RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

An opposing molecular gradient axis underlies primate cortical organization

Zhi Huang†, Qianqian Yang†, Shenglong Li†, Xiaojia Zhu†, He Wang†, Jixuan Lin† *et al.*

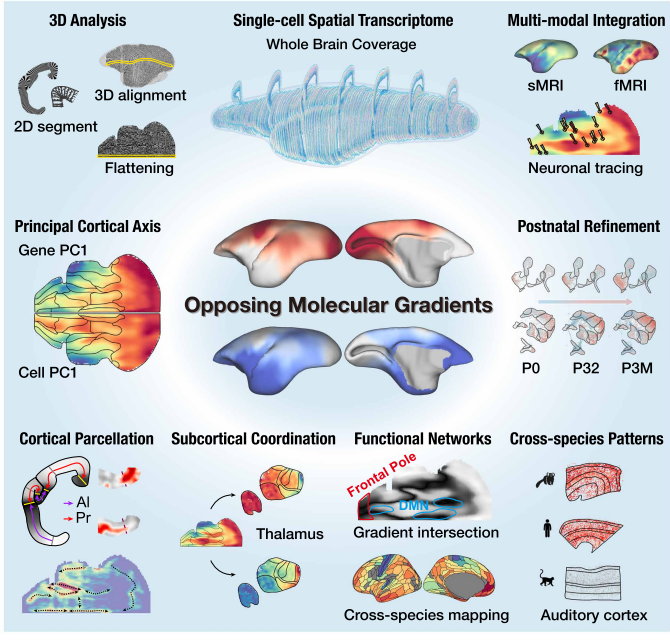
Full article and list of author affiliations: https://doi.org/10.1126/science.aea2673

INTRODUCTION: The primate cerebral cortex has expanded into a complex mosaic of specialized areas and networks that support advanced cognition. However, the fundamental principles governing this complex cortical organization remain elusive, which has led to conflicting hypotheses regarding cortical expansion. The dual-origin hypothesis posits that the cortex expanded from two evolutionarily ancient allocortical regions and progressively differentiated toward the most specialized six-layered primary sensory areas (the koniocortex). Conversely, alternative theories propose that primary sensory areas served as the early anchors for cortical expansion. Resolving this conflict requires a holistic view that integrates gene expression, cellular architecture, and brain-wide connectivity.

RATIONALE: The common marmoset (*Callithrix jacchus*) offers a solution to this challenge. It retains key primate brain characteristics yet has a smooth (lissencephalic) cortex, bypassing the analytical difficulties caused by the complex folding seen in the cortices of larger primates.

Leveraging this advantage, we integrated whole-brain, single-cell–resolution spatial transcriptomics and single-nucleus RNA sequencing with magnetic resonance imaging (MRI) and neuronal tracing data. This multimodal synthesis allowed us to uncover the principles underlying cortical organization.

RESULTS: Our analysis uncovered a fundamental axis of cortical organization defined by two opposing molecular gradients that capture the dominant patterns of cortical gene expression, cell composition, and hierarchy. One gradient radiates from allocortical and periallocortical regions (e.g., the piriform and entorhinal cortices), whereas the opposing gradient originates from primary sensory areas, with the association cortex residing at their intersection. Conserved across humans, macaques, marmosets, and mice, these gradients reconcile the conflicting hypotheses by demonstrating that both the allocortical and primary sensory areas act as anchors at opposite ends of a single organizational axis. Although present at birth, these gradients undergo prominent postnatal refinement, which suggests that they are actively shaped by sensory experience. Molecular shifts along these gradients underpin cortical parcellation, where sharp transitions in cell composition and gene expression align with cortical area borders and reveal previously unrecognized subdivisions. Notably, these cortical gradients are mirrored in thalamic gene expression patterns and align with thalamocortical connectivity. This molecular coupling between the cortex and thalamus is significantly stronger in marmosets than it is in mice, which highlights the evolutionary advances in primate-specific thalamocortical integration. At the gradient convergence zone, the putative default mode network (DMN) and the frontal pole (area 10) in marmosets share similar molecular signatures despite their weak functional connectivity, suggesting that this molecular identity evolved before the strong connectivity observed in humans. Finally, comparative analysis of gradient-related genes revealed that the marmoset auditory cortex more closely resembles the human auditory cortex compared with that of the macaque, likely reflecting shared neural mechanisms for complex vocal communication.



Opposing molecular gradients of the cerebral cortex. By integrating whole-brain spatial transcriptomics with MRI and neuronal tracing in marmosets, we revealed an opposing molecular gradient axis as a fundamental principle of cortical organization. Undergoing active postnatal refinement, these gradients serve as a key organizational backbone for delineating cortical boundaries, elucidating cortical-subcortical relationships, characterizing functional networks, and identifying species-specific molecular patterns. 3D, three-dimensional; PC1, principal component 1; Al, allocortical/periallocortical; Pr, primary sensory; P0, postnatal day 0; P32, postnatal day 32; P3M, postnatal 3 months.

CONCLUSION: Through multimodal analysis, we identify an opposing molecular gradient axis as a fundamental principle of primate cortical organization that resolves debates regarding cortical expansion. This axis links molecular profiles to anatomical and functional architecture, offering a precise biological basis for delineating cortical boundaries, elucidating cortical-subcortical relationships, characterizing functional networks, and identifying species-specific molecular specializations. Together, this work establishes the opposing gradient axis as a key organizational backbone and presents a foundational multimodal resource for understanding primate brain organization and evolution. □

Corresponding authors: Cirong Liu (crliu@ion.ac.cn); Shijie Hao (haoshijie@genomics.cn); Yidi Sun (ydsun@ion.ac.cn); Marcello G. P. Rosa (marcello.rosa@monash.edu) †These authors contributed equally to this work. **Cite this article as Z. Huang *et al.*, *Science* 392, eaea2673 (2026). DOI: 10.1126/science.aea2673**

NEUROSCIENCE

An opposing molecular gradient axis underlies primate cortical organization

Zhi Huang[1,2,3]†, Qianqian Yang[1,4,5]†, Shenglong Li[1,2,3]†, Xiaojia Zhu[1,6,7]†, He Wang[1]†, Jixuan Lin[1,8]†, Yafeng Zhan[1], Yan Wu[2], Zefang Wang[1], Piotr Majka[9], Haichao Qu[1], Nafiseh Atapour[10], Tao Yang[2], Youning Lin[2,11], Luman Cui[2], Yong-Gang Yao[6,7], Zhifeng Liang[1], Zhen Liu[1,12], Chao Li[1], Wu Wei[4], Yi Zhou[1], Shaojie Ma[1], Zhiming Shen[1,12], Xiaoyu Wei[2], Xun Xu[2,13], Shiping Liu[2,11,13], Chengyu Li[4], Muming Poo[1,12], Longqi Liu[2,11,13], Marcello G. P. Rosa[10]*, Yidi Sun[1,14]*, Shijie Hao[2,3,11,13]*, Cirong Liu[1,8,15]*

The principles organizing cellular diversity and connectivity in primate brains remain elusive. By integrating spatial transcriptomics, magnetic resonance imaging, and retrograde labeling in marmosets, we identified two opposing molecular gradients that undergo postnatal refinement, emanating from allocortices and primary sensory cortices, respectively. These gradients reconcile conflicting hypotheses on cortical expansion and characterize distinct cortical areas. Cortical gradients align with thalamic gene expression and thalamocortical projection patterns. At gradient intersections, the default mode network and frontal pole exhibited similar molecular features in humans and marmosets, despite species-specific differences in functional connectivity. Comparative analysis of gradient-related genes showed that marmoset and human auditory cortices are highly similar but differ from those of macaques, potentially reflecting complex vocalization. Together, these opposing gradients represent a fundamental organizing principle of the primate cortex.

Primate cortical expansion produced diverse areas and networks essential for cognitive functions. However, principles governing this complex cortical organization remain unclear, prompting several divergent hypotheses. The dual-origin hypothesis posits cortical expansion from allocortical areas (the archicortex and paleocortex) into specialized six-layered primary sensory regions (the koniocortex) through progressive laminar differentiation (*1*). Conversely, the molecular anchors (*2*) and neo-associationism (*3*) hypotheses suggest that primary sensory areas were specified early as anchors, guiding adjacent cortical formation during evolutionary and developmental expansion. These hypotheses offer diverse, often contradictory explanations for cortical expansion, area specification, and hierarchical organization. Reconciling these views requires investigating the cellular and connectivity architecture, especially in primates, where cortical expansion is most pronounced.

Understanding the cortex's diverse cellular composition and complex connectivity (*4*) requires integrating transcriptomic and connectomic data (*5*). Although new technologies enable multiscale brain mapping (*6*, *7*), developing an integrated multimodal view of the primate brain remains challenging owing to large brain sizes and cortical folding (*8*). The common marmoset (*Callithrix jacchus*), with its small, lissencephalic brain, offers important advantages. It retains essential primate cortical regions and architectural features (*2*, *9*), which simplifies the study of cortical organization. This has driven extensive brain mapping studies, yielding detailed neuroanatomical atlases (*10*–*13*), multimodal connectomes (*14*, *15*), and transcriptomic datasets (*16*–*19*). However, existing marmoset transcriptomic resources (*16*–*19*) lack comprehensive, whole-brain, and single-cell spatial mapping. Multimodal integration is also hindered by dimensional mismatches between three-dimensional (3D) connectomic and 2D transcriptomic data (*20*).

To address these gaps, we constructed a 3D multimodal atlas of the marmoset brain, integrating whole-brain, single-cell spatial transcriptomes with multimodal magnetic resonance imaging (MRI) (*9*, *14*, *15*, *21*–*25*) and retrograde tracing data (*14*). This integrated atlas reveals a unifying organizational principle of the primate cortex that reconciles previously divergent hypotheses. Guided by this principle, we linked transcriptomic profiles with the connectome to elucidate cortical-subcortical topographic alignment, the molecular signatures of functional networks, and species-specific specializations across primate species.

An integrated multimodal 3D atlas of the marmoset brain

We performed spatial transcriptome and single-nucleus RNA sequencing (snRNA-seq) of the whole marmoset brain using spatial enhanced resolution omics sequencing (Stereo-seq) (*26*) and DNBelab C4 methods (*27*), respectively (Fig. 1A). Our snRNA-seq dataset profiled 496,718 cells from 116 brain areas, and the Stereo-seq data comprised 125 coronal sections from one adult male and 27 parasagittal sections from one adult female (Fig. 1B and table S1, A to C). We reconstructed the Stereo-seq data for whole-cortical analysis using two approaches (fig. S1): (i) segment flattening for a 2D whole-cortex flat map and (ii) a 3D reconstruction framework coregistered with MRI templates (*9*, *10*) to enable integrative analysis with anatomical and functional MRI (fMRI) and retrograde tracing data (table S1D).

Unsupervised clustering using gene expression profiles from snRNA-seq data revealed 10 neuronal glutamatergic, 7 neuronal γ-aminobutyric acid–releasing (GABAergic), and 6 nonneuronal subclasses (fig. S2, A and B). The subclasses, annotated with known markers (*6*, *28*), were further divided into 112 neuronal glutamatergic, 68 neuronal GABAergic, and 27 nonneuronal clusters (referred to tentatively as "cell types") (fig. S2, C to F, and table S1E). For the spatial data, Stereo-seq coronal and parasagittal sections covered all cortical lobes (fig. S3A), yielding 6.37 and 3.82 billion DNA nanoballs (DNBs) (400 million per square centimeter), respectively, with an average of 200 (bin25) and 700 (bin50) genes per bin (fig. S3B). To achieve single-cell segmentation, we applied a deep learning–based cell segmentation method (*6*) on the single-strand DNA (ssDNA) staining images (fig. S3, C to E), generating 12.65 million (coronal) and 8.83 million (parasagittal) segmented cells (fig. S3, E and F). We then annotated these cells using DestVI (*29*), integrating the

[1]Institute of Neuroscience, State Key Laboratory of Genetic Evolution and Animal Models, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China. [2]BGI Research, Hangzhou, China. [3]School of Biology and Biological Engineering, South China University of Technology, Guangzhou, China. [4]Lingang Laboratory, Shanghai, China. [5]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [6]State Key Laboratory of Genetic Evolution and Animal Models, National Research Facility for Phenotypic and Genetic Analysis of Model Animals (Primate Facility), National Resource Center for Non-Human Primates, Yunnan Key Laboratory of Animal Models and Human Disease Mechanisms, and KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China. [7]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China. [8]Sino-Danish College, University of Chinese Academy of Sciences, Beijing, China. [9]Laboratory of Neuroinformatics, Nencki Institute of Experimental Biology of the Polish Academy of Sciences, Warsaw, Poland. [10]Neuroscience Program, Biomedicine Discovery Institute and Department of Physiology, Monash University, Clayton, VIC, Australia. [11]Key Laboratory of Brain Cell Mapping of Zhejiang Province, BGI Research, Hangzhou, China. [12]Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China. [13]State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Hangzhou, China. [14]Shanghai Key Laboratory of Precision Gene Editing and Clinical Translation, Shanghai, China. [15]Shanghai Key Laboratory of Child Brain and Development, Shanghai, China. ***Corresponding author. Email: crliu@ion.ac.cn (C.L.); haoshijie@genomics.cn (S.H.); ydsun@ion.ac.cn (Y.S.); marcello.rosa@monash.edu (M.G.P.R.)** †These authors contributed equally to this work.

snRNA-seq–identified cell types (fig. S3, G and H). The resulting cell type composition was highly reproducible between adjacent sections, both within and between animals, validating our annotations (fig. S3, I to K).

These cell type spatial maps revealed distinct subclass laminar preferences (Fig. 1C and fig. S3L), forming a basis for layer parcellation. We constructed a spatial cell-cell adjacency matrix based on the relative proportions of cell types within each cell's k nearest neighbors (KNNs). Unsupervised clustering of this matrix identified seven clusters corresponding to the six isocortical layers and allocortical layer 2 (Fig. 1D). We refined this parcellation using a supervised method incorporating laminar markers and ssDNA staining (Fig. 1E). This approach accurately segmented diverse cortical regions, capturing simpler structures in the piriform (Pir) and entorhinal (Ent) cortices, the absence of layer 4 (L4) in the anterior cingulate (ACC) and motor (Mot) cortices, and an enlarged L4 in the primary visual cortex (V1) (Fig. 1, F and G). Spatial visualization confirmed L4 presence in the primary somatosensory cortex (S1) and absence in the primary motor cortex (M1) and a relatively larger L4 in V1 compared with the secondary visual cortex (V2) (Fig. 1H), patterns further supported by distinct glutamatergic cell distributions (Fig. 1I). These analyses produced a comprehensive marmoset cortical cell type atlas that enabled laminar parcellation.

Two opposing molecular gradients underlie cortical organization

Analysis of cortical cell type distribution identified 39 (out of 180 total) neuronal cell types with significant regional enrichment, and these were largely partitioned between either allocortical and periallocortical regions (14/39) or primary sensory cortices (21/39) (Fig. 2A and fig. S4A). For example, L3/4_12, L4_2, and PV_5 were enriched in V1, primary auditory cortex (AuA1), and S1, respectively. Conversely, L2_3, L5_1, and RELN_2 were predominantly localized to allocortical/periallocortical regions, including the piriform and entorhinal cortices (fig. S4B). These distinct patterns are likely relevant to competing hypotheses on cortical expansion, which posit either primary sensory (*2*, *3*) or allocortical (*1*) origins.

To test whether cell type composition reflects these origins, we established two reference profiles on a flattened cortex using the segment flattening approach (fig. S1D): a primary sensory ("Pr") profile (from the averaged cell type compositions of primary visual, auditory, and somatosensory areas) and an allocortical/periallocortical ("Al") profile (from the entorhinal and piriform cortices). Correlating the cell type composition of each cortical segment with the two references revealed a strong anticorrelation (Pearson's correlation coefficient $r = -0.84$, $P < 0.001$) (Fig. 2, B and C). The strong anticorrelation remained consistent using individual Pr and Al cortical areas as the reference (fig. S5A). Intermediate segments, mostly association cortices, exhibited mixed composition profiles with weaker correlations to both Pr and Al (table S2B). To visualize this bimodal gradient pattern, we computed a differential Pr-Al index (Pr correlation minus Al correlation), defining a continuous gradient axis between Pr and Al poles (Fig. 2D). The middle temporal area (MT/V5) exhibited a strong Pr signature on this axis, consistent with its primary sensory-like characteristics (*30*). Using MT/V5 alone as the Pr reference yielded a similar Pr-related gradient (fig. S5A), reinforcing the robustness of the axis.

The Pr-Al index captured the dominant spatial organizational axis of cell type composition, as shown by its similarity to the first principal component (PC1; 40.6% variance) of a principal components analysis (PCA) on whole-cortex cell compositions (Fig. 2, E and F; $r = 0.90$, $P < 0.001$). Layer-specific PCA further corroborated this finding, except for layer 1 (fig. S5C). Analysis of mouse (*31*), macaque (*6*), and human (*32*) data at the region level (Fig. 2G and fig. S5, B and C) revealed similar patterns, suggesting that the Pr-Al index is a fundamental, conserved organizational principle. Correlating cell type spatial distributions with the index (fig. S5, D and E) revealed different cell type preferences for Pr or Al patterns (Fig. 2, H and I, and fig. S6A). For example, glutamatergic upper-layer L2/3_11 and all layer 4 neurons displayed Pr-like distributions, whereas deep-layer L6_6 was Al-like. Among GABAergic interneurons, most PV neuron subtypes were Pr-like, consistent with their enrichment in primary sensory areas (*33*), whereas many SST and VIP subtypes exhibited Al-like distribution.

Cortical gene expression PC1 (36.88% variance) also correlated strongly with the Pr-Al index (Fig. 2J), a relationship conserved in mice, macaques, and humans (fig. S6B). By correlating individual genes with the index, we identified 6980 Pr-associated (positive correlation) and 8437 Al-associated (negative correlation) genes (Fig. 2K and table S2C). The top gene sets from these two groups exhibited opposing expression in both glutamatergic and GABAergic cell types, consistent with their spatial preference along the Pr-Al axis (Fig. 2L). Functionally, these genes were implicated in brain development and neurological disorders (*34*) (fig. S6C). KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis revealed that although both sets shared enrichment for core synaptic functions, such as GABAergic, glutamatergic, and serotonergic signaling, they exhibited distinct functional specializations (fig. S6, D and E). For example, the Al-associated gene set was enriched for neuroactive ligand-receptor interactions, consistent with the known role of the allocortex in integrating and relaying diverse signals (*1*), whereas the Pr-associated gene set was enriched for the mitogen-activated protein kinase (MAPK) signaling pathway, critical for sensory processing across the auditory (*35*), visual (*36*), and somatosensory (*37*) systems. These opposing molecular profiles thus may underpin the functional specializations of areas along the Pr-Al axis.

To visualize the gradient intersection, we transformed the Pr-Al index to an intersection index, mapping the endpoints (Pr-Al index = ±1) to 0 and the center (Pr-Al index = 0) to 1 with a smooth transition (Fig. 2M). The intersection index map closely resembled the second principal component (PC2) of the cell composition PCA (fig. S7, A to C). Thus, the Pr-Al index and intersection index captured the first two PCA components, supporting the Pr-Al axis as the major organizational axis. Spatial analysis of 2D Stereo-seq sections also demonstrated opposing molecular gradients radiating from Pr and Al reference areas (Fig. 2N and fig. S7, D and E). These spatial patterns were also shown in the distribution of individual cell types, where Pr- or Al-enriched cell types showed opposite profiles (Fig. 2, O to Q, and fig. S7, F and G). Sharp cell type distribution boundaries, marked by abrupt changes in cell density and marker gene expression, emerged at the intersection zone (Fig. 2Q), likely corresponding to cortical area boundaries.

Together, the Pr-Al index revealed that opposing molecular gradients anchored in Al and Pr regions underlie primate cortical organization, with association cortices at their intersection (fig. S7H). By positing both regions as anchors, our results reconcile the opposing hypotheses of cortical expansion (*1*–*3*).

