



OPEN

## Cerebral cortical processing time is elongated in human brain evolution

Kosuke Itoh<sup>1</sup>✉, Naho Konoike<sup>2</sup>, Masafumi Nejime<sup>2,4</sup>, Haruhiko Iwaoki<sup>2</sup>, Hironaka Igarashi<sup>1</sup>, Satoshi Hirata<sup>3</sup> & Katsuki Nakamura<sup>2</sup>

An increase in number of neurons is presumed to underlie the enhancement of cognitive abilities in brain evolution. The evolution of human cognition is then expected to have accompanied a prolongation of net neural-processing time due to the accumulation of processing time of individual neurons over an expanded number of neurons. Here, we confirmed this prediction and quantified the amount of prolongation *in vivo*, using noninvasive measurements of brain responses to sounds in unanesthetized human and nonhuman primates. Latencies of the N1 component of auditory-evoked potentials recorded from the scalp were approximately 40, 50, 60, and 100 ms for the common marmoset, rhesus monkey, chimpanzee, and human, respectively. Importantly, the prominent increase in human N1 latency could not be explained by the physical lengthening of the auditory pathway, and therefore reflected an extended dwell time for auditory cortical processing. A longer time window for auditory cortical processing is advantageous for analyzing time-varying acoustic stimuli, such as those important for speech perception. A novel hypothesis concerning human brain evolution then emerges: the increase in cortical neuronal number widened the timescale of sensory cortical processing, the benefits of which outweighed the disadvantage of slow cognition and reaction.

“Bigger is better” is a simple and widely accepted principle in brain evolution. The uniquely developed cognitive abilities of humans, such as language and abstract thinking, are supported by the disproportionate increase in the human brain size, which occurred to an extent that surpasses the general allometric relationship between body size and brain size in mammals<sup>1,2</sup>. An important assumption behind this hypothesis is that a larger brain contains a greater number of neurons, particularly cortical neurons, which augments the capacity for complex information processing. Although recent studies have shown that brain size is not a reliable predictor of the number of cortical neurons when compared across clade<sup>3</sup>, cerebral neuronal number increases with brain size within a clade, and primates have the highest density of cortical neurons among other mammals of similar brain size<sup>4</sup>. Consequently, the evolutionarily scaled-up human brain contains a greater number of cerebral cortical neurons than the African elephant, which has a brain three times larger than that of an average human<sup>5</sup>.

However, there is a sometimes overlooked disadvantage to having an increased number of cortical neurons, namely, a prolonged neural-processing time. An increase in the total number of cortical neurons implies that the number of neurons that support a specific cortical function, e.g., detecting a sound event or identifying a familiar face, also increases. Then, the total neural-processing time for executing a brain function is expected to become longer unless (1) all evolutionarily acquired cortical neurons operate in parallel with previously existing neural circuits, or unless (2) the temporal integration time of individual neurons (i.e., the time required by a neuron to summate synaptic inputs before it elicits an action potential) is shortened with the increase in the total neuronal number. To the best of our knowledge, there is no evidence to support these assumptions.

Some mechanisms have been hypothesized to mitigate the expected lengthening of the neural-processing time in the human brain, namely, cortical folding<sup>1,6</sup>, high neuronal density<sup>4</sup>, and thickening of myelination<sup>7</sup> that would contribute to maintaining a short axonal conduction time. However, to the best of our knowledge, these hypotheses have not been tested *in vivo* because of the difficulty to measure the speed of neural processing in the cortical tissues of living humans and nonhuman primates.

<sup>1</sup>Center for Integrated Human Brain Science, Brain Research Institute, Niigata University, Niigata, Japan. <sup>2</sup>Cognitive Neuroscience Section, Primate Research Institute, Kyoto University, Inuyama, Aichi 951-8585, Japan. <sup>3</sup>Wildlife Research Center, Kyoto University, Kyoto, Japan. <sup>4</sup>Present address: Division of Biomedical Science, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. ✉email: itoh@bri.niigata-u.ac.jp

Species	Brain mass (g) <sup>34</sup>	Number of cortical neurons <sup>36</sup>	Estimated number of auditory cortical neurons <sup>a</sup>
Common marmoset	7.64	244,720,000	1,713,04
Rhesus monkey	88.94	1,710,000,000	11,970,000
Chimpanzee	376.13	4,900,821,995*	31,933,022
Human	1250.43	16,340,000,000	102,678,300

**Table 1.** Brain size and number of neurons in the four primate species investigated in this study. \*Predicted from brain mass by using equation in<sup>37</sup>. <sup>a</sup>See “Supplementary analysis of allometry” in the “Results” section for the estimation method.

Did the human brain trade off neural-processing speed for the benefit of elaborate computation? To address this problem, we used scalp electroencephalogram (EEG) to measure the speed at which neural circuits in the cerebral cortex process sensory information in four living primates: the human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), rhesus monkey (*Macaca mulatta*), and common marmoset (*Callithrix jacchus*). These species have vastly different brain sizes and hence considerably different number of neurons (Table 1). Scalp EEG has recently been applied to nonhuman primates as a novel and powerful tool in comparative neurophysiology, as it reveals species differences in neural processing that underpin the evolution of cognition and behavior<sup>8–14</sup>.

This study focused on audition, which represents one of the most time-critical sensory modalities in the brain: sub-second temporal processing of sounds is fundamental to virtually all forms of audition, including speech perception. Specifically, the peak latencies of the P1 and N1 components of the cortical auditory evoked potential (CAEP) were exploited as a neurophysiological measure of neural processing time in the auditory cortex. EEG signals reflect the sum of postsynaptic potentials of many neurons that are activated in synchrony. Therefore, the peak latency of P1 or N1 indicates the timepoint at which a specific group of neurons subserving a particular stage of auditory processing were activated together in a functionally connected network.

A major generator of the human P1, which peaks around 50 ms after sound onset, is located in the primary auditory area<sup>15</sup>. Considering that single neurons in the human auditory core fire as early as 10–20 ms in latency following auditory stimulus<sup>16,17</sup>, the delayed latency of P1 at approximately 50 ms provides evidence of the first stages of neuronal processing in the neocortical circuit. Such delay appears to be a common property of the mammalian auditory cortex, as epicranial CAEP in mice indicated auditory cortical activities between 18 and 25 ms and between 25 and 54 ms in latency<sup>18</sup>, although intracranial recordings revealed that neural activities in the same cortical region begin earlier at approximately 10 ms<sup>18,19</sup>.

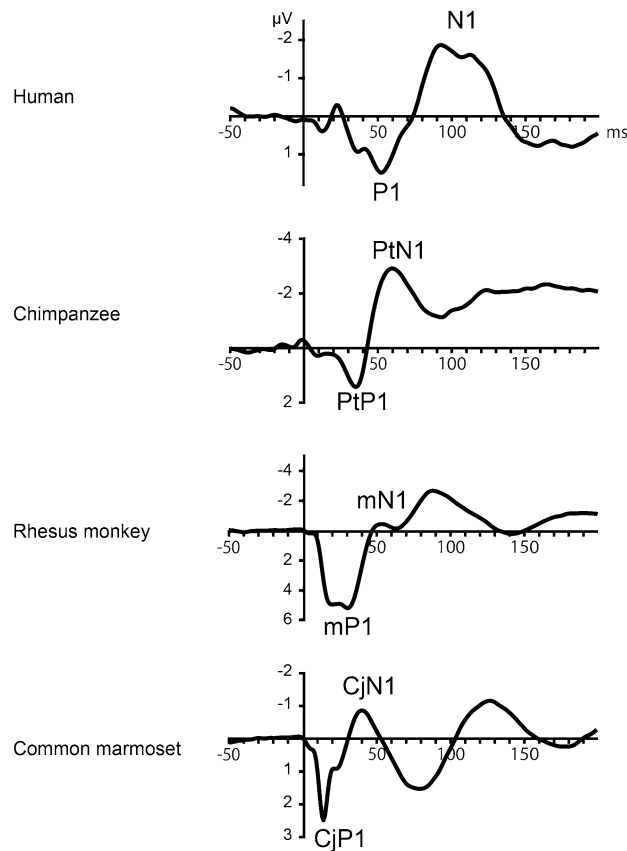
After approximately 50 ms following P1, further cortical processing culminates in N1, which is hypothesized to be generated in the primary auditory cortex as well as in other cortical areas<sup>20</sup>. N1 reflects a higher level of auditory processing than P1 and marks the beginning of a sound event as segmented in the auditory cortex<sup>20,21</sup>. Segmentation is based on the detection of abrupt changes in acoustic features, such as intensity and spectrum<sup>20,22,23</sup>, as well as higher-order changes in spectrotemporal patterns<sup>24</sup>. Such segmentations require analyzing sound inputs over a time window that is sufficiently long to embrace the acoustic change.

