



Organization of auditory areas in the superior temporal gyrus of marmoset monkeys revealed by real-time optical imaging

Masataka Nishimura¹ · Makoto Takemoto¹ · Wen-Jie Song^{1,2}

Received: 11 May 2017 / Accepted: 18 November 2017 / Published online: 28 November 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

The prevailing model of the primate auditory cortex proposes a core–belt–parabelt structure. The model proposes three auditory areas in the lateral belt region; however, it may contain more, as this region has been mapped only at a limited spatial resolution. To explore this possibility, we examined the auditory areas in the lateral belt region of the marmoset using a high-resolution optical imaging technique. Based on responses to pure tones, we identified multiple areas in the superior temporal gyrus. The three areas in the core region, the primary area (A1), the rostral area (R), and the rostrotemporal area, were readily identified from their frequency gradients and positions immediately ventral to the lateral sulcus. Three belt areas were identified with frequency gradients and relative positions to A1 and R that were in agreement with previous studies: the caudolateral area, the middle lateral area, and the anterolateral area (AL). Situated between R and AL, however, we identified two additional areas. The first was located caudoventral to R with a frequency gradient in the ventrocaudal direction, which we named the medial anterolateral (MAL) area. The second was a small area with no obvious tonotopy (NT), positioned between the MAL and AL areas. Both the MAL and NT areas responded to a wide range of frequencies (at least 2–24 kHz). Our results suggest that the belt region caudoventral to R is more complex than previously proposed, and we thus call for a refinement of the current primate auditory cortex model.

Keywords Frequency gradient · Optical imaging · Primate auditory cortex · Tonotopy

Introduction

In recent years, the common marmoset, a highly vocal New World monkey, has become an increasingly important primate model for auditory research (reviewed in Wang 2013). A three-region model of the primate auditory cortex has been proposed, consisting of the core region, the belt region surrounding the core, and the parabelt region lateral to the belt (reviewed in Kaas and Hackett 2000). Each region has multiple areas or subdivisions; the core region, for example,

has three areas, the primary area (A1), the rostral area (R), and the rostrotemporal area (RT) (Kaas and Hackett 2000). The belt region caudal to A1 is called the caudal belt and consists of two areas, the caudomedial area (CM) and the caudolateral (CL) area, and the belt region lateral/medial to the core is called the lateral/medial belt region, each with three areas (Kaas and Hackett 2000). The parcellation of the primate auditory cortex is mainly based on the cyto-, myelo-, and chemoarchitectonic features of each area, the anatomical connections between areas, and inputs from the auditory thalamus (Kaas and Hackett 2000). The same three-region model has been proposed for the marmoset auditory cortex based on cyto- and chemoarchitectonic features and anatomical connections (de la Mothe et al. 2006a, b; Paxinos et al. 2012).

Electrophysiological studies have also been performed extensively on the marmoset core region to examine the organization and function of each area (Aitkin et al. 1986; Wang et al. 1995; Lu et al. 2001; Wang and Kadia 2001; Nagarajan et al. 2002; Lu and Wang 2004; Philibert et al. 2005; Bendor and Wang 2005, 2008; Feng and Wang 2017).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00429-017-1574-0>) contains supplementary material, which is available to authorized users.

✉ Wen-Jie Song
song@kumamoto-u.ac.jp

¹ Department of Sensory and Cognitive Physiology, Graduate School of Medical Sciences, Kumamoto University, Kumamoto 860-8556, Japan

² Program for Leading Graduate Schools HIGO Program, Kumamoto University, Kumamoto 860-8556, Japan

The results of these studies on the organization of the core region are consistent with the current model. The belt regions of marmosets, however, have rarely been targeted for electrophysiological studies, except for the CM area, which lies in the caudal belt located on the ventral bank of the lateral sulcus (LS) (Kajikara et al. 2005). Studies on the belt areas of other primates are also limited in number (Rauschecker et al. 1995; Bieser and Müller-Preuss 1996; Recanzone et al. 2000; Tian et al. 2001; Rauschecker and Tian 2004; Tian and Rauschecker 2004; Sweet et al. 2005; Petkov et al. 2006, 2008). The pioneering work by Rauschecker et al. (1995) identified the CL area and two further areas in the macaque lateral belt and revealed contrasting response features of neurons in the three regions. However, in this and later studies by the same group, the belt areas were examined only along a single, narrow trajectory (Tian et al. 2001; Rauschecker and Tian 2004; Tian and Rauschecker 2004). Although dense measurements of activity across the cortical sheet are essential to produce a reliable functional map of a cortical region, no high-density mapping has been carried out on the lateral belt region. Functional magnetic resonance imaging (fMRI) studies have been carried out on the macaque auditory cortex (Petkov et al. 2006, 2008), but fMRI has limited spatial and temporal resolution.

In marmosets, all of the known lateral belt areas, as well as a large portion of all three core areas and part of the parabelt, are on the superior temporal gyrus [STG; the gyrus between the LS and superior temporal sulcus (STS)] (de la Mothe et al. 2006a, b). Exploiting this structural feature, we aimed to map the auditory areas on the STG of marmosets with a focus on the lateral belt region, using a real-time, high-resolution optical imaging technique. Identification of all auditory areas in the lateral belt should help in clarifying the functional specialization of the region. Optical imaging is a powerful tool for the identification and characterization of cortical auditory areas in rodents because of its high temporal and spatial resolution (Horikawa et al. 1996, 2001; Song et al. 2006; Nishimura et al. 2007; Nishimura and Song 2014). In addition to confirming the three core areas, our results revealed five areas in the ventral belt region, two of which we believe to be newly identified areas of the lateral belt.

Materials and methods

Animals

All experiments were performed according to the Guidelines for Use of Animals in Experiments of Kumamoto University. The protocol was approved by the Committee of Animal Experiments of Kumamoto University. Three male common marmosets (referred to as marmosets L, M, and N; 20–22

months old) were used in this study. For surgery, anesthesia was induced with an intramuscular injection of a mixture of ketamine (40 mg/kg) and xylazine (4 mg/kg) and maintained by supplemental doses (half of the initial dose per hour). Further ketamine (15 mg/kg) was added for immobilization 30 min after every supplemental dose. To suppress bradycardia, atropine sulfate (0.25 mg/kg) was subcutaneously administered twice at a 6-hour interval, with the initial injection given 20 min after the induction of anesthesia. During recording, ketamine (20 mg/kg) and xylazine (2 mg/kg) were injected intramuscularly every 45 min. A small metal plate was mounted on the dorsal surface of the skull with screws and dental cement for fixation of the head. The skull and dura mater over the left superior temporal cortex were removed for staining the cortex with voltage-sensitive dye. A heating pad was used to maintain the body temperature around 37 °C. Heart rate and cardiac rhythm were monitored throughout the experiments (100–140 bpm).

Voltage-sensitive dye imaging

The cortex was stained with the voltage-sensitive dye RH-795 (0.5 mg/ml in saline; Invitrogen, Grand Island, NY, USA) for 2 h (Grinvald et al. 1994; Song et al. 2006). A piece of Cupro unwoven cloth (Asahi Kasei, Osaka, Japan) slightly smaller than the skull opening was placed gently on the cortical surface, to which a drop of RH-795 solution had been applied. More RH-795 solution was added until the cloth was submerged. The cloth was replaced every 30 min, and RH-795 solution was added to every replacement. Fluorescent optical signals from the cortex were recorded using a 100 × 100 pixel CMOS imaging system (MiCAM Ultima, Brain Vision, Tokyo, Japan). The sampling interval of the optical signal was 2 ms. The size of the recording field was 6.25 × 6.25 mm² for marmosets L and M, and 10.0 × 10.0 mm² for marmoset N. One edge of the recording field was set parallel to the midline so that the orientation of the recording fields was similar across animals. The noise from the heartbeat has been shown to significantly interfere with optical recordings in vivo (Maeda et al. 2001; Inagaki et al. 2003). To suppress heartbeat noises, recordings were triggered when the R wave of the electrocardiogram exceeded a threshold, and evoked responses were obtained by subtracting recordings without stimulus presentation from those with stimulus presentation (Song et al. 2006; Nishimura et al. 2007). Fractional optical signals ($\Delta F/F_0$ s) were calculated and analyzed offline with custom-made software written in C++. Because we used a sequence of tones for stimulation, the subtraction technique had a limited effect on the extraction of evoked responses by tones other than the first one, due to the fluctuation in heart rate. We suppressed slow fluctuations in $\Delta F/F_0$ by differentiating the subtracted signal [$\Delta(\Delta F/F_0)$; see “Results”].

Acoustic stimuli

We used a broadband noise to determine the extent of the auditory cortex on the STG, and pure tones to functionally tessellate the auditory cortex. The hardware, software, and calibration methods for sound stimulation have been described previously (Nishimura and Song 2014). Briefly, tones and broadband noise were presented via two earphones (ATH-C602, Audio-Technica, Tokyo, Japan) driven separately by a two-channel headphone buffer (HB7, Tucker-Davis Technologies, Alachua, FL, USA). An ear speculum was attached to each earphone to spatially confine the sounds. The frequency response of each earphone was measured with a sound level meter (type 2610 with a model 4191 microphone, Brüel & Kjær, Nærum, Denmark) at a distance of 5 mm from the tip of the speculum.

To generate noise with a flat frequency spectrum, we measured the frequency response characteristic of our sound presentation system 1000 times with a synthesized noise consisting of 2^{18} sampling points. Using the averaged frequency response characteristic, we inversely equalized the power of each frequency component in the broadband noise using discrete Fourier transform. The equalized band noise with components ranging from 250 to 32,000 Hz at 70 dB SPL was sampled at 195,312.5 Hz, and used as a frozen waveform for stimulation. The broadband noise was set to a window of 50-ms duration (10 ms cosine rise/fall and 30 ms plateau). To identify tonotopic areas in the cortex, for each recording, tone pips with various frequencies were presented at 60 dB SPL. We used a pseudorandom sequence of 9 tones with octave spacing (0.125–32 kHz) plus four additional frequencies (6, 12, 24 and 38.1 kHz) to cover the full frequency range. Consecutive tones in the sequence always were spaced wider than 1.25 octaves to minimize adaptation (Nieto-Diego and Malmierca 2016). Stimulation using such tone sequences is referred to as multiple-tone-pip stimulation. The measured frequency response characteristic of our sound delivery system was used for intensity control. The duration of each tone pip was also 50 ms (10 ms cosine rise/fall and 30 ms plateau), and the inter-tone interval was 150 ms.