The opposing gradients are refined during postnatal development

To determine whether the opposing gradients are established innately or refined postnatally, we collected spatial transcriptomics and snRNA-seq data from marmosets at newborn (P0), infant [day 32 (P32)], and juvenile [3 months (P3M)] ages (Fig. 3A and fig. S8, A and B). We calculated a Pr-Al index for each age based on gene expression from the spatial transcriptomics data, using Pr (auditory/somatosensory) and Al (entorhinal) cortices as references. This analysis revealed that the gradients are present at birth (P0) but undergo prominent postnatal refinement (Fig. 3, B and C, and fig. S8C). This refinement was evident in the spatial coexpression of the top 500 adult-identified Pr- and Al-associated genes, which showed progressively stronger opposing patterns over time (Fig. 3C). We quantified this segregation with a Pr-Al gene expression pattern difference (PAD) score, measuring the difference in correlation strengths between within-group (Pr-Pr, Al-Al) and between-group (Pr-Al) gene pairs. This score steadily increased

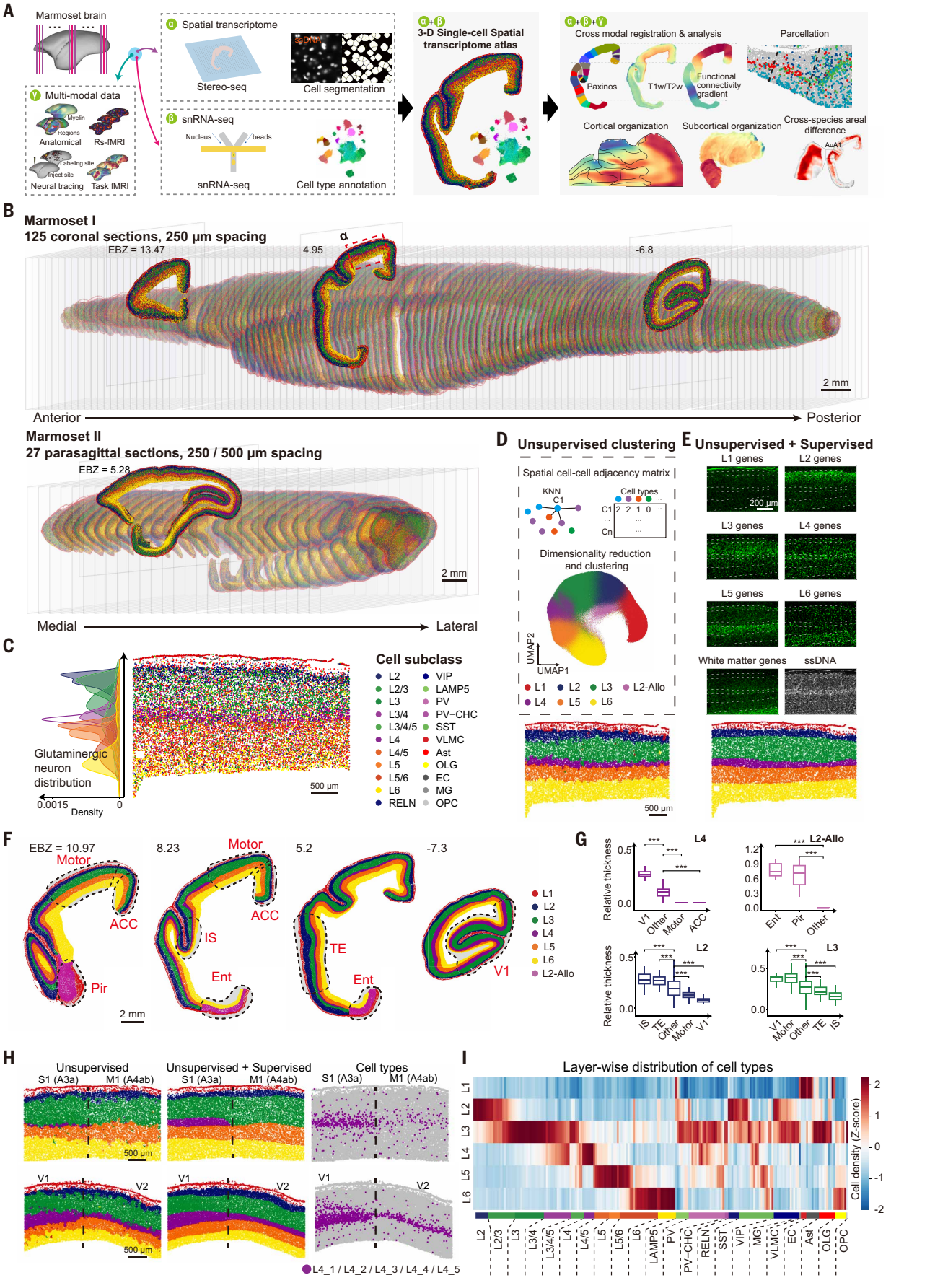


Fig. 1. Single-cell spatial transcriptomic atlas of the marmoset cortex. (**A**) Schematic flow charts of data collection and integration. Spatial transcriptomics (Stereo-seq),

snRNA-seq, multimodal MRI, and retrograde neuronal tracing datasets were combined to create a comprehensive 3D atlas for downstream analyses. (**B**) Spatial distribution of cell subclasses across coronal (top) and parasagittal (bottom) sections of two marmosets. Three coronal sections and one parasagittal section are displayed in the front view, with section numbers indicated. EBZ, ear-bar zero. (**C**) Cell-subclass distribution in an example cortical region marked α in (B), visualized at higher spatial resolution (right), together with a cell density plot for this region (left). Color-coding of cell types is the same as in (B). (**D**) Schematic of the *k* nearest-neighbors (KNN) method and a uniform manifold approximation and projection (UMAP) plot of the seven resulting spatial clusters (marked by colors; top) and spatial map of the seven clusters, corresponding to cortical layers (bottom). (**E**) Examples of cortical layer parcellation improved by supervised learning. (**F**) Cortical layer parcellation in four representative sections, coded in colors as in (E). (**G**) Boxplots showing the average thickness of four example layers in different cortical regions, including the primary visual cortex (V1; $n = 47$ per layer group; sections treated as independent samples), motor cortex (Mot; $n = 60$), anterior cingulate cortex (ACC; $n = 98$), entorhinal cortex (Ent; $n = 36$), piriform cortex (Pir; $n = 11$), insular cortex (IS; $n = 40$), inferior temporal cortex (TE; $n = 74$), and other regions ($n = 1525$, 894, 1511, and 1509 for L4, L2-Allo, L2, and L3 groups respectively) as indicated in (F). Statistical significance was determined by one-way analysis of variance (ANOVA) with post hoc two-sided Student's *t* test. $^{***}P < 0.001$. (**H**) Spatial maps showing cortical layers in four cortical regions (S1, M1, V1, and V2) with layer classification by unsupervised learning (left) and by both unsupervised and supervised learning (middle), together with the distribution of L4 cell types (right). (**I**) Heatmap showing the relative abundance (*z*-scored density) of various cell types in six cortical layers across the whole cortex. Each column represents one cell type, and the cell type subclass is marked below.

from 0.25 at P0 to 1.4 in adults, demonstrating the postnatal sharpening of the gradients.

To assess the cellular basis of this refinement, we integrated snRNA-seq data from all ages, mapping cell type annotation from the adult to earlier ages (fig. S8, D and E). We represented Pr and Al signatures using their top 500 associated genes, a number at which the anticorrelation between these gene sets stabilized (fig. S8F). Calculating expression scores for these sets in each cell revealed that Pr and Al molecular signatures become progressively more distinct with age (Fig. 3D). This trend was clear in glutamatergic and GABAergic neurons but not in nonneuronal cells (fig. S8G), suggesting a primarily neuron-specific refinement process.

Further cell-subclass analysis, quantified by PAD scores, showed distinct developmental trajectories (Fig. 3E). Most glutamatergic neurons exhibited monotonic strengthening of their Pr-Al pattern, with the L3 subclass being the most prominent (Fig. 3E and fig. S8H). Other subclasses, such as SST interneurons, followed a nonmonotonic, "valley-shaped" trajectory (fig. S8H). To identify molecular drivers of these patterns, we tracked each gene's spatial correlation with the Pr-Al index across development, identifying Pr–up-regulated (consistently increasing positive correlation) and Al–up-regulated (increasing negative correlation) genes (Fig. 3F and table S3). For example, the Pr–up-regulated gene *RNFI44A* transitioned from a diffuse pattern at birth into a concentration in sensory cortices (Fig. 3G), whereas the Al–up-regulated gene *TIMP2*, a key regulator of extracellular matrix remodeling vital for learning and memory (*38*), refined from a broad pattern to a localized expression toward the entorhinal cortex (Fig. 3H). These gene sets include many developmental pattern formation genes (*39*), which suggests that they have roles in the spatiotemporal patterning of the developing cortex (fig. S8, I and J). Thus, cortex-wide refinement of the Pr-Al axis is not a monolithic process but a composite of highly specific, temporally regulated gene programs across diverse cell subclasses.

Our findings indicate that the Pr-Al axis is not fully predetermined at birth but undergoes postnatal maturation. Although a primordial axis is innately specified, it may be actively refined by sensory input and maturing neural connectivity after birth.

Molecular shifts along the opposing gradients delineate cortical parcellation

To test whether sharp Pr-Al index changes align with anatomical cortical boundaries, we computed the rate of change (ROC) in cell type density, gene expression, and the Pr-Al index on the flat map (Fig. 4A). The mean ROC of cell density (all cell types) strongly correlated with the ROC of the Pr-Al index (Fig. 4B; $r = 0.55$), which shows that sharp cell composition shifts coincide with rapid index changes. To characterize these shifts, we delineated nine putative boundaries ("a" to "i") on the basis of high mean ROC values (Fig. 4C). These boundaries were heterogeneous: Cell type ROC profiles showed low correlation between boundary pairs (Fig. 4D), and the top 15 contributing cell types for each boundary included distinct combinations of neuronal glutamatergic, neuronal GABAergic, and nonneuronal cells (Fig. 4E and fig. S9A). Gene expression ROC analysis yielded similar results (fig. S9, B and C). Thus, sharp transitions in cell composition and gene expression provide distinct molecular signatures for these putative cortical boundaries.

To investigate the alignment of these sharp cellular transitions with established cortical boundaries (*40*), we analyzed 2D sections. Regions with high mean ROC showed abrupt density shifts across numerous cell types, including glutamatergic, GABAergic, and glial cells (Fig. 4G and fig. S9, D to H). Although the glutamatergic subclass was most abundant and had the highest cumulative ROC, normalizing by cell number revealed similar ROC profiles across all three major cell subclasses. These high-ROC boundaries frequently coincided with high intersection index values and aligned with a subset of Paxinos atlas histological boundaries (*40*). However, some Paxinos-defined boundaries showed low mean ROC, indicating gradual changes in overall cell type composition. Even within these gradual transitions, a few specific cell types exhibited sharp ROC values aligned with Paxinos boundaries (*40*). For example, in one such low-ROC region (Fig. 4H), a subset of L4 neurons sharply defined the A3a-A4ab boundary, whereas most other cell types showed gradual or no change. These findings demonstrate that cortical boundaries are defined either by collective, sharp changes across many cell types or by abrupt changes in a few highly specific cell types.

Similar to cell composition, gene expression analysis revealed coordinated changes in the expression of many genes or a small subset of specific genes at Paxinos area boundaries (*40*). High-ROC regions are enriched at area boundaries of various atlases, including those based on cytoarchitecture (*13*, *40*) and connectivity (*11*, *15*), with alignment being stronger for the cytoarchitecture-based boundaries (fig. S9I). However, known cortical boundaries only account for 30 to 40% of ROC peaks (fig. S9J), suggesting that sharp molecular changes may delineate many previously unrecognized cortical areas. For example, sharp changes in *EYA2* and *SNN* expression delineated a putative previously unidentified boundary within the entorhinal cortex (Fig. 4I), agreeing with a recent cyto- and myeloarchitectural reanalysis of this region (*41*). Similarly, changes in *NOL4* and *NKAIN3* expression highlighted a potential previously undetermined boundary within the primary motor cortex (area A4ab) (Fig. 4J), possibly corresponding to the 4a-4b transition previously reported in baboons (*42*) and suggested in marmosets (*43*). Thus, these molecular and cellular features provide a refined characterization of cortical parcellation beyond traditional methods.

The opposing gradients are mirrored in thalamic gene expression and connectivity

Given that the thalamus and cerebral cortex coevolved as a tightly interconnected system (*44*), we hypothesized that the cortical molecular gradient reflects a shared organizational principle with the thalamus. Supporting this hypothesis, we found that cortical Pr- and Al-enriched gene sets were anticorrelated in the thalamus ($r = -0.96$; Fig. 5A and fig. S10A). To relate the Pr-Al pattern to the thalamus's intrinsic spatial

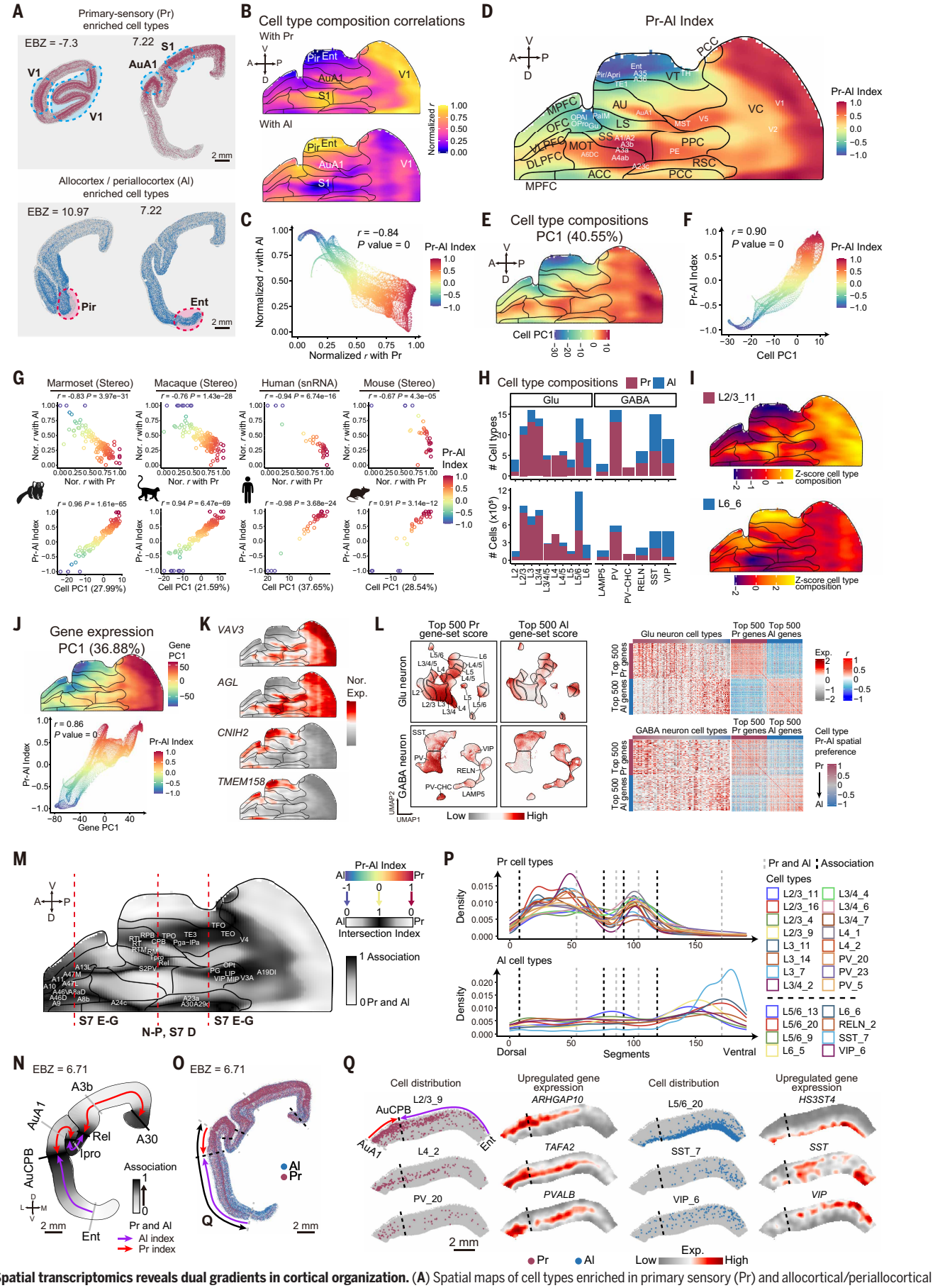


Fig. 2. Spatial transcriptomics reveals dual gradients in cortical organization. (**A**) Spatial maps of cell types enriched in primary sensory (Pr) and allocortical/periallocortical

(AI) regions. (**B**) Pearson correlation coefficients of cell type composition profiles between cortical segments and the Pr or AI reference cortices. (**C**) Scatter plot of correlation between each segment and Pr versus AI reference regions. Color represents the Pr-AI index, calculated as the difference between the Pr correlation and the AI correlation of each segment. (**D**) Spatial pattern of the Pr-AI index on the cortical flat map. (**E**) The first principal component (PC1; percentage indicates variance explained) derived from PCA of cell type compositions across all cortical segments. (**F**) Correlation between PC1 [from (E)] and the Pr-AI index. (**G**) Cross-species conservation of the Pr-AI axis and its relation to cell composition PC1. (Top) Scatter plots illustrating regional correlations with Pr and AI references across species. Each dot represents a cortical region in the marmoset (Stereo-seq), macaque (Stereo-seq), human (snRNA-seq), and mouse (Stereo-seq), respectively. (Bottom) The correlation between cell composition PC1 and the Pr-AI index within each dataset. (**H**) Number of cell subclasses (top) and the corresponding cell numbers (bottom) exhibiting statistically significant correlations ($P < 0.05$) with either Pr or AI reference. (**I**) Spatial distribution of two example cell types: L2/3_11 (correlated with Pr) and L6_6 (correlated with AI). (**J**) PC1 derived from PCA of gene expression profiles across all cortical segments. (**K**) Spatial maps of example Pr-associated (*VAV3* and *AGL*) and AI-associated (*CNIH2* and *TMEM158*) genes. (**L**) UMAPs and heatmaps showing Pr and AI gene-set scores (top 500 genes) in glutamatergic (top) and GABAergic (bottom) neurons from snRNA-seq data. For each neuron subclass, left heatmaps display individual gene expression across cell types, ordered by spatial correlation with the Pr-AI index, and the corresponding right heatmaps show the gene-gene correlation matrix. (**M**) Intersection-index map, which is derived by rescaling the Pr-AI index: Values of +1 or −1 are mapped to 0, and a value of 0 is mapped to 1. The red dashed lines indicate the positions of the coronal sections shown in (N) to (P) and fig. S7, F and G. (**N**) Spatial visualization of the intersection index map in a representative Stereo-seq section (EBZ = 6.71 mm). Arrows point to intersection zones of Pr- and AI-originated gradients. (**O**) Spatial distribution of cell types significantly enriched in either Pr or AI reference regions within the Stereo-seq section (EBZ = 6.71 mm). (**P**) Density distribution of Pr-associated (top) and AI-associated (bottom) cell subclasses along the representative dorsal-ventral (D-V) axis path indicated in (N). Vertical dashed lines mark approximate regional boundaries. (**Q**) Examples of cell subclass distributions and the spatial expression patterns of their corresponding marker genes.

gradients (*45*, *46*), we performed PCA on the thalamic gene expression profile. The analysis revealed that PC1 (Fig. 5B) and PC2 (fig. S10B) corresponded to the medial-lateral and dorsal-ventral axes, respectively, consistent with previous studies in humans (*46*). This thalamocortical molecular alignment was reciprocal: The thalamic PC1 strongly correlated with the thalamic expression of the cortical Pr- and AI-enriched gene sets (Fig. 5, A and B), and conversely, genes from the thalamic PC1 poles ("Tha-P," "Tha-N") showed cortical expression patterns that strongly aligned with the cortical Pr-AI index map (Fig. 5C). These findings suggest that the principal thalamic gene expression axis likely corresponds to the cortical Pr-AI axis.

We found strong overlap between thalamic (Tha-P, Tha-N) and cortical (Pr-AI) gene sets (Fig. 5D), suggesting topographic correspondence in their molecular organization. Several neuronal communication gene sets showed similar gradients in both structures (Fig. 5E, fig. S10C, and table S5B). For example, specific genes showed opposing patterns in AI and Tha-N regions versus Pr and Tha-P regions (Fig. 5E and fig. S10, D to F), such as glutamatergic and GABAergic receptors (*GRIN3A*, *GRIA1*, *GRIK5*, *GABRG1*, *GABRA5*, *GABRB3*, and *GABRA2* versus *GRIN2A*, *GABRB2*, *GABRD*, and *GABRA1*); ion channels (*CACNB3*, *SCN3B*, and *KCNA3* versus *CACNB4*, *SCN1B*, and *KCNA1*); and axon-guidance genes (*PCDH19* versus *SPOCK3*). This opposing pattern extends to transcription factors (for example, *FOSL2* versus *KLF16*) and their respective downstream targets (*STAC2* versus *FOXO6*) (fig. S10G and table S5C), suggesting that a common regulatory program governs the molecular identity of corresponding cortical and thalamic areas.

To examine whether the coordinated gene expression reflects thalamocortical connectivity (*45*), we performed retrograde neuronal tracing at 16 cortical sites along the Pr-AI axis (Fig. 5F). Pr cortical injections (V1, A1-2, and AuA1) labeled neurons predominantly in positive-PC1 thalamic nuclei, including the dorsolateral geniculate nucleus (DLG), the ventral posterior nucleus (VP), and the dorsal part of the medial geniculate nucleus (MGD). Conversely, periallocortical injections (A32, A32V, and A24b) labeled neurons in negative-PC1 thalamic nuclei, particularly the medial division of the mediodorsal nucleus (MDM). Intersection area injections (association cortical areas LIP, AIP, and PG) labeled neurons in intermediate-PC1 thalamic regions, such as the ventral lateral nucleus (VL). These findings support the notion that cortical and thalamic gene expression patterns relate to orderly thalamocortical connectivity (*45*, *47*).

This tight relationship is likely shaped by shared developmental interactions (*44*, *48*). Given that thalamic axons arrive earlier in the developing primate cortex than in the rodent cortex (*49*), this timing difference implies potential primate-specific coupling. Comparing marmosets and mice, we found that the opposing Pr-AI gradient pattern is a conserved axis, evident in the cortex and thalamus of both species (fig. S11, A and B). However, the gene expression correlation between corresponding thalamic and cortical regions along this axis is stronger in marmosets (Fig. 5G and fig. S11, C and D), suggesting enhanced thalamocortical interdependence in primates.

To determine whether molecular coupling extends to other regions, we conducted PCA on gene expression profiles of the dorsal striatum (caudate and putamen), a primary recipient of widespread cortical projections (*50*). This analysis revealed that PC1 captures a pattern similar to that of the cortex and thalamus (fig. S11E), exhibiting a strong positive correlation with the Pr-enriched gene set ($r = 0.92$) and a negative correlation with the AI-enriched gene set ($r = -0.91$) (Fig. 5H). Similar opposing patterns were observed in the dorsal striatum for glutamatergic and GABAergic receptors, ion channels, and axon guidance genes (fig. S11, F and G). Further analysis across multiple subcortical regions revealed that this pronounced Pr- and AI-associated anticorrelation was most prominent in the thalamus and dorsal striatum, whereas it was attenuated or absent in other nuclei with weaker or less widespread cortical connections, such as the amygdala, globus pallidus, hypothalamus, substantia nigra, nucleus accumbens, subthalamic nucleus, and cerebellar nuclei (Fig. 5I and fig. S11H). Collectively, these findings demonstrate that the cortical Pr-AI gradient represents a fundamental organizational axis shared by cortically connected subcortical areas, underpinned by topographic anatomical connectivity and coordinated molecular features.