Importantly, despite the large differences in brain size, there are only a few milliseconds of difference in the response latencies of brainstem and midbrain neurons between humans, apes, and monkeys, as reflected by waves I (latency: 1–2 ms) through V (latency: 5–7 ms) of auditory brainstem responses<sup>25–28</sup>. Moreover, invasive intracranial recordings have revealed that the minimum response latencies of primary auditory cortical neurons are strikingly similar across primates: these are approximately 10–20 ms in humans<sup>16,17</sup>, macaques<sup>29,30</sup>, and marmosets<sup>31,32</sup> alike, although there are variations depending on experimental conditions. Accordingly, any differences in P1 and N1 latency between species would represent differences in auditory cortical processing time, rather than differences in neural conduction time from the cochlea to the cortex.

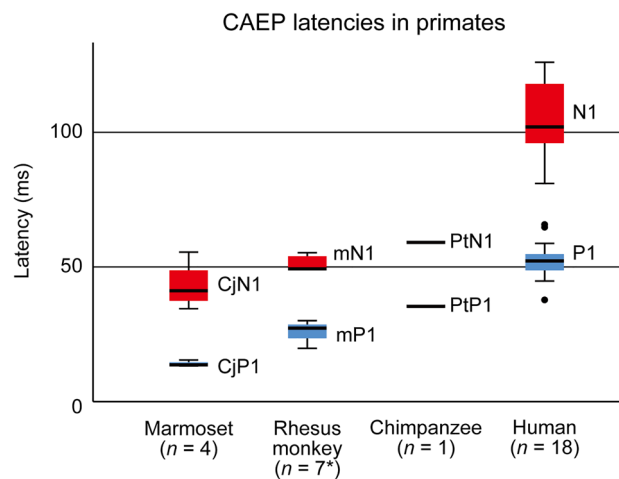
## Results

**Main.** We measured the latencies of P1 and N1 in humans ( $n = 18$ ), a chimpanzee ( $n = 1$ ), rhesus monkeys ( $n = 7$ ), and marmosets ( $n = 4$ ) to investigate this issue. The measurements of P1 and N1 latencies were obtained by reanalyzing the raw data from published and unpublished experiments by our group<sup>9,11–14</sup>, which were selected so that the stimuli and recording conditions were matched as closely as possible between the species. All recordings were made in a passive listening condition without a behavioral task in an awake state, to ensure that the CAEPs reflected exogenous and automatic response to sounds with minimal confounding effects of attention and other higher functions. The obtained CAEP waveforms in all species shared the same feature that the first prominent peak was positive in polarity, representing the P1 and, naturally, the negative peak that immediately followed it represented the N1 (Fig. 1). The homologs of the human P1 and N1 have been identified in the rhesus monkey (labeled mP1 and mN1, “m” standing for macaque)<sup>12,13</sup> and in the common marmoset<sup>11</sup> (CjP1 and CjN1, “Cj” standing for *Callithrix jacchus*). Precisely, the CjP1 is a composite of two positive waves, CjPa and CjPb, and the detailed analyses of these subcomponents are presented in the next section (Supplementary analysis of P1); these yielded consistent results. This is the first study to identify the homologs of P1 and N1 in the chimpanzee, which were labeled here as PtP1 and PtN1, respectively, where “Pt” stands for *Pan troglodytes*.

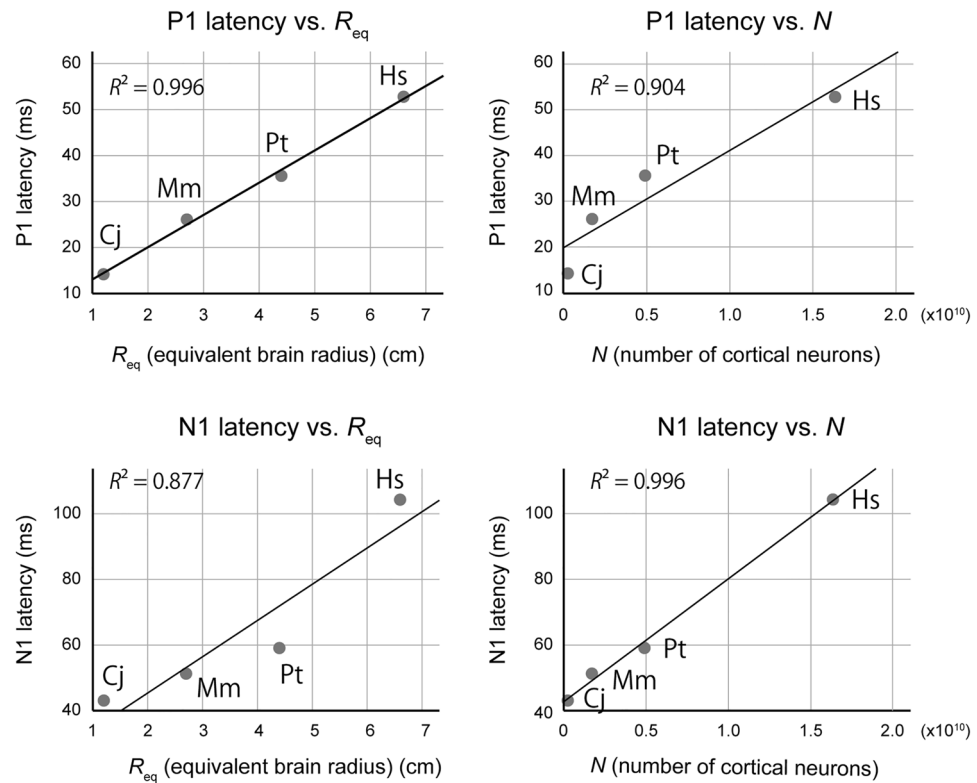
The peak latencies of P1 and N1 homologs increased as the brain enlarged in size (Figs. 1 and 2), and the differences between the species were statistically significant for P1 (Kruskal–Wallis  $H = 22.4$ ,  $P = 0.0001$ ) and N1 ( $H = 19.4$ ,  $P = 0.0003$ ). As the latencies of auditory brainstem responses have been relatively preserved in primate evolution<sup>25–28</sup>, the lengthening of the P1 and N1 latencies indicated an extension of the dwell time of cortical



**Figure 1.** Scalp-recorded CAEPs in the four primate species investigated in the study. The homologs of human P1/N1 were identified as PtP1/PtN1 in the chimpanzee, mP1/mN1 in the rhesus monkey, and CjP1/CjN1 in the common marmoset. The latencies of P1/N1 homologs are elongated as the brain size increases. The electrode was placed at the vertex. CAEP cortical auditory-evoked potential.



**Figure 2.** Box plot of the latencies of P1/N1 homologs in the four primate species. The thick horizontal line indicates the median, and the top and bottom box lines show the first and third quartiles. The whiskers show the maximum and minimum values, excluding outliers (circles) that were at least 1.5 box length from the median. Numbers in parentheses below the species names indicate the number of subjects (*n*). \*Two rhesus monkeys were excluded for the evaluation of mN1, as the peak was not clearly identified in the individual CAEP waveforms. The mean  $\pm$  standard deviation of the peak latencies was as follows:  $14 \pm 1$  (CjP),  $26 \pm 4$  (mP1),  $36 \pm 3$  (PtP1), and  $53 \pm 7$  (P1) for the P1 homologs;  $43 \pm 9$  (CjN1),  $51 \pm 3$  (mN1),  $59 \pm 3$  (PtN1), and  $104 \pm 15$  (N1) for the N1 homologs. The standard deviations were calculated over subjects in the common marmosets, rhesus monkeys, and humans, and over experimental conditions and sessions in the chimpanzee.



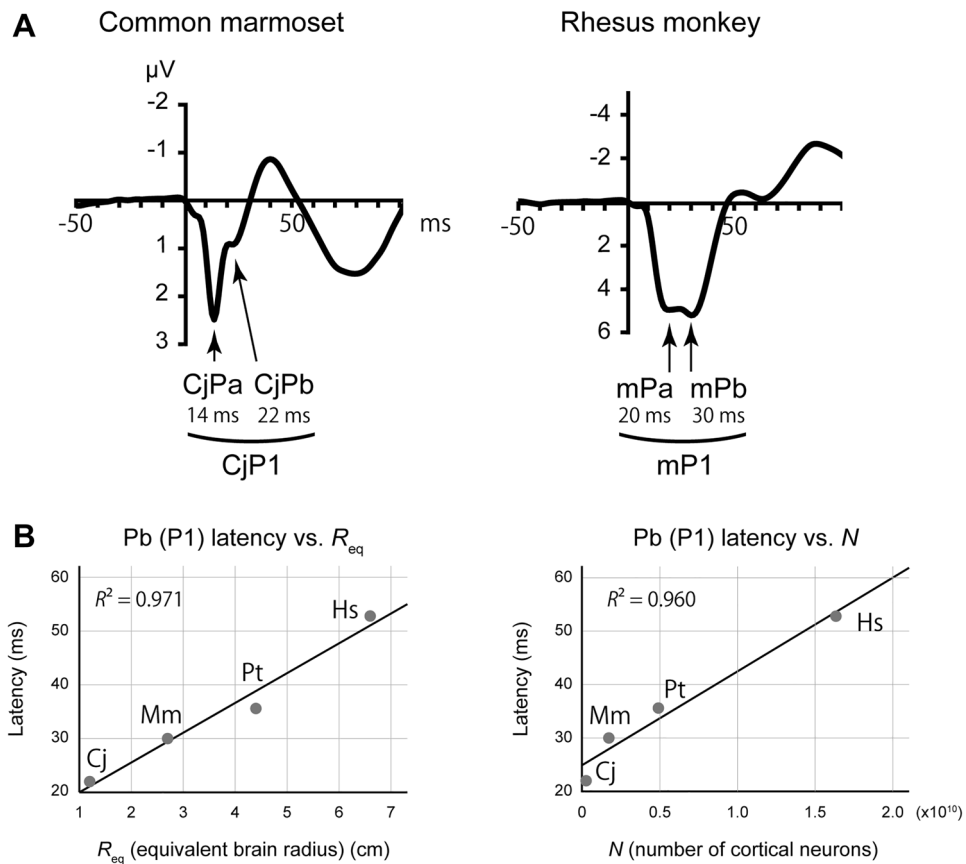
**Figure 3.** Allometry of P1/N1 latencies with respect to brain size (equivalent brain radius,  $R_{eq}$  cm) and the number of neurons in the brain ( $N$ ). The P1 latency showed a higher correlation with  $R_{eq}$  than with  $N$ , while the N1 latency showed a higher correlation with  $N$  than with  $R_{eq}$ . Cj common marmoset, Mm rhesus monkey, Pt chimpanzee, Hs human.

