

Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography

Joshua B. Burt¹, Murat Demirtaş², William J. Eckner¹, Natasha M. Navejar³, Jie Lisa Ji⁴, William J. Martin⁵, Alberto Bernacchia⁶, Alan Anticevic² & John D. Murray^{1,2}

¹*Department of Physics, Yale University, New Haven, CT*

²*Department of Psychiatry, Yale University School of Medicine, New Haven, CT*

³*Tulane University, New Orleans, LA*

⁴*Interdepartmental Neuroscience Program, Yale University, New Haven, CT*

⁵*BlackThorn Therapeutics*

⁶*University of Cambridge, Cambridge, United Kingdom*

Hierarchy provides a unifying principle for the macroscale organization of anatomical and functional properties across primate cortex, yet the microscale bases of specialization across human cortex are poorly understood. Anatomical hierarchy is conventionally informed by invasive histological measurements, creating the need for a principled proxy measure in humans. Moreover, cortex exhibits marked interareal variation in gene expression, yet organizing principles of its transcriptional architecture remain unclear. We hypothesized that specialization of human cortical microcircuitry involves hierarchical gradients of gene expression. We found that a noninvasive neuroimaging measure, the MRI-derived T1w/T2w map, reliably indexes anatomical hierarchy, and captures the dominant pattern of transcriptional variation across human cortex. We found strong hierarchical gradients in expression profiles of genes related to microcircuit function, [in line with monkey microanatomical measurements](#), and genes implicated in neuropsychiatric disorders. Our findings suggest that a hierarchical axis links the transcriptional and anatomical architectures of cortex, along which gradients of microscale properties contribute to macroscale specialization of cortical function.

The neocortex of human and nonhuman primates exhibits interareal patterns of structural and functional variation. Cortical areas are distinguished by differences in their cellular composition, laminar differentiation, and long-range anatomical connectivity. Primate cortex is characterized by large-scale gradients of specialization in physiology and function, including in representational selectivity^{1–3} and dynamics of intrinsic activity^{4,5}. Recent advances in large-scale high-throughput transcriptomics, which can produce genome-wide atlases of spatially distributed gene expression profiles, have also revealed a rich transcriptional architecture in humans characterized by spatially heterogeneous gene expression levels across brain areas^{6–8}. Interareal transcriptional diversity has been related to differences in cortical function, including the spatiotemporal structure of intrinsic network activity^{7,9–11}, and to spatially heterogeneous patterns of anatomical connectivity^{11,12}. Yet unifying principles for the macroscale organization of structural, functional, and transcriptional differences across primate cortex are still unknown.

A parsimonious principle for the large-scale anatomical and functional organization of nonhuman primate cortex is the concept of cortical hierarchy^{2–4,13–16}. Anatomical hierarchy, defined as a globally self-consistent ordering of cortical areas according to characteristic laminar patterns of interareal projections, has been studied extensively in monkeys through histological tract-tracing methods^{13–15}. The ordering of cortical areas along the anatomical hierarchy, which situates early sensory areas toward the bot-

tom and higher-order association areas toward the top of the hierarchy, has also been found to align with areas' functional organization in sensory processing hierarchies^{13,15}. We hypothesized that the transcriptional architecture of human cortex is also hierarchically organized, such that the functional specialization of human cortical microcircuitry involves hierarchical gradients of gene expression levels. However, the highly invasive nature of the tract-tracing data acquisition procedures which are required to index hierarchy in nonhuman primates has thus far precluded analogous investigations of cortical organization in humans, thereby creating the need for noninvasive alternative measures.

To address these open questions, we analyzed transcriptional, anatomical, and neuroimaging data from humans and monkeys to study the hierarchical organization of cortical microcircuit specialization. We found that a structural neuroimaging measure, the MRI-derived T1w/T2w map¹⁷, provides a noninvasive proxy for anatomical hierarchy in primate cortex. To test for hierarchical gradients in gene expression, we then compared the spatial expression profiles of genes in the Allen Human Brain Atlas (AHBA) to the topography of the human T1w/T2w map. We found strong hierarchical gradients in expression profiles of genes related to synaptic physiology, cell-type specificity, and cortical cytoarchitecture, in line with monkey microanatomical measurements. Furthermore, we observed a remarkably close topographic correspondence between the T1w/T2w map and the dominant spatial pattern of gene expression variation across human cortex. Fi-

nally, we found that hierarchically patterned genes are preferentially associated with functional processes and brain disorders. Overall, these findings suggest that the transcriptional and anatomical architectures of human cortex share a common principal axis related to hierarchy, and that hierarchical gradients of microscale properties shape the macroscale specialization of cortical function.

Results

T1w/T2w maps noninvasively capture anatomical hierarchy

To enable the study of hierarchy in human cortex, we first sought to establish a neuroimaging measure that can serve as a noninvasive proxy for indexing anatomical hierarchy. One measure we examined was the cortical T1w/T2w map, a structural neuroimaging map defined by the contrast ratio of T1- to T2-weighted (T1w/T2w) magnetic resonance images. The cortical T1w/T2w map has been proposed as an *in vivo* measure that is sensitive to regional variation in gray-matter myelin content, based on its close correspondence with myelin stained sections in histological validation studies, and its recapitulation of known neuroanatomical borders between cyto- and myelo-architecturally delineated areas^{17,18} (see Discussion). Motivated by the empirical observation that T1w/T2w map values are high in primary sensory cortex (visual, somatosensory, auditory) and low in association cortex, homologously in human and monkey (Fig. 1a–c, Extended Data

Fig. 1), and stably across individuals, we hypothesized that the group-averaged cortical T1w/T2w map, through an inverse relationship with hierarchy, provides a noninvasive correlate for areas' hierarchical positions.

We validated the T1w/T2w map as a proxy for hierarchy in monkey cortex by comparing T1w/T2w map values to model-estimated anatomical hierarchy levels, derived from conventional tract-tracing approaches that quantify long-range interareal projections and their laminar specificity¹⁵. These laminar connectivity data, based only on direct cortico-cortical projections, are used to specify a globally optimal hierarchical ordering of cortical areas, such that lower areas send feedforward projections to higher areas, and higher areas send feedback projections to lower areas^{13–15,19} (Extended Data Fig. 2). Feed-forward and feedback projections primarily originate from the supragranular and infragranular cortical layers, respectively^{13,15}. At the level of individual projections, we found that the difference in T1w/T2w map values between connected areas is correlated with the laminar feedforward/feedback structure of the connection (Fig. 1d), more strongly in high-T1w/T2w sensory areas than in low-T1w/T2w association areas (Extended Data Fig. 3). Global anatomical hierarchy levels are estimated by fitting a generalized linear model to pairwise laminar projection data^{15,19} (see Methods). We found a strong negative correlation between model-estimated anatomical hierarchy level and T1w/T2w map value ($r_s = -0.76, P < 10^{-5}$; Spearman rank correlation) (Fig. 1e,f).

How well does the T1w/T2w map capture estimated anatomical hierarchy levels, relative to other putative proxy measures? We compared the performance of the T1w/T2w map against two alternative proxy candidates derived from structural MRI²⁰: the map of cortical thickness, as cortex is generally thicker in association cortex than sensory cortex; and the map of geodesic distance from primary visual cortex, which defines a posterior-anterior gradient. We found that the T1w/T2w map was more strongly correlated with model-estimated hierarchy than were either of the two other candidate proxy measures (Fig. 2). Furthermore, we found that the parcellated T1w/T2w topography is highly stable across individuals. The mean pairwise Spearman rank correlation between maps of individuals ($N = 339$) was 0.94 for the T1w/T2w map, in contrast to a correlation of 0.76 for the thickness map. These findings, and the observed inter-species homology¹⁸, support the use of the T1w/T2w map as a noninvasive correlate of hierarchy across human cortex, for which lack of invasive tract-tracing data precludes a more direct characterization of anatomical hierarchy per conventional approaches.

Hierarchical gradients in cortical microcircuit specialization

We hypothesized that the large-scale organization of cortical microcircuit specialization—that is, regional variation in synaptic and cellular composition with functional relevance—may involve hierarchical gradients of gene expression levels across human cortex. To test

this hypothesis, we examined areal patterns of cortical gene expression variation from the AHBA in relation to the T1w/T2w map. The AHBA is a transcriptional atlas that contains gene expression levels measured with DNA microarray probes and sampled from hundreds of neuroanatomical structures in the left cortical hemisphere across six normal post-mortem human brains⁶. From these data, we calculated group-averaged gene expression profiles across 180 unilateral cortical areas using a multimodal parcellation from the Human Connectome Project²¹ (Fig. 3, see Methods). Because of the strong inverse relationship found between the T1w/T2w map and model-estimated anatomical hierarchy (Fig. 1), if gene expression levels negatively correlate with T1w/T2w map values across cortical areas, then by extension, gene expression levels increase along the anatomical hierarchy—i.e., from sensory to association cortex—and thus exhibit a positive hierarchical gradient; conversely, a positive correlation between T1w/T2w map values and expression levels indicates decreasing gene expression along the hierarchy, or a negative hierarchical gradient. To support the validity of our interpretations, we compared the T1w/T2w map correlation (TMC) with microcircuitry-related gene expression profiles in human cortex to TMCs with more direct anatomical measures in monkey cortex, with a focus on cytoarchitecture, inhibitory interneuron densities, and synaptic processes (Fig. 4).

An established feature of microcircuit specialization that varies along cortical hier-

archy is the degree of laminar differentiation in local cytoarchitecture²²: primary sensory cortex is highly laminated and exhibits a thick and well-defined granular layer, whereas association cortex is characterized by decreasing laminar differentiation and a gradual loss of the granular layer with progression up the hierarchy. In monkey cortex, we found that areas' cytoarchitectural types²² correlate strongly with their T1w/T2w map values (Fig. 4a). In human cortex, we examined average expression profiles of genes reported to be preferentially expressed in specific cortical layers²³. Consistent with the cytoarchitectural trends observed in monkey cortex, we found a positive TMC for granular (L4) layer-specific genes, and negative TMCs for supra- (L1–3) and infra-granular (L5/6) layer-specific genes (Fig. 4b,c). These findings demonstrate that the noninvasive T1w/T2w map captures anatomical gradients related to cortical hierarchy in humans and nonhuman primates.

To gain further insight into microcircuit bases of hierarchical specialization, we examined the spatial distributions of markers for different inhibitory interneuron cell types. Inhibitory interneuron cell types fall into several biophysically distinct classes which differ in their synaptic connectivity patterns, morphology, electrophysiology, and functional roles^{24,25}. In monkey cortex, we found that immunohistochemically measured densities of parvalbumin- and calretinin-expressing interneurons exhibit positive and negative TMCs, respectively (Fig. 4d). Consistent with these results, in human cortex we found

corresponding positive and negative hierarchical gradients in the expression profiles for the genes which code for parvalbumin and calretinin (Fig. 4e). In general, we observed strong hierarchical gradients in transcriptional markers for a number of inhibitory interneuron cell types²⁴ (Fig. 4f), as well as for composite gene expression profiles associated with specific neuronal cell types derived from RNA sequencing in individual human neurons²⁶ (Extended Data Fig. 4). These findings suggest that hierarchical gradients in neuronal cell-type distributions may contribute to sensory–association specialization of cortical microcircuit function.

Gradients in the composition of synapses may endow cortical areas with diverse physiological properties required to perform the various computations which underlie specialized cognitive and behavioral functions. For instance, local increases in the strength of recurrent excitatory connectivity may endow cortical circuits in association cortex with extended temporal integration supporting cognitive computations^{4,19,27}. One putative microanatomical correlate for the strength of recurrent synaptic excitation in local cortical microcircuits is the number of excitatory synapses on pyramidal neurons, which can be quantified by counting the number of spines on pyramidal cell dendrites. In monkey cortex, we found a strong negative TMC for basal-dendritic spine counts on cortical pyramidal neurons²⁸ (Fig. 4g). This finding suggests a gradient of increasing local recurrent excitation strength along the cortical hierarchy in primates¹⁹.

Distinct subunits of synaptic receptor proteins that mediate neurotransmission are differentially expressed across neuronal cell types and produce physiologically diverse synaptic properties. In the AHBA dataset, we examined expression profiles of genes that code for various excitatory and inhibitory synaptic receptor subunits (Fig. 4h–j). The gene *GRIN2B*, which codes for a glutamatergic NMDA receptor subunit that mediates synaptic excitation preferentially in association cortex²⁹, exhibited a strong negative TMC. This results suggests increased recurrent excitatory strength in association cortical areas and is consistent with the spine count gradient observed in monkey. Gene sets coding for neuromodulatory synaptic receptor subunits also contain strong positive and negative hierarchical gradients (Extended Data Fig. 5). The positive and negative TMCs reported in Fig. 4i,j suggest that gradients in local excitatory and inhibitory synaptic machinery contribute to the functional specialization of cortical microcircuitry^{4,19}.

T1w/T2w topography captures the dominant axis of transcriptional variation across human cortex

How well does the T1w/T2w map capture areal variation in the transcriptional architecture of human cortex in general? We performed principal component analysis (PCA) to identify the dominant areal patterns underlying gene expression variation (Fig. 5a–e, Extended Data Fig. 6). To test for generality of effects, we analyzed categorical sets of genes which are preferentially expressed in human brain tissue, neurons, oligodendrocytes, and

synaptic compartments^{30,31}. To assess statistical significance of effects, we developed a novel method for spatial autocorrelation-preserving permutation testing to generate surrogate maps (Extended Data Fig. 8, see Methods). The first principal component (PC1) is the spatial map that captures the maximal amount of total gene expression variance across cortical areas (Fig. 5a). Across all five gene sets, PC1 captures a large fraction of gene expression variance (range: 21–27%, more than twice PC2) (Fig. 5b, Extended Data Fig. 6), revealing that cortical gene expression patterns are effectively low-dimensional.

Remarkably, we found that T1w/T2w map topography is strongly correlated with PC1, i.e., the dominant axis of gene expression variation, across all tested gene sets (TMC range: 0.80–0.81; $P < 10^{-5}$) (Fig. 5c,d). Like the T1w/T2w map, PC1 exhibits relatively high values in primary sensorimotor areas and low values in association areas (Extended Data Fig. 7), consistent with a prior report on a subset of the AHBA dataset⁶. We can also quantify how much gene expression variance is captured by the T1w/T2w map (see Methods). We found that across all gene sets the T1w/T2w map captures more than half as much variance as PC1, which by construction is the spatial map that captures the maximum possible gene expression variation (Fig. 5e). We then compared performance of the T1w/T2w map against the two alternative candidate proxy maps, cortical thickness and geodesic distance from primary visual cortex (Fig. 6). Across all gene sets, the T1w/T2w map was more strongly correlated with PC1 and captured more gene expression variance

than either alternative map. The close alignment between T1w/T2w map topography and spatial gene expression variation suggests that the dominant axis of transcriptional variation in human cortex relates to hierarchy. Furthermore, the robustness of our findings across gene sets demonstrates that this axis captures areal variation in general across a number of neurobiological processes.