Data analyses

First, a spatial low-pass filter (Gaussian windowed-sinc filter; < 0.4 cycle/pixel) was applied to the raw fluorescent optical signals and a temporal low-pass filter (Gaussian windowed-sinc filter; < 83.3 Hz) was then applied to $\Delta F/F_0$ s. To detect statistically significant responses, the t score for each pixel in each frame was calculated from $\Delta(\Delta F/F_0)$ s according to the following formula:

$$t(x, y, n) = \text{Mean}(\Delta(\Delta F/F_0)(x, y, n)) / \text{SEM}(\Delta(\Delta F/F_0)(x, y, n))$$

where x and y are coordinates in a recording frame, and n is the index of the recording frame. Mean() and SEM() denote the average and standard error across trials at location (x, y) and in frame (n) . There were 50–80 trials for each tone. The calculated t scores were encoded in color and superimposed on the basal fluorescent image F_0 with an opacity level function $O(p)$ using the following formula:

$$O(p) = \begin{cases} 0 & \text{if } p > p_{\text{threshold}} \\ 1.0 - p/p_{\text{threshold}} & \text{if } p \leq p_{\text{threshold}} \end{cases}$$

where p is the percentage of the Student's t distribution given by the calculated t score, two tailed. The $p_{\text{threshold}}$ in the analyses was 0.01 or 0.001. To illustrate the response amplitudes in a frame, such as the $\Delta F/F_0$ or $\Delta(\Delta F/F_0)$ image, we applied the same methods as described in our previous report (Nishimura and Song 2014).

Tonotopic gradient analysis To measure the direction and steepness of the tonotopic gradient in a cortical field, the centroid of significant responses for each tone frequency was first calculated from the t scores in that field, according to the following formulae:

$$C_x = \frac{\sum t_i x_i}{\sum t_i}, \quad C_y = \frac{\sum t_i y_i}{\sum t_i}$$

where t_i is the t score that reached significance at pixel (x_i, y_i) in the target field of an image frame, and (C_x, C_y) is the centroid. The centroids for responses to different frequencies were calculated, and a smooth curve was drawn on the centroids in the low-to-high-frequency direction, to measure the distance of tonotopic centroid shift. The steepness of the tonotopic centroid shift was calculated as the distance on the curve representing the frequency range of the target area, divided by the width of the frequency range in octaves.

Response latency We initially defined the response latency in a pixel as the interval from the onset of the tone stimulus to the point when $\Delta(\Delta F/F_0)$ first became significant ($p < 0.01$). Such a definition, however, introduced false short latencies when one or two sporadic frames with a significant $\Delta(\Delta F/F_0)$ were not followed by frames with significant responses. To suppress such errors, we defined latency as the interval from the onset of the tone stimulus to the point at which more than three frames showed significant responses, within a time window of five frames (10 ms). Thus, a five-frame window was shifted from the onset of each stimulus towards the end of the stimulus, until a latency that satisfied the definition was found. Paired t tests were used to statistically compare the response latencies in different auditory areas.

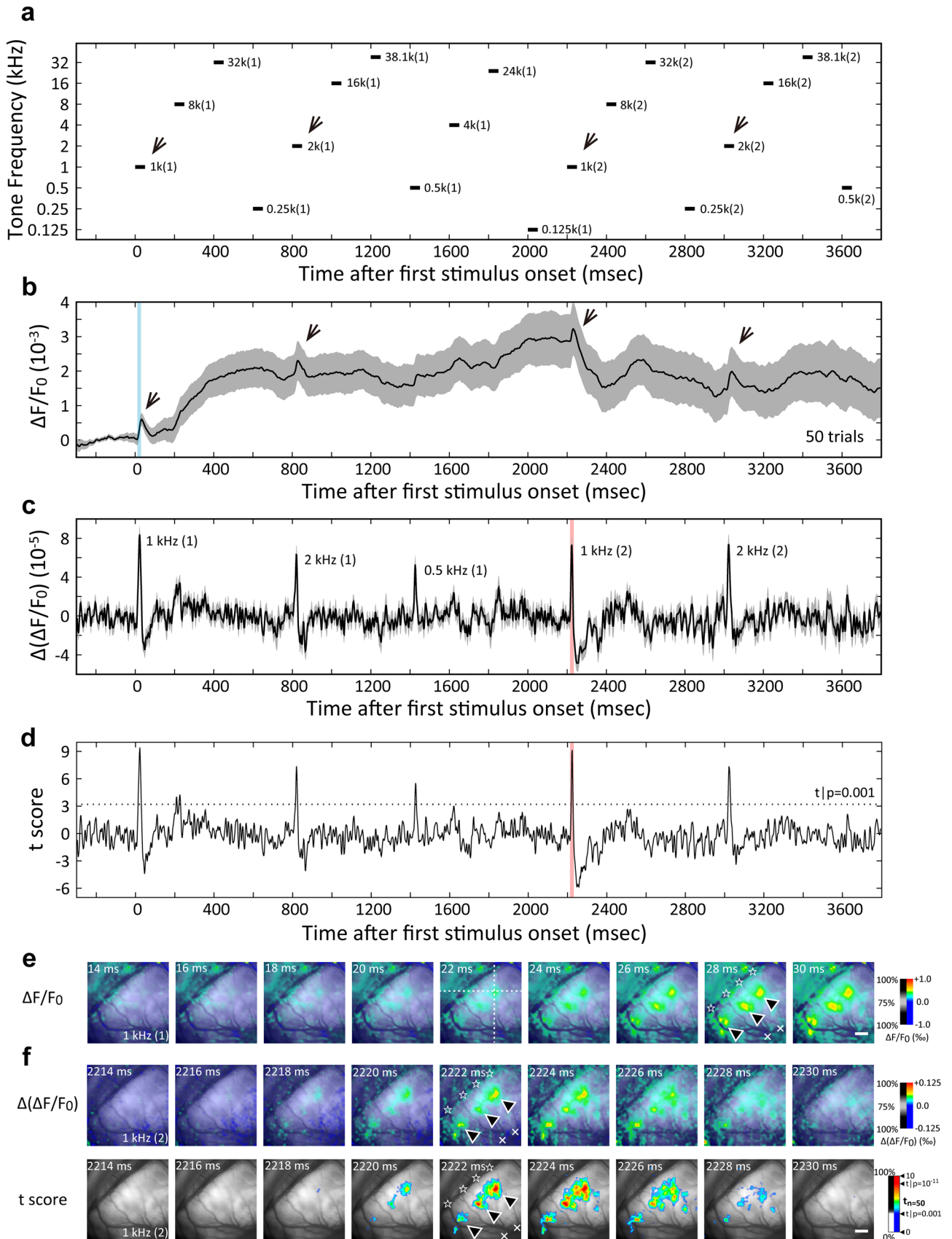


Fig. 1 Statistical analysis of cortical activity in response to multiple-tone-pip stimulation. **a** Order and timing of tone pips in a multiple-tone-pip stimulation. Each bar indicates the timing and frequency of a tone pip. **b** Example of recorded optical signals in response to repeated stimulation, recorded from the epicenter of the response shown in **(e)** (the cross-point of the dotted lines in the center graph). The black trace and gray-shaded area are the mean and SEM, respectively, of the fractional fluorescence change, $\Delta F/F_0$, of 50 recordings. The bluish area indicates the epoch illustrated in **(e)**. **c** Differentiation of optical signals [$\Delta(\Delta F/F_0)$]. The tone frequency and its number (1st or 2nd) are labeled next to the response peak. The black trace and gray-shaded area are the mean and SEM of the same 50 recordings as in **(b)**. **d** t score calculated from the mean and SEM in **c** for each sampling point. The dotted line indicates t scores at the significance level of $p=10^{-3}$. The reddish areas in **(c)** and **(d)** indicate the epoch illustrated in **f**. **e** Spatio-temporal patterns of $\Delta F/F_0$ during the epoch shown in **(b)**. $\Delta F/F_0$ was encoded in color and superimposed on the cortical surface image with the color scale shown to the right. The intersection point of the vertical and horizontal dotted lines marks the position of the recording shown in **b**, **c**, and **(d)**. Black arrowheads point to the three spots of activity. The time shown in the corner of each graph in this and all other figures indicates the time after the tone pip onset. Stars and crosses mark the lateral sulcus and the superior temporal sulcus, respectively, in this and all other figures. Scale bar is 1 mm. **f** Upper panel shows the spatial patterns of $\Delta(\Delta F/F_0)$ during the epoch shown in **c**. $\Delta(\Delta F/F_0)$ was encoded in color and superimposed on the cortical surface image with the color scale shown to the right. Lower panel shows the spatio-temporal patterns of t scores that reached significance ($p<0.001$). t scores were encoded in color and superimposed on the cortical surface image with the color scale shown to the right. This method of showing the spatio-temporal pattern of t scores also applies to Figs. 2, 3, 4, 5 and 6. Scale bar is 1 mm. Other figure conventions follow those in **(e)**

Results

Spatio-temporal patterns of tone-evoked activities in the auditory cortex

We first show an example to illustrate our methods of extraction and presentation of tone-evoked responses in the auditory cortex. We recorded 50–80 trials in each animal to test statistical significance of the responses. This design also reduced the number of marmosets required for the experiment. From our experience with rodents (Nishimura and Song 2014), RH-795 signals start to deteriorate 2–3 h after staining. To obtain a large number of recordings in a limited time, we repeatedly presented a multiple-tone-pip stimulation with tones ranging from 0.125 to 38.1 kHz in a sequence (Fig. 1a), and recorded the cortical responses to all stimuli in one file instead of recording the responses to each tone. The order of tones in the sequence was designed to separate the frequencies of neighboring tones to minimize adaptation. The fractional fluorescence signal ($\Delta F/F_0$) from a representative pixel (the cross-point of the two dotted lines in Fig. 1e) is shown in Fig. 1b; the line and the gray area indicate the mean and SEM, respectively, for 50 trials. The four arrows from left to right in Fig. 1b mark responses to the first 1 kHz, first 2 kHz, second 1 kHz, and second 2 kHz, respectively.

