PNAS RESEARCH ARTICLE | NEUROSCIENCE

Evolutionary continuity and divergence of auditory dorsal and ventral pathways in primates revealed by ultra-high field diffusion MRI

Yang Zhang[a,1], Sherry Xinyi Shen[a], Adnan Bibic[b,c], and Xiaoqin Wang[a,1]

Edited by Jon Kaas, Vanderbilt University, Nashville, TN; received August 10, 2023; accepted January 22, 2024

Auditory dorsal and ventral pathways in the human brain play important roles in supporting speech and language processing. However, the evolutionary root of the dual auditory pathways in the primate brain is unclear. By parcellating the auditory cortex of marmosets (a New World monkey species), macaques (an Old World monkey species), and humans using the same individual-based analysis method and tracking the pathways from the auditory cortex based on multi-shell diffusion-weighted MRI (dMRI), homologous auditory dorsal and ventral fiber tracks were identified in these primate species. The ventral pathway was found to be well conserved in all three primate species analyzed but extend to more anterior temporal regions in humans. In contrast, the dorsal pathway showed a divergence between monkey and human brains. First, frontal regions in the human brain have stronger connections to the higher-level auditory regions than to the lower-level auditory regions along the dorsal pathway, while frontal regions in the monkey brain show opposite connection patterns along the dorsal pathway. Second, the left lateralization of the dorsal pathway is only found in humans. Moreover, the connectivity strength of the dorsal pathway in marmosets is more similar to that of humans than macaques. These results demonstrate the continuity and divergence of the dual auditory pathways in the primate brains along the evolutionary path, suggesting that the putative neural networks supporting human speech and language processing might have emerged early in primate evolution.

auditory dorsal and ventral pathways | evolution | primate brains | diffusion MRI

Significance

Speech and language are unique abilities of humans, how the neural pathways for processing speech and language evolve in primates remains largely unknown. We compared the auditory dorsal and ventral pathways in three primate species using multi-shell diffusion-weighted MRI. We identified homologous auditory dorsal and ventral fiber tracks in these three primate species. Further analyses showed several similarities and divergences in auditory dorsal and ventral pathways between humans and non-human primates. This study sheds light on the evolutionary root of the auditory dorsal and ventral pathways in primates and provides insights into the human-specific brain specialization for speech and language processing.

Studies on the evolution of speech and language networks are important to understand the emergence of human-specific specializations in the brain. In non-human primates, species-specific vocalizations are thought to be processed through two parallel processing streams, including a ventral "what" pathway and a dorsal "where" pathway (1–4). In humans, speech and language processing is also thought to be supported by a cortical dorsal–ventral dual stream architecture (5–9). This architecture involves two cortical pathways of information flow between the auditory cortex and the frontal cortex (5–9). The dorsal pathway, via the arcuate fasciculus (AF), connects to Wernicke's area (the speech comprehension region) in the temporal lobe, the angular gyrus in the parietal cortex, and Broca's area (the speech production region) in the frontal lobe. This fiber tract is considered the most important tract for speech and language processing in the human brain (10–15). The ventral pathway is less understood than the dorsal pathway. Two major fiber tracts have been found along the ventral pathway in the human brain that are related to speech and language processing (16–22): the inferior fronto-occipital fasciculus (IFOF) and the uncinate fasciculus (UF).

The use of MRI techniques in non-human primates studies is becoming increasingly common (23–25). The development of comparative MRI has enabled anatomical and functional comparisons of the auditory dual streams across humans and various non-human primate species, such as macaques and chimpanzees (7, 12–15). Originally, the absence of AF that connects posterior superior temporal to inferior frontal cortex found in macaques was considered as the reason for the lack of speech in monkeys (12). However, such a connection had been demonstrated later by Romanski et al. (4). In addition, the existence of AF homologs in macaques and chimpanzees that interconnects the posterior temporal areas and inferior frontal areas was confirmed (7, 14), and the weaker connections of AF homologs in macaques than those in human were considered as the reason for the lack of speech in monkeys, making it into a quantitative than a categorical difference. Furthermore, expansions of AF's temporal terminations to the middle temporal gyrus (MTG) were found in humans as compared with macaques and chimpanzees (7, 13). However, little is known about whether non-human primates have homologous ventral fiber tracts as humans. In order to make comparisons across primate species, homologous brain structures need to be defined. For example, the auditory cortex of non-human primates is divided into core,

Author affiliations: [a]Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205; [b]Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD 21205; and [c]Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, F. M. Kirby Center, Baltimore, MD 21205

Author contributions: Y.Z. and X.W. designed research; Y.Z., S.X.S., and A.B. performed research; Y.Z. analyzed data; and Y.Z. and X.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: yzhan142@jhu.edu or xiaoqin.wang@jhu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2313831121/-/DCSupplemental.

Published February 20, 2024.

belt, and parabelt sub-regions (26, 27). However, the parcellations of the human auditory cortex are quite different from those of non-human primates, which makes it difficult to directly compare the connectivity of these auditory cortical sub-regions along the dorsal–ventral dual streams between humans and non-human primates.

Other than chimpanzees and macaques (Old World monkeys), it is unclear whether the dorsal–ventral dual streams exist in New World monkeys, such as the common marmoset (Callithrix jacchus), which separated from Old World monkeys at about 40 Mya (28). The answer to this question is critical for pinpointing the evolutionary origin of the dorsal–ventral dual streams in the primate brain. The marmoset is a highly vocal and social primate species with a sophisticated vocal repertoire for vocal communication (29, 30). In recent years, marmosets have gained increasing interest in neuroscience and preclinical research (31–34). Understanding the extent to which marmosets have similar brain structures and connectivity as humans and other primate species is important for translational research.

In the current study, we compared probabilistic diffusion-weighted MRI (dMRI) tractography among marmosets, macaques, and humans. To establish homologous brain structures across these three species, we used an identical individual-based quantitative parcellation method based on MRI contrasts to parcellate the auditory cortex into comparable sub-regions for each individual subject. The connectivity strengths were compared for each sub-region along the dorsal–ventral dual streams across all three primate species. In addition, the connectivity strengths from regions beyond the auditory cortex (anterior temporal lobe and middle temporal regions) to the frontal cortex were also computed to explore the expansions of dorsal and ventral fiber tracks in the temporal lobe. Our analyses revealed homologous auditory dorsal and ventral fiber tracks in marmosets and macaques as in humans. In contrast to marmosets and macaques, the human dorsal pathway showed an expansion in the temporal lobe to lateral regions beyond the auditory cortex and the human ventral pathway showed an expansion in the temporal lobe to anterior regions beyond the auditory cortex. Additionally, the dorsal pathways in marmosets were found to be more similar to those in humans as compared with macaques. These results demonstrate the evolutionary continuity and divergence of auditory dual pathways in primates and illustrate the human-specific specializations for speech and language processing.

Results

Individual-Based Quantitative Parcellation of Auditory Cortex in Marmosets, Macaques, and Humans. We obtained whole-brain multi-shell dMRI data from nine ex vivo marmosets and three in vivo marmosets from our colony at Johns Hopkins University. Multi-shell dMRI datasets of five macaques and four awake humans were obtained from publically available online databases (Table 1). To visualize fiber tracks from the auditory cortex to the frontal cortex in marmosets, macaques, and humans, deterministic tractography was performed using DSI-Studio software (https://dsi-studio.labsolver.org/) with the auditory cortex as the region of interest for each individual subject (see *Methods* for details). The deterministic tractography reconstructs fiber tracks by applying an algorithm to link the voxels with similar diffusion directions (35), assuming that there is one main diffusion direction for each voxel. This technique is used primarily for visualization due to its high false-negative rates. Our exploratory analyses (*SI Appendix*, Figs. S1–S3) uncovered homologous auditory dorsal and ventral fiber tracks in marmosets, macaques, and humans.

Table 1. Summary of all tested subjects

	Subject ID	Gender	Age (year)	Imaging region	Data source	Note
Marmoset (ex vivo)	M16E	Female	2	Auditory cortex	Acquired in this study	
	M31W	Male	6	Whole brain	Acquired in this study	Array implantation before the animal was killed
	MF19	Female	6	Whole brain	Acquired in this study	
	M8E	Male	4	Whole brain	Acquired in this study	
	M101D	Male	5	Whole brain	Acquired in this study	
	M126D	Female	5	Whole brain	Acquired in this study	
	M11X	Male	12	Whole brain	Acquired in this study	
	MF52	Female	4	Whole brain	Acquired in this study	
	MF53	Female	7	Whole brain	Acquired in this study	
	MF33	Male	6	Whole brain	Acquired in this study	
Marmoset (in vivo)	M9E	Male	5	Whole brain	Acquired in this study	
	M40E	Male	5	Whole brain	Acquired in this study	
	M133E	Male	5	Whole brain	Acquired in this study	
Macaque	Subject 1	Male	4	Whole brain	From Ambrosen et al. (36)	https://www.drcmr.dk/rhesus-macaque-brain
	Subject 2	Male	7 to 8	Whole brain	From PRIMatE Data Exchange	https://fcon_1000.projects.nitrc.org/indi/PRIME/amu.html
	Subject 3	Male	7 to 8	Whole brain	From PRIMatE Data Exchange	https://fcon_1000.projects.nitrc.org/indi/PRIME/amu.html
	Subject 4	Male	7 to 8	Whole brain	From PRIMatE Data Exchange	https://fcon_1000.projects.nitrc.org/indi/PRIME/amu.html
	Subject 5	Female	7 to 8	Whole brain	From PRIMatE Data Exchange	https://fcon_1000.projects.nitrc.org/indi/PRIME/amu.html
Human	Subject 1	Male	26 to 30	Whole brain	From WU-Minn HCP data	https://db.humanconnectome.org/app/template/Index.vm
	Subject 2	Female	26 to 30	Whole brain	From WU-Minn HCP data	https://db.humanconnectome.org/app/template/Index.vm
	Subject 3	Female	26 to 30	Whole brain	From WU-Minn HCP data	https://db.humanconnectome.org/app/template/Index.vm
	Subject 4	Female	31 to 35	Whole brain	From WU-Minn HCP data	https://db.humanconnectome.org/app/template/Index.vm

To quantitatively compare fiber connectivity strengths along the auditory dorsal and ventral pathways across species, we sought to parcellate the auditory cortex of individual subjects using the same quantitative parcellation method. Five MRI contrasts (see *Methods* for details) were analyzed for each subject. 1) T2 (T1 in humans) refers to signal amplitudes of the T2w (T1w in humans) MRI images; 2) mean diffusivity (MD) measures the average mobility of water molecules with higher MD values indicating more space between axons; 3) fractional anisotropy (FA) measures the degree of anisotropy that reflects fiber density and myelination; 4) orientation dispersion index (odi) measures the neurite density and neurite orientation dispersion of brain microstructures; 5) directionally encoded color map (DEC) represents fiber orientation distribution weighted principal fiber direction information.

