nature neuroscience

Article

https://doi.org/10.1038/s41593-024-01812-2

The cell-type underpinnings of the human functional cortical connectome

Received: 27 July 2023

Accepted: 26 September 2024

Published online: 21 November 2024

Check for updates

Xi-Han Zhang[1] ✉, Kevin M. Anderson[1], Hao-Ming Dong[1], Sidhant Chopra[1], Elvisha Dhamala[2], Prashant S. Emani[3], Mark B. Gerstein[3,4,5,6], Daniel S. Margulies[7,8] & Avram J. Holmes[9] ✉

The functional properties of the human brain arise, in part, from the vast assortment of cell types that pattern the cerebral cortex. The cortical sheet can be broadly divided into distinct networks, which are embedded into processing streams, or gradients, that extend from unimodal systems through higher-order association territories. Here using microarray data from the Allen Human Brain Atlas and single-nucleus RNA-sequencing data from multiple cortical territories, we demonstrate that cell-type distributions are spatially coupled to the functional organization of cortex, as estimated through functional magnetic resonance imaging. Differentially enriched cells follow the spatial topography of both functional gradients and associated large-scale networks. Distinct cellular fingerprints were evident across networks, and a classifier trained on postmortem cell-type distributions was able to predict the functional network allegiance of cortical tissue samples. These data indicate that the in vivo organization of the cortical sheet is reflected in the spatial variability of its cellular composition.

A core goal of research in the brain sciences is to understand the multiscale relationships that link molecular and cellular processes with the in vivo functional organization of the human cerebral cortex. Historically, the definition of borders and associated areal parcels along the cortical sheet were determined by invasive techniques including histology, anatomical tract tracing, electrophysiology and lesion methods. Through these approaches, neuroscientists and histologists produced landmark maps that divide territories based on regional patterns of cytoarchitecture[1–4], revealing the presence of both serial and parallel information processing hierarchies[5,6]. Recently, the development of dense spatial transcriptomic atlases has enabled the study of cellular correlates of brain function in humans, for instance, as estimated through functional magnetic resonance imaging (fMRI)[7]. Initial work established molecular correlates of large-scale network organization[7–9], including genes encoding ion channels[10] and those enriched in supragranular layers of cortex[11], as well as associations between the spatial distribution of interneuron-linked genes and regional differences in fMRI signal variability[12,13]. However, the extent to which associated ex vivo cellular architectures mirror the hierarchical functional properties of the human cerebral cortex as measured by resting-state fMRI (rs-fMRI) has yet to be systematically investigated.

From sensation through cognition and action, the human cortex is organized into a multiscale system composed of areal units situated along segregated processing streams[6,14,15]. These units, or parcels, are embedded within corresponding networks evident through anatomical projections, task-evoked functioning and patterns of intrinsic activity[16–18]. Supporting this network architecture, converging evidence indicates the presence of a broad division separating unimodal

[1]Department of Psychology, Yale University, New Haven, CT, USA. [2]Institute of Behavioral Science, Feinstein Institutes for Medical Research, Manhasset, NY, USA. [3]Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA. [4]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA. [5]Department of Computer Science, Yale University, New Haven, CT, USA. [6]Department of Statistics and Data Science, Yale University, New Haven, CT, USA. [7]Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. [8]Cognitive Neuroanatomy Lab, Université Paris Cité, INCC UMR 8002, CNRS, Paris, France. [9]Department of Psychiatry, Brain Health Institute, Rutgers University, Piscataway, NJ, USA. ✉e-mail: xihan.zhang@yale.edu; avram.holmes@rutgers.edu

somatosensory/motor (somato/motor) and visual territories from the heteromodal association areas that integrate long-distance projections across distributed brain systems[5,19]. This hierarchical organizational property is reflected in the presence of functional gradients that span the cortical sheet, situating distinct networks and corresponding areal parcels along a continuous spectrum[20,21]. These gradients reflect low-dimensional representations of functional connectivity (FC), with the first, or primary, gradient anchored at one end by unimodal regions supporting primary sensory or motor functions and at the other end by the association cortex[22]. The second gradient reflects a sensory organization anchored at each end by either visual or somato/motor cortex[20,23]. However, while recent evidence suggests a genetic basis for the macroscale organization of the cortex[24–26], the extent to which diverse cellular profiles may underpin aspects of functional brain organization remains to be established.