No obvious responses were observed for other frequencies, indicating the frequency selectivity of this pixel. The similar size of the responses to the two 1 kHz stimuli and the two 2 kHz stimuli, judged by eye, suggest our tone sequences induced little adaptation.

It is also clear from Fig. 1b that the slow fluctuations were larger in magnitude than the evoked responses. It is thus difficult to define a baseline for responses other than the first one. The fluctuations were largely caused by interference from the heartbeat (Maeda et al. 2001; Inagaki et al. 2003; Song et al. 2006). The signal around the first stimulus appeared to be relatively stable because the interferences were suppressed by subtracting a non-stimulus trial from a stimulus trial, with the phase of the heartbeat matched. Taking advantage of the fast rise of evoked responses, we calculated the differentiated signal [$\Delta(\Delta F/F_0)$] to suppress the slow fluctuation and stress the evoked response, as illustrated in Fig. 1c (also see Nishimura and Song 2014). After differentiation, a response to 0.5 kHz also became evident (Fig. 1c). Because we had 50–80 recordings for each tone sequence, we could calculate the t score of the signals for each pixel (e.g., Fig. 1d) and for each sampling time (i.e., each frame; e.g., Fig. 1f bottom panel; see below for illustration methods), to reveal statistically significant responses.

To illustrate the responses across pixels (space), we encode the magnitude of the response in color and show the time course of responses as a time series of frames. Figure 1e shows the spatio-temporal pattern of the response as a fractional fluorescence change ($\Delta F/F_0$) to the first stimulus (1 kHz; the blue bar region in Fig. 1b). Three spots of activity were observed (Fig. 1e, arrowheads). Clouds of pseudo-activity, however, also appeared in regions dorsal to the LS (the LS is marked with stars in Fig. 1e). These signals were judged as pseudo-activity because (1) no auditory response is expected in regions dorsal to the LS (de la Mothe et al. 2006a, b), and (2) they appeared without delay after the tone stimulation. It was impossible to illustrate the spatial pattern of $\Delta F/F_0$ activity evoked by tones later in the sequence because of the large fluctuation (see Fig. 1b). However, the spatio-temporal pattern of the differentiated signals [$\Delta(\Delta F/F_0)$], even those evoked by the second 1 kHz tone in the sequence, clearly exhibited three activation spots and much less pseudo-activity in the region dorsal to the LS (the upper panel of Fig. 1f illustrates the activity during the period marked by the red bar in Fig. 1c), as expected from the temporal data shown in Fig. 1c. Further, using t scores with a threshold for significance suppresses all non-significant values, sparing the three activity spots only (arrowheads in Fig. 1f, lower panel). In the following, we identify areas in the marmoset auditory cortex using the spatio-temporal patterns of significant t scores and p values. In general, the strength of tone-evoked responses should be positively related to t scores. To confirm this, we examined the

relationship between the response magnitude ($\Delta F/F_0$) and the t score for the first response shown in Fig. 1b, and found a significant positive correlation ($r=0.547$, $p < 10^{-150}$; supplementary Figure S1). We, therefore, also positively relate the strength of the tone-evoked responses to the t scores (= negatively with p values) in the following.

The three areas in the core region

Next, we aimed to identify the three areas in the core region, which will serve as references for analysis of belt areas; the three core areas are known to be immediately inferior to the LS, and can be identified using the frequency gradient and gradient reversal. We thus targeted a wide area of the cortex ($6.25 \times 6.25 \text{ mm}^2$) containing the LS, STS, and a large portion of the STG (Fig. 2a1). Based on a previous report on tonotopic gradients in the core region (Bendor and Wang 2005), the border between A1 and R in the core can be identified as a gradient reversal in the low-frequency region (Fig. 2a2). Marmoset L's early responses to tones $\leq 4 \text{ kHz}$ are shown in Fig. 2a3. Two spots of activity were observed in response to the 250 Hz tone, one in the center region of the STG (white arrowhead; Fig. 2a3, leftmost graph) and the other at a rostroventral location (double white arrowhead). When the tone frequency was increased to 500 Hz, an additional spot appeared at a dorsocaudal location (arrowhead in Fig. 2a3, second graph from left), and the locations of the first two spots shifted. Responses were also clearly observed in three spots for tones of 1 and 2 kHz (Fig. 2a3). As the frequency increased, the two rostroventral spots shifted toward each other and toward the LS, while the dorsocaudal spot shifted in a dorsocaudal direction (Fig. 2a3). For tones of 4 kHz, the two rostroventral spots appeared to merge (double arrowhead in Fig. 2a3, rightmost graph.). To examine the frequency gradient, we calculated the centroid of each activity spot (see Methods for definition), and superimposed all centroids together in Fig. 2a4, with frequencies encoded in color (as the two dorsal spots to 500 Hz were not well separated, they were treated as one spot; Fig. 2a4, arrow). Three gradients and two frequency reversals could be identified from Fig. 2a4, which defined three tonotopically organized areas (Fig. 2a5). Figure 2a5 also illustrates the approximate contour of activity in each area (dotted lines), obtained from the superimposition of ellipses approximating the contour of each activity spot in Fig. 2a3. Based on previous reports on areas in the core region (Bendor and Wang 2005, 2008; Fig. 2a2), the three areas were identified as A1, R, and RT, respectively, from the direction of their frequency gradients and their locations relative to the LS (Fig. 2a5; compare to Fig. 2a2). The steepness of the frequency gradient, estimated as the cortical distance traveled by the activity centroid (see "Materials and methods" for definition), was 0.42 mm/

octave in A1, 0.46 mm/octave in R, and 0.56 mm/octave in RT (Table 1).

The imaging field was located more caudally in marmoset M (Fig. 2b1–3). Only one spot was activated in response to a tone of 0.25–2 kHz (Fig. 2b1), with the activity centroid shifting smoothly with frequency (Fig. 2b2). Because the trajectory of the centroid and the steepness of the frequency gradient (0.43 mm/octave) agreed with that of A1 in marmoset L (compare Fig. 2b2 with Fig. 2a4), the frequency gradient probably corresponds to A1, although the low-frequency end may also contain R (Fig. 2b3). An additional spot was activated in response to 4 kHz tones at a rostroventral location (Fig. 2b1, rightmost graph, double arrowhead), which is probably a part of RT (Fig. 2b3). The frequency gradient of RT (Fig. 2a5) suggests that responses to lower frequencies may have been outside of the imaging field in marmoset L.

The above results on the core areas and their frequency gradients are consistent with those of previous studies on New World monkeys, including marmosets (Imig et al. 1977; Luethke et al. 1989; Aitkin et al. 1986; Morel and Kaas 1992; Bendor and Wang 2005, 2008; Feng and Wang 2017) and macaques (Merzenich and Brugge 1973; Morel et al. 1993; Kosaki et al. 1997; Petkov et al. 2006, 2008).

Areas in the caudal and lateral belt

To identify the caudal end of A1, we examined the early responses to higher frequency tones ($\geq 4 \text{ kHz}$). Such responses were successfully obtained in marmoset N, in which our imaging field covered a wider dorsocaudal area (Fig. 3a1). Two activity spots adjacent to the LS were observed in response to higher frequency tones (Fig. 3a1; arrowhead and double arrowhead). As the tone frequency increased, the two spots shifted toward each other along the direction of the LS (Fig. 3a1); the rostroventral spot shifted much further than the dorsocaudal one. The two spots appeared to merge into one at 32 kHz (Fig. 3a1). Superimposing the activity centroids together revealed two gradients and one reversal (Fig. 3a2), leading to the identification of two areas (Fig. 3a3). The direction of the frequency gradients and the location of these areas suggest that the frequency reversal represents the border between A1 and the CL area (Fig. 3a3; compare to Fig. 2a2). The steepness of the frequency gradient was 1.5 mm/octave in A1, which was greater than the steepness of the low-frequency region in A1 (Table 1). In CL, the steepness of the frequency gradient was 0.39 mm/octave (Table 1).

In marmoset M, responses to tones of 4–32 kHz were also observed in the dorsocaudal region (Fig. 3b1), forming one smooth gradient (Fig. 3b2). Both the direction and the steepness of the gradient (1.1 mm/octave, Table 1) were comparable to those of A1 in marmoset N. Thus, the frequency gradient is likely to represent the

