A voice region in the monkey brain

Christopher I Petkov[1], Christoph Kayser[1], Thomas Steudel[1], Kevin Whittingstall[1], Mark Augath[1] & Nikos K Logothetis[1,2]

For vocal animals, recognizing species-specific vocalizations is important for survival and social interactions. In humans, a voice region has been identified that is sensitive to human voices and vocalizations. As this region also strongly responds to speech, it is unclear whether it is tightly associated with linguistic processing and is thus unique to humans. Using functional magnetic resonance imaging of macaque monkeys (Old World primates, *Macaca mulatta*) we discovered a high-level auditory region that prefers species-specific vocalizations over other vocalizations and sounds. This region not only showed sensitivity to the 'voice' of the species, but also to the vocal identify of conspecific individuals. The monkey voice region is located on the superior-temporal plane and belongs to an anterior auditory 'what' pathway. These results establish functional relationships with the human voice region and support the notion that, for different primate species, the anterior temporal regions of the brain are adapted for recognizing communication signals from conspecifics.

The human voice is a sound that carries an acoustical signature of our species and is rich in acoustical features that are useful for numerous vocal recognition abilities. For example, we can distinguish human vocal sounds from other sounds or vocalizations, discriminate the voice of different speakers and recognize the voice of someone that we know. Because much of our social communication depends on recognizing voices in various contexts, scientists have wondered whether the human brain has regions that are specialized for processing the human voice. Earlier lesion studies in humans have revealed deficits in vocal recognition and discrimination, together called phonagnosia[1,2]. Then a voice-preferring region was identified in humans with functional imaging[3], which is located anteriorly on the temporal lobe, on the upper bank of the superior-temporal sulcus (STS). This region has been shown to prefer human voices and vocalizations over other animal vocalizations, acoustical controls and natural sounds[3–6]. These studies provide evidence that the voice region is specialized for processing the acoustical features that distinguish our vocalizations from other sounds: the voice of our species. Further evidence suggests that the human voice region is also sensitive to the acoustics that identify different human voices[7], a sensitivity that could be used to recognize the voices of different individuals.

The human voice region might constitute an auditory analog to the face-preferring visual regions that have been identified in both humans and monkeys[8–11]. However, it does not immediately follow that the voice region that has been described for humans exists in other animals. For instance, the processing of human vocalizations could depend on linguistic processing. After all, humans use their voices as a vehicle for expressing vocal language, and speech strongly activates the voice region or nearby regions[3,4,12,13]. Studies with nonhuman primates are important because they can reveal whether voice regions were evolutionarily conserved in primates. These animals can also shed light on whether such regions depend on linguistic processing, as nonhuman primates, like many other vocal animals, lack the expressive vocal range and a number of linguistic capabilities of humans[14]. Yet many primates readily orient to species-specific vocalizations and can recognize conspecifics by their vocalizations[15–17], suggesting that their auditory system might include regions dedicated to processing species-specific vocalizations.

A first step in localizing and evaluating the functional properties of brain regions in the nonhuman primate, whose activity could be related to the animal's capacity to recognize verbal communications from members of their species, would be to identify brain regions that strongly prefer species-specific vocalizations. Electrophysiological studies are useful to consider, as many have used species-specific sounds as meaningful sounds to evaluate neuronal responses (for example, see refs. 18–21). However, few studies have been able to quantify the neuronal preference for species-specific vocalizations in relation to other categories of sounds (including natural sounds and acoustical controls)[22]. Also, as most of the recordings have targeted primary auditory cortex and the neighboring (belt) regions, it is uncertain how neurons in other auditory regions compare with these regions. Imaging studies are ideal for mapping many brain regions, yet recent reports of positron emission tomography (PET) imaging in monkeys using species-specific vocalizations provide conflicting results and a lack of evidence for a voice region, which can only be interpreted as evidence that the human voice region is unique[23–25].

Here, we used high-resolution functional magnetic resonance imaging (fMRI) with different auditory procedures to demonstrate the existence of a voice region in the rhesus macaque monkey. We first localized numerous auditory cortical fields in each subject by using a

ARTICLES
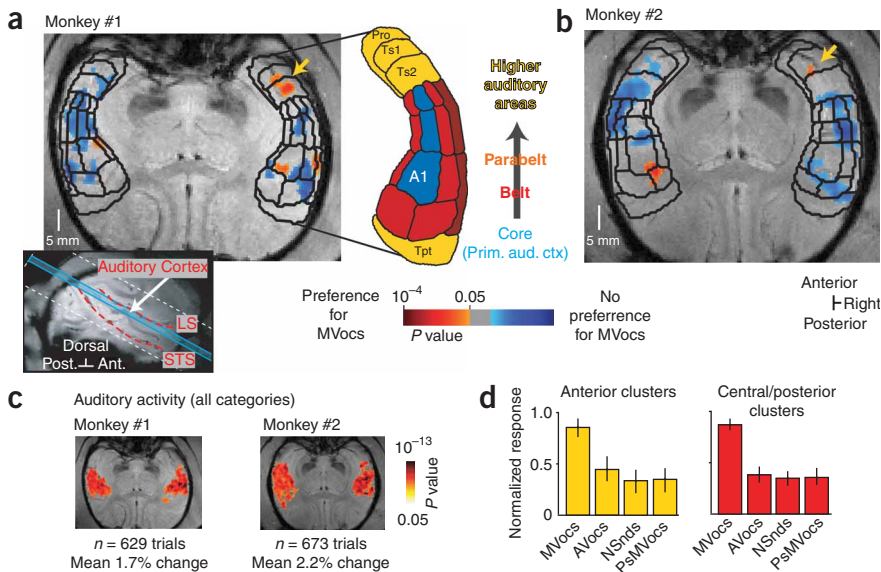
Figure 1 Auditory cortex regions preferring species-specific vocalizations in two awake monkeys. (**a,b**) Combined and coregistered data from six of the experiments with each animal. The color code from orange to red indicates voxels with a clear and significant preference for macaque vocalizations. The cyan-to-blue color code identifies voxels with no preference for MVocs. The slice orientation and position are shown in the lower inset of **a**. LS, lateral sulcus. The black contours outline the functionally or anatomically parcellated regions obtained independently of the main experiments for each animal; the blow-up identifies fields and regions of auditory cortex[26,34–36]. (**c**) The mean auditory cortex activity for each animal as the percentage of signal in relation to the silent baseline across the sound categories. (**d**) Voxel-based normalized response for the anterior clusters (orange arrows, $n = 24$ voxels; **a,b**) and the more distributed central/posterior clusters ($n = 58$ voxels) that prefer MVocs. The MVocs response was significantly higher than the other conditions in these regions of auditory cortex (bars show the mean and the 5 and 95% confidence intervals, bootstrap procedure). A1, field in primary auditory cortex; Tpt, temporoparietal; Pro, proisocortex of the temporal pole.

