Orthogonal spectral and temporal envelope representation in auditory cortex

Kuniyuki Takahashi[1,*], Tianrui Guo[2], Tatsuya Yamagishi[2], Shinsuke Ohshima[2], Hiroaki Tsukano[3,4,*], Arata Horii[2]

1. Department of Otolaryngology Head and Neck Surgery, Faculty of Medicine, University of Miyazaki, Miyazaki 889-1692, Japan
2. Department of Otolaryngology, Graduate School of Medicine and Dental Sciences, Niigata University, Niigata 951-8510, Japan
3. Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
4. Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Correspondence to: kuniyuki_takahashi@med.miyazaki-u.ac.jp and hiroakit@email.unc.edu

Abstract

Speech sounds are acoustically composed of two main components: spectral and temporal. The spectral component or carrier frequency is important for pitch detection and is perceived through activity at specific locations on the tonotopic map in the auditory cortex. The temporal envelope—the dynamic changes in sound wave amplitude over time—plays an essential role in speech perception by mediating phoneme recognition. However, how the steepness of the envelope, which shapes the temporal structure, is represented in the auditory cortex remains poorly understood. This study investigates how the steepness of the envelope of sound waves is represented as a functional organization in the mouse auditory cortex, in relation to spectral tonotopic organization. Using macroscale calcium imaging in GCaMP6f-expressing mice, we mapped the auditory cortex's responses to tones with systematically varying rise-ramp steepness and frequencies. Our findings reveal that, in two primary-like regions of the auditory cortex, the steepness of the rise ramp is orderly mapped orthogonally to the tonotopic gradient, forming a two-dimensional representation on the cortical surface. Conversely, a significant representation of rise-ramp steepness was not detected in higher-order-like auditory regions. These findings suggest that a dichotomous functional organization independently processes temporal envelopes and spectral features within mammalian auditory cortical areas, with functional structures mirroring the two acoustic structures.

Introduction

Speech sounds are acoustically composed of two main components: spectral and temporal (Figure 1A). The spectral components generate pitch perception, which contributes to identifying individuals or intonation. On the other hand, the temporal components, particularly the sound envelope—representing the dynamic fluctuation in sound amplitude over time (Drullman, 1995)—plays a crucial role in speech perception as a mediator of voiced phonemes (Rosen, 1992; Shannon et al., 1995). This is supported by multiple lines of evidence: sound waves with similar envelope shapes produce similar percepts, allowing accurate phoneme recognition even when spoken by individuals with different fundamental frequencies (Aizawa and Eggermont, 2006; Frye et al., 2007). Additionally, the envelope provides stronger cues for speech perception than carrier frequency, as shown in psychological studies using synthesized sounds with selectively weighted components (Smith et al., 2002). Since the temporal envelope is a feature shared across vocalizations in various mammalian species (Figure 1B), this dichotomy in acoustic elements likely facilitates effective vocal communication among individuals across many species.

Functional organization in the cortex is a fundamental principle of mammalian sensory processing, and the auditory cortex exhibits a prominent topographic organization. It is well-established that the auditory cortex exhibits tonotopy, where spectral information are spatially arranged from low to high (Bizley et al., 2005). A recent study demonstrated that pitch is perceived through activity at corresponding locations on the tonotopic map (Ceballo et al., 2019). In contrast, much less is known whether envelopes are represented as a functional organization, although it is already known that envelope perception is mediated by auditory cortical activity from clinical cases with the damage in the human auditory cortex (Maffei et al., 2007). Many studies have investigated neural responses to temporal envelopes in animal brains using sparse single-neuron recordings (Eggermont, 2001; Fishbach et al., 2001; Heil, 2001; Phillips et al., 2002), revealing that auditory neurons are sensitive to temporal envelopes (Nagarajan et al., 2002); however, comprehensive visualization of the underlying principles remains a challenge. It could be reasoned that the limited exploration of envelope-related functional organization may be due to technical challenges associated with sparse recordings, which lack the spatial resolution needed to map out potential fine functional organizations.

To address this issue, here we characterized temporal envelope representation in

the auditory cortex using macroscale calcium imaging with GCaMP6f-expressing mice. This technique has achieved mapping fine cortical functional organizations with its high spatial resolution (Issa et al., 2014; Romero et al., 2020; Calhoun et al., 2023). Our investigation focused on the steepness of the envelope, a key feature in shaping the envelope, and sought to determine the presence or absence of a functional organization related to the envelope. We employed simplified acoustic stimuli systematically varying onset ramp steepness, that has been used in many previous physiological investigation (Suga, 1971; Phillips, 1998; Heil, 2001; Peterson and Heil, 2021). This stimulus is particularly advantageous for GCaMP macroscale imaging, as GCaMP reliably elicits responses during the onset phase of sound presentation, with less sustained activity during the ongoing phase (Issa et al., 2014; Romero et al., 2020; Chen et al., 2021). We found that among the four tonotopic regions studied, two primary-like regions exhibited a two-dimensional map that orthogonally represents both envelope steepness and tonotopic organization; however, higher-order-like regions represent only tonotopy but not envelope steepness maps. These findings suggest functional specialization across auditory areas and reveal an alignment between two key acoustic and functional structures in the mammalian auditory cortex.

Results

Visualizing auditory cortex map with calcium imaging

Calcium signals were simultaneously recorded from the bilateral auditory cortices of GCaMP6f-expressing mice, with our initial analysis focused on the right auditory cortex (Figure 2A and 2B). To visualize the tonotopic maps, mice were exposed to tones with frequencies ranging from 5 to 60 kHz, following the parameters used in our previous studies (Takahashi et al., 2006; Tsukano et al., 2015). Fluorescent changes were consistently observed in response to the onset of tones across all frequencies (Figure 2C). Although various naming conventions exist for areas of the mouse auditory cortex, in this study, we refer to auditory areas according to the classification presented by Tsukano et al., 2017. Clear tonotopic gradients were observed in the anterior auditory field (AAF), dorsal medial field (DM), primary auditory field (A1), and secondary auditory field (A2), and the directions of tonotopic gradients were consistent with many previous studies (Figure 2D and 2E) (Tsukano et al., 2016; Liu et al., 2019; Narayanan et al., 2022).

In order to evaluate the impact of the temporal envelope steepness on the auditory cortex, we employed simplified stimuli that systematically varied onset ramp steepness, following previous studies (Phillips, 1998; Heil, 2001). We varied the rise-ramp steepness of tones presented to the same mouse, with ramp durations ranging from 0.01 to 100 msec, and carrier frequencies of 5, 20, and 40 kHz (Figure 2F and 2G). Surprisingly, we observed a clear shift in the peak location of tonal responses especially in AAF and DM, occurring in a direction orthogonal to the tonotopic gradient.

Presence of two-dimensional map in primary-like regions

To examine these maps in greater detail, we performed a quantitative analysis of group data from AAF. In the 2-D plots, tonotopic organization is seen in the dorso-ventral direction (Figure 3A), while the shorter rise time, i.e., steeper rise ramp, shifted peak response locations in the rostro-caudal direction regardless of carrier frequency (Figure 3A). Notably, tonotopic organization remained the same across different rise ramp steepness. These findings suggest that AAF represents a two-dimensional map on the cortical surface. The mean angle between the tonotopy and rise-ramp map was approximately 82 degrees (Figure 3B), suggesting that these two maps were distinct and orthogonal to each other. A statistically significant correlation was found between rise-ramp steepness and the distance orthogonal to tonotopy ($r = 0.63$, $p = 5.8 \times 10^{-10}$; Figure 3C, left). It is worthwhile to note that the regression line aligns well on the logarithmic-scale horizontal axis for rise-ramp times, which suggests that the variations in steepness, particularly when the onset ramp is steep, are captured with high spatial resolution. In contrast, no correlation was observed between rise-ramp steepness and the direction along tonotopy ($r = -0.064$, $p = 0.57$; Figure 3C, right).