processing that cooccurred with brain enlargement. A possible anatomical basis for the selective delay of higher cortical processing, compared to lower subcortical processing, is a disproportionate increase in the number of auditory cortical neurons. In the primate auditory system, the number of neurons in the auditory cortex and subcortical structures (i.e., inferior colliculus and medial geniculate nucleus) both increase with brain size, but the rate of the increase is greater in the former<sup>33</sup>; hence, larger primate brains have increased ratios of cortical to subcortical neurons that are involved in processing sounds. Accordingly, the observed lengthening of the P1 and N1 latencies is consistent with the hypothesis that neural-processing time increases with the accumulation of the number of neurons leading to the production of the neural response.

To further investigate this hypothesis, we studied the allometric relationships between the P1/N1 latencies and two metrics of brain size, namely, the number of cortical neurons  $N$  as reported in the literature, and the equivalent brain radius  $R_{eq}$  (mm). For an alternative analysis using estimated numbers of auditory cortical neurons,  $N_{auditory}$ , instead of  $N$ , please refer to the section below: Supplementary analysis of allometry. The equivalent brain radius  $R_{eq}$  was obtained by solving the equation,  $V = \frac{3}{4}\pi R_{eq}^3$ , where the brain volume  $V$  was calculated by dividing the brain mass (g) as reported in the literature (Table 1)<sup>34</sup> by the specific gravity of the brain tissue, 1.036 g/cm<sup>335</sup>. As the auditory pathway length would scale with the cube root of brain volume, the CAEP latencies would increase linearly with  $R_{eq}$ , if these mostly reflected the conduction time of neural signals along axons in subcortical and cortical structures. Alternatively, latencies would scale linearly with  $N$  if these mostly reflected the sum of processing time of individual neurons, as each neuron requires time to summate excitatory and inhibitory synaptic inputs before it elicits an action potential, and this delay would accumulate with increasing  $N$ .

The results supported the latter hypothesis regarding N1, as  $N$  was a better predictor of its latency ( $R^2 = 0.996$ ,  $P = 0.002$ ) than  $R_{eq}$  ( $R^2 = 0.877$ ,  $P = 0.063$ ) (Fig. 3). However, the opposite results were obtained for P1, as the effect size was greater with  $R_{eq}$  ( $R^2 = 0.996$ ,  $P = 0.002$ ) than with  $N$  ( $R^2 = 0.904$ ,  $P = 0.049$ ). Thus, although the number of neurons and physical brain size both contributed to prolonging the CAEP latency, the former factor became more dominant at a higher level of cortical processing, as manifested by the prominent delay of human N1 latency (Fig. 2).

**Supplementary analysis of P1.** A careful inspection of the CAEP waveforms in Fig. 1 suggested that the CjP1 in the marmoset comprised two subcomponents that were not clearly resolved in time (Fig. 4A). In fact, our previous analysis of the high frequency components of marmoset auditory-evoked potential (AEP) revealed that CjP1 is a complex of two middle latency responses (MLRs), CjPa and CjPb, which putatively correspond to the human MLR components Pa and Pb, respectively<sup>11</sup>. MLRs are AEPs that occur after brainstem responses



**Figure 4.** Detailed waveforms of the putative P1 homologs. **(A)** The marmoset CjP1 and rhesus mP1 comprised two subcomponents that putatively corresponded to the human MLR components Pa and Pb. **(B)** Allometry of P1 latencies with respect to brain size (equivalent brain radius,  $R_{eq}$  cm) and the number of neurons in the brain ( $N$ ), in which the subcomponents of CjP1 and mP1 were distinguished. The results were similar to those of the original analysis (Fig. 1, main text). MLR, middle latency response. *Cj* common marmoset, *Mm* rhesus monkey, *Pt* chimpanzee, *Hs* human.

(< 10 ms in humans) and before the CAEP (> 50 ms in humans); they can be recorded using a higher EEG sampling rate and a higher passband frequency than those used for CAEPs. In the present study, the Pa and Pb were lumped together to form a single positive peak P1, owing to the low-pass filter that we used to focus on the late cortical components.

According to widely accepted nomenclature, the human MLR component Pb is equivalent to the human CAEP component P1. However, in the marmoset MLR, CjP1 (which is a lowpass-filtered CjPa–CjNa–CjPb complex) was dominated by CjPa, because it was much larger in amplitude than the CjPb (Fig. 4A), as has been previously described in detail<sup>11</sup>. Accordingly, there was a concern that our analysis of CjP1 latency reflected the latency of CjPa rather than that of CjPb, although it is the latter that represents the putative homolog of the human P1.

The situation was somewhat similar in macaques. Although, to our knowledge, the component structure of the rhesus monkey scalp MLR has not been clarified, the waveforms in the present study suggested that it also comprised two positive peaks: mPa and mPb. Assuming that mPa and mPb are homologs of CjPa and CjPb, respectively, then the second subcomponent mPb represents the candidate homolog of human P1. Nevertheless, as mPa and mPb were comparable in amplitude (Fig. 4A), our measurement of the mP1 latency in the above analysis probably did not accurately represent mPb latency.

By contrast, the PtP1 in the chimpanzee was similar in waveform to human P1, in that it only had one clear peak. Therefore, we could tentatively assume, although without concrete evidence, that PtP1 corresponded to human Pb (or P1) rather than to human Pa.

Table 2 summarizes the proposed correspondences of P1 and its subcomponents between species, as described above.

To take into account these considerations, the allometric analyses in the main analyses (Fig. 3) were repeated by using the latencies of CjPb and mPb, rather than those of CjP1 and mP1, to aim at a more accurate cross-species correspondence of components. The latencies of CjPb and mPb were obtained by visually inspecting the grand-average waveforms (Fig. 4A, marked with arrows).

	Human		Chimpanzee		Rhesus monkey		Common marmoset	
	MLR	CAEP	MLR	CAEP	MLR	CAEP	MLR	CAEP <sup>a</sup>
Putative homologs	Pa	Small, usually not seen	n.a	Was not seen	mPa	First peak of mP1	CjPa	Main peak of CjP1
Putative homologs	Pb	P1	n.a	PtP1	mPb	Second peak of mP1	CjPb	Small hump in CjP1

**Table 2.** Putative homologies of Pa/Pb (MLR) and their relationship with P1 (CAEP). *n.a.* information not available, to our knowledge. *CAEP* cortical auditory-evoked potential, *MLR* middle latency response. <sup>a</sup>The first positive CAEP component, referred to here as CjP1, was denoted as CjP1 in our previous study to distinguish it from the “true” CjP1, or CjPb.

As a result, the findings were similar to those of the original analyses of P1 homologs. That is, the effect size was greater for the equivalent brain radius  $R_{eq}$  ( $R^2 = 0.971$ ,  $P = 0.015$ ) than for the number of cortical neurons  $N$  ( $R^2 = 0.960$ ,  $P = 0.020$ ) (Fig. 4B), although the effect of  $N$  was slightly greater (and that of  $R_{eq}$  was slightly smaller) than that of the original analysis.

**Supplementary analysis of allometry.** In the main analysis above, elongations of P1 and N1 latencies were analyzed with respect to differences between species in the total number of cortical neurons  $N$ . However, since the P1 and N1 were auditory cortical responses, the number of auditory cortical neurons  $N_{auditory}$  may be a more appropriate independent variable.

Regarding common marmosets and rhesus monkeys, the size of auditory cortex scales proportionally with that of the entire cerebral cortex, maintaining a relative number of neurons of 0.7%<sup>33</sup>. Therefore, the  $N_{auditory}$  in these species could be estimated from the  $N$ s reported in the literature<sup>36</sup> as listed in Table 1, referred to as Estimation 1.