recently developed technique[26,27], which allowed us to evaluate the multiple processing stages and auditory cortical pathways in the macaque[28–31]. Then, following an approach employed in human studies[3–5], we identified several brain regions that showed a clear response preference for species-specific vocalizations in comparison to a large number of other animal vocalizations and sounds. Out of this set of regions, only one, a higher-level auditory region, was reliably observed across experiments and animals, which we identified as a candidate monkey voice region. Additional experiments established further functional relationships with the known human voice region. First, this candidate monkey voice region retained its preference for macaque vocalizations when tested on a set of familiar sounds (including the vocalizations of several familiar conspecifics). Second, only this anterior monkey voice region showed selectivity for the vocal identity of the conspecific individuals that elicited the vocalizations.
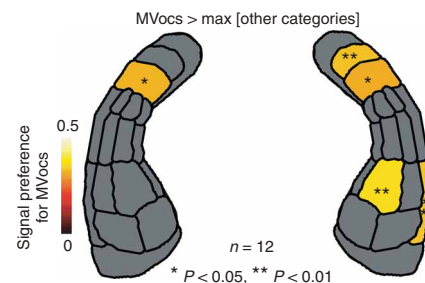
RESULTS
To identify which brain regions preferentially respond to species-specific vocalizations in Experiment 1, we used a large repertoire of acoustical stimuli. These included four categories of sounds: species-specific macaque vocalizations (MVocs), heterospecific animal vocalizations (AVocs), natural sounds (NSnds) and an acoustical control that preserved the spectrum and duration of the species-specific vocalizations (PsMVocs). The animal-vocalization categories (MVocs and AVocs) contained different vocalization calls from many animals that were unknown to the subjects, and all sounds were unfamiliar to

the subjects prior to the experiments (see **Supplementary Fig. 1** online for the acoustical characteristics of the sound categories).

Two trained animals were scanned while being stimulated with sounds from each category using a previously described auditory fMRI procedure[26,27] (**Supplementary Fig. 2** online). The imaging procedure aligns the imaging slices with the monkey superior-temporal plane (STP; **Fig. 1**). As a necessary prerequisite for precise localization, we first mapped numerous auditory fields for each animal independently of the main experiments[26]. Our mapping technique functionally localizes 11 auditory fields from the first and second stages of the auditory cortical processing hierarchy (core and belt regions, respectively; **Fig. 1**). Briefly, the mapping technique uses pure-tone versus band-passed noise responses and reversals of best-center frequency of the sounds[26], which has a basis in monkey electrophysiology[32] and human imaging[33]. In relation to these fields, we localized additional auditory regions on the STP and superior-temporal gyrus (STG) using anatomical landmarks[34–36], demarcating a total of 17 auditory cortical fields in each hemisphere (**Fig. 1a,b**, black contour lines).

Throughout the brains of the two macaques, we observed several regions with an activity preference for macaque vocalizations. These are first described for auditory cortex using a voxel-based analysis, whereby patches or clusters of voxels identify the MVocs-preferring regions (**Fig. 1**). To identify these regions we used the conservative criterion that MVocs elicit significantly greater activity than the maximum response to any of the other three categories. The results from multiple coregistered experiments with each animal are summarized in **Figure 1** (6 experiments and >600 imaging trials each; **Fig. 1c**). The color code from orange to red indicates voxels with a clear

Figure 2 Auditory cortical field summaries of MVocs preference. Across experiment results are summarized for the localized auditory cortical fields shown here as outlined ROIs on hemispheric models of the fields (see **Fig. 1** for the labels of auditory regions and fields). The models stereotype the shape and location of the many auditory cortical fields that were localized independently for each animal. These analyses are across the 12 experiments with the two awake animals (Monkeys 1 and 2). Only regions with significant MVocs-preferring signal are color coded (MVocs > max [AVocs, NSnds, PsMVocs]); the others are shown in gray (see Methods).
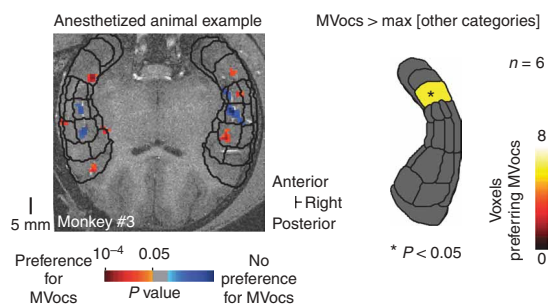
Figure 3 Experiments with anesthetized animals. Left, an example experiment with an anesthetized animal (see **Supplementary Fig. 3** for additional examples). Format and statistics are the same as in **Figure 1a,b**. Right, summary across experiments on auditory cortical field ROIs, averaged across both hemispheres (see **Fig. 1** for the labels of auditory fields and regions). Only regions with significant MVocs-preferring voxel counts are color coded (MVocs > max [AVocs, NSnds, PsMVocs]); the others are shown in gray.

and significant preference for macaque vocalizations. The cyan-to-blue color code identifies voxels with no preference for MVocs: voxels that, because of our statistical (max) comparison, preferred one of the other three sound categories. In the blue regions, however, we observed no consistent preference for the acoustics of the AVocs, NSnds or PsMVocs sound categories.