Similarly, analysis of DM revealed a rise-ramp map that was a mirror image of that in AAF (Figure 3D–F). Due to the instability of low-frequency responses in DM, we conducted a quantitative analysis of DM maps, excluding the 5 kHz responses. The peak response locations shifted caudo-rostrally with increasing rise ramp steepness, irrespective of carrier frequency (Figure 3D). The angle between the tonotopic map and the rise-ramp map in DM was approximately 74 degrees (Figure 3E). A statistically significant correlation was also observed between rise-ramp steepness and the distance orthogonal to tonotopy in DM ($r = 0.58$, $p = 2.0 \times 10^{-8}$; Figure 3F, left), while no correlation was found between rise-ramp steepness and the direction along tonotopy ($r = 0.13$, $p = 0.26$; Figure 3F, right). In summary, both AAF and DM exhibit mirror-imaged rise-ramp maps that are orthogonal to each area's tonotopic organization.

We conducted similar analyses in A1 and A2. Interestingly, neither A1 nor A2 exhibited detectable rise-ramp maps. In A1, tonotopic organization runs in the caudo-rostral direction, curving in the high frequency part (Figure 4A), as reported previously (Tsukano et al., 2015). Interestingly, there was no shift in response peaks perpendicular to the tonotopic axis, even when the rise-ramp steepness was altered (Figure 4A and 4B). No statistically significant correlation was observed between rise-ramp steepness and the distance orthogonal to tonotopy ($r = 0.0052$, $p = 0.96$; Figure 4C, left), and there was no significant correlation between rise-ramp steepness and the direction along tonotopy ($r = 0.21$, $p = 0.059$; Figure 4C, right). Similarly, in A2, altering the rise-ramp steepness did not cause any shift in response peaks perpendicular to the tonotopic axis (Figure 4D and 4E). No statistically significant correlation was found between rise-ramp steepness and the distance orthogonal to tonotopy ($r = -0.053$, $p = 0.64$; Figure 4F, left), and there was no significant correlation between rise-ramp steepness and the direction along tonotopy ($r = 0.044$, $p = 0.70$; Figure 4F, right). These results suggest that A1 and A2 primarily represent one-dimensional maps reflecting tonotopy, without distinct representation of rise-ramp steepness, indicating region-specific processing of envelope steepness.

Bilateral representation of two-dimensional map

We performed simultaneous imaging of the bilateral auditory cortex in the same group of mice, allowing us to examine whether the rise-ramp map is represented in both hemispheres. We found similar functional organizations in the left auditory cortex. As rise-ramp steepness increases, the responses shifted caudally in AAF, while the responses in DM shifted rostrally (Figure S1A), as seen in the right hemisphere. On the other hand, clear shifts were not seen in A1 or A2. A statistically significant correlation was found between rise-ramp steepness and the distance orthogonal to tonotopy in AAF ($r = 0.60$, ****$p = 3.9 \times 10^{-9}$) and DM ($r = -0.40$ ***$p = 2.1 \times 10^{-4}$), but not in A1 ($r = 0.026$, $p = 0.83$) or A2 ($r = -0.11$, $p = 0.33$) (Figure S1B). No correlation was found between rise-ramp steepness and the direction along tonotopy in any areas (AAF, $r = 0.076$, $p = 0.50$; DM, $r = 0.15$, $p = 0.19$; A1, $r = 0.012$, $p = 0.92$; A2, $r = 0.11$, $p = 0.35$) (Figure S1C). These data suggest that the AAF and DM symmetrically represent rise-ramp maps across hemispheres.

Orthogonal shifting during the ramping phase

Although we observed shifts in response location with changes in rise-ramp steepness in

AAF and DM, the GCaMP signals were slow, with peak responses occurring after the tonal waveform had reached its plateau phase. To verify that response locations differ during the ramping phase, we visualized responses specifically during the ramping phase by performing imaging at a higher sampling rate (Figure S2). We used tones with rise ramp times of 75 msec and 200 msec and compared the location of peak responses in AAF using data at the timing of 40–60 msec after the tonal onset to ensure evaluation of the rise-ramp map during the ramping phase. The location of peak responses for 200 msec rise-ramp time was significantly more rostral than that for the 75 msec ramp time ($p = 0.039$, two-sided paired t-test). This data confirms that positional shifts occurred definitely during the ramping phase, suggesting that response location is purely sensitive to rise-ramp steepness. Overall, our study demonstrates that AAF and DM represent two-dimensional maps encompassing both tonotopy and rise-ramp steepness, with the rise-ramp maps displayed in a mirror-imaged fashion and separated by the center area (the dorsoanterior field) (Figure 5).

Discussion

The current study reveals that mouse auditory cortex contains well-defined two-dimensional maps where tonotopy and rise-ramp steepness are represented orthogonally. While tonotopy is a prominent feature that is topographically represented in the auditory cortex, the existence of the information encoded in the secondary dimension has been a longstanding question. In the mouse auditory cortex, previous research has identified functional organizations distinct from tonotopy, such as representation of frequency modulation sweeps (Issa et al., 2016) and offset responses (Baba et al., 2016; Liu et al., 2019); however, these structures were typically aligned with or located apart from the tonotopic organization. Although the overall sound pressure level may vary depending on the slope of the onset—decreasing as the slope becomes more gradual—it is already known that the intensity is not represented as a functional organization in the mouse auditory cortex (Bandyopadhyay et al., 2010), making it unlikely that the observed steepness map arise from intensity variations. Therefore, to the best of our knowledge, this is the first demonstration of two-dimensional maps in the mouse auditory cortex where tonotopy and envelope steepness are orthogonally organized.

Systematic spatial representation of temporal envelope steepness suggests that the

mouse auditory cortex encodes selectivity for envelope steepness. This finding advances the understanding of temporal envelope processing, building on earlier electrophysiological studies, which demonstrated neurons consistently responding to the same steepness (Heil, 1997, 1998; Heil and Irvine, 1998) and neurons selective for the direction of envelope slope (Lu et al., 2001; Zhou and Wang, 2010). While the auditory cortex is known to be critical for perceiving both of temporal envelopes (Nourski et al., 2009) and pitch (frequency information) (Warrier and Zatorre, 2004), the orthogonal representation of envelope steepness and tonotopy suggests that temporal envelopes are processed independently of spectral features in the auditory cortex, supporting the psychological observation that temporal envelope and carrier frequency work independently in verbal communication (Smith et al., 2002). Our study underscores a fundamental biological principle: dichotomous functional structures mirror two acoustic features critical for vocalization, linking neural circuitry with acoustic signals of vocal communication. While this principle should apply broadly across mammals, identifying it in the mouse auditory cortex is especially valuable given the increasing focus on vocalization processing in mice. This finding lays the groundwork for exploring how animals perceive conspecific vocalizations (Stiebler et al., 1997; Agarwalla et al., 2023).

Previous studies reported that the orthogonal map to tonotopy represents periodotopy—the cycle of peaks in amplitude-modulated sound waveforms. This has been observed in several species other than mice, throughout different stages of the auditory pathway (Schreiner and Langner, 1988; Krishna and Semple, 2000; Langner et al., 2002, 2009; Baumann et al., 2011, 2015; Barton et al., 2012). In contrast, our findings suggest that the presentation of a one-time rise-ramp steepness, which does not inherently involve periodicity, is sufficient to elucidate an orthogonal map in the mouse auditory cortex. Two possible explanations could reconcile our findings with previous reports. First, the mouse auditory system may differ from that of other animals, with cortical representations being particularly sensitive to envelope steepness. As a result, the mouse auditory cortex developed a specialized functional organization to process this aspect in detail. Second, envelope steepness might contribute to the visualization of the orthogonal map in previous studies using other animals. In nearly all previous studies, amplitude-modulated tones were used as periodic acoustic stimuli, which involve repeated ramps with consistent rise and fall steepness, making it difficult to determine whether periodicity or repetitive slope plays a greater role in visualizing the orthogonal map. To clarify that the periodicity alone is sufficient to visualize the second map, click trains, which lack envelope ramping, would need to be used as a periodic

stimulus; however, few studies—except one that could not visualize orderly periodotopy in the inferior colliculus (Schnupp et al., 2015)—have employed click trains. Future research can be directed to explore how ramp steepness and periodicity contribute to the visualization of secondary dimensional maps. Although periodicity and envelopes are sometimes discussed interchangeably, this study suggests that envelope steepness and periodicity should be more clearly understood as biologically distinct concepts (Rosen, 1992).