What effects may driving the deviating outliers from the otherwise strikingly strong correspondence between gene expression PC1 and the T1w/T2w map shown in Fig. 5c? We constructed a map of the absolute deviation (i.e., residual) of each cortical area from the best-fit line illustrated in Fig. 5c (Extended Data Fig. 9a), and noticed that the large residuals were preferentially located in areas of cortex with large gradients in local T1w/T2w map values (Extended Data Fig. 1a). Due to sparse cortical sampling in the AHBA (203 ± 27 samples per subject, 1220 total across six subjects), substantial spatial interpolation was required to produce our parcellated gene expression maps (Fig. 3). We therefore hypothesized that large discrepancies were due to relatively poor gene expression estimates in regions with large T1w/T2w gradients. To test this quantitatively, we computed for each area a measure of local T1w/T2w gradient and compared these values to the PC1 residuals. We indeed found a strong correlation between local T1w/T2w gradient and PC1 residual magnitude ($r_p = 0.70, P < 10^{-5}$; Pearson correlation) (Extended Data Fig. 9b,c). Our prediction was further supported by a validation our key results with

the older group-averaged ($N = 69$) T1w/T2w map, from the Conte69 dataset¹⁷. Results were highly consistent, and the Conte69 T1w/T2w map tended to yield stronger TMC values than did the HCP T1w/T2w map (Extended Data Fig. 10). The HCP map contains more sharply separated T1w/T2w values among neighboring cortical parcels, compared to the smoother Conte69 map (autocorrelation space constant: 6.16 mm for HCP vs. 7.39 mm for Conte69), likely due to differences in areal registration²¹ and smoothing¹⁷. Together, these findings suggest that the remarkable relationships reported in this study, between cortical structure and transcription in humans, may be systematically underestimated due to the limited spatial resolution in the AHBA dataset.

Stably expressed genes preferentially exhibit hierarchical gradients

Genes that are especially vital to normal healthy cortical function may be more likely to have consistent spatial expression profiles across individual subjects. Hawrylycz and colleagues defined differential stability (DS) as the mean pairwise correlation between subjects' individual gene expression profiles, which they found predicts association with key neurobiological functions when computed across all (i.e., cortical and subcortical) brain structures⁷. We found a strong nonlinear and positive relationship between cortical DS (DS_c) and TMC magnitude (Fig. 7a). To gain additional insight into this relationship, we explored the impact of filtering genes through progressively higher DS_c thresholds on the

TMC distribution. Exclusion of low-DS_c genes greatly alters the shape of the TMC distribution, collapsing the prominent peak centered near zero while progressively producing two roughly symmetric bimodal peaks at strong TMCs (Fig. 7b). Furthermore, exclusion of low-DS_c genes strongly increases the fraction of transcriptional variance captured by PC1 (Fig. 7c), rendering gene expression patterns more quasi-one-dimensional. **Together, these results suggest that high-DS_c genes, i.e. genes which exhibit highly stable spatial patterns of expression within cortex across individuals, preferentially exhibit strong positive and negative hierarchical gradients.**

Hierarchically expressed genes are enriched for functional and disease annotations

To examine the functional roles of genes with strong hierarchical gradients, we tested for their preferential enrichment in gene sets defined by functional and disease ontologies. We found that genes with stronger TMCs are enriched in more functional categories, relative to genes with weaker TMCs, for all functional gene ontologies tested^{7,32}: biological processes, cellular components, molecular functions, microRNA binding sites, and drug targets (Fig. 8a). These results suggest that diverse key cell-biological processes contribute to hierarchical differentiation of cortical microcircuitry. **Finally, we examined whether hierarchical expression is a preferential property found in group-averaged profiles of genes associated with psychiatric and neurological disorders.** For instance, we

found that the genes *APOE* and *SNCA*, which are strongly linked to Alzheimer's and Parkinson's diseases, respectively³³, exhibit robust negative TMCs and are therefore more highly expressed in association cortex (Fig. 8b,c). For a systematic examination, we statistically quantified the enrichment of genes with strong hierarchical variation in disease-related gene sets⁷, obtained from the DisGeNet database³⁴. We found that genes with strongly negative TMCs were significantly over-represented across multiple disease-related gene sets (Fig. 8d). In particular, gene sets for schizophrenia, bipolar disorder, autistic disorders, and depressive disorders are significantly enriched with strongly negative TMC genes which are more highly expressed in association cortex. These findings suggest that brain disorders involve differential impacts to areas along the cortical hierarchy.

Discussion

Taken together, our findings show that multiple complementary measurement approaches reveal a robust hierarchical organization of microscale variation that may contribute to the macroscale specialization of primate cortical function. First, the MRI-derived T1w/T2w map provides a noninvasive neuroimaging proxy for anatomical hierarchy in the absence of axonal tract-tracing data. Second, the principal axis of transcriptional variation across human cortex aligns with cortical hierarchy as captured by the T1w/T2w map. Third, this hierarchical axis reflects a gradient of local microcircuit specialization involving synapses

and cell types, with relevance to brain disease pathophysiology. Strong similarities between the patterns of anatomical, functional, and transcriptional variation suggest that hierarchical gradients of microcircuit properties play key roles in the functional specialization of large-scale networks across the human cortex. Moreover, the agreement between human transcriptional and monkey anatomical measurements suggests that hierarchy is a general conserved organizing principle in human and nonhuman primate cortex.

Specialization of cortical function may derive in part from the multiple features of microcircuitry identified here to exhibit hierarchical gradients. For instance, stronger recurrent excitation in association cortex can endow association circuits with longer timescales of intrinsic activity^{19,35}, as observed empirically^{4,5}, which subserve the prolonged integration of signals in these areas^{3,16,27}. Furthermore, computational modeling of cortical circuits identifies recurrent excitation strength as a key property governing functional specialization across areas for core cognitive computations such as working memory and decision making^{27,35}. Hierarchical gradients of inhibitory interneuron cell types can additionally shape regional specialization of dynamics and function, due to cell-type differences in physiology and synaptic connectivity^{24,25}. For example, parvalbumin-expressing inhibitory interneurons preferentially target the perisomatic areas of pyramidal neurons where they can gate pyramidal-neuron outputs. In contrast, calretinin-expressing in-

hibitory interneurons preferentially target distal dendrites of pyramidal neurons and other inhibitory interneurons, where they may play key computational roles in disinhibition-mediated gating of dendritic inputs³⁶. Cytoarchitectural differences between areas correlate with their pairwise laminar projection profiles^{14,22}, linking local microcircuit specialization of areas to their hierarchical long-range interactions.

Our study adds to a growing understanding of how transcriptomic specialization shapes cortical function. Transcriptional diversity, particularly of genes which regulate synaptic function and ion channel activity, relates to the spatiotemporal organization of intrinsic activity in large-scale cortical networks^{7,9–11}, and transcriptional markers for synaptic, neuronal, and axonal structure relate to patterns of anatomical connectivity^{11,12}. Of note, Hawrylycz et al. (2015) found that genes most strongly predictive of functional connectivity patterns in cortex were shifted toward high DS_c, and that across all brain regions, high-DS genes were significantly enriched in gene sets related to functional ontologies and brain diseases, leading the authors to suggest these genes constitute a “canonical transcriptional blueprint” for the human brain⁷. We found that high-DS_c genes exhibit strong hierarchical gradients (Fig. 7a), and that these strong-TMC genes exhibit similar functional and brain disease-related enrichments. These results suggest that hierarchically and stably expressed genes across the cortex contribute significantly to the transcriptional regulation of cortical function and to its pathophysiological disruption in

disease.

Our findings demonstrate that the T1w/T2w map generally captures an axis of hierarchical differentiation across cortex that reflects multiple features of interareal variation. The T1w/T2w map—an MR contrast map that removes shared imaging intensity biases and increases image contrast—is sensitive to gray-matter myelin content¹⁷, which may itself contribute to functional specialization in several ways¹⁸. However, both T1- and T2-weighted image intensities depend on multiple MRI parameters, each of which is sensitive to several other brain microstructural properties, including cell size and density, degree of dendritic arborization, iron, and water^{37,38}. Further *in vivo* characterization of microstructural variation can be provided by quantitative MRI techniques such as T1 mapping^{sereno13, lutti14, 39,40}. Thus the T1w/T2w map provides a readily acquired, noninvasive neuroimaging measure which is sensitive to areal variation in not one but several structural components of local cortical microarchitecture. We note that there are interesting deviations between the topographies of the T1w/T2w map and other hierarchical features. For instance, primary motor cortex and retrosplenial cortex exhibit high T1w/T2w map values yet differ from primary sensory areas in their laminar structure^{17,18}.

Multiple functionally defined hierarchies in human and nonhuman primate cortex have been proposed, none of which are mutually exclusive with the anatomical hierarchy informed by long-range projection patterns. For instance, studies have found hierarchical

differences across areas in the temporal selectivity of spontaneous dynamics and sensory processing^{4,5}, but it remains unclear how these differences relate to hierarchies of microcircuit specialization. In this study, we have identified multiple microanatomical and transcriptional properties of cortical microcircuitry which exhibit hierarchical gradients and may contribute to physiological and functional specialization. Importantly, cortical function has a complex multidimensional organization, and only one axis of its areal variation can be captured by a single scalar-valued map. Distinct information processing hierarchies can be defined for different sensory modalities, with within a modality, such as the dorsal and ventral processing streams within the primate visual system^{wandell07, 13,15}. Future studies can investigate how integration of multiple neuroimaging measures, for instance combining T1w/T2w imaging with diffusion weighted imaging, can reveal new multidimensional principles of cortical organization, both within and across functionally specialized networks.

Multiple lines of evidence point to a transcriptional basis for disease phenotypic variation, linking white matter dysconnectivity⁴¹ and developmental changes in structural topology⁴² to genes implicated in schizophrenia. Further characterization of the developmental trajectory of hierarchical transcriptomic specialization,^{41,43,44} and structural brain tissue degeneration⁴⁵, may inform the progression of neurodevelopmental disorders. Strong hierarchical gradients in highly differentially stable drug targets, such as

receptor subunits, could enable preferential modulation of sensory or association cortical areas, at the group level, through targeted pharmacology. This may guide future rational design of drug treatments to target specific macroscale cortical circuits. Large-scale mapping of the cortical transcriptome at finer spatial resolution will further elucidate the microcircuit basis of hierarchical specialization with laminar²³ and cell-type^{8,26} specificity.

Acknowledgements We thank B.D. Fulcher, X.-J. Wang, R. Chaudhuri, and D.C. Glahn for useful discussions. This research was supported by NIH grants R01MH112746, R01MH108590, and TL1TR000141, and BlackThorn Therapeutics.

Author Contributions J.B.B., W.J.M., A.B., A.A. and J.D.M. designed the research. J.B.B., M.D., W.J.E., N.N., and L.J. analyzed the data. J.D.M. supervised the project. J.B.B. and J.D.M. wrote the manuscript and prepared the figures. All authors contributed to editing the manuscript.

Author Information Correspondence and requests for materials should be addressed to J.D.M. (john.murray@yale.edu).

1. Lennie, P. Single units and visual cortical organization. *Perception* **27**, 889–935 (1998).
2. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A hierarchy of temporal receptive windows in human cortex. *J Neurosci* **28**, 2539–2550 (2008).
3. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* **31**, 2906–15 (2011).
4. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* **17**, 1661–3 (2014).
5. Honey, C. J. *et al.* Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
6. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–9 (2012).
7. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat Neurosci* **18**, 1832–44 (2015).
8. Lein, E. S., Belgard, T. G., Hawrylycz, M. & Molnár, Z. Transcriptomic perspectives on neocortical structure, development, evolution, and disease. *Annu Rev Neurosci* **40**, 629–652 (2017).

9. Wang, G.-Z. *et al.* Correspondence between resting-state activity and brain gene expression. *Neuron* **88**, 659–66 (2015).
10. Krienen, F. M., Yeo, B. T. T., Ge, T., Buckner, R. L. & Sherwood, C. C. Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proc Natl Acad Sci U S A* **113**, E469–78 (2016).
11. Richiardi, J. *et al.* Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–4 (2015).
12. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc Natl Acad Sci U S A* **113**, 1435–40 (2016).
13. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* **1**, 1–47 (1991).
14. Barbas, H & Rempel-Clower, N. Cortical structure predicts the pattern of corticocortical connections. *Cereb Cortex* **7**, 635–646 (1997).
15. Markov, N. T. *et al.* Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J Comp Neurol* **522**, 225–59 (2014).
16. Badre, D. & D'Esposito, M. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci* **10**, 659–669 (2009).
17. Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J Neurosci* **31**, 11597–616 (2011).
18. Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E. & Van Essen, D. C. Trends and properties of human cerebral cortex: correlations with cortical myelin content. *Neuroimage* **93 Pt 2**, 165–75 (2014).
19. Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H. & Wang, X.-J. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* **88**, 419–31 (2015).
20. Wagstyl, K., Ronan, L., Goodyer, I. M. & Fletcher, P. C. Cortical thickness gradients in structural hierarchies. *Neuroimage* **111**, 241–50 (2015).
21. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–8 (2016).
22. Hilgetag, C. C., Medalla, M., Beul, S. F. & Barbas, H. The primate connectome in context: Principles of connections of the cortical visual system. *Neuroimage* **134**, 685–702 (2016).
23. Zeng, H. *et al.* Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–96 (2012).
24. Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* **5**, 793–807 (2004).

25. Kepcs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–26 (2014).
26. Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–90 (2016).
27. Wang, X.-J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* **24**, 455–463 (2001).
28. Elston, G. N. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb Cortex* **13**, 1124–38 (2003).
29. Wang, H., Stradtman 3rd, G. G., Wang, X.-J. & Gao, W.-J. A specialized NMDA receptor function in layer 5 recurrent microcircuitry of the adult rat prefrontal cortex. *Proc Natl Acad Sci U S A* **105**, 16791–16796 (2008).
30. Genovese, G. *et al.* Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci* **19**, 1433–1441 (2016).
31. Pirooznia, M. *et al.* SynaptomeDB: an ontology-based knowledgebase for synaptic genes. *Bioinformatics* **28**, 897–9 (2012).
32. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **37**, W305–11 (2009).
33. Bras, J. *et al.* Genetic analysis implicates APOE, SNCA and suggests lysosomal dysfunction in the etiology of dementia with Lewy bodies. *Hum Mol Genet* **23**, 6139–46 (2014).
34. Piñero, J. *et al.* DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res* **45**, D833–D839 (2017).
35. Murray, J. D., Jaramillo, J. H. & Wang, X.-J. Working memory and decision making in a fronto-parietal circuit model 2017.
36. Yang, G. R., Murray, J. D. & Wang, X.-J. A dendritic disinhibitory circuit mechanism for pathway-specific gating. *Nat Commun* **7**, 12815 (2016).
37. Lorio, S. *et al.* Neurobiological origin of spurious brain morphological changes: a quantitative MRI study. *Human brain mapping* **37**, 1801–1815 (2016).
38. Stuber, C *et al.* Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. *Neuroimage* **93** (Pt. 1), 95–106 2014.
39. Carey, D. *et al.* Quantitative MRI provides markers of intra-, inter-regional, and age-related differences in young adult cortical microstructure. *NeuroImage* (2017).
40. Gomez, J. *et al.* Microstructural proliferation in human cortex is coupled with the development of face processing. *Science* **355**, 68–71 (2017).
41. Whitaker, K. J. *et al.* Adolescence is associated with genetically patterned consolidation of the hubs of the human brain connectome. *Proc Natl Acad Sci U S A* **113**, 9105–10 (2016).

42. Romme, I. A. C., de Reus, M. A., Ophoff, R. A., Kahn, R. S. & van den Heuvel, M. P. Connectome Disconnection and Cortical Gene Expression in Patients With Schizophrenia. *Biol Psychiatry* **81**, 495–502 (2017).
43. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
44. Bakken, T. E. *et al.* A comprehensive transcriptional map of primate brain development. *Nature* **535**, 367–75 (2016).
45. Yeatman, J. D., Wandell, B. A. & Mezer, A. A. Lifespan maturation and degeneration of human brain white matter. *Nature communications* **5**, 4932 (2014).