These MRI contrasts were used to parcellate the auditory cortex individually. First, we calculated columnar anisotropy by local probabilistic fiber tracking to find the best fiber route from each voxel on the GM (gray matter) surface of the auditory cortex to the GM–WM (white matter) boundary with the highest connectivity probability. These fiber routes were defined as cortical columns (*SI Appendix*, Fig. S4, *Upper Left*). Second, the GM of the auditory cortex was divided into an upper layer part and a lower layer part based on the cortical columns (*SI Appendix*, Fig. S4, *Upper Right*). Finally, the five MRI contrasts averaged across the upper layer and the lower layer were used to construct the feature space (10 features for each cortical column, *SI Appendix*, Fig. S4, *Lower*). Unsupervised distance-based clustering (K-means clustering) was used to group the cortical columns. Fig. 1 *A*, *Left* showed a confusion matrix from a representative ex vivo marmoset (M16E) in which four clusters were observed. To determine the optimal number of clusters, we calculated the WCSS (within-cluster sum of square) which is the sum of the squared distance between each point and the centroid in a cluster. The WCSS values across cluster numbers were plotted

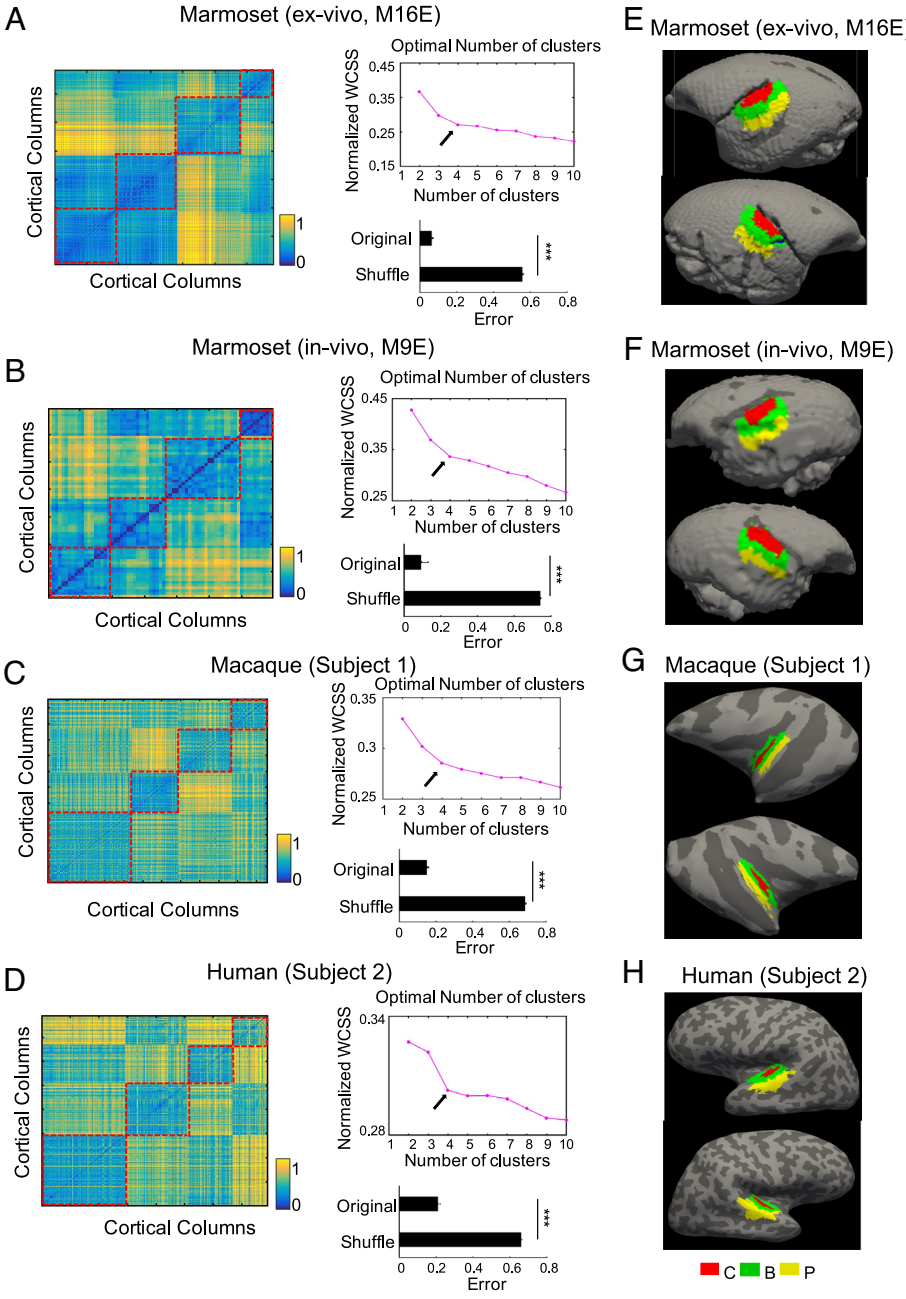

Fig. 1. Individual-based parcellation of auditory cortex in marmosets, macaques, and humans. (*A–D*) *Left*: Classification confusion matrices from a representative ex vivo marmoset (*A*), a representative in vivo marmoset (*B*), a representative macaque (*C*), a representative human (*D*); *Upper Right*: The normalized WCSS (within-cluster sum of square) values across cluster numbers; *Lower Right*: Comparisons between 10-fold cross-validations and shuffled 10-fold cross-validations (***$P < 0.001$; WRS-test). (*E–H*) Mapping the parcellation results onto the individual-based 3D brain surface for a representative ex vivo marmoset (*E*), a representative in vivo marmoset (*F*), a representative macaque (*G*), a representative human (*H*).

(Fig. 1 *A*, *Upper Right* showed the WCSS curve of M16E) in which the optimal cluster number is four. The demonstration that the multi-contrast MRI could be used to identify and segment different cortical areas was originally proposed in a previous study (37).

In addition to ex vivo marmosets, the auditory cortices of in vivo marmosets (Fig. 1*B* for a representative in vivo marmoset), macaques (Fig. 1*C* for a representative macaque), and humans (Fig. 1*D* for a representative human) were all parcellated using the same method. In vivo marmosets remained awake in a customized restraint tube (*SI Appendix*, Fig. S5) during the MRI scans. The confusion matrices and the WCSS curves all indicated that the optimal cluster number is four in in vivo marmoset (Fig. 1 *B*, *Left and Upper Right*), macaque (Fig. 1 *C*, *Left and Upper Right*), and human (Fig. 1 *D*, *Left and Upper Right*). All parcellation results were mapped onto the individual-based 3D brain surface for visualization (Fig. 1*E* for a representative ex vivo marmoset; Fig. 1*F* for a representative in vivo marmoset; Fig. 1*G* for a representative macaque; Fig. 1*H* for a representative human), the core (C), belt (B), and parabelt (P) sub-regions were labeled based on the anatomic locations (lateral belt and medial belt were merged into belt).

We next sought to compare the MRI contrasts across the auditory sub-regions in the three primate species based on the parcellation results. There were marked differences in the five MRI contrast measures among auditory core, belt, and parabelt regions as shown in Fig. 2 *A*–*D*. Auditory core regions across the species showed low T2 values in marmosets and macaques (high T1 values in humans), low MD values, high FA values, and low odi values. In contrast, auditory belt regions showed high T2 values in marmosets and macaques (low T1 values in humans), high MD values, low FA values, and high odi values. In auditory parabelt regions, the upper layers showed low T2 values in marmosets and macaques (high T1 values in humans), high FA values, and low odi values, whereas the lower layers showed low FA values and high odi values. The differences in the MRI contrasts among auditory core, belt, and parabelt regions are similar across the three primate species (Fig. 2 *A*–*D*). *SI Appendix*, Fig. S6 shows different fiber orientations among core (C), belt (LB: lateral belt; MB: medial belt) and parabelt (P) regions of the three species. Similar fiber orientation patterns were found for each auditory sub-region in all three primate species (*SI Appendix*, Fig. S6). The above findings of the auditory sub-regions are consistent with the prior knowledge on these cortical regions: 1) Auditory core regions have the highest myelination structures and the predominant connections with the auditory thalamus (26, 27, 38, 39), 2) auditory belt regions have diverse connections to both core and parabelt regions, and 3) auditory parabelt regions have the predominant connections to the frontal cortex (26, 40–42).

One advantage of this parcellation method is the invariance to brain distortions. *SI Appendix*, Fig. S7 showed the parcellation results from an example ex vivo marmoset in which brain distortion was seen on the right hemisphere due to electrode array implantation (*SI Appendix*, Fig. S7*A*). The confusion matrix (*SI Appendix*, Fig. S7 *B*, *Left*) and the WCSS curve (*SI Appendix*, Fig. S7 *B*, *Upper Right*) both showed that four clusters are optimal. MRI contrasts (*SI Appendix*, Fig. S7*E*) and fiber orientations (*SI Appendix*, Fig. S7*F*) showed similar patterns as normal ex vivo and in vivo marmosets (Fig. 2 and *SI Appendix*, Fig. S6). These results suggest auditory cortices of marmosets, macaques, and humans share similar microstructures and can be parcellated into comparable core, belt (including lateral belt and medial belt), and parabelt regions. The

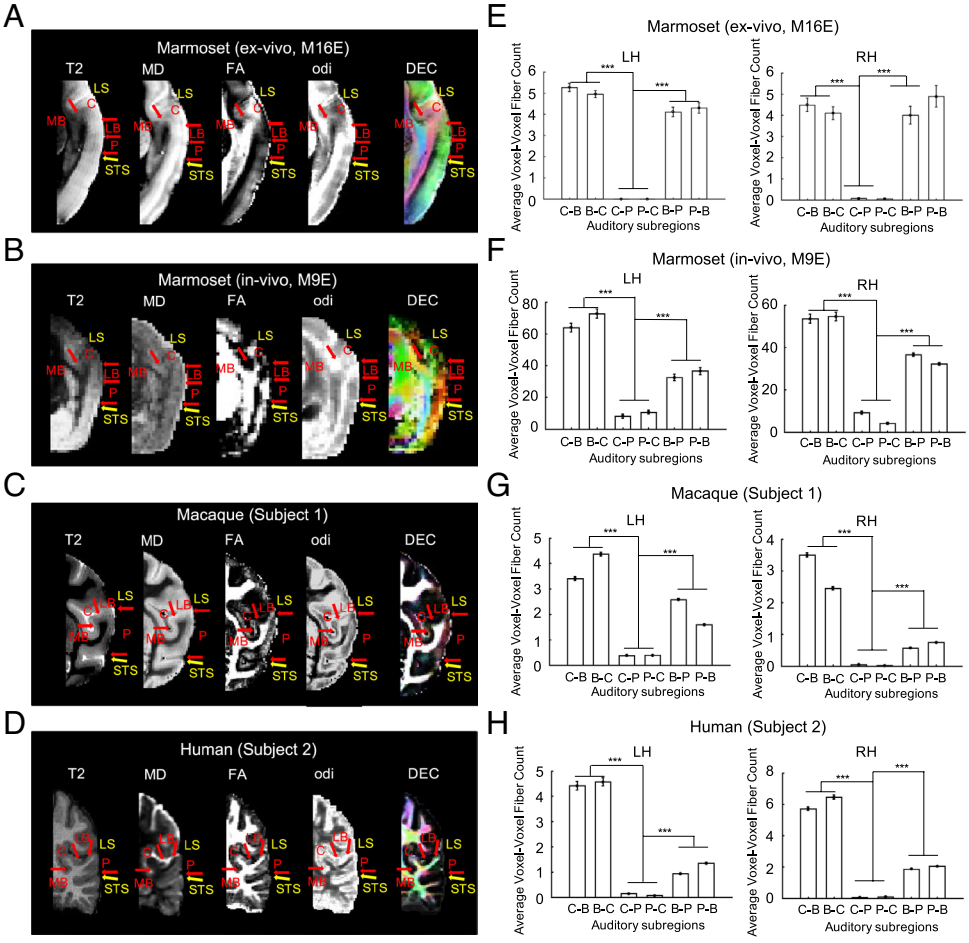

Fig. 2. MRI contracts and fiber connectivity among auditory sub-regions in marmosets, macaques, and humans. (*A*–*D*) Five MRI contrasts (T2/T1, MD, FA, odi, and DEC) in a representative ex vivo marmoset (*A*), a representative in vivo marmoset (*B*), a representative macaque (*C*), and a representative human (*D*). (*E*–*H*) Connectivity between the parcellated auditory sub-regions (C: core; B: belt; P: parabelt) in both hemispheres for a representative ex vivo marmoset (*E*), a representative in vivo marmoset (*F*), a representative macaque (*G*), a representative human (*H*) (***$P < 0.001$; WRS-test).

invariance of the parcellation method to brain distortions highlights the importance of remapping the anatomical brain connectivity following the surgical implant of chambers for electrophysiological recordings or optical imaging windows.