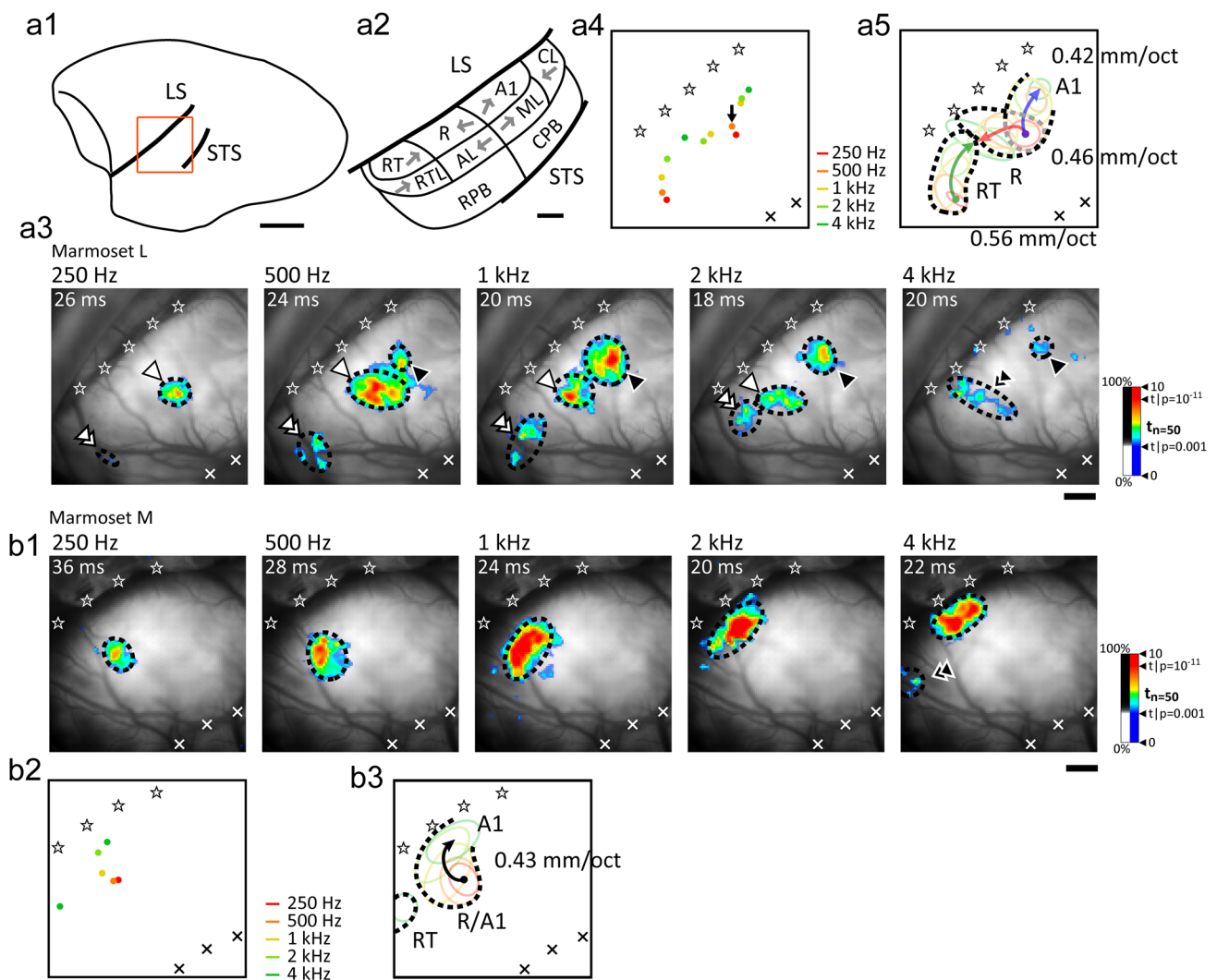


Fig. 2 Core areas identified with early responses to lower frequency tones. **a1** A schematic drawing of the lateral view of the marmoset brain, with the imaging field (red square) superimposed. Rostral is to the left and dorsal is upward. **a2** Current model of the marmoset auditory cortex on the STG, adapted from de la Mothe et al. 2006a. Gray arrows indicate the frequency gradient in the high-frequency direction. Frequency gradients in core areas are based on marmoset data (Bendor and Wang 2005), and those in the belt are based on macaque data (Rauschecker et al. 1995; Petkov et al. 2006, 2008). **a3** Significant cortical responses to tones of 250 Hz, 500 Hz, 1 kHz, 2 kHz, and 4 kHz, in marmoset L at 4 ms after the onset of the earliest response. t scores satisfying $p < 0.001$ were encoded in color and shown as spatio-temporal patterns. The arrowhead, white arrowhead, and double white arrowhead point to three activity spots. The double arrowhead points to the merged result of the two rostral spots. Each ellipse approximates the outer contour of an activity spot. **a4** Super-

imposition of centroids of activities shown in (**a3**), with tone frequency encoded in color, as shown in the legend. The arrow points to the centroid of the two activity spots indicated by the white and black arrowheads in the second to the left graph in (**a3**). Tonotopic gradients and reversals defined three areas. **a5** The response contours in (**a3**) belonging to each of the three frequency gradients in (**a4**) were merged as one area, whose contour was then approximated with either a closed or open dashed curve. The area enclosed by the gray dashed curves corresponds to the dorsocaudal activity spot in response to 250 Hz (**a3**, leftmost graph), and thus may belong to both R and A1. The arrow in each area approximates the trajectory of activity centroid from low to high frequency in that area. **b1–3** Significant cortical responses to tones of different frequencies in marmoset M. All panels are illustrated using the same methods as in (**a**). Scale bar is 1 mm for **a2**, **a3**, and **b1**, 5 mm for **a1**

high-frequency region of A1 in marmoset M (Fig. 3b3). The location of the response to 32 kHz tones in marmoset M probably represents the caudal end of A1, although the CL area was not observed here due to the limited spatial extent of the craniotomy.

Two additional belt areas were confirmed when we examined the responses to tones of 0.25–4 kHz in marmosets L and M, with a less stringent significance criterion ($p < 0.01$) than that used in Figs. 2 and 3 ($p < 0.001$). A less stringent criterion makes responses with a smaller amplitude available

Table 1 Steepness of tonotopic gradient

Area	Frequency range for measurement (kHz)	Marmoset L (mm/octave)	Marmoset M (mm/octave)	Marmoset N (mm/octave)
RT	0.25–4	0.56		
R ^a	0.25–4	0.46		
A1 (low) ^b	0.25–4	0.42	0.43	
A1 (high)	4–32		1.1	1.5
CL	4–24			0.39
ML	250–4	0.25	0.29	
AL	500–4	0.17	0.16	
MAL	8–24		0.51	0.60

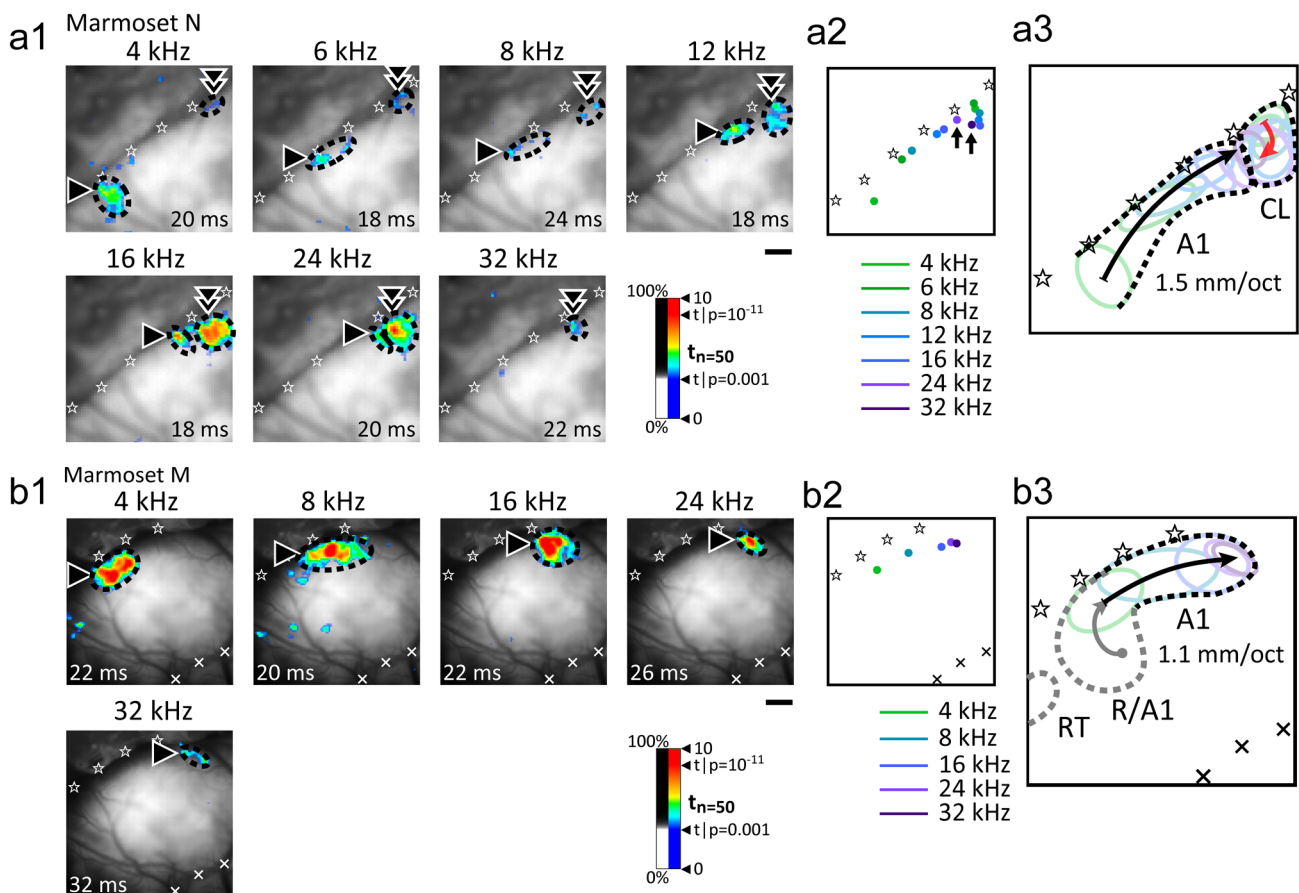
^aIt may include A1 in the low-frequency part^bIt may include R in the low-frequency part