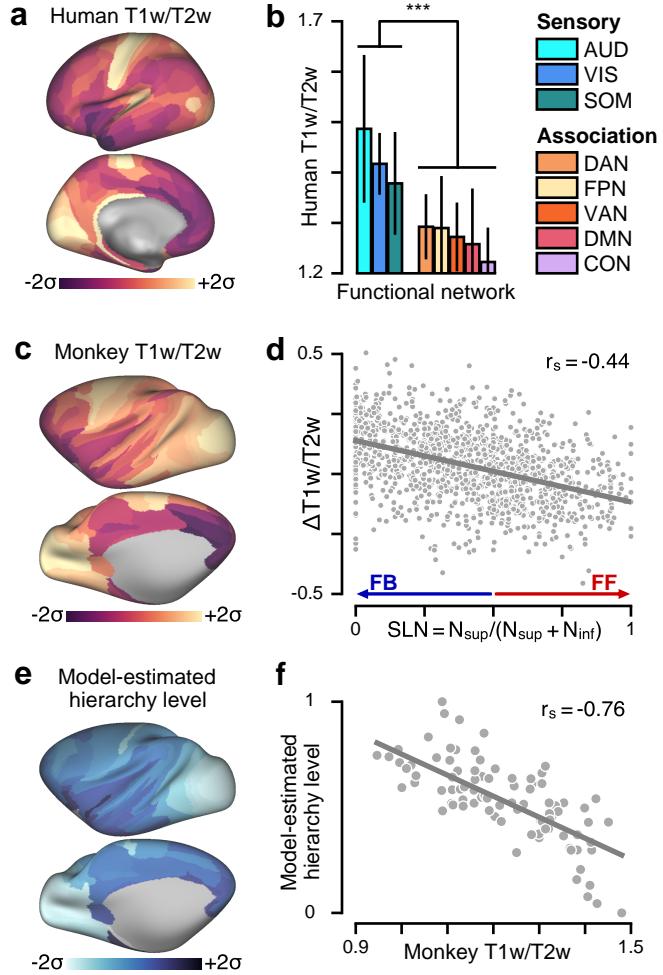


Figure 1: T1w/T2w neuroimaging maps noninvasively capture the hierarchical organization of primate cortex. **(a)** The parcellated group-averaged ($N = 339$) human T1w/T2w map exhibits high values in primary sensory cortical areas relative to association areas. **(b)** Human T1w/T2w map values are significantly lower in functionally defined association networks than in sensory networks ($P < 10^{-3}$; Wilcoxon signed-rank test) (Extended Data Fig. 1c,d). Error bars mark the std. dev. across areas. **(c)** The parcellated group-averaged ($N = 19$) macaque monkey T1w/T2w map topography is similar to that of the human. **(d)** Interareal variation in the T1w/T2w map correlates with the laminar specificity of directed feedforward (FF) and feedback (FB) projections in monkey cortex, as quantified by the fraction of labeled supragranular layer neurons (SLN) in the source area. High and low SLN correspond to FF and FB projection motifs, respectively. SLN significantly correlates with the difference (target minus source) in areal T1w/T2w map values ($r_s = -0.44, P < 10^{-5}$; Spearman rank correlation). **(e)** Anatomical hierarchy levels across cortical areas are estimated by fitting a generalized linear model to predict projections' SLNs as a function of pairwise hierarchical distance. **(f)** Model-estimated anatomical hierarchy levels are highly anti-correlated with T1w/T2w map values in monkey cortex ($r_s = -0.76, P < 10^{-5}$).

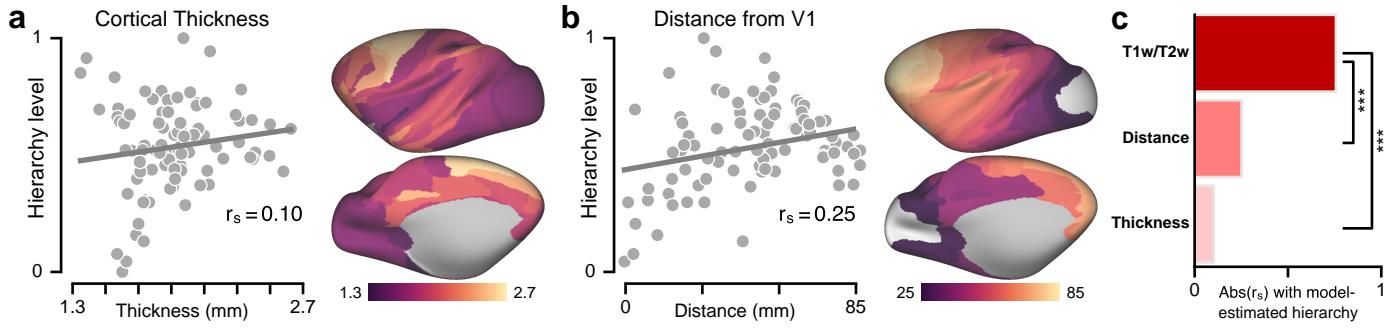


Figure 2: Model-estimated anatomical hierarchy in monkey cortex is better captured by the group-averaged T1w/T2w map than by two other candidate proxy measures derived from structural MRI. (a) Correlation between hierarchy and cortical thickness. (b) Correlation between hierarchy and geodesic distance from primary visual cortex (V1), which follows a rostro-caudal gradient. (c) Comparison of hierarchy correlation values for the T1w/T2w map, thickness map, and distance from area V1. The T1w/T2w map is much more strongly correlated with model-estimated anatomical hierarchy than the other two maps ($P < 10^{-3}$). Statistical significance is calculated by a test of the difference between dependent correlations (*, $P < 0.05$; **, $P < 10^{-2}$; *, $P < 10^{-3}$).**

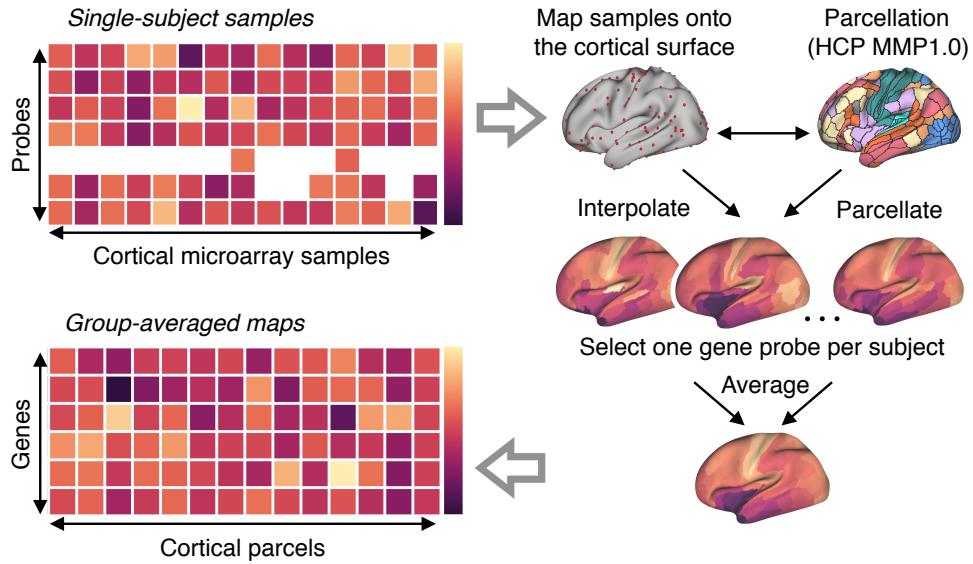


Figure 3: Procedure for generating group-averaged parcellated maps of gene expression levels. All analyses of gene expression patterns used group-averaged parcellated expression maps derived from the Allen Human Brain Atlas (AHBA) (see Methods for details). The AHBA contains genes expression levels measured with DNA microarray probes and sampled from hundreds of neuroanatomical structures in the left hemisphere across six normal post-mortem human brains. First, cortical samples for each subject were mapped from volumetric space onto that subject's native reconstructed two-dimensional cortical surface. Second, parcellated gene expression maps were constructed, for each subject, [using the Human Connectome Project's \(HCP\) Multi-Modal Parcellation \(MMP1.0\)](#) of the left cortical surface into 180 contiguous areas. For genes profiled by multiple microarray probes, we selected a single representative probe for each subject. Finally, a group-level parcellated expression map for each unique gene was computed by averaging parcellated expression levels across subjects' selected gene probes (see Methods).

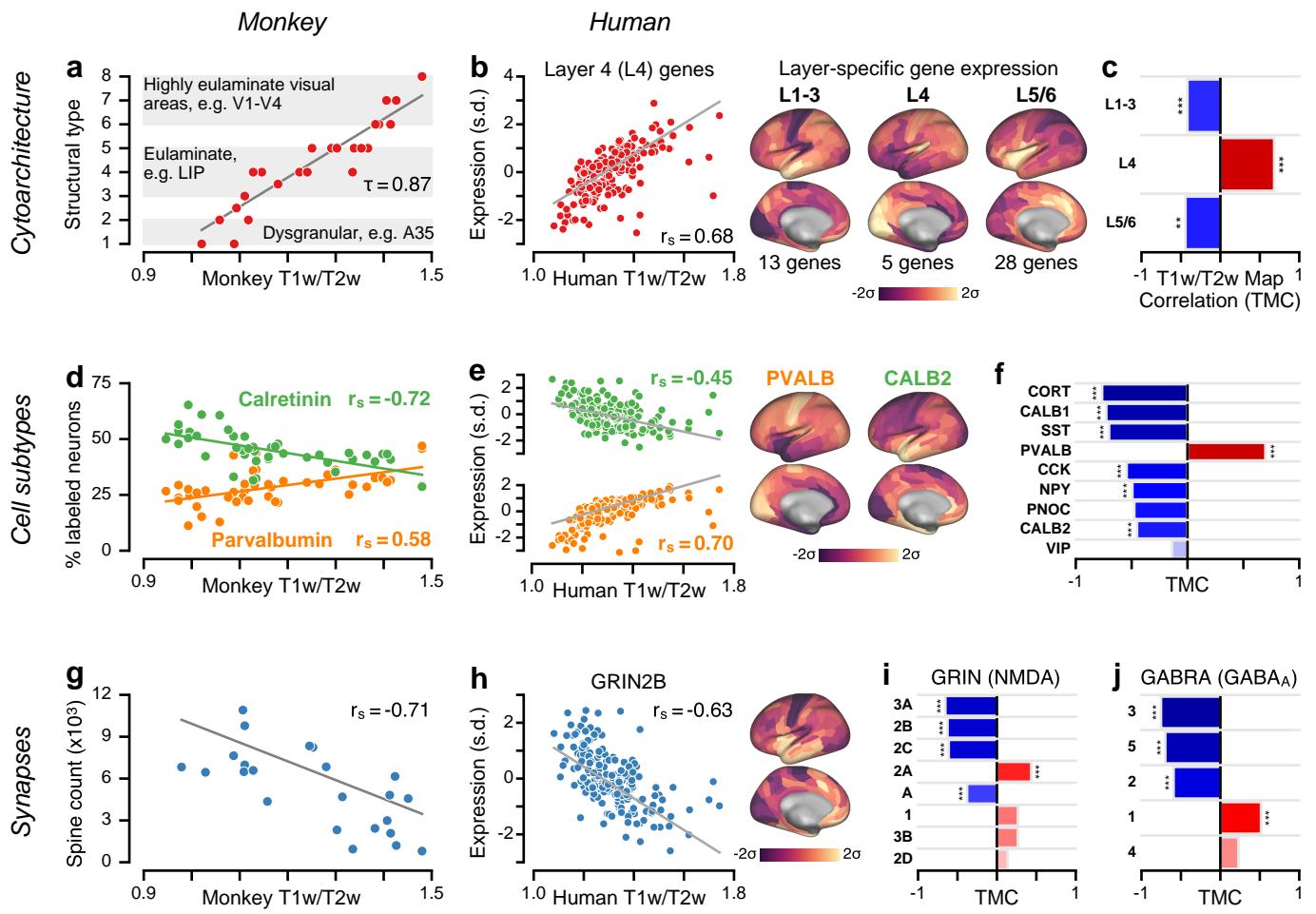


Figure 4: [Caption on next page]

Figure 4: Group-averaged T1w/T2w maps capture specialization of cortical microcircuitry in humans and nonhuman primates. **(a)** Cortical cytoarchitectural type is very strongly correlated with the macaque monkey T1w/T2w map ($\tau = 0.87, P < 10^{-5}$; Kendall's tau-b correlation) ($r_s = 0.96, P < 10^{-5}$; Spearman's rho). **(b)** The average expression map of 5 genes preferentially expressed in human granular layer 4 (L4) is positively correlated with the human cortical T1w/T2w map ($r_s = 0.68, P < 10^{-5}$; Spearman rank correlation), consistent with a more prominent granular L4 in sensory than association cortex. Expression is plotted in units of std. dev. (s.d.) from the mean. **(c)** Average expression maps of laminar-specific genes show significant T1w/T2w map correlations (TMCs). L1-3: supragranular layers 1-3; L5/6: infragranular layers 5 and 6. **(d)** The T1w/T2w map captures areal variation in the relative proportions of calretinin- ($r_s = -0.72, P < 10^{-5}$) and parvalbumin-expressing ($r_s = 0.58, P < 10^{-4}$) inhibitory interneurons across monkey cortex. **(e)** Genes coding for calretinin (*CALB2*; $r_s = -0.45, P < 10^{-5}$) and parvalbumin (*PVALB*; $r_s = 0.70, P < 10^{-5}$) exhibit homologous hierarchical gradients in human cortex. **(f)** TMCs of genes coding for markers of specific inhibitory interneuron cell types. **(g)** Basal-dendritic spine counts on pyramidal cells are significantly anti-correlated with the monkey T1w/T2w map ($r_s = -0.71, P < 10^{-4}$). **(h)** The gene coding for the NMDA receptor subunit NR2B (*GRIN2B*) exhibits a negative TMC ($r_s = -0.63, P < 10^{-5}$). **(i, j)** TMCs of genes coding for distinct subunits of the excitatory NMDA receptor and inhibitory GABA_A receptor. Statistical significance is calculated using a spatial autoregressive model to account for spatial autocorrelation (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$).