Interestingly, we observed region-specificity in the presence of envelope steepness maps: clear two-dimensional maps were observed in AAF and DM, but were unable to be visualized in A1 and A2. This may indicate the presence of functional specialization across areas in the mouse auditory cortex, with AAF and DM specifically processing envelopes. Alternatively, it is possible that, although envelopes are processed at a single-neuron levels, the orthogonal maps in higher-order become unclear due to hierarchy of the mouse auditory cortex. Based on the latency (Guo et al., 2012), the broadness of frequency tuning (Guo et al., 2012; Joachimsthaler et al., 2014), and the connectivity patterns with the auditory thalamus (Ohga et al., 2018), AAF and DM exhibit more primary-like characteristics, while A1 is intermediate, and A2 functions as a higher-order area (Tsukano et al., 2017; Kline et al., 2021). Given that the auditory cortex's limited temporal resolution (Ma et al., 2013), it is likely that envelop slope information is initially processed and represented as a spatial map in subcortical nuclei, such as in inferior colliculus, which has better temporal resolution (Khouri et al., 2011). Previous research has demonstrated that primary auditory areas receive well-organized topographic projections from the ventral division of the medial geniculate body (MGv), whereas higher-order areas tend to receive more anatomically and physiologically convergent inputs from MGv (Vasquez-Lopez et al., 2017; Ohga et al., 2018). Even if these envelope steepness maps are topographically projected from MGv to higher-order areas, they may become more heterogeneous due to the convergent projections, similar to how tonotopic maps in higher-order areas become less clear (Issa et al., 2014; Ohga et al., 2018). This transformation of functional organization could render the maps too disordered to be clearly represented with the spatial resolution of macroscale calcium imaging, making it difficult to determine the presence or absence of orthogonal maps in A1 and A2. This underscores the need for higher-resolution mapping at the single-neuron scale.

In the field of otolaryngology, the phenomenon of elderly individuals with normal hearing thresholds struggling with speech perception, despite adjustments in sound

pressure levels (Eisenberg et al., 1995; Peters et al., 1998; Summers and Molis, 2004; George et al., 2006), is a significant clinical concern. This issue underscores the importance of temporal processing in the central nervous system beyond the sensitivity of peripheral hearing receptors (Moore et al., 1992; Noordhoek et al., 2001; Bernstein et al., 2013). Studies using mouse models have explored changes in the central auditory system following the onset of hearing loss, reporting qualitative changes in the arrangement of tonotopic maps in the auditory cortex (Yang et al., 2011; Persic et al., 2020; Resnik and Polley, 2021). Therefore, it is also possible that the functional organization representing temporal features is degraded in individuals with impaired speech perception. The current study provides an experimental foundation for future research using animal models to investigate the pathology of the auditory central nervous system associated with speech intelligibility disorders, where cochlear function appears normal.

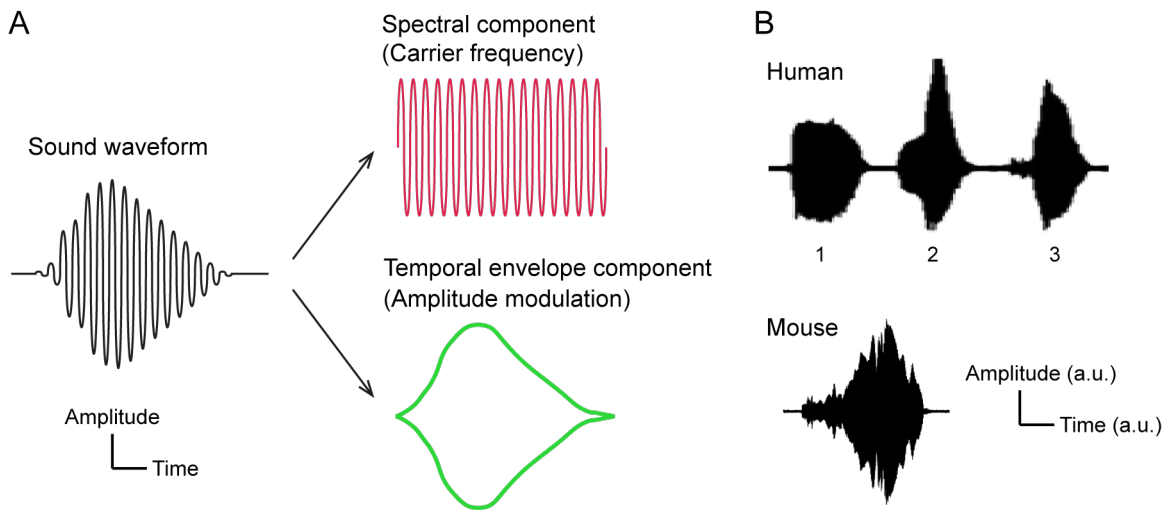

Figure 1. Explanation of two acoustic components in sound waveforms.
(A) Two primary acoustic components of sound waveforms: spectral and temporal. The spectral component refers to the rapid oscillations in sound waves, corresponding to the frequency of sounds. This component, also known as carrier frequencies or temporal fine structures, is crucial for perceiving pitch. The temporal envelope component represents the slow, overall changes in the amplitude of sound signals over time, and is important for perceiving vocal sounds. **(B)** Example waveforms of vocalized sounds of mammals. The top waveform shows examples of human vocalization. The waveform labeled 1, 2, and 3 correspond to the phonemes “go”, “le”, and “ku”, respectively. The bottom waveform is an example of mouse vocalization. Both examples demonstrate clear temporal envelopes, a feature known to be shared across vocalization of various mammalian species. The top images are adapted from Milne et al., 2016 under a Creative Commons license. The bottom image was made by reusing our past data (Aponte et al., 2021).