Fig. 3 Tonotopic early responses to higher frequencies. **a1** Significant responses to tones of 4, 6, 8, 12, 16, 24, and 32 kHz in marmoset N, at 4 ms after the onset of the earliest response. The arrowhead and double arrowhead point to the two activity spots, which shifted toward each other as the frequency increased. t scores satisfying $p < 0.001$ were encoded in color and shown as spatio-temporal patterns. Each ellipse approximates a response contour. Scale bar is 1 mm. **a2** Superimposition of the centroids of responses shown in **(a1)**, with tone frequency encoded in color, as shown in the legend. The two arrows point to the centroid of activities evoked by the 24 kHz tone and the 32 kHz tone, respectively. **a3** The response contours in **(a1)** belonging to the two frequency gradients in **(a2)** were merged into one area, and its contour was then approximated

with either a closed or open dashed curve. The arrow in each area approximates the trajectory of the activity centroid from low to high frequency in that area. The open dashed curve outlines the higher frequency region of A1 and the closed dashed curve is the belt area CL. **b1** Significant responses to tones of 4, 8, 16, 24, and 32 kHz in marmoset M, at 4 ms after the onset of the earliest response. **b2** Superimposition of the centroids of activities shown in **(b1)**. **b3** The response contours in **(b1)** were merged into one area, and its contour was then approximated with an open dashed curve. The arrow approximates the trajectory of the activity centroid from low to high frequency. The open dashed curve outlines the higher frequency region of **(a1)**. The gray dashed contours and arrows are the results shown in Fig. 2**b3** obtained from the same animal. Scale bar is 1 mm

for analysis. The significant responses of marmosets L and M to tones of 0.25–4 kHz are shown in Fig. 4a, b, respectively, where the criterion for significance was $p < 0.01$. In marmoset L, in addition to activity in A1, R, and RT, activity spots caudal to R/A1 were observed (Fig. 4a1, arrowhead

and white arrowhead). Activity appeared in one spot in response to the 0.25 kHz tone, but in two spots in response to tones of 1 kHz and above (Fig. 4a1). The two spots diverged from each other as the tone frequency increased, roughly along the dorsoventral direction (Fig. 4a1). Superimposition

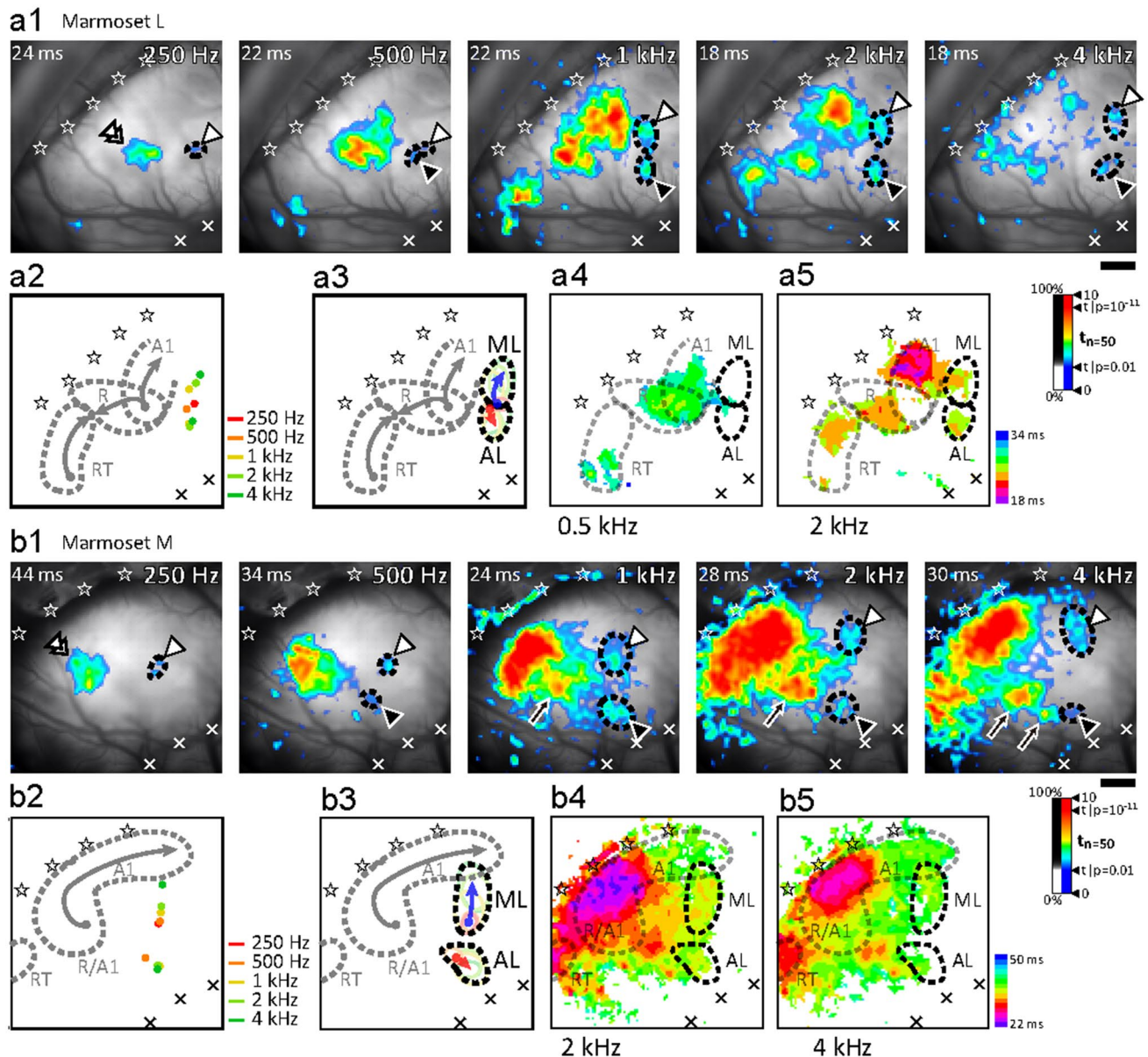


Fig. 4 Two mirror-imaged tonotopic areas in the lateral belt. **a1** Significant responses to tones of 0.25, 0.5, 1, 2, and 4 kHz in marmoset L, at 2 ms after the onset of the earliest response. The arrowhead and the white arrowhead point to two tonotopic areas, which diverged from each other as the frequency increased. t scores satisfying $p < 0.01$ were encoded in color and shown as spatio-temporal patterns. Each ellipse approximates the contour of a response area. **a2** Superimposition of the centroids of responses marked by the arrowheads and the white arrowheads shown in (**a1**), with frequency encoded in color. The gray dashed contours and arrows are the results shown in Fig. 2a5, obtained from the same animal. **a3** The response

contours in (**a1**) belonging to each of the two frequency gradients in (**a2**) were merged into one area, and its contour was then approximated with a closed dashed curve. The color arrow in each area approximates the trajectory of activity centroid from low to high frequency. The two areas were identified as ML and AL, as indicated. **a4** Spatial distribution of latency of the response to 0.5 kHz tones. The response latency in each pixel was encoded in color with the color scale shown to the right of (**a5**). **a5** The distribution of latency of the response to 2 kHz tones. **b1–5** Significant responses and response latencies obtained from marmoset M. All panels are illustrated in the same manner as (**a1–5**). Scale bar in **a** and **b** = 1 mm

of the centroids of activities evoked by all tones revealed two gradients and one reversal (Fig. 4a2, 3), suggesting the existence of two areas. Based on previous reports on the lateral belt (Kaas and Hackett 2000; Fig. 2a2), the two areas probably correspond to ML and AL (Fig. 4a3) based on the direction of their frequency gradients, their locations relative to A1 and R, and the location of the border between ML and AL relative to the location of the border between A1 and R. The values for the steepness of the frequency gradients in these areas were small: 0.25 mm/octave in ML and 0.17 mm/octave in AL (Table 1). To further test the identification of ML and AL, we examined maps of response latencies. We calculated the latency in each pixel that showed a significant response ($p < 0.01$; see Methods for definition), encoded the latency in color, illustrated the latencies as a spatial map, and superimposed the areas identified in Fig. 4a3. Figure 4a4 shows the results for marmoset L in response to tones of 0.5 kHz; pixels having a response latency were observed at the border between ML and AL. The latency map of the response evoked by 2 kHz tones exhibited an isolated trough at a ventral location of AL, and another trough in the middle region of ML (Fig. 4a5). These results support our identification of ML and AL and the frequency gradients in these areas. Similar observations on the locations, the direction of frequency gradient, and the steepness of frequency gradient were obtained for ML and AL in marmoset M (Fig. 4b1–5; Table 1). In the latency maps of responses to 2 kHz tones and 4 kHz tones (Fig. 4b4–5), both ML and AL appeared as latency troughs, largely separated from other areas by latency ridges; further, the direction of shift of the latency troughs within ML and AL with increasing tone frequency (Fig. 4b4–5) agreed with the direction of shift of activity spots (Fig. 4b1).

Additional areas of the lateral belt

Figure 4b1 also shows areas of strong activity ($t > 5$) between the core area R and the belt area AL, when the stimulus tone frequency was 1 kHz and above (Fig. 4b1, arrows). These areas were isolated from the core areas and the ML and AL areas. To further examine these areas, we set a more stringent significance criterion ($p < 0.001$) to focus on stronger cortical activity after the onset of responses in the core areas, at times later than those shown in Figs. 2 and 3. The results from marmoset M are shown in Fig. 5a. Two isolated spots of activity were identified in response to a 2 kHz tone (Fig. 5a1, leftmost graph, arrowhead and white arrowhead). As the frequency increased, the rostradorsal spot shifted in a ventrocaudal direction, while the caudoventral spot remained largely at the same location (Fig. 5a1). This is also clear from the trajectory of the activity centroids (Fig. 5a2). These observations suggest the existence of a larger tonotopic area and a smaller non-tonotopic area between R and AL

(Fig. 5a3). We refer to the non-tonotopic area as NT, and the tonotopic area as the medial anterolateral area (MAL). The steepness of frequency gradient was 0.51 mm/octave in MAL (Table 1). In the latency maps of responses to 4 kHz tones (Fig. 5a4), both MAL and NT appeared as a latency trough. The latency trough in MAL shifted towards the ventral direction when the tone frequency was increased to 24 kHz (Fig. 5a5), in agreement with the shift of activity spots shown in Fig. 5a1. No clear shift was observed for the latency trough in NT (Fig. 5a5). Similar observations were obtained in marmoset N (Fig. 5b1–5). The steepness of the frequency gradient in the MAL area of marmoset N was 0.60 mm/octave (Table 1).