We observed clusters that strongly preferred MVocs in two auditory cortex regions, one anterior and another that was central or more posterior (orange to red, **Fig. 1a,b**). Specifically, in both animals we found a similarly localized right-hemisphere cluster in the anterior auditory regions temporalis superior 1 and 2 (Ts1 and Ts2, respectively) both of which are hierarchically beyond the first several stages of auditory cortex processing: namely, beyond the auditory 'core' (primary auditory cortex, including field A1), the 'belt' and the 'parabelt' regions (**Fig. 1a**)[36]. The second, central/posterior MVocs-preferring region was composed of clusters in the hierarchically earlier regions of both hemispheres, including field A1. No other significant MVocs-preferring clusters were seen anywhere else on these slices (the outer contour of auditory cortex merely delineates the cyan-to-blue activity that trivially predominated throughout the brain).
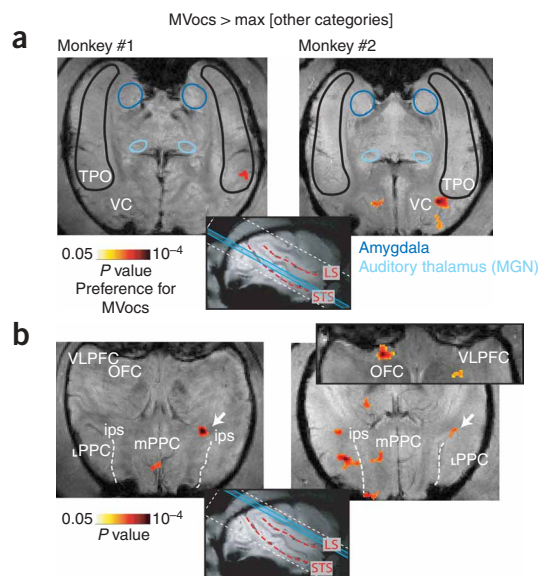
We confirmed our findings of two main regions in auditory cortex preferring the conspecific vocalizations by region of interest (ROI) analyses. These analyses used the 17 independently localized auditory cortical fields to summarize the 12 experiments with both animals (**Fig. 2**). The ROI analyses showed a significant preference for MVocs in the anterior region (localized to fields Ts1 and Ts2 in the right hemi-
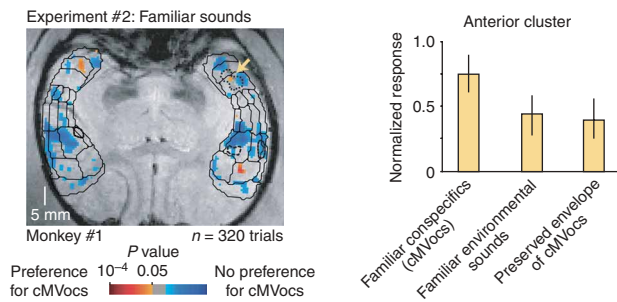
Figure 4 Brain clusters outside of auditory cortex that prefer MVocs. Shown are voxel clusters significantly preferring MVocs in the two awake monkeys (Monkeys 1 and 2), using the previously described criterion. Middle insets show the slice alignment and position. (**a**) We observed preferential MVocs activity in the lower slices in the temporal-parietal occipital area (TPO) on the upper bank of the STS, but only in one of the animals (posterior in the right hemisphere, Monkey 1). Monkey 2 showed clusters in higher visual cortical (VC) regions, which may have resulted from his less controlled eye movements (see Methods). (**b**) An upper slice is shown here that summarizes the observed clusters of MVocs preference around the intra-parietal sulcus (IPS) in both animals (regions of the medial and lateral PPC). Monkey 2 showed some orbito-frontal cortex (OFC) clusters in a higher slice (offset on top of image) and a cluster near the ventro-lateral prefrontal cortex (VLPFC). Outside of auditory cortex, only one region presented with colocalized clusters in both animals; see the arrows pointing to a PPC cluster.

sphere and to Ts2 alone in the left, all $P < 0.05$, **Fig. 2**). The central/posterior MVocs-preferring region was only significant in the right A1 and caudal parabelt fields (all $P < 0.01$, compare **Fig. 2** with **Fig. 1**).

The anterior MVocs-sensitive region that we described for the awake animals was the only one that was also identified in anesthetized animals using both voxel-based (example experiment, left panel, **Fig. 3**; see **Supplementary Fig. 3** online for additional examples) and ROI analyses (group analysis, right panel, **Fig. 3**). This region's preference for macaque vocalizations under anesthesia provides evidence that its preference is not the result of greater attention by the awake animals to the MVocs category of sounds. The group analyses of the anesthetized animal data did not show any other brain regions with a significant preference for MVocs (all $P > 0.1$). These observations indicate that the mapped anterior region that was observed in the seven awake and anesthetized animals is a good candidate for a monkey voice region; it is located in the hierarchically higher anterior regions of auditory cortex and has a strong preference for species-specific vocalizations, regardless of the vigilance of the animal.

These results suggest that a monkey voice region is located in the middle of the anterior STP. However, the human voice region, despite being also present in the anterior portion of the temporal lobe, is observed right below the STP[3,5,7], in the upper bank of the STS. We investigated the selectivity to MVocs beyond auditory cortex to see what other MVocs-preferring regions are present in the other parts of the monkey brain. The remaining clusters preferring MVocs in the awake animals are shown in **Figure 4**. Unlike the human reports, we did not find a reliable preference for species-specific vocalizations in the upper bank of the STS, although one animal had a cluster in posterior portions (**Fig. 4a**, see the outlined field in the temporal-parietal occipital area). Regions preferring MVocs that were consistent across experiments and animals were located only in the posterior-parietal cortex (PPC; green arrow in **Fig. 4b**). Anatomical landmark-based ROI analyses across experiments corroborated these observations, showing no significant MVocs-preferring voxels in the STS, nor in other identified regions of the amygdala and auditory thalamus (**Fig. 4**). The only regions outside of the STP that met this criterion were again in the PPC, and, only with this analysis, in the ventro-lateral prefrontal cortex and in higher visual cortical areas (all $P < 0.001$, **Fig. 4**).