The translational challenge of linking molecular and cellular processes with properties of functional organization is addressable, in part, by integrating transcriptomic data from ex vivo tissue samples with estimates of in vivo brain function[7]. Classic neuroanatomical discoveries revealed the evolutionary processes and developmental mechanisms that constrain the layout of cortical areas, their corresponding microstructure and anatomical connectivity[22,27]. Recent work supports the presence of broad axes of cortical organization[22,28], for instance, as reflected in the spatial distribution of receptor densities[29], intracortical myelination[30] and regional estimates of transcriptomic distinctiveness[31]. Preliminary studies associating gene expression to functional networks have revealed shared enrichment of genes among anatomical regions that are functionally coupled[8,10,11], perhaps reflecting the network-preferential presence of particular cell types. In line with this conjecture, subsequent work indicates that spatial profiles of fMRI signal variability follow the relative distribution of parvalbumin (PVALB)-expressing and somatostatin (SST)-expressing inhibitory interneurons[12], broadly separating unimodal and association cortices. Intriguingly, select areal boundaries derived from rs-fMRI have also been shown to correspond to histologically and structurally defined architectonic areas[17,23,32,33]. However, while these data suggest a link between the microscale (molecular and cellular) and macroscale (gradients and networks) properties of brain organization, our understanding of how the complex functional architecture of the cerebral cortex is supported by diverse cell types remains fragmentary. One possibility is that the relative preponderance of certain cell classes may spatially couple to gradually shifting gradient patterns across the cortical sheet. An alternative, but not mutually exclusive, hypothesis is that the spatial distribution of cortical cell types tracks the topographic organization of functionally connected but spatially distinct networks.

Cell types exhibit divergent properties, encompassing their morphology, electrophysiology, cell–cell connectivity and molecular content. Although these multidimensional features do not necessarily vary concurrently, recent developments in high-throughput single-nucleus RNA-sequencing (snRNA-seq) data have played a key role in characterizing broad profiles of cellular gene expression[34]. From technologies characterizing multimodal features of single cells, such as patch-seq[35], transcriptionally defined cell types demonstrate electrophysiological and morphological distinction[36,37]. Here we examine the association between functional gradients, networks and the spatial distribution of cell types, inferred through single nuclei from eight cortical regions[38] and patterns of gene transcription from bulk-tissue samples resected from donor tissue across cortex[7]. We demonstrate that imputed cell-type distributions spatially track the gradient organization of cortex, both at the level of individual cell types and multivariate cellular profiles. Suggesting the presence of a complementary network structure of cellular organization, distinct cellular enrichment patterns were also evident across large-scale networks. These cellular fingerprints can be used to predict the network allegiance of cortical parcels from their corresponding cell-type abundance measured in independent postmortem brain tissue. These data help address a key challenge in neuroscience to understand how cell-type distributions may underlie the in vivo functional properties of the human brain, establishing spatial correspondence between regional cellular profiles and the hierarchical organization cortex.

Results

The large-scale functional organization of the human cortex

The multiscale and hierarchical organization of cortex was characterized through both gradients and large-scale networks (Fig. 1). Diffusion map embedding[20,39,40] was applied to decompose vertex-level rs-fMRI FC into continuous gradients that capture the maximum variance in data from the Human Connectome Project (HCP, $n = 820$)[41]. To allow for functional measures to be associated with cellular profiles, the vertex-level gradient values were parceled into the Schaefer 400 atlas[32]. Parcel-level network assignment was obtained as detailed in ref. 23 (Fig. 1a). Next, we use these complementary gradient- and network-based approaches as a foundation to establish the cellular associates of the in vivo functional organization of the cerebral cortex.

The current analyses focused on the first two primary gradients, reflecting the canonical information processing hierarchies in the human cortex[6]. Unlike linear methods that reduce geometric dimensionality, diffusion map embedding allows topologically similar local and long-distance connections to be placed into common spaces with interpretable architectures[20,40]. The resulting gradients are unitless and reflect the position of vertices along an associated embedding axis that captures the primary differences in FC patterns. Consistent with previous work[20,40], the architecture of the first gradient (gradient 1) spans from unimodal areas (including primary visual, auditory, somatosensory and motor cortex) through transmodal, or association (default network), territories (Fig. 1b). The peak values in the second gradient (gradient 2) were evident along the central and calcarine sulcus, differentiating the somato/motor cortex from the primary visual system (Fig. 1c). The first two gradients account for a substantial proportion of the variance in FC (gradient 1, 26%; gradient 2, 12%). As initially reported in ref. 20, large-scale networks[23,32] are distributed across the cortical sheet and spatially ordered along these first two primary gradients (Fig. 1d), a property of cortical organization reflected in the repeating transitions between networks across cortical lobes. These data highlight complementary analytic frameworks that situate large-scale cortical networks and functions in separate domains along overlapping organizing axes.