However, the proportion  $N_{auditory}/N$  was expected to be smaller than 0.7% in the chimpanzee and the human, which have an evolutionarily well-developed prefrontal cortex. Therefore, the number of auditory cortical neurons in these species were estimated by a different method. With  $P$  as the relative volume of the prefrontal cortex, the relative volume of “the rest of the cortex” would be  $1 - P$ . Assuming that a fixed proportion,  $p$ , of “the rest of the cortex” is occupied by the auditory area in all four species, the relative volume of the auditory cortex with respect to the whole cortex would be  $(1 - P) * p$ . Here,  $P$  has been measured to be 0.377, 0.354, and 0.306 for the human, chimpanzee, and macaque, respectively<sup>37</sup>. Although the measurement was not available for the common marmoset, we considered this to be identical to that of the macaque. A small error in this estimation would only have an insignificant effect on the results, as discussed below.

By using these values of  $P$ , the number of auditory cortical neuron  $N_{auditory}$  was estimated as  $0.694 * p * N_{marmoset}$  for marmosets,  $0.694 * p * N_{macaque}$  for macaques,  $0.646 * p * N_{chimpanzee}$  for chimpanzees, and  $0.623 * p * N_{human}$  for humans, where  $N_{species}$  was the total number of cortical neurons in that species: we refer to these as Estimation 2. Here, by equating Estimations 1 and 2 for the marmoset and the macaque,  $p$  was calculated as 0.010, and  $N_{auditory}$  in the human and the chimpanzee were obtained by inserting this value of  $p$  in Estimation 2, as shown in Table 1.

The analysis of allometry was performed with these estimates of  $N_{auditory}$ . As a result, a higher correlation was observed between the N1 latency and  $N_{auditory}$  ( $R^2 = 0.996$ ,  $P = 0.002$ ) than between  $N_{auditory}$  and P1 ( $R^2 = 0.918$ ,  $P = 0.042$ ). These results closely replicated the findings of the main analysis above. The robustness of this finding was anticipated because in the scope of a comparative study that spans several orders of magnitude, differences between  $N$  and  $N_{auditory}$  are minor. For the same reason, small variations in  $P$  or  $p$  would only have an insignificant effect.

## Discussion

An elongation of the auditory cortical processing time entails both merits and demerits. Slow processing is disadvantageous in situations where fast identifications of sounds are required, such as when detecting the sound of a predator. On the other hand, a widened processing time has the advantage of allowing the auditory input to be integrated over a longer time window<sup>13</sup>, which is beneficial for analyzing the time-varying temporal structures of sounds, such as those important for the segmentation and categorization of phonemes. It may be that the evolution of speech in humans was supported by the enhanced capacity for sound perception that was made possible by the longer cortical processing time.

In fact, the hypothesis that the auditory cortical time window elongated in humans explains many of the previously reported differences between species in primate sound perception, particularly regarding those involving the processing of time. Discrimination of sound duration is more difficult in monkeys than in humans<sup>38</sup>, and detectability of small frequency changes improved with stimulus duration more in humans than in monkeys<sup>39</sup>. Monkeys are less sensitive than humans to temporally global features of spectrotemporal variation in sounds and focus on local features in cognitive tasks<sup>40</sup>. The cutoff frequency of the temporal modulation transfer function, which reflects the low-pass filter characteristic of an auditory system, is lower in humans than in macaques<sup>41</sup>. These diverse species differences in sound perception are consistent with a longer auditory cortical time window in humans.

Although our findings highlighted the impact of increased neuronal number in human brain evolution, we also acknowledge the importance of brain size enlargement. For example, the long-range neural communication

between the two cerebral hemispheres is constrained in speed as brain size increases, because of the limit of the action potential propagation speed<sup>42</sup>. Such constraint has been proposed to have promoted the evolution of hemispheric specialization in enlarged brains, so that fast, time-critical computations are performed by short and small circuits in one hemisphere<sup>43,44</sup>. However, brain enlargement would also increase transmission delays in long-range intra-hemispheric connections such as the arcuate fasciculus, and communications via these fibers may be slower than those across the commissure between homologous cortical areas. Thus, neuronal number and brain size have complicated effects on the evolution of neural circuitries that support human cognition, the details of which remain to be elucidated.

Although the neural generators of the P1s and N1s investigated in this study were mostly unknown, a putative macaque homolog of P1 at 28 ms was generated by current sources located in lamina II of the primary auditory cortex, which were balanced by sinks located in upper lamina III<sup>45</sup>. This sink-source configuration was consistent with polysynaptic activation of supragranular pyramidal cell elements that followed the earliest current sinks that occurred within the thalamorecipient zone (layer IV and III) at a latency of 17 ms<sup>45</sup>. Subsequently, a putative macaque homolog of N1 at around 50 ms was coincident in time with a superficial sink in the supragranular layer of the primary auditory cortex, balanced by a deeper source<sup>45,46</sup>. Therefore, assuming that our proposed homologies were valid, the CAEP components analyzed in this study likely reflected initial stages of the neural processing of sounds in the primary auditory cortex, although sources outside the primary auditory cortex may also have contributed to the CAEPs. As further evidence, all P1s and N1s were recorded maximally around the vertex, and this scalp topography was consistent with localization of the sources on the superior temporal plane within the lateral sulcus. Although there are anatomical variations across and within species, most of the primary auditory cortex is located on the upper bank of the lateral sulcus in humans, chimpanzees, and macaques<sup>47</sup>, whereas it extends laterally from within the lateral sulcus onto the lateral surface of the superior temporal gyrus in the marmoset<sup>48</sup>.

As a limitation of the study, it is important to acknowledge that we assumed that the human and nonhuman primates' P1s and N1s were functionally homologous to each other. Nevertheless, two lines of evidence support this assumption. First, the polarities of the CAEP components were preserved among the species. It seemed more natural to interpret the first and second prominent peaks as corresponding to P1 and N1, rather than pursue more complex interpretations. Second, and more importantly, the mismatch negativity (MMN), which reflects the detection of contextual deviance in sounds<sup>49</sup>, exhibits a similar evolutionary elongation of latency as the putative N1 homologs. The peak MMN latency is 150–250 ms in humans<sup>49</sup>, 125–210 ms in chimpanzee<sup>14</sup>, 80–110 ms in macaques<sup>8,10</sup>, and 60–70 ms in marmosets<sup>50</sup>, all of which are slightly longer than the latencies of their respective putative N1 homologs (Fig. 5). Therefore, it is reasonable to assume that N1, PtN1, mN1, and CjN1 reflect functionally comparable stages of auditory processing.

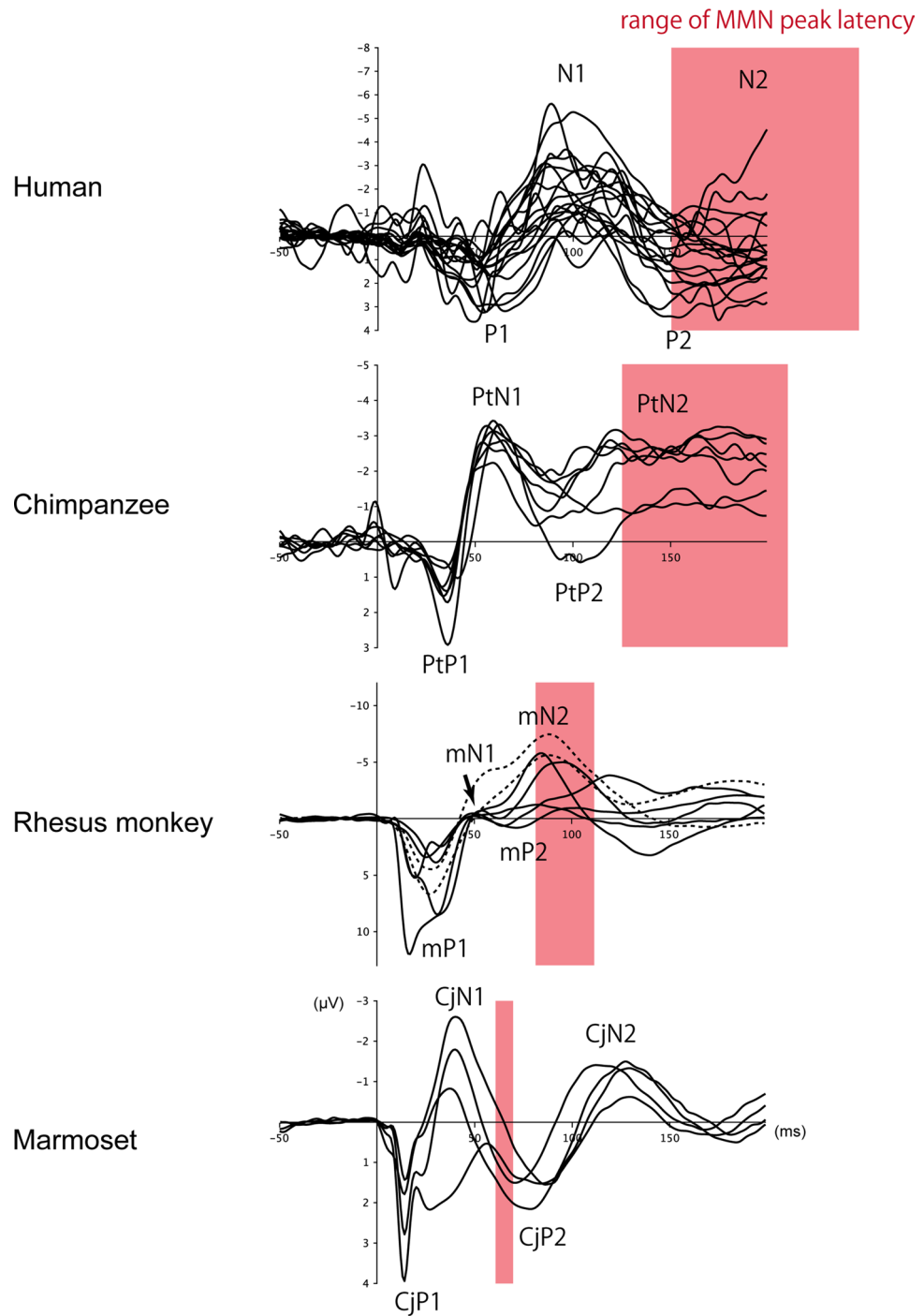
Another limitation was that only four species were included in the analyses. More accurate calculations of the allometric relationships would be possible, as future technical advances enable noninvasive CAEP recordings in other nonhuman primates. Moreover, future inclusions of non-primate species will provide a critical test of our hypothesis. Since primates have the highest density of cortical neurons among other mammals of similar brain size<sup>4</sup>, our hypothesis predicts that the CAEP latencies will be shorter in non-primate brains than in primate brains, when their brains are comparable in size. Nevertheless, notable strengths of this study were that the analysis covered a wide range of primate brain sizes, from the marmoset to the human, and that the analyses included data from a chimpanzee, which are extremely rare and valuable.