The opposing gradients reflect multimodal anatomical and functional cortical maps

We next sought to determine how the Pr-AI molecular gradient relates to other spatial patterns of the cerebral cortex (Fig. 6, A and B; fig. S12A; and table S1D), including cytoarchitecture, connectivity, and fMRI activation maps (*9*, *10*, *14*, *15*, *22*–*25*). The Pr-AI index correlated strongly with the PV+/SST+ interneuron ratio and the T1w/T2w myelin map, which mark the sensory-to-association transition (*33*, *51*). The index also correlated strongly with cortical hierarchical levels ($r = -0.74$) derived from marmoset anatomical connectivity data (*52*) (Fig. 6B) as well as with cortical hierarchies in macaques and mice (*53*–*55*) (Fig. 6C). Analysis of marmoset task fMRI data (*22*, *24*, *25*) revealed a strong positive correlation between the Pr-AI index and activation patterns engaging widespread cortical networks ($r = 0.57$) but only a weak correlation with activations triggered by vocal or visual stimuli ($r < 0.04$) (fig. S12A and table S1D). Additionally, we assessed the correlation with two simulated gradients along the anterior-posterior (AP) and dorsal-ventral (DV) axes (fig. S12A). The Pr-AI index showed comparable correlation strengths with both axes (AP: $r = -0.58$; DV: $r = 0.60$), which suggests that it captures components of both spatial dimensions rather than aligning with a single primary axis.

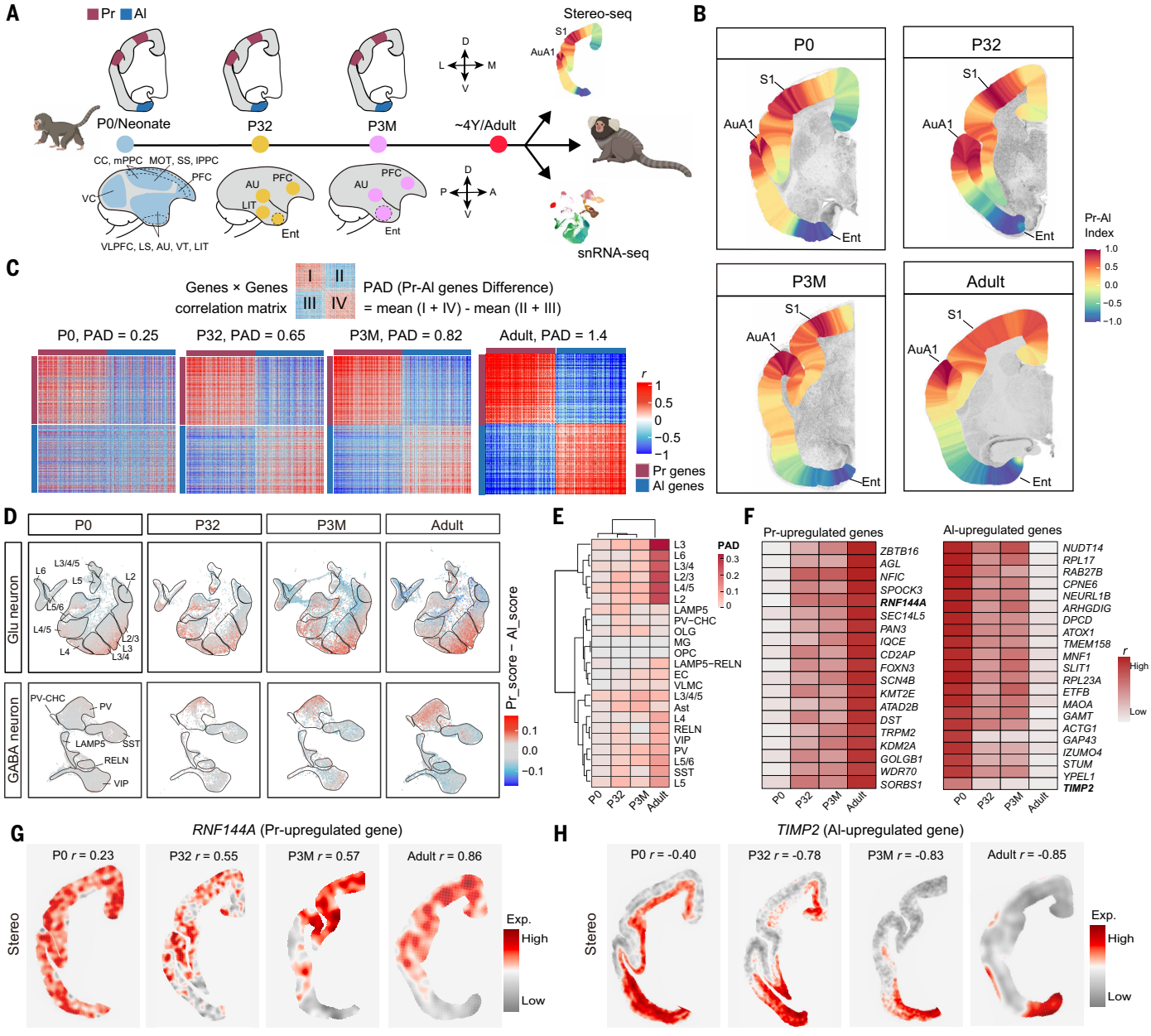


Fig. 3. Postnatal refinement of the molecular gradients. (**A**) Schematic illustration of snRNA-seq and Stereo-seq data collection from marmosets at postnatal day 0 (P0), day 32 (P32), and 3 months (P3M). The cartoon marmoset image was adapted from BioRender.com (https://www.biorender.com/). (**B**) Spatial maps of the Pr-AI index on representative coronal sections at each age. To calculate the index for each cortical segment shown, the primary auditory cortex (AuA1) and primary somatosensory cortex (S1) were used as the Pr reference, and the entorhinal cortex (Ent) served as the AI reference. (**C**) Correlation matrices for the top 500 Pr- and AI-associated genes at each developmental stage, calculated from their expression in corresponding spatial transcriptomics sections. The Pr-AI gene expression PAD score was calculated as the difference between the mean within-group (Pr-Pr, AI-AI) and between-group (Pr-AI) correlation coefficients. (**D**) UMAP projections of glutamatergic (Glu) and GABAergic (GABA) cells derived from snRNA-seq data. Cells are colored by the differences in gene-set scores between the top Pr- and AI-associated gene sets. (**E**) Heatmap showing the developmental trajectory of the PAD score for individual neuronal subclasses from snRNA-seq data. (**F**) Heatmaps showing the correlation coefficients of selected example Pr–up-regulated and AI–up-regulated genes with the Pr-AI index across the four stages in spatial transcriptomics data. (**G** and **H**) Spatial expression maps showing the postnatal refinement of a Pr–up-regulated gene *RNF144A* (G) and an AI–up-regulated gene *TIMP2* (H), with the corresponding correlation coefficient to the Pr-AI index noted above each map.

Given the links between gene expression and connectivity (*5*), we then examined marmoset structural and functional connectivity gradients (*23*), which are low-dimensional representations (*56*) of neuronal tracing (*14*) and resting-state fMRI (*15*) data, respectively. Although the first connectivity gradient distinguishes sensorimotor from visual areas and the second separates transmodal from unimodal regions (*23*), our analysis revealed only a weak correlation between these connectivity gradients and the Pr-AI index (Fig. 6A and fig. S12A). We next determined which pattern yielded the strongest correlation for each gene and cell type. Clustering pattern-enriched cell types and genes ($|r| > 0.5$ with any pattern) revealed that 93.2% of cell types (109 of 117) and 89.6% of genes (6812 of 7605) correlated most strongly with the Pr-AI index (Fig. 6, D and E, and fig. S12B). By contrast, only a small gene subset correlated more strongly with the connectivity gradients (Fig. 6,

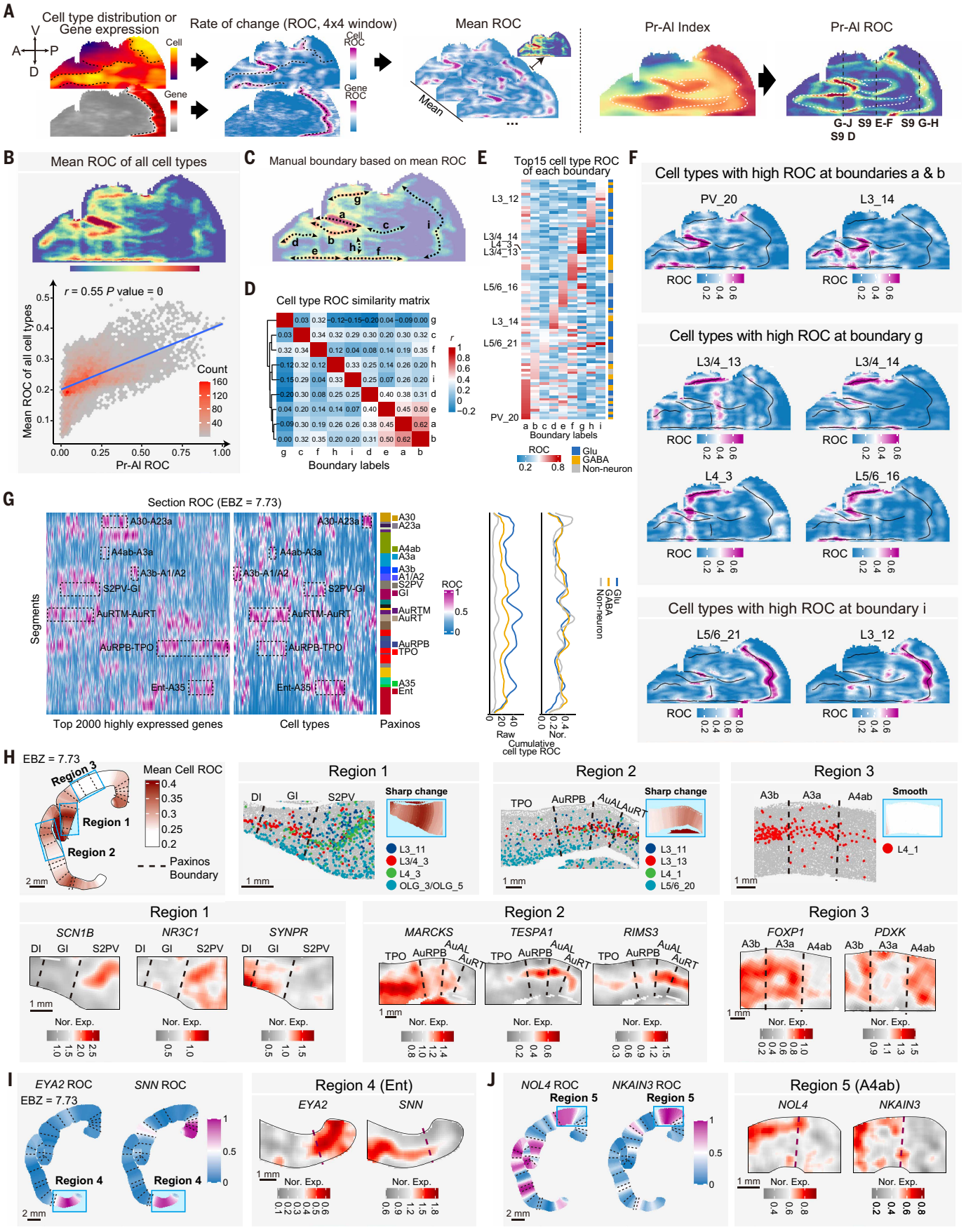


Fig. 4. Cell type and gene expression changes delineate boundaries of cortical areas. (**A**) Schematic illustration of the procedure for calculating the mean ROC for cell type

distribution, gene expression, and the Pr-Al index. Higher mean ROC values indicate sharper transitions. Annotations (lines and labels) on the Pr-Al ROC map indicate the locations of sections detailed in (G) to (J) and in fig. S9, D to H. (**B**) The mean ROC of all cell types across the cortex (top) and its correlation with the Pr-Al ROC (bottom). (**C**) Nine putative boundaries (labeled "a" to "i") manually delineated along ridges of high mean ROC [from (B)], indicating zones where cell type composition changes most sharply. (**D**) Heatmap showing the similarity relationships between the nine putative boundaries, quantified by the Pearson correlation coefficients of their respective cell type ROC profiles. Hierarchical clustering reveals groups of boundaries characterized by similar cell type transition patterns. (**E**) Heatmap showing the ROC for selected cell types across the nine putative boundaries. Displayed cell types are those ranking among the top 15 highest ROC values for at least one boundary. Cell types explicitly labeled are examples that have ROC patterns visualized in (F). The sidebar on the right indicates the cell subclasses (glutamatergic, GABAergic, and nonneuronal) for each cell type. (**F**) Visualization of ROC patterns for selected cell types. (**G**) (Left) Heatmaps showing normalized ROC values for the top 2000 most highly expressed genes and each cell type. Anatomical regional labels are provided alongside the heatmaps. (Right) Cumulative ROC (summed across cell types) and normalized cumulative ROC (cumulative ROC divided by cell type count) across all segments. (**H**) (Top) Three regions with varying mean ROC (left) and spatial distribution of specific cell types exhibiting sharp changes in regions 1 and 2 with high mean ROC and region 3 with low mean ROC (right). (Bottom) Normalized expression of genes with sharp changes across boundaries in regions 1 to 3 (EBZ = 7.73 mm). A list of full names of abbreviated regions is provided in table S1B. (**I** and **J**) Spatial visualization of ROC and expression patterns for the example genes showing sharp changes within the entorhinal cortex (Ent) (I) and the primary motor cortex (A4ab) (J).

D to F). These results suggest that connectivity gradients, as low-dimensional representations of global connectivity patterns, may obscure the distinct molecular signatures of specific brain networks, motivating our subsequent investigation into individual functional networks.

The default mode network resides at the intersection of opposing gradients

We analyzed the molecular basis of 15 functional networks previously identified using independent components analysis (ICA) on resting-state fMRI in awake marmosets (*15*) (fig. S13A and table S7A). Assessing their transcriptomic identity using the Pr-Al intersection index revealed a spectrum (Fig. 7A): Networks in Pr (primary visual LatV1, MedV1) or Al (parahippocampal, ParaHipp) regions exhibited the lowest indices, whereas the putative default mode-like network (DMN) displayed the highest. This peak value situates the DMN at the zone of maximum gradient intersection, consistent with its proposed role at the apex of the cortical hierarchy (*57*).

The human DMN, critical for self-referential thought and social cognition, features the expanded frontal pole (A10) as a central network hub (*58*–*60*). By contrast, the marmoset frontal pole exhibits weak functional connectivity (*21*, *61*) (fig. S13B) and, unlike the DMN, does not occupy an extreme position in the connectivity gradient (Fig. 7B). However, the marmoset frontal pole and DMN shared similarly high Pr-Al intersection indices (Fig. 7A), suggesting similar gene expression. Correlating gene expression with network maps (Fig. 7C) revealed that numerous genes were enriched in networks with localized spatial patterns, particularly in Pr (LatV1) and Al (ParaHipp) regions. By contrast, networks with long-range connectivity, such as the DMN, frontal-parietal network (FPN), and high-level visual network (HighVisC), showed enrichment for far fewer genes. We identified a gene subset enriched in both the DMN and frontal pole (Fig. 7, D and E, and fig. S13D). These include the orphan receptor gene *GPR83* regulating energy metabolism (*62*), the synaptic organizer gene *CBLN2* essential for long-term synapse maintenance (*63*), and *ALOX15* regulating DHA metabolism and spatial working memory (*64*). Given limited gene enrichment in long-range networks, we used partial least squares (PLS) regression (*65*) to identify gene sets that reconstruct network spatial patterns (Fig. 7, F and G). The DMN required the largest gene set for reconstruction (Fig. 7F), reflecting its molecular heterogeneity and likely composition of functionally distinct yet spatially interdigitated subnetworks (*66*). Analysis of DMN PLS weights revealed that the top-ranked genes for the DMN also exhibited high fitting scores in the frontal pole (Fig. 7G). Thus, both gene enrichment and PLS analyses highlighted transcriptomic similarity between the frontal pole and the DMN in marmosets.

To examine whether the molecular signature of the marmoset DMN is evolutionarily conserved in humans, we projected the PLS weights of the top 200 genes from the marmoset DMN reconstruction (123 human homologs; table S7E) onto the human Allen Brain Atlas (*67*). These marmoset-based gene weights effectively recapitulated the spatial pattern of the human DMN, including the expanded frontal pole (Fig. 7, H and I). Quantitatively, the marmoset-based DMN gene set yielded significantly higher fitting scores in the human DMN compared with non-DMN regions (Fig. 7J; $P < 0.001$). These results support the notion of evolutionarily conserved gene expression underlying the DMN, comprising functionally distinct (*66*) but molecularly overlapping subnetworks, whereas the expanded human frontal pole likely enhances network connectivity to support more complex functions (*68*).

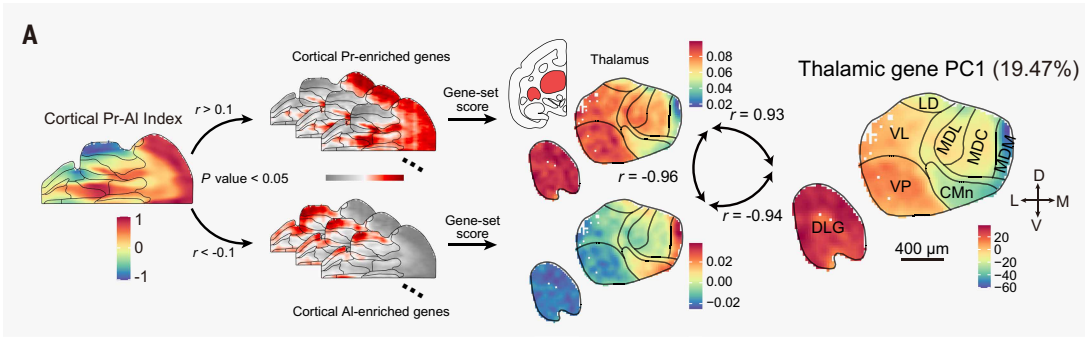
Shared molecular signatures in auditory cortices of humans and marmosets

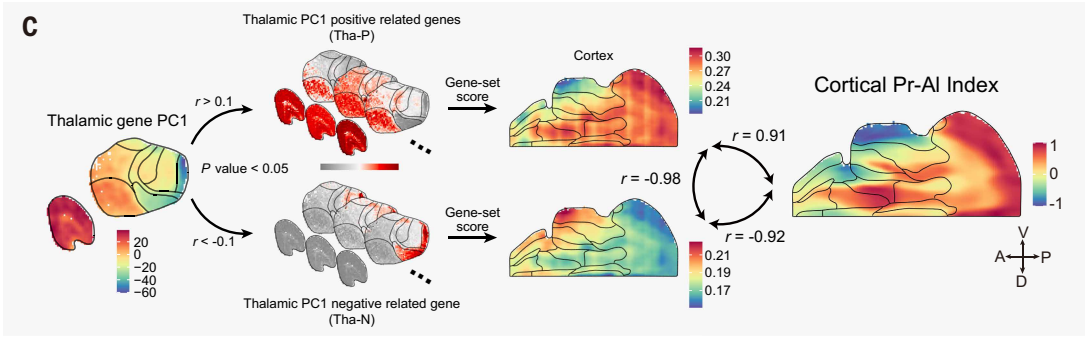
The comparative analysis of the DMN and frontal pole suggests that there are conserved gene expression patterns alongside species-specific variations. To further characterize such variations within the conserved Pr-Al gradient, we examined genes consistently correlated positively (Pr-shared) or negatively (Al-shared) with the Pr-Al index across marmosets, macaques, and humans. To facilitate interspecies comparison, we merged existing cortical parcellations for humans (*69*), macaques (*70*), and marmosets (*40*) into 15 broad regions (fig. S14A and table S8A). Using spatial transcriptomic (macaque and marmoset) and human Allen Brain Atlas microarray data (*6*, *67*, *71*), we identified Pr- and Al-shared genes and determined their regional enrichment across species (fig. S14, A and B, and table S8A).

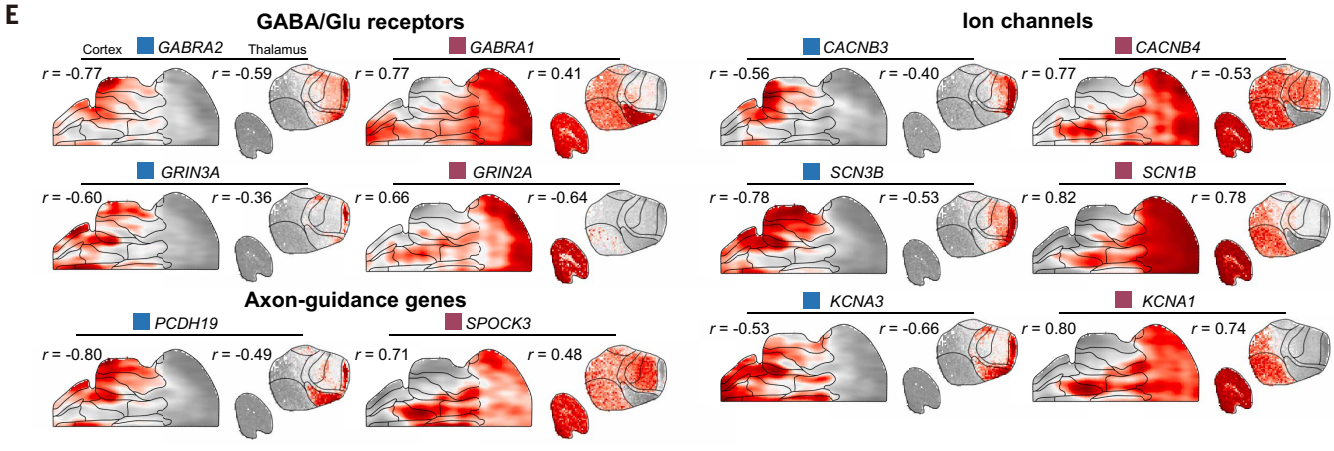
Although Pr- and Al-shared genes showed largely conserved expression across species in most cortical regions (Fig. 8A), species differences emerged. The marmoset somatosensory cortex (SS) exhibited a distinct gene expression profile, potentially reflecting its fewer subdivisions and less defined functional boundaries compared with those of humans and macaques (*72*, *73*). Furthermore, prefrontal areas (VLPFC, DLPFC, and MPFC) exhibited greater transcriptomic similarity between humans and macaques (Fig. 8A and fig. S14C). Gene ontology (GO) enrichment analysis linked these human-macaque shared genes to neuronal projection development, synaptic plasticity, and cognition (fig. S14, D and E), consistent with the more developed prefrontal cortex in humans and macaques.