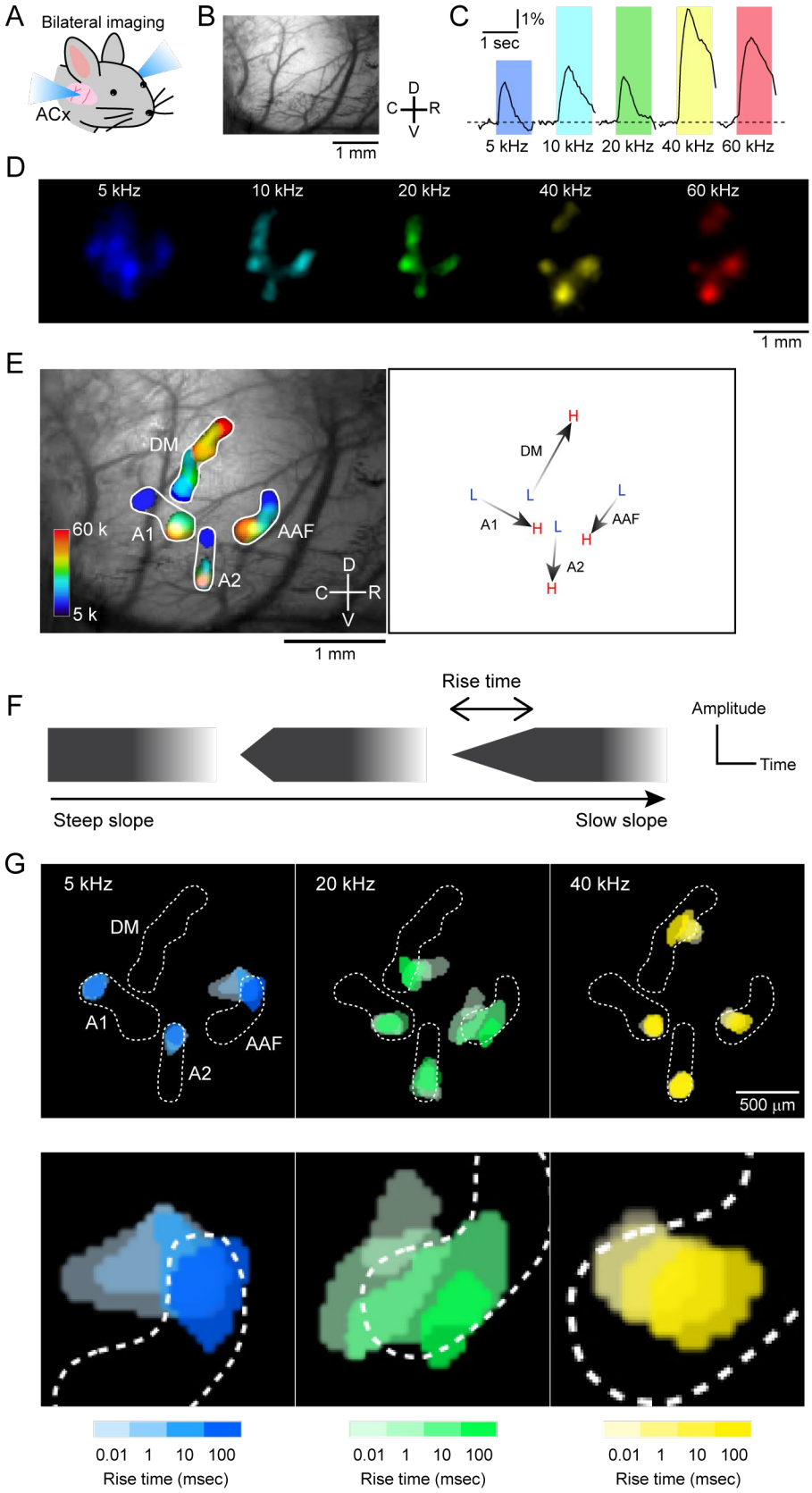

Figure 2. Visualizing envelope steepness maps orthogonal to tonotopy in right auditory cortex

(A) Illustration of bilateral auditory cortex imaging. ACx, auditory cortex. **(B)** An example field of view of the right auditory cortex. C, caudal; D, dorsal; R, rostral; V, ventral. **(C)** Example temporal profiles of GCaMP6f signals in AAF in response to tones with frequencies of 5–60 kHz. Tonal presentation periods were shown in color. **(D)** Example images of peak responses for 5–60 kHz tones. **(E)** Left: Tonotopic maps of the right auditory cortex visualized by superimposing thresholded response images shown in (D). AAF, anterior auditory field; A1, primary auditory field; A2, secondary auditory field; DM, dorsomedial field. Right: Low to high tonotopic directions of each auditory area in the example mouse. **(F)** Schematic explaining the change in rise ramp of tonal onset. **(G)** Example maps of rise-ramp steepness, obtained from the same mouse as above. Thresholded peak responses visualized using tones with rise-ramp times ranging from 0.01 to 100 msec were superimposed. The gradation in color indicates rise rimes of 0.01, 1, 10, and 100 msec from light to dark. Areas enclosed by white dots indicate the boundaries of tonotopic maps shown in (E). Magnified AAF images were shown at the bottom.

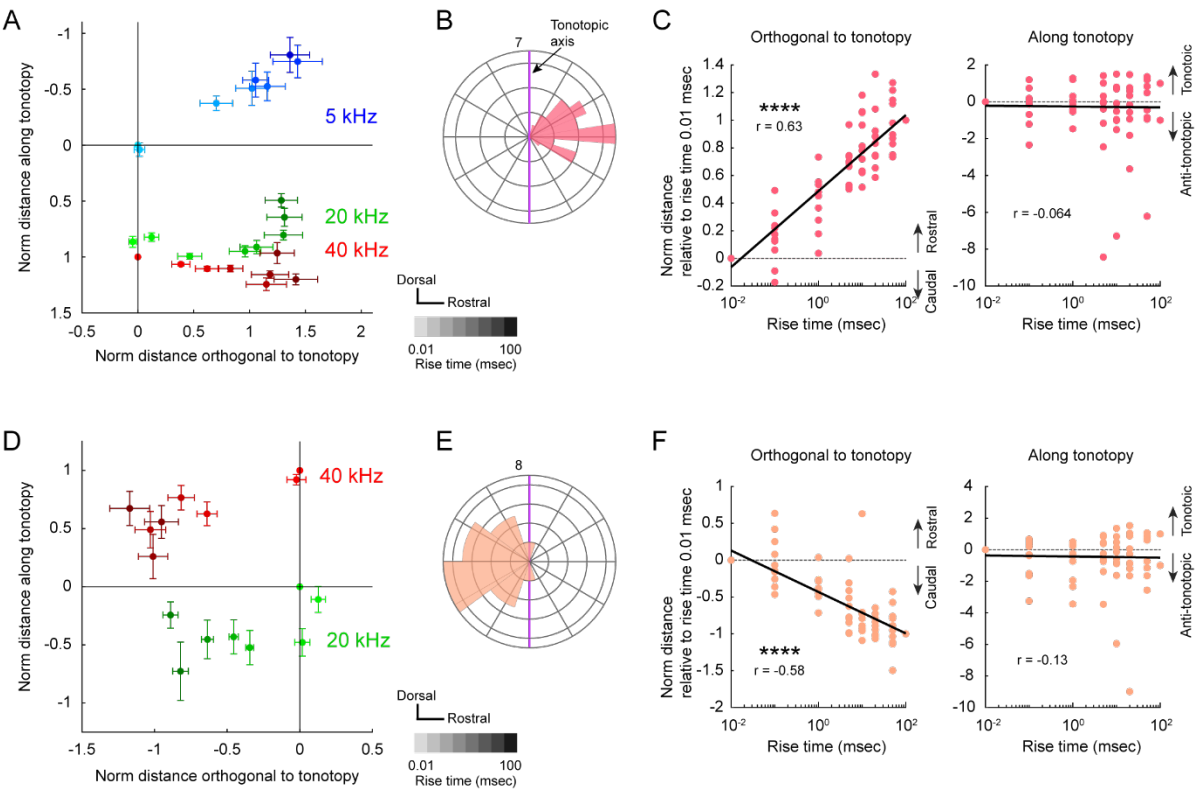

Figure 3. Presence of envelope maps in AAF and DM

(A) Group data showing the location of response peaks for tones with various rise-ramp times in AAF. The response location is evaluated using 5, 20, and 40 kHz tones with rise-ramp times ranging from 0.01 to 100 msec. The gradation in color indicates rise rimes of 0.01, 0.1, 1, 5, 10, 20, 50, and 100 msec from light to dark. Coordinates are normalized using the distance between peaks for 5 kHz and 40 kHz tones with a rise-ramp time of 0.01 msec, applied to both horizontal and vertical axes. Data are presented in mean ± SEM. **(B)** Circular histogram showing the direction of responses of rise-ramp times of 100 msec relative to those for rise-ramp times of 0.01 msec. Purple lines indicate the tonotopic axis. Data for 5, 20, 40 kHz are co-plotted. Orientation is consistent to (A). n = 30 plots from 10 mice. **(C)** Left: Relationship between rise-ramp time and the shift relative to the peak location for a rise-ramp time of 0.01 msec in the direction orthogonal to tonotopy in AAF. r = 0.63, ****p = 5.8×10^{-10} (Pearson's correlation); n = 80 plots from 10 mice. Right: Relationship between rise-ramp time and the shift from the peak location for rise-ramp time of 0.01 msec in the tonotopic direction. r = -0.064, p = 0.57. **(D)** Same as (A) but for DM. Coordinates are normalized using the distance between peaks for 20 kHz and 40 kHz tones with a rise-ramp time of 0.01 msec for horizontal and vertical axes. n = 30 plots from 10 mice. **(E)** Same as (B) but for DM.

(F) Same as (C) but for DM. Right, r = -0.58, ****p = 2.0×10^{-8}. Left, r = -0.13, p = 0.26; n = 80 plots from 10 mice.