Validation of the Parcellation Results. The parcellation results were validated in three aspects. First, a 10-fold cross-validation was performed to check the robustness of the parcellation results. A support vector machine (SVM) was used to train the model and compute the misclassification rate. The original error rate was computed by averaging the misclassification rates. In addition, to compute a null distribution of the misclassification rates, 10-fold cross-validations were performed on the same dataset by shuffling the corresponding labels. The distribution of shuffled error rates was computed by shuffling the labels 1,000 times. The original error rates were significantly lower than the shuffled error rates (Wilcoxon rank-sum test: WRS-test, $P < 0.001$) in ex vivo marmoset (Fig. 1 *A*, *Lower Right*), in vivo marmoset (Fig. 1 *B*, *Lower Right*), macaque (Fig. 1 *C*, *Lower Right*), and human (Fig. 1 *D*, *Lower Right*).

Second, based on the prior knowledge that auditory belt regions have reciprocal connections with auditory core and parabelt regions whereas auditory core and parabelt regions have fewer connections (27, 40), we sought to compare the connectivity between the parcellated sub-regions. Probabilistic fiber tracking was performed for each subject. Voxel-voxel fiber count was defined as the connected fiber number between a given voxel in one sub-region and a given voxel in another sub-region. Reciprocal connections between core and belt regions were found to be significantly higher than the reciprocal connections between core and parabelt regions, and reciprocal connections between belt and parabelt regions were also significantly higher than the reciprocal connections between core and parabelt regions (WRS-test, $P < 0.001$). These connectivity patterns were seen for both hemispheres in ex vivo marmoset (Fig. 2*E*), in vivo marmoset (Fig. 2*F*), macaque (Fig. 2*G*), and human (Fig. 2*H*), suggesting that consistent connectivity patterns for all three primate species were evident between our parcellations and the results from previous studies.

Finally, our parcellation results were compared with existing brain atlases. For marmosets, the Nencki-Monash (NM) template (43) and the NIH-MBA v1.1 (NIH) template (37) were registered to the T2w images of marmosets using affine transformations. The inconsistency rate was calculated as the percentage of inconsistent voxels across the total number of voxels between our parcellation results and the atlases. The registration results for one of the ex vivo marmosets (M31W) are shown in *SI Appendix*, Fig. S8. In general, the left hemisphere showed consistent parcellation results using our method and the two atlases, whereas parcellations using our method are much better than those of the two atlases in the right hemisphere due to the distortion on this side (*SI Appendix*, Fig. S8*A*). Inconsistency rates between our parcellations and the two atlases were similar to the benchmark rate measured between the two atlases in the left hemisphere but were higher than the benchmark rate in the right hemisphere (*SI Appendix*, Fig. S8*B*). Additionally, connectivity across the parcellated sub-regions showed similar patterns between our parcellations and the two atlases in the left hemisphere, whereas significant differences were observed in the right hemisphere in which the connections between belt and parabelt regions were not significantly different from the connections between core and parabelt regions (*SI Appendix*, Fig. S8*C*). For macaque, the D99 macaque template (44) was registered to the T2w images of macaque. As shown in *SI Appendix*, Fig. S9, similar parcellation results were observed using our method and the atlas.

Together, these validation results suggest that the parcellations using MRI contrasts are robust and valid for all three primate species tested. In addition, our parcellations showed better performance for distorted brains compared with the atlases, which indicates that our method takes individual differences into consideration. So it is vital to redo individual-based parcellations rather than rely on applying templates for brains with distortions. Furthermore, the brain distortions may also cause everlasting topographic changes to the cortex and affect local connectivity.

Connectivity of Auditory Cortex along Dorsal and Ventral Pathways in Marmosets, Macaques, and Humans. After parcellating the auditory cortex into core, belt, and parabelt sub-regions, we sought to compare the connectivity strengths of these sub-regions along the dorsal and ventral pathways. Whole-brain probabilistic fiber trackings were performed with the auditory sub-regions used as seed regions. Dorsal pathway, ventral pathway, and connections to the parietal cortex were observed in ex vivo marmosets (Fig. 3*A* and *SI Appendix*, Fig. S10), in vivo marmosets (Fig. 3*B* and *SI Appendix*, Fig. S11), macaques (Fig. 3*C* and *SI Appendix*, Fig. S11), and humans (Fig. 3*D* and *SI Appendix*, Fig. S11). To quantify the connectivity strength, we defined the dorsal, ventral, and parietal waypoints as target ROIs (regions of interest) based on the probabilistic tracking results in each subject. Fiber count was defined as the connected fiber number between the seed region and a given voxel in the target region and was averaged across all voxels in the target region.

For the dorsal pathway, in ex vivo marmosets, connectivity from core and belt to dorsal target ROIs was significantly stronger than connectivity from parabelt (Fig. 3*E*; paired *t* test across subjects, $P < 0.05$). In in vivo marmosets, connectivity from belt to dorsal target ROIs was significantly stronger than connectivity from core and parabelt (Fig. 3*F*; paired *t* test across subjects, $P < 0.05$). In macaques, connectivity from core and belt to dorsal target ROIs was significantly stronger than connectivity from parabelt (Fig. 3*G*; paired *t* test across subjects, $P < 0.05$). In humans, connectivity from parabelt to dorsal target ROIs was significantly stronger than connectivity from core and belt (Fig. 3*H*; paired *t* test across subjects, $P < 0.05$). In addition, all connections to the dorsal target ROIs were found from ipsilateral auditory sub-regions (i.e., left auditory sub-regions connect to left dorsal target ROIs, while right auditory sub-regions connect to right dorsal target ROIs). The results, which show that the strongest connectivity to dorsal target ROIs was from belt in ex vivo marmosets, in vivo marmosets, and macaques, while the strongest connectivity was from parabelt in humans, suggest the connectivity along the dorsal pathway extended to more lateral regions of the auditory cortex (higher-level auditory cortex) during primate evolution.

For the ventral pathway, connectivity from parabelt to ventral target ROIs was significantly stronger than connectivity from core and belt in ex vivo marmosets (Fig. 3*I*; paired *t* test across subjects, $P < 0.05$), in vivo marmosets (Fig. 3*J*; paired *t* test across subjects, $P < 0.05$), macaques (Fig. 3*K*; paired *t* test across subjects, $P < 0.05$), and humans (Fig. 3*L*; paired *t* test across subjects, $P < 0.05$), suggesting the connectivity along the ventral pathway is conserved during the evolution in primates. Additionally, all connections to the ventral target ROIs were found from ipsilateral auditory sub-regions as well.

For the connections to the parietal cortex, similar connectivity patterns as the dorsal pathways were found in ex vivo marmosets (Fig. 3*M*; paired *t* test across subjects, $P < 0.05$), in vivo marmosets (Fig. 3*N*; paired *t* test across subjects, $P < 0.05$), and humans (Fig. 3*P*; paired *t* test across subjects, $P < 0.05$). In macaques, connectivity from belt and parabelt to parietal target ROIs was significantly stronger than connectivity from core (Fig. 3*O*; paired *t* test across subjects, $P < 0.05$). Overall, these results suggest the connections to the parietal cortex also extend to more lateral regions of the auditory cortex during the evolution of primates.

Moreover, by comparing the connectivity strengths among marmosets, macaques, and humans along the dorsal and ventral

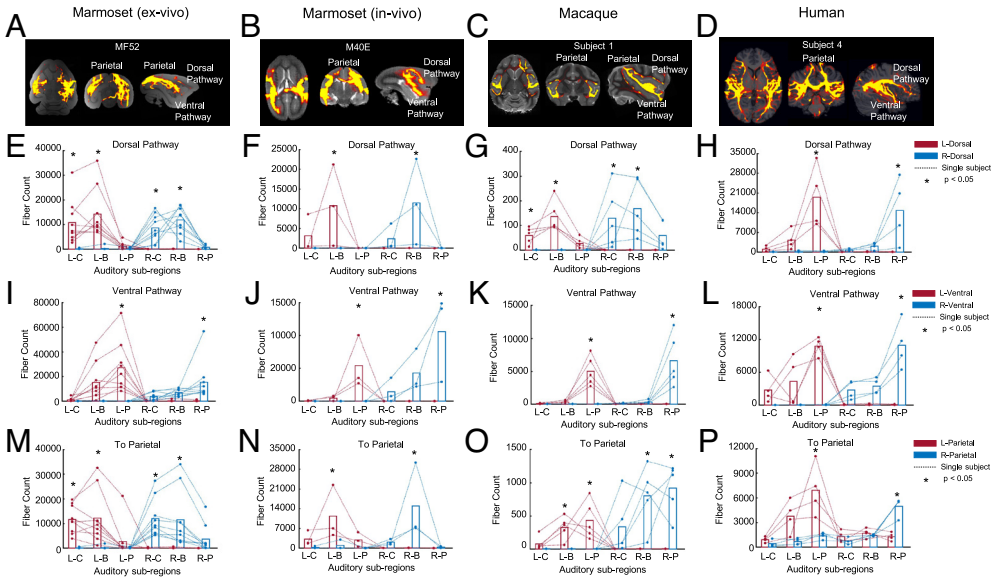

Fig. 3. Probabilistic tractography of auditory dorsal and ventral pathways in marmosets, macaques, and humans. (*A*–*D*) Probabilistic tractography showing the dorsal pathway, ventral pathway, and connections to the parietal cortex in a representative ex vivo marmoset (*A*, MF52), a representative in vivo marmoset (*B*, M40E), a representative macaque (*C*, Subject 1), and a representative human (*D*, Subject 4). (*E*–*H*) Connectivity strength from the parcellated auditory sub-regions to frontal cortex through the dorsal pathway in ex vivo marmosets (*E*), in vivo marmosets (*F*), macaques (*G*), and humans (*H*). (*I*–*L*) Connectivity strength from the parcellated auditory sub-regions to frontal cortex through the ventral pathway in ex vivo marmosets (*I*), in vivo marmosets (*J*), macaques (*K*), and humans (*L*). (*M*–*P*) Connectivity strength from the parcellated auditory sub-regions to the parietal cortex in ex vivo marmosets (*M*), in vivo marmosets (*N*), macaques (*O*), and humans (*P*). L-C: left core regions; L-B: left belt regions; L-P: left parabelt regions; R-C: right core regions; R-B: right belt regions; R-P: right parabelt regions; L-dorsal: left dorsal target ROIs; R-dorsal: right dorsal target ROIs; L-ventral: left ventral target ROIs; R-ventral: right ventral target ROIs; L-parietal: left parietal target ROIs; R-parietal: right parietal target ROIs; dot lines represent each subject; *$P < 0.05$, paired *t* test across subjects.

pathways (Fig. 3), we found that connectivity strengths in macaques along the dorsal pathway (Fig. 3*G*) and to the parietal cortex (Fig. 3*O*) are much lower than those of marmosets and humans (~100 times lower, Fig. 3 *E*, *F*, *H*, *M*, *N*, and *P*), whereas the connectivity strengths of marmosets along the dorsal pathway and to the parietal cortex are more similar to those in humans. In addition, the connectivity strengths along the ventral pathway are similar across all three primate species (Fig. 3 *I*–*L*). These results suggest that compared with macaques, the auditory dorsal pathways of marmosets are more similar to those of humans.