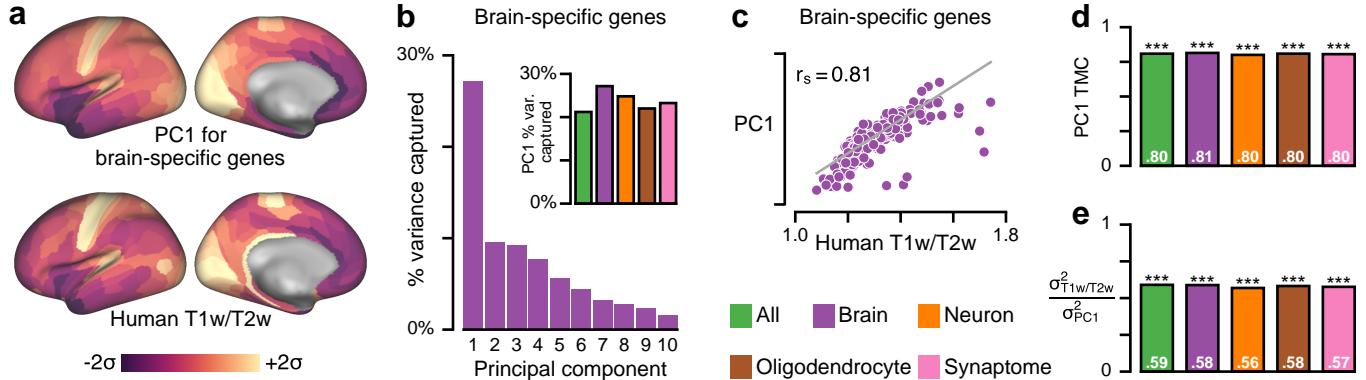


Figure 5: The group-averaged T1w/T2w map captures the dominant axis of gene expression variation across human cortex. **(a)** The first principal component (PC1), here for a set of brain-specific genes, is the areal map that linearly captures the maximum variation in gene expression. **(b)** PC1 captures a large fraction of total gene expression variance. *Inset:* Variance captured by PC1 for five gene sets: all genes, and genes preferentially expressed in brain, neurons, oligodendrocytes, and synaptic processes. **(c)** PC1 for this gene set is highly correlated with the T1w/T2w map ($r_s = 0.81$; $P < 10^{-4}$). **(d)** Across all sets, PC1 exhibits a highly similar areal topography to the T1w/T2w map (TMC range: 0.80–0.81; $P < 10^{-5}$ for each). **(e)** Gene expression variance captured by the T1w/T2w map ($\sigma_{T1w/T2w}^2$) relative to PC1 (σ_{PC1}^2). Statistical significance is calculated through permutation testing with surrogate maps that preserve spatial autocorrelation structure (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$).

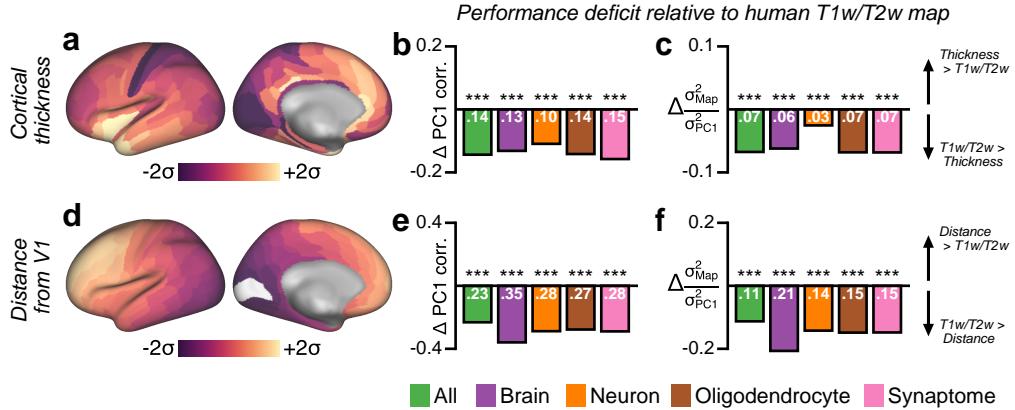


Figure 6: Principal component analysis (PCA) shows that the dominant mode of gene expression variation (PC1) is better captured by the group-averaged T1w/T2w map than by other candidate proxies. (a) Parcellated group-averaged ($N = 339$) map of human cortical thickness. **(b)** The difference in correlation with PC1 between the T1w/T2w map and the cortical thickness map, i.e., $(r_s(\text{T1w/T2w}, \text{PC1}) - r_s(\text{Thickness}, \text{PC1}))$, across several categorical gene sets. Negative values indicate that the T1w/T2w map is more strongly correlated with PC1 than is the thickness map. Statistical significance is calculated by a test of the difference between dependent correlations (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$). **(c)** The difference in the fraction of gene expression variance captured, relative to the variance captured by PC1, between the T1w/T2w map and the cortical thickness map, i.e., $(\sigma_{\text{T1w/T2w}}^2 - \sigma_{\text{Thickness}}^2) / \sigma_{\text{PC1}}^2$, across several categorical gene sets. Negative values indicate that the T1w/T2w map captures more gene expression variance than does the thickness map. Statistical significance is calculated through permutation testing with surrogate maps that preserve spatial autocorrelation structure (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$). **(d)** Parcellated map of geodesic distance from primary visual cortical area V1. **(e)** The difference in correlation with PC1 between the T1w/T2w map and the map of distance from area V1. **(f)** The difference in the fraction of gene expression variance captured, relative to the variance captured by PC1, between the T1w/T2w map and the map of distance from V1.

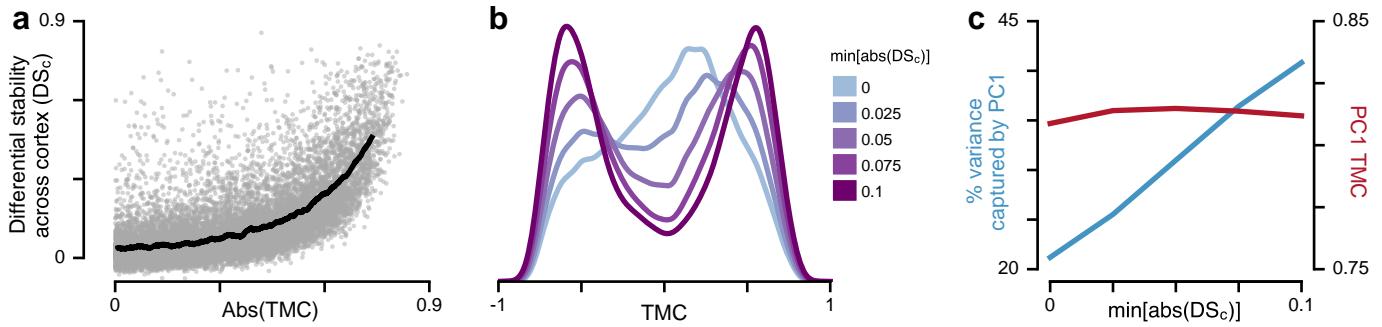


Figure 7: Expression profiles of genes which exhibit strong hierarchical gradients tend to be relatively stable across individuals. **(a)** **Differential stability across cortex (DS_c)**, defined as the mean pairwise Spearman rank correlation between subjects' cortical gene expression maps, as a function of the magnitude of the T1w/T2w map correlation (TMC) ($r_s = 0.66$, $P < 10^{-5}$; Spearman rank correlation). Each gray dot represents a single gene. The black line indicates the average value in a sliding window of size 600 points. **(b)** Filtering genes by a threshold on DS_c alters the shape of the TMC distribution. Increasing the DS_c threshold filters out genes whose cortical expression profiles are not relatively consistent across subjects. The trough which develops near $TMC=0$ suggests that high- DS_c genes preferentially exhibit strong hierarchical gradients. **(c)** Thresholding genes by DS_c substantially increases variance captured by the first principal component (PC1) of all gene expression variation (blue), whereas it has little effect on PC1's TMC (red).

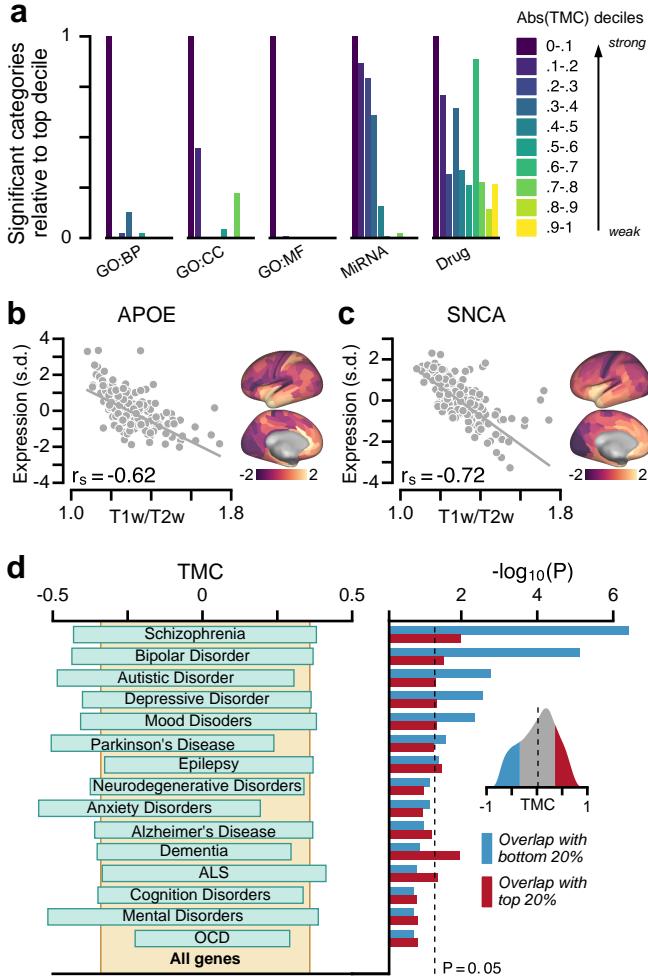


Figure 8: Hierarchical variation relates to enrichment in neurobiological function and brain disorders. **(a)** Genes with strong TMCs are overrepresented in functional annotations across multiple gene ontologies (GOs). BP, biological process; CC, cellular component; MF, molecular function; MiRNA, microRNA binding sites; Drug, drug targets. **(b, c)** Two key risk genes for neurodegenerative disorders, *APOE* for Alzheimer’s disease and *SNCA* for Parkinson’s disease, exhibit strongly negative TMCs, with higher expression levels in association cortex relative to sensory cortex (*APOE*: TMC = -0.62 , $P < 10^{-15}$; *SNCA*: TMC = -0.72 , $P < 10^{-25}$). *APOE* is a leading risk gene for Alzheimer’s disease. The $\epsilon 4$ allele of *APOE* is the largest genetic risk factor for late-onset Alzheimer’s disease. *SNCA* (*PARK1/PARK4*) is a key risk gene for Parkinson’s disease. Duplication of *SNCA* is risk factor for familial Parkinson’s disease with dominant inheritance. *SNCA* codes for the alpha-synuclein protein which is the primary component of Lewy bodies, a biomarker of Parkinson’s disease. **(d)** Genes with strong negative TMCs are overrepresented in multiple gene sets associated with neuropsychiatric disorders. *Left panel:* 20–80% ranges of TMC for all genes. *Right panel:* Enrichment is quantified by the hypergeometric test, which assesses the statistical significance of overlap between each gene set and the top (red) or bottom (blue) 20% TMC genes. *Inset:* Distribution of TMCs across all genes.

Methods

Parcellated structural neuroimaging maps. The human T1w/T2w and cortical thickness maps in the surface-based CIFTI file format²¹ were obtained from the Human Connectome Project (HCP)⁴⁶. To produce the T1w/T2w maps, high resolution T1- and T2-weighted images were first registered to a standard reference space using a state-of-the-art areal-feature-based technique^{21,47}, which precluded the need for spatial smoothing, and then corrected for bias-field intensity inhomogeneities, yielding dimensionless quantities defined with respect to a reference group-averaged map (for more details, see [17], [48], and [21]). Group-averaged ($N = 339$) left-hemispheric T1w/T2w and thickness maps were parcellated into 180 areas using the HCP’s Multi-Modal Parcellation (MMP1.0)²¹. Assignment of MMP1.0 parcels to functional networks (Fig. 1b, Extended Data Fig. 1d) was performed through community detection analysis [49] on time-series correlations in the HCP resting-state fMRI dataset.

For validation, key findings reported for human cortex were replicated using group-averaged ($N = 69$) T1w/T2w maps from the publicly available Conte69 dataset¹⁷ (Extended Data Fig. 10). In contrast to the HCP maps, both individual and group-averaged T1w/T2w maps in the Conte69 dataset were smoothed using Gaussian filters weighted by geodesic distance to reduce high frequency spatial artifacts¹⁷.

The group-averaged T1w/T2w and thickness maps for macaque monkey cortex were obtained from the publicly available BALSA database⁵⁰ ($N = 19$) (<https://balsa.wustl.edu/study/show/W336>) and were produced by adapting the HCP pre-processing pipelines to work with monkey MRI data (see [50] for more details). Monkey T1w/T2w map values for the left cortical hemisphere were parcellated into 91 areas using the M132 parcellation, which was used for the anatomical tract-tracing dataset⁵⁰.

To construct maps of geodesic distance from primary visual area V1 in human and monkey cortex, pairwise geodesic distance between two parcels i and j was calculated as the average of all pairwise surface-based distances between grayordinate vertices in parcel i and vertices in parcel j .

Anatomical hierarchy levels in monkey cortex. To assess whether macaque cortical T1w/T2w maps could reliably capture the laminar-specific interareal projection patterns conventionally used to define anatomical hierarchy, we fit a generalized linear model (GLM) to quantitative laminar projection data, yielding ordinal hierarchy values in 89 cortical areas, following the procedure of ref. [15]. Anatomical tract-tracing data, derived from retrograde tracers, was obtained from the publicly available Core-Nets database (<http://core-nets.org>). Retrograde tracer was injected into a target area i , and the number of tracer-labeled neurons in source area j were counted. The fraction of external labeled neurons, $FLNe_{ij}$, provides a quantitative measure of connection strength defined

as the number of labeled neurons in the source area normalized by the total number of labeled neurons in all external cortical source areas for a given injection⁵¹. Labeled neurons in source areas are classified by their location in either supragranular or infragranular layers. For a given projection, the proportion of supragranular labeled neurons, SLN_{ij} , is defined as the ratio of N_{supra} to $N_{\text{supra}} + N_{\text{infra}}$ for neurons labeled in source area j . As feedforward and feedback connections preferentially originate in supragranular and infragranular layers, respectively^{13–15}, SLN is a quantitative measure of hierarchical distance between two cortical areas¹⁵: within this paradigm for laminar-specific projection motifs, a pure feedforward connection from source area j to target area i would originate entirely in the superficial layers, resulting in an SLN of 1. Conversely, a pure feedback projection originating entirely in deep infragranular layers would result in an SLN of 0.

The GLM procedure for estimating hierarchy levels from SLN data is described in detail in ref. [15]. In brief, the hypothesis that SLN is a measure of hierarchical distance can be expressed as $g(SLN_{ij}) = H_i - H_j$, where H_i corresponds to the hierarchical level of area i , and g is an arbitrary and possibly nonlinear function linking SLN values on the unit interval (0, 1) to their corresponding hierarchical distance. We used a logit link function to map SLN values from the unit interval to the entire real number line following the procedure of ref. [19]. Fitting linear predictors (i.e. hierarchical levels) to logit-transformed SLN values constitutes a type of generalized linear model, with maximum likelihood estimation assuming a binomial family probability distribution for the supra-

and infra-granular neuron counts. To assign more weight to stronger connections during model estimation of hierarchical levels, we also weight each pathway in the model by the negative logarithm of the FLNe value. We clip SLN values to lie in the interval (0.01, 0.99) so the logit-transformed SLN value is well-defined for all pathways used to fit the model. Furthermore, to reduce the impact of noise on model parameter estimation, we only included pathways which contained at least 100 projection neurons when fitting the GLM; we confirmed that results were generally robust to the choice of neuron count threshold.

Maximum likelihood estimation of model parameters was done in the R programming language using the `glm` function. The model-estimated hierarchy levels, invariant under linear transformations, were shifted and rescaled to span the unit interval [0, 1]. To assess the statistical relationship between T1w/T2w map value and hierarchy level, we calculated the Spearman rank correlation between the 89 ordinal hierarchy values and their corresponding parcellated T1w/T2w map values (Fig. 1f). For visual clarity in Fig. 1e,f we remove the nonlinear logit transformation by displaying model-estimated hierarchy levels after applying the inverse-logit (i.e., logistic) transformation. This rescaling preserves the ordering of areas and therefore does not affect the reported Spearman rank correlations.