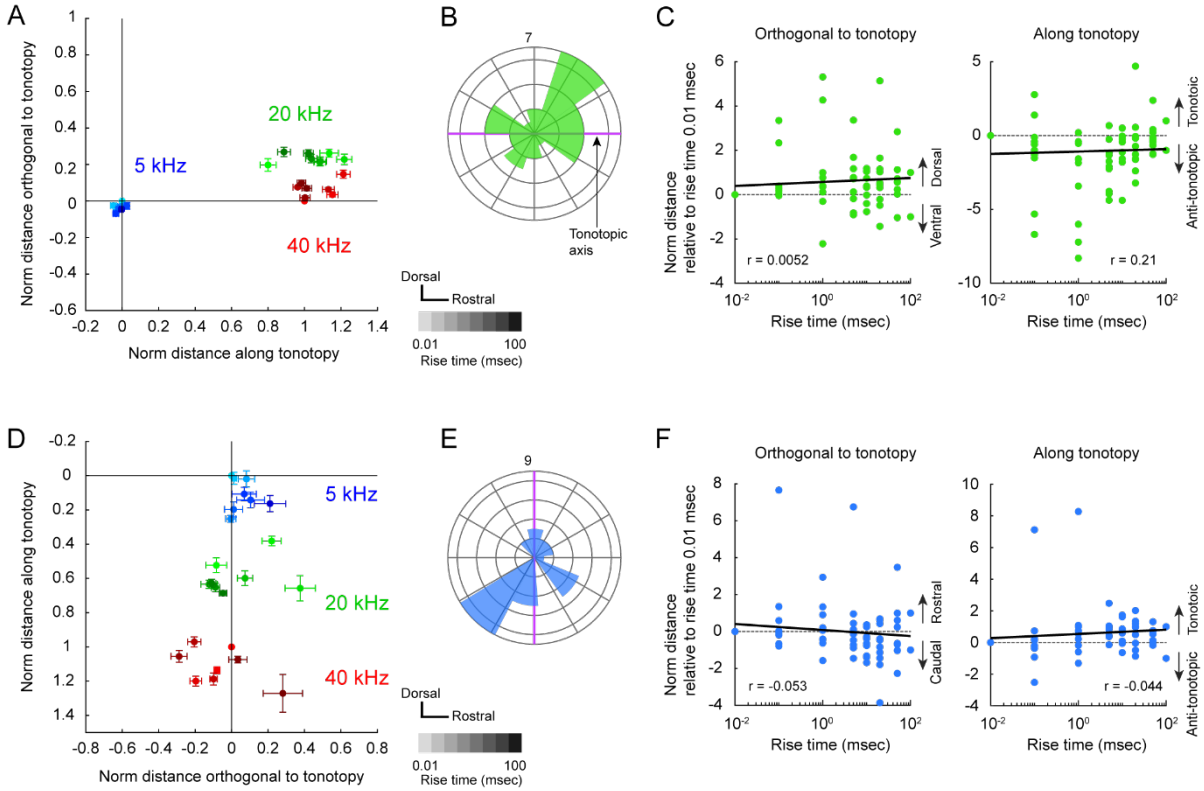

Figure 4. Absence of envelope maps in A1 and A2

(A) Group data showing the location of response peaks for tones with various rise-ramp times in A1. The response location is evaluated using 5, 20, and 40 kHz tones with rise-ramp times ranging from 0.01 to 100 msec. The gradation in color indicates rise rimes of 0.01, 0.1, 1, 5, 10, 20, 50, 100 msec from light to dark. Coordinates are normalized using the distance between peaks for 5 kHz and 40 kHz tones with a rise-ramp time of 0.01 msec, applied to both horizontal and vertical axes. Data are presented in mean ± SEM. **(B)** Circular histogram showing the direction of responses of rise-ramp times of 100 msec relative to those for rise-ramp times of 0.01 msec. Purple lines indicate the tonotopic axis. Data for 5, 20, 40 kHz are co-plotted. Orientation is consistent to (A). n = 30 plots from 10 mice. **(C)** Left: Relationship between rise-ramp time and the shift from the peak location for a rise-ramp time of 0.01 msec in the direction orthogonal to tonotopy in A1. r = 0.0052, p = 0.96 (Pearson's correlation); n = 80 plots from 10 mice. Right: Relationship between rise-ramp time and the shift from the peak location for rise-ramp time of 0.01 msec in the tonotopic direction. r = 0.21, p = 0.059. **(D)** Same as (A) but for A2. n = 29 plots from 10 mice. **(E)** Same as (B) but for A2. **(F)** Same as (C) but for A2. Right, r = -0.053, p = 0.64. Left, r = 0.044, p = 0.70; n = 77 plots from 10 mice.

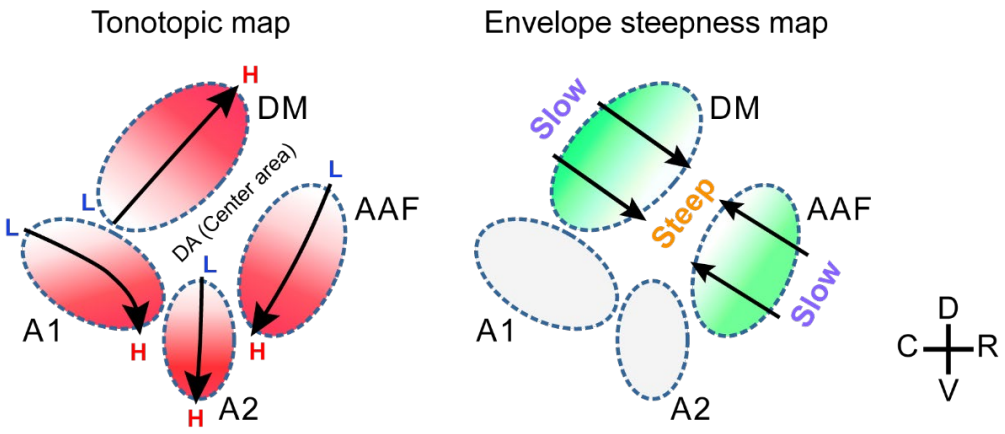

Figure 5. Summary illustration of two-dimensional maps present in mouse auditory cortex.

Left: Tonotopic maps in the mouse auditory cortex. The mouse auditory cortex is thought to comprise at least four tonotopic regions. In addition, there is a center area (annotated as the dorsoanterior field (DA) in Tsukano et al., 2017) between AAF and DM. Many literatures show that the center area does not respond to simple presentation of pure tones (Honma et al., 2013; Issa et al., 2014, 2016; Tsukano et al., 2015; Ceballo et al., 2019; Liu et al., 2019; Narayanan et al., 2022). Right: The envelope steepness map revealed in the current study. This map represents orthogonal to the tonotopic maps in AAF and DM. In AAF, the envelope steepness map progresses from slow to steep as it moves from rostral to caudal, while in DM, it progresses from slow to steep as it moves from caudal to rostral. The rise steepness is represented on a logarithmic scale.