To further explore the connectivity patterns of the auditory cortex along dorsal and ventral pathways, the auditory cortex was divided into a rostral part and a caudal part in ex vivo marmosets (Fig. 4*A* for a representative ex vivo marmoset), in vivo marmosets (Fig. 4*B* for a representative in vivo marmoset), macaques (Fig. 4*C* for a representative macaque), and humans (Fig. 4*D* for a representative human). Connectivity strengths along dorsal and ventral pathways were then compared between the rostral part and the caudal part. For the dorsal pathway, connectivity from the caudal part was significantly stronger than connectivity from the rostral part in ex vivo marmosets (Fig. 4*E*; paired *t* test across subjects, $P < 0.05$), in vivo marmosets (Fig. 4*F*; paired *t* test across subjects, $P < 0.05$), macaques (Fig. 4*G*; paired *t* test across subjects, $P < 0.05$), and humans (Fig. 4*H*; paired *t* test across subjects, $P < 0.05$). For the ventral pathway, opposite connectivity patterns were observed as compared with the dorsal pathway, that connectivity from the rostral part was significantly stronger than connectivity from the caudal part in ex vivo marmosets (Fig. 4*I*; paired *t* test across subjects, $P < 0.05$), in vivo marmosets (Fig. 4*J*; paired *t* test across subjects, $P < 0.05$), macaques (Fig. 4*K*; paired *t* test across subjects, $P < 0.05$), and humans (Fig. 4*L*; paired *t* test across subjects, $P < 0.05$). While for the connections to the parietal cortex, similar connectivity patterns were observed as the dorsal pathway, that connectivity from the caudal part was significantly stronger than connectivity from the rostral part in ex vivo marmosets (Fig. 4*M*; paired *t* test across subjects, $P < 0.05$), in vivo marmosets (Fig. 4*N*; paired *t* test across subjects, $P < 0.05$), macaques (Fig. 4*O*; paired *t* test across subjects, $P < 0.05$), and humans (Fig. 4*P*; paired *t* test across subjects, $P < 0.05$). These results suggest that the dorsal pathway and the connections to the parietal cortex originate from the caudal part of the auditory cortex, while the ventral pathway originates from the rostral part of the auditory cortex, indicating the existence of two separate fiber pathways from the auditory cortex in the three primate species.

Comparisons across Species for Human-Specific Specializations. To further demonstrate the human-specific specializations of the dorsal and ventral pathways, we made comparisons across the three species in four aspects. First, the relative sizes of core, belt, and parabelt regions were compared across species (Fig. 5*A*). In ex vivo and in vivo marmosets, belt regions occupied significantly larger areas than core and parabelt regions in both hemispheres (paired *t* test across subjects, $P < 0.05$). In macaques and humans, parabelt regions occupied significantly larger areas than core and belt regions in both hemispheres (paired *t* test across subjects, $P < 0.05$). These results demonstrated the expansion of higher-level auditory regions (parabelt regions) from marmosets, to macaques, and then humans.

Second, the connectivity strengths along the dorsal and ventral pathways were compared between hemispheres in the three primate species (Fig. 5*B*). For connectivity along the dorsal pathway, no significant differences were observed between hemispheres in ex vivo marmosets, in vivo marmosets, and macaques (paired *t* test across subjects, $P > 0.05$). While in humans, connectivity in the left hemisphere was significantly stronger than that in the right hemisphere (paired *t* test across subjects, $P < 0.05$; Fig. 5 *B*, *Left*). For connectivity along the ventral pathway, no significant differences were observed between hemispheres in ex vivo marmosets, in vivo marmosets, macaques, and humans (paired *t* test across subjects, $P > 0.05$; Fig. 5 *B*, *Middle*). Similarly, for connectivity

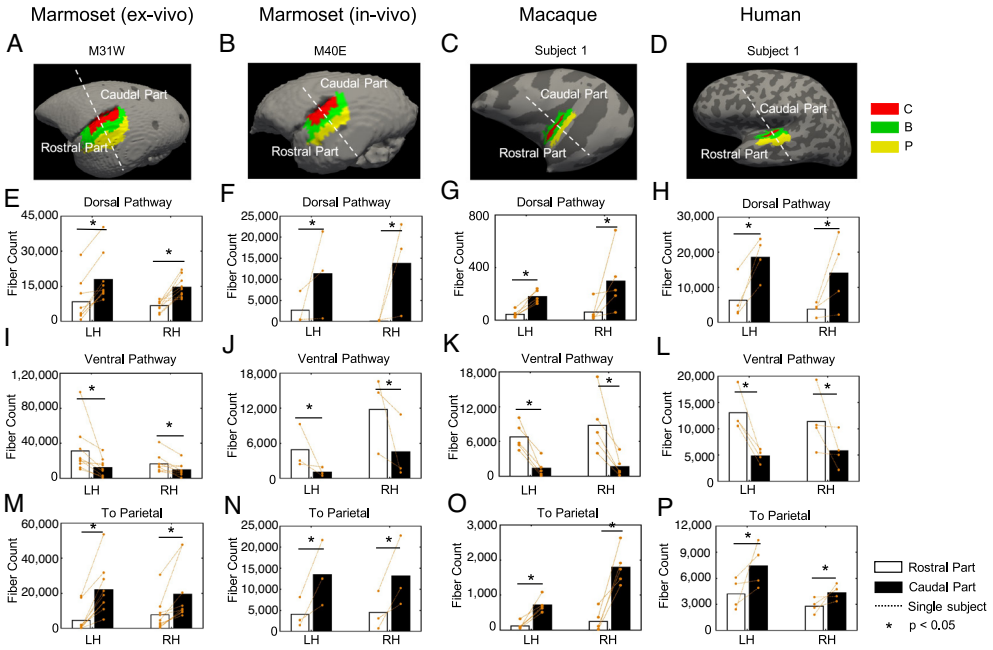

Fig. 4. Differences of connectivity for rostral and caudal parts of the auditory cortex. (*A–D*) Demonstrations of dividing the auditory cortex into a rostral part and a caudal part in a representative ex vivo marmoset (*A*, M31W), a representative in vivo marmoset (*B*, M40E), a representative macaque (*C*, Subject 1), and a representative human (*D*, Subject 1). (*E–H*) Connectivity strength from the rostral and caudal auditory cortex to frontal cortex through the dorsal pathway in both hemispheres of ex vivo marmosets (*E*), in vivo marmosets (*F*), macaques (*G*), and humans (*H*). (*I–L*) Connectivity strength from the rostral and caudal auditory cortex to frontal cortex through the ventral pathway in both hemispheres of ex vivo marmosets (*I*), in vivo marmosets (*J*), macaques (*K*), and humans (*L*). (*M–P*) Connectivity strength from the rostral and caudal auditory cortex to the parietal cortex in both hemispheres of ex vivo marmosets (*M*), in vivo marmosets (*N*), macaques (*O*), and humans (*P*). Dots represent each subject. *$P < 0.05$, paired *t* test across subjects.

to the parietal cortex, no significant differences were observed between hemispheres in ex vivo marmosets, in vivo marmosets, and macaques (paired *t* test across subjects, $P > 0.05$; Fig. 5 *B*, *Right*), while in humans, the connectivity in the left hemisphere was slightly stronger than that in the right hemisphere (paired *t* test across subjects, $P < 0.05$; Fig. 5 *B*, *Right*). These results suggest that the dorsal pathway is left dominated in humans but not in marmosets and macaques, while the ventral pathway is equally distributed in both hemispheres of all three primate species.

Third, based on the results that the connectivity along the dorsal pathway extends to more lateral regions of the auditory cortex in humans as compared with marmosets and the macaque (Fig. 3 *E–H*), we sought to compare the connectivity of the middle temporal region (MTR): including the superior temporal sulcus and middle temporal gyrus) with those of the parcellated auditory sub-regions along the dorsal and ventral pathways in humans. MTR is lateral to the parabelt region and is known for processing lexical and syntactic information in human language (5, 45–48). The MTR ROI was defined based on the anatomic location of each human subject (Fig. 5*C*). For the dorsal pathway, comparable connectivity strength was observed between parabelt and MTR in both hemispheres (Fig. 5 *D*, *Left*). For the ventral pathway, the connectivity from MTR was the lowest and was significantly lower than that from the parabelt region in both hemispheres (Fig. 5 *D*, *Middle*; paired *t* test across subjects, $P < 0.05$). Similarly, for the connections to the parietal cortex, the connectivity from MTR was the lowest and was also significantly lower than that from the parabelt region in both hemispheres (Fig. 5 *D*, *Right*; paired *t* test across subjects, $P < 0.05$). These results suggest that the termination areas of the dorsal pathway in the temporal cortex extend to MTRs in humans, while no corresponding areas to MTRs are found in marmosets and macaques. These extensions are not observed in the ventral pathway and the connections to the parietal cortex in humans, which are similar to marmosets and macaques.

Finally, the connectivity strength was compared between the anterior temporal lobe (ATL) and the parcellated auditory sub-regions along the dorsal and ventral pathways in all three primate species. ATL is a hub for semantic processing of speech and language in humans (49–51), moreover, vocalization-specific areas were found in ATL in marmosets (52) and macaques (53). The ATL ROI was defined based on the anatomic location of each subject (Fig. 6 *A–D*). For the dorsal pathway, the connectivity from ATL was the lowest and significantly lower than that from core and belt regions in ex vivo marmosets (Fig. 6*E*; paired *t* test across subjects, $P < 0.05$), was the lowest and significantly lower than that from belt region in in vivo marmosets (Fig. 6*F*; paired *t* test across subjects, $P < 0.05$) and macaques (Fig. 6*G*; paired *t* test across subjects, $P < 0.05$), and was significantly lower than that from the parabelt region in humans (Fig. 6*H*; paired *t* test across subjects, $P < 0.05$). For the ventral pathway, the connectivity from ATL was the highest in all species (Fig. 6 *I–L*) with significant differences observed between ATL with core and belt regions in ex vivo (Fig. 6*I*; paired *t* test across subjects, $P < 0.05$) and in vivo marmosets (Fig. 6*J*; paired *t* test across subjects, $P < 0.05$), and with significant differences observed between ATL and all parcellated auditory sub-regions (core, belt, and parabelt) in macaques (Fig. 6*K*; paired *t* test across subjects, $P < 0.05$) and humans (Fig. 6*L*; paired *t* test across subjects, $P < 0.05$). Similar to the dorsal pathway, in ex vivo marmosets, the connectivity to the parietal cortex from ATL was the lowest and significantly lower than that from core and belt regions (Fig. 6*M*; paired *t* test across subjects, $P < 0.05$); in in vivo marmosets, the connectivity to the parietal cortex from ATL was the lowest and significantly lower than that from belt region (Fig. 6*N*; paired *t* test across subjects, $P < 0.05$); in macaques, the connectivity to the parietal cortex from ATL was significantly lower than that from belt and parabelt regions (Fig. 6*O*; paired *t* test across subjects, $P < 0.05$); and in humans, the connectivity to the parietal cortex from ATL was significantly lower than that from the parabelt region (Fig. 6*P*; paired *t* test across subjects, $P < 0.05$). These results