Macaque monkey anatomical data: cytoarchitectural types, inhibitory interneuron densities, and pyramidal neuron spine counts. To quantify the statistical relationship be-

tween T1w/T2w map value and categorical cytoarchitectural type (Fig. 4a), we compared T1w/T2w map values to structural classification values reported for 29 regions of primate visual cortex, obtained from ref. [22]. To characterize hierarchical distributions of cortical inhibitory interneuron cell types (Fig. 4d), we compiled, from multiple immunohistochemical studies, the relative densities of inhibitory interneurons which are immunoreactive (ir) to the calcium-binding proteins parvalbumin (PV) and calretinin (CR)^{52–55}. To characterize hierarchical variation in pyramidal neuron excitatory synaptic connectivity (Fig. 4g), we compiled, from multiple studies by Elston and colleagues^{56–61}, the number of spines of basal-dendritic trees of layer-3 pyramidal neurons.

For each of these three analyses, we produced a mapping between the 91 areas in the M132 atlas parcellation, which was used to calculate parcellated T1w/T2w map values in monkey cortex, to the architectonic areas reported in these collated studies (Supplementary Table 1). Where the anatomical mapping was not a one-to-one correspondence, we mapped the reported architectonic area onto the set of all M132 parcels with nonzero spatial overlap, and the corresponding T1w/T2w map value was calculated as the average across these M132 parcels.

Gene expression preprocessing. The Allen Human Brain Atlas (AHBA) is a publicly available transcriptional atlas containing gene expression data, measured with DNA microarrays, and sampled from hundreds of histologically validated neuroanatomical struc-

tures across six normal post-mortem human brains⁶. After no significant interhemispheric transcriptional differences were observed in the first two bilaterally profiled brains⁶, the remaining four donor brains were profiled only in the left cortical hemisphere⁷. To construct parcellated group-averaged gene expression profiles, we therefore restricted all analyses to microarray data sampled from the left cortical hemisphere in each of the six brains. Microarray expression data and all accompanying metadata were downloaded from the AHBA (<http://human.brain-map.org>)^{6,7}. The raw microarray expression data for each of the six donors includes expression levels of 20,737 genes, profiled by 58,692 microarray probes. These data were preprocessed according to the following procedure:

1. Gene probes without a valid Entrez Gene ID were excluded.
2. Microarray samples exhibiting exceptionally low inter-areal similarity were excluded.

We first computed the spatial correlation matrix of expression values between samples using the remaining 48,170 probes, then summed this matrix across all samples. Samples whose similarity measure was more than five standard deviations below the mean across all samples were excluded. At most, this step excluded three samples within a subject.

3. Samples whose annotations did not indicate that they originated in the left hemisphere of the cerebral cortex were excluded. To focus analysis on neocortex, we also excluded

samples taken from cortical structures that are cytoarchitecturally similar to the hippocampus, including the rhinal sulcus, piriform cortex, parahippocampal gyrus, and the hippocampal formation.

4. Samples whose measured expression level was not well above background, as provided in the AHBA dataset, were excluded⁷. Samples surviving this step i) belonged to a probe whose mean signal was significantly different from the corresponding background, and ii) had a background-subtracted signal which was at minimum 2.6 times greater than the standard deviation of the background.
5. The remaining cortical samples were mapped from volumetric space to the two dimensional cortical surface by minimizing the pairwise Euclidean distance between stereotaxic MNI coordinates reported for each cortical sample, and coordinates of grayordinate vertices in each subject's native cortical surface mesh (which was constructed using the procedure described in the following section below). Samples whose Euclidean distance to the nearest surface vertex was more than two standard deviations above the mean distance computed across all samples were excluded (excluding between 2 and 15 samples per subject). An average of 203 ± 27 samples per subject, yielding 1220 total samples across all six subjects, remained at this stage.
6. Expression levels for samples mapped onto the same surface vertex were averaged. Then expression levels within each remaining sample were z-scored across all gene

probes.

7. Using cortical samples mapped onto subjects' native surface meshes, expression profiles for each of the 180 unilateral parcels in the HCP's MMP1.0 cortical parcellation²¹ were computed in one of the two following ways. (I) For parcels which had at least one sample mapped directly onto one of their constituent surface vertices, parcellated expression values were computed by averaging expression levels across all samples mapped directly onto the parcel. (II) For parcels which had no samples mapped onto any of their constituent vertices, we first created densely interpolated expression maps, in which each vertex in the native surface mesh was assigned the expression level associated with the most proximal surface vertex onto which a sample had been directly mapped, determined using surface-based geodesic distance along each subject's cortical surface mesh (i.e., a Voronoi diagram approach); the average of expression levels across parcels' constituent vertices was then computed to obtain parcellated expression values, effectively equivalent to performing a weighted average.
8. A coverage score was also assigned to each gene probe, defined as the fraction of 180 parcels that had at least one sample mapped directly onto one of its constituent surface vertices. Probes with coverage below 0.4 (i.e., probes for which fewer than 72 of the 180 parcels contained samples) were excluded from further analysis.
9. For each gene profiled by multiple gene probes, we selected and used the expression

profile of a single representative probe. If two probes were available, we selected the probe with maximum gene expression variance across sampled cortical structures, in order to more reliably capture spatial patterns of areal heterogeneity. If three or more probes were available, we computed a correlation matrix of parcellated gene expression values across the available gene probes, summed the resultant matrix along one of its dimensions to obtain a quantitative similarity measure for each probe, relative to the other gene probes, and selected the probe with the highest similarity measure, as it is most highly representative among all available gene probes.

10. Each subject-level gene expression profile was z-scored before we computed group-level expression profiles, which were obtained by computing the mean across subjects which were assigned a probe for that gene. Group-level gene expression profiles were not computed if fewer than four subjects had an available gene probe. Finally, group-level expression profiles were z-scored across all 180 areas for each gene.

These steps yielded group-averaged expression values for 16,088 genes across 180 cortical areas, which were used for all reported analyses. The T1w/T2w map correlation (TMC) for each gene is reported in Supplementary Table 2. We also replicated all reported findings after mapping subjects' gene expression samples to the HCP's group-averaged surface mesh instead of subjects' native surface meshes in step 5 above. However, we found that native surface-based expression sample mapping yielded slightly stronger TMCs and

improved spatial registration in general (not shown).

Native surface mesh construction. Single-subject surface registration for each of the six subjects in the AHBA was performed following a procedure adapted from the HCP’s minimal preprocessing pipelines [48]. Briefly, the T1w image was first rigidly aligned to the MNI coordinate axes to produce a native space volume, which was then nonlinearly registered to the standard MNI template using FSL’s FLIRT and FNIRT. The native space image was run through FreeSurfer’s *recon-all* pipeline, which performs automated segmentation of brain structures to reconstruct the white matter and pial surfaces. The FreeSurfer output surface was then converted to standard GIFTI format to produce each subject’s native surface mesh. Finally, subjects’ native surface meshes were registered to the standard HCP surface mesh.

Categorical gene sets. We conducted analyses on biologically and physiologically meaningful gene sets extracted from existing databases and neuroscientific literature, reported below (Supplementary Table 2):

1. **Brain-specific.** Genes with expression specific to human brain tissue, relative to other tissues, were obtained from supplementary data set 1 of ref. [62]. Following ref. [30], brain-specific genes were selected for which expression in brain tissue was four times higher than the median expression across all 27 different tissues.

2. Neuron- and oligodendrocyte-specific. Brain genes with expression specific to neurons or oligodendrocytes, relative to other central nervous system (CNS) cell types, were obtained from supplementary data set S3b of ref. [63]. Following ref. [30], neuron-specific genes were selected for which log-expression in neurons of P7n cell type in the mouse was 0.5 greater than the median log-expression across 11 CNS cell types.

3. Synaptome. We aggregated four sets of synaptic genes encoding proteins found in the presynaptic nerve terminal, presynaptic active zone, synaptic vesicles, and postsynaptic density, which were obtained from SynaptomeDB, an ontology-based database of genes in the human synaptome³¹.

4. Neuron subtype-specific. Gene sets representing distinct classes of neuronal subtypes were obtained from ref. [26], in which clustering and classification analyses yielded 16 distinct neuron subtypes, on the basis of differential gene expression measured by RNA sequencing from single neurons in human cortex. The fraction of positive values using exon-only derived transcripts per million (TPM) associated with each subtype-specific gene were obtained from supplementary table S5; within each neuronal subtype cluster, the TPM values for the cluster genes were normalized and used to create a weighted gene expression profile representative of each subtype's spatial topography (Extended Data Fig. 4).

5. Layer-specific. Sets of laminar-specific genes localized to different layers of human neocortex were obtained from supplementary table S2 of ref. [23]. Genes were broadly grouped into sets representative of supragranular (L1–3), granular (L4), and infragranular (L5/6) layers.

Spatial autoregressive modeling. Significance values indicated by the number of stars reported on bar plots for T1w/T2w map correlations (TMCs) were corrected to account for spatial autocorrelation structure in parcellated T1w/T2w maps and gene expression maps. Because physical quantities like microstructural tissue composition and gene expression must vary smoothly and continuously in space, measurements recorded from proximal cortical areas tend to be more similar than measurements recorded from distal areas of cortex. This departure from the assumption of independent observations biases calculations of statistical significance. To model this spatial autocorrelation, we used a spatial lag model (SLM) commonly applied in the spatial econometrics literature⁶⁴, of the form $y = \rho W y + X\beta + \nu$, where W is a user-defined weight matrix implicitly specifying the form of spatial structure in the data, and ν is normally distributed.

To implement a spatial lag model in the python programming language, we used the maximum likelihood estimation routine defined in the Python Spatial Analysis Library (*pysal*)⁶⁵. We first determined the surface-based spatial separation between each pair of cortical parcels by computing the mean of the pairwise geodesic distances between each

vertex in parcel i and each vertex in parcel j , from which we constructed a pairwise parcel distance matrix, D .

Similarity of gene expression profiles was well-approximated by an exponential decaying spatial autocorrelation function (Extended Data Fig. 8a,b), as was found in mouse cortex¹². We fit the correlation of gene expression profiles between two areas with the exponential function $\text{Corr}(x_i, x_j) \sim \exp(-D_{ij}/d_0)$, where x_i and x_j are vectors containing the parcellated gene expression values at parcels i and j , D_{ij} is the geodesic distance between the parcels, and d_0 is the characteristic spatial scale of autocorrelation. We empirically determined d_0 by first computing the pairwise gene co-expression matrix $C_{ij} \equiv \text{Corr}(x_i, x_j)$. We then fit the free parameter d_0 using ordinary least squares (OLS) regression on the off-diagonal (upper-triangular) elements of the gene co-expression and parcel distance matrices, so as to minimize the sum-of-squared-residuals between empirical and model-estimated gene co-expression values over all pairs of cortical parcels,

$$S = \sum_{i>j} r_{ij}^2 = \sum_{i>j} [C_{ij} - \exp(-D_{ij}/d_0)]^2.$$

This empirical fit was performed on the gene co-expression matrix computing using the set of brain-specific genes. Using the OLS estimate of the spatial autocorrelation scale from the fit to the empirical gene expression data, we calculated the elements of the spatial weight matrix, $W_{ij} = \exp(-D_{ij}/d_0)$. Finally, we fit the SLM to parcellated gene expression profiles, using the maximum likelihood estimator routine (`pysal.spreg.ml_lag.ML_Lag`) in *pysal*. P-values indicated by the number of stars in the bar plots of T1w/T2w map correlations (TMCs) correspond to p-values for

model parameter β defined above.

Of note, spatial autoregressive model parameters do not have the same interpretation as they do in OLS regression. The parameter β reflects the direct (i.e. local) impact on the dependent variable y due to a unit change in the independent variable x . In addition, because of the underlying spatial structure, the direct impact of x_i on y_i results in an indirect effect of y_i on neighboring y_j . Therefore β cannot be interpreted as a corrected, global correlation coefficient, and we restrict our use of the SLM to correcting for the biasing effect of spatially autocorrelated samples on reported significance values.

Theil-Sen estimator. Trend lines in figures are calculated using the Theil-Sen estimator, which is a nonparametric estimator of linear slope based on Kendall's tau rank correlation, that is insensitive to the underlying distribution and robust to statistical outliers⁶⁶. It is defined as the median of the set of slopes computed between all pairs of points.

Multiple comparisons corrections. Significance values indicated by the number of stars reported on bar plots for T1w/T2w map correlations (TMCs) were Bonferroni-corrected for multiple comparisons. Statistical significance thresholds (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$) were divided by the number of null hypotheses tested, i.e., by the number of constituent bars contained in each bar plot.

Principal components analysis. We used principal component analysis (PCA) to identify the dominant modes of spatial variation in the transcriptional profiles of gene expression in the human cortex. For a set of N genes, each with group-averaged expression values for P cortical parcels, we constructed a gene expression matrix \mathbf{G} with one row for each cortical parcel and one column for each unique gene (i.e. with dimensions $P \times N$). The $P \times P$ spatial covariance matrix \mathbf{C} was constructed by computing the covariance between vectors of gene expression values for each pair of cortical parcels: $C_{ij} = \text{Cov}(G_i, G_j)$, where G_i is the i -th row in the matrix \mathbf{G} , corresponding to the vector of N gene expression values for the i -th cortical parcel. Eigen-decomposition is performed on the spatial covariance matrix to obtain the matrix eigenvectors (i.e., the principal components, PCs) and their corresponding eigenvalues, which are proportional to the amount of variance captured by the corresponding PC. To enumerate each principal component, eigenvalues are ranked in descending order of absolute magnitude, with larger magnitudes indicating a greater proportion of the total variance captured by the associated PC (i.e., the associated mode of spatial covariation). PCA therefore allows for simultaneous identification of spatial patterns of covariation and quantification of the extent to which these spatial modes capture variance in cortical gene expression profiles.

To quantify the overlap of these spatial PCs with the cortical T1w/T2w map, we compute the Spearman rank correlation coefficient between each P -dimensional PC and the P -dimensional vector of T1w/T2w map values for each cortical parcel. We can quan-

tify the amount of gene expression variance that is captured by any given spatial map, such as the T1w/T2w map (Fig. 5e, Extended Data Fig. 6f-j): from the spatial covariance matrix \mathbf{C} , the variance captured along a unit-length vector \mathbf{a} , here a demeaned and normalized map, is given by $\mathbf{a}^\top \mathbf{C} \mathbf{a}$.