Response latencies in different areas

Previous studies have demonstrated differences in response latencies among areas in the core region as well as between core and belt areas (Recanzone et al. 2000; Bendor and Wang 2008; Camalier et al. 2012). To provide further support for the current identification of auditory areas, we compared the response latencies in different areas to the same stimulus tone. We calculated the latency in each pixel that showed a significant response ($p < 0.01$), and identified the minimum response latencies to each tone, in each field, and in each monkey (Table 2). We further compared all pairs of fields with sufficient data for statistical analysis, and the results are presented in Table 3. Within the core region, A1 had shorter latencies than RT ($p < 0.01$, Table 3). Comparison of the belt areas with the core areas revealed that MAL and NT, the two areas identified here, had longer latencies than both A1 and RT (MAL/NT vs A1, $p < 0.01$; MAL/NT vs RT, $p < 0.05$; Table 3), and both CL and ML had longer latencies than A1 (CL vs A1, $p < 0.05$; ML vs A1, $p < 0.01$; Table 3). Among the belt areas, MAL/NT had longer latencies than CL ($p < 0.01$), but shorter latencies than ML ($p < 0.01$; Table 3).

Cortical areas activated by broadband noise

Despite the numerous auditory areas identified above, some areas of the imaged cortical field were not activated by pure tones, especially regions close to the STS. Previous studies have found that the auditory cortex extends to the upper bank of the STS (de la Mothe et al. 2006b). To determine the cortical region of the STG that responds to acoustic stimulation, we applied a broadband noise to marmoset N. We used broadband noise instead of pure tones because previous studies have shown that tones are less effective in driving neurons than spectrally broader stimuli in the belt and parabelt regions (Rauschecker et al. 1995; Petkov et al. 2006). The results are shown in Fig. 6. Because of the transient nature of optical signals, the differentiated optical signal had

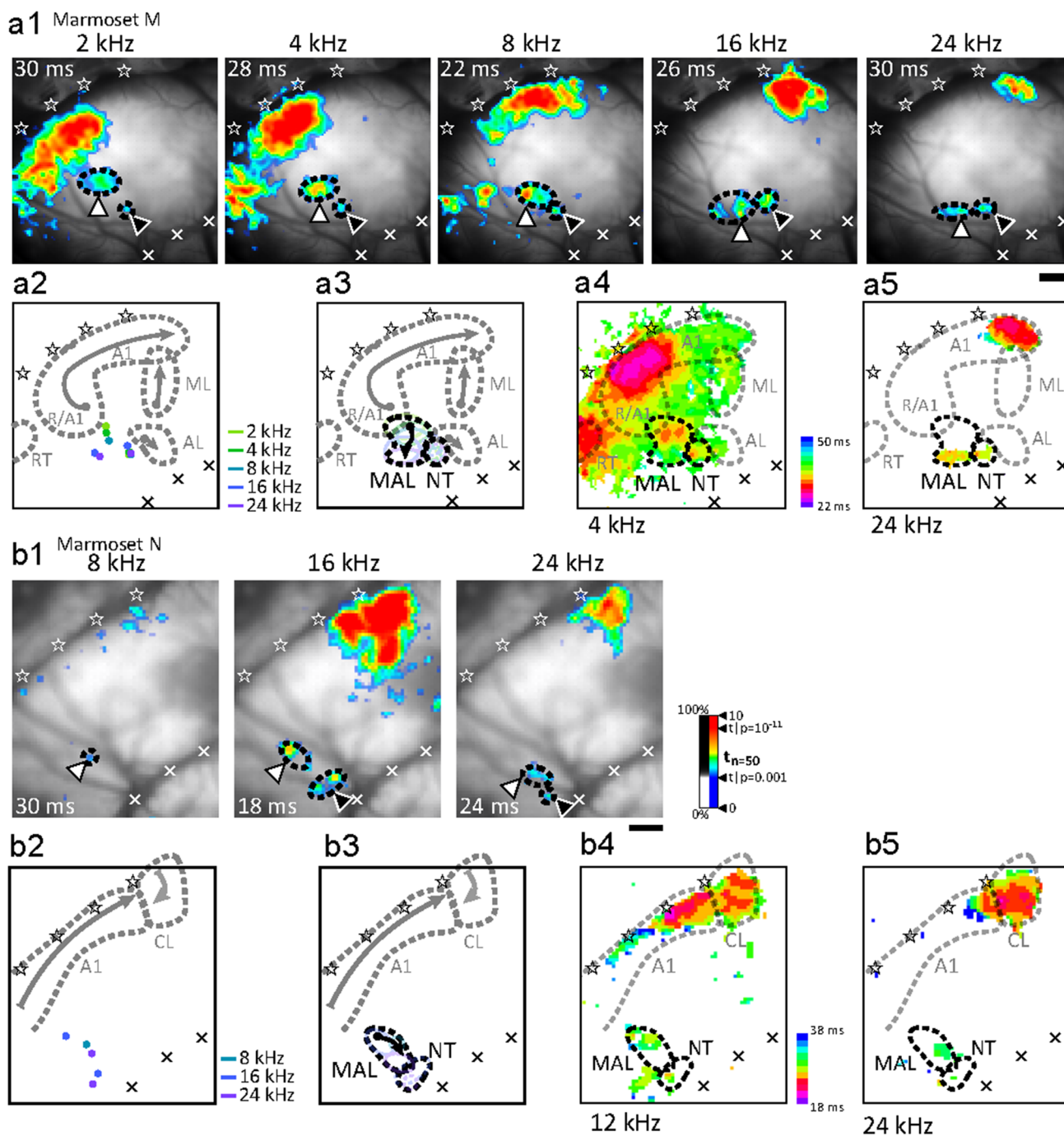


Fig. 5 Two novel areas between R and AL. **a1** Significant responses to tones of 2, 4, 8, 16, and 24 kHz in marmoset M, at 2–4 ms after the onset of the earliest response. The arrowhead and the white arrowhead point to the two activity spots inferior to the strong activity in **(a1)**. Each ellipse approximates a response contour. *t* scores satisfying $p < 0.001$ were encoded in color and shown as spatio-temporal patterns. While the spot indicated by the white arrowhead showed tonotopy, the one indicated by the arrowhead did not, suggesting the existence of a tonotopic area and a non-tonotopic area. **a2** Superimposition of the centroids of responses marked by the arrowheads and white arrowheads shown in **(a1)**, with frequencies encoded in color as depicted in the legend. The gray dashed contours and gray arrows

are the results shown in Fig. 4b3, obtained from the same animal. **a3** The response contours belonging to each of the two areas in **(a1)** were merged into one area, and its contour was then approximated with a closed dashed curve. The two areas were named MAL and NT, as indicated. Note their locations between R and AL. The arrow in MAL approximates the trajectory of activity centroid from low to high frequency. **a4** Spatial distribution of latency of the response to 4 kHz tones. The response latency in each pixel was encoded in color with the color scale to the right. **a5** The distribution of latency of the response to 24 kHz tones. **b1–5** Significant responses and response latencies obtained from marmoset N. All panels are illustrated in the same manner as **(a1–5)**. Scale bar in **a** and **b** = 1 mm

Table 2 Minimum response latency in identified auditory areas (in ms)

Frequency	RT	R	A1	CL	ML	AL	MAL	NT
Marmoset L								
250 Hz		26 ^a	26 ^a		30 ^c	30 ^c		
500 Hz	26	26 ^a	26 ^a		28	30		
1 kHz	24	24	24		24	24		
2 kHz	24	22	18		22	20		
4 kHz	18 ^b	18 ^b	16					
Marmoset M								
250 Hz		30 ^a	30 ^a		46			
500 Hz		30 ^a	30 ^a		40	42	34	38
1 kHz			24 ^d		24	26	24	24
2 kHz	26		22		34	34	30	32
4 kHz	28		24		36	34	32	32
8 kHz	32	30	26				32	32
16 kHz			24		32	34	26	28
24 kHz							32	32
Marmoset N								
4 kHz			22	22				
6 kHz			20	20				
8 kHz			22	26			32	
12 kHz			20	22			26	26
16 kHz			18	22			24	24
24 kHz			16	20			20	24
32 kHz				24				

^aLatency measured in the merged area of R and A1

^bLatency measured in the merged area of RT and R

^cLatency measured in the merged area of ML and AL

^dLatency might have been affected by latency of R

Table 3 Comparison of latency difference in identified auditory fields

	A1	MAL or NT ^a
A1		-5.8 ± 0.6 ms ($n=19$) ^c
RT	$+3.4 \pm 0.7$ ms ($n=7$) ^c	-3.0 ± 1.0 ms ($n=6$) ^b
CL	$+2.3 \pm 0.8$ ms ($n=6$) ^b	-3.1 ± 0.7 ms ($n=7$) ^c
ML	$+6.8 \pm 1.8$ ms ($n=10$) ^c	$+3.2 \pm 0.7$ ms ($n=10$) ^c

The mean \pm SEM of the latency of a column field subtracted from the latency of a row field are shown

^aLatency differences against MAL and against NT were pooled together

^bDifference was statistically significant ($p < 0.05$)

^cDifference was statistically significant ($p < 0.01$)

an initial sharp positive component followed by a slower negative component (see Fig. 1c). The upper panel of Fig. 6a shows the statistically significant positive changes in the differentiated optical signals evoked by broadband noise, and the lower panel shows the negative changes. The proportion of the total area of pixels showing positive changes to the STG area imaged here (the area between the LS and STS)

reached a maximum of 44.4% at 26 ms after sound stimulation (Fig. 6b, left graph), and the proportion of the total area of pixels showing negative changes to the STG area imaged here reached a maximum of 24.4% at 62 ms (Fig. 6b, right graph). It is clear from the figures that most of the STG was activated by the broadband stimulus, including cortical areas immediately adjacent to the STS.