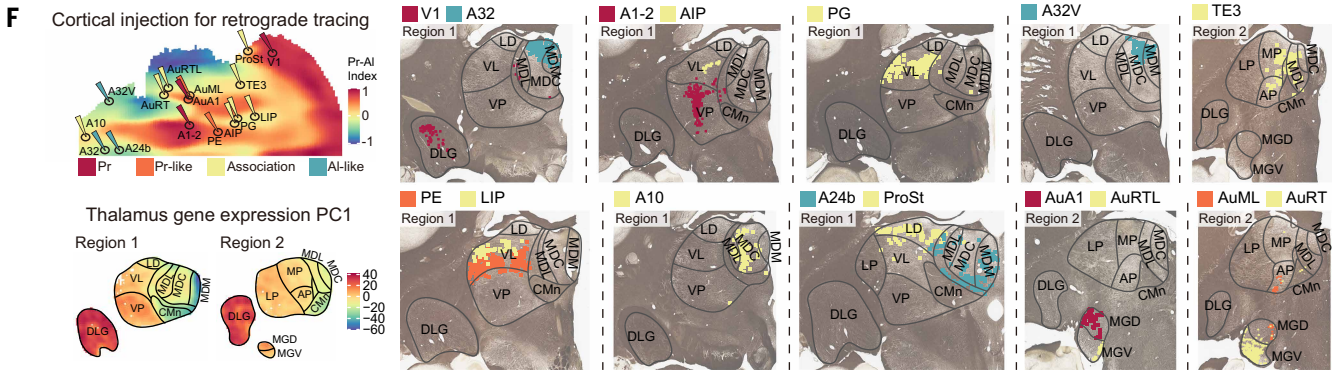
Although the macaque is phylogenetically closer to humans, marmosets shared more region-enriched shared genes in the auditory cortex with humans than macaques did (Fig. 8A). These shared genes were associated with neural signaling and circuit plasticity, such as GO terms related to ion channel activity and axonogenesis (fig. S14F). This similarity peaked for Pr-shared genes in core auditory areas (AuA1, AuR, and AuRT) and Al-shared genes in the superior temporal gyrus (STG) (Fig. 8B), aligning with the Pr-Al index map. Analysis of all Pr- and Al-shared genes revealed that human-marmoset similarity was most pronounced in auditory core regions (Fig. 8C and table S8B). Hierarchical clustering placed human AuA1 closest to marmoset core regions (Fig. 8D), and spatial transcriptomic maps confirmed high expression of these shared genes in marmoset AuA1, but not in that of the macaque (Fig. 8E and fig. S14, G and H).

To investigate the cellular basis for auditory cortex similarities, we performed cross-species integrative clustering of AuA1 neurons (Fig. 8F









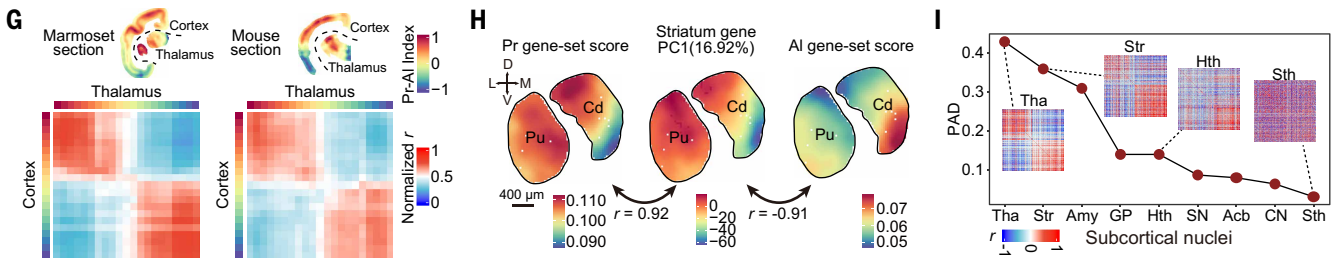


Fig. 5. Coordinated gene expression patterns in the cortex and the thalamus. (**A**) Thalamic expression of gene sets defined in the cortex. (Left) Identification of Pr-enriched

and AI-enriched gene sets based on the Pearson correlation coefficients with the cortical Pr-AI index ($P < 0.05$, $|r| > 0.1$). (Middle) Expression amounts for cortical Pr- and AI-enriched gene sets in the thalamus. (Right) Correlation of Pr- and AI-enriched gene expression pattern in the thalamus with the first principal component (PC1) of thalamic gene expression. DLG, dorsolateral geniculate nucleus; VP/VL, ventral posterior/lateral nucleus; LD, laterodorsal nucleus; MDL/MDC/MDM, lateral/central/medial division of the mediodorsal nucleus. (**B**) 3D spatial visualization of PC1 of thalamic gene expression. (**C**) Cortical expression of gene sets defined in the thalamus. (Left) Identification of genes positively (Tha-P) and negatively (Tha-N) correlated with thalamic gene expression PC1 ($P < 0.05$, $|r| > 0.1$). (Middle) Expression amounts for Tha-P and Tha-N genes in the cortex. (Right) Correlation of Tha-P and Tha-N gene expression amounts in the cortex with the cortical Pr-AI index map. (**D**) Overlap between cortical Pr- and AI-enriched genes and thalamic PC1-correlated genes (Tha-P and Tha-N). (**E**) Coordinated spatial expression of example genes encoding GABA or glutamate receptors, ion channels, and axon-guidance molecules in the cortex and thalamus. (**F**) Thalamocortical connectivity revealed by retrograde tracing. (Left) Cortical retrograde tracer injection sites color-coded based on the Pr-AI index. (Right) Retrogradely labeled neurons in the thalamus after injections into cortical areas V1, A32, A1-2, AIP, PG, A32V, TE3, PE, LIP, A10, A24b, ProSt, AuA1, AuRTL, AuML, and AuRT. Labeled neurons are color-coded according to injection sites. (**G**) (Top) Pr-AI index maps for the cortical and thalamic regions of the section analyzed. (Bottom) Heatmaps showing the gene expression correlation (normalized correlation) between cortical and thalamic parcels binned along their respective Pr-AI axes (21 bins each) for the corresponding marmoset (left) and mouse (right) sections. Additional examples are shown in fig. S11C. (**H**) Spatial distribution of Pr- and AI-enriched gene-set scores in the dorsal striatum, shown alongside the PC1 of overall striatal gene expression. (**I**) The PAD value, reflecting the difference between Pr- and AI-enriched gene spatial expression patterns (see Materials and methods), is plotted for subcortical regions, including the amygdala (Amy), globus pallidus (GP), hypothalamus (Hth), substantia nigra (SN), nucleus accumbens (Acb), cerebellar nuclei (CN), striatum (Str), and subthalamic nucleus (Sth). The inset heatmaps show the correlation matrix of the top 500 Pr and top 500 AI genes for select regions (Tha, Str, Hth, and Sth). Spatial expression patterns of Pr and AI genes of each region are shown in fig. S11H.

and fig. S15A) using published snRNA-seq data from humans (*32*) and marmosets (*16*) alongside newly generated macaque data. We identified eight conserved glutamatergic and seven GABAergic cell subclasses across the three species based on known marker genes (Fig. 8G). Hierarchical clustering of their average expression profiles revealed that most human and marmoset glutamatergic subclasses formed species-paired clusters, separating from macaque homologs (except for a subset of L5/6 and L6 neurons) (Fig. 8H). This indicates higher overall transcriptomic similarity between human and marmoset AuA1 glutamatergic neurons compared with those in macaques. Conversely, GABAergic subclasses predominantly clustered by conserved cell type identity rather than by species, although subtle similarities often positioned human and marmoset subclasses closer to each other than to macaques.

Differentially expressed gene (DEG) analysis across homologous glutamatergic subclasses identified 7996 DEGs (table S8C), with a core set of 4469 DEGs conserved across all three species (Fig. 8I). A substantial divergence emerged in humans, having 1888 species-specific DEGs, outnumbering those specific to macaques ($n = 117$) or marmosets ($n = 220$), which indicates molecular specialization in the human AuA1 (fig. S15B). Pairwise comparisons revealed 984 shared human-marmoset DEGs, substantially exceeding the overlaps between humans and macaques ($n = 269$) or between macaques and marmosets ($n = 49$). These patterns, replicated in our Stereo-seq data (fig. S15C), confirmed that human AuA1 glutamatergic neurons exhibit greater transcriptomic similarity to those of marmosets than to those of macaques.

We next examined the spatial distribution of the 984 shared human-marmoset DEGs (H&C DEGs) (table S8C) using our data and published macaque and human Stereo-seq data (*6*, *74*). Spatial mapping revealed that H&C DEGs were highly expressed in the human and marmoset AuA1 but not in that of macaques (Fig. 8J and fig. S15D). Layer-specific H&C DEGs were identified across all cortical layers (Fig. 8K, fig. S15E, and table S8D), including layer 4, the target of medial geniculate thalamic input (*75*), which suggests a role in early cortical auditory processing. Layer 2 ($n = 192$) and layer 3 ($n = 110$) contained the highest number of enriched genes, likely reflecting their role as key layers for intracortical auditory processing. Functional enrichment analysis linked these layer-enriched H&C DEGs to axonogenesis, synapse maturation, and ion transport (fig. S15F). For example, the neurotransmitter-associated gene *NSG2* was enriched in L2 and L3, and the neurodifferentiation-associated gene *RORA* was enriched in L4 of humans and marmosets (Fig. 8, L and M, and fig. S15G). In conclusion, we identified an unexpected molecular convergence between human and marmoset auditory cortices, suggesting a potential convergent evolution that likely reflects shared auditory processing mechanisms.

Discussion

In this study, we integrated spatial transcriptomics, neuroimaging, and retrograde tracing data to investigate the organizational principles of the primate brain. The smooth marmoset cortex facilitated this multimodal analysis, revealing an opposing molecular gradient governing cortical organization. This finding reconciles conflicting hypotheses regarding cortical expansion. The dual-origin hypothesis posits that the cortex expanded from two allocortical poles, the hippocampus (archicortex) and the piriform cortex (paleocortex), toward the most specialized primary sensory areas (*1*). Our findings provide a molecular basis for this view, demonstrating that these ancient regions function as a collective "Al" pole in the opposing-gradient axis. Conversely, other hypotheses propose that early-specified primary sensory areas act as foundational anchors of cortical expansion (*2*, *3*). Our "Pr" gradient, anchored in primary sensory cortices, supports this model. Thus, our study suggests that these theories are not mutually exclusive but describe the opposing ends of a single, dominant organizational axis. Furthermore, the Pr-Al axis refines the standard sensory-association model by revealing the molecular heterogeneity of the association cortex, which comprises distinct allocortical-like areas and gradient intersection zones. The gradient intersection likely fosters functional diversity in the expanded cortex, particularly primate higher-order association areas (*76*).

The opposing-gradient model offers a refined view of the molecular and cellular basis of cortical parcellation. Many established boundaries are marked by sharp, coordinated changes involving diverse cell types and genes, often coinciding with gradient intersections. These sharp transitions support the view of cortical areas as discrete domains with distinct molecular identities (*77*). However, not all boundaries exhibit such broad molecular shifts; instead, some are defined by localized changes restricted to specific cell types or genes, which suggests that these elements play pivotal roles in area specialization. The distinction between sharp and gradual transitions may indicate different stages of evolutionary specialization, consistent with the hypothesis that more recently developed areas can have interdigitating or gradual borders at some points of their evolutionary differentiation (*78*). Thus, these cell type and gene expression transitions offer additional features beyond traditional methods for cortical parcellation, potentially linking anatomical boundaries to specific molecular and cellular functions.

The opposing gradients are mirrored in the thalamus, aligning with thalamocortical connectivity. Whereas intrinsic thalamic gradients are known to correlate with connectivity (*45*–*47*), our findings extend this organization by demonstrating cross-structure molecular coordination. This molecular coupling between the cortex and thalamus is stronger in marmosets than in mice, suggesting a primate-specific enhancement. We hypothesize that this enhancement stems from a distinct primate developmental timeline, specifically the early arrival and prolonged interaction of thalamic axons with the developing cortex (*49*). Together with intrinsic cortical mechanisms (*79*), this interaction may influence the proliferation and specification of cortical progenitor

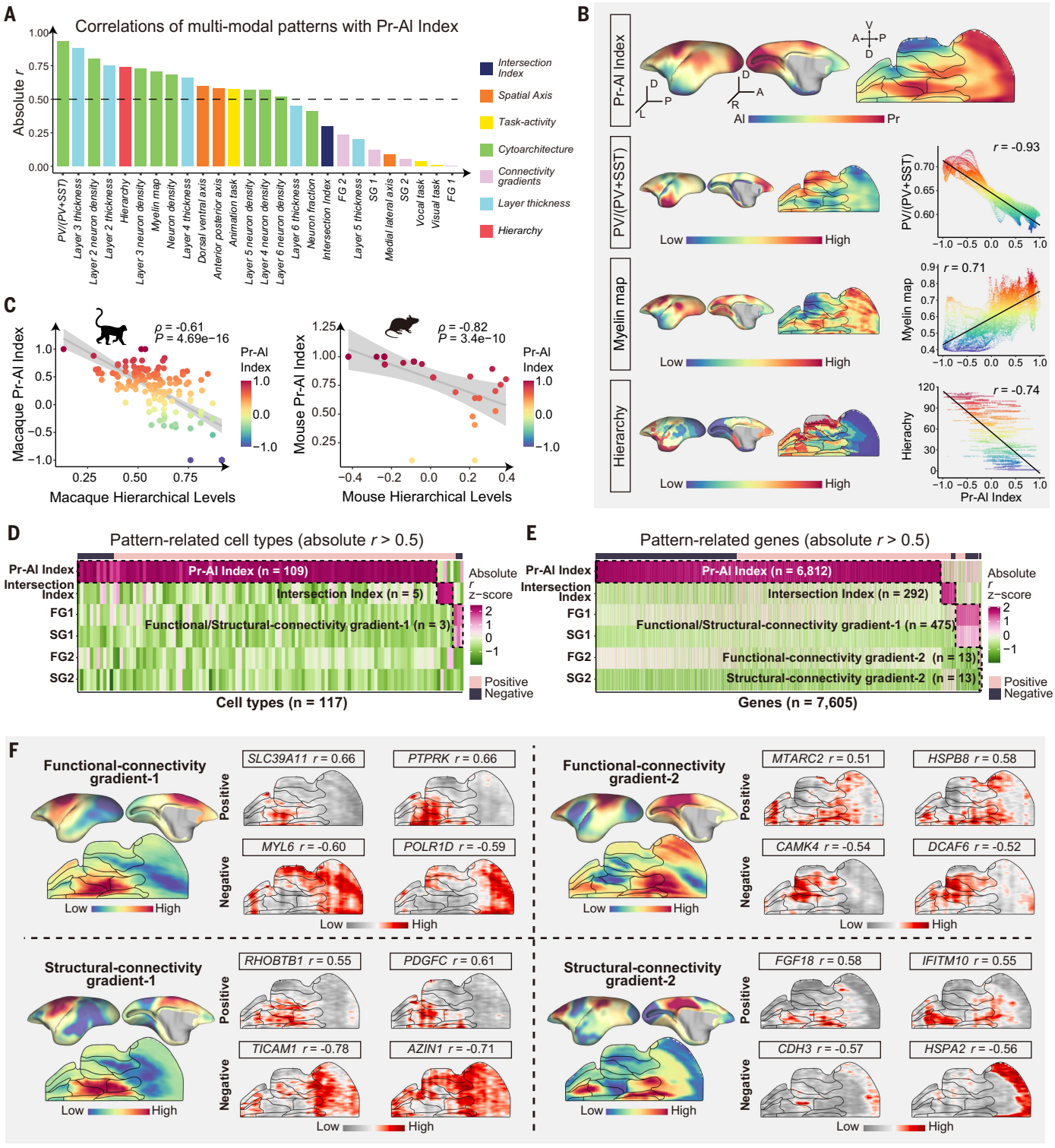


Fig. 6. Correlation between Pr-AI index and multimodal cortical patterns. (**A**) Bar plot showing the absolute Pearson correlation coefficients (|*r*|) between the Pr-AI index and various multimodal cortical patterns (table S1D). Dashed line indicates |*r*| = 0.5. (**B**) Spatial maps of the Pr-AI index and selected highly correlated patterns from (A) on the 3D cortical surface and segment-based flat map representations. More patterns are shown in fig. S12A. (**C**) Spearman correlation (ρ) of hierarchical levels in mice (*53*) and macaques (*54*,*55*) with the Pr-AI index of corresponding species. (**D** and **E**) Heatmaps show cell types or genes with |*r*| > 0.5 with at least one pattern: Pr-AI index, intersection index, functional connectivity gradients (FG1/FG2), and structural connectivity gradients (SG1/SG2). Cell types or genes are clustered according to the pattern with which they show the highest absolute correlation. The number of cell types or genes in each cluster is indicated. (**F**) Spatial expression for the example genes exhibiting the strongest positive or negative correlations with functional-connectivity gradients (top) or structural-connectivity gradients (bottom) (*23*).

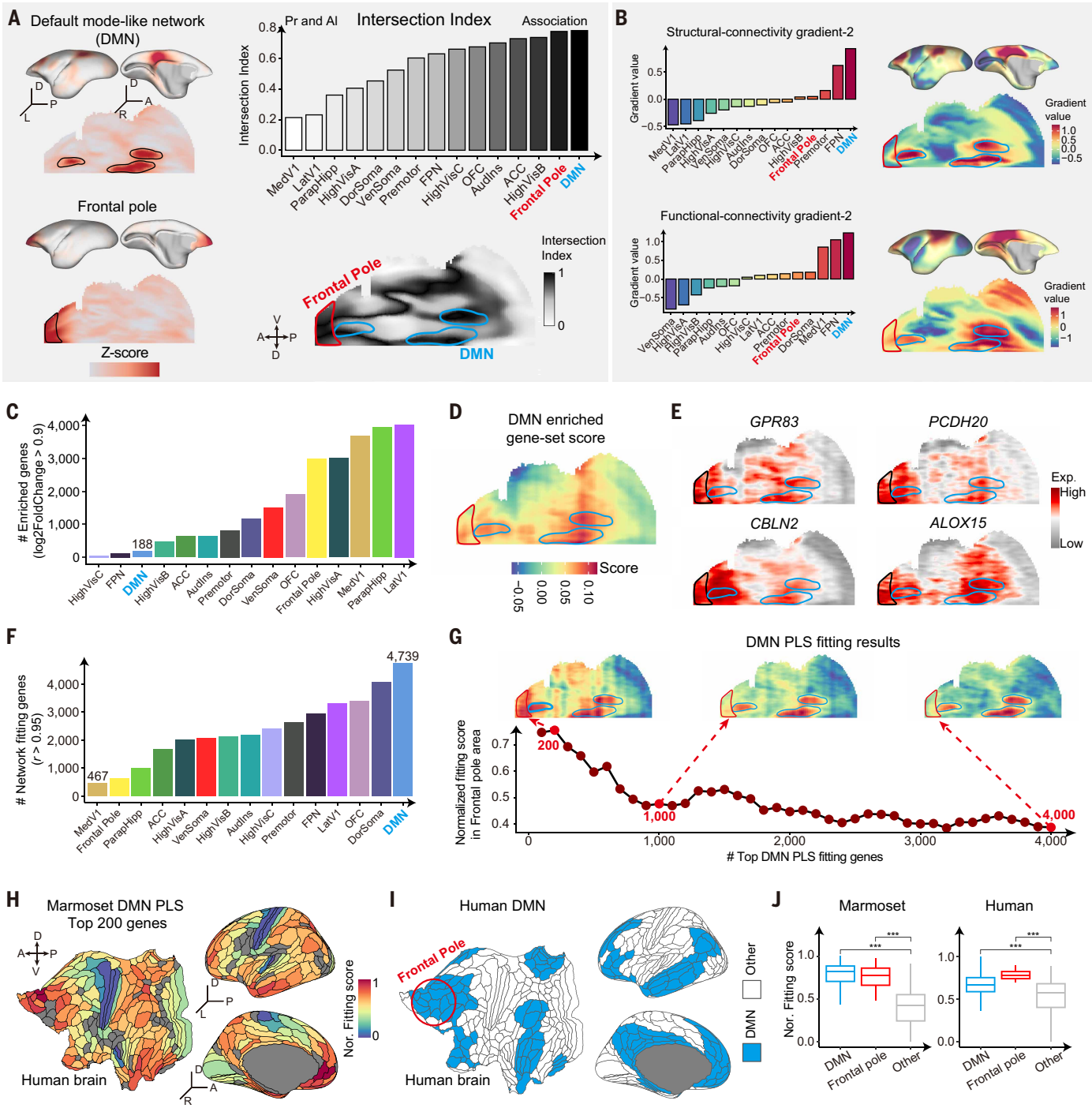


Fig. 7. Transcriptomic characteristics of functional brain networks. (**A**) Spatial patterns and the Pr-AI intersection-index values of different brain networks. (Left) Spatial maps of two example ICA-derived functional networks (DMN and frontal pole) from resting-state fMRI (*15*), shown on the cortical surface and segment-based flat map. See fig. S13A and table S7A for all networks. (Right top) Pr-AI intersection index for each network. (Right bottom) The intersection-index map highlighting the relative positions of the DMN (blue) and frontal pole (red) near the Pr-AI gradient's apex. (**B**) Positioning of brain networks on the second structural (SG2; top) and functional (FG2; bottom) connectivity gradients. (Left) Bar plots show the gradient value for each of the 15 networks. (Right) Spatial maps highlight the positions of the DMN and frontal pole on connectivity gradient maps. (**C**) Number of genes significantly enriched ($\log_2$ fold change > 0.9, $P < 0.05$) in each of the 15 networks (table S7B). (**D**) Spatial map showing the gene-set score, representing the aggregate expression pattern of DMN-enriched genes identified in (C). (**E**) Spatial expression patterns of four example genes significantly enriched in the DMN. Note that these genes are also highly expressed in the frontal pole. (**F**) Number of genes required for robust ($r > 0.95$) reconstruction of each network's spatial pattern using PLS regression (table S7, C and D). (**G**) Impact of gene number on PLS fitting. (Top) Spatial maps illustrating the PLS-based reconstruction (fitting results) of the DMN using increasing numbers of top-ranked PLS-weighted genes (200, 1000, and 4000). (Bottom) Plot showing the normalized fitting score for the frontal pole as a function of the number of top DMN-fitting genes included in the PLS model. (**H**) Projection of the marmoset DMN PLS model onto the human brain. The map shows normalized fitting scores across human cortical regions, generated by applying the PLS weights derived from the top 200 marmoset DMN-fitting genes (table S7E) to homologous human gene expression data (*67*). Scores are visualized on the human cortical flat map and 3D surface views. (**I**) Spatial map of the human DMN derived from resting-state fMRI. (**J**) Comparison of

normalized fitting scores for the DMN (excluding frontal pole), the frontal pole, and other cortical regions ("Other") in both marmoset and human cortices. Individual segments (marmoset) or regions (human) were treated as independent samples. For marmosets, $n = 1753$ (DMN), 1753 (frontal pole), and 14,133 (other) segments; for humans, $n = 55$ (DMN), 6 (frontal pole), and 76 (other) regions. Scores were calculated using a model based on 138 marmoset-human homologous genes (from the top 200 marmoset DMN PLS gene set). Normalized fitting scores represent the sum of products of PLS weights and gene expression for these genes, scaled to [0, 1]. Statistical significance was determined by one-way ANOVA with post hoc two-sided Student's *t* test. ***$P < 0.001$.

cells (*44*, *80*), forging the tightly coupled gradients shared by both structures. Because these gradients are enriched in genes crucial for neuronal communication (Fig. 5E), this coordination may be essential for regulating information integration within reciprocal thalamocortical loops across the cortical hierarchy (*81*).