Surrogate data generation. To nonparametrically determine significance values in our PCA results, in Fig. 5 and Extended Data Fig. 6, we generated surrogate maps with a spatial autocorrelation structure matched to the empirical data (Extended Data Fig. 8c). Parameters characterizing the empirical spatial autocorrelation were determined numerically for the cortical T1w/T2w map, cortical thickness map, and the map of surface-based geodesic distance from area V1; in each case, we fit the data using a spatial lag model of the form $\mathbf{y} = \rho \mathbf{W} \mathbf{y}$, where \mathbf{y} is a vector of first Box-Cox transformed and then mean-subtracted map values. The Box-Cox transformation was first applied to the maps so their values were approximately normally distributed. \mathbf{W} is the row-normalized weight matrix with zero diagonal and off-diagonal elements proportional to $W_{ij} = z_i^{-1} \exp(-D_{ij}/d_0)$, where D_{ij} is the surface-based geodesic distance between cortical areas i and j , and $z_i \equiv \sum_j \exp(-D_{ij}/d_0)$ is a row-wise normalization factor. Weights W_{ij} define the fraction of spatial influence on area i attributable to area j . Two free parameters ρ and d_0 are estimated by minimizing the residual sum-of-squares⁶⁴. Using best-fit parameter values $\hat{\rho}$ and \hat{d}_0 , surrogate maps \mathbf{y}_{surr} are generated according to $\mathbf{y}_{\text{surr}} = (\mathbb{I} - \hat{\rho} \mathbf{W}[\hat{d}_0])^{-1} \mathbf{u}$,

where $\mathbf{u} \sim \mathcal{N}(0, 1)$. To match surrogate map values distributions to the distribution of values in the corresponding empirical map (e.g. the T1w/T2w map), rank-ordered surrogate map values were re-assigned the corresponding rank-ordered values in the empirical map. Note that this approach to surrogate data generation approximates a spatial autocorrelation-preserving permutation test of the empirical neuroimaging map.

Using these surrogate maps we construct null distributions for $N = 10,000$ statistics and report significance values as the proportion of samples in the null distributions whose absolute value is greater than or equal to the absolute value of the test statistic. To compute significance values reported in Fig. 6c,f, we first constructed null distributions of the statistic $\sigma_{\text{Map}}^2 / \sigma_{\text{PC1}}^2$ using surrogate maps constructed for each neuroimaging map. For each neuroimaging map, we then computed distributions of the difference between the test statistic and each sample statistic in the null distribution. Finally, we used the non-parametric Wilcoxon signed-rank test on these difference distributions, one for the T1w/T2w map and one for either the cortical thickness or geodesic distance map, were statistically different from zero. The interpretation of our statistically significant results reported in Fig. 6c,f is that the T1w/T2w map tends to capture a more appreciable fraction of gene expression variance, relative to its surrogate maps, than do either of the other two candidate neuroimaging maps.

Differential stability. Differential stability (DS) is a correlation-based metric which quantifies the consistency of spatial gene expression patterns across individual brains. DS was originally defined in ref. [7] as "the tendency for a gene to exhibit reproducible differential expression relationships across brain structures." To compute DS for a gene, we calculated the average pairwise Spearman rank correlation (r_s) across all subject-level gene expression profiles, for the (four to six) AHBA subjects with an available gene probe, for a maximum of 15 possible pairs (Supplementary Table 2). That is, for gene g whose expression profile across 180 cortical areas in brain i is the vector $b_i(g)$, we define the DS in cortex (DS_c) by: $DS_c(g) \equiv \frac{1}{15} \sum_{i=1}^6 \sum_{j>i} r_s(b_i(g), b_j(g))$. We note that DS is therefore defined with respect to a specified set of brain structures, in this case 180 unilateral cortical areas. We note that any differences between cortical DS values shown in our Fig. 7a and those shown in Fig. 7b of ref. [7] are due to i) different cortical parcellations (containing 180 vs. 52 parcels, respectively); ii) different pre-processing procedures; and iii) different correlation coefficients (Spearman vs. Pearson, respectively).

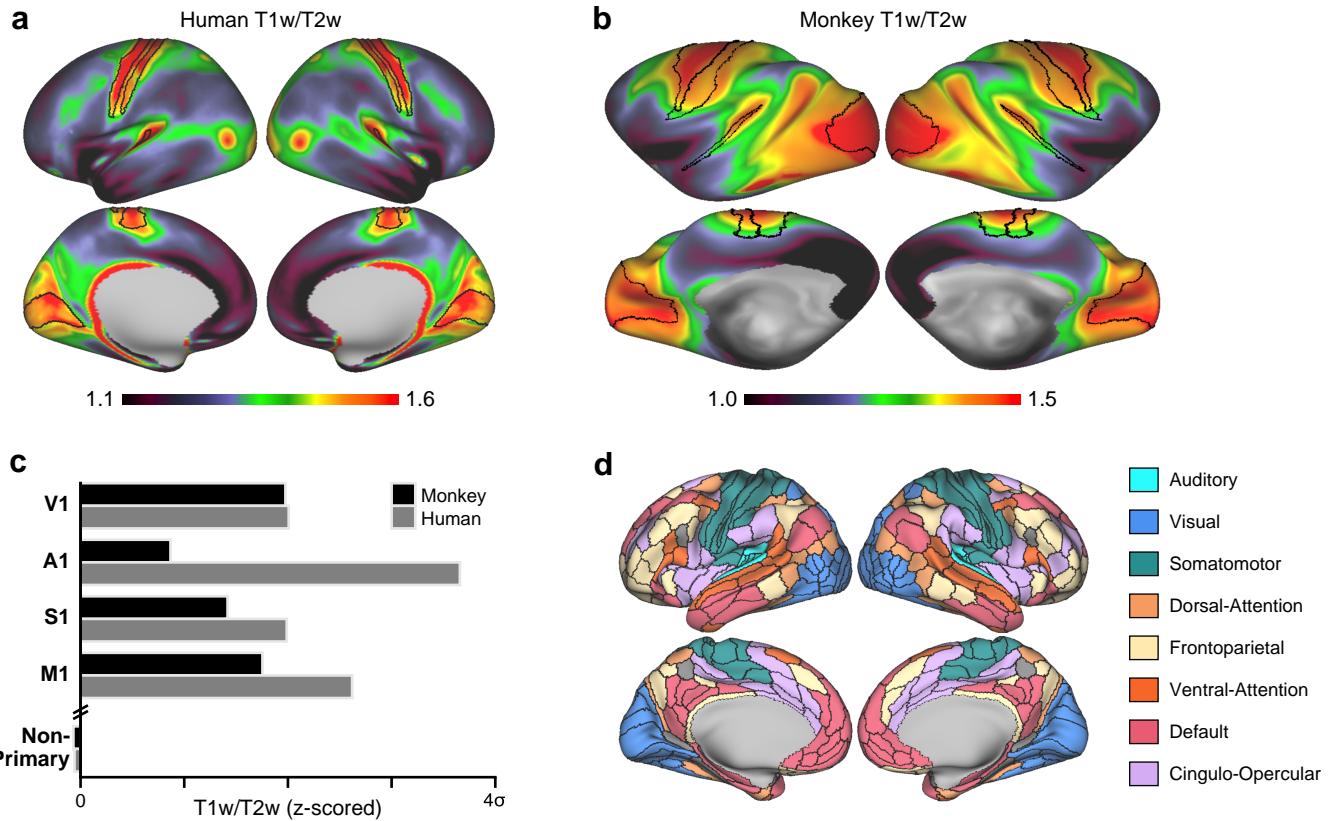
Functional enrichment analyses. Functional enrichments were determined using the ToppGene (<https://toppgene.cchmc.org/>) web portal³², including gene ontology annotations (biological process, cellular component, and molecular function); microRNA targets (from all sources indicated on <https://toppgene.cchmc.org/navigation/database.jsp>); and drug annotations (from DrugBank, Comparative Toxicogenomics Database, includ-

ing marker and therapeutic, and Broad Institute CMAP). Significant genes in each category were identified using the ToppFun utility. Disease annotations were determined using curated disease gene associations in the DisGeNet database³⁴ (<http://www.disgenet.org/web/DisGeNET/menu/home>). Hypergeometric testing was used to determine significant over-representation of brain-related disease genes in the top and bottom gene quintiles (20%, 3218 genes) ranked by T1w/T2w map correlation, following ref. [7].

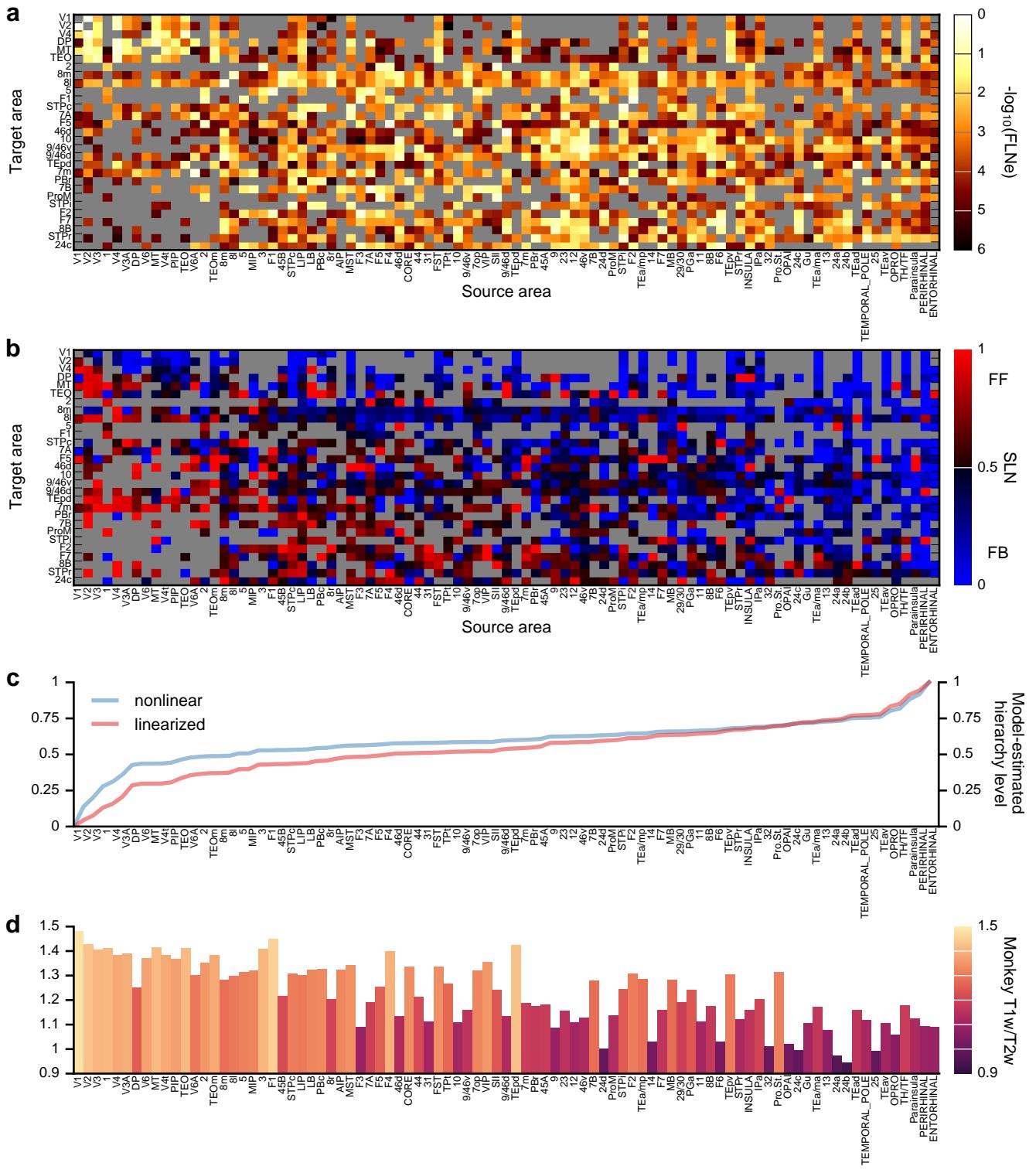
Code and data availability. Custom analysis codes written in Python, and data supporting the findings of this study, are available from the corresponding author upon reasonable request. All results derive from data that is publicly available from sources described above.

46. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80**, 62–79 (2013).
47. Robinson, E. C. *et al.* MSM: a new flexible framework for Multimodal Surface Matching. *Neuroimage* **100**, 414–426 (2014).
48. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–24 (2013).
49. Ito, T. *et al.* Cognitive task information is transferred between brain regions via resting-state network topology. *Nat Commun* **8**, 1027 (2017).
50. Donahue, C. J. *et al.* Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey. *J Neurosci* **36**, 6758–70 (2016).
51. Markov, N. T. *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex* (2012).
52. Condé, F., Lund, J. S., Jacobowitz, D. M., Baimbridge, K. G. & Lewis, D. A. Local circuit neurons immunoreactive for calretinin, calbindin D-28k or parvalbumin in monkey prefrontal cortex: distribution and morphology. *J Comp Neurol* **341**, 95–116 (1994).
53. Gabbott, P. L. & Bacon, S. J. Local circuit neurons in the medial prefrontal cortex (areas 24a,b,c, 25 and 32) in the monkey: II. Quantitative areal and laminar distributions. *J Comp Neurol* **364**, 609–36 (1996).
54. Kondo, H., Tanaka, K., Hashikawa, T. & Jones, E. G. Neurochemical gradients along monkey sensory cortical pathways: calbindin-immunoreactive pyramidal neurons in layers II and III. *Eur J Neurosci* **11**, 4197–203 (1999).
55. Dombrowski, S. M., Hilgetag, C. C. & Barbas, H. Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cereb Cortex* **11**, 975–88 (2001).
56. Elston, G. N. & Rosa, M. G. The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cereb Cortex* **7**, 432–52 (1997).
57. Elston, G. N. & Rosa, M. G. Morphological variation of layer III pyramidal neurones in the occipitotemporal pathway of the macaque monkey visual cortex. *Cereb Cortex* **8**, 278–94 (1998).
58. Elston, G. N., Tweedale, R & Rosa, M. G. Cortical integration in the visual system of the macaque monkey: large-scale morphological differences in the pyramidal neurons in the occipital, parietal and temporal lobes. *Proc Biol Sci* **266**, 1367–74 (1999).
59. Elston, G. N. & Rockland, K. S. The pyramidal cell of the sensorimotor cortex of the macaque monkey: phenotypic variation. *Cereb Cortex* **12**, 1071–8 (2002).

60. Elston, G. N., Benavides-Piccione, R. & Defelipe, J. A study of pyramidal cell structure in the cingulate cortex of the macaque monkey with comparative notes on inferotemporal and primary visual cortex. *Cereb Cortex* **15**, 64–73 (2005).
61. Elston, G. N., Oga, T., Okamoto, T. & Fujita, I. Spinogenesis and Pruning in the Anterior Ventral Inferotemporal Cortex of the Macaque Monkey: An Intracellular Injection Study of Layer III Pyramidal Cells. *Front Neuroanat* **5**, 42 (2011).
62. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* **13**, 397–406 (2014).
63. Cahoy, J. D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* **28**, 264–78 (2008).
64. Anselin, L. in *A companion to theoretical econometrics* (ed Baltagi, B. H.) 310–330 (Blackwell, Malden, Mass., 2001).
65. Fischer, M. M. & Getis, A. *Handbook of applied spatial analysis: software tools, methods and applications* (Springer, Berlin, 2010).
66. Sen, P. K. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63**, 1379–1389 (1968).