In conclusion, the enlargement of human brain size, or, more accurately, the increase in the number of cortical neurons developing over the course of primate evolution, correlated with an increase in the dwell time of auditory cortical processing in humans. This finding was logically expected, although the results were against some previous speculations that the human brain maintained a short processing time due to extensive cortical folding<sup>1,6</sup>, high neuronal density<sup>4</sup>, and thickened myelination<sup>7</sup> that kept the signal conduction time short.

A longer timescale of cortical processing allows neural inputs to the cortex to be integrated over an extended time window. In the case of audition, such elongation would enable analyses of the temporal features of time-varying acoustic stimuli that are important for the perception of speech and other complex sounds. It is possible that the evolutionary prolongation of cortical processing time might have occurred not only in audition, but also in other sensory modalities, as the latencies of visual evoked potentials are also shorter in nonhuman primates than in humans<sup>51</sup>. These findings lead to a novel hypothesis concerning human brain evolution, such that the increase in the number of cortical neurons widened the time window of neural computation over the entire cerebral cortex. In other words, if the benefit of perceptual or cognitive enhancement due to the widened time window was sufficiently large, the number of cortical neurons would be biased to increase during evolution despite the cost of longer processing time. The hypothesis that higher brain functions are supported by longer timescales of processing is consistent with the previous findings that the timescale of spontaneous firing at the level of single neurons is longer in higher association cortices compared to sensory cortices in the non-human primate brain<sup>52,53</sup>, and that higher-tier cortical areas shows extended timescales of neural processing in the human brain<sup>54</sup>. Further clarification of differences between species in how the brain handles time may provide a key to understanding the evolution of human intelligence.

## Materials and methods

**Experimental design.** The raw data obtained in several published<sup>9,11–14</sup> and unpublished experiments from our group were reanalyzed for the purpose of this study. Methods that were common to all experiments were as follows: EEGs were measured noninvasively from the scalp, subjects listened to the sounds passively without performing a behavioral task; and the sound pressure levels were medium to loud. There were also minor differences, such as the experiment site, EEG amplifier, and specific stimulus parameters, which are described below.



**Figure 5.** CAEP waveforms of all individual subjects (marmosets, rhesus monkeys, and humans), or all individual experiments (chimpanzee), compared with the MMN. Although there were variations in the waveforms within species or between experiments, within-species variations in the peak latencies of P1/N1 homologs were small compared to the between-species differences. Two rhesus monkeys (dotted trace) were excluded from the analysis of mN1, as their peak latency was not clearly determined. The red bands indicate the time range of the reported peak latencies of MMN in each species. Note that the MMNs occur slightly after the N1 homologs and that they increase in latency with brain size. CAEP cortical auditory-evoked potential, MLR middle latency response, MMN mismatch negativity.



The study is reported in accordance with the recommendations in the ARRIVE guidelines (<https://arriveguidelines.org>) for all nonhuman primate species.

**Marmoset EEG recording.** Three experiments were conducted on four adult common marmoset monkeys, consisting of three males (subject names: Cj399, Cj419, Cj459; body weight: 260–330 g; 5–6 years old) and one female (Cj368, 320 g, 7 years old). Not all subjects were used in all of the experiments. The details of animal preparation and EEG recording method are described elsewhere<sup>11</sup>. Briefly, the monkey sat in a primate chair and free movement of the head was restricted by a thermoplastic mask that was adjusted to the size and shape of each animal. Silver dish electrodes with a diameter of 4 mm (Unique Medical, Tokyo, Japan) were placed on the shaved scalp according to the International ten-twenty system. The electrodes were fixed using adhesives, and electrode gel was applied to lower the impedance to below 5 k $\Omega$ . All sounds were presented at an intensity of 65–70 dB SPL.

In Experiment 1<sup>11</sup>, the auditory stimulus was a pure tone (1500 Hz) with a duration of 300 ms (6 ms linear rise/fall times). The sound was presented at two different stimulus rates: fast and slow. In the fast-rate condition, the stimulus-onset asynchrony (SOA; the time interval between the onsets of two consecutive sounds) was 330 ms, and in the slow-rate condition, the SOA was 1130 ms. Only the slow-rate condition was analyzed because the CAEPs were not clearly elicited in the fast-rate condition. All four monkeys were used in this experiment. A total of 2000–4000 trials were presented depending on the subject, and 97.9–99.2% of those were accepted for analysis after artifact rejection (see below).

In Experiment 2 (unpublished), the auditory stimulus was a pure tone (1500 Hz) that had various durations in the range of 2–150 ms (1 ms linear rise/fall times). Only the CAEPs to sounds with a duration of 50 ms or longer were analyzed, so that the stimulus condition matched that of the human version of this experiment (see below). The sound was presented with an inter-trial silent interval of 300–400 ms. Other aspects of the experiment were the same as in Experiment 1. All four monkeys were used in this experiment. A total of 2800–7000 stimuli were presented depending on the subject, and 96.3–99.5% of those were accepted for analysis.

In Experiment 3 (unpublished), five stimuli of different frequencies (500, 750, 1125, 1688, and 2531 Hz) were presented randomly and equiprobably in sequence. The sound duration was 200 ms, and the SOA was 600 ms. The CAEPs for all frequency conditions were pooled for analysis. Two monkeys (Cj399 and Cj459) were used in this experiment. A total of 2000 trials were presented for each subject, and 98.6–99.9% of those were accepted for analysis.

After obtaining the CAEPs, the latencies of CjP1 and CjN1 were measured at the Cz electrode for each experiment and each subject. Next, the latencies were averaged across the experiments for each subject. Finally, the measurements were averaged over all subjects to obtain the CjP1 and CjN1 latencies for the marmosets. Figure 1 shows the grand-averaged CAEPs of all the experiments and all the subjects. The waveforms of the individual subjects are shown in Fig. 5.

Presentation software (Neurobehavioral Systems, Berkeley, CA) and Sound Blaster audio hardware (Creative Technology, Jurong East, Singapore) were used for stimulus presentation. Sound playback latencies with respect to the trigger signal were measured at the point of signal output from the computer, and these were corrected in post-processing. There was also a small delay of ~2 ms that reflected the traveling time of the sound wave from the loudspeaker to the subject (0.6 m); this was also corrected in post-processing. EEG was recorded using BrainAmp MR plus (Brain Products, Munich, Germany).

All marmoset experiments were conducted at the Primate Research Institute of Kyoto University (Aichi, Japan) and adhered to the Guide for the Care and Use of Laboratory Animals of the National Research Council (1996) and the Guide for Care and Use of Laboratory Primates published by the Primate Research Institute of Kyoto University (2010). The experiment adhered to the legal requirements that apply to animal research conducted in Japan and was approved by Kyoto University and Niigata University.