The opposing-gradient model illuminates the molecular architecture of large-scale functional networks, particularly the DMN. Situated at the zone of maximum gradient intersection, the marmoset DMN exhibits distinctive molecular complexity, requiring a substantially larger gene set for reconstruction compared with other networks. This aligns with the view of the DMN as a composite of functionally distinct, interdigitated subnetworks rather than a monolithic entity (*66*). Our findings contribute to elucidating the relationship between the frontal pole and the DMN (*21*, *61*, *82*, *83*). Despite weak functional connectivity, the marmoset frontal pole shares a pronounced transcriptomic identity with the DMN. This suggests that the frontal pole has a fundamental DMN molecular signature independent of connectivity, supporting the idea that its evolutionary expansion in humans (*68*) represents the connectomic integration of a preexisting, molecularly specified DMN.

The molecular similarity between the human and marmoset auditory cortices potentially provides a compelling example of convergent evolution. Despite large evolutionary divergence, both species independently developed sophisticated social vocal repertoires (*84*), which suggests that the cortical molecular gradients may adapt to shared selective pressures. Like humans, marmosets exhibit rich vocal behaviors, including vocal learning, turn-taking, and social accommodation (*85*–*87*). Coupled with neurophysiological specializations for complex auditory processing (*88*, *89*), these shared molecular features likely support the complex neural mechanisms of vocal interaction. Because this similarity peaks in the auditory core regions, adaptation likely arose at an early processing stage. Thus, the marmoset might represent a valuable model for studying the molecular basis of vocal communication and potentially the origins of language.

We acknowledge several limitations in the study. First, because we focused on postnatal development and adulthood, future prenatal profiling is essential to elucidate the mechanisms underlying the emergence of these gradients (*90*). Second, our sample size precluded the assessment of interindividual or sex-related variability. Third, our data inform cortical boundaries; however, establishing a new parcellation atlas requires advanced computational frameworks and functional validation of boundaries. Finally, our connectivity analysis lacked single-cell resolution. Emerging tools such as barcode-based tracing (*91*), cell type–specific viral approaches (*92*), and optogenetic fMRI (*93*) offer promising avenues for systematically linking gene expression with specific neuronal connections.

Materials and methods

Animals

All animal studies and procedures were approved (ION-2019011) by the Animal Care and Use Committee of the Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Chinese Academy of Sciences (Shanghai, China). All animals were individually housed under CEBSIT's institutional guidelines, with a 12-hour light-dark cycle (7 a.m. to 7 p.m.), a humidity- and temperature-controlled environment (27° to 30°C), and ad libitum access to food and water. Seven adult marmosets (*C. jacchus*; 5 female, 2 male; average age: 2 years and 6 months) received retrograde neuronal tracing injections. Four postnatal day 0 marmosets (P0; 2 males, 2 females), two postnatal day 32 marmosets (P32; 1 male, 1 female), two 3-month-old marmosets (P3M; 1 male, 1 female), one 4-year-old male marmoset (marmoset I), and one 5-year-old female marmoset (marmoset II) were used for Stereo-seq and snRNA-seq analyses. One 19-year-old female macaque (*Macaca fascicularis*) was used for snRNA-seq of auditory regions. Detailed information for these animals can be found in table S1A.

Retrograde neuronal tracing

Marmoset CJ173 (male, 2 years) was injected with four tracers using a protocol described in (*14*): Fast blue (FB) spanning somatosensory cortical areas 3b and 1/2 (A1/2), diamidino yellow (DY) restricted to the supragranular layers of primary motor cortex (cytoarchitectural area 4a/b, A4ab), CTB-green in the rostral posterior parietal cortex centered on the anterior intraparietal area (AIP), and CTB-red near the rostral border of parietal area PE.

Marmoset CJ178 (female, 2 years and 2 months) received FB in the dorsal bank of the calcarine sulcus (attempting peripheral V1 representation), DY in medial prefrontal area 32 (A32), and both CTB-green and CTB-red in the caudal frontopolar cortex (area 10, A10).

Marmoset CJ203 (female, 2 years and 11 months) received DY in the posterior parietal area PG (spanning most layers).

Marmoset CJ148 (female, 1 year and 7 months) received a single, very restricted DY injection in area 32V (A32V) of the right hemisphere with no tracer leakage up the track.

Marmoset CJ215 (female, 3 years and 5 months) received two tracer injections: FB centered on the middle layers of the rostral anterior cingulate area 24b (A24b); DY centered on area prostriata (ProSt), with some tracer deposited in the adjacent white matter.

Marmoset CJ187 (female, 2 years and 7 months) received two effective injections: CTB-green on the lateral rostral temporal lobe, crossing from area AuRTL into the rostral parabelt; and CTB-red in the lateral part of primary auditory area (AuA1).

Marmoset CJ180 (male, 2 years and 8 months) received four tracer injections in the temporal lobe: a large FB injection centered in area AuRT; DY in the dorsal part of area TE3; CTB-green centered in the ventral auditory cortex (AuML); and CTB-red in the inferior temporal cortex, preliminarily located in area TE3.

Histological sections stained for Nissl substance, myelin, and cytochrome oxidase revealed cytoarchitectural boundaries. Injection sites were mapped relative to neuroanatomical landmarks and validated against laminar staining patterns. CJ173 exhibited accidental tissue loss in the ventral temporal cortex during sectioning. CJ178 showed FB-induced necrosis in the occipital white matter that may mask connections with V2. No major technical issues were noted in the other cases.

Tracer spread analysis indicated: potential FB invasion of A1/2 in CJ173; minimal DY halo spread toward prefrontal area 8b (A8b) in CJ178; in CJ203, potential DY encroachment into PFG; the restricted nature of the DY injection in CJ148; minimal or absent FB halo contamination of areas 8b and 32 in CJ215; in CJ187, a CTB-green injection that did not involve infragranular layers, and a CTB-red injection with no white matter involvement and only possible invasion of the lateral belt; and in CJ180, a CTB-red injection with possible invasion of adjacent polysensory cortex.

Tissue collection for Stereo-seq and snRNA-seq

The animals were anesthetized via intramuscular injection with a combination of tiletamine hydrochloride, zolazepam hydrochloride (25 mg/kg), and xylazine hydrochloride (20 mg/kg). After anesthesia,

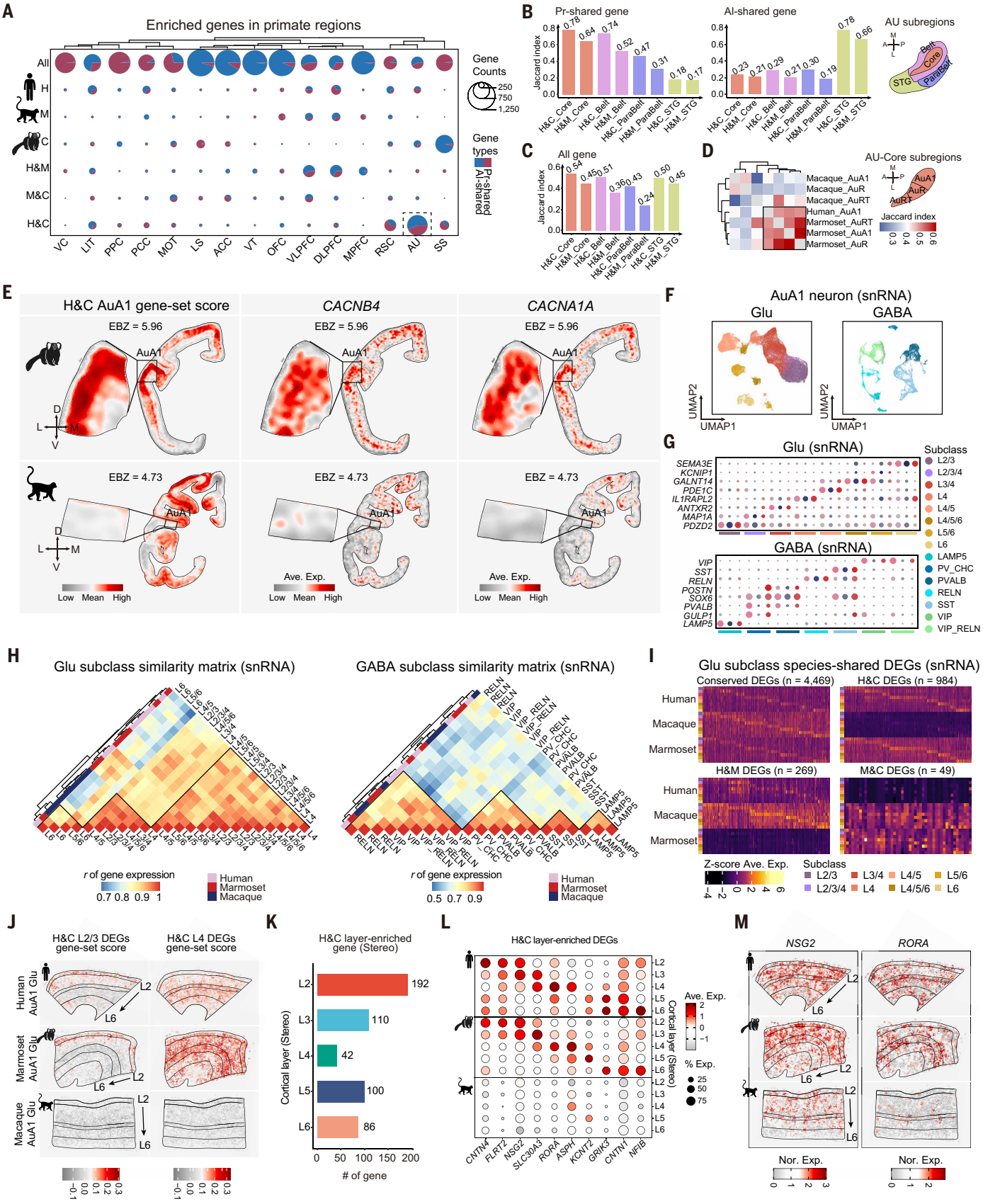


Fig. 8. Transcriptomic similarity in the auditory cortex between humans and marmosets. (**A**) Pie plots showing the proportion and number of regionally enriched Pr-shared ($r > 0$ with Pr-AI index across species; blue) and AI-shared ($r < 0$; red) genes. Rows define the species for enrichment: marmoset (*C. jacchus*; C), macaque (*M. fascicularis*; M), human (*Homo sapiens*; H), all species (all), or shared pairs (H&M, M&C, and H&C). See table S8A for full names of brain region abbreviations and the Pr- and AI-shared gene lists.

(**B**) Jaccard index comparing species-pair similarity based on regionally enriched Pr-shared or AI-shared genes within different auditory subregions (Core, Belt, ParaBelt, and STG regions; see table S8B for region details). (Left) Pr-shared gene. (Right) AI-shared gene. (**C**) Jaccard index as in (B) but using all regionally enriched genes (Pr- and AI-shared combined). (**D**) Hierarchical clustering of Jaccard index matrix, comparing species-pair similarity based on enriched genes within auditory core subregions (human AuA1; macaque and marmoset AuA1, AuR, and AuRT). Heatmap shows Jaccard index values. (**E**) Comparison of gene expression patterns in AuA1 cortex across species. (Left) Spatial expression map (Stereo-seq) showing the aggregated score for a gene set enriched in human and marmoset (H&C) AuA1. (Right) Spatial expression map (Stereo-seq) for H&C-enriched gene examples. More gene examples are provided in fig. S14H. (**F**) UMAP showing the integrated clustering of AuA1 glutamatergic neurons (Glu) and GABAergic neurons (GABA) for the three species. (**G**) Dot plot showing the expression of selected subclass marker genes in the integrated snRNA-seq data. (**H**) Heatmaps showing the cross-species similarity (Pearson correlation coefficient) of homologous neuronal subclasses, calculated from average snRNA-seq expression. The dendrograms illustrate the hierarchical clustering of these subclasses. (**I**) Heatmaps showing *z* score average expression of DEGs in homologous glutamatergic subclasses, grouped by conservation pattern: conserved (across all three species) or species-pair shared (H&C, H&M, and M&C). Species-specific DEGs are shown in fig. S15B. (**J**) Spatial maps (Stereo-seq) in AuA1 showing aggregated scores for H&C shared DEGs [from (I)] with layer-specific patterns. (Left) L2/3 subclass DEGs. (Right) L4 subclass DEGs. (**K**) Bar chart quantifying the number of H&C shared DEGs significantly enriched in specific cortical layers based on the Stereo-seq data ($\log_2$ fold change > 0.25, $P < 0.05$). (**L**) Dot plot showing expression characteristics of selected H&C layer-enriched DEGs in AuA1 Stereo-seq data across the three species. (**M**) Spatial expression maps (Stereo-seq) in AuA1 for selected H&C layer-enriched gene examples from (L) across the three species.

the brains were rapidly perfused first with room-temperature artificial cerebrospinal fluid (ACSF) saturated with 95% O_2 and 5% CO_2, and then with 4°C bubbled ACSF at 100 ml/min. The whole brains of 10 marmosets were dissected, and using a stereotaxic micromanipulator (SMM-200, Narishige), each brain was divided into two hemispheres. The left hemispheres were embedded in 4°C OCT (Sakura no. 4583). Subsequently, they were snap-frozen in 4°C OCT within metal molds on dry ice, and finally stored at −80°C. The entire procedure from anesthesia to storage was completed within 30 min. Tissue cryosection was carried out at −20°C using precooled tools. From adult marmoset I, 125 coronal sections, each 10-μm thick with a 250-μm spacing, were obtained. From adult marmoset II, 27 parasagittal sections, each 10-μm thick, were collected for biological replication. The remaining sections from Marmoset I and Marmoset II were reserved for staining and snRNA-seq. From postnatal marmosets, using similar procedures, one coronal section was obtained from each of two anesthetized animals at each time point. For snRNA-seq at P0, we performed brain-wide sampling from two marmosets, including the prefrontal cortex (PFC), cingulate cortex (CC), medial and lateral parietal cortices (mPPC, lPPC), motor cortex (MOT), somatosensory areas (SS), lateral sulcus (LS), auditory cortex (AU), ventral temporal areas (VT), lateral and inferior temporal areas (LIT), and visual cortex (VC). At P32 and P3M, the PFC, AU, and entorhinal cortex (Ent) were sampled, whereas the LIT was sampled only at P32. For adult marmosets, a total of 24 samples were collected, covering all major anatomical regions of the brain. For the macaque, the auditory core region was dissected for snRNA-seq. All snRNA-seq samples were stored at −80°C. Table S1, A and B, details the animal and sampling information for both Stereo-seq and snRNA-seq.

Stereo-seq library building and sequencing

After cryosection, in a dry environment, each tissue section was first flattened on a −20°C metal plane of the cryosection platform and then carefully adhered to a precooled Stereo-seq chip to prevent air bubbles and tissue folding and incubated at 37°C for 5 to 8 min. Then, they were fixed in methanol (Sigma, 34860, precooled for 30 min at −20°C; 40 ml methanol added in a 10 cm dish for each section) and incubated at −20°C for 30 min. Next, the tissue section on the chip was stained with a mixture of nucleic acid reagent (Invitrogen, Q10212) and conA Rhodamine (Vector, RL-1002) for 5 min, washed with 0.1x SSC buffer [Ambion, AM9770; containing 0.05 U/ml ribonuclease (RNase) inhibitor], and section images were captured using Zeiss Axio Scan Z1 microscope [at enhanced green fluorescent protein (EGFP) wavelength, 10-ms exposure]. After that, the tissue sections were permeated by incubating in 0.1% pepsin (Sigma, P7000) prewarmed at 37°C for 3 min in 0.01M HCl buffer (pH 2) at 37°C for 15 min, and then washed with 0.1x SSC buffer (containing 0.05 U/ml RNase inhibitor) to remove pepsin, during which RNAs were released from the permeated tissue and captured by the Stereo-seq chip. RNAs were reverse transcribed for 2 hours at 42°C using SuperScript II [Invitrogen, 18064-014, 10 U/ml reverse transcriptase, 1 mM deoxynucleotide triphosphates (dNTPs), 1 M betaine solution polymerase chain reaction (PCR) reagent, 7.5 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 2 U/ml RNase inhibitor, 2.5 mM Stereo-seq–TSO, and 1x First-Strand buffer]. After reverse transcription, tissue sections were washed twice with 0.1x SSC buffer and digested with tissue removal buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 55°C for 30 min. The cDNA-containing chips were then subjected to Exonuclease I (NEB, M0293L) treatment for 1 hour at 37°C and were washed once with 0.1x SSC buffer. The resulting cDNAs were amplified with Hot Start DNA Polymerase (QIAGEN) with 0.8 mM cDNA-PCR primer. The PCR reaction protocol was: first incubation at 95°C for 5 min, 15 cycles at 98°C for 20 s, 58°C for 20 s, 72°C for 3 min, and a final incubation at 72°C for 5 min. The concentrations of the resulting PCR products were quantified by Qubit double-stranded DNA (dsDNA) HS assay kit (Invitrogen, Q32854).

A total of 20 ng of DNA was fragmented with in-house Tn5 transposase at 55°C for 10 min. The reaction was stopped by adding 5 ml of 0.02% SDS and gently mixing at 37°C for 5 min. Fragmented products were amplified with 25 ml of fragmentation product, 1×KAPA HiFi Hotstart Ready Mix, and 0.3 mM Stereo-seq–Library-F primer, 0.3 mM Stereo-seq–Library-R primer in a total volume of 100 ml with the addition of nuclease-free H_2O. The amplification protocol was 1 cycle of 95°C for 5 min, 13 cycles of 98°C for 20 s, 58°C for 20 s, and 72°C for 30 s, and 1 cycle of 72°C for 5 min. PCR products were purified using 0.6× VAHTSTM DNA Clean Beads, used for DNB (DNA Nano Ball) generation, and finally sequenced on the DNBSEQ TM T10 sequencing platform (MGI, Shenzhen, China) with 50–base pair (bp) read1 and 100-bp read2.

snRNA-seq library building and sequencing

First, we prepared nucleus suspensions with a prechilled Dounce homogenizer (Sigma, no. D8938-1SET) and filtered the homogenate with a cell strainer (Sigma, no. CLS431752-50EA). Then, we used the DNBelab C Series High-throughput Single-Cell RNA Library Preparation Kit (MGI, no. 940-000047-00) to build snRNA-seq libraries, following the manufacturer's protocol. Through microfluidics, we combined them with gel beads [bearing unique molecular identifiers (UMIs) and cell barcodes] to form droplets. After in-droplet transcripts capture and reverse transcription, we broke the emulsion using a chemical demulsifier and collected the gel beads via centrifugation and filtration. Next, we did reverse transcription with SuperScript II, then synthesized the second-strand cDNA. Using KAPA HiFi Hotstart Ready Mix, we amplified the cDNA with a set PCR program. After that, we further amplified the products to get barcoded libraries. We quantified the libraries with the Qubit ssDNA Assay Kit (Thermo Fisher Scientific, no. Q10212) and sequenced them on the DNBSEQ T1 or DNBSEQ T10 at the China National GeneBank in Shenzhen, China.

Stereo-seq data preprocessing

The analysis of Stereo-seq raw data adhered to the Stereo-seq Analysis Workflow (SAW v4.1.1) (*94*). Initially, the CID sequence was aligned to

the whitelist. This step filters out reads with incorrect or low-quality barcodes. Next, STAR (bcSTAR v1.0.2) was used to align the cDNA sequences to the marmoset reference genome (GCA_009663435.2, Callithrix_jacchus_cj1700_1.1). Subsequently, reads with a mapping quality of less than 10, multiple alignments, and duplicate alignments were removed. The raw spot × gene expression matrix was then generated. All these SAW steps were performed in a wrapped pipeline in the DCS cloud platform (https://cloud.stomics.tech/). Subsequently, UMIs from segmented cells were aggregated to create the cell × gene expression matrix (see "Image-based single-cell segmentation using Stereo-seq data" section for details). Additionally, for analysis of gene expression, UMIs from adjacent 50×50 DNBs or 25×25 DNBs were aggregated to generate the bin50 and bin25 expression matrices. Genes expressed in fewer than three cells were removed from the matrix. The genes *ND6*, *COX3*, *COX1*, *ND5*, *ND4*, *ND2*, *ND4L*, *ATP8*, *CYTB*, *COX2*, *ND3*, *ATP6*, and *ND1* were used to calculate the proportion of transcripts mapping to mitochondrial genes. Cells with fewer than 50 genes, fewer than 100 UMIs, or with a mitochondrial gene proportion greater than 10% were excluded from subsequent analysis.