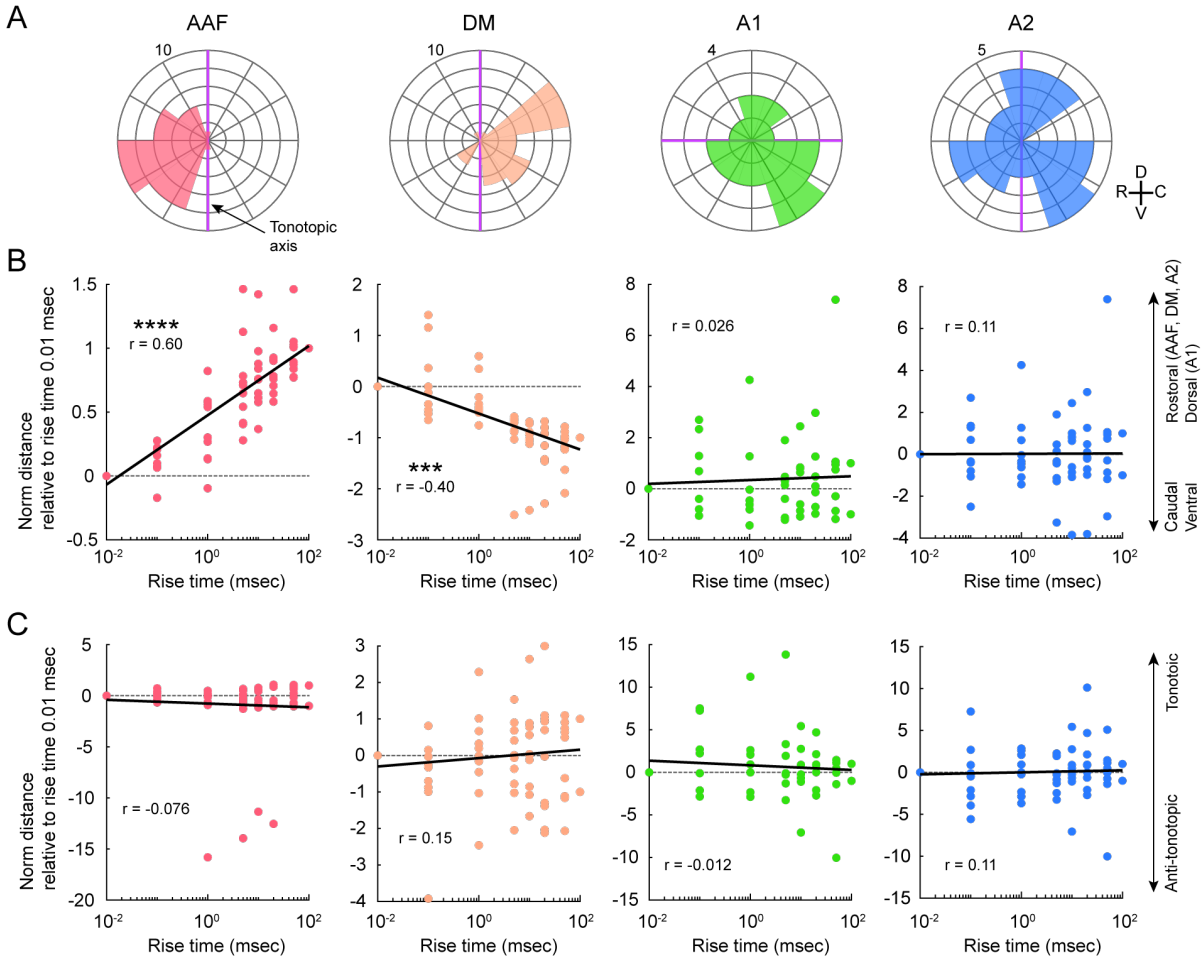

Figure S1. Envelope maps in the left auditory cortex.

(A) Circular histogram showing the direction of responses of rise-ramp times of 100 msec relative to those for rise-ramp times of 0.01 msec in the left auditory cortical areas. Purple lines indicate the tonotopic axis. Data for 5, 20, 40 kHz are co-plotted. AAF, n = 30 plots; DM, n = 30 plots; A1, n = 26 plots; A2, n = 30 plots from 10 mice. **(B)** Relationship between rise-ramp time and the shift from the peak location for a rise-ramp time of 0.01 msec in the direction orthogonal to tonotopy. AAF, r = 0.60, ****p = 3.9×10^{-9}, n = 80 plots (Pearson's correlation); DM, r = -0.40, ***p= 2.1×10^{-4}, n = 80 plots; A1, r = 0.026, p = 0.83, n = 72 plots; A2, r = 0.11, p = 0.33, n = 80 plots. **(C)** Relationship between rise-ramp time and the shift from the peak location for rise-ramp time of 0.01 msec in the tonotopic direction. AAF, r = -0.076, p = 0.50, n = 80 plots; DM, r = 0.15, p = 0.19, n = 80 plots; A1, r = -0.012, p = 0.92, n = 72 plots; A2, r = 0.11, p = 0.35, n = 80 plots.

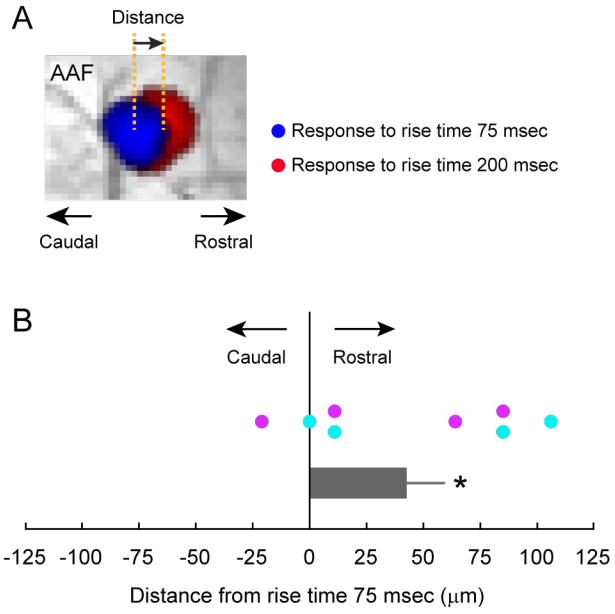

Figure S2. Envelope maps during the ramping phase.

(A) Explanation of analysis. Positional shifts of 5 kHz responses are evaluated between rise times of 75 and 200 msec. **(B)** Plots of rostro-caudal deviation between the peak responses for a rise-ramp time of 200 msec relative to that for a rise-ramp time of 75 msec in AAF. Positive values indicate that the response for a rise-ramp time of 200 msec is rostral to that for 75 msec, while negative values indicate it is caudal. Data for AAF in the right (cyan) and left hemisphere (purple) are co-plotted (n = 8 plots from 4 mice). Data are shown in mean ± SEM. *p = 0.039 (two-sided paired t-test).

Methods

Animals

The experimental procedures in the present study were approved by the Committee for Animal Care at Niigata University. All the experiments were performed in accordance with the approved guidelines and regulations. We used 5–7-wk-old Thy1-GCaMP6f mice (C57BL/6J-Tg(Thy1-GCaMP6f)GP5.5Dkim/J) acquired from Jackson Laboratories. The animals were housed in cages with *ad libitum* access to food pellets and water and kept on a 12-h light/dark cycle.

Auditory stimuli

Sound waveforms were made using an application (RPvdsEx) of a sound generator (System3, Tucker-Davis Technologies, Alachua, FL) at a sampling rate of 130 kHz. Waveforms were low-pass filtered at 100 kHz (3625, NF, Kanagawa, Japan). The auditory stimuli were transmitted to both ear canals simultaneously. Each ear canal received sound through a dedicated sound guide tube, which was connected to a speaker (EC1, Tucker-Davis Technologies). The sound intensity was set at ~60 dB SPL at the location of the ear canal. Sound intensity was calibrated using a sound level meter (NA-42S, RION, Kokubunji, Japan) with a microphone and pre-amplifier (UC54 and NH-17, RION). The sound duration was 1 sec with a rise/fall time ranging 0.01–200 msec across 8 conditions (0.01, 0.1, 1, 5, 10, 20, 50, and 100 msec; Figures 2–4, S1) or 2 conditions (75 and 200 msec; Figure S2). To visualize tonotopy in Figure 2, we used 1-sec tones with amplitude modulation at 20 Hz.

Surgery

Mice were deeply anesthetized using urethane (1.65 g/kg, i.p.; Wako, Osaka, Japan) and administered atropine (0.5 mg/kg, s.c.; Terumo, Tokyo, Japan). After local anesthesia using 1% Xylocaine (Aspen Pharmacare, Durban, South Africa), the skin and temporal muscle over the auditory cortex in both hemispheres were incised, leaving the skull intact. A piece of metal was attached to the skull with dental resin, and the head was fixed by screwing the metal piece onto a manipulator. The exposed surface of the skull was covered with liquid paraffin (Wako) to keep the skull transparent during imaging. The rectal temperature was maintained at 37°C during the surgery using a heat pad throughout experiments.