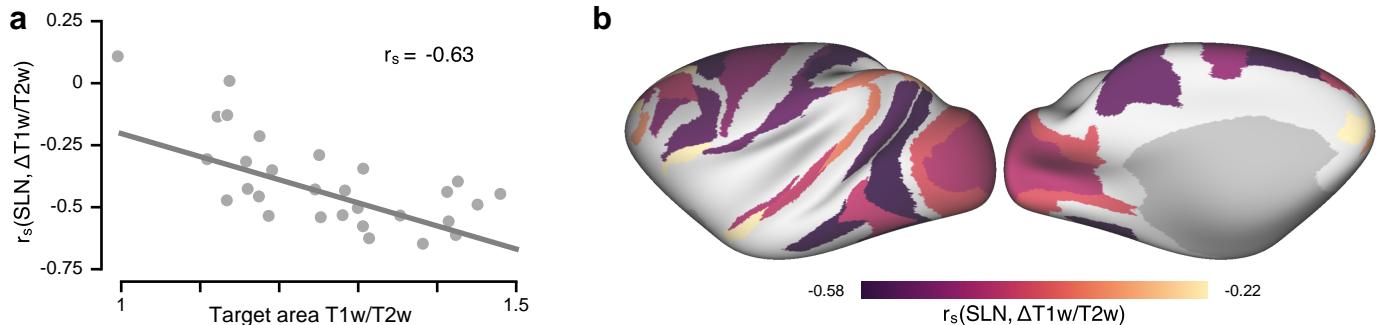


Extended Data Figure 1: Group-averaged cortical T1w/T2w maps exhibit inter-species homology and inter-hemispheric symmetry. **(a)** Unparcellated group-averaged T1w/T2w map in human cortex visualized bilaterally on an inflated cortical surface. **(b)** Unparcellated group-averaged T1w/T2w map in monkey cortex visualized bilaterally on an inflated cortical surface. **(c)** Primary sensory (visual, V1; auditory, A1; somatosensory, S1) and primary motor (M1) homologues in human and monkey cortex, corresponding to the four bordered areas in panels (a) and (b), exhibit high T1w/T2w map values relative to the average value computed across all higher-order sensory and association (Non-Primary) areas. The low T1w/T2w map value found in monkey area A1 (relative to human area A1) is likely driven by the size and lack of spatial specificity of the corresponding M132 parcel which extends deep down into temporal cortex. Note that human Brodmann area 3 is subdivided into areas 3a and 3b in the HCP parcellation, the latter which was used for this analysis. **(d)** Functional networks derived from resting-state functional connectivity from the Human Connectome Project (HCP). Cortical areas are parcellated using the HCP multi-modal parcellation (MMP1.0). We assigned each region to a functional network using a community detection method applied to resting-state fMRI data from the HCP, and designated functional labels to networks, including three sensory and five association, that align with previously reported functional networks (with abbreviations labeled in Fig. 1b): Auditory (AUD), Visual (VIS), Somatomotor (SOM), Dorsal Attention (DAN), Frontoparietal (FPN), Ventral Attention (VAN), Default (DMN), and Cingulo-Opercular (CON).

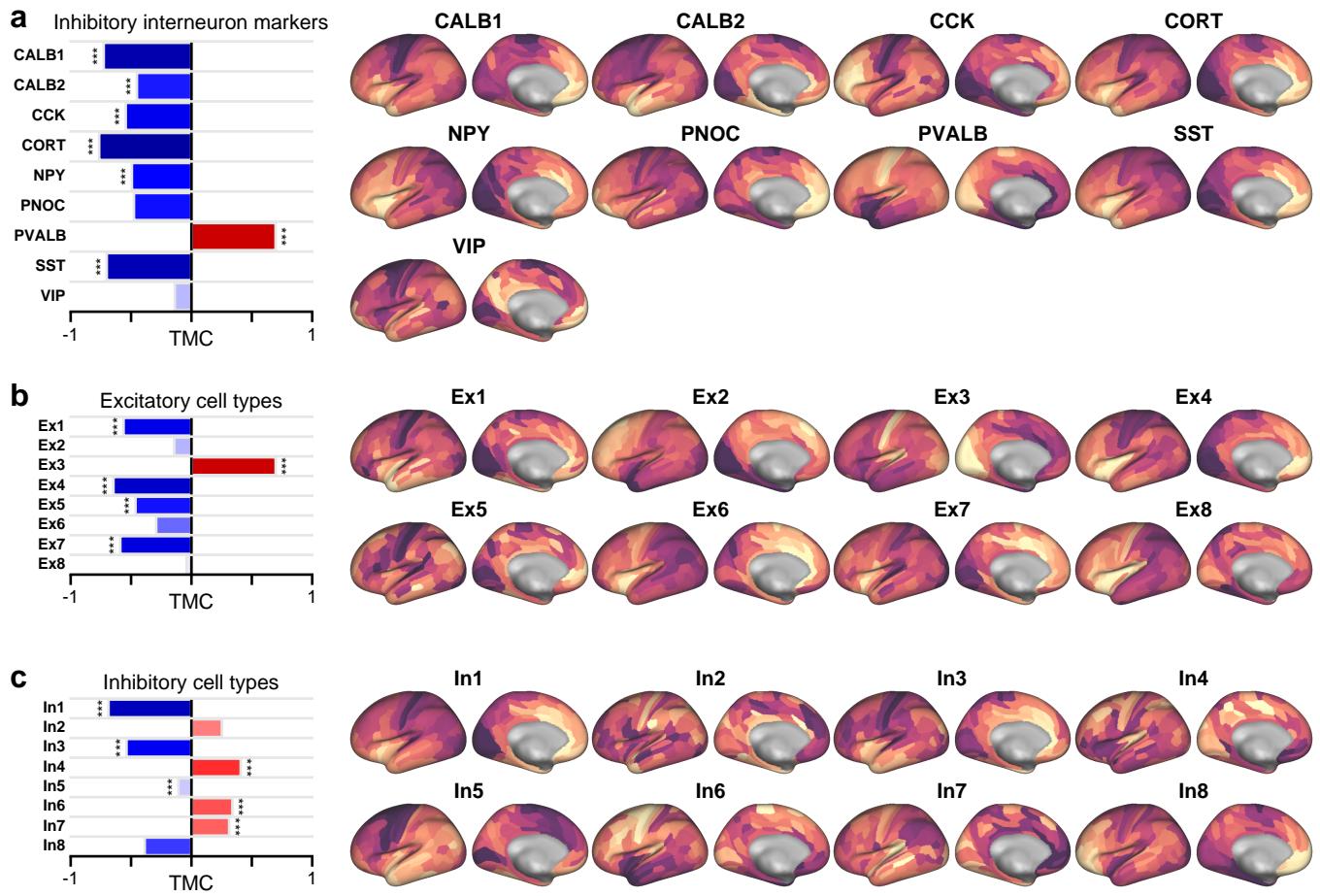


Extended Data Figure 2: [Caption on next page]

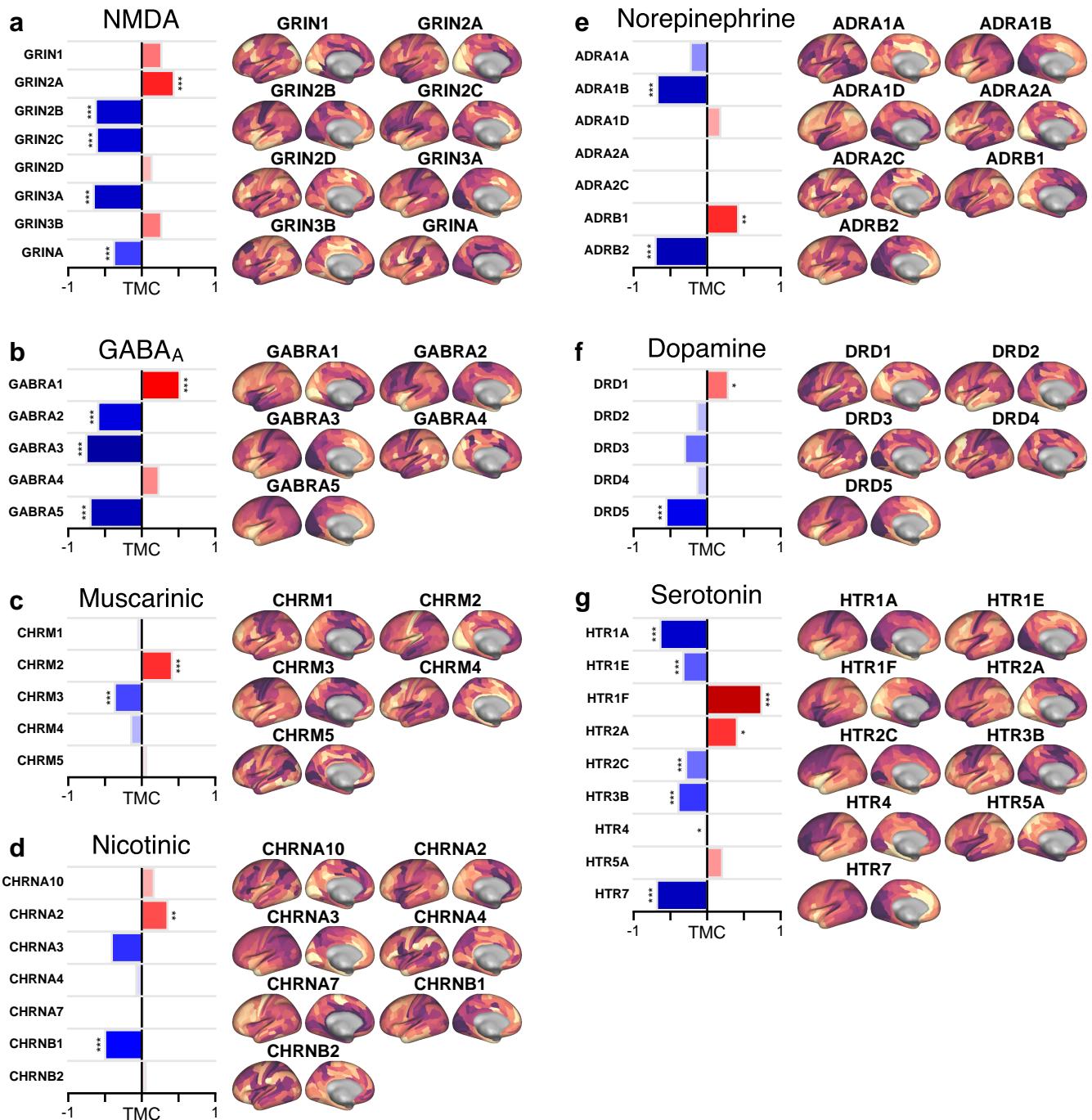
Extended Data Figure 2: Anatomical cortical hierarchy derived from laminar-specific interareal projections in monkey cortex. **(a)** Fraction of external labeled neurons ($FLNe$). Target area i is injected with a retrograde tracer that labels neurons in many source areas; the $FLNe$ in source area j is then defined as the fraction of all external labeled neurons terminating in area i that originated in source area j . Each row of the FLN matrix is therefore normalized to 1. Measurements which yielded no labeled neurons are marked in grey. **(b)** Fraction of supragranular layer neurons (SLN), defined as the fraction of neurons in an interareal projection (to target area i from source area j) originating in supragranular layers. An SLN of 1 indicates that all labeled projection neurons were of supragranular origin, reflecting a pure feedforward connection; an SLN of 0 indicates that all projection neurons originated in deep infragranular layers, reflecting a pure feedback connection. Measurements which yielded no labeled neurons are marked in grey. **(c)** Model-estimated hierarchy levels for 89 cortical regions. The blue line indicates hierarchy levels estimated by the model after shifting and re-scaling them to lie on the unit interval. The red indicates hierarchy values passed through a logistic function to remove the nonlinearity introduced by the logit link function in the GLM fitting procedure. The monotonicity of this transformation preserves the order of the cortical regions and therefore does not affect the Spearman rank correlations reported in the main text. **(d)** Group-averaged T1w/T2w map values for 89 cortical areas in the monkey.



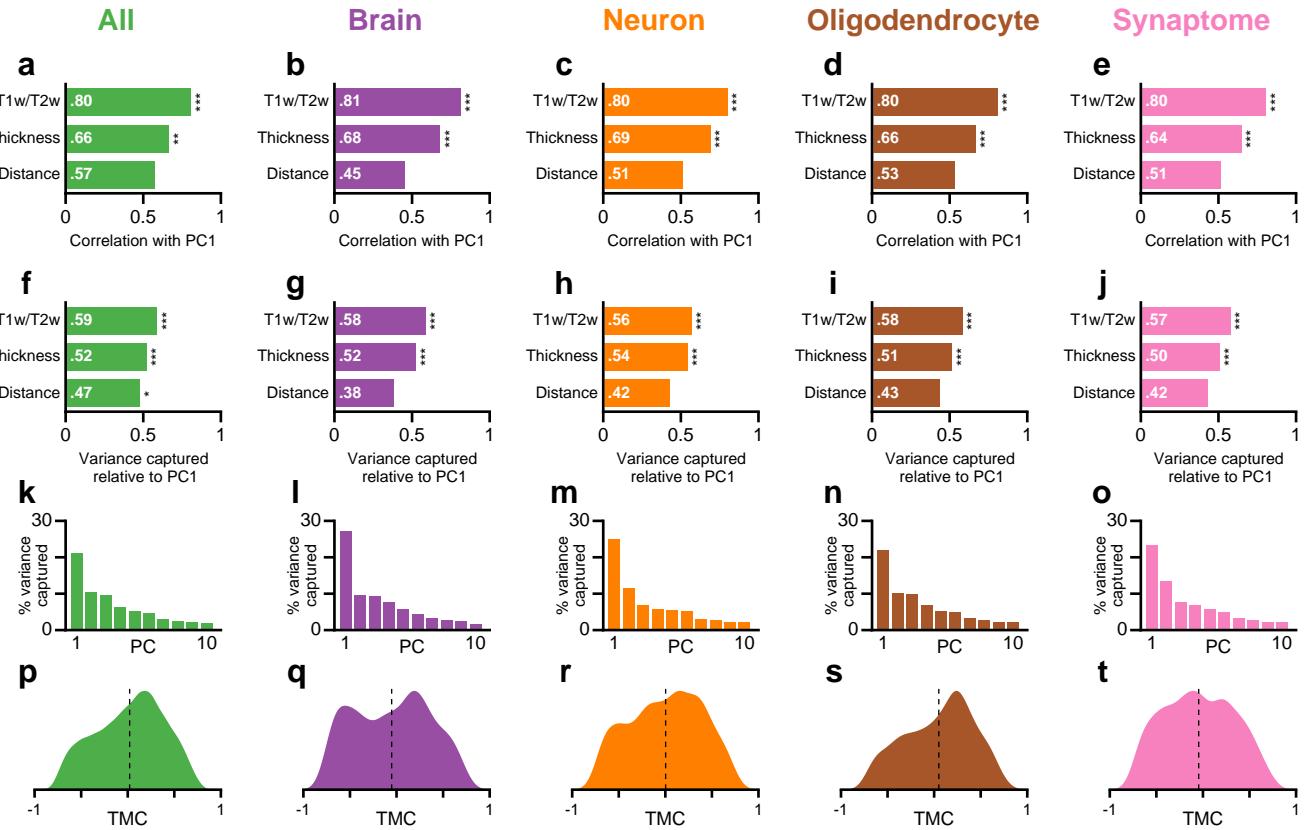
Extended Data Figure 3: The strength of the correspondence between laminar projection profiles and T1w/T2w value is strongest in high-T1w/T2w areas. **(a)** Retrograde tracer-labelled projections are separated by target area (i.e., by $N = 29$ unique injection sites). The laminar profile of a projection is quantified by the fraction of labeled supragranular layer neurons (SLN) in the source area. We found the Spearman rank correlation between SLN and the areal difference in T1w/T2w map values ($\Delta T1w/T2w$; target minus source) is stronger for source areas with higher T1w/T2w values ($r_s = -0.63$; $P < 10^{-3}$; Spearman rank correlation). This indicates that the relationship between laminar specificity and structural dissimilarity holds more strongly in sensory areas at lower levels of the anatomical hierarchy, which have high T1w/T2w values. **(b)** Hierarchically organized structure in the strength of the relationship between laminar specificity (SLN) and structural dissimilarity ($\Delta T1w/T2w$) for the 29 unique target areas, visualized on the inflated monkey cortical surface.



