the sample in our system was 117 mW, significantly higher than that in previous OCT systems [12–18, 20–25, 36–39]. Because the total acquisition time is 0.1 s, total irradiation energy in one session of vibrometry is $\sim 11.7 \text{ mW}\cdot\text{s}$. This total energy is equivalent to or less than that in other recent studies [13, 18, 19, 24, 25]. For example, in the recording with another commercial SD-OCT system in the same location and same animal, the total irradiation energy was $\sim 100 \text{ mW}\cdot\text{s}$ (3.7 mW with the total acquisition time of 27.2 s). Together with the minor hearing threshold shifts before and after the 2 h irradiation in Fig. S3, the power of the light source in our study seems to be harmless to the functions of animals.

Our current understanding of biophysical function in *ex vivo* OHCs has expanded in recent decades [5, 8, 9]. Electrophysiological and molecular biological approaches have shown two representative functions. First, tilting the hair bundle atop OHCs allows nonlinear flow of electrical current through the mechanoelectrical channel mainly composed of transmembrane channel-like (TMC) proteins [47, 48]. Second, the stimulated OHCs produce force owing to the conformational change of the motor protein, prestin [5, 6, 10, 49]. Thus, the motility of OHCs is dependent on the current, and is known as electromotility [5, 8, 9]. As for the *in vivo* function of OHCs, recent theoretical and experimental approaches have suggested that the physical environment surrounding OHCs can change their responses *in vivo* [50, 51]. In *in vivo* recording of the motion of electromotility of OHC is, therefore, one of the biggest issues in the field of auditory physiology. Our data in the cluster of three OHCs in Fig. 7 contains the motion in different OHCs. Conversely, all three rows of OHCs are likely to vibrate as a single cluster at the basal turn in the *in vivo* guinea pigs [17]. Consequently, our data clarify one aspect of the *in vivo* OHCs' electromotility.

It has been reported that the vibration amplitude of OHCs is significantly amplified, and its phase leads that at BM under the near-BF stimuli in guinea pigs, gerbils, and mice [15, 18, 19, 31]. Also in our study, vibration was amplified and its phase led in apiOHCs. These observations are consistent with previous discoveries. By contrast, the amplitude in basoOHCs is not actively amplified, and its phase lags relative to the phase of BM in our data (Fig. 8i). These results suggest that the electromotility force is unlikely to directly transmitted to BM via basal region of OHCs. Inner/outer pillar cells, the phalangeal process of Deiters' cells, and other cellular structures are possible candidates for other pathway. Accordingly, the force transmission mechanism of the electromotility inside the epithelium may differ among species. Also, in reference to the phase of basoOHCs, apiOHCs motion leads by ~30 degrees under near-BF stimulation in the weak intensities. This phase lead was consistent with the active phase lead recorded during *in vitro* hair bundle behavior [3, 52, 53], and seems to contribute to the active amplification through the negative stiffness of the bundle. Similarly to our data, the phase lead of the bundle is not remarkable in strong stimuli.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10043-021-00654-8>.

Acknowledgements We thank the members of our research group for their critical discussion, technical advice, and Mr. Samuel Asare and Editage (www.editage.jp) for English language editing. This study was supported by Grants-in-Aid for Young Scientists A 15H05683 (to F.N.); Grants-in-Aid for Scientific Research B 18H02950 (to F.N.), 19H02151 (to S.C.); Grants-in-Aid for Scientific Research A 18H04062 (to H.H.);