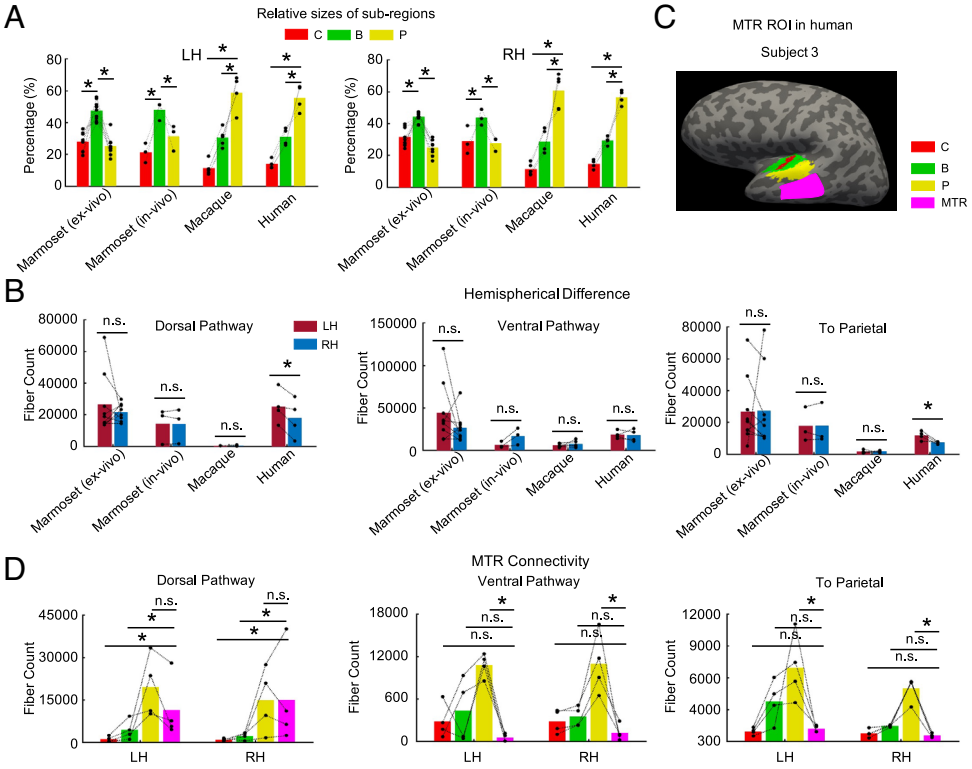

Fig. 5. Comparisons of auditory dorsal and ventral pathways in marmosets, macaques, and humans. (*A*) Comparisons of the relative sizes of core, belt, and parabelt regions in both hemispheres across all three primate species. (*B*) Comparisons of connectivity strength between hemispheres for dorsal pathway (*Left*), ventral pathway (*Middle*), and connections to the parietal cortex (*Right*) across all three primate species (*$P < 0.05$; n.s.: not significant; paired *t* test across subjects). (*C*) Demonstration of the location of MTR (middle temporal region) ROI in the human brain (human subject 3). (*D*) Comparisons of connectivity strength between MTR and the parcellated auditory sub-regions through the dorsal pathway (*Left*), the ventral pathway (*Middle*), and to the parietal cortex (*Right*) in both hemispheres of humans. Dots represent each subject. *$P < 0.05$; n.s.: not significant; paired *t* test across subjects.

demonstrated that ATL connects to the frontal cortex through the auditory ventral pathway in all three primate species, suggesting the ATL connections to frontal cortex were highly conserved in primates. In addition, the connectivity strengths from ATL through the ventral pathway are significantly stronger than those from parabelt in macaques and humans but not in marmosets, suggesting connectivity patterns of the ventral pathway in macaques are more like those in humans than those in marmosets, indicating that macaques have a more similar ventral pathway as humans than marmosets. Furthermore, connectivity strengths from ATL through the ventral pathway in humans are much stronger than those from the parcellated auditory sub-regions (four times stronger in the left hemisphere and three times stronger in the right hemisphere. Fig. 6*L*). This is a sharp contrast compared with marmosets and macaques, suggesting the termination areas of the auditory ventral pathway in the temporal cortex extend to more anterior regions which are beyond the auditory cortex in humans as compared with marmosets and macaques.

Discussion

Comparative analysis can help elucidate what is unique in humans and what is conserved across species. In the current study, we provided evidence to suggest evolutionary continuity in the auditory dorsal and ventral pathways between a New World monkey, an Old World monkey and human. Our results showed several similarities among the three analyzed primate species: 1) Similar MRI contrasts of the auditory cortex were found in marmosets, macaques, and humans, which can be used for quantitative parcellation (Fig. 2). 2) Homologous white matter tracts for auditory dorsal and ventral pathways were found in marmosets, macaques, and humans (*SI Appendix*, Figs. S1–S3). 3) In all three primate species, the frontal cortex has stronger connections to the caudal part of the auditory cortex through the dorsal pathway and stronger connections to the rostral part of the auditory cortex through the ventral pathway (Fig. 4). 4) The anterior temporal lobe connects to the frontal cortex through the ventral pathway in all three primate species (Fig. 6).

Our results also showed divergence in the auditory dorsal and ventral pathways among the three analyzed primate species: 1) Compared with marmosets and macaques, higher-level auditory regions in humans (parabelt) occupy a larger area of the auditory cortex (Fig. 5*A*). 2) In humans, the termination areas of the dorsal pathway on the temporal lobe extend to higher-level auditory regions (parabelt) and further to the middle temporal regions, whereas the termination areas of dorsal pathways in marmosets and macaques are restricted to lower-level auditory regions (core and belt; Figs. 3 and 5*D*). 3) Left lateralization of the auditory dorsal pathway was found in humans, but not in marmosets and macaques (Fig. 5*B*). 4) In humans, the termination areas of the ventral pathway in the temporal cortex extend to more anterior regions beyond the auditory cortex as compared with marmosets and macaques (Fig. 6).

Diffusion MRI calculates direct fiber pathways linking different brain regions, but it does not provide information about the directions or causalities of the connections. Modern effective connectivity is measured in a very different way. Modern effective connectivity measurements, including using a Hopf algorithm and fMRI with electric stimulation, are the ways to measure the causal connectivity between different brain areas with directional information (54–56). In humans, results of modern effective connectivity showed the connections to the inferior frontal gyrus (IFG) are from the higher-level auditory cortex beyond the parabelt regions (56). Our results that the dorsal pathway in humans extends to regions lateral to parabelt regions and the ventral pathway in humans extends to the regions anterior to parabelt regions are consistent with the previous findings using effective connectivity, suggesting the diffusion MRI results can provide indications for the effective connectivity of the auditory dorsal and ventral pathways.

Language is one of the most distinctive attributes of humans. The findings of homologous white matter tracts for auditory dorsal and ventral pathways in marmosets, macaques, and humans provide the anatomical basis for flexibility and plasticity in vocal production. Previous studies have shown that marmosets and macaques can alter the amplitude and other aspects of their vocalizations in

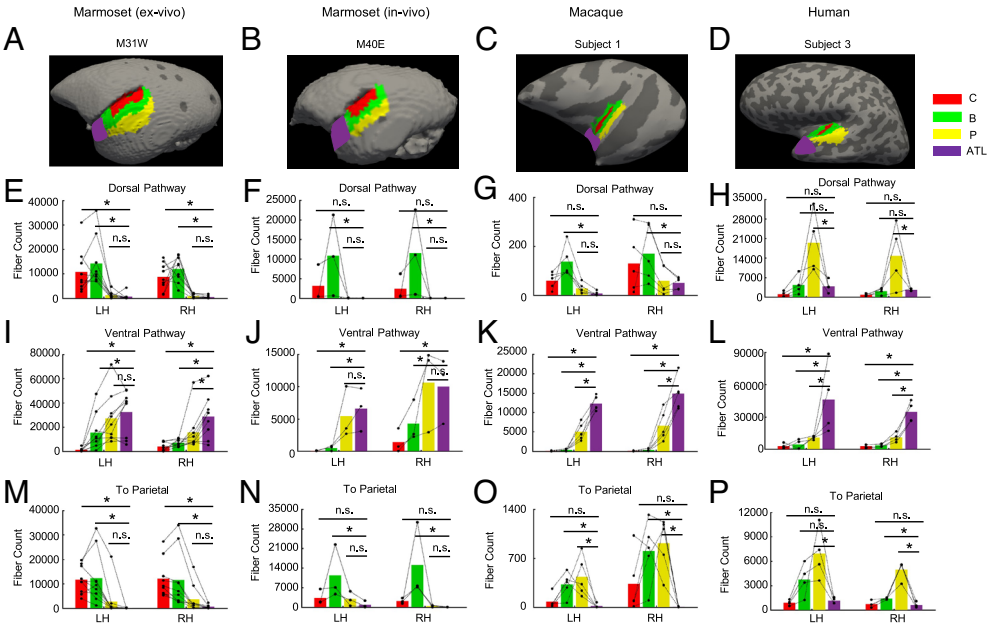

Fig. 6. Comparisons of ATL connectivity in marmosets, macaques, and humans. (*A–D*) Demonstration of the locations of ATL (anterior temporal lobe) ROIs in a representative ex vivo marmoset (*A*, M31W), a representative in vivo marmoset (*B*, M40E), a representative macaque (*C*, Subject 1), and a representative human (*D*, Subject 3). (*E–H*) Comparisons of connectivity strength between ATL and the parcellated auditory sub-regions through the dorsal pathway in both hemispheres of ex vivo marmosets (*E*), in vivo marmosets (*F*), macaques (*G*), and humans (*H*). (*I–L*) Comparisons of connectivity strength between ATL and the parcellated auditory sub-regions through the ventral pathway in both hemispheres of ex vivo marmosets (*I*), in vivo marmosets (*J*), macaques (*K*), and humans (*L*). (*M–P*) Comparisons of connectivity strength between ATL and the parcellated auditory sub-regions to the parietal cortex in both hemispheres of ex vivo marmosets (*M*), in vivo marmosets (*N*), macaques (*O*), and humans (*P*). Dots represent each subject.*$P < 0.05$; n.s.: not significant; paired *t* test across subjects.

the presence of background noises (57–59) (Lombard effect) or interfering noises (60, 61), suggesting the ability of vocal control in non-human primates. However, while non-human primate species including marmosets and macaques can communicate through species-specific vocalizations, they do not have the syntax-like abilities to combine symbols according to a defined set of rules to generate phrases with a variety of meanings (9, 62). The capacity of humans in language is likely supported by human-specific specializations in the brain. In the current study, first we found the expansion of higher-level auditory regions (parabelt) in humans compared with marmosets and macaques, suggesting more resources for higher-level auditory processing in humans. Previous studies have shown that compared with the primary sensory cortex, the association cortex has expanded greater in size and exhibits more diverse connectivity patterns in humans than in non-human primates, providing the anatomical basis for more complex cognitive and behavioral functions in humans (63–66). Second, consistent with the previous findings (7, 12, 14), we found that the terminations of the dorsal pathway in the temporal cortex are located in higher-level auditory regions which extend to the middle temporal region in humans, which has been shown to support lexical-semantic, syntactic, and prosodic processes (5, 45–48). The connections from the middle temporal region to the frontal cortex via the dorsal pathway might be important for carrying lexical-semantic and syntactic information during speech production since the dorsal pathway in humans is proposed for auditory-motor integration (5). Last, the terminations of the ventral pathway in the temporal cortex were found to extend to the anterior temporal lobe in humans, which is found to be involved in semantic and identity processing (49–51). The connections from the anterior temporal lobe to the frontal cortex via the ventral pathway are important for speech and language comprehension (5). These human-specific specializations described above may serve as the foundations for supporting the uniqueness of human speech and language.