Experiment #2: Familiar sounds

Monkey #1   n = 320 trials

Anterior cluster

Preference for cMVocs P value 10^{-4} 0.05   No preference for cMVocs

Figure 5 Auditory cortex regions preferring familiar conspecific vocalizations (Experiment 2). Left, combined and coregistered data from multiple experiments with Monkey 1. The format of this figure is similar to that of **Figure 1a**. Here, the preference for cMVocs was evaluated as cMVocs > max [ESnds, PeMVocs]. To facilitate comparisons of the results of Experiment 1 with this animal, the black dashed lines outline the anterior and central/posterior clusters that preferred MVocs in Experiment 1 (**Fig. 1a**). Right, using the clusters of voxels outlined by the dashed lines resulted in only the anterior region (orange arrow) having a significant preference for cMVocs. This result is shown in the bar graph as the voxel-based normalized response for the outlined anterior cluster ($n = 18$ voxels). The cMVocs response is significantly higher than the other conditions (bars show the mean and the 5 and 95% confidence intervals, bootstrap procedure).

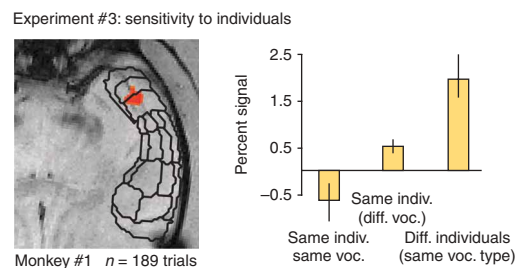## Sensitivity to the vocalizations of familiar conspecifics

The prior experiments used a large number of unfamiliar sounds to evaluate which brain regions could be important for recognizing a sound as a vocalization from a member of the species, even if the vocalizing individual is unknown. To determine whether these or other regions might be involved in evaluating the vocalizations from familiar conspecifics and the acoustical identity of the callers, we conducted additional experiments with Monkey 1.

For Experiment 2, we evaluated which brain regions showed a preference for conspecific vocalizations relative to a set of familiar sounds. We recorded the vocalizations of conspecifics from the animal's colony for use as stimuli (cMVocs, see Methods and **Supplementary Fig. 2**). The two comparison sound categories were environmental sounds that the animal would be familiar with (ESnds, including animal-generated, non-vocal sounds) and an acoustical control that preserved the envelope and duration of the cMVocs (PeMVocs). Using these sounds as stimulation, we analyzed 320 trials from the behaving animal (**Supplementary Fig. 2**). We then evaluated the data for evidence of an activity preference for cMVocs, using the criterion that cMVocs elicit significantly greater activity than the maximum response to the other two categories. We observed clusters of voxels that significantly preferred cMVocs in the previously noted anterior and central/posterior regions (all $P < 0.05$, **Fig. 5**). We compared these results with the monkey's results from Experiment 1 (the black dashed lines in **Fig. 5** outline the voxels from Experiment 1, see **Fig. 1a**, that significantly preferred MVocs, all $P < 0.05$). The comparison revealed that the central/posterior auditory cluster, which was previously localized to A1, was located outside of A1, 5 mm away (see the more posterior red cluster in **Fig. 5**). Thus, ROIs defined by any of the central/posterior clusters from Experiment 1 showed no significant preference for cMVocs in Experiment 2 (all $P > 0.1$, **Fig. 5**), suggesting that the earlier regions of auditory cortex are sensitive to the specific acoustical features of the sound categories used in the different experiments. The anterior region, on the other hand, maintained a significant preference for cMVocs even when using an ROI as defined from Experiment 1 ($P < 0.05$, **Fig. 5**). Although the preference for cMVocs of this anterior region seemed to be more prominent in the left

hemisphere, all tested hemispheric differences in the monkeys were subtle at best (**Supplementary Fig. 4** online).

Having shown that comparable auditory regions (including at least the anterior region) prefer familiar or unfamiliar species-specific vocalizations over comparison sound categories, in Experiment 3 we used an fMRI adaptation procedure to evaluate which brain regions are sensitive to the identity of the individuals that elicited the vocalizations. For stimulation we used only the familiar conspecific vocalizations (the same cMVocs used in Experiment 2; two 'coos' and two 'grunts' each from three conspecific individuals). In each scanning trial, we then repeated three sounds, varying only how the sounds were repeated (**Supplementary Fig. 2**). The repetition conditions were (i) repeating the same vocalization from one individual, (ii) repeating the different vocalizations from one of the individuals, and (iii) repeating a similar vocalization from different individuals. We reasoned that neuronal populations insensitive to a repetition condition would respond minimally to that condition, whereas those that maintained a strong response could be considered to be sensitive for that condition[7,37].

To objectively evaluate a sensitivity for the third condition (repeating a similar vocalization from different individuals), we split the data obtained for Experiment 3 and conducted a test-retest analysis (**Fig. 6**). We used the first part of the data to identify the brain clusters that showed an activity preference for the third condition. These clusters were then retested with the second part of the data, and only one cluster maintained a significant preference for the third condition (**Fig. 6**). This cluster was in the previously noted anterior auditory region, in areas Ts1 and Ts2. The cluster shows a signal drop below baseline when we repeated the same vocalization (first condition; **Fig. 6**). There was a higher signal when we repeated different vocalizations from the same individual (second condition), indicative of some sensitivity to the acoustical differences between the coos and grunts. But the condition with the highest sensitivity (signal, see **Fig. 6**) was the third, where we repeated a similar vocalization (coo or grunt) from different individuals (the third condition was greater than both the first, $P < 0.001$, and second, $P = 0.005$, conditions; paired samples t-tests, $n = 13$ voxels). This observation reveals that the anterior, higher-level auditory region is sensitive to the vocal acoustics that distinguish the identity of different individuals, more so than the rather large acoustical differences between the coo and grunt calls[16,38] uttered by the same speaker.