Grants-in-Aid for Exploratory Research 20K21883 (to H.H.); Early-Career Scientists 20K20164 (to T.O.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan; AMED-CREST JP20gm0810004 AMED (to H.H., and S.C.); and Grants from Inoue Foundation for Science (to F.N.), the Tsukada Grant for Niigata University Medical Research (to F.N.), Takeda Science Foundation (to F.N.), Nakatani Foundation for Advancement of Measuring Technologies in Biomedical Engineering (to F.N.), Senri Life Science Foundation S-25011 (to F.N.), the Ichiro Kanehara Foundation for the Promotion of Medical Sciences 13KI010 (to F.N.), Kanae Foundation for the Promotion of Medical Science (to F.N.), Nagai NS Promotion Foundation for Science of Perception (to F.N. and T.O.), Grand for Basic Science Research Projects from the Sumitomo Foundation no. 150362 (to F.N.), the Kurata Grant awarded by the Hitachi Global Foundation no. 1353 (to F.N.), Shimadzu Science Foundation (to F.N.), Toray Science Foundation No. 20-6109 (to F.N.), Yamaguchi Educational and Scholarship Foundation J18G0050 (to T.O.), the Naito Foundation (to H.H.), and Joint Research Program of the Biosignal Research Center, Kobe University 281003 (to H.H.).

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by FN, SC, TO and HH. The first draft of the manuscript was written by FN and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declaration

Conflict of interest The authors declare that there are no conflicts of interest related to this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Von Békésy, G.: Experiments in hearing. McGraw-Hill series in psychology. McGraw-Hill, New York (1960)
2. Dallos, P., Popper, A.N., Fay, R.R.: The cochlea. Springer handbook of auditory research, vol 8. Springer, New York (1996)
3. Hudspeth, A.J.: Integrating the active process of hair cells with cochlear function. *Nat. Rev. Neurosci.* **15**(9), 600–614 (2014). <https://doi.org/10.1038/nrn3786>
4. Fettiplace, R.: Hair cell transduction, tuning, and synaptic transmission in the mammalian cochlea. *Compr. Physiol.* **7**(4), 1197–1227 (2017). <https://doi.org/10.1002/cphy.c160049>
5. Brownell, W.E., Bader, C.R., Bertrand, D., de Ribaupierre, Y.: Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227**(4683), 194–196 (1985). <https://doi.org/10.1126/science.3966153>