Univariate cellular associates of cortical gradients

Areal parcellations derived from rs-fMRI have been shown to follow boundaries of select histologically and structurally defined architectonic areas, for instance, within somato/motor territories and putative language areas[23,32]. However, the extent to which regional cellular profiles spatially covary with the hierarchical organization of cortex has yet to be fully established in vivo. To relate inferred spatial distributions of ex vivo measured cell types to the two primary functional gradients, we estimated cell-type abundances using Allen Human Brain Atlas (AHBA) postmortem bulk gene expression samples via a previously validated method (Methods; Extended Data Fig. 1a)[7]. Transcriptional signatures identifying cell class in the AHBA bulk samples were derived from cortical snRNA-seq data of eight cortical areas as reported in the Jorstad dataset[38] (Extended Data Figs. 1 and 2). The present analyses focused on 24 cellular classes with distinct laminar specialization, developmental origins, morphology, spiking patterns and broad projection targets[42] (details in Extended Data Table 1). These cells include nine GABAergic inhibitory interneurons (PAX6, SNCG, VIP, LAMP5, LAMP5 LHX6, Chandelier, PVALB, SST CHODL and SST), nine glutamatergic excitatory neurons (L2/3 IT, L4 IT, L5 IT, L6 IT, L5 ET, L5/6 NP, L6 CT, L6b and L6 IT Car3), and six non-neuronal cells (Astro, Endo, VLMC, Oligo, OPC and Micro/PVM). The widely used snRNA-seq data provided in Lake