Experiment #3: sensitivity to individuals

Monkey #1   n = 189 trials

Figure 6 Sensitivity in the anterior STP region to the identity of the individuals producing the vocalizations (Experiment 3). Left, only one cluster in auditory cortex showed a test-retest sensitivity for the identity of the vocalizing individuals. This cluster was identified with the first part of the dataset (the test). Right, retest sensitivity of this cluster in the second part of the data (as signal magnitude, percentage from the silent baseline) for each of the three repetition conditions. Left bar, repeating the same vocalization (first condition); middle bar, repeating different vocalization types (coos and grunts) from the same individual (second condition); right bar, repeating a similar vocalization type (coo or grunt) as produced by three different individuals (third condition). All bars show the cluster's mean and the standard error for each condition ($n = 13$ voxels).

DISCUSSION

The vocal expressions of social animals carry considerable meaning for all members of the species and often elicit immediate behavioral reactions to dangerous or socially important events, such as the presence of a predator or the movement of the group. This is also true for human vocalizations, although humans, unlike other animals, can greatly enhance their vocal expressions linguistically. Our high-resolution fMRI experiments in the rhesus macaque provide evidence for a voice region in the monkey auditory cortex and establish functional relationships with the described human voice region.

Evidence for a voice region in monkeys

The first set of experiments with the two awake macaque monkeys (Experiment 1) identified several brain regions that showed a clear preference for species-specific (macaque) vocalizations. Here, a preference for species-specific vocalizations was defined as significantly increased activity to a large set of macaque vocalizations (many call types from many unfamiliar callers) in relation to the maximal activity response to several sound level–matched categories. Our approach is similar to others that have been used to identify the human voice region, where the response to many human vocalizations (from unfamiliar speakers) is evaluated in relation to multiple sound categories containing natural sounds and acoustical controls[3–5]. In the awake animals, only three brain regions strongly preferred the species-specific vocalizations. We observed two regions in auditory cortex and a third region surrounding the intra-parietal sulcus of the PPC. Of the two auditory regions, the anterior was localized to a hierarchically high-level processing region in auditory cortex, in the anterior STP. The central or more posterior region was well within the first few auditory cortical-processing stages, which included the primary field, A1 (refs. 34–36).

The same experiment (Experiment 1) with five anesthetized animals only revealed the anterior region that was observed with the awake animals, constraining the interpretation of the other more posterior auditory region. Possible interpretations of the more posterior (central/posterior) auditory region and its preference for macaque vocalizations are that its preference reflects either what electrophysiological recordings have described as diffuse clusters of strong neuronal responses to species-specific vocalizations in the earlier auditory regions[18,20,21] (that would have been subdued by anesthesia), or top-down influences in the awake animals. In all cases, these experiments emphasized the role of the anterior auditory region, and the results allow us to conclude that the monkey brain has an anterior auditory region that prefers the acoustical features that distinguish species-specific vocalizations from other categories of sounds, a region that is sensitive to the voice of the species. Similar interpretations are made of the results obtained from imaging the human voice region, which have revealed its preference for species-specific human vocalizations over other sound categories[3–5,22]. These voice regions in monkeys and humans could serve an important role in a general vocal-recognition ability: namely, to recognize that a sound is a vocalization from a member of the species, even if the vocalization is from an unfamiliar individual.

Additional experiments with one of the awake animals again emphasized the role of the anterior auditory region. In Experiment 2, we confirmed that the preference of the anterior region for species-specific vocalizations extends to vocalizations from familiar conspecifics. Then in Experiment 3, we used only the category of familiar conspecific vocalizations as stimuli to establish another functional relationship to the human voice region[7]. The result here was that only the anterior monkey voice region was sensitive for the vocal identity of the conspecific monkeys that elicited the vocalizations.

This experiment showed that the anterior auditory region is sensitive to the identity of the callers, which supports the notion that this region serves another vocal recognition function: namely, to recognize the voices of individuals.

Altogether our findings provide multiple sources of evidence that a monkey voice region exists with functional properties that are comparable to those that have been described for the known human voice region. In both primate species, these voice regions show preferential activation by species-specific vocalizations[3–6], are sensitive to the identity of the individuals that elicited the vocalizations[7], and belong to higher-level auditory processing regions[3,5,6,34,36] (although the mapping of the human auditory cortical fields has been difficult[33,39]; also see ref. 26). We next consider our findings in a broader context.

Enhanced MVocs processing in the auditory what pathway

We are surprised that the anterior voice region in monkeys was not identified by the previous monkey imaging studies using species-specific sounds. There is a rough correspondence with a monkey PET imaging study that proposed that the anterior STP is important for species-specific vocalizations[23]. That study suggested that an anterior temporal-lobe region tended to unilaterally represent the species-specific vocalizations, with stronger responses to these vocalizations in the left hemisphere[23]. The other PET imaging reports[24,25] were unable to replicate or elaborate on this finding[40], emphasizing either that multiple regions throughout the monkey brain preferred monkey coos and screams over carefully controlled nonbiological sounds[24], or proposing that two regions in particular were monkey homologs of the classical language territories in the human brain[25]. In the more recent report, the authors proposed that a ventral premotor region that preferred monkey vocalizations is the monkey homolog of Broca's territory, whereas a posterior high-level auditory area, Tpt, is the homolog of Wernicke's territory[25]. We also imaged these regions, but observed no strong preference for species-specific vocalizations in many of the previously highlighted brain regions[24,25], except for the PPC. We believe that our results help to reconcile and clarify the previous monkey imaging studies[23–25]. We show that the strongest and most reliable specificity for species-specific vocalizations is in an anterior auditory STP region. The other more posterior and less reliable auditory region that we localized closer to the primary auditory cortex may have been what the most recent PET study attributed to field Tpt[25]. It could be that our imaging resolution, design, stimulation and analyses allowed for a more selective description of areas that preferentially process species-specific calls. Yet our findings are easier to interpret, including in the context of the auditory processing pathway to which the presumed monkey voice region would belong.