6. Ashmore, J.F.: A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol.* **388**, 323–347 (1987). <https://doi.org/10.1113/jphysiol.1987.sp016617>
7. Gold, T.: Hearing II. The physical basis of the action of the cochlea. *Proc. R. Soc. Lond. B Biol. Sci.* **131**(881), 492–498 (1948)
8. Robles, L., Ruggero, M.A.: Mechanics of the mammalian cochlea. *Physiol. Rev.* **81**(3), 1305–1352 (2001). <https://doi.org/10.1152/physrev.2001.81.3.1305>
9. Ashmore, J.: Cochlear outer hair cell motility. *Physiol. Rev.* **88**(1), 173–210 (2008). <https://doi.org/10.1152/physrev.00044.2006>
10. Liberman, M.C., Gao, J., He, D.Z., Wu, X., Jia, S., Zuo, J.: Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* **419**(6904), 300–304 (2002). <https://doi.org/10.1038/nature01059>
11. Dallos, P., Wu, X., Cheatham, M.A., Gao, J., Zheng, J., Anderson, C.T., Jia, S., Wang, X., Cheng, W.H., Sengupta, S., He, D.Z., Zuo, J.: Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron* **58**(3), 333–339 (2008). <https://doi.org/10.1016/j.neuron.2008.02.028>
12. Hong, S.S., Freeman, D.M.: Doppler optical coherence microscopy for studies of cochlear mechanics. *J. Biomed. Opt.* **11**(5), 054014 (2006). <https://doi.org/10.1117/1.2358702>
13. Chen, F., Choudhury, N., Zheng, J., Matthews, S., Nuttall, A.L., Jacques, S.L.: In vivo imaging and low-coherence interferometry of organ of Corti vibration. *J. Biomed. Opt.* **12**(2), 021006 (2007). <https://doi.org/10.1117/1.2717134>
14. Wang, R.K., Nuttall, A.L.: Phase-sensitive optical coherence tomography imaging of the tissue motion within the organ of Corti at a subnanometer scale: a preliminary study. *J. Biomed. Opt.* **15**(5), 056005 (2010). <https://doi.org/10.1117/1.3486543>
15. Chen, F., Zha, D., Fridberger, A., Zheng, J., Choudhury, N., Jacques, S.L., Wang, R.K., Shi, X., Nuttall, A.L.: A differentially amplified motion in the ear for near-threshold sound detection. *Nat. Neurosci.* **14**(6), 770–774 (2011). <https://doi.org/10.1038/nn.2827>
16. Warren, R.L., Ramamoorthy, S., Ciganovic, N., Zhang, Y., Wilson, T.M., Petrie, T., Wang, R.K., Jacques, S.L., Reichenbach, T., Nuttall, A.L., Fridberger, A.: Minimal basilar membrane motion in low-frequency hearing. *Proc. Natl. Acad. Sci. USA* **113**(30), E4304–4310 (2016). <https://doi.org/10.1073/pnas.1606317113>
17. Ramamoorthy, S., Zha, D., Chen, F., Jacques, S.L., Wang, R., Choudhury, N., Nuttall, A.L., Fridberger, A.: Filtering of acoustic signals within the hearing organ. *J. Neurosci.* **34**(27), 9051–9058 (2014). <https://doi.org/10.1523/JNEUROSCI.0722-14.2014>
18. Lee, H.Y., Raphael, P.D., Park, J., Ellerbee, A.K., Applegate, B.E., Oghalai, J.S.: Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea. *Proc. Natl. Acad. Sci. USA* **112**(10), 3128–3133 (2015). <https://doi.org/10.1073/pnas.1500038112>
19. Cooper, N.P., Vavakou, A., van der Heijden, M.: Vibration hot-spots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea. *Nat. Commun.* **9**(1), 3054 (2018). <https://doi.org/10.1038/s41467-018-05483-z>
20. Zha, D., Chen, F., Ramamoorthy, S., Fridberger, A., Choudhury, N., Jacques, S.L., Wang, R.K., Nuttall, A.L.: In vivo outer hair cell length changes expose the active process in the cochlea. *PLoS ONE* **7**(4), e32757 (2012). <https://doi.org/10.1371/journal.pone.0032757>
21. Gao, S.S., Raphael, P.D., Wang, R., Park, J., Xia, A., Applegate, B.E., Oghalai, J.S.: In vivo vibrometry inside the apex of the mouse cochlea using spectral domain optical coherence tomography. *Biomed. Opt. Express* **4**(2), 230–240 (2013). <https://doi.org/10.1364/BOE.4.00230>