Registration of Stereo-seq data with MRI data

Spatial maps of total RNA based on the bin50 matrix, 25-μm resolution, were generated for all Stereo-seq sections. These maps were then converted into the NIFTI format. Subsequently, the converted images underwent processing with the AFNI software package (*95*). Specifically, the "3dzeropad command" in AFNI was used to enlarge the images, followed by manual rotation to correct their orientation. After that, the "3dresample" and "3drefit" functions of AFNI were applied. This ensured standardization of both the image orientation and the dimensional information stored within the image headers. The 3D volumetric MRI (T2) template (*9*) was resampled to achieve an isotropic resolution of 25 μm and then rotated to align with the spatial transcriptomic data (6° clockwise around the x axis and 1° clockwise around the z axis). Based on the anteroposterior axis location of the Stereo-seq sections, the most analogous cortical MRI slices were manually selected. The resulting 2D MRI slices and the aforementioned total RNA spatial maps served as templates for the subsequent registration of gene expression and cell type data with multimodal data. The "antsRegistration" method from the ANTs (*96*) software was used to perform both affine and nonlinear (SyN) registration from the total RNA spatial maps to the MRI template, thereby generating transformation matrices that ensured spatial correspondence between the two datasets. To improve registration accuracy, the registration used only the cortical voxels to calculate the transformation. The derived transformation matrices were subsequently applied to the entire brain section.

The 3D reconstruction of gene expression profile and cell type distribution

The previously generated transformation matrices were applied to register the spatial distributions of gene expression and cell type from the spatial transcriptomic space to the MRI space. For cell type registration, the bin5 map was used. Cell type density was estimated using the "ksdensity" method in MATLAB. Subsequently, all 2D slice data for gene expression and cell type density were integrated into a 3D spatial transcriptomic framework. To generate high-resolution maps, missing values within the cortical region were imputed using linear interpolation along the slice dimension (row-wise) via MATLAB's "fillmissing" function, with end values filled using the nearest-neighbor method. This process generated high-resolution 3D maps of gene expression and cell type distributions. These 3D maps were mapped to the surface template (*10*) and smoothed with a 1-mm full width at half maximum (FWHM) kernel within the cortical mask using the "wb_command" tool from the Connectome Workbench (*97*).

Registration of multimodal data to Stereo-seq data

The multimodal datasets included cortical parcellation, myelin map, awake resting-state fMRI time series and corresponding functional gradients, retrograde tracing data and associated structural gradients, and task-fMRI activation maps.

The cortical parcellation and myelin map were derived from the Marmoset Brain Mapping project (*10*). The cortical parcellation was constructed using the Paxinos atlas, resulting in a 3D MRI-based atlas encompassing 116 cortical regions. The myelin map was generated from the ratio of T1-weighted to T2-weighted templates (*51*), quantifying myelination across the adult marmoset neocortex.

Resting-state fMRI time series data and functional brain networks were obtained from the Marmoset Brain Mapping project version-4 (*15*). The data consisted of 39 adult marmosets (31 males, 8 females; mean age = 4 years) with 113 imaging sessions, 710 valid fMRI runs, and 12,117 min of total scan time. Fifteen functional brain networks were identified through ICA of resting-state fMRI data, comprising: ventral somatomotor (VenSoma), dorsal somatomotor (DorSoma), frontal pole, parahippocampus/temporal pole (ParaHipp), orbital frontal (OFC), auditory and insular cortex (AudIns), frontoparietal-like network (FPN), DMN, anterior cingulate cortex (ACC), premotor, and two visual-related networks: medial primary visual and MT/MST (MedV1) and lateral primary visual and MT/MST (LatV1), three hierarchical high-level visual networks: high-level 1 (HighVisA), high-level 2 (HighVisB) and high-level 3 (HighVisC) networks. Cortical retrograde tracing data were acquired from the Marmoset Brain Connectivity Atlas (*14*), encompassing 143 cerebral cortical tracer injections across 52 adult marmosets (31 males, 21 females; median age = 2.5 years). The structural and functional gradients were obtained from Tong *et al.* (*23*), whose study computed these gradients using diffusion map embedding methodology (*57*). Specifically, structural gradients were derived by applying this method to connectivity matrices from retrograde tracing data, whereas functional gradients were generated from resting-state fMRI connectivity matrices. Both gradients represented the principal axes of variation in their respective structural and functional connectivity patterns.

Task-evoked activation maps included three paradigms. Vocal task: cortical activation during auditory processing of conspecific vocalizations compared with silent baseline periods (*22*). Visual task: activation patterns during free-choice trials where marmosets selected between new and familiar visual targets after movie clip exposure, without external reinforcement (*25*). Frith-Happe animation task: neural correlates of social cognition assessed through gaze patterns and cortical activation during viewing of Frith-Happe animations, probing theory-of-mind processing in marmosets (*24*).

For datasets originally mapped to 3D cortical surfaces, such as task-evoked activation maps and connectivity gradients, surface-based information was first projected to the volumetric MRI template using the "wb_command." All multimodal MRI data were then coregistered to the volumetric MRI template via affine and nonlinear transformations implemented in the ANTs toolkit (antsRegistrationSyN.sh command). Leveraging the inverse transformations derived from the Stereo-seq and MRI Data Registration workflow (see Materials and methods, section "Registration of Stereo-seq data with MRI data"), the aligned multimodal MRI data were further warped into the spatial transcriptomic coordinate space. This step enabled the integration of multimodal MRI features within the spatial transcriptomic slice space for cross-modal correlation analyses.

Image-based single-cell segmentation using Stereo-seq data

A previous deep learning–based method developed by Tencent AI Lab (https://github.com/TencentAILabHealthcare/Cell_Segmentation_Fluorescence) (*6*) was used for single-cell segmentation in Stereo-seq data, which adopted the Mask R-CANN framework. This method primarily comprised two components: it first annotated cells in ConA images through an active learning process, and then trained a cell segmentation model for ssDNA images using labels transferred from the ConA images. To mitigate computational complexity, the staining

images were cropped into small patches for subsequent training. Manual correction was applied to refine the cell segmentation model until it attained a high accuracy. After obtaining the single-cell segmentation of Stereo-seq based on the ssDNA imaging data, we evaluated the performance of this automated single-cell segmentation method using manually annotated single cells from ssDNA images of the marmoset cortex as the ground truth. The results showed that the method achieved high recall and precision rates, as illustrated in fig. S3, C to E.

snRNA-seq data preprocessing

Leveraging the power of DCS cloud, the public data analysis pipeline (https://github.com/MGI-tech-bioinformatics/DNBelab_C_Series_scRNA-analysis-software) specific to MGI eLab C4 scRNA-seq platform (*98*) was used to process the raw data into a cell×gene matrix. In brief, we first used the "parse" function from PISA (https://github.com/shiquan/PISA) to extract cell barcodes (1 to 20 bp and 11 to 20 bp) and UMIs (21 to 30 bp) from the sequencing reads and placed them into the sequence identifier line. Subsequently, the cDNA sequences were aligned to the marmoset reference genome (GCA_009663435.2, Callithrix_jacchus_cj1700_1.1) using STAR (v2.7.3), with the corresponding GTF file as the annotation reference (*99*). To determine the actual number of cells, the "barcodeRanks" function from the R package DropletUtils (*100*) was used to identify the threshold value at the sharp drop point in the total UMI count distribution. Finally, barcodes were merged into single cells, and gene expression amounts were quantified using PISA to generate the cell × gene matrix for downstream analysis.

snRNA-seq annotation

For cell clustering and cell type annotation, we used the R package Seurat (v4.1) (*101*) along with custom scripts and parameter sets. Initially, cells with fewer than 1000 genes, or with a UMI-to-gene ratio less than 1.2, or with a mitochondrial gene percentage greater than 5% were discarded. Additionally, doublets were filtered out using the R package DoubletFinder (*102*). The gene expression data for each cell were normalized and scaled using the "SCTransform" function from Seurat, with 3000 top variable genes selected for downstream analysis. Then, PCA, nearest-neighbor graph construction, and clustering were performed on the SCT-scaled gene expression data using the Seurat functions "RunPCA," "FindNeighbors," and "FindClusters." We then categorized the cells into three groups based on class markers (level 1): glutamatergic neurons, GABAergic neurons, and nonneuronal cells. Subsequently, we further iteratively classified these three cell groups into higher-resolution subpopulations, resulting in 27 subclasses (level 2) and 206 cluster-level cell types (level 3). Using the "FindAllMarkers" function from the Seurat package, we performed DEG analysis between clusters. Genes with a $\log_2$ fold change > 1 and a false discovery rate (FDR) < 0.05 were retained as significant marker genes for each cluster. Based on the top 50 marker genes for each cluster, we calculated the Jaccard similarity between clusters. Clusters with a Jaccard similarity greater than 0.75 were merged.

For the machine learning–based cell type analysis, we used the top 50 DEGs as features. From each cluster, we randomly sampled 100 cells and split them into an 80% training dataset and a 20% validation dataset. This was implemented using the "tuneRF," "randomForest," and "predict" functions from the RandomForest R package (v.4.7.1.1). Finally, we identified and annotated specific cell clusters using the predetermined marker genes, ensuring a comprehensive and accurate cell type annotation for our dataset.

Stereo-seq data cell type annotation

To deconvolute the cell types of segmented cells from Stereo-seq data based on snRNA-seq cell types, we used DestVI (fig. S3G) (*29*). First, a preprocessing step was carried out on the snRNA-seq data. Cells with fewer than 50 genes or fewer than 100 UMIs were removed. To ensure a representative sample, 1000 cells were randomly sampled for each cell type present in the snRNA-seq dataset. The gene set of these sampled cells was intersected with the genes expressed across all Stereo-seq sections. Subsequently, the "CondSCVI" function was used to train a model (scmodel) on the preprocessed snRNA-seq data. The training process was configured with specific parameters: weight_obs set to True, n_latent at 4, n_layers at 2, n_hidden at 128, and max_epochs set to 250. After training the scmodel, it was transferred to analyze Stereo-seq data using the "DestVI.from_rna_model" function, generating a new model (stmodel). For this transfer, the parameters max_epochs were set to 500 and the learning rate (lr) to 0.001. Through the "st_model.get_proportions" function, the scmodel was effectively adapted to Stereo-seq data. As a result, each cell in the Stereo-seq dataset was assigned a probability distribution over potential cell type. To enhance the accuracy of cell type annotations, we calculated the second-order derivative of the probability distribution. This enabled us to determine appropriate thresholds for the distribution of each cell type. Finally, every single cell on the Stereo-seq map was assigned the cell type with the highest probability within its distribution.

In addition, we evaluated the similarity of cell type distributions across Stereo-seq sections. Owing to overclustering and high similarities among certain cell types, some clusters exhibited similar or complementary distribution patterns. Specifically, four interneuron clusters and four nonneuronal clusters were manually merged. After the merging process, we validated the distributions of all cell types across sections. This step was crucial for ensuring the consistency and robustness of cell type annotations.

Cortical segmentation and production of the segment flat map

To visualize Stereo-seq sections comprehensively while minimizing spatial information loss, we adopted the streamline-based approach detailed in the references (*103*, *104*). In short, the perimeters of the pial surface and white matter boundary were digitized along the same orientation. Both were assigned potential values within the same range. Next, each perimeter, that of the pial surface (outer boundary) and the white matter boundary (inner boundary), was divided into equal intervals. Equipotential streamlines were then drawn to connect points of the same potential value across these two perimeters. Through this process, each cortex in Stereo-seq sections was effectively segmented into a series of segments of the same length. Then we stacked the segments from one section into a bar and aligned the bars into a segment flat map.

We initiated the process by leveraging the MATLAB "bwboundaries" function. Using 8-connectivity, this function enabled us to extract all boundaries from the binary cortex tissue masks. When dealing with ring-shaped sections, differentiating between inner and outer boundaries was straightforward; we simply compared their lengths. However, for strip-shaped sections, owing to their more complex morphology, we had to manually predefine the start and end points of the inner and outer boundaries. This was done by carefully examining the morphological characteristics of each strip-shaped section. After precisely defining the boundaries, we implemented a semiautomated pipeline (https://github.com/chrisadamsonmcri/CCSegThickness) (*105*). This pipeline was applied to the marked boundaries, facilitating the generation of equipotential streamlines. The pipeline functioned by solving for equipotential streamlines between the defined boundaries. Using these streamlines, we created a series of evenly spaced segments within the cortex. The midline equipotential contour, where $\varphi = 0.5$, was used as a reference for cross section alignment. This ensured that the segment widths remained consistent across the entire sections. Finally, we filled in the streamlines and assigned sequential numbers to each segment. Completing these steps finalized the segmentation process, resulting in a well-defined, organized representation of the cortex structure.

Beginning with the stripe-shaped sections that covered the cingulate cortex, we designated the segment at the top of the cingulate gyrus as the zero value. Subsequently, we stacked the remaining segments above

it to form bars. Building on this foundation, we then aligned the nearly ring-shaped sections in the occipital lobe and frontal lobe, which did not cover the cingulate gyrus, with the adjacent stripe-shaped section according to their positions. We also manually defined their zero points and generated their stacked segment bars. For the ring-shaped parts of the temporal lobe that covered the auditory region, we conducted separate segment calculations and chose the first segment of the PaIM region to stack the segments and concatenate them to the main parts. Finally, to create the segment flat map, we arranged the segment stacks sequentially, starting from the first coronal section at the frontal pole and ending with the last coronal section at the occipital lobe (fig. S1D).

Stereo-seq data smoothing

To interpolate and smooth the spatial expression data, we initially leveraged the R package RANN (*106*) to construct a kd-tree for query bins (or segments) and reference bins (or segments) according to their spatial positions. For each query point, we estimated its expression profile by calculating the average of the expression profiles of its nearest neighboring reference points and itself. To enhance the smoothing effect, we repeated this averaging procedure for a predefined number of additional iterations. To ensure reproducibility, we developed a smoothing function named "smooth_kNN" (see our GitHub repository: https://github.com/marmosetbrainmapping/Code_Huang2026) to encapsulate this process. Regarding the smoothing of discrete data, we also computed a kd-tree. Subsequently, we counted the occurrences of all categories among the nearest neighboring reference points and the query point itself, and assigned the most prevalent category to the query point. The "winner_kNN" function encapsulated this operation.

Visualization of gene expression, cell type distribution, and multimodal data on the flat map

After the construction of the segment flat map, we aggregated the gene expression and the quantity of each cell type within every individual segment. This aggregation process was carried out to generate the segment × gene and segment × cell matrices. Subsequently, we normalized both matrices using the "NormalizeData" function from the Seurat package. After normalization and smoothing, as detailed in the preceding section, we visualized these matrices with the help of the ggplot2 package (*107*) integrated in R. For continuous data derived from other modalities, such as T1-weighted and T2-weighted data, we computed the average of all pixels within each segment. When it came to discontinuous data, such as the partitions in the Paxinos atlas (*40*), we assigned the most prevalent category within each segment as its representative value. Similar to the previous data, this was visualized using the ggplot2 package. We projected the hierarchy level of the marmoset cortex (*52*) onto the segment flat map, along with the Paxinos area labels. Furthermore, in the visualization of the segment flat map, we set the axis aspect ratio to 3.5 based on the actual dimensions of the brain.

Stereo-seq data unsupervised clustering with cell-cell adjacency

First, we used the kd-tree algorithm from the sklearn package (*108*) in Python (v3.11.6) to calculate the 30 nearest-neighbor cell-bins for each cell-bin. We then counted the number of different cell types within these 30 neighboring cell bins. Subsequently, we constructed an adjacent cell frequency matrix (Fig. 1D). In this matrix, each row corresponds to a cell bin, each column represents a cell type, and the values denote the cell counts. To reduce the matrix size and expedite subsequent computations, we aggregated the cell bins within each bin150 (75 μm) region.

Using this matrix, we carried out dimensionality reduction and clustering following the standard procedures of the Seurat package. This included data normalization, scaling, variable feature selection, PCA, and Louvain clustering, which were consistent with the methods described previously for scRNA-seq data clustering. By setting the dimension parameter to 10 and the resolution to 0.5 in the "FindClusters" function, we ultimately identified seven spatially distinct clusters. These clusters were then annotated as L1, L2, L3, L4, L5, L6, and L2-Allocortex.

Moreover, to acquire the annotations at the cell-bin level, we combined the cell-bin adjacency matrix for each section with the previously mentioned annotated bin150 adjacency matrix. Then, we conducted a PCA analysis and transferred the bin150 annotations to the cell-bins based on the annotation of the most adjacent bin-150 in the principal component space (the first 15 principal components).

Manual parcellation of cortex layers

First, we used the script "gem_visual.py" from our GitHub repository (https://github.com/marmosetbrainmapping/Code_Huang2026). Taking the bin1 matrix file as input, this script generated a spatial map at a bin50 resolution for each cortical layer (L1-L6) and white-matter up-regulated gene. Subsequently, we manually selected the genes that exhibited the most distinct laminar structure (table S1G). Then, we summed the expression amounts of these seven up-regulated gene groups and generated spatial maps at a bin50 resolution. Using the information of gene expression maps, unsupervised clustering results, and ssDNA staining images, we manually parcellated the six layers of the isocortex region and the two layers of the allocortex region using ImageJ software (*109*).

Region- or cortical layer–level matrix processing

First, we mapped the brain region information onto the Stereo-seq data of macaque (*6*), marmoset (this study), and mouse (*31*) matrices at the bin1 resolution. We then aggregated the total expression of each gene and the quantity of each cell type in all regions into a region-level matrix. Based on these matrices (region × gene, region × cell type), we created Seurat objects. When analyzing the cell distribution at the cortical layer level, we began by calculating the proportions of cell types within each layer. Then, we normalized their abundances by computing z scores. Subsequently, we excluded those cell types whose z scores were less than 1 in any layer. Finally, based on the filtered matrix, we also created a Seurat object with the region-layer × cell type format. For the human snRNA-seq data (*32*), we adopted the region information from the sampling process. We aggregated the cell types across brain regions to obtain a region × cell type matrix and created a corresponding Seurat object. Regarding the human microarray data (*67*) with finer brain region information, to enhance the data quality, we directly generated Seurat objects using the filtered and preprocessed region × gene expression matrix from the reference (*110*). After preparing all the datasets, we conducted unified preprocessing on all the data using the "NormalizeData" function in Seurat.

Region-enriched cell type identification

We conducted a hypergeometric test for each cell type to evaluate the statistical significance of enrichment across all brain regions. To validate the cell type enrichment results, we first calculated the proportion of each cell type across all brain regions. A permutation test was further carried out by randomly permuting the cell type labels 1000 times to generate a background distribution. We then compared the observed frequency of cell types with this background distribution. The empirical P value was computed by determining the proportion of background values that were greater than or equal to the observed frequency.

Finally, we identified the enriched cell types according to the following criteria: (i) The P value from the hypergeometric test was less than 0.05. (ii) No more than 20% of the regions demonstrated significant enrichment for the cell type. (iii) The number of cells of the cell type was greater than 30. (iv) The enriched cell type passed the permutation test with a P value of 0.

Pr-AI index calculation

Before calculating the Pr (primary sensory)–AI (allocortical/periallocortical) index, we need to define the Pr regions and AI regions for each species. For marmoset, the Pr regions include A3a/A3b (S1), V1, and AuA1, and

the Al regions include Ent (Ent) and Pir/Apir (Pir). For macaque (*6*), the Pr regions are 3a/b (S1), V1, and AuA1, and the Al regions include Pir and ER/EO/EI/ELr/ELc/ECL/EC (Ent). For human (*32*), the Pr regions include V1C (V1), A1C (AuA1), and S1C (S1), and the Al regions include Pir and LEC/MEC (Ent). For mouse (*31*), the Pr regions include VISp (V1), SSp (S1), and AUDp (AuA1), and the Al regions include PIR (Pir) and ENTl/ENTm/ENTmv (Ent).

For region-level Pr-Al index calculation, we first calculated the Pearson correlation coefficients of cell compositions between each brain region and the Pr regions as well as the Al regions, respectively. We took the maximum correlation coefficient of the two types of brain regions as Cor_{Pr} and Cor_{Al} for each brain region, and normalized Cor_{Pr} and Cor_{Al} independently by setting the maximum value to 1, the mean value to 0.5, and the minimum value to 0. We finally calculated the Pr-Al index with the following normalization formula for each region

$$\text{Pr-Al index} = \frac{Cor_{Pr} - Cor_{Al}}{(1 - Cor_{Al}) + (1 - Cor_{Pr})}$$

For the segment-level data, we used the same calculation process to calculate Cor_{Pr} and Cor_{Al} with the Pr and Al regions, and calculated the Pr-Al index for each segment.

PCA of cell types and genes

We used the "RunPCA" function from the Seurat package to conduct PCA on the segment × gene data, segment × cell type data, region × gene data, and region × cell type data. For the PCA of the segment × gene data, we selected the top 5000 variable genes, which were identified by the "FindVariableFeatures" function. In the cross-species region-level comparison, we used all the expressed homologous genes ($n = 5422$) as variable features for the PCA.

Identification of Pr or Al cell types and genes

To investigate the cell types and genes underlying the Pr-Al index, we first standardized the regional distribution of cell types and gene expression. Subsequently, we calculated their Pearson correlation coefficients with the regional-level Pr-Al index. By setting a significance threshold of $P < 0.05$ and using the correlation coefficient values, we classified the cell types and genes into either Pr or Al types.