To illustrate the spatial extent of the cortical response to broadband noise, the approximate contours of the positive and negative signal areas in Fig. 6b were estimated by eye and are shown in Fig. 6c as dashed lines (red for positive changes and blue for negative changes); these were further imposed on the results in Fig. 5b3, which were obtained from the same animal. To illustrate the spatial extent of the acoustically driven cortical area, the contours of broadband noise-driven activities (dashed lines in Fig. 6c) were merged into one contour covering the largest area. This contour is represented by the dashed line in Fig. 6d. To illustrate all of the auditory areas in the core and belt regions identified in the current study, we further incorporated the results of marmoset M, shown in Fig. 5a3, into Fig. 6d, using the position and orientation of the A1, MAL, and NT areas. Figure 6d

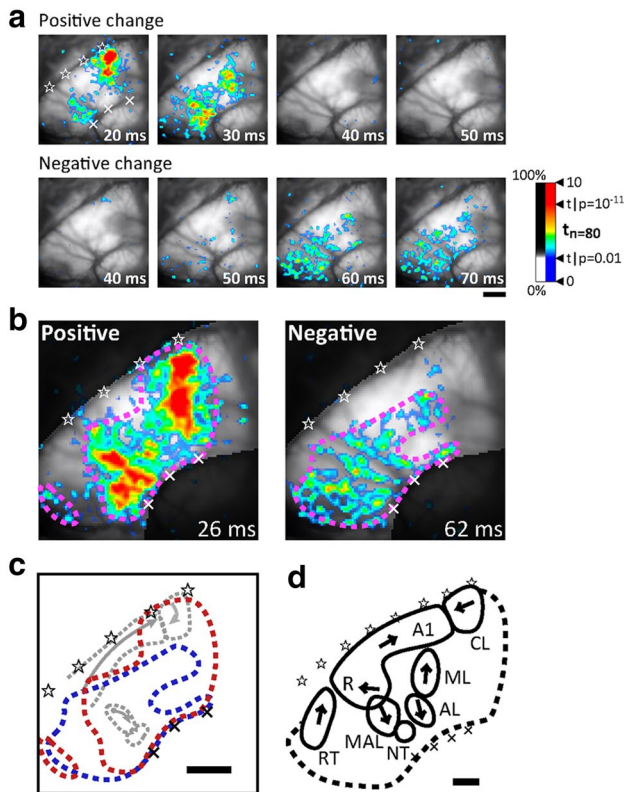


Fig. 6 Cortical areas activated by a broadband noise in marmoset N. **a** Areas showing significant positive (upper panel) or negative (lower panel) changes in $\Delta(\Delta F/F_0)$ in response to the noise. t scores satisfying $p < 0.01$ were encoded in color and shown as spatio-temporal patterns. Scale bar is 1 mm. **b** Maximum spatial extent of significant positive change (left) and negative change (right). The pink closed dashed curves approximate the contours of significant responses. Areas other than the STG were masked for calculation of the STG area and the ratio of the significant response area to the STG area. **c** Superimposition of the response contours in (b) (dashed lines; red for positive change, blue for negative change), and superimposition of the results shown in Fig. 5b3 in gray, obtained from the same animal; area names (A1, CL, MAL, and NT) are omitted for clarity. Scale bar is 2 mm. **d** The contours in (c) were merged into one area, and its contour was then approximated with a dashed curve. Results from all three monkeys were further incorporated using the position and orientation of A1, MAL, and NT. Each arrow indicates the tonotopic gradient of the area from low to high frequency. Scale bar is 1 mm

thus summarizes all of the auditory areas identified here, and the spatial extent of the noise-driven cortical area.

Discussion

Using an optical imaging technique with high spatial and temporal resolution, we confirmed three areas in the auditory core region in marmosets: A1, R, and RT. We also confirmed three known belt areas, CL, ML, and AL, with frequency gradients and relative positions to A1/R that were in agreement with those of previous studies. Further, we found two

novel areas between the R and AL areas: MAL, which was caudoventral to R and had a frequency gradient in the ventrocaudal direction, and NT, which was positioned between MAL and AL and appeared to have no tonotopy. The MAL and NT areas responded to tones with frequencies up to at least 24 kHz.

Comparison with previous studies

Aitkin et al. (1986) conducted the first electrophysiological study on the marmoset auditory cortex, and found only one area, A1, with a clear frequency gradient in the range of 0.6–30 kHz. Importantly, the frequency axis is not parallel to the LS, but the neurons with low best frequencies are found in areas of the STG away from the LS, whereas those with high best frequencies are located in areas of the STG close to the LS (Aitkin et al. 1986; reviewed in Kaas and Hackett 2000). These features have been repeatedly confirmed in subsequent electrophysiological mapping studies (Bendor and Wang 2005, 2008; Feng and Wang 2017). Our results are also consistent with these findings (see the trajectory of activity centroids in marmoset L in Fig. 2). The frequency axes of the R and RT areas identified here also agree well with previous studies (Bendor and Wang 2005, 2008; Feng and Wang 2017). Thus, both the area composition and frequency gradients of the core region in marmosets identified here are in good agreement with the current model of the primate auditory cortex (Kaas and Hackett 2000).

The location and frequency gradient of the CL belt area identified here are in agreement with a previous finding in macaques (Rauschecker et al. 1995). Our observation that CL had a short latency close to that of A1 is also consistent with a previous report (Camalier et al. 2012). From our observation that CL had a shorter latency than MAL, and MAL had a shorter latency than ML (there are no data showing co-activation in CL and ML), it is highly likely that CL has a much shorter latency than ML. This is also consistent with the results of Camalier et al. (2012). Recanzone et al. (2000), however, reported long latencies for area L, which is located lateral to CM and thus may correspond to CL. The reason for the long latency in area L is unclear, although this area might contain ML, which has long latencies (Camalier et al. 2012). While no physiological evidence has been provided for the lateral ML and AL belt areas in marmosets, our results on the locations and frequency gradients of these areas are in agreement with those reported in macaques (Rauschecker et al. 1995; Petkov et al. 2006, 2008). The finding that the border between the ML and AL areas was at a similar dorsoventral level to the border between A1 and R (see Fig. 4a) is also in agreement with previous studies (reviewed in Kaas and Hackett 2000). The observation of longer latencies in ML and AL than in A1 is also consistent with a previous report (Camalier et al. 2012).

It has previously been shown that belt neurons respond poorly and inconsistently to pure tones, in contrast to the robust response of core neurons (Miller et al. 1972; Brugge and Merzenich 1973; Rauschecker et al. 1995). Consistent with this, we found that the lateral ML and AL belt areas were only revealed when the significance criterion was lower than that for the core areas (see Fig. 4). Another feature of the belt areas is that they are driven by core areas, despite the existence of direct thalamic inputs (Rauschecker et al. 1997). The longer response latencies we found for the ML, AL, MAL, and NT lateral belt areas, compared with A1, are consistent with this proposal. The shorter response latency of A1 compared with the belt areas has also been reported in macaques (Recanzone et al. 2000; Camalier et al. 2012). However, other interpretations of the latency difference between core and belt areas, such as differences in the conduction velocity of thalamocortical axons projecting to the core and belt, cannot be excluded at this time.

The novel finding of the current study is the identification of the MAL and NT areas in the lateral belt. The longer latencies of these areas compared with those of A1 are consistent with the notion that they are in the belt region. We considered the MAL and NT areas as part of the belt region, rather than the parabelt region, because they were located medial to the belt area AL (see Fig. 5). Because MAL and NT appeared as isolated activity spots upon tone stimulation (see Fig. 5), their identification is rather unequivocal. Further, MAL and NT also appeared as short latency spots in the latency map (see Fig. 5). The successful identification of MAL and NT here is attributable to the high spatial and temporal resolution of the imaging technique. Rauschecker and colleagues (Rauschecker et al. 1995; Tian et al. 2001; Rauschecker and Tian 2004; Tian and Rauschecker 2004) successfully mapped three belt areas in macaques, but the mapping was only performed along a single row of electrode penetrations. Petkov et al. (2006, 2008) conducted fMRI studies in macaque with a voxel resolution of $1 \times 1 \times 2 \text{ mm}^3$. Because the diameters of MAL and NT in marmoset are around 1 mm, it would be difficult for an fMRI study to reveal these areas. The high sound pressure level (70–85 dB) used in fMRI studies may also obscure small auditory areas.

For the broadband noise activated cortical areas, we counted both the area showing positive changes and the area showing negative changes of the differentiated optical signals. Although these areas overlapped considerably, they did not agree with each other (see Fig. 6c). The reason for the difference is not clear at this time, but location-dependent excitation/inhibition strength and lateral inhibition may in part contribute to the difference. The positive and negative changes should reflect the rise phase and the decay phase of the optical signal, respectively. It has been shown in guinea pig auditory cortex that the rise phase is primarily a glutamatergic excitatory response, and the decay phase includes

a large GABAergic inhibitory component (Horikawa et al. 1996). The strength of excitation and inhibition may depend on cortical location, and thus the areas of positive changes may differ from the areas of negative changes in Fig. 6. Further, lateral inhibition is well demonstrated in the auditory cortex (Horikawa et al. 1996), which may also make the spatial extent of inhibition differ from that of excitation.

The broadband noise activated cortical area, counted as both the area showing positive changes and the area showing negative changes, occupied a large portion of the STG, and was larger than the sum of all of the areas activated by tones (see Fig. 6d). We speculate that there may be two major reasons for this difference. First, belt- and parabelt-areas are known to be less sensitive to narrowband stimuli (Miller et al. 1972; Brugge and Merzenich 1973; Rauschecker et al. 1995; Recanzone et al. 1999; Kajikawa et al. 2005; Petkov et al. 2006). This might explain, in part, why we saw smaller activation areas in the belt areas, and no activation in areas close to the STS (presumably the parabelt areas), in response to pure tones. Second, we used the early responses to tones in our analyses, to better reveal the tonotopy; however, it is well known that in optical imaging, activated areas in the sensory cortices increase in size over time (Grinvald et al. 1994; Song et al. 2006). Thus, considering only the early responses to tones may have made the response area to tones appear much smaller. Conversely, the response evoked by the broadband noise was considered when the activated area reached its maxima (see Fig. 6).