**Macaque EEG recording.** Four experiments were conducted on seven adult or young adult rhesus monkeys, consisting of five males (subject names: Mm1979, Mn1961, Mm1963, Mm1866, and Mm1789; body weight: 4.1–7.5 kg; 3–10 years old) and two females (Mm1892 and Mm2033; body weight, 4.5–6.2 kg, 4–7 years old). Not all subjects were used in all of the experiments. The details of the animal preparation and EEG recording method are described elsewhere<sup>12,13</sup>. Briefly, the monkey sat in a primate chair and its head was fixed by using bars or thermoplastic masks that were adjusted to the size and shape of each animal. Silver electrodes for human sleep recording (NE-136A, Nihon Kohden, Japan) were placed on the shaved scalp according to the International ten-twenty system. The electrodes were fixed using gauze soaked with collodion, and electrode gel was applied between the electrode and the scalp to lower the impedance to below 5 k $\Omega$ . All sounds were presented at an intensity of 65–70 dB SPL, except in Experiment 4.

Experiment 1 was a reanalysis of our previously published data<sup>12</sup>, in which a single pure tone (1500 Hz) was presented repeatedly. The stimulation conditions were identical to those of Experiment 1 in the marmoset study, except that the SOA was 300 ms and 1100 ms for the fast- and slow-rate conditions, respectively. Only the slow-rate condition was analyzed, as in the marmoset experiment. Three monkeys (Mm1892, Mm1963, and Mm1789) were used in this experiment. A total of 923–7039 stimuli were presented depending on the subject, and 88.0–94.1% of those were accepted for analysis.

Experiment 2 was a reanalysis of the data obtained in a previous study<sup>13</sup>, in which a pure-tone stimulus (1500 Hz) of varying durations (2–100 ms) was presented, similar to Experiment 2 in marmosets. Only CAEPs with a duration of 50 ms or longer were analyzed, as in the marmoset study. Six monkeys (Mm1979, Mm1892, Mm1961, Mm1963, Mm2033, and Mm1866) were used in the original experiment (13), but another subject

(Cj2033) was added for the present study. A total of 8250–12,000 stimuli were presented depending on the subject, and 85.0–94.7% of those were accepted for analysis.

Experiment 3 (unpublished) was similar to that of Experiment 3 in marmosets. Five stimuli of different frequencies (667, 1000, 1500, 2250, and 3375 Hz) were presented randomly and equiprobably in sequence. The sound duration was 300 ms, and the SOA was 570 ms. All frequency conditions were pooled. Two monkeys (Mm1892 and Mm1866) were used in this experiment. A total of 4049–7990 stimuli was presented depending on the subject, and 76.9–84.7% of those were accepted for analysis.

Experiment 4 (unpublished) was an MMN experiment, in which a click stimulus (0.1 ms duration) was presented at two levels of intensity in a random order: (95 dB  $LC_{peak}$ ) or soft (75 dB  $LC_{peak}$ ). In the standard soft condition, the probability of stimulus presentation was 0.8 and 0.2 for the soft and loud clicks, respectively; the probabilities were reversed in the standard loud condition. Only the CAEPs to the standard stimuli were analyzed, indistinguishing the soft and loud standards. A total of 5435–5440 standard stimuli were presented depending on the subject, and 79.2–91.1% of those were accepted for analysis.

After obtaining the CAEPs, the mP1 and mN1 latencies were measured at Cz for each experiment and for each subject. Next, the latencies were averaged across the experiments for each subject. Finally, the measurements were averaged over all subjects to obtain the mP1 and mN1 latencies for rhesus monkeys. Subjects Mm1961 and Mm1866 were excluded from the analysis of mN1 because they did not elicit a clear mN1. The waveform in Fig. 1 was obtained by averaging the CAEPs across all experiments and all subjects, excluding those that did not elicit a clear mN1. Individual waveforms of all subjects, including Mm1961 and Mm1866, are shown in Fig. 5.

Presentation software (Neurobehavioral Systems, Berkeley, CA) and Sound Blaster audio hardware (Creative Technology, Jurong East, Singapore) were used for stimulus presentation. Sound playback latencies with respect to the trigger signal were measured at the point of signal output from the computer, and these were corrected in post-processing. Even after this correction, there remained a small delay of ~2 ms that reflected the traveling time of the sound wave from the loudspeaker to the subject (0.6 m), which was also corrected in post-processing. EEG was recorded using BrainAmp MR plus (Brain Products, Munich, Germany).

All rhesus monkey experiments were conducted at the Primate Research Institute of Kyoto University (Aichi, Japan) and adhered to the Guide for the Care and Use of Laboratory Animals of the National Research Council (1996) and the Guide for Care and Use of Laboratory Primates published by the Primate Research Institute of Kyoto University (2010). The experiment adhered to the legal requirements for the conduct of animal research in Japan and was approved by Kyoto University and Niigata University.

**Chimpanzee EEG recording.** Three experiments were performed on a single nine-year-old chimpanzee named Mizuki. A larger number of chimpanzees would have made the findings more robust. However, such opportunity is extremely rare, and the inclusion of the chimpanzee data is a strength of this study rather than a limitation. Details of animal preparation and the EEG recording method are described elsewhere<sup>9,14</sup>. Briefly, the method of electrode application was similar to that for macaque monkeys, except that the animal was not physically restrained from free body and head movements. All sounds were presented at an intensity of 80 dB HL. Seven CAEPs were obtained from the chimpanzee, as described below, and the waveform in Fig. 1 shows their grand average.

Experiment 1 was a reanalysis of a previously published MMN experiment<sup>14</sup>, in which a pure-tone stimulus of one frequency was presented as the standard stimulus ( $p=0.8$ ) and another pure-tone stimulus of a different frequency as the deviant stimulus ( $p=0.2$ ). Only the CAEPs to the standard stimulus were analyzed. The sound duration was 100 ms, and the SOA was 700 ms. The stimulus frequency varied across four sessions (500 Hz vs. 1500 Hz, 500 Hz vs. 2000 Hz, 1500 Hz vs. 500 Hz, 2000 Hz vs. 500 Hz), and the different sessions were analyzed separately; that is, four independent CAEPs were obtained to measure the PtP1 and PtN1 latencies. A total of 1460–1605 standard stimuli were presented depending on the session, and 96.2–97.7% of those were accepted for analysis.

Experiment 2 was a standard CAEP experiment, in which a 500 Hz pure tone (duration 100 ms) was presented repeatedly with a fixed SOA of 700 ms. A total of 1615 trials was presented, and 96.8% of those were accepted for analysis.

In Experiment 3, the auditory stimulus was the subject's own name (Mizuki) that was uttered by a human in an angry or a friendly voice. This experiment was a variation of another study that recorded CAEP to the subject's own name<sup>9</sup>. The name stimuli were presented at a probability of  $p=0.1$  each, and a pure-tone stimulus (500 Hz, 100 ms duration) was presented at a probability of  $p=0.8$ . Two CAEPs were obtained one for the name stimuli without distinguishing the voices, and the other for the pure-tone stimulus. A total of 3128 trials was presented for the name stimuli, and 97.0% of the trials were accepted for analysis. A total of 12,458 trials was presented for the pure tone, and 97.0% of those were accepted for analysis. Although the name stimulus was acoustically more complex than pure-tones, the latencies of P1 and N1 were resilient to the stimulus difference, as shown in Fig. 5.

After obtaining the CAEPs, the PtP1 and PtN1 latencies were measured at Cz for each experiment. Next, the latencies were averaged across experiments to obtain the PtP1 and PtN1 latencies for the chimpanzee. Individual waveforms of all experiments are shown in Fig. 5, and the grand average is shown in Fig. 1.

The stimulus delivery hardware (STIM2, Neuroscan, Inc.) was configured to have no sound playback latency with respect to the trigger signal. There was a delay of ~6 ms that reflected the traveling time of the sound wave from the loudspeaker to the subject (2 m), which was corrected in post-processing. The EEG signals were obtained using a NuAmp-40 amplifier (NeuroScan Inc., El Paso, USA).

All chimpanzee experiments were conducted at the Great Ape Research Institute, Hayashibara Biochemical Laboratories, Inc. (Okayama, Japan). The experiment adhered to the legal requirements of Japan and the research protocol was approved by the Research Ethics Committee of the Hayashibara Great Ape Research Institute.

**Human EEG recording.** Human EEG data were obtained as an integrated part of the corresponding non-human primate experiments described above for species comparisons. Therefore, the stimuli and recording conditions were matched as closely as possible within those experiments.

Experiment 1 was a human version of the Experiment 2 in rhesus monkeys, the details of which are described above and elsewhere<sup>13</sup>. Twelve subjects (two males and ten females, aged 18–23 years) participated in the study. Pure-tone stimuli (1500 Hz) of varying durations (2–200 ms) were presented. However, only the CAEPs to sounds with a duration of 50 ms or longer were analyzed, because the latencies of P1 and N1 were clearly longer for shorter sounds; see Fig. 1 in<sup>13</sup>. A total of 1794–1800 stimuli were presented depending on the subject, and 97.4–99.9% of those were accepted for analysis.