22. Nuttall, A.L., Ricci, A.J., Burwood, G., Harte, J.M., Stenfelt, S., Caye-Thomassen, P., Ren, T., Ramamoorthy, S., Zhang, Y., Wilson, T., Lunner, T., Moore, B.C.J., Fridberger, A.: A mechanoelectrical mechanism for detection of sound envelopes in the hearing organ. *Nat. Commun.* **9**(1), 4175 (2018). <https://doi.org/10.1038/s41467-018-06725-w>
23. Dewey, J.B., Xia, A., Muller, U., Belyantseva, I.A., Applegate, B.E., Oghalai, J.S.: Mammalian auditory hair cell bundle stiffness affects frequency tuning by increasing coupling along the length of the cochlea. *Cell Rep.* **23**(10), 2915–2927 (2018). <https://doi.org/10.1016/j.celrep.2018.05.024>
24. Fallah, E., Strimbu, C.E., Olson, E.S.: Nonlinearity and amplification in cochlear responses to single and multi-tone stimuli. *Hear. Res.* **377**, 271–281 (2019). <https://doi.org/10.1016/j.heares.2019.04.001>
25. Vavakou, A., Cooper, N.P., van der Heijden, M.: The frequency limit of outer hair cell motility measured in vivo. *Elife* (2019). <https://doi.org/10.7554/eLife.47667>
26. Choi, S., Nin, F., Ota, T., Sato, K., Muramatsu, S., Hibino, H.: In vivo tomographic visualization of intracochlear vibration using a supercontinuum multifrequency-swept optical coherence microscope. *Biomed. Opt. Express* **10**(7), 3317–3342 (2019). <https://doi.org/10.1364/BOE.10.003317>
27. Khanna, S.M., Willemin, J.F., Ulfendahl, M.: Measurement of optical reflectivity in cells of the inner ear. *Acta Otolaryngol. Suppl.* **467**, 69–75 (1989)
28. Ramamoorthy, S., Zhang, Y., Petrie, T., Fridberger, A., Ren, T., Wang, R., Jacques, S.L., Nuttall, A.L.: Minimally invasive surgical method to detect sound processing in the cochlear apex by optical coherence tomography. *J. Biomed. Opt.* **21**(2), 25003 (2016). <https://doi.org/10.1117/1.JBO.21.2.025003>
29. Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M., Altman, D.G.: Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**(6), e1000412 (2010). <https://doi.org/10.1371/journal.pbio.1000412>
30. Ota, T., Nin, F., Choi, S., Muramatsu, S., Sawamura, S., Ogata, G., Sato, M.P., Doi, K., Doi, K., Tsuji, T., Kawano, S., Reichenbach, T., Hibino, H.: Characterisation of the static offset in the travelling wave in the cochlear basal turn. *Pflugers Arch.* **472**(5), 625–635 (2020). <https://doi.org/10.1007/s00424-020-02373-6>
31. Ren, T., He, W., Kemp, D.: Reticular lamina and basilar membrane vibrations in living mouse cochleae. *Proc. Natl. Acad. Sci. USA* **113**(35), 9910–9915 (2016). <https://doi.org/10.1073/pnas.1607428113>
32. Ren, T., He, W.: Two-tone distortion in reticular lamina vibration of the living cochlea. *Commun. Biol.* **3**(1), 35 (2020). <https://doi.org/10.1038/s42003-020-0762-2>
33. He, W., Kemp, D., Ren, T.: Timing of the reticular lamina and basilar membrane vibration in living gerbil cochleae. *Elife* (2018). <https://doi.org/10.7554/eLife.37625>
34. Ren, T.: Longitudinal pattern of basilar membrane vibration in the sensitive cochlea. *Proc. Natl. Acad. Sci. USA* **99**(26), 17101–17106 (2002). <https://doi.org/10.1073/pnas.262663699>
35. Welzel, J.: Optical coherence tomography in dermatology: a review. *Skin Res. Technol.* **7**(1), 1–9 (2001)
36. Lee, H.Y., Raphael, P.D., Xia, A., Kim, J., Grillet, N., Applegate, B.E., Ellerbee Bowden, A.K., Oghalai, J.S.: Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of Corti. *J. Neurosci.* **36**(31), 8160–8173 (2016). <https://doi.org/10.1523/JNEUROSCI.1157-16.2016>
37. Dewey, J.B., Applegate, B.E., Oghalai, J.S.: Amplification and suppression of traveling waves along the mouse organ of Corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces. *J. Neurosci.* **39**(10), 1805–1816 (2019). <https://doi.org/10.1523/JNEUROSCI.2608-18.2019>
38. Iyer, J.S., Batts, S.A., Chu, K.K., Sahin, M.I., Leung, H.M., Tearney, G.J., Stankovic, K.M.: Micro-optical coherence