To the best of our knowledge, Bendor and Wang (2005) conducted the highest density mapping over a large extent of the marmoset STG reported in the literature. We superimposed our summary results (Fig. 6d) on their mapping and the result is shown in Fig. 7. In addition to the rough agreement between the locations and frequency gradients of the core areas, the partial mapping results of Bendor and Wang (2005) overlapping with the MAL and ML areas are also in agreement with the current observations. We must stress that although the borders determined here using tonotopy reversal are reliable, other borders were determined by the contours of the activated cortical areas and thus depend on the time window of analysis, and do not necessarily represent the boundary of an area. These latter borders are outlined by gray lines in Fig. 7.

Limitations of the current study

This was the first study to apply a voltage-sensitive dye imaging technique to study the parcellation of the marmoset auditory cortex. While the high spatial and temporal resolution helped to identify known and novel areas, there are certainly limitations. First, auditory areas occupy not only the STG, but also the lower bank of the LS and the upper bank of the STS (de la Monthe et al. 2006a, b); areas within the LS and STS

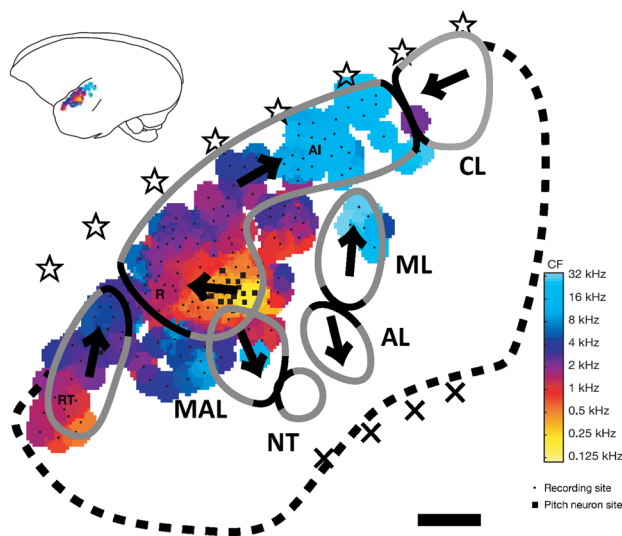


Fig. 7 Superimposition of the current results with a previous high-density mapping result (Bendor and Wang 2005). Arrows indicate frequency gradients. Solid curves illustrate borders identified by tonotopy reversal, and gray curves show only the contour of early activity observed here, not necessarily the border of a field. Note the rough agreement in the frequency gradient in the MAL area, in addition to the three core areas. Scale bar is 1 mm

were not visualized in this study. Further, we were unable to identify three areas of the STG in the current model of the marmoset auditory cortex: the two areas in the parabelt region and the lateral rostrotemporal area in the lateral belt region (see Fig. 2a2). The reason for this failure is unclear, but it might be attributable to the weak responses of these areas to pure tones. Second, because we used a limited number of stimulus types, i.e., pure tones and a broadband noise, cortical areas that were non-responsive to these sounds were not revealed. Third, because our results were based on differentiated optical signals, rapidly rising/falling responses could have been stressed, while sluggish responses might have been suppressed. Last, for unknown reasons, not all of the areas reported here were fully identified in each animal; for example, the A1 response of marmoset L was observed only for lower frequency tones. The lack of responses to higher frequency tones might have been due to functional damage to the cortex during surgery or incomplete staining of the cortex. Nevertheless, all conclusions drawn in the current study, with the exception of those for the R and CL area, were based on consistent data from at least two marmosets.

Conclusions

Our results suggest the existence of two novel areas in the belt region caudoventral to R: the tonotopic MAL area and the non-tonotopic NT area. The current model of the primate

auditory cortex proposes that all areas in the belt region exhibit tonotopy (Kaas and Hackett 2000; Rauschecker et al. 1995; Petkov et al. 2006, 2008). Our results add a new tonotopic area to the lateral belt. The NT area identified here is the first finding of a non-tonotopic area in the belt. If the observation that NT responded to a wide range of frequencies is also true at the cellular level, it is tempting to speculate that the NT area processes temporal information by integrating spectral information. Our findings call for a refinement of the current model of the area organization of the primate auditory cortex. The exact functional roles of MAL and NT in sound processing await explication in future studies, as do the roles of other belt areas.

Acknowledgements This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas “Adaptive Circuit Shift” (#15H01442) and “Dynamic Regulation of Brain Function by Scrap & Build System” (#17H05749) of the Ministry of Education, Culture, Sports, Science and Technology of Japan, and JSPS Grants (#25290006).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aitkin LM, Merzenich MM, Irvine DR, Clarey JC, Nelson JE (1986) Frequency representation in auditory cortex of the common marmoset (*Callithrix jacchus jacchus*). *J Comp Neurol* 252:175–185
- Bendor D, Wang X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature* 436:1161–1165
- Bendor D, Wang X (2008) Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J Neurophysiol* 100(2):888–906
- Bieser A, Müller-Preuss P (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 108:273–284
- Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. *J Neurophysiol* 36(6):1138–1158
- Camalier CR, D’Angelo WR, Sterbing-D’Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc Natl Acad Sci USA* 109(44):18168–18173
- de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2006a) Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *J Comp Neurol* 496(1):27–71
- de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2006b) Thalamic connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *J Comp Neurol* 496:72–96
- Feng L, Wang X (2017) Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc Natl Acad Sci USA* 114(5):E840–E848
- Grinvald A, Lieke EE, Frostig RD, Hildesheim R (1994) Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J Neurosci* 14:2545–2568

- Horikawa J, Hosokawa Y, Kubota M, Nasu M, Taniguchi I (1996) Optical imaging of spatiotemporal patterns of glutamatergic excitation and GABAergic inhibition in the guinea-pig auditory cortex in vivo. *J Physiol* 497(Pt 3):629–638
- Horikawa J, Hess A, Nasu M, Hosokawa Y, Scheich H, Taniguchi I (2001) Optical imaging of neural activity in multiple auditory cortical fields of guinea pigs. *Neuroreport* 12:3335–3339
- Imig TJ, Ruggero MA, Kitzes LM, Javel E, Brugge JF (1977) Organization of auditory cortex in the owl monkey (*Aotus trivirgatus*). *J Comp Neurol* 171:111–128
- Inagaki S, Katura T, Kawaguchi H, Song W-J (2003) Isolation of neural activity from respiratory and heartbeat noises for in vivo optical recordings using independent component analysis. *Neurosci Lett* 352:9–12
- Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci USA* 97:11793–11799
- Kajikawa Y, de la Mothe LA, Blumell S, Hackett TA (2005) A comparison of neuron response properties in areas A1 and CM of the marmoset monkey auditory cortex: tones and broad band noise. *J Neurophysiol* 93:22–34
- Kosaki H, Hashikawa T, He J, Jones EG (1997) Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J Comp Neurol* 386:304–316
- Lu T, Wang X (2004) Information content of auditory cortical responses to time-varying acoustic stimuli. *J Neurophysiol* 91:301–313
- Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4:1131–1138
- Luethke LE, Krubitzer LA, Kaas JH (1989) Connections of primary auditory cortex in the New World monkey, *Saguinus*. *J Comp Neurol* 285(4):487–513
- Maeda S, Inagaki S, Kawaguchi H, Song W-J (2001) Separation of signal and noise from in vivo optical recording in guinea pig using independent component analysis. *Neurosci Lett* 302:137–140
- Merzenich MM, Brugge JF (1973) Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res* 50:275–296
- Miller JM, Sutton D, Pflugst B, Ryan A, Beaton R, Gourevitch G (1972) Single cell activity in the auditory cortex of Rhesus monkeys: behavioral dependency. *Science* 177(4047):449–451
- Morel A, Kaas JH (1992) Subdivisions and connections of auditory cortex in owl monkeys. *J Comp Neurol* 318:27–63
- Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol* 335:437–459
- 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
- Nieto-Diego J, Malmierca MS (2016) Topographic distribution of stimulus-specific adaptation across auditory cortical fields in the anesthetized rat. *PLoS Biol* 14(3):e1002397
- Nishimura M, Song W-J (2014) Greenwood frequency-position relationship in the primary auditory cortex in guinea pigs. *NeuroImage* 89:181–191
- Nishimura M, Shirasawa H, Kaizo HH, Song H W-J (2007) New field with tonotopic organization in guinea pig auditory cortex. *J Neurophysiol* 97:927–932
- Paxinos G, Watson C, Petrides M, Rosa M, Tokuno H (2012) The marmoset brain in stereotaxic coordinates, 1st edn. Academic Press, New York
- Petkov CI, Kayser C, Augath M, Logothetis NK (2006) Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol* 4(7):e215
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK (2008) A voice region in the monkey brain. *Nat Neurosci* 11(3):367–374
- Philibert B, Beitel RE, Nagarajan SS, Bonham BH, Schreiner CE, Cheung SW (2005) Functional organization and hemispheric comparison of primary auditory cortex in the common marmoset (*Callithrix jacchus*). *J Comp Neurol* 487:391–406
- Rauschecker JP, Tian B (2004) Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 91:2578–2589
- Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114
- Rauschecker JP, Tian B, Pons T, Mishkin M (1997) Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol* 382:89–103
- Recanzone GH, Schreiner CE, Sutter ML, Beitel RE, Merzenich MM (1999) Functional organization of spectral receptive fields in the primary auditory cortex of the owl monkey. *J Comp Neurol* 415(4):460–481
- Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 83:2315–2331
- Song W-J, Kawaguchi H, Totoki S, Inoue Y, Katura T, Maeda S, Inagaki S, Shirasawa H, Nishimura M (2006) Cortical intrinsic circuits can support activity propagation through an isofrequency strip of the guinea pig primary auditory cortex. *Cereb Cortex* 16:718–729
- Sweet RA, Dorph-Petersen KA, Lewis DA (2005) Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol* 491:270–289
- Tian B, Rauschecker JP (2004) Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 92:2993–3013
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293
- Wang X (2013) The harmonic organization of auditory cortex. *Front Syst Neurosci*. <https://doi.org/10.3389/fnsys.2013.00114>
- Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86:2616–2620
- Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685–2706