Calcium imaging

About 30 min after the completion of the surgery, macroscale calcium imaging was performed transcranially to reveal functional maps in the auditory cortex. Bilateral cortical images (128 × 168 pixels) were recorded using a cooled CCD camera system (ORCA-R2, Hamamatsu Photonics, Hamamatsu, Japan) attached to a microscope (M165FC, Leica Microsystems), which was controlled with RatioImaging Recorder (E.I.SOL, Minato-ku, Japan). GCaMP6f was excited by LED light (λ = 460 nm; UHP-Mic-LED-460, Opto-Line, Warabi, Japan), and its fluorescence was detected through band-bass filters (λ = 500–550 nm). The area covered by one pixel was 21.2 × 21.2 mm. Images were taken at the sampling rate of 20 (Figures 3–4, S1) or 50 Hz (Figure S2).

Image analyses were performed using custom-made MATLAB codes (MathWorks, Natick, MA). Images of the same stimuli were averaged over 15 trials. These images were converted to $\Delta F/F_0$, with the baseline intensity (F_0) obtained by averaging the intensity values during the prestimulus period of 250 msec. A spatial averaging of 5 × 5 pixels was applied. Temporal averaging was performed on images over frames from 0–400 msec (Figures 3–4, S1) or 40–60 msec (Figure S2) following tonal onset. The locations of pixels with the maximum responses were identified within each auditory area in a semi-automated manner: Reference pixels were put within each auditory areas according to the overall spatial pattern, and pixels with the largest $\Delta F/F_0$ value were automatically found within a radius of 8 pixels from the reference point. When multiple pixels were identified as peak pixels with identical values, the averaged coordinate across these pixels was considered the peak location.

For the analysis presented in Figures 3–4, S1, a coordinate matrix of [3 frequencies × 8 rise steepness × 10 mice] was obtained for each auditory area. For each mice, the coordinate matrix of [3 frequencies × 8 rise steepness] was rotated so that the tonotopic direction between 5 kHz and 40 kHz (AAF, A1, and A2) or 20 kHz and 40 kHz (DM) with a rise time of 0.01 msec was aligned vertical (AAF, DM, and A2) or horizontal (A1). When creating two-dimensional plots, the distance between 5 kHz and 40 kHz (AAF, A1, and A2) or 20 kHz and 40 kHz (DM) was normalized to 1 to average across mice. After averaging coordinate data across frequencies, angles of each data point relative to the tonotopic axis were plotted in a circular histogram, with the data distributed across 10 bins. In plots showing the relationships between rise time and distance, the distance between the peak locations for rise times of 0.01 and 100 ms was set to 1 for each

frequency for normalization. NaN was put in data points when clear peak location of auditory areas was not identified. To provide the temporal profile of GCaMP signals in Figure 2C, the intensity values within a 5 × 5 pixel area centered around the peak location of the response were averaged. For the analysis of Figure S2, the rostrocaudal difference in the peak locations for 5 kHz between rise times of 75 msec and 200 msec was measured.

Statistics

Statistical analyses were conducted using MATLAB. Correlations were evaluated using Pearson's correlation. The two-sided paired t-test was used to evaluate whether data distribution was significantly different than zero. Only p-values less than 0.05 were highlighted by asterisks.

Author Contributions

K.T. and H.T. conceptualized the project. K.T., T.G., T.Y., and S.O. performed experiments. K.T., T.G., T.Y., S.O., and H.T. analyzed the data. K.T. and H.T. wrote the manuscript. K.T., H.T., and A.H. revised the manuscript. All the authors approved the final version of the manuscript.

Acknowledgements

This work was supported by JSPS KAKENHI No. 20K09750 (T.K.). We thank Michellee M. Garcia at the University of North Carolina at Chapel Hill for her helpful comments and for editing the English in this manuscript.

Conflict of Interests

Authors declare no conflict of interest.

Data and codes availability

Data and codes supporting this study are available from corresponding authors upon reasonable requests.

References

Agarwalla S, De A, Bandyopadhyay S (2023) Predictive Mouse Ultrasonic Vocalization Sequences: Uncovering Behavioral Significance, Auditory Cortex Neuronal Preferences, and Social-Experience-Driven Plasticity. J Neurosci 43:6141–6163.

Aizawa N, Eggermont JJ (2006) Effects of noise-induced hearing loss at young age on voice onset time and gap-in-noise representations in adult cat primary auditory cortex. JARO - J Assoc Res Otolaryngol 7:71–81.

Aponte DA, Handy G, Kline AM, Tsukano H, Kato HK (2021) Recurrent network dynamics shape direction selectivity in primary auditory cortex. Nat Commun:1–15.

Baba H, Tsukano H, Hishida R, Takahashi K, Horii A, Takahashi S, Shibuki K (2016) Auditory cortical field coding long-lasting tonal offsets in mice. Sci Rep 6:34421.

Bandyopadhyay S, Shamma SA, Kanold PO (2010) Dichotomy of functional organization in the mouse auditory cortex. Nat Neurosci 13:361–368 Available at: http://dx.doi.org/10.1038/nn.2490.

Barton B, Venezia JH, Saberi K, Hickok G, Brewer AA (2012) Orthogonal acoustic dimensions define auditory field maps in human cortex. Proc Natl Acad Sci U S A 109:20738–20743.

Baumann S, Griffiths TD, Sun L, Petkov CI, Thiele A, Rees A (2011) Orthogonal representation of sound dimensions in the primate midbrain. Nat Neurosci 14:423–425.

Baumann S, Joly O, Rees A, Petkov CI, Sun L, Thiele A, Griffiths TD (2015) The topography of frequency and time representation in primate auditory cortices. Elife 2015:1–15.

Bernstein JGW, Mehraei G, Shamma S, Gallun FJ, Theodoroff SM, Leek MR (2013) Spectrotemporal Modulation Sensitivity as a Predictor of Speech Intelligibility for Hearing-Impaired Listeners. J Am Acad Audiol 24:293–306.

Bizley JK, Nodal FR, Nelken I, King AJ (2005) Functional organization of ferret auditory cortex. Cereb Cortex 15:1637–1653.

Calhoun G, Chih-Ting C, Kanold PO (2023) Bilateral widefield calcium imaging reveals circuit asymmetries and lateralized functional activation of the mouse

auditory cortex. Proc Natl Acad Sci 120:e2219340120.

Ceballo S, Piwkowska Z, Bourg J, Daret A, Bathellier B (2019) Targeted Cortical Manipulation of Auditory Perception. Neuron 104:1–12.

Chen G, Kang B, Lindsey J, Druckmann S, Li N (2021) Modularity and robustness of frontal cortical networks. Cell 184:3717-3730.e24.

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. J Acoust Soc Am 97:585–592.

Eggermont JJ (2001) Between sound and perception: Reviewing the search for a neural code. Hear Res 157:1–42.

Eisenberg LS, Dirks DD, Bell TS (1995) Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. J Speech Hear Res 38:222–233.

Fishbach A, Nelken I, Yeshurun Y (2001) Auditory edge detection: A neural model for physiological and psychoacoustical responses to amplitude transients. J Neurophysiol 85:2303–2323.

Frye RE, Fisher JMG, Coty A, Zarella M, Liederman J, Halgren E (2007) Linear coding of voice onset time. J Cogn Neurosci 19:1476–1487.

George ELJ, Festen JM, Houtgast T (2006) Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. J Acoust Soc Am 120:2295–2311.

Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG, Polley DB (2012) Robustness of Cortical Topography across Fields, Laminae, Anesthetic States, and Neurophysiological Signal Types. J Neurosci 32:9159–9172.

Heil P (1997) Auditory cortical onset responses revisited. I. First-spike timing. J Neurophysiol 77:2616–2641.

Heil P (1998) Further observations on the threshold model of latency for auditory neurons. Behav Brain Res 95:233–236.

Heil P (2001) Representation of sound onsets in the auditory system. Audiol Neuro-Otology 6:167–172.

Heil P, Irvine DRF (1998) Functional specialization in auditory cortex: Responses to frequency- modulated stimuli in the cat's posterior auditory field. J Neurophysiol 79:3041–3059.