Previous studies have demonstrated the left lateralization for speech and language processing in humans (5, 67, 68). In the current study, we found left lateralization of the auditory dorsal pathway but not the auditory ventral pathway in humans, which is consistent with the proposed dual-processing streams for speech processing (5). There are two models proposed to explain the origin of lateralization. One model proposes that the lateralization is due to differences in what kinds of information are processed between the different hemispheres, with the left hemisphere processing temporal information and the right hemisphere processing spectral information (69, 70). The other model proposes that lateralization derives from the differences in what information is stored in the different hemispheres, with the left hemisphere storing lexical information and the right hemisphere storing affective prosodic information (5). Moreover, no lateralization was observed in marmosets and macaques. A hypothesis has been proposed to explain the differences in brain asymmetry between humans and non-human primates. This hypothesis claims that brain size expansion during evolution increased the time for neuron signal transmission, resulting in decreased interhemispheric connectivity and increased intrahemispheric connectivity. Therefore, more cortical neurons with similar characteristics are located in one hemisphere than the other, which led to the asymmetries in brain functions (71–73). The lateralized frameworks in humans provide the increased computational capacity for processing speech and language and may play a facilitating role in the human cognitive and behavioral specializations.

Comparing marmosets with macaques, we found that the connectivity strengths of the dorsal pathways in marmosets are similar to those in humans, but much stronger than those in macaques, while the connectivity strengths of ventral pathways are similar among all three primate species. In humans, auditory dorsal stream supports an interface with the motor system, which suggests an auditory-motor integration function (5, 6). In non-human primates, the earliest proposals argued for a "where" function for the auditory dorsal pathway (1–3). More recently, people speculate that this auditory dorsal stream is involved in not just spatial processing in the classical sense but also sensorimotor control and vocal production (14, 74, 75). Previous studies have shown functional differences between marmosets and macaques in vocal communication and auditory perception. Macaque vocalization is generally spontaneous utterances related to its emotional states (76). It is difficult for macaques to generate vocalizations and suppress their spontaneous calls in a laboratory setting (77, 78). In sharp contrast, marmosets readily vocalize in laboratory settings and can take turns back and forth to exchange their vocalizations with conspecifics (antiphonal calling) just as in human conversations (79–82). A recent study showed that marmosets perceive their species-specific vocalizations in a manner similar to how humans perceive speech

(83). Furthermore, marmosets can voluntarily control the amplitude (58), timing (58, 80), duration (58), and pitch (84) of their vocalizations. Marmosets and macaques are also different in their pitch perception abilities. Humans have specialized cortical regions (pitch regions) with enhanced responses to harmonic tones versus noises to support pitch perception (85–87). In marmosets, a putative cortical pitch center was found located in the anterolateral low-frequency border of the primary auditory cortex (88). Moreover, human-like pitch perception also exists in marmosets (89). However, no regions with response preferences for harmonic tones versus noises were found in macaques (90). Furthermore, compared with macaques, expansion of the descending output from premotor areas was found in marmosets, which may in part explain the enhanced vocal motor skills of marmosets (91). Overall, marmosets demonstrate more similarities to humans than macaques in above mentioned aspects, although marmosets and macaques separated on the evolutionary tree at around 40 Mya (28), suggesting the dorsal–ventral dual streams have originated before the lineages of marmosets diverged, and these streams in marmosets continue to evolve in a direction that is closer to humans than to macaques after the marmoset lineage diverged.

Methods

Ten ex vivo marmosets, three in vivo marmosets, five macaques, and four humans (Table 1) were used in the current study. All procedures conformed to local and US NIH guidelines. All experiments were performed with the approval of the Johns Hopkins University Animal Use and Care Committee.

MRI Data Acquisitions.

Marmoset (ex vivo). Marmoset ex vivo MRI data were acquired on an 11.7T (Bruker Biospin) MRI platform running Paravision 6.0.1 with an 8-channel surface coil. The marmoset brains were immersed in a 1:200 dilution of a 0.5 mmol/mL gadolinium MR contrast agent (Gadoteridol, Bracco Diagnostics Inc.) mixed with phosphate buffer saline (PBS) solution for 7 d before imaging to reduce the T1 relaxation time (37, 92). Multi-modal MRI data were collected for each subject, including T2-weighted (T2w) images, and multi-shell diffusion MRI. The T2w images were acquired using a 3D FLASH sequence with the following parameters: TE = 3.5 ms, TR = 25 ms, flip angle = 20°, matrix size = 193 × 202 × 234, resolution = 0.15-mm isotropic, number of average = 4. The multi-shell diffusion MRI (dMRI) was collected with a 3D diffusion-weighted (DW) echo-planar imaging (EPI) sequence with the following parameters: TE = 25 ms, TR = 200 ms, flip angle = 90°, matrix size = 193 × 202 × 234, resolution = 0.15 mm isotropic, 3 shells (b = 1,000, 3,000, 5,000 s/mm^2) were acquired with 5 b = 0 images and 40 directions per shell.

Marmoset (in vivo). Marmoset in vivo MRI data were acquired on an 11.7 T (Bruker Biospin) MRI platform running Paravision 6.0.1 with a volume coil. The marmosets were trained to sit in a customized restraint tube (*SI Appendix*, Fig. S5) with heads fixed in the sphinx position. The training sessions lasted 2 wk. The marmosets remained awake during the MRI scans. Before scanning, 0.1 mL/kg cerenia (10 mg/mL maropitant antiemetic; Zoetis Petcare) was subcutaneously injected to the marmosets. During scanning, the bedding temperature was fixed at 30 to 35 °C. Multi-modal MRI data were collected from each subject, including T2w images, and multi-shell dMRI. The T2w images were acquired using a T2 TurboRARE sequence with the following parameters: TE = 21.4 ms, TR = 9,060 ms, flip angle = 90°, matrix size = 167 × 120, number of slice = 128, resolution = 0.3 mm × 0.3 mm × 0.3 mm, number of average = 5. The multi-shell dMRI was collected with a diffusion-weighted (DW) EPI sequence with the following parameters: TE = 16.2 ms, TR = 1,100 ms, flip angle = 65°, matrix size = 125 × 90, number of slice = 48, resolution = 0.4 mm × 0.4 mm × 0.8 mm, 2 shells (b = 500, 1,000 s/mm^2) were acquired with 5 b = 0 images and 42 directions per shell. All scans were acquired with right-to-left and left-to-right phase encoding polarities once, respectively.

Macaque. Macaque MRI data were downloaded from public datasets (36) (one is from https://www.drcmr.dk/rhesus-macaque-brain and the other four are from https://fcon_1000.projects.nitrc.org/indi/PRIME/amu.html). These are the macaque datasets we found with the highest resolution of multi-shell dMRI data. The one macaque brain was scanned on an experimental 4.7T Agilent scanner with a quadrature volume RF coil and a maximum gradient strength of 600 MT/m. The multi-shell dMRI was collected with a pulsed gradient spin-echo sequence with the following parameters: TE = 30 ms, matrix size = 128 × 256, number of slice = 91, resolution = 0.5 mm isotropic, 3 shells (b = 1,477, 4,102, 8,040 s/mm^2) were acquired by varying the gradient strength, G = [150, 250, 350] mT/m, TR = [8,000, 7,900, 8,600] ms. 3 b = 0 images and 180 directions per shell were acquired. The T2w images were acquired by averaging the b = 0 images. The other four macaques brains were scanned on a Siemens Prisma 3T scanner with a body transmit array and 11-cm loop receiving coil. The multi-shell dMRI was collected with a spin-echo EPI sequence with the following parameters: TE = 87.6 ms, TR = 7,520 ms, resolution = 1 mm isotropic, 2 shells (b = 300, 1,000 s/mm^2) were acquired with 5 b = 0 images. All scans were acquired twice with reversed phase encoding directions. The T2w images were acquired with a 3D SPACE sequence with a resolution of 0.8 mm isotropic.

Human. Four human subjects between 26 and 35 y old were randomly selected from WU-Minn Human Connectome Project (HCP) data (93) (https://db.humanconnectome.org/app/template/Index.vm). All subjects were scanned on a customized Siemens 3T "Connectome Skyra," using a standard 32-channel Siemens receive head coil and a "body" transmission coil designed by Siemens. T1 weighted (T1w) images were acquired using a 3D MRRAGE sequence with the following parameters: TE = 2.14 ms, TR = 2,400 ms, angle = 8°, resolution = 0.7 mm isotropic. The multi-shell dMRI was collected with a spin-echo EPI sequence with the following parameters: TE = 89.5 ms, TR = 5,520 ms, flip angle = 78°, matrix size = 168 × 144, number of slice = 111, resolution = 1.25 mm isotropic, 3 shells (b = 1,000, 2,000, 3,000 s/mm^2) were acquired with 6 b = 0 images and 90 directions per shell. All scans were acquired with right-to-left and left-to-right phase encoding polarities once, respectively.

MRI Data Analysis.

dMRI data preprocessing and MRI contrasts. The dMRI data of marmosets and macaque were denoised using the dwidenoise function implemented in MRtrix3 (94) and eddy current corrected using the eddy_correct function of FSL (95). The in vivo marmoset dMRI data were also preprocessed using DR-BUDDI of TORTOISE (96) to correct EPI distortion using two datasets with opposite phase encoding directions. The dMRI data of humans were already preprocessed with the HCP processing pipeline.

The preprocessed data were fitted with the diffusion tensor imaging (DTI) model using the DTIFIT of FSL, the multi-shell multi-tissue constrained spherical deconvolution (for reconstructing the fiber pattern of each voxel) by MRtrix3 (97), and the neurite orientation dispersion and density model (98) (NODDI) using NODDI MATLAB toolbox (https://mig.cs.ucl.ac.uk/index.php?n=Tutorial.NODDImatlab). These fittings generated multiple MRI contrasts including fractional anisotropy (FA), mean diffusivity (MD), fiber orientation distribution weighted principal fiber direction information (37) (directionally encoded color map, DEC), and orientation dispersion index (odi). These MRI contrasts provided alternative structural features that complemented the T1w and T2w images.

Deterministic fiber tracking. Deterministic fiber tracking was performed using DSI-Studio software (https://dsi-studio.labsolver.org/), which was used primarily for visualization purposes. The tractography was computed for each of the nerves with the auditory cortex used as the region of interest and the frontal cortex used as terminations. The analyses were performed with an angular threshold of 60°, an FA threshold of 0.1, a step size of 0.2 mm, and a fiber length of 0 to 300 mm.

Probabilistic fiber tracking. Probabilistic fiber tracking (99) was performed using FSL. Diffusion parameters were computed using BEDPOSTX with default settings (two fibers, weight 1). Fiber tracking was performed using PROBTRACKX with 5,000 samples, a curvature threshold of 0.2, a step length of 0.05 mm, and a subsidiary fiber volume fraction threshold of 0.01. The probabilistic fiber tracking provides a fiber count connectivity index map across all target regions for a given seed voxel or region.