Given that marmosets (this study), macaques (*6*), and mice (*31*) all have 3D Stereo-seq data encompassing the entire brain along with brain region annotations, we aggregated the cell types and genes of the corresponding data at the regional level. Subsequently, we used the identical calculation method and threshold as previously described to identify Pr or Al genes and Pr or Al cell types. For human samples, we used the snRNA-seq data (*32*) to identify Pr or Al cell types. Regarding the identification of Pr or Al genes, we used the microarray data (*67*) because of its finer sampling resolution, using the same methods as mentioned above. As depicted in fig. S6B, the gene PC1 exhibits a strong gradient from Pr to Al regions in humans, and the gene PC1 in other species is highly correlated with the Pr-Al index. Therefore, we first recalculated the Pr-Al index for the microarray data using the PC1 of gene expression and then used the same approach to identify whether genes were of the Pr or Al type.

Integration and cell type annotation for developmental snRNA-seq datasets

For cell clustering and identification of glutamatergic, GABAergic, and nonneuronal subclasses, we followed the procedure described in the "snRNA-seq annotation" section. To integrate developmental datasets, we applied reciprocal PCA (RPCA) integration using the Seurat package with 3000 highly variable genes. Independently preannotated snRNA-seq datasets from different stages were first combined. Integration anchors were identified with Seurat's "FindIntegrationAnchors" function, and label transfer was carried out using "TransferData." The transferred labels were further refined based on canonical marker genes provided in the script "00_snRNA_annotation.md" available on our GitHub repository (https://github.com/marmosetbrainmapping/Code_Huang2026).

To enable direct comparison with adult cortical cell types, the adult dataset was adopted as the reference, and its annotations were transferred to other stages. Cell types were assigned using the "winner_kNN" approach described previously in the "Stereo-seq data smoothing" section. Specifically, for each query cell, its k nearest neighbors ($knn = 15$) in the reference dataset were first identified in PCA space (npcs = 5). The labels of these nearest neighbors were summarized, and then the most prevalent label (the "winner") was assigned to the query cell.

Calculation of PAD and gene coexpression matrix

To evaluate the degree of opposite expression between Pr and Al genes, we calculated the PAD (Pr-Al gene Difference) value using the gene coexpression matrix. Correlation values in the coexpression matrix were determined as follows: In snRNA-seq data, they were calculated via the Pearson correlation coefficient; in spatial transcriptomics data, they were derived from the local correlation method used in Hotspot (*111*). The calculation process of local correlation is similar to that of Pearson correlation coefficient, except that spatial neighboring information is incorporated as a weighting factor. The calculation formula for PAD is shown in Fig. 3C: it is computed by taking the average of the Pr-Pr gene coexpression matrix and the Al-Al gene coexpression matrix, then subtracting the average of the Pr-Al gene coexpression matrix.

Cell type similarity calculation across species

The cell type comparison analysis was based on the cortical snRNA-seq data of human (*32*), macaque (*6*), and marmoset (this study), with two biological replicates from macaque and marmoset. First, the Metacell-2 software (*112*) was used to generate pseudo-bulk Meta cells within each species. Subsequently, the homologous genes for the three species were obtained from the ENSEMBL BioMart database (version 9.1), and the expression matrices for all species were filtered based on this list. To achieve comparability of the data across species, the SCT integration workflow of the Seurat software was adopted: After SCTransform normalization on the expression matrices of each species, the union of highly variable genes was extracted as the feature gene set. Then, functions of "PrepSCTIntegration" for preprocessing, "FindIntegrationAnchors" for anchor genes searching, and "IntegrateData" for data integration were sequentially carried out to finally generate an aligned cross-species cell profile. On this basis, the MetaNeighbor algorithm (*113*) was used to calculate the AUROC values between cell types to quantify the similarities of cell types among species. In this study, the cell types of the marmoset were used as a reference framework.

Correlation analysis of marmoset multimodal data with cell types and genes

For the marmoset multimodal data, which encompassed cytoarchitecture, intersection index, cortical architecture, spatial axis, hierarchy, connectivity gradients, and task activity, we computed the Pearson correlation coefficients at the segment level. These coefficients were calculated between the cell type distribution and gene expression, as well as among different patterns. When determining the genes and cell types associated with the patterns, we set the threshold of the Pearson correlation coefficient at 0.5. Subsequently, we designated the pattern with the highest correlation coefficient as the label for the corresponding cell type or gene. During the calculation of Pearson correlation coefficients between other patterns and the Pr-Al index, we excluded several regions of other modalities during pairwise computations. Such regions included the Pir/Ent regions within the hierarchical pattern and the regions where the layer thickness was zero.

Gene expression patterns analysis of the thalamus

For the thalamus, we carried out an analysis of gene expression patterns at a bin200 (100-μm) resolution. Initially, we used the Seurat package's

RPCA method to integrate all sections. This process involved the selection of 5000 integration features with the "SelectIntegrationFeatures" function (*101*). Subsequently, based on the integrated matrix, the "smooth_kNN" function within the RANN package was used (*106*). This function was applied to smooth the values for each bin200 spot. The smoothing was conducted in two rounds, considering nine bins in the vicinity of each spot. After the smoothing step, the "RunPCA" function was executed to perform PCA on the integrated and smoothed results. The Seurat "AddModuleScore" function was used to calculate the gene-set scores of Pr/Al genes specifically for Pr and Al genes.

For the identification of thalamic PC1-related genes, Pearson correlation coefficients were computed between the smoothed gene expressions and PC1. Genes with a P value less than 0.05 and a Pearson correlation coefficient (PCC) greater than 0.15 were recognized as being positively correlated with PC1. Conversely, genes with a PCC less than −0.15 were identified as negatively correlated with PC1. Analogously, the gene-set scores of both the positively and negatively related genes were calculated on the segment flat map of the cortex. This calculation was performed using the "AddModuleScore" function. To better visualize the thalamic genes, as shown in Fig. 5E, we additionally used the Rmagic package to perform imputation on these genes with the default parameters (*114*).

ROC calculation

For the ROC calculation, we defined an R function "ROC_kNN" (https://github.com/marmosetbrainmapping/Code_Huang2026). The underlying principle was to compute the quotient of the difference between the maximum and minimum expression values or cell abundance and the distance between them in the vicinity of each spot. This method mitigated the instability of the ROC that is attributable to the inadequate smoothness of the expression amounts. By using this method, we calculated ROCs of cell type abundances and gene expressions, respectively. The calculation was based on the segment flat map, which determined the whole-brain ROC. In this mode, we took the 16 surrounding points in the calculation.

We also carried out ROC calculations at the segment level for single sections. In this approach, initially, we used the "NormalizeData" function in Seurat to normalize the data. Subsequently, we used the "smooth_kNN" function to smooth the data across nine surrounding segments for more than four rounds. Finally, we executed the "ROC_KNN" function to compute the ROC within nine surrounding segments as well.

Multiatlas boundary comparison and ROC-boundary relationship analysis

We compared ROC values against four marmoset brain atlases: Paxinos (116 regions) (*11*, *40*), MBMv4 (96 regions) (*15*), Riken (114 regions) (*13*), and MBMv1 (105 regions) (*11*). Boundaries were detected using k nearest neighbors ($k = 3$), where segments were classified as boundary points if their neighbors belonged to exactly two different regions.

We quantified ROC-boundary relationships through three complementary metrics. For peak proportion analysis, we identified ROC peaks as local maxima and calculated the proportion of peaks located within 1 spatial unit of boundaries using Euclidean distance. For enrichment analysis, we defined high ROC segments as those exceeding the 90th percentile threshold and calculated fold enrichment as the ratio of observed to expected overlap between high ROC segments and boundaries

$$\mathrm{FE} = \left(n_{\mathrm{overlap}}/n_{\mathrm{border}}\right) / \left(n_{\mathrm{high}}/n_{\mathrm{total}}\right)$$

Statistical significance was assessed by the hypergeometric test.

Functional network up-regulated genes identification

After the aforementioned multimodal data registration, we obtained the continuous functional-network value, displayed on the segment flat map. Subsequently, we defined the dominant regions of each network by selecting the top 90% of values (the 90th percentile and above) of the network. Finally, the up-regulated genes of each network are computed by the function "FindMarkers" in Seurat, and genes with a $\log_2$ fold change greater than 0.9 and a P value less than 0.05 were selected as network up-regulated genes. The gene-set score of the default mode network up-regulated genes was calculated by the "AddModuleScore" function.

PLS regression analysis

To investigate the relationship between gene expression and brain functional networks, we performed PLS regression, a multivariate technique suitable for high-dimensional data. The gene expression data (X) and brain functional networks (Y) were used as predictors and response variables, respectively. Before model construction, we systematically determined the optimal number of latent components for each network through an iterative optimization procedure, with a preset threshold of explained variance (95%). The maximum number of components was constrained to 20 based on preliminary stability analyses, ensuring comprehensive model fitting across all networks while preventing overparameterization. The PLS regression coefficients (BETA) were subsequently estimated through covariance maximization between X and Y. The fitted values of the brain network maps (Y_{fit}) were computed using the augmented predictor matrix that incorporated an intercept term

$$Y_{\mathrm{fit}} = [1, X] \times \mathrm{BETA}$$

where [1, X] represents the augmented predictor matrix with an intercept term.

To validate model robustness, we conducted sensitivity analyses with alternative component configurations (16, 18, 22, and 25 latent components). Notably, the gene weight patterns derived from PLS models with varying component configurations demonstrated strong consistency with those generated by the 20-component reference model. The cross-component agreement within the tested range suggested that the biological interpretations derived from our PLS implementation were robust to moderate variations in the number of latent components. To determine the number of genes required to reconstruct each brain network map, we performed iterative PLS regression fitting using top-ranked genes based on the absolute values of their regression weights. Genes were sequentially added to the model in descending order of weight magnitude until the Pearson correlation coefficient between the fitted maps (Y_{fit}) and the corresponding original network maps reached a threshold of 0.95. The minimal gene count achieving this criterion was defined as the optimal number of genes required for network reconstruction.

Functional connectivity analysis

The marmoset awake resting-state fMRI data were obtained from the Marmoset Brain Mapping project (The preprocessed data "regrBMWC0" of the MBMv4) (*15*). The dataset included 39 marmosets (8 females, 31 males) with an average age of 4 years. Each subject underwent repeated scanning, resulting in a total of 710 fMRI runs (17 min per run). For human awake resting-state fMRI data, we used the CoRR-HNU dataset (*115*), part of the Consortium for Reliability and Reproducibility. This dataset comprises neuroimaging data from 30 healthy young adults (15 females, 15 males) with an average age of 24 years.

To assess the functional connectivity, voxel-wise time series were extracted from the A10, DMN, and other cortical regions for each subject. Pairwise functional connectivity between voxels in the A10, DMN, and other cortical regions was computed using Pearson correlation coefficients, reflecting voxel-level connectivity. Correlation matrices were generated for each subject and subsequently averaged across subjects to obtain group-level mean functional connectivity matrices. The resulting correlation matrices were averaged across subjects at the group level to obtain mean functional connectivity between each pair of regions (A10-DMN, A10-other cortical regions, and DMN-other cortical regions).

Region–up-regulated gene identification

Regarding the spatial transcriptomic data, based on the region each cell belongs to, we computed the aggregate expression amounts of every gene in each region. Subsequently, we built a Seurat object (V4) where genes were in rows and regions were in columns. After getting this matrix, we ran the "NormalizeData" function for data normalization. Postnormalization, the expression counts of each gene across diverse regions are scaled within the range from 1 to 2. After the scaling step, for each specific target region, we calculated the fold change (FC) value. This was achieved by dividing the scaled aggregated expression of the target region by the scaled aggregated expressions of all other regions. The mean value of these calculated FC values for a particular gene is then determined and designated as the FC value for that gene. If a gene had an FC value greater than 1, it was designated as an up-regulated gene in the target region. This approach was systematically implemented for each brain region and subregion.

Cross-species integration of snRNA-seq data in the auditory cortex

To integrate auditory cortical cells from humans, macaques, and marmosets, the snRNA-seq datasets for the auditory areas of humans (*32*) and marmosets (*16*) were retrieved from public databases. Coupled with an in-house generated snRNA-seq dataset from the macaque auditory cortex, we conducted RPCA analysis implemented within the Seurat package. This was done to integrate the snRNA-seq datasets across species, using 4000 integrating genes.

To attain a unified annotation for independently preannotated snRNA-seq datasets, we initiated the process by combining the snRNA-seq data and Stereo-seq data of macaques. Subsequently, we harnessed the "FindIntegrationAnchors" function within Seurat to identify the anchors, which were then used as a basis for using the "TransferData" function to perform label transfer. To refine the accuracy of the transferred labels, we used marker genes to rectify the label-transfer results. With the macaque data comprehensively processed and annotated, we then adopted the macaque data as the reference, and its labels were transferred to the other datasets. Specifically, the previously mentioned "winner_kNN" function was used to identify the most similar cells for each query cell within the top five principal components of the PCA space. Subsequently, the most frequently occurring cell label among the neighbors was assigned to the query cell. If the frequency of the most common label surpassed a predefined threshold, the label was transferred; otherwise, the original label was retained.

Correlation of cell subclass in snRNA-seq

Using the integrated snRNA-seq datasets of humans, macaques, and marmosets, we calculated the average expression of each gene within each subclass of each species. This was achieved by using the "AverageExpression" function in Seurat. Subsequently, the "cor" function in R was used to compute the Pearson correlation coefficient of average gene expression across groups. Glutamatergic neurons and GABAergic neurons were analyzed independently.

Species-wise and layer-wise DEGs identification

We computed the species-specific or cortical layer–enriched markers using the "FindMarkers" function in Seurat. Only significant DEGs with a P value less than 0.05, a relatively high expression ($\log_2$ fold change > 0.25), and a notable difference in expression percentage (pct.1 − pct.2 > 0.15) were retained. Subsequently, the DEGs of different species were intersected for comparison and visualization. Next, taking the shared DEGs between humans and marmosets as input, we further applied the "FindMarkers" function to the human and marmoset spatial transcriptome data across cortical layers ($\log_2$ fold change > 0.25) to obtain the layer-enriched markers.

REFERENCES AND NOTES

1. D. Pandya, B. Seltzer, M. Petrides, P. B. Cipolloni, *Cerebral Cortex: Architecture, Connections, and the Dual Origin Concept* (Oxford Univ. Press, 2015). doi: 10.1093/med/9780195385151.001.0001
2. M. G. Rosa, R. Tweedale, Brain maps, great and small: Lessons from comparative studies of primate visual cortical organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 665–691 (2005). doi: 10.1098/rstb.2005.1626; pmid: 15937007
3. N. Geschwind, Disconnexion syndromes in animals and man. I. *Brain* **88**, 237–294 (1965). doi: 10.1093/brain/88.2.237; pmid: 5318481
4. M. M. Poo, Transcriptome, connectome and neuromodulation of the primate brain. *Cell* **185**, 2636–2639 (2022). doi: 10.1016/j.cell.2022.05.011; pmid: 35732175
5. A. Fornito, A. Arnatkevičiūtė, B. D. Fulcher, Bridging the Gap between Connectome and Transcriptome. *Trends Cogn. Sci.* **23**, 34–50 (2019). doi: 10.1016/j.tics.2018.10.005; pmid: 30455082
6. A. Chen *et al.*, Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. *Cell* **186**, 3726–3743.e24 (2023). doi: 10.1016/j.cell.2023.06.009; pmid: 37442136
7. Y. Shen *et al.*, Mapping big brains at subcellular resolution in the era of big data in zoology. *Zool. Res.* **43**, 597–599 (2022). doi: 10.24272/j.issn.2095-8137.2022.138; pmid: 35726585
8. Q. Wang, S. Zhao, T. Liu, J. Han, C. Liu, Temporal fingerprints of cortical gyrification in marmosets and humans. *Cereb. Cortex* **33**, 9802–9814 (2023). doi: 10.1093/cercor/bhad245; pmid: 37434368
9. 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). doi: 10.1038/s41593-019-0575-0; pmid: 31932765
10. C. Liu *et al.*, Marmoset Brain Mapping V3: Population multi-modal standard volumetric and surface-based templates. *Neuroimage* **226**, 117620 (2021). doi: 10.1016/j.neuroimage.2020.117620; pmid: 33307224
11. C. Liu *et al.*, A digital 3D atlas of the marmoset brain based on multi-modal MRI. *Neuroimage* **169**, 106–116 (2018). doi: 10.1016/j.neuroimage.2017.12.004; pmid: 29208569
12. X. Zhu *et al.*, An anatomical and connectivity atlas of the marmoset cerebellum. *Cell Rep.* **42**, 112480 (2023). doi: 10.1016/j.celrep.2023.112480; pmid: 37163375
13. A. Woodward *et al.*, The Brain/MINDS 3D digital marmoset brain atlas. *Sci. Data* **5**, 180009 (2018). doi: 10.1038/s41597-022-01247-z; pmid: 35304470
14. P. Majka *et al.*, Open access resource for cellular-resolution analyses of corticocortical connectivity in the marmoset monkey. *Nat. Commun.* **11**, 1133 (2020). doi: 10.1038/s41467-020-14858-0; pmid: 32111833
15. X. Tian *et al.*, An integrated resource for functional and structural connectivity of the marmoset brain. *Nat. Commun.* **13**, 7416 (2022). doi: 10.1038/s41467-022-35197-2; pmid: 36456558
16. F. M. Krienen *et al.*, A marmoset brain cell census reveals regional specialization of cellular identities. *Sci. Adv.* **9**, eadk3986 (2023). doi: 10.1126/sciadv.adk3986; pmid: 37824615
17. Y. Kita *et al.*, Cellular-resolution gene expression profiling in the neonatal marmoset brain reveals dynamic species- and region-specific differences. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2020125118 (2021). doi: 10.1073/pnas.2020125118; pmid: 33903237
18. J. P. Lin *et al.*, Transcriptomic architecture of nuclei in the marmoset CNS. *Nat. Commun.* **13**, 5531 (2022). doi: 10.1038/s41467-022-33140-z; pmid: 36130924
19. M. E. Schroeder *et al.*, A transcriptomic atlas of astrocyte heterogeneity across space and time in mouse and marmoset. *Neuron* **113**, 3942–3965.e19 (2025). doi: 10.1016/j.neuron.2025.09.011; pmid: 41270736
20. S. Hao *et al.*, Cross-species single-cell spatial transcriptomic atlases of the cerebellar cortex. *Science* **385**, eado3927 (2024). doi: 10.1126/science.ado3927; pmid: 39325889
21. C. Liu *et al.*, Anatomical and functional investigation of the marmoset default mode network. *Nat. Commun.* **10**, 1975 (2019). doi: 10.1038/s41467-019-09813-7; pmid: 31036814
22. A. Jafari *et al.*, A vocalization-processing network in marmosets. *Cell Rep.* **42**, 112526 (2023). doi: 10.1016/j.celrep.2023.112526; pmid: 37195863
23. C. Tong *et al.*, Multimodal analysis demonstrating the shaping of functional gradients in the marmoset brain. *Nat. Commun.* **13**, 6584 (2022). doi: 10.1038/s41467-022-34371-w; pmid: 36329036
24. A. Dureux, A. Zanini, J. Selvanayagam, R. S. Menon, S. Everling, Gaze patterns and brain activations in humans and marmosets in the Frith-Happé theory-of-mind animation task. *eLife* **12**, e86327 (2023). doi: 10.7554/eLife.86327; pmid: 37449983
25. X. Tian, A. C. Silva, C. Liu, The Brain Circuits and Dynamics of Curiosity-Driven Behavior in Naturally Curious Marmosets. *Cereb. Cortex* **31**, 4220–4232 (2021). doi: 10.1093/cercor/bhab080; pmid: 33839768
26. A. Chen *et al.*, Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022). doi: 10.1016/j.cell.2022.04.003; pmid: 35512705
27. L. Han *et al.*, Cell transcriptomic atlas of the non-human primate *Macaca fascicularis*. *Nature* **604**, 723–731 (2022). doi: 10.1038/s41586-022-04587-3; pmid: 35418686
28. S. Ma *et al.*, Molecular and cellular evolution of the primate dorsolateral prefrontal cortex. *Science* **377**, eabo7257 (2022). doi: 10.1126/science.abo7257; pmid: 36007006
29. R. Lopez *et al.*, DestVI identifies continuums of cell types in spatial transcriptomics data. *Nat. Biotechnol.* **40**, 1360–1369 (2022). doi: 10.1038/s41587-022-01272-8; pmid: 35449415