In both humans and monkeys, the ventral or anterior auditory processing pathway that would include our anterior region is thought to be a 'what' pathway for sound identification[13,20,28–31,41,42]. This underscores the suitability of the anterior auditory region as a monkey voice region, where the region's observed preference for species-specific vocalizations in the what pathway is a likely substrate for vocal recognition. Certainly, no brain region functions alone. Thus, the interpretation that the monkey voice region belongs to the auditory what pathway suggests that this region relies on a network of other regions for analyzing the acoustics of sounds that give rise to its particular sensitivity to species-specific vocalizations. It is also possible that what we term a voice region (human or monkey) involves multiple anatomical areas and/or different neuronal populations. It may be that the sensitivities of these primate voice regions are the results of different closely localized populations of neurons that cannot be resolved with fMRI. Alternatively, a common neuronal mechanism might be able to

explain the results, such as stronger responses to species-specific vocalizations combined with tuning for the vocalizations of individuals. Electrophysiological recordings from the monkey voice areas can evaluate these alternatives by describing how the observed fMRI sensitivities of the voice region arise at the neuronal level.

Is it possible that the other more posterior auditory region and the PPC, both of which showed some preference for species-specific vocalizations, belong to a different pathway? The identified central/posterior auditory and PPC regions anatomically belong to the early part of the dorsal 'where' pathway[20,28–31,41] (the second of the two pathways proposed by the dual-pathway hypothesis[28]). However, we are careful not to over-interpret these results or the relationship of these regions to the where pathway for the following reasons. First, unlike the anterior auditory region, the more posterior region showed a less reliable preference for species-specific vocalizations and its location was less stable, perhaps because its acoustical sensitivity was tied to the particular sound categories used in each experiment. Second, the dorsal auditory where pathway that the more posterior auditory and the PPC regions would belong to is thought to involve sound localization[20,28–31,41] and crossmodal registration, presumably to a spatial frame of reference[27,43–46]. Because we didn't vary the sound location of our stimuli, we cannot be certain that the central/posterior auditory and PPC regions are involved in the where pathway unless this is explicitly demonstrated.

Social animals and voice regions: evolutionary pressures

Social animals depend on members of their species for survival. To the extent that similar evolutionary pressures influenced the brain specialization of other social animals, species-specific specializations similar to the described primate voice regions should exist in many more animal species. If the human and macaque voice regions are homologous, it is likely that many other primate species possess voice regions. Animals more distantly related to primates could have evolved voice regions as independent adaptations to support social abilities, and such analogous adaptations could differ considerably from those in primates. However, even in primates there appear to be differences, as the following observation illustrates.

We have thus far described the similarities in function between the human voice region and its presumed homolog in monkeys. However, a comparison of the anatomical location of the monkey voice region and that of the known human voice region suggests a discrepancy. The monkey voice region resides in the middle of the anterior portion of the STP, whereas the human voice region, although also located anteriorly, is in the upper bank of the STS, possibly also in the STG[3–6]. Given that the primary auditory cortex is similarly located in macaques, chimpanzees and humans on the STP[47], our comparison could mean that at least one of these higher-level voice regions in the two primates repositioned following the split of the lineages from their common ancestor. In all cases, the observation of similar function, yet different anatomical location, might not be surprising in light of a recent anatomical study that used the same comparative approach to parcellate many areas in the human and macaque STP and STG[48]. The study shows evidence that a number of human areas (from earlier to higher-level regions) are present in macaques, but seem to have differentiated into multiple areas in the human brain, which the authors suggest could relate to speech and language functions[48]. Also many areas appear to have, by comparison, expanded or diminished in one of the species, leaving a strong possibility that an area with comparable functional properties in two species could occupy different anatomical regions of the brain. Further comparative studies of structure and function using similar approaches in multiple species will be critical for advancing our understanding of the relationship between the many auditory areas in different species.

An animal model for human vocal recognition?

The classical face-recognition deficit called prosopagnosia is an excellent example of a longstanding collaboration between neurology and neuroscience. The neurologist Joachim Bodamer named the disorder and was the first to propose a neural substrate for it 60 years ago. This helped neuroscientists to identify human face-preferring regions in the ventral portions of the temporal lobe[9], including their presumed homologs in monkeys, which are the target of electrophysiological studies aiming to describe their function at the neuronal level[8,11,49].

Although less well known, a voice-recognition disorder called phonagnosia[1,2] also seems to have inspired the discovery of the human voice-preferring region[3]. Building on many studies, our results now propose that there is a close functional correspondence between the human and monkey voice regions. Although cross-species brain homologies require considerable evidence to establish[14,50], it could well be that an animal model can now be developed for neuronal studies of the vocal-recognition functions that were evolutionarily conserved in different primate species, including humans.

METHODS

Animals. We obtained fMRI data from seven male macaque monkeys (*Macaca mulatta*), weighing 6–12 kg, from a group-housed male colony. Two monkeys were trained to complete trials of a sparse-imaging/stimulation sequence with visual fixation and minimal body movement (**Supplementary Fig. 2**)[26,27]. The behaving animal procedure was modified for Monkey 2 because the monkey was not able to complete full fixation trials even after extensive training. He was allowed to fixate for shorter durations while maintaining the required minimal body movement during the trials. Another five monkeys were scanned while anesthetized[26,27,43]. All studies were in full compliance with the guidelines of the European Community for the care and use of laboratory animals (EUVD/86/609/EEC) and were approved by the local authorities (Regierungspraesidium Tuebingen).

Magnetic resonance imaging. Measurements of the fMRI blood oxygen level–dependant signal were made on vertical 7- or 4.7-T scanners (Bruker Medical). The monkeys sat in a primate chair in the magnet. During the experiment, the monkey's head was positioned with a plastic head holder (Tecapeek, Ensinger). Signals were acquired using 150-mm diameter radiofrequency saddle coils, or with a 70-mm diameter surface coil in some anesthetized animal experiments.