Seed regions were selected based on the individual-based auditory cortex parcellation results (see below for details). The auditory cortex of all subjects was divided into three major sub-regions: core, belt, and parabelt. The anterior temporal lobe (ATL) was manually defined for all subjects, and the middle temporal region (MTR, including the superior temporal sulcus and middle temporal gyrus) was manually defined for all humans. The probabilistic fiber tracking was performed separately with core, belt, parabelt, and ATL as seed regions for each individual subject. For humans, probabilistic fiber tracking was also performed with MTR as the seed region. Target regions of interest (ROIs) were defined manually based on the probabilistic fiber tracking results in each subject for dorsal, ventral, and parietal waypoints, respectively.

Parcellation of the Auditory Cortex.

Individual-based parcellation. To compare the connectivity patterns of the auditory cortex across different species, we sought to parcellate the auditory cortices of all species using the same individual-based quantitative parcellation method. In detail, 1) For each subject, we manually localized the auditory cortex, and defined the gray matter (GM) surface and GM–WM (white matter) boundary. 2) The columnar anisotropy was estimated by local probabilistic fiber tracking using FSL (37). The columnar anisotropy analysis was applied to find the best fiber route from each voxel on the GM surface to the GM–WM boundary with the highest connectivity probability. These fiber routes were defined as cortical columns. 3) The GM of the auditory cortex for each subject was divided into the upper layer part and lower layer part based on the cortical columns. 4) Five MRI contrasts including T2 (T1 for humans), MD, FA, odi, and DEC across two different layers were used to construct the feature space. For each subject, each cortical column was represented by 10 features (*SI Appendix*, Fig. S4). 4) Unsupervised distance-based clustering (K-means clustering) was used to group the cortical columns into a certain number of clusters (in this case, 4 clusters are the best, see details below), the classification labels of the cortical columns were corrected based on their spatial locations. The labels for the clusters (core, lateral belt, medial belt, and parabelt) were also defined based on their spatial locations. The above methodologies for segmenting the auditory cortex including MRI contrast calculation, columnar anisotropy computation, and K-means clustering have been previously developed and published systematically by Dr. Afonso Silva's group using both ex vivo (37, 100) and in vivo marmoset data (101), all resources are open-accessed and available online (https://marmosetbrainmapping.org).

To demonstrate the optimal number of clusters for the parcellation, WCSS (within-cluster sum of square) for each subject was calculated. WCSS is the sum of the squared distance between each point and the centroid in a cluster. WCSS curve was plotted across different numbers of clusters. The cluster number corresponding to the elbow point of the WCSS curve is the optimal number of clusters.

To check the robustness of the parcellation results, a 10-fold cross-validation was performed. A training dataset which consisted of 90% of the total dataset (cortical columns) was randomly selected for each subject, with the remaining 10% used as the testing dataset for validation. The label for each cortical column of the dataset was derived from the clustering results shown above. Support vector machine (SVM) was used to train the model and compute the misclassification rate. This fitting and testing procedure was performed for 10 times, and the original error rate was computed by averaging the misclassification rates. In addition, to compute a null distribution of the misclassification rates, 10-fold cross-validations were performed on the same dataset by shuffling the corresponding labels. The distribution of shuffled error rates was computed by shuffling the labels for 1,000 times. The Wilcoxon rank-sum test (WRS-test) was used to compare the original error rates with the shuffled error rates.

Previous studies have shown that auditory belt regions have reciprocal connections with auditory core and parabelt regions, whereas no connections were found between auditory core and parabelt regions (26, 40). These findings provided us with an alternative criterion to validate our parcellation results. The connectivity strength of the three sub-regions (core, belt, and parabelt) from our parcellation results were estimated by probabilistic fiber tracking using FSL for each subject. Voxel-voxel fiber count was defined as the connected fiber number between a given voxel in one sub-region and a given voxel in another sub-region. Voxel-voxel fiber count was calculated for all voxels in both regions respectively.

3D surface reconstruction and atlas registration. The 3D cortical surface of the brain was reconstructed using Freesurfer (https://surfer.nmr.mgh.harvard.edu/). T2w images of marmosets and macaque and T1w images of humans were used for the reconstruction. The parcellation results were mapped onto the individual brain surface for visualization. To compare our parcellation results with the existing brain atlases, the Nencki–Monash (NM) template (43) and the NIH-MBA v1.1 (NIH) template (37) for marmosets were registered to the T2w images of marmosets, and the D99 macaque template (44) was registered to the T2w images of macaque. Affine transformations implemented in Freesurfer were used for all registrations. The inconsistency rate was calculated as the percentage of inconsistent voxels across the total number of voxels between our parcellation results and the atlas. The human brain atlas was not used in this study since there is no common agreement for the auditory core, belt, and parabelt regions in humans.

Statistical Analysis for Cross-Species Comparisons. The statistical analyses were performed across subjects for each species (ex vivo marmoset, in vivo marmoset, macaque, and human) in six aspects. Paired *t* tests were used for the comparisons.

1) Comparisons of the relative size among the auditory core, belt, and parabelt regions.
2) Comparisons of the connectivity strength from the auditory core, belt, and parabelt regions to the three target regions (dorsal target ROIs, ventral target ROIs, and parietal target ROIs).
3) Comparisons of the connectivity strength from the rostral and caudal parts of the auditory cortex to the three target regions (dorsal target ROIs, ventral target ROIs, and parietal target ROIs).
4) Comparisons of the connectivity strength from the left and right auditory cortex to the three target regions (dorsal target ROIs, ventral target ROIs, and parietal target ROIs).
5) Comparisons of the connectivity strength between MTR and auditory cortex to the three target regions (dorsal target ROIs, ventral target ROIs, and parietal target ROIs) in humans.
6) Comparisons of the connectivity strength between ATL and auditory cortex to the three target regions (dorsal target ROIs, ventral target ROIs, and parietal target ROIs).

Data, Materials, and Software Availability. The raw data supporting the current study (.nii format) have been deposited in Open Science Foundation repository (https://doi.org/10.17605/OSF.IO/CZGEF) (102).

ACKNOWLEDGMENTS. This research was supported by NIH grants DC003180 and DC005808. We thank J. Lynch, K. Schonvisky, S. Miller, and J. Izzi for their assistance with animal care, Jiadi Xu and Kazi Akhter for their assistance on MRI scans, Hanzhang Lu and all Wang's lab members especially Xindong Song for their comments on the manuscript.

1. J. P. Rauschecker, Parallel processing in the auditory cortex of primates. *Audiol. Neurotol.* **3**, 86–103 (1998).
2. B. Tian, D. Reser, A. Durham, A. Kustov, J. P. Rauschecker, Functional specialization in rhesus monkey auditory cortex. *Science* **292**, 290–293 (2001).
3. J. P. Rauschecker, K. S. Sophie, Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **12**, 718–724 (2009).
4. L. M. Romanski *et al.*, Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* **2**, 1131–1136 (1999).
5. G. Hickok, D. Poeppel, The cortical organization of Speech Processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
6. D. Saur *et al.*, Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18035–18040 (2008).
7. J. K. Rilling, M. F. Glasser, S. Jbabdi, J. Andersson, T. M. Preuss, Continuity, divergence, and the evolution of brain language pathways. *Front Evol. Neurosci.* **3**, 11 (2012).
8. A. D. Friederici, The cortical language circuit: From auditory perception to sentence comprehension. *Trends Cogn. Sci.* **16**, 262–268 (2012).
9. A. D. Friederici, Evolution of the neural language network. *Psych. Bull Rev.* **24**, 41–47 (2016).
10. N. Geschwind, The organization of language and the brain. *Science* **170**, 940–944 (1970).
11. M. Catani, M. Mesulam, The arcuate fasciculus and the disconnection theme in language and aphasia: History and current State. *Cortex* **44**, 953–961 (2008).
12. J. K. Rilling *et al.*, The evolution of the arcuate fasciculus revealed with comparative dti. *Nat. Neurosci.* **11**, 426–428 (2008).
13. N. Eichert *et al.*, What is special about the human arcuate fasciculus? Lateralization, projections, and expansion. *Cortex* **118**, 107–115 (2019).
14. F. Balezeau *et al.*, Primate auditory prototype in the evolution of the arcuate fasciculus. *Nat. Neurosci.* **23**, 611–614 (2020).
15. Y. Becker, K. K. Loh, O. Coulon, A. Meguerditchian, The arcuate fasciculus and language origins: Disentangling existing conceptions that influence evolutionary accounts. *Neurosci. Biobehav. Rev.* **134**, 104490 (2022).
16. G. J. M. Parker *et al.*, Lateralization of ventral and dorsal auditory-language pathways in the human brain. *NeuroImage* **24**, 656–666 (2005).
17. C. Weiller, T. Bormann, D. Saur, M. Musso, M. Rijntjes, How the ventral pathway got lost – and what its recovery might mean. *Brain Lang.* **118**, 29–39 (2011).
18. A. S. Dick, P. Tremblay, Beyond the arcuate fasciculus: Consensus and controversy in the Connectional Anatomy of language. *Brain* **135**, 3529–3550 (2012).
19. C. J. Bajada, M. A. Lambon Ralph, L. L. Cloutman, Transport for language south of the sylvian fissure: The routes and history of the main tracts and stations in the ventral language network. *Cortex* **69**, 141–151 (2015).
20. J. Hau *et al.*, Cortical terminations of the inferior fronto-occipital and uncinate fasciculi: Anatomical stem-based virtual dissection. *Front. Neuroanat.* **10**, 58 (2016).
21. T. Nakae *et al.*, Connectivity gradient in the human left inferior frontal gyrus: Intraoperative cortico-cortical evoked potential study. *Cerebral Cortex* **30**, 4633–4650 (2020).
22. C. Weiller *et al.*, The ventral pathway of the human brain: A continuous association tract system. *NeuroImage* **234**, 117977 (2021).