30. J. A. Bourne, M. G. Rosa, Hierarchical development of the primate visual cortex, as revealed by neurofilament immunoreactivity: Early maturation of the middle temporal area (MT). *Cereb. Cortex* **16**, 405–414 (2006). doi: 10.1093/cercor/bhi119; pmid: 15944371
31. L. Han *et al.*, Single-cell spatial transcriptomic atlas of the whole mouse brain. *Neuron* **113**, 2141–2160.e9 (2025). doi: 10.1016/j.neuron.2025.02.015; pmid: 40132589
32. K. Siletti *et al.*, Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023). doi: 10.1126/science.add7046; pmid: 37824663
33. Y. Kim *et al.*, Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. *Cell* **171**, 456–469.e22 (2017). doi: 10.1016/j.cell.2017.09.020; pmid: 28985566
34. G. Stelzer *et al.*, The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinform.* **54**, 1.30.1–1.30.33 (2016). doi: 10.1002/cpbi.5; pmid: 27322403
35. I. Intskirveli, R. Metherate, Nicotinic neuromodulation in auditory cortex requires MAPK activation in thalamocortical and intracortical circuits. *J. Neurophysiol.* **107**, 2782–2793 (2012). doi: 10.1152/jn.01129.2011; pmid: 22357798
36. B. Kaminska, L. Kaczmarek, S. Zangenehpour, A. Chaudhuri, Rapid phosphorylation of Elk-1 transcription factor and activation of MAP kinase signal transduction pathways in response to visual stimulation. *Mol. Cell. Neurosci.* **13**, 405–414 (1999). doi: 10.1006/mcne.1999.0757; pmid: 10383826
37. E. Dupont *et al.*, ERK is involved in the reorganization of somatosensory cortical maps in adult rats submitted to hindlimb unloading. *PLOS ONE* **6**, e17564 (2011). doi: 10.1371/journal.pone.0017564; pmid: 21408155
38. A. Dityatev, M. Schachner, Extracellular matrix molecules and synaptic plasticity. *Nat. Rev. Neurosci.* **4**, 456–468 (2003). doi: 10.1038/nrn1115; pmid: 12778118
39. A. Kicheva, M. Cohen, J. Briscoe, Developmental pattern formation: Insights from physics and biology. *Science* **338**, 210–212 (2012). doi: 10.1126/science.1225182; pmid: 23066071
40. G. Paxinos, C. Watson, M. Petrides, M. Rosa, H. Tokuno, *The Marmoset Brain in Stereotaxic Coordinates* (Academic Press, 2012).
41. P. L. A. de Góis Morais *et al.*, Cytoarchitecture and myeloarchitecture of the entorhinal cortex of the common marmoset monkey (*Callithrix jacchus*). *J. Comp. Neurol.* **528**, 1307–1320 (2020). doi: 10.1002/cne.24814; pmid: 31765000
42. K. Watanabe-Sawaguchi, K. Kubota, T. Arikuni, Cytoarchitecture and intrafrontal connections of the frontal cortex of the brain of the hamadryas baboon (*Papio hamadryas*). *J. Comp. Neurol.* **311**, 108–133 (1991). doi: 10.1002/cne.903110109; pmid: 1719042
43. K. J. Burman, S. M. Palmer, M. Gamberini, M. W. Spitzer, M. G. Rosa, Anatomical and physiological definition of the motor cortex of the marmoset monkey. *J. Comp. Neurol.* **506**, 860–876 (2008). doi: 10.1002/cne.21580; pmid: 18076083
44. N. Antón-Bolaños, A. Espinosa, G. López-Bendito, Developmental interactions between thalamus and cortex: A true love reciprocal story. *Curr. Opin. Neurobiol.* **52**, 33–41 (2018). doi: 10.1016/j.conb.2018.04.018; pmid: 29704748
45. J. W. Phillips *et al.*, A repeated molecular architecture across thalamic pathways. *Nat. Neurosci.* **22**, 1925–1935 (2019). doi: 10.1038/s41593-019-0483-3; pmid: 31527803
46. S. Oldham, G. Ball, A phylogenetically-conserved axis of thalamocortical connectivity in the human brain. *Nat. Commun.* **14**, 6032 (2023). doi: 10.1038/s41467-023-41722-8; pmid: 37758726
47. D. S. Roy, Y. Zhang, M. M. Halassa, G. Feng, Thalamic subnetworks as units of function. *Nat. Neurosci.* **25**, 140–153 (2022). doi: 10.1038/s41593-021-00996-1; pmid: 35102334
48. C. Dehay, G. Horsburgh, M. Berland, H. Killackey, H. Kennedy, Maturation and connectivity of the visual cortex in monkey is altered by prenatal removal of retinal input. *Nature* **337**, 265–267 (1989). doi: 10.1038/337265a0; pmid: 2536139
49. I. H. Smart, C. Dehay, P. Giroud, M. Berland, H. Kennedy, Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb. Cortex* **12**, 37–53 (2002). doi: 10.1093/cercor/12.1.37; pmid: 11734531
50. G. E. Alexander, M. R. DeLong, P. L. Strick, Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986). doi: 10.1146/annurev.ne.09.030186.002041; pmid: 3085570
51. M. F. Glasser, D. C. Van Essen, Mapping human cortical areas *in vivo* based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011). doi: 10.1523/JNEUROSCI.2180-11.2011; pmid: 21832190
52. P. Theodoni *et al.*, Structural Attributes and Principles of the Neocortical Connectome in the Marmoset Monkey. *Cereb. Cortex* **32**, 15–28 (2021). doi: 10.1093/cercor/bhab191; pmid: 34274966
53. J. A. Harris *et al.*, Hierarchical organization of cortical and thalamic connectivity. *Nature* **575**, 195–202 (2019). doi: 10.1038/s41586-019-1716-z; pmid: 31666704
54. N. T. Markov *et al.*, Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* **522**, 225–259 (2014). doi: 10.1002/cne.23458; pmid: 23983048
55. D. J. Felleman, D. C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991). doi: 10.1093/cercor/1.1.1-a; pmid: 1822724
56. R. R. Coifman *et al.*, Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7426–7431 (2005). doi: 10.1073/pnas.0500334102; pmid: 15899970
57. D. S. Margulies *et al.*, Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12574–12579 (2016). doi: 10.1073/pnas.1608282113; pmid: 27791099
58. V. Menon, 20 years of the default mode network: A review and synthesis. *Neuron* **111**, 2469–2487 (2023). doi: 10.1016/j.neuron.2023.04.023; pmid: 37167968
59. R. L. Buckner, L. M. DiNicola, The brain's default network: Updated anatomy, physiology and evolving insights. *Nat. Rev. Neurosci.* **20**, 593–608 (2019). doi: 10.1038/s41583-019-0212-7; pmid: 31492945
60. J. Smallwood *et al.*, The default mode network in cognition: A topographical perspective. *Nat. Rev. Neurosci.* **22**, 503–513 (2021). doi: 10.1038/s41583-021-00474-4; pmid: 34226715
61. C. M. Garin *et al.*, An evolutionary gap in primate default mode network organization. *Cell Rep.* **39**, 110669 (2022). doi: 10.1016/j.celrep.2022.110669; pmid: 35417698
62. T. D. Müller *et al.*, The orphan receptor Gpr83 regulates systemic energy metabolism via ghrelin-dependent and ghrelin-independent mechanisms. *Nat. Commun.* **4**, 1968 (2013). doi: 10.1038/ncomms2968; pmid: 23744028
63. E. Seigneur, T. C. Südhof, Genetic Ablation of All Cerebellins Reveals Synapse Organizer Functions in Multiple Regions Throughout the Brain. *J. Neurosci.* **38**, 4774–4790 (2018). doi: 10.1523/JNEUROSCI.0360-18.2018; pmid: 29691328
64. S. M. Shalini *et al.*, Distribution of Alox15 in the Rat Brain and Its Role in Prefrontal Cortical Resolvin D1 Formation and Spatial Working Memory. *Mol. Neurobiol.* **55**, 1537–1550 (2018). doi: 10.1007/s12035-017-0413-x; pmid: 28181190
65. A. R. McIntosh, F. L. Bookstein, J. V. Haxby, C. L. Grady, Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* **3**, 143–157 (1996). doi: 10.1006/nimg.1996.0016; pmid: 9345485
66. R. M. Braga, R. L. Buckner, Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron* **95**, 457–471.e5 (2017). doi: 10.1016/j.neuron.2017.06.038; pmid: 28728026
67. M. J. Hawrylycz *et al.*, An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012). doi: 10.1038/nature11405; pmid: 22996553
68. F. A. Mansouri, E. Koechlin, M. G. P. Rosa, M. J. Buckley, Managing competing goals — A key role for the frontopolar cortex. *Nat. Rev. Neurosci.* **18**, 645–657 (2017). doi: 10.1038/nrn.2017.111; pmid: 28951610
69. M. F. Glasser *et al.*, A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016). doi: 10.1038/nature18933; pmid: 27437579
70. K. S. Saleem, N. K. Logothetis, *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic Press, 2012).
71. A. Arnatkevičiūtė, B. D. Fulcher, A. Fornito, A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353–367 (2019). doi: 10.1016/j.neuroimage.2019.01.011; pmid: 30648605
72. M. Niu *et al.*, Multimodal mapping of macaque monkey somatosensory cortex. *Prog. Neurobiol.* **239**, 102633 (2024). doi: 10.1016/j.pneurobio.2024.102633; pmid: 38830482
73. J. H. Kaas, Comparative Functional Anatomy of Marmoset Brains. *ILAR J.* **61**, 260–273 (2020). doi: 10.1093/ilar/ilaa026; pmid: 33550381
74. L. Han *et al.*, A Population-scale Single-cell Spatial Transcriptomic Atlas of the Human Cortex. bioRxiv 2025.10.13.681959 [Preprint] (2025). doi: 10.1101/2025.10.13.681959
75. L. A. De La Mothe, S. Blumell, Y. Kajikawa, T. A. Hackett, Thalamic connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *J. Comp. Neurol.* **496**, 72–96 (2006). doi: 10.1002/cne.20924; pmid: 16528728
76. 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). doi: 10.1523/JNEUROSCI.2909-13.2013; pmid: 24048842
77. S. E. Petersen, B. A. Seitzman, S. M. Nelson, G. S. Wig, E. M. Gordon, Principles of cortical areas and their implications for neuroimaging. *Neuron* **112**, 2837–2853 (2024). doi: 10.1016/j.neuron.2024.05.008; pmid: 38834069
78. L. Krubitzer, J. Kaas, The evolution of the neocortex in mammals: How is phenotypic diversity generated? *Curr. Opin. Neurobiol.* **15**, 444–453 (2005). doi: 10.1016/j.conb.2005.07.003; pmid: 16026978
79. M. Heide *et al.*, Human-specific *ARHGAP11B* increases size and folding of primate neocortex in the fetal marmoset. *Science* **369**, 546–550 (2020). doi: 10.1126/science.abb2401; pmid: 32554627
80. M. Betizeau *et al.*, Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron* **80**, 442–457 (2013). doi: 10.1016/j.neuron.2013.09.032; pmid: 24139044
81. S. Li, X. J. Wang, Hierarchical timescales in the neocortex: Mathematical mechanism and biological insights. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2110274119 (2022). doi: 10.1073/pnas.2110274119; pmid: 35110401
82. R. L. Buckner, D. S. Margulies, Macroscale cortical organization and a default-like apex transmodal network in the marmoset monkey. *Nat. Commun.* **10**, 1976 (2019). doi: 10.1038/s41467-019-09812-8; pmid: 31036823

83. G. N. Ngo, Y. Hori, S. Everling, R. S. Menon, Joint-embeddings reveal functional differences in default-mode network architecture between marmosets and humans. *Neuroimage* **272**, 120035 (2023). doi: 10.1016/j.neuroimage.2023.120035; pmid: 36948281
84. D. M. Grijseels, B. J. Prendergast, J. C. Gorman, C. T. Miller, The neurobiology of vocal communication in marmosets. *Ann. N. Y. Acad. Sci.* **1528**, 13–28 (2023). doi: 10.1111/nyas.15057; pmid: 37615212
85. C. T. Miller *et al.*, Marmosets: A Neuroscientific Model of Human Social Behavior. *Neuron* **90**, 219–233 (2016). doi: 10.1016/j.neuron.2016.03.018; pmid: 27100195
86. S. Wirth, Encoding identity in the marmoset. *Science* **382**, 372–373 (2023). doi: 10.1126/science.adk8413; pmid: 37883556
87. D. Y. Takahashi *et al.*, The developmental dynamics of marmoset monkey vocal production. *Science* **349**, 734–738 (2015). doi: 10.1126/science.aab1058; pmid: 26273055
88. H. H. Zeng *et al.*, Distinct neuron populations for simple and compound calls in the primary auditory cortex of awake marmosets. *Natl. Sci. Rev.* **8**, nwab126 (2021). doi: 10.1093/nsr/nwab126; pmid: 34876995
89. 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). doi: 10.1073/pnas.2221756120; pmid: 37276391
90. J. Tsyporin *et al.*, Competing Programs Shape Cortical Sensorimotor-Association Axis Development. bioRxiv 2025.06.26.660775 [Preprint] (2025). doi: 10.1101/2025.06.26.660775
91. J. M. Kebschull *et al.*, High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. *Neuron* **91**, 975–987 (2016). doi: 10.1016/j.neuron.2016.07.036; pmid: 27545715
92. L. Li *et al.*, Identification and application of cell-type-specific enhancers for the macaque brain. *Cell* **188**, 4382–4400.e27 (2025). doi: 10.1016/j.cell.2025.07.026; pmid: 40695278
93. J. H. Lee, Q. Liu, E. Dadgar-Kiani, Solving brain circuit function and dysfunction with computational modeling and optogenetic fMRI. *Science* **378**, 493–499 (2022). doi: 10.1126/science.abq3868; pmid: 36327349
94. C. Gong *et al.*, SAW: An efficient and accurate data analysis workflow for Stereo-seq spatial transcriptomics. *Gigabyte* **2024**, gigabyte111 (2024). doi: 10.46471/gigabyte.111; pmid: 38434930
95. R. W. Cox, AFNI: What a long strange trip it's been. *Neuroimage* **62**, 743–747 (2012). doi: 10.1016/j.neuroimage.2011.08.056; pmid: 21889996
96. B. B. Avants, C. L. Epstein, M. Grossman, J. C. Gee, Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **12**, 26–41 (2008). doi: 10.1016/j.media.2007.06.004; pmid: 17659998
97. D. S. Marcus *et al.*, Informatics and data mining tools and strategies for the human connectome project. *Front. Neuroinform.* **5**, 4 (2011). doi: 10.3389/fninf.2011.00004; pmid: 21743807
98. C. Liu *et al.*, A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling. bioRxiv 818450 [Preprint] (2019). doi: 10.1101/818450
99. Marmoset Genome Sequencing and Analysis Consortium, The common marmoset genome provides insight into primate biology and evolution. *Nat. Genet.* **46**, 850–857 (2014). doi: 10.1038/ng.3042; pmid: 25038751
100. J. A. Griffiths, A. C. Richard, K. Bach, A. T. L. Lun, J. C. Marioni, Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat. Commun.* **9**, 2667 (2018). doi: 10.1038/s41467-018-05083-x; pmid: 29991676
101. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021). doi: 10.1016/j.cell.2021.04.048; pmid: 34062119
102. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329–337.e4 (2019). doi: 10.1016/j.cels.2019.03.003; pmid: 30954475
103. P. Majka *et al.*, Histology-Based Average Template of the Marmoset Cortex With Probabilistic Localization of Cytoarchitectural Areas. *Neuroimage* **226**, 117625 (2021). doi: 10.1016/j.neuroimage.2020.117625; pmid: 33301940
104. S. E. Jones, B. R. Buchbinder, I. Aharon, Three-dimensional mapping of cortical thickness using Laplace's equation. *Hum. Brain Mapp.* **11**, 12–32 (2000). doi: 10.1002/1097-0193(200009)11:1<12::AID-HBM20>3.0.CO;2-K; pmid: 10997850
105. C. Adamson, R. Beare, M. Walterfang, M. Seal, Software pipeline for midsagittal corpus callosum thickness profile processing: Automated segmentation, manual editor, thickness profile generator, group-wise statistical comparison and results display. *Neuroinformatics* **12**, 595–614 (2014). doi: 10.1007/s12021-014-9236-3; pmid: 24968872
106. G. Jefferis, S. E. Kemp, K. Müller, S. Arya, D. Mount, RANN: Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric (2024); https://CRAN.R-project.org/package=RANN.
107. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2015).
108. F. Pedregosa *et al.*, Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
109. J. Schindelin *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019; pmid: 22743772
110. R. Dear *et al.*, Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. *Nat. Neurosci.* **27**, 1075–1086 (2024). doi: 10.1038/s41593-024-01624-4; pmid: 38649755
111. D. DeTomaso, N. Yosef, Hotspot identifies informative gene modules across modalities of single-cell genomics. *Cell Syst.* **12**, 446–456.e9 (2021). doi: 10.1016/j.cels.2021.04.005; pmid: 33951459
112. O. Ben-Kiki, A. Bercovich, A. Lifshitz, A. Tanay, Metacell-2: A divide-and-conquer metacell algorithm for scalable scRNA-seq analysis. *Genome Biol.* **23**, 100 (2022). doi: 10.1186/s13059-022-02667-1; pmid: 35440087
113. S. Fischer, M. Crow, B. D. Harris, J. Gillis, Scaling up reproducible research for single-cell transcriptomics using MetaNeighbor. *Nat. Protoc.* **16**, 4031–4067 (2021). doi: 10.1038/s41596-021-00575-5; pmid: 34234317
114. D. van Dijk *et al.*, Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716–729.e27 (2018). doi: 10.1016/j.cell.2018.05.061; pmid: 29961576
115. X. N. Zuo *et al.*, An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci. Data* **1**, 140049 (2014). doi: 10.1038/sdata.2014.49; pmid: 25977800

ACKNOWLEDGMENTS

We thank the following individuals and platforms: the Single Cell Typing Platform and Non-human Primate Anatomical Research Platform of CEBSIT for assistance with adult transcriptomic data collection; the Single Cell Typing Platform of Shanghai Center for Brain Science and Brain-Inspired Technology as well as M. Li, Z. Hao, J. Wang, and H. Liu for assistance with developmental transcriptomic data collection and quality control; the 9.4T MRI core facility of CEBSIT for assistance with MRI imaging; and the Mouse, Marmoset, and Macaque Animal Facility of CEBSIT for animal care. **Funding:** This study was supported by National Science and Technology Innovation 2030 Major Program (grant nos. 2022ZD0205000 to Ci.L.; 2021ZD0200100 to Y.S., L.L., Z.Liu, Z.Lia., and Che.L.; and 2021ZD0203900 to Y.Zha.); the National Key Research and Development Program of China (2022YFC3400400 to X.X., 2022YFC3400405 to L.L., 2024YFC3408000 to Y.S., 2024YFF1206600 to S.H., and 2024YFA1803400 to Y.Zho.); the National Natural Science Foundation of China Outstanding Youth Foundation (T2422026 to Y.S.); the National Natural Science Foundation of China (nos. 32427802 and W2512020 to Ci.L., no. U23A6010 to Y.S., no. 32500590 to S.H., no. 32200921 to Y.Zha., and no. 32471024 to S.M.); China National Postdoctoral Program for Innovative Talents (BX20250145 to S.H.); Shanghai Science and Technology Development Funds (no. 23QA1410400 to Y.S.); a Yunnan Province grant (202305AH340006 to Y.-G.Y.); NHMRC grants (APP1194206 and APP2041019 to M.G.P.R. and APP2019011 to N.A.); a National Science Centre grant (2019/35/D/NZ4/03031 to P.M.); Lingang Laboratory grants (LGL-6672-08 to Ci.L. and LGL-6672-07 to Y.S.); the Natural Science Foundation of Shanghai (24ZR1473700 to S.M.); the Shanghai Pujiang Program (24PJA144 to S.M.); and the Shanghai Key Laboratory of Child Brain and Development (24dz2260100 to Ci.L.). **Author contributions:** Conceptualization: S.H., Ci.L., Y.S., Z.H., M.G.P.R.; Formal analysis: Z.H., Q.Y., She.L., X.Z., H.W., J.L.; Funding acquisition: Y.S., Ci.L., L.L., X.X., S.H., Z.Liu, Che.L., M.G.P.R., Y.-G.Y., N.A., P.M.; Investigation: Z.H., S.H., Y.S., Ci.L., Q.Y., She.L., X.Z., H.Q., H.W., J.L., M.G.P.R., P.M., N.A., Z.W., T.Y., Y.L., L.C., Cha.L., Che.L., W.W., Z.S.; Methodology: Z.H., Y.S., S.H., Ci.L., Q.Y., She.L., X.Z., H.W., J.L., Y.Zha., Y.W., P.M., N.A., W.W., Y.-G.Y., Y.Zho., S.M., Z.Lia., Shi.L., X.X., Z.Liu, X.W.; Project administration: Ci.L., S.H., Y.S.; Supervision: Ci.L., S.H., Y.S., M.P.; Writing – original draft: Ci.L., Y.S., S.H., Z.H.; Writing – review & editing: Ci.L., M.P., Y.S., M.G.P.R. **Competing interests:** The chip, procedure, and applications of Stereo-seq are covered in pending patents ("Array and method for detecting spatial information of nucleic acids," patent no. PCT/CN2020/090340). Employees of BGI have stock holdings in BGI. All the other authors declare no competing interests. **Data, code, and materials availability:** All data have been deposited in the database of China National GeneBank (https://db.genomics.cn/stomics/mccsta/). Raw gene expression matrices of snRNA-seq and Stereo-seq data for marmosets have been deposited in the CNGB Nucleotide Sequence Archive (access no. CNP0005746, https://db.cngb.org/search/project/CNP0005746). All scripts and codes used to generate our results are publicly accessible through a GitHub repository (https://github.com/marmosetbrainmapping/Code_Huang2026). Additionally, the pipeline for processing Stereo-seq raw data is available and regularly updated by the BGI group (https://github.com/STOmics/SAW).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aea2673
Figs. S1 to S15; Tables S1 to S8; MDAR Reproducibility Checklist

Submitted 6 July 2025; accepted 8 January 2026

10.1126/science.aea2673

An opposing molecular gradient axis underlies primate cortical organization

Zhi Huang, Qianqian Yang, Shenglong Li, Xiaojia Zhu, He Wang, Jixuan Lin, Yafeng Zhan, Yan Wu, Zefang Wang, Piotr Majka, Haichao Qu, Nafiseh Atapour, Tao Yang, Youning Lin, Luman Cui, Yong-Gang Yao, Zhifeng Liang, Zhen Liu, Chao Li, Wu Wei, Yi Zhou, Shaojie Ma, Zhiming Shen, Xiaoyu Wei, Xun Xu, Shiping Liu, Chengyu Li, Muming Poo, Longqi Liu, Marcello G. P. Rosa, Yidi Sun, Shijie Hao, and Cirong Liu

Science **392** (6795), eaea2673. DOI: 10.1126/science.aea2673

Editor's summary

The expansion of the cerebral cortex during primate evolution has resulted in remarkable complexity in the cellular organization and connectivity of this brain region, yet the fundamental principles governing the organization of the expanded cortex remain unclear. Huang *et al*. developed a single-cell whole-brain spatial transcriptome database, applied retrograde neuronal tracing, and used functional magnetic resonance imaging datasets to construct a three-dimensional integrated multimodal atlas of the entire marmoset brain. Two distinct molecular gradients emanated from allocortical and primary sensory regions, converging at association cortices. Cross-species comparisons revealed similarities and differences among marmosets, macaques, and humans. The study provides a comprehensive, multimodal characterization of primate cortical organization. —Mattia Maroso