tomography of the mammalian cochlea. *Sci. Rep.* **6**, 33288 (2016). <https://doi.org/10.1038/srep33288>

39. Kim, S., Oghalai, J.S., Applegate, B.E.: Noise and sensitivity in optical coherence tomography based vibrometry. *Opt. Express* **27**(23), 33333–33350 (2019)

40. Recio-Spinoso, A., Oghalai, J.S.: Mechanical tuning and amplification within the apex of the guinea pig cochlea. *J. Physiol.* **595**(13), 4549–4561 (2017). <https://doi.org/10.1113/JP273881>

41. Tan, B., Hosseinaee, Z., Han, L., Kralj, O., Sorbara, L., Bizheva, K.: 250 kHz, 1.5 micron resolution SD-OCT for in-vivo cellular imaging of the human cornea. *Biomed. Opt. Express* **9**(12), 6569–6583 (2018). <https://doi.org/10.1364/BOE.9.006569>

42. Werkmeister, R.M., Sapeta, S., Schmidl, D., Garhofer, G., Schmidinger, G., Dos Santos, V.A., Aschinger, G.C., Baumgartner, I., Pircher, N., Schwarzhans, F., Pantalon, A., Dua, H., Schmetterer, L.: Ultra-high-resolution OCT imaging of the human cornea. *Biomed. Opt. Express* **8**(2), 1221–1239 (2017). <https://doi.org/10.1364/BOE.8.001221>

43. Kolb, J.P., Pfeiffer, T., Eibl, M., Hakert, H., Huber, R.: High-resolution retinal swept source optical coherence tomography with an ultra-wideband Fourier-domain mode-locked laser at MHz A-scan rates. *Biomed. Opt. Express* **9**(1), 120–130 (2018). <https://doi.org/10.1364/BOE.9.000120>

44. Choma, M., Sarunic, M., Yang, C., Izatt, J.: Sensitivity advantage of swept source and Fourier domain optical coherence tomography. *Opt. Express* **11**(18), 2183–2189 (2003). <https://doi.org/10.1364/oe.11.002183>

45. de Boer, J.F., Cense, B., Park, B.H., Pierce, M.C., Tearney, G.J., Bouma, B.E.: Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography. *Opt. Lett.* **28**(21), 2067–2069 (2003). <https://doi.org/10.1364/ol.28.002067>

46. Leitgeb, R., Hitzenberger, C., Fercher, A.: Performance of Fourier domain vs. time domain optical coherence tomography. *Opt. Express* **11**(8), 889–894 (2003). <https://doi.org/10.1364/oe.11.000889>

47. Hudspeth, A.J.: How the ear's works work. *Nature* **341**(6241), 397–404 (1989). <https://doi.org/10.1038/341397a0>

48. Kurima, K., Ebrahim, S., Pan, B., Sedlacek, M., Sengupta, P., Millis, B.A., Cui, R., Nakanishi, H., Fujikawa, T., Kawashima, Y., Choi, B.Y., Monahan, K., Holt, J.R., Griffith, A.J., Kachar, B.: TMC1 and TMC2 localize at the site of mechanotransduction in mammalian inner ear hair cell stereocilia. *Cell Rep.* **12**(10), 1606–1617 (2015). <https://doi.org/10.1016/j.celrep.2015.07.058>

49. Zheng, J., Shen, W., He, D.Z., Long, K.B., Madison, L.D., Dallos, P.: Prestin is the motor protein of cochlear outer hair cells. *Nature* **405**(6783), 149–155 (2000). <https://doi.org/10.1038/35012009>

50. Maoiléidigh, D.O., Hudspeth, A.J.: Effects of cochlear loading on the motility of active outer hair cells. *Proc. Natl. Acad. Sci. USA* **110**(14), 5474–5479 (2013). <https://doi.org/10.1073/pnas.1302911110>

51. Salvi, J.D., Maoiléidigh, D.O., Fabella, B.A., Tobin, M., Hudspeth, A.J.: Control of a hair bundle's mechanosensory function by its mechanical load. *Proc. Natl. Acad. Sci. USA* **112**(9), E1000–E1009 (2015). <https://doi.org/10.1073/pnas.1501453112>

52. Martin, P., Mehta, A.D., Hudspeth, A.J.: Negative hair-bundle stiffness betrays a mechanism for mechanical amplification by the hair cell. *Proc. Natl. Acad. Sci. USA* **97**(22), 12026–12031 (2000). <https://doi.org/10.1073/pnas.210389497>

53. Martin, P., Hudspeth, A.J.: Compressive nonlinearity in the hair bundle's active response to mechanical stimulation. *Proc. Natl. Acad. Sci. USA* **98**(25), 14386–14391 (2001). <https://doi.org/10.1073/pnas.251530498>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.