For the behaving animals, functional data were acquired using a multi-shot (two segments) gradient-recalled echo planar imaging sequence with the following typical parameters: echo time, TE: 22 ms; sparse imaging: volume acquisition time, 2 s; volume repetition time, TR: 10 s; flip angle, 45°; 9–12 slices, 2-mm thick and centered on auditory cortex; field of view, 9.6 × 8.0 cm, on a grid of 96 × 80 voxels, with voxel resolution of 1 × 1 × 2 mm³. Anatomical images were acquired in register with each functional scanning experiment using a FLASH sequence with parameters TE: 12 ms, TR: 750 ms, flip angle: 45°, 192 × 160 voxels. The anesthetized animal imaging paradigm and the typical parameters used are described elsewhere[26,27,43].

Sound stimulation. Experiment 1 used four categories of sounds (33 sounds per category): (i) species-specific, rhesus macaque vocalizations consisting of screams (33%), coos (18%), grunts (15%), barks (12%), pant threats (9%), girney/geckers (6%) and harmonic arches (6%) (MVocs), (ii) heterospecific animal vocalizations from domestic and wild animals, including other primate species (AVocs), (iii) natural sounds such as insects, thunder, rain and water (NSnds), and (iv) preserved frequency spectrum (scrambled phase in the Fourier domain) and duration of the MVocs (PsMVocs).

Experiment 2 used three categories of sounds (12 sounds per category): (i) vocalizations recorded from familiar conspecifics in the colony of the imaged animal, two coo and two grunt vocalizations[38] each from three individuals (all males; ages: 6, 6 and 7 years; weighing: 6, 6 and 13 kg,

respectively) (cMVocs), (ii) environmental sounds that the animal was accustomed to hearing, such as cage rattling, doors closing, rain and music (ESnds), (iii) acoustical controls for the envelope and duration of cMVocs that were created by extracting the envelope of the cMVocs and using it to shape pink noise ($1/f$ amplitude spectrum; PeMVocs). Experiment 3 used only the 12 cMVocs (**Supplementary Fig. 2**).

During an imaging trial, in between volume acquisitions (sparse imaging[26]), a category was randomly selected (Experiments 1 and 2 only) and sounds from that category were each randomly presented with an interstimulus interval of 300 ms. For Experiment 3, three sounds were repeated with a stimulus onset asynchrony of 1,500 ms (**Supplementary Fig. 2**). All sounds were resampled at 22,050 Hz, root-mean-square equated in amplitude and presented through magnetic-resonance compatible headphones (MR Confon) at ~80 dB sound pressure level (for additional details see **Supplementary Fig. 2**, **Supplementary Methods** online and refs. 26,27).

Functional and anatomical parcellation of auditory cortex. Numerous auditory core and belt fields (that is, primary auditory cortex and surrounding belt regions) were functionally localized. We used a functional imaging technique to segregate 11 core and belt fields using topographical gradients of sound frequency in the antero-posterior direction and sound bandwidth in the medio-lateral direction[26,27]. Surrounding these 11 core and belt fields we defined 6 additional regions, relying on the outer borders of the functionally localized fields and the local anatomical landmarks of the remaining regions on the STP and STG[34-36].

Data analysis. We coregistered and concatenated the datasets across experiments with each animal. Then the multislice data (volumes) were converted into voxel time points and linear drifts were removed. Functional maps of auditory activity were computed using cross-correlation with a boxcar model (P values were Bonferroni corrected; **Fig. 1c**). For the voxel-based analyses comparing different categories, we computed the macaque vocalization preference by regressing a biphasic model with voxels' responses to MVocs or cMVocs and their maximal response to the other sound categories (evaluation criterion: Experiment 1, MVocs > max [AVocs, NSnds, PsMVocs], Experiment 2, cMVocs > max [ESnds, PeMVocs]). The P values from this analysis were used to make statistical maps, with the direction of the associated regression B value determining the color code and the preference. ROI analyses for Experiment 1 were conducted independently of the voxel-based analyses as follows. For each experiment, we computed each voxel's MVocs preference using a permutation test[26,27], yielding the significance of the MVocs preference versus the maximal activity to the other sounds. Then, two numbers were obtained from each experiment and each ROI: the number of voxels significantly preferring MVocs ($P < 0.01$) and the 'signal preference' (**Fig. 2**), defined as the difference between the mean activity of the voxels preferring MVocs ($P < 0.01$) and those preferring any of the other categories ($P < 0.01$). We evaluated whether each ROI distribution of values differed from 0 using a t-test uncorrected for multiple comparisons, which seemed appropriate given the rather conservative evaluation criterion (MVocs or cMVocs > max [other conditions]). For the awake animals, both the activity (signal preference) and voxel count findings yielded similar results. For the anesthetized animals, only the voxel count analysis revealed any significant MVocs preferring fields, probably because of the fewer experiments in the dataset ($n = 6$, right, **Fig. 3**). The remaining analyses are explained as they are presented.

Note: Supplementary information is available on the Nature Neuroscience website.

AUTHOR CONTRIBUTIONS
C.I.P. conceived and designed the experiments. C.I.P., C.K., T.S. and M.A. carried out the experiments. C.I.P. analyzed the data and wrote the paper. C.K. and N.K.L. contributed conceptually and with data analysis. C.I.P., C.K., T.S., K.W., M.A. and N.K.L. provided materials and analysis tools.