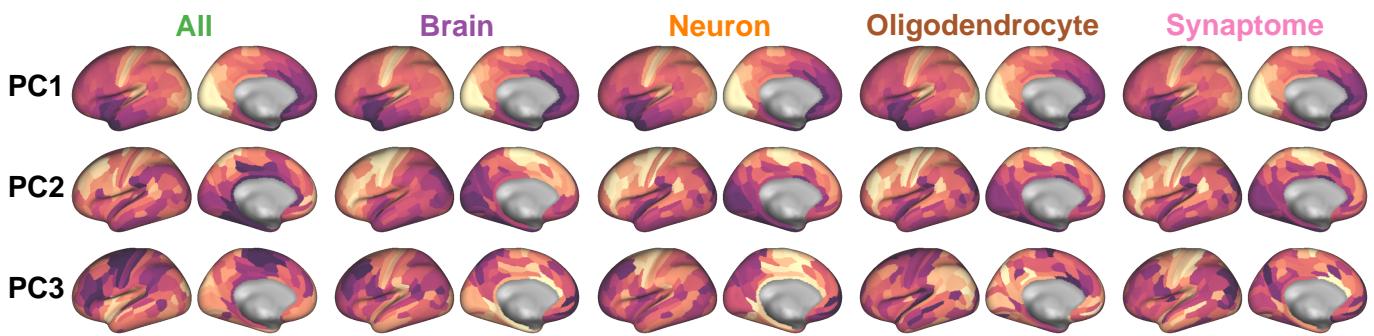
Extended Data Figure 4: Expression maps and T1w/T2w map correlations (TMCs) for genes that code for markers of distinct inhibitory interneuron cell types, and for weighted profiles characteristic of distinct neuronal cell types derived from single-cell RNA sequencing of human cortical neurons. **(a)** Markers for inhibitory interneuron cell types. **(b)** Weighted gene sets for excitatory neuronal cell types, derived from single-cell RNA sequencing. **(c)** Weighted gene sets for inhibitory neuronal cell types, derived from single-cell RNA sequencing.



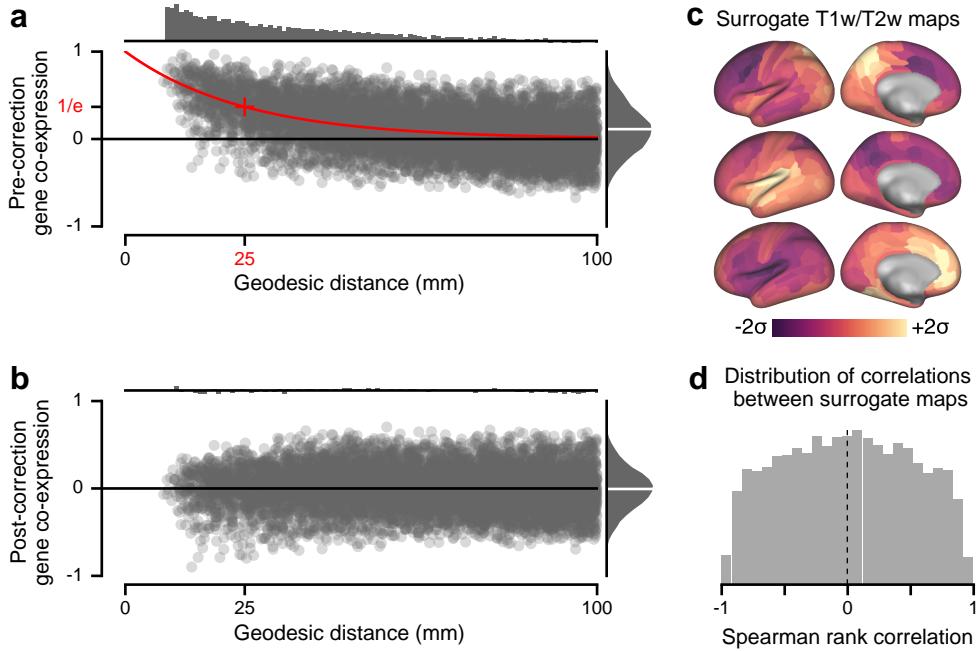
Extended Data Figure 5: Expression maps and T1w/T2w map correlations (TMCs) for genes coding for synaptic receptor subunits and neuromodulator receptor subunits. **(a)** NMDA receptor subunits. **(b)** GABA_A receptor subunits. **(c)** Muscarinic acetylcholine receptors (CHRM). **(d)** Nicotinic acetylcholine receptors (CHRN). **(e)** Norepinephrine receptors (ADR). **(f)** Dopamine receptors (DRD). **(g)** Serotonin receptors (HTR).



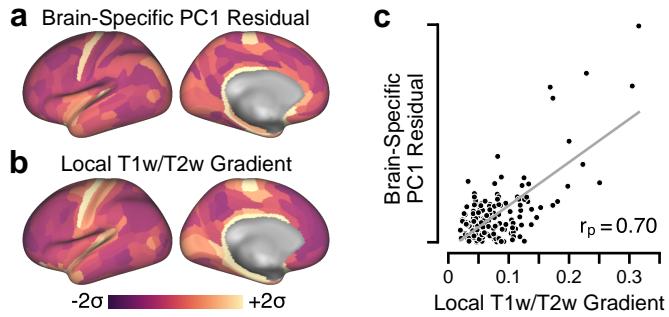
Extended Data Figure 6: The dominant axis of gene expression variation is better captured by the group-averaged T1w/T2w map than by two alternative candidate proxies: cortical thickness and distance from primary visual cortex (V1). **(a–e)** For five gene sets, the Spearman rank correlation between the first principal component (PC1) and the T1w/T2w map, the map of cortical thickness, and the map of geodesic distance from primary visual cortex. For each gene set, PC1 is more strongly correlated with the T1w/T2w map than with the two other candidate maps. **(f–j)** For five gene sets, the amount of gene expression variance captured, relative to PC1, for the three candidate maps. For each gene set, the T1w/T2w map captures more gene expression variance than do the other two maps. **(k–o)** Percentage of gene expression variance captured by the top 10 PCs, out of 179 total PCs (due to 180 cortical areas in our parcellation). For all five gene sets, PC1 captures between 21% and 27% of the variance, roughly two to three times the amount captured by PC2. **(p–t)** Distribution of T1w/T2w map correlations (TMCs) across genes for the five gene sets. Dashed lines mark the mean of the distribution. For all five gene sets, the distributions are broad, containing large fractions of strong positive and negative TMCs, and centered near zero, with a range of means ($-0.05, +0.05$).



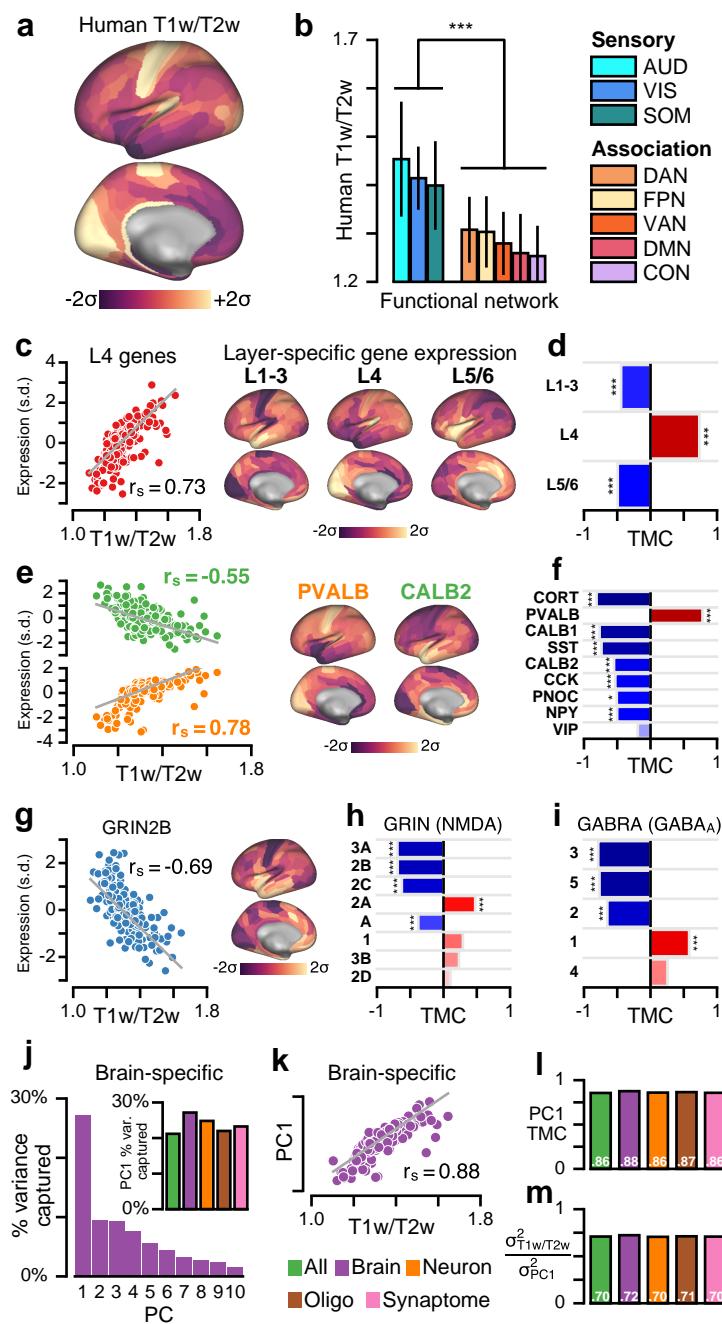
Extended Data Figure 7: Principal component analysis (PCA) reveals that the dominant axis of gene expression variation (i.e., the first principal component, PC1) is conserved across categorical gene sets. PC1, which aligns with the T1w/T2w map, combines sensory areas (e.g., primary visual, somatosensory, and auditory cortical areas), and separates sensory areas from association areas. In contrast, the secondary mode of gene expression tends to fractionate sensorimotor cortical areas by modality, separating early visual cortex from somatomotor cortex. Rows correspond to the first three PCs, respectively, across gene sets.



Extended Data Figure 8: Autocorrelation structure in gene expression and group-averaged T1w/T2w maps. **(a)** Spatial autocorrelation structure in the parcellated cortical gene expression data is well-approximated by a decaying exponential. Gene co-expression is defined as the pairwise Spearman rank correlation between cortical parcels' gene expression values, here for the brain-specific gene set. Proximal cortical parcels exhibit more similar gene expression values compared to distal parcels. All pairs of parcels with geodesic distance less than 100 mm were used to fit the characteristic scale of spatial autocorrelation, illustrated in red (i.e., $\exp(-d/d_0)$), where d is geodesic distance and $d_0 = 25$ mm. Each data point corresponds to the co-expression of a pair of cortical parcels. *Top:* Mean co-expression value as a function of geodesic distance bin. **(b)** Gene co-expression values after correcting for spatial autocorrelation structure by subtraction of the fitted exponential decay. After correction, the mean co-expression value is near zero across all geodesic distance bins. **(c)** Example randomized surrogate maps with spatial autocorrelation structure matched to the empirical T1w/T2w map (see Methods). Autocorrelation structure-preserving surrogate T1w/T2w maps are used for nonparametric calculation of statistical significance for PCA results in Extended Data Fig. 6, Figs. 5 and 6. **(d)** Distribution of pairwise Spearman rank correlations between pairs of surrogate T1w/T2w maps.



Extended Data Figure 9: Evidence for spatial structure in residual gene expression variance not captured by the group-averaged T1w/T2w map. **(a)** Map of the brain-specific PC1 residual, defined as the magnitude of the vertical distance between an area's brain-specific PC1 weight and the best-fit Theil-Sen estimator of linear slope for the PC1 vs. T1w/T2w relationship (i.e., absolute vertical distance between scatter points and best-fit line in Fig. 5c). **(b)** Map of the local T1w/T2w gradient, defined as the mean absolute difference between an area's T1w/T2w map value and the values of all neighboring areas, i.e. $x_i = \frac{1}{N} \sum_{n \in \{N_i\}} |T_i - T_n|$, where T_i is the T1w/T2w value for area i , and $\{N_i\}$ is the set of all areas bordering area i . **(c)** The brain-specific PC1 residual is highly correlated with the magnitude of the local T1w/T2w gradient ($r_p = 0.70$; $P < 10^{-5}$; Pearson correlation). This relationship shows that areas with the strongest T1w/T2w gradient tend to be those driving the discrepancy between the T1w/T2w and brain-specific gene PC1 topographies. This is consistent with the reported correspondence shown in Fig. 5c being limited by the relatively poor spatial resolution of the AHBA sampling.



Extended Data Figure 10: [Caption on next page]

Extended Data Figure 10: Validation of main results using the group-averaged (N=69) Conte69 T1w/T2w map. **(a)** The parcellated T1w/T2w map. **(b)** Human T1w/T2w map values are significantly lower in functionally defined association networks than in sensory networks ($P < 10^{-3}$; Wilcoxon signed-rank test). Error bars mark the std. dev. across areas. **(c)** The average expression map of genes preferentially expressed in human granular layer 4 (L4) is positively correlated with the human cortical T1w/T2w map ($r_s = 0.73$, $P < 10^{-5}$; Spearman rank correlation). **(d)** Average expression maps of laminar-specific genes show significant T1w/T2w map correlations (TMCs). L1-3: supragranular layers 1-3; L5/6: infragranular layers 5 and 6. **(e)** Genes coding for calretinin (*CALB2*; $r_s = -0.55$, $P < 10^{-5}$) and parvalbumin (*PVALB*; $r_s = 0.78$, $P < 10^{-5}$) exhibit homologous hierarchical gradients in human cortex. **(f)** TMCs of genes coding for markers of specific inhibitory interneuron cell types. **(g)** The gene coding for the NMDA receptor subunit NR2B (*GRIN2B*) exhibits a negative TMC ($r_s = -0.69$, $P < 10^{-5}$). **(h, i)** TMCs of genes coding for distinct subunits of NMDA & GABA_A receptors. Statistical significance for panels **d**, **f**, **h**, and **i** is calculated using a spatial autoregressive model (*, $P < 0.05$; **, $P < 10^{-2}$; ***, $P < 10^{-3}$). **(j)** PC1 captures a large fraction of total gene expression variance. *Inset:* Variance captured by PC1 for five gene sets: all genes, and genes preferentially expressed in brain, neurons, oligodendrocytes, and synaptic processes. **(k)** PC1 for this gene set is highly correlated with the T1w/T2w map ($r_s = 0.88$; $P < 10^{-4}$). **(l)** Across all sets, PC1 exhibits a highly similar areal topography to the T1w/T2w map (TMC range: 0.86–0.88; $P < 10^{-5}$ for each). **(m)** Gene expression variance captured by the T1w/T2w map ($\sigma_{\text{T1w/T2w}}^2$) relative to PC1 (σ_{PC1}^2).

Caption for Data File S1: Compiled monkey microanatomical data for cytoarchitectural type, Spearman rank correlations for projections terminating in each target area, inhibitory interneuron proportions, and pyramidal neuron dendritic spine counts (Figs. 2a,d,g; Extended Data Fig. 3).

Caption for Data File S2: T1w/T2w map correlation (TMC) values and cortical differential stability (DS_c) for all genes in our dataset, and sets of brain-, neuron-, oligodendrocyte-, and synaptome-specific genes.