The Presentation software (Neurobehavioral Systems, Berkeley, CA) and Sound Blaster audio hardware (Creative Technology, Jurong East, Singapore) were used for stimulus presentation in Experiment 1. Sound playback latencies with respect to the trigger signal were measured at the point of signal output from the computer, and these were corrected in post-processing. In addition, there was a small delay of ~4 ms, which reflected the traveling time of the sound wave from the loudspeaker to the subject (1.2 m); this was also corrected in post-processing. The EEG was recorded using a BrainAmp DC (Brain Products, Munich, Germany).

The experiment was conducted at Niigata University after the Internal Review Board of the University of Niigata approved the study. All experiments were performed in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained from all participants.

Experiment 2 was a human version of the chimpanzee's Experiment 2, and the stimuli were identical. Six subjects (two males and four females, aged 23–53 years) participated in the study. A total of 100–1000 trials was presented depending on the subject, and 97.4–100% of those were accepted for analysis. The EEG signals were obtained using NuAmp-40 (NeuroScan Inc., El Paso, USA).

The stimulus delivery hardware (STIM2, Neuroscan, Inc.) was configured to have no sound playback latency with respect to the trigger signal. There was a delay of ~6 ms that reflected the traveling time of the sound wave from the loudspeaker to the subject (2 m), which was corrected in post-processing.

The experiment was conducted at the Great Ape Research Institute, Hayashibara Biochemical Laboratories, Inc. (Okayama, Japan). The research protocol was approved by the Research Ethics Committee of the Hayashibara Great Ape Research Institute. Informed consent was obtained from all participants.

**Data analyses.** The EEG data were analyzed similarly for all experiments in all species to obtain the CAEPs, utilizing the EEGLAB software<sup>55</sup> with the ERPLAB extension<sup>56</sup>. Only the Fz, Cz, Pz, and ear channels were used, where Fz, Cz, and Pz refer to the electrode positions as specified by the 10–20 system. The data were first band-pass filtered (1–100 Hz), and re-referenced to the average of both ears in the marmoset, macaque, and human, or to the left ear in the chimpanzee (as data were not recorded from the right ear in the chimpanzee). The continuous EEG data were then segmented, time locked to the onset of the stimulus (–50 to 200 ms), corrected for baseline by subtracting the pre-stimulus period average, screened for artifacts, and accordingly rejected if they contained an artifact of  $\pm 100 \mu\text{V}$  relative to the baseline. One CAEP was obtained for each experimental condition for each subject. These waveforms were used for two further analyses: inspection of waveform morphology and latency measurement.

For the inspection of waveforms, the CAEPs were averaged over experimental conditions for each subject, and then over subjects within species to obtain a grand-averaged CAEP for each species (Fig. 1).

For the latency measurement, the latencies were recorded, where an inflection point of the CAEP waveform represented the local minimum (N1) or the local maximum (P1) within the respective time windows that were defined by inspecting the grand-average waveforms (Fig. 1). The time window for P1 was defined as 0–100 ms in humans and 0–50 ms in the other species, as measured at Cz. The time window for N1 was 50–150 ms in humans and 30–80 ms in the other species, as measured at Cz. Two rhesus monkeys were excluded from the analysis of mN1, as the CAEP amplitude decreased monotonically and showed no inflection point within the mN1 time window (Fig. 5).

For the allometric analysis of P1/N1 latencies and brain size (Fig. 3), the latencies were averaged over subjects to obtain a single latency measure for each species, so that the number of cases matched between the species for equal weighting in regression. Then, the data were fitted by linear regression, using the equivalent brain radius ( $R_{eq}$ ) or the number of neurons ( $N$ ) as the independent variable, using the CURVEFIT function of IBM SPSS Statistics version 25 (IBM, Armonk, NY, United States).

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 13 September 2021; Accepted: 4 January 2022

Published online: 20 January 2022

## References

- Hofman, M. A. Evolution of the human brain: When bigger is better. *Front. Neuroanat.* **8**, 15. <https://doi.org/10.3389/fnana.2014.00015> (2014).
- Jernison, H. J. *Evolution of the Brain and Intelligence* (Academic Press, 1973).
- Herculano-Houzel, S. Numbers of neurons as biological correlates of cognitive capability. *Curr. Opin. Behav. Sci.* **16**, 1–7. <https://doi.org/10.1016/j.cobeha.2017.02.004> (2017).
- Herculano-Houzel, S., Manger, P. R. & Kaas, J. H. Brain scaling in mammalian evolution as a consequence of concerted and mosaic changes in numbers of neurons and average neuronal cell size. *Front. Neuroanat.* <https://doi.org/10.3389/fnana.2014.00077> (2014).