23. N. K. Logothetis, MR imaging in the non-human primate: Studies of function and of dynamic connectivity. *Curr. Opin. Neurobiol.* **13**, 630–642 (2003).
24. M. P. Milham *et al.*, An open resource for non-human primate imaging. *Neuron* **100**, 61–74 (2018).
25. X. Zhang, Y. Zhang, A. W. Roe, Ultra-high-field MRI studies of brain structure and function in humans and nonhuman primates: A collaborative approach to precision medicine. *Curr. Opin. Biomed. Eng.* **20**, 100320 (2021).
26. T. A. Hackett, I. Stepniewska, J. H. Kaas, Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J. Comp. Neurol.* **394**, 475–495 (1998).
27. J. H. Kaas, T. A. Hackett, Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11793–11799 (2000).
28. M. E. Steiper, N. M. Young, Primate molecular divergence dates. *Mol. Phylogenet. Evol.* **41**, 384–394 (2006).
29. X. Wang, On cortical coding of communication sounds in primates. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11843–11849 (2000).
30. A. L. Pistorio, B. Vintch, X. Wang, Acoustic analysis of vocal development in a New World primate, the common marmoset (Callithrix jacchus). *J. Acoust. Soc. Am.* **120**, 1655–1670 (2006).
31. T. Hashikawa, R. Nakatomi, A. Iriki, Current models of the marmoset brain. *Neurosci. Res.* **93**, 116–127 (2015).
32. J. F. Mitchell, D. A. Leopold, The marmoset monkey as a model for visual neuroscience. *Neurosci. Res.* **93**, 20–46 (2015).
33. C. T. Miller *et al.*, Marmosets: A neuroscientific model of human social behavior. *Neuron* **90**, 219–233 (2016).
34. J. Homman-Ludiye, J. A. Bourne, The marmoset: An emerging model to unravel the evolution and development of the primate neocortex. *Dev. Neurobiol.* **77**, 263–272 (2016).
35. D. K. Jones, Studying connections in the living human brain with diffusion MRI. *Cortex* **44**, 936–952 (2008).
36. K. S. Ambrosen *et al.*, Validation of structural brain connectivity networks: The impact of scanning parameters. *NeuroImage* **204**, 116207 (2020).
37. C. Liu *et al.*, A digital 3D atlas of the marmoset brain based on multi-modal MRI. *NeuroImage* **169**, 106–116 (2018).
38. A. L. Pistorio, S. H. Hendry, X. Wang, A modified technique for high-resolution staining of myelin. *J. Neurosci. Methods* **153**, 135–146 (2006).
39. N. A. Bock, A. Kocharyan, J. V. Liu, A. C. Silva, Visualizing the entire cortical myelination pattern in marmosets with magnetic resonance imaging. *J. Neurosci. Methods* **185**, 15–22 (2009).
40. J. H. Kaas, T. A. Hackett, Subdivisions of auditorycortex and levels of processing in primates. *Audiol. Neurotol.* **3**, 73–85 (1998).
41. T. A. Hackett, I. Stepniewska, J. H. Kaas, Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res.* **817**, 45–58 (1999).
42. P. Majka *et al.*, Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. *J. Comp. Neurol.* **524**, 2161–2181 (2016).
43. P. Majka *et al.*, Histology-based average template of the marmoset cortex with probabilistic localization of cytoarchitectural areas. *NeuroImage* **226**, 117625 (2021).
44. K. S. Saleem *et al.*, High-resolution mapping and Digital Atlas of subcortical regions in the macaque monkey based on matched map-MRI and histology. *NeuroImage* **245**, 118759 (2021).
45. J. F. Démonet *et al.*, The anatomy of phonological and semantic processing in normal subjects. *Brain* **115**, 1753–1768 (1992).
46. M. Vigneau *et al.*, Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage* **30**, 1414–1432 (2006).
47. T. Ethofer *et al.*, Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *NeuroImage* **30**, 580–587 (2006).
48. L. K. Tyler, P. Wright, B. Randall, W. D. Marslen-Wilson, E. A. Stamatakis, Reorganization of syntactic processing following left-hemisphere brain damage: Does right-hemisphere activity preserve function? *Brain* **133**, 3396–3408 (2010).
49. K. Patterson, P. J. Nestor, T. T. Rogers, Where do you know what you know? the representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* **8**, 976–987 (2007).
50. C. Perrodin, C. Kayser, T. J. Abel, N. K. Logothetis, C. I. Petkov, Who is that? Brain networks and mechanisms for identifying individuals. *Trends Cogn. Sci.* **19**, 783–796 (2015).
51. Y. Zhang *et al.*, Hierarchical cortical networks of "voice patches" for processing voices in human brain. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2113887118 (2021).
52. S. Sadagopan, N. Z. Temiz-Karayol, H. U. Voss, High-field functional magnetic resonance imaging of vocalization processing in marmosets. *Sci. Rep.* **5**, 1–15 (2015).
53. C. I. Petkov *et al.*, A voice region in the monkey brain. *Nat. Neurosci.* **11**, 367–374 (2008).
54. E. Premereur, J. Taubert, P. Janssen, R. Vogels, W. Vanduffel, Effective connectivity reveals largely independent parallel networks of face and body patches. *Curr. Biol.* **26**, 3269–3279 (2016).
55. D. Archakov *et al.*, Auditory representation of learned sound sequences in motor regions of the macaque brain. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15242–15252 (2020).
56. E. T. Rolls, J. P. Rauschecker, G. Deco, C. C. Huang, J. Feng, Auditory cortical connectivity in humans. *Cerebral Cortex* **33**, 6207–6227 (2023).
57. J. M. Sinnott, W. C. Stebbins, D. B. Moody, Regulation of voice amplitude by the monkey. *J. Acoust. Soc. Am.* **58**, 412–414 (1975).
58. H. Brumm, K. Voss, D. Köllmer Ireen, Todt, Acoustic communication in noise: Regulation of call characteristics in a new world monkey. *J. Exp. Biol.* **207**, 443–448 (2004).
59. S. J. Eliades, X. Wang, Neural correlates of the lombard effect in primate auditory cortex. *J. Neurosci.* **32**, 10737–10748 (2012).
60. L. Zhao, B. B. Rad, X. Wang, Long-lasting vocal plasticity in adult marmoset monkeys. *Proc. R. Soc. B: Biol. Sci.* **286**, 20190817 (2019).
61. T. Pomberger, C. Risueno-Segovia, J. Löschner, S. R. Hage, Precise motor control enables rapid flexibility in vocal behavior of Marmoset Monkeys. *Curr. Biol.* **28**, 788–794.e3 (2018).
62. M. D. Hauser, N. Chomsky, W. T. Fitch, The faculty of language: What is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).
63. T. A. Chaplin, H.-H. Yu, J. G. Soares, R. Gattass, M. G. Rosa, A conserved pattern of differential expansion of cortical areas in Simian Primates. *J. Neurosci.* **33**, 15120–15125 (2013).
64. R. B. Mars, R. E. Passingham, F.-X. Neubert, L. Verhagen, J. Sallet, "Evolutionary specializations of human association cortex" in *Evolution of Nervous Systems*, J. H. Kaas, T. M. Preuss, Eds. (Elsevier, New York, NY, 2017), pp. 185–205, 10.1016/b978-0-12-804042-3.00118-4.
65. D. J. Ardesch *et al.*, Evolutionary expansion of connectivity between Multimodal Association areas in the human brain compared with chimpanzees. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7101–7106 (2019).
66. D. C. Van Essen *et al.*, Cerebral cortical folding, parcellation, and connectivity in humans, nonhuman primates, and mice. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 26173–26180 (2019).
67. R. J. Zatorre, P. Belin, V. B. Penhune, Structure and function of auditory cortex: Music and speech. *Trends Cogn. Sci.* **6**, 37–46 (2002).
68. M. F. Glasser, J. K. Rilling, DTI tractography of the human brain's language pathways. *Cerebral Cortex* **18**, 2471–2482 (2008).
69. R. J. Zatorre, P. Belin, Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* **11**, 946–953 (2001).
70. P. Albouy, L. Benjamin, B. Morillon, R. J. Zatorre, Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and Melody. *Science* **367**, 1043–1047 (2020).
71. J. L. Ringo, R. W. Doty, S. Demeter, P. Y. Simard, Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cerebral Cortex* **4**, 331–343 (1994).
72. K. A. Phillips *et al.*, The corpus callosum in primates: Processing speed of axons and the evolution of hemispheric asymmetry. *Proc. R. Soc. B: Biol. Sci.* **282**, 20151535 (2015).
73. L. Cheng *et al.*, Connectional asymmetry of the inferior parietal lobule shapes hemispheric specialization in humans, chimpanzees, and rhesus macaques. *eLife* **10**, e67600 (2021).
74. Y. S. Zhang, A. Asif, Ghazanfar., A hierarchy of autonomous systems for vocal production. *Trends Neurosci.* **43**, 115–126 (2020).
75. J. P. Rauschecker, An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Res.* **271**, 16–25 (2011).
76. U. Jürgens, Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* **26**, 235–258 (2002).
77. G. Coudé *et al.*, Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One* **6**, e26822 (2011).
78. S. R. Hage, N. Gavrilov, A. Nieder, Cognitive control of distinct vocalizations in Rhesus Monkeys. *J. Cogn. Neurosci.* **25**, 1692–1701 (2013).
79. C. T. Miller, K. Beck, B. Meade, X. Wang, Antiphonal call timing in marmosets is behaviorally significant: Interactive playback experiments. *J. Comp. Physiol. A* **195**, 783–789 (2009).
80. S. Roy, C. T. Miller, D. Gottsch, X. Wang, Vocal control by the common marmoset in the presence of interfering noise. *J. Exp. Biol.* **214**, 3619–3629 (2011).
81. D. Y. Takahashi, D. Z. Narayanan, A. Ghazanfar, A., Coupled oscillator dynamics of vocal turn-taking in Monkeys. *Curr. Biol.* **23**, 2162–2168 (2013).
82. J. Y. Choi, D. Y. Takahashi, A. A. Ghazanfar, Cooperative vocal control in marmoset monkeys via vocal feedback. *J. Neurophysiol.* **114**, 274–283 (2015).
83. M. S. Osmanski, X. Wang, Perceptual specializations for processing species-specific vocalizations in the common marmoset (*callithrix jacchus*). *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2221756120 (2023).
84. L. Zhao, S. Roy, X. Wang, Rapid modulations of the vocal structure in Marmoset Monkeys. *Hearing Res.* **384**, 107811 (2019).
85. R. D. Patterson, S. Uppenkamp, I. S. Johnsrude, T. D. Griffiths, The processing of temporal pitch and melody information in auditory cortex. *Neuron* **36**, 767–776 (2002).
86. H. Penagos, A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J. Neurosci.* **24**, 6810–6815 (2004).
87. S. Norman-Haignere, N. Kanwisher, J. H. McDermott, Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. *J. Neurosci.* **33**, 19451–19469 (2013).
88. D. Bendor, X. Wang, The neuronal representation of pitch in primate auditory cortex. *Nature* **436**, 1161–1165 (2005).
89. X. Song, M. S. Osmanski, Y. Guo, X. Wang, Complex pitch perception mechanisms are shared by humans and a new world monkey. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 781–786 (2015).
90. S. V. Norman-Haignere, N. Kanwisher, J. H. McDermott, B. R. Conway, Divergence in the functional organization of human and macaque auditory cortex revealed by fmri responses to Harmonic Tones. *Nat. Neurosci.* **22**, 1057–1060 (2019).
91. C. M. Cerkevich, J. A. Rathelot, P. L. Strick, Cortical basis for skilled vocalization. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122345119 (2022).
92. S. B. Sébille, A.-S. Rolland, M.-L. Welter, E. Bardinet, M. D. Santin, Post mortem high resolution diffusion MRI for large specimen imaging at 11.7 T with 3D segmented echo-planar imaging. *J. Neurosci. Methods* **311**, 222–234 (2019).
93. D. C. Van Essen *et al.*, The Wu-minn human connectome project: An overview. *NeuroImage* **80**, 62–79 (2013).
94. J. Veraart *et al.*, Denoising of diffusion MRI using random matrix theory. *NeuroImage* **142**, 394–406 (2016).
95. M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, S. M. Smith, FSL. *NeuroImage* **62**, 782–790 (2012).
96. M. O. Irfanoglu *et al.*, DR-Buddi (diffeomorphic registration for blip-up blip-down Diffusion Imaging) method for correcting echo Planar Imaging distortions. *NeuroImage* **106**, 284–299 (2015).
97. B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, J. Sijbers, Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* **103**, 411–426 (2014).
98. H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, D. C. Alexander, Noddi: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* **61**, 1000–1016 (2012).
99. T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, M. W. Woolrich, Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* **34**, 144–155 (2007).
100. C. Liu *et al.*, A resource for the detailed 3D mapping of white matter pathways in the marmoset brain. *Nat. Neurosci.* **23**, 271–280 (2020).
101. C. Liu *et al.*, Marmoset Brain Mapping V3: Population multi-modal standard volumetric and surface-based templates. *Neuroimage* **226**, 117620 (2021).
102. Y. Zhang, Evolutionary continuity and divergence of auditory dorsal and ventral pathways in primates revealed by ultra-high field diffusion MRI. Open Science Framework. https://doi.org/10.17605/OSF.IO/CZGEF. Deposited 24 January 2024.