Honma Y, Tsukano H, Horie M, Ohshima S, Tohmi M, Kubota Y, Takahashi K,

Hishida R, Takahashi S, Shibuki K (2013) Auditory cortical areas activated by slow frequency-modulated sounds in mice. PLoS One 8:e68113.

Issa JB, Haeffele BD, Agarwal A, Bergles DE, Young ED, Yue DT (2014) Multiscale Optical Ca2+ Imaging of Tonal Organization in Mouse Auditory Cortex. Neuron 83:944–959.

Issa JB, Haeffele BD, Young ED, Yue DT (2016) Multiscale mapping of frequency sweep rate in mouse auditory cortex. Hear Res 344:207–222.

Joachimsthaler B, Uhlmann M, Miller, Frank, G Ehret, Kurt S (2014) Quantitative analysis of neuronal response properties in primary and higher-order auditory cortical fields of awake house mice (Mus musculus). Eur J Neurosci 39:904–918.

Khouri L, Lesica NA, Grothe B (2011) Impaired auditory temporal selectivity in the inferior colliculus of aged mongolian gerbils. J Neurosci 31:9958–9970.

Kline AM, Aponte DA, Tsukano H, Giovannucci A, Kato HK (2021) Inhibitory gating of coincidence-dependent sensory binding in secondary auditory cortex. Nat Commun 12:1–17.

Krishna BS, Semple MN (2000) Auditory temporal processing: Responses to sinusoidally amplitude- modulated tones in the inferior colliculus. J Neurophysiol 84:255–273.

Langner G, Albert M, Briede T (2002) Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla (Chinchilla laniger). Hear Res 168:110–130.

Langner G, Dinse HR, Godde B (2009) A map of periodicity orthogonal to frequency representation in the cat auditory cortex. Front Integr Neurosci 3:1–11.

Liu J, Whiteway MR, Sheikhattar A, Butts DA, Babadi B, Kanold PO (2019) Parallel Processing of Sound Dynamics across Mouse Auditory Cortex via Spatially Patterned Thalamic Inputs and Distinct Areal Intracortical Circuits. Cell Rep 27:872–885.

Lu T, Liang L, Wang X (2001) Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. J Neurophysiol 85:2364–2380.

Ma L, Tai X, Su L, Shi L, Wang E, Qin L (2013) The Neuronal Responses to Repetitive Acoustic Pulses in Different Fields of the Auditory Cortex of Awake Rats. PLoS One 8:1–10.

Maffei C, Capasso R, Cazzolli G, Colosimo C, Dell'Acqua F, Piludu F, Catani M, Miceli G (2007) Pure word deafness following left temporal damage: Behavioral and neuroanatomical evidence from a new case. Cortex 97:240–254.

Milne AE, Mueller JL, Männel C, Attaheri A, Friederici AD, Petkov CI (2016) Evolutionary origins of non-adjacent sequence processing in primate brain potentials. Sci Rep 6:1–10.

Moore BC, Peters RW, Glasberg BR (1992) Detection of temporal gaps in sinusoids by elderly subjects with and without hearing loss. J Acoust Soc Am 92:1923–1932.

Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2002) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. J Neurophysiol 87:1723–1737.

Narayanan DP, Tsukano H, Kline AM, Onodera K, Kato HK (2022) Biological constraints on stereotaxic targeting of functionally-defined cortical areas. Cereb Cortex 33:3293–3310.

Noordhoek IM, Houtgast T, Festen JM (2001) Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. J Acoust Soc Am 109:1197–1212.

Nourski K V., Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29:15564–15574.

Ohga S, Tsukano H, Horie M, Terashima H, Nishio N, Kubota Y, Takahashi K, Hishida R, Takebayashi H, Shibuki K (2018) Direct Relay Pathways from Lemniscal Auditory Thalamus to Secondary Auditory Field in Mice. Cereb Cortex 28:4424–4439.

Persic D, Thomas ME, Pelekanos V, Ryugo DK, Takesian AE, Krumbholz K, Pyott SJ (2020) Regulation of auditory plasticity during critical periods and following hearing loss. Hear Res 397:107976.

Peters RW, Moore BC, Baer T (1998) Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. J Acoust Soc Am 103:577–587.

Peterson AJ, Heil P (2021) A simplified physiological model of rate-level functions of

645 auditory-nerve fibers. Hear Res 406:108258.
646 Phillips DP (1998) Factors shaping the response latencies of neurons in the cat's
647 auditory cortex. Behav Brain Res 93:33–41.
648 Phillips DP, Hall SE, Boehnke SE (2002) Central auditory onset responses, and
649 temporal asymmetries in auditory perception. Hear Res 167:192–205.
650 Resnik J, Polley DB (2021) Cochlear neural degeneration disrupts hearing in
651 background noise by increasing auditory cortex internal noise. Neuron
652 109:984-996.e4.
653 Romero S, Hight A, Clayton K, Resnik J, Williamson R, Hancock K, Polley D (2020)
654 Cellular and Widefield Imaging of Sound Frequency Organization in Primary and
655 Higher Order Fields of the Mouse Auditory Cortex. Cereb Cortex 30:1603–1622.
656 Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic
657 aspects. Philos Trans R Soc Lond B Biol Sci 336:367–373.
658 Schnupp JWH, Garcia-Lazaro JA, Lesica NA (2015) Periodotopy in the gerbil inferior
659 colliculus: Local clustering rather than a gradient map. Front Neural Circuits
660 9:1–21.
661 Schreiner CE, Langner G (1988) Periodicity coding in the inferior colliculus of the cat.
662 II. Topographical organization. J Neurophysiol 60:1823–1840.
663 Shannon R V., Zeng F-G, Kamath V, Wygonski J, Ekelid M (1995) Speech
664 Recognition with Primarily Temporal Cues. Science (80-) 270:303–304.
665 Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in
666 auditory perception. Nature 416:87–90.
667 Stiebler I, Neulist R, Fichtel I, Ehret G (1997) The auditory cortex of the house
668 mouse: Left-right differences, tonotopic organization and quantitative analysis of
669 frequency representation. J Comp Physiol - A Sensory, Neural, Behav Physiol
670 181:559–571.
671 Suga N (1971) Responses of inferior collicular neurones of bats to tone bursts with
672 different rise times. J Physiol 217:159–177.
673 Summers V, Molis MR (2004) Speech Recognition in Fluctuating and Continuous
674 Maskers: Effects of Hearing Loss and Presentation Level. J Speech, Lang Hear
675 Res 47:245–256.
676 Takahashi K, Hishida R, Kubota Y, Kudoh M, Takahashi S, Shibuki K (2006)
677 Transcranial fluorescence imaging of auditory cortical plasticity regulated by

acoustic environments in mice. Eur J Neurosci 23:1365–1376.

Tsukano H, Horie M, Bo T, Uchimura A, Hishida R, Kudoh M, Takahashi K, Takebayashi H, Shibuki K (2015) Delineation of a frequency-organized region isolated from the mouse primary auditory cortex. J Neurophysiol 113:2900–2920.

Tsukano H, Horie M, Hishida R, Takahashi K (2016) Quantitative map of multiple auditory cortical regions with a stereotaxic fine-scale atlas of the mouse brain. Sci Rep:1–12.

Tsukano H, Horie M, Ohga S, Takahashi K, Kubota Y, Hishida R, Takebayashi H, Shibuki K (2017) Reconsidering Tonotopic Maps in the Auditory Cortex and Lemniscal Auditory Thalamus in Mice. Front Neural Circuits 11:14.

Vasquez-Lopez SA, Weissenberger Y, Lohse M, Keating P, King AJ, Dahmen JC (2017) Thalamic input to auditory cortex is locally heterogeneous but globally tonotopic. Elife 6:e25141.

Warrier CM, Zatorre RJ (2004) Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. Brain 127:1616–1625.

Yang S, Weiner BD, Zhang LS, Cho SJ, Bao S (2011) Homeostatic plasticity drives tinnitus perception in an animal model. Proc Natl Acad Sci U S A 108:14974–14979.

Zhou Y, Wang X (2010) Cortical processing of dynamic sound envelope transitions. J Neurosci 30:16741–16754.