5. Herculano-Houzel, S. The human brain in numbers: A linearly scaled-up primate brain. *Front. Hum. Neurosci.* **3**, 31. <https://doi.org/10.3389/neuro.09.031.2009> (2009).
6. Mota, B. & Herculano-Houzel, S. How the cortex gets its folds: An inside-out, connectivity-driven model for the scaling of Mammalian cortical folding. *Front. Neuroanat.* **6**, 3. <https://doi.org/10.3389/fnana.2012.00003> (2012).
7. Roth, G. & Dicke, U. Evolution of the brain and intelligence. *Trends Cogn. Sci.* **9**, 250–257. <https://doi.org/10.1016/j.tics.2005.03.005> (2005).
8. Gil-da-Costa, R., Stoner, G. R., Fung, R. & Albright, T. D. Nonhuman primate model of schizophrenia using a noninvasive EEG method. *Proc. Natl. Acad. Sci. USA* **110**, 15425–15430. <https://doi.org/10.1073/pnas.1312264110> (2013).
9. Hirata, S. *et al.* Event-related potentials in response to subjects' own names: A comparison between humans and a chimpanzee. *Commun. Integr. Biol.* **4**, 321–323. <https://doi.org/10.4161/cib.4.3.14841> (2011).
10. Honing, H., Merchant, H., Haden, G. P., Prado, L. & Bartolo, R. Rhesus monkeys (*Macaca mulatta*) detect rhythmic groups in music, but not the beat. *PLoS ONE* **7**, e51369. <https://doi.org/10.1371/journal.pone.0051369> (2012).
11. Itoh, K., Iwaoki, H., Konoike, N., Igarashi, H. & Nakamura, K. Noninvasive scalp recording of the middle latency responses and cortical auditory evoked potentials in the alert common marmoset. *Hear. Res.* **1**, 108229. <https://doi.org/10.1016/j.heares.2021.108229> (2021).
12. Itoh, K., Nejime, M., Konoike, N., Nakada, T. & Nakamura, K. Noninvasive scalp recording of cortical auditory evoked potentials in the alert macaque monkey. *Hear. Res.* **327**, 117–125. <https://doi.org/10.1016/j.heares.2015.05.007> (2015).
13. Itoh, K., Nejime, M., Konoike, N., Nakamura, K. & Nakada, T. Evolutionary elongation of the time window of integration in auditory cortex: Macaque vs. human comparison of the effects of sound duration on auditory evoked potentials. *Front. Neurosci.* **13**, 630. <https://doi.org/10.3389/fnins.2019.00630> (2019).
14. Ueno, A. *et al.* Auditory ERPs to stimulus deviance in an awake chimpanzee (*Pan troglodytes*): Towards hominid cognitive neurosciences. *PLoS ONE* **3**, e1442. <https://doi.org/10.1371/journal.pone.0001442> (2008).
15. Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P. & Chauvel, P. Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalogr. Clin. Neurophysiol. Evok. Potent. Sect.* **92**, 204–214. [https://doi.org/10.1016/0168-5597\(94\)90064-7](https://doi.org/10.1016/0168-5597(94)90064-7) (1994).
16. Brugge, J. F. *et al.* Functional localization of auditory cortical fields of human: Click-train stimulation. *Hear. Res.* **238**, 12–24. <https://doi.org/10.1016/j.heares.2007.11.012> (2008).
17. Nourski, K. V. *et al.* Functional organization of human auditory cortex: Investigation of response latencies through direct recordings. *Neuroimage* **101**, 598–609. <https://doi.org/10.1016/j.neuroimage.2014.07.004> (2014).
18. Khani, A. *et al.* Large-scale networks for auditory sensory gating in the awake mouse. *ENeuro* <https://doi.org/10.1523/ENEURO.0207-19.2019> (2019).
19. Joachimsthaler, B., Uhlmann, M., Miller, F., Ehret, G. & Kurt, S. 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. <https://doi.org/10.1111/ejn.12478> (2014).
20. Näätänen, R. & Picton, T. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology* **24**, 375–425. <https://doi.org/10.1111/j.1469-8986.1987.tb00311.x> (1987).
21. Picton, T. W., Hillyard, S. A., Krausz, H. I. & Galambos, R. Human auditory evoked potentials I: Evaluation of components. *Electroencephalogr. Clin. Neurophysiol.* **36**, 179–190. [https://doi.org/10.1016/0013-4694\(74\)90155-2](https://doi.org/10.1016/0013-4694(74)90155-2) (1974).
22. Itoh, K., Okumiya-Kanke, Y., Nakayama, Y., Kwee, I. L. & Nakada, T. Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *Eur. J. Neurosci.* **36**, 3580–3592. <https://doi.org/10.1111/j.1460-9568.2012.08278.x> (2012).
23. Jerger, J. & Jerger, S. Evoked response to intensity and frequency change. *Arch. Otolaryngol.* **91**, 433–436. <https://doi.org/10.1001/archotol.1970.00770040627007> (1970).
24. Weise, A., Schröger, E. & Horváth, J. The detection of higher-order acoustic transitions is reflected in the N1 ERP. *Psychophysiology* **55**, e13063. <https://doi.org/10.1111/psyp.13063> (2018).
25. Allen, A. R. & Starr, A. Auditory brain stem potentials in monkey (*M. mulatta*) and man. *Electroencephalogr. Clin. Neurophysiol.* **45**, 53–63. [https://doi.org/10.1016/0013-4694\(78\)90341-3](https://doi.org/10.1016/0013-4694(78)90341-3) (1978).
26. Harada, T. & Tokuriki, M. Brain-stem auditory evoked potentials in the common marmoset (*Callithrix jacchus*). *Electroencephalogr. Clin. Neurophysiol.* **104**, 43–50. [https://doi.org/10.1016/s0168-5597\(96\)96015-3](https://doi.org/10.1016/s0168-5597(96)96015-3) (1997).
27. Kraus, N., Smith, D. I., Reed, N. L., Willott, J. & Erwin, J. Auditory brainstem and middle latency responses in non-human primates. *Hear. Res.* **17**, 219–226. [https://doi.org/10.1016/0378-5955\(85\)90066-8](https://doi.org/10.1016/0378-5955(85)90066-8) (1985).
28. Legatt, A. D., Arezzo, J. C. & Vaughan, H. G. Jr. Short-latency auditory evoked potentials in the monkey. I. Wave shape and surface topography. *Electroencephalogr. Clin. Neurophysiol.* **64**, 41–52 (1986).
29. Benson, D. A. & Hienz, R. D. Single-unit activity in the auditory cortex of monkeys selectively attending left vs. right ear stimuli. *Brain Res.* **159**, 307–320. [https://doi.org/10.1016/0006-8993\(78\)90537-1](https://doi.org/10.1016/0006-8993(78)90537-1) (1978).
30. Camalier, C. R., D'Angelo, W. R., Sterbing-D'Angelo, S. J., de la Mothe, L. A. & Hackett, T. A. Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc. Natl. Acad. Sci. USA* **109**, 18168–18173. <https://doi.org/10.1073/pnas.1206387109> (2012).
31. Bender, D. A., Ni, R. & Barbour, D. L. Spontaneous activity is correlated with coding density in primary auditory cortex. *J. Neurophysiol.* **116**, 2789–2798. <https://doi.org/10.1152/jn.00474.2016> (2016).
32. Kajikawa, Y., de La Mothe, L., Blumell, S. & Hackett, T. A. A comparison of neuron response properties in areas A1 and CM of the marmoset monkey auditory cortex: Tones and broadband noise. *J. Neurophysiol.* **93**, 22–34. <https://doi.org/10.1152/jn.00248.2004> (2005).
33. Wong, P. *et al.* Faster scaling of auditory neurons in cortical areas relative to subcortical structures in primate brains. *Brain Behav. Evol.* **81**, 209–218. <https://doi.org/10.1159/000350709> (2013).
34. DeCasien, A. R., Williams, S. A. & Higham, J. P. Primate brain size is predicted by diet but not sociality. *Nat. Ecol. Evol.* **1**, 0112. <https://doi.org/10.1038/s41559-017-0112> (2017).
35. Gompertz, R. H. Specific gravity of the brain. *J. Physiol.* **27**, 459–462. <https://doi.org/10.1113/jphysiol.1902.sp000884> (1902).
36. Herculano-Houzel, S., Catania, K., Manger, P. R. & Kaas, J. H. Mammalian brains are made of these: a dataset of the numbers and densities of neuronal and nonneuronal cells in the brain of glires, primates, scandentia, eulipotyphlans, afrotherians and artiodactyls, and their relationship with body mass. *Brain Behav. Evol.* **86**, 145–163. <https://doi.org/10.1159/000437413> (2015).
37. Semendeferi, K., Lu, A., Schenker, N. & Damasio, H. Humans and great apes share a large frontal cortex. *Nat. Neurosci.* **5**, 272–276. <https://doi.org/10.1038/nn814> (2002).
38. Sinnott, J. M. & Brown, C. H. Effects of varying signal and noise levels on pure-tone frequency discrimination in humans and monkeys. *J. Acoust. Soc. Am.* **93**, 1535–1540. <https://doi.org/10.1121/1.406811> (1993).
39. Sinnott, J. M., Owren, M. J. & Petersen, M. R. Auditory duration discrimination in Old World monkeys (*Macaca*, *Cercopithecus*) and humans. *J. Acoust. Soc. Am.* **82**, 465–470. <https://doi.org/10.1121/1.395447> (1987).
40. Izumi, A. Effect of temporal separation on tone-sequence discrimination in monkeys. *Hear. Res.* **175**, 75–81 (2003).
41. Moody, D. B., May, B., Cole, D. M. & Stebbins, W. C. The role of frequency modulation in the perception of complex stimuli by primates. *Exp. Biol.* **45**, 219–232 (1986).

42. Phillips, K. A. *et al.* The corpus callosum in primates: processing speed of axons and the evolution of hemispheric asymmetry. *Proc. Biol. Sci.* **282**, 20151535. <https://doi.org/10.1098/rspb.2015.1535> (2015).
43. Aboitiz, F., López, J. & Montiel, J. Long distance communication in the human brain: Timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biol. Res.* **36**, 89–99. <https://doi.org/10.4067/s0716-97602003000100007> (2003).
44. Ringo, J. L., Doty, R. W., Demeter, S. & Simard, P. Y. Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cereb. Cortex* **4**, 331–343. <https://doi.org/10.1093/cercor/4.4.331> (1994).
45. Fishman, Y. I., Reser, D. H., Arezzo, J. C. & Steinschneider, M. Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. *J. Acoust. Soc. Am.* **108**, 247–262 (2000).
46. Fishman, Y. I. & Steinschneider, M. Searching for the mismatch negativity in primary auditory cortex of the awake monkey: Deviance detection or stimulus specific adaptation?. *J. Neurosci.* **32**, 15747–15758. <https://doi.org/10.1523/JNEUROSCI.2835-12.2012> (2012).
47. Hackett, T. A., Preuss, T. M. & Kaas, J. H. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J. Comp. Neurol.* **441**, 197–222. <https://doi.org/10.1002/cne.1407> (2001).
48. de la Mothe, L. A., Blumell, S., Kajikawa, Y. & Hackett, T. A. Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *J. Comp. Neurol.* **496**, 27–71. <https://doi.org/10.1002/cne.20923> (2006).
49. Naatanen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin. Neurophysiol.* **118**, 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026> (2007).
50. Komatsu, M., Takaura, K. & Fujii, N. Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms. *Sci. Rep.* **5**, 15006. <https://doi.org/10.1038/srep15006> (2015).
51. Woodman, G. F., Kang, M.-S., Rossi, A. F. & Schall, J. D. Nonhuman primate event-related potentials indexing covert shifts of attention. *Proc. Natl. Acad. Sci.* **104**, 15111–15116. <https://doi.org/10.1073/pnas.0703477104> (2007).
52. Cavanagh, S. E., Hunt, L. T. & Kennerley, S. W. A diversity of intrinsic timescales underlie neural computations. *Front. Neural Circuits* <https://doi.org/10.3389/fncir.2020.615626> (2020).
53. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663. <https://doi.org/10.1038/nn.3862> (2014).
54. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31**, 2906–2915. <https://doi.org/10.1523/JNEUROSCI.3684-10.2011> (2011).
55. Delorme, A. & Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009> (2004).
56. Lopez-Calderon, J. & Luck, S. J. ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* **8**, 213. <https://doi.org/10.3389/fnhum.2014.00213> (2014).
57. Herculano-Houzel, S. & Kaas, J. H. Gorilla and orangutan brains conform to the primate cellular scaling rules: Implications for human evolution. *Brain Behav. Evol.* **77**, 33–44. <https://doi.org/10.1159/000322729> (2011).

## Author contributions

Conceptualization: K.I. Methodology: K.I., N.K., M.N., H.Iw., S.H., K.N. Investigation: K.I., N.K., M.N., H.Iw., S.H., K.N. Funding acquisition: K.I., S.H., K.N. Supervision: H.Ig., K.N. Writing—original draft: K.I., N.K., M.N., H.Iw., S.H., H.Ig., K.N.

## Funding

JPS KAKENHI Grant 19H05309 (KI). JPS KAKENHI Grant 21H00304 (KI). Cooperative Research Program, Primate Research Institute, Kyoto University (KI). JPS KAKENHI grant 18H05524 (SH). Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) (KN).

## Competing interests

The authors declare no competing interests.

## Additional information

Correspondence and requests for materials should be addressed to K.I.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2022