1. Van Lancker, D.R. & Canter, G.J. Impairment of voice and face recognition in patients with hemispheric damage. *Brain Cogn.* **1**, 185–195 (1982).
2. Van Lancker, D.R., Cummings, J.L., Kreiman, J. & Dobkin, B.H. Phonagnosia: a dissociation between familiar and unfamiliar voices. *Cortex* **24**, 195–209 (1988).
3. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. & Pike, B. Voice-selective areas in human auditory cortex. *Nature* **403**, 309–312 (2000).
4. Belin, P., Zatorre, R.J. & Ahad, P. Human temporal-lobe response to vocal sounds. *Brain Res. Cogn. Brain Res.* **13**, 17–26 (2002).
5. Fecteau, S., Armony, J.L., Joanette, Y. & Belin, P. Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* **23**, 840–848 (2004).
6. Altmann, C.F., Doehrmann, O. & Kaiser, J. Selectivity for animal vocalizations in the human auditory cortex. *Cereb Cortex* **17**, 2601–2608 (2007).
7. Belin, P. & Zatorre, R.J. Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* **14**, 2105–2109 (2003).
8. Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189 (1991).
9. Sergent, J., Ohta, S. & MacDonald, B. Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* **115**, 15–36 (1992).
10. Kanwisher, N., McDermott, J. & Chun, M.M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
11. Tsao, D.Y., Freiwald, W.A., Tootell, R.B. & Livingstone, M.S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
12. Uppenkamp, S., Johnsrude, I.S., Norris, D., Marslen-Wilson, W. & Patterson, R.D. Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* **31**, 1284–1296 (2006).
13. Obleser, J. *et al.* Vowel sound extraction in anterior superior temporal cortex. *Hum. Brain Mapp.* **27**, 562–571 (2006).
14. Hauser, M.D., Chomsky, N. & Fitch, W.T. The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).
15. Masataka, N. Development of vocal recognition of mothers in infant Japanese macaques. *Dev. Psychobiol.* **18**, 107–114 (1985).
16. Rendall, D., Owren, M.J. & Rodman, P.S. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *J. Acoust. Soc. Am.* **103**, 602–614 (1998).
17. Ghazanfar, A.A. *et al.* Vocal-tract resonances as indexical cues in rhesus monkeys. *Curr. Biol.* **17**, 425–430 (2007).
18. Winter, P. & Funkenstein, H.H. The effect of species-specific vocalization on the discharge of auditory cortical cells in the awake squirrel monkey. *(Saimiri sciureus).* *Exp. Brain Res.* **18**, 489–504 (1973).
19. Suga, N., Niwa, H., Taniguchi, I. & Margoliash, D. The personalized auditory cortex of the mustached bat: adaptation for echolocation. *J. Neurophysiol.* **58**, 643–654 (1987).
20. Tian, B., Reser, D., Durham, A., Kustov, A. & Rauschecker, J.P. Functional specialization in rhesus monkey auditory cortex. *Science* **292**, 290–293 (2001).
21. Wang, X. & Kadia, S.C. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J. Neurophysiol.* **86**, 2616–2620 (2001).
22. Belin, P. Voice processing in human and non-human primates. *Phil. Trans. R. Soc. Lond. B* **361**, 2091–2107 (2006).
23. Poremba, A. *et al.* Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature* **427**, 448–451 (2004).
24. Gil-da-Costa, R. *et al.* Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. *Proc. Natl. Acad. Sci. USA* **101**, 17516–17521 (2004).
25. Gil-da-Costa, R. *et al.* Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. *Nat. Neurosci.* **9**, 1064–1070 (2006).
26. Petkov, C.I., Kayser, C., Augath, M. & Logothetis, N.K. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol.* **4**, e215 (2006).
27. Kayser, C., Petkov, C.I., Augath, M. & Logothetis, N.K. Functional imaging reveals visual modulation of specific fields in auditory cortex. *J. Neurosci.* **27**, 1824–1835 (2007).
28. Rauschecker, J.P. Parallel processing in the auditory cortex of primates. *Audiol. Neurootol.* **3**, 86–103 (1998).
29. Romanski, L.M. *et al.* Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* **2**, 1131–1136 (1999).
30. Kaas, J.H. & Hackett, T.A. 'What' and 'where' processing in auditory cortex. *Nat. Neurosci.* **2**, 1045–1047 (1999).
31. Rauschecker, J.P. & Tian, B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. USA* **97**, 11800–11806 (2000).
32. Rauschecker, J.P., Tian, B. & Hauser, M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* **268**, 111–114 (1995).
33. Wessinger, C.M. *et al.* Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J. Cogn. Neurosci.* **13**, 1–7 (2001).
34. Pandya, D.N. & Sanides, F. Architectonic parcellation of the temporal operculum in rhesus monkey and its projection pattern. *Z. Anat. Entwicklungsgesch.* **139**, 127–161 (1973).
35. Hackett, T.A., Stepniewska, I. & Kaas, J.H. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J. Comp. Neurol.* **394**, 475–495 (1998).

36. Kaas, J.H. & Hackett, T.A. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA* **97**, 11793–11799 (2000).

37. Grill-Spector, K. & Malach, R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst.)* **107**, 293–321 (2001).

38. Hauser, M.D. & Marler, P. Food-associated calls in rhesus macaques (*Macaca mulatta*). I. Socioecological factors. *Behav. Ecol.* **4**, 194–205 (1993).

39. Formisano, E. *et al.* Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* **40**, 859–869 (2003).

40. Ghazanfar, A.A. & Miller, C.T. Language evolution: loquacious monkey brains? *Curr. Biol.* **16**, R879–R881 (2006).

41. Alain, C., Arnott, S.R., Hevenor, S., Graham, S. & Grady, C.L. "What" and "where" in the human auditory system. *Proc. Natl. Acad. Sci. USA* **98**, 12301–12306 (2001).

42. Scott, S.K., Blank, C.C., Rosen, S. & Wise, R.J. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* **123**, 2400–2406 (2000).

43. Kayser, C., Petkov, C.I., Augath, M. & Logothetis, N. Integration of touch and sound in auditory cortex. *Neuron* **48**, 373–384 (2005).

44. Ghazanfar, A.A., Maier, J.X., Hoffman, K.L. & Logothetis, N.K. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* **25**, 5004–5012 (2005).

45. Mullette-Gillman, O.A., Cohen, Y.E. & Groh, J.M. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J. Neurophysiol.* **94**, 2331–2352 (2005).

46. Schroeder, C.E. & Foxe, J. Multisensory contributions to low-level, 'unisensory' processing. *Curr. Opin. Neurobiol.* **15**, 454–458 (2005).

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 (2001).

48. Fullerton, B.C. & Pandya, D.N. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *J. Comp. Neurol.* **504**, 470–498 (2007).

49. Leopold, D.A., Bondar, I.V. & Giese, M.A. Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* **442**, 572–575 (2006).

50. Sereno, M.I. & Tootell, R.B. From monkeys to humans: what do we now know about brain homologies? *Curr. Opin. Neurobiol.* **15**, 135–144 (2005).