https://doi.org/10.1038/s41593-024-01812-2

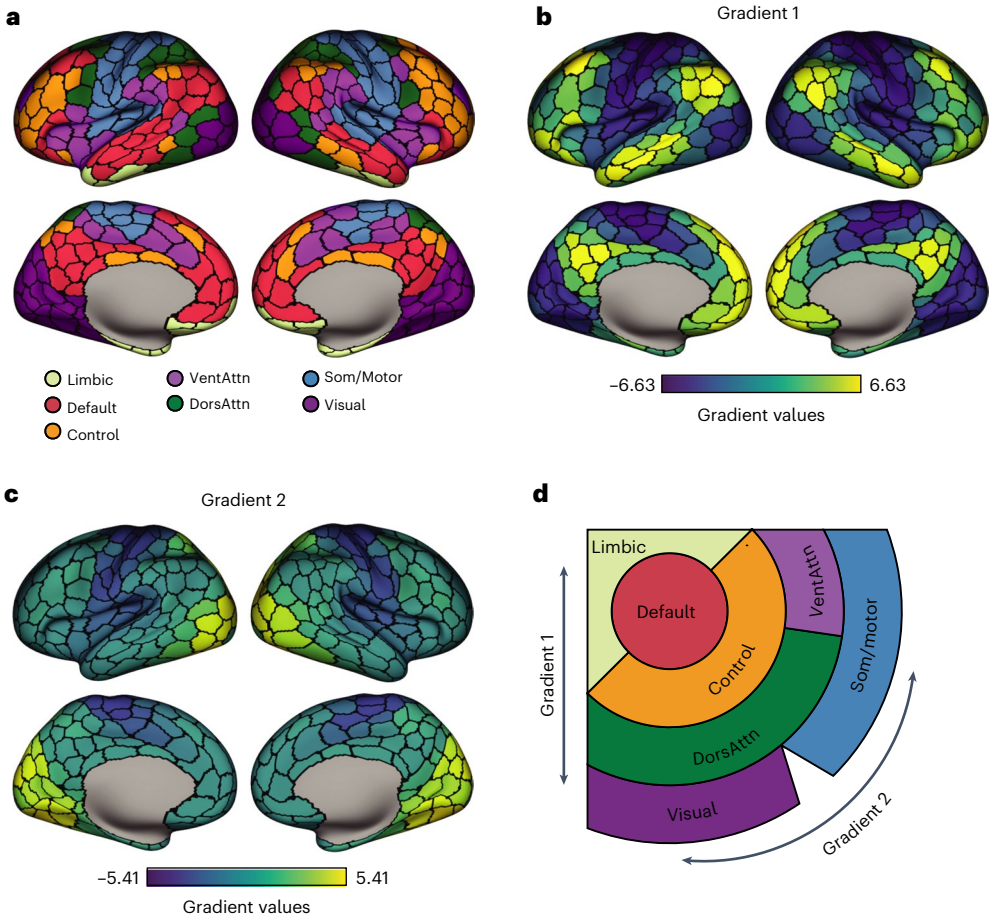


Fig. 1 | Large-scale functional networks are embedded along two principal gradients. a, The network organization of the human cerebral cortex is revealed through patterns of intrinsic functional connectivity. Network assignments were determined based on the seven-network solution from Yeo et al.[23] averaged across the 400-parcel functional atlas of Schaefer et al.[32]. Colors reflect regions estimated to be within the same network. **b**, The first principal gradient of intrinsic FC. Parcels are colored by their relative topological position spanning between the association cortex (bright yellow) and the unimodal cortex (dark blue). Scale bar reflects *z*-transformed principal gradient values derived from connectivity matrices using diffusion map embedding[39]. **c**, The second gradient of intrinsic FC is anchored within unimodal areas, including the primary visual cortex (bright yellow) at one end and the somatomotor/auditory cortex at the other (dark blue). **d**, The spatial organization of the seven networks along the two primary gradients. VentAttn, salience/ventral attention network; DorsAttn, dorsal attention network; Som/motor, somatosensory/motor network. Panel **d** adapted from ref. 20 under a Creative Commons licence CC BY.

dataset[43] were compared with these cell types (Supplementary Fig. 3), as a technical replication of the observed cell-to-function relationships (Supplementary Information). The abundance of each cell type was estimated in available bulk samples, which were further aggregated into the 400 cortical parcels (Methods). The resulting distributions of parcel-level cell-type abundances were examined relative to the in vivo functional gradient organization of the cortex. Statistical significance was established using permuted spin tests accounting for the spatial autocorrelation[44,45], False discovery rate (FDR) was corrected for 24 comparisons per gradient.

The associations between gradients and cell types are displayed in Fig. 2a (Lake dataset in Extended Data Fig. 4). Figure 2b,c shows univariate cell-type associations, and the full results are displayed in Extended Data Figs. 5 and 6 (Lake dataset in Supplementary Figs. 1 and 2). The estimated preponderances of imputed cell types are displayed across cortical parcels (Fig. 2b), demonstrating that certain cell-type abundances gradually increase/decrease across the functional networks distributed along the gradients, with their relative enrichment or absence evident within certain networks (Fig. 2c). The value of cell-type fractions is always positive, and gradient values represent the parcels' topological position relative to each other. Accordingly, both positive and negative correlations indicate that the abundance distribution of a given cell type follows the corresponding gradient values. Gradient 1 (Fig. 2b), spanning the unimodal-transmodal axis, was positively associated with the imputed spatial distributions of SNCG-like interneuron (SNCG; $\rho = 0.402$, $P < 0.001$), the interneuron SST ($\rho = 0.308$, $P = 0.036$) and layer 5 intratelencephalic-projecting excitatory neuron (L5 IT; $\rho = 0.281$, $P < 0.001$), indicating the three cell types are preferentially enriched in the transmodal end of gradient 1. Suggesting the abundance in unimodal areas, layer 4 intratelencephalic-projecting excitatory neuron (L4 IT; $\rho = -0.423$, $P < 0.001$) was negatively associated with gradient 1, enriched at the somato/motor and visual end.

SNCG is densely distributed within the temporoparietal junction, temporal lobe, medial prefrontal cortex (mPFC) and superior prefrontal through inferior frontal gyrus. SNCG is localized in the most superficial layers of the cerebral cortex (mainly L1 and partially L2/3)[38], where the majority of GABAergic interneurons that regulate the deeper-layer excitatory neurons through local inhibition are situated[46]. Originating from the caudal ganglionic eminence[47] (CGE) in the telencephalon, SNCG is matched with cholecystokinin-expressing layer 1 interneuron (CCK/In1 (ref. 43)) in the Lake dataset (Extended Data Fig. 3). In Lake dataset replication, In1 also displays the strongest positive correlation with gradient 1 (Extended Data Fig. 4; $\rho = 0.309$, $P = 0.011$). SNCG's homolog in rodents reflects the classic CCK basket interneuron[36,48,49]. Interneurons SST peak around the mPFC, temporal pole and inferior frontal gyrus and are densely distributed near the intraparietal sulcus.

https://doi.org/10.1038/s41593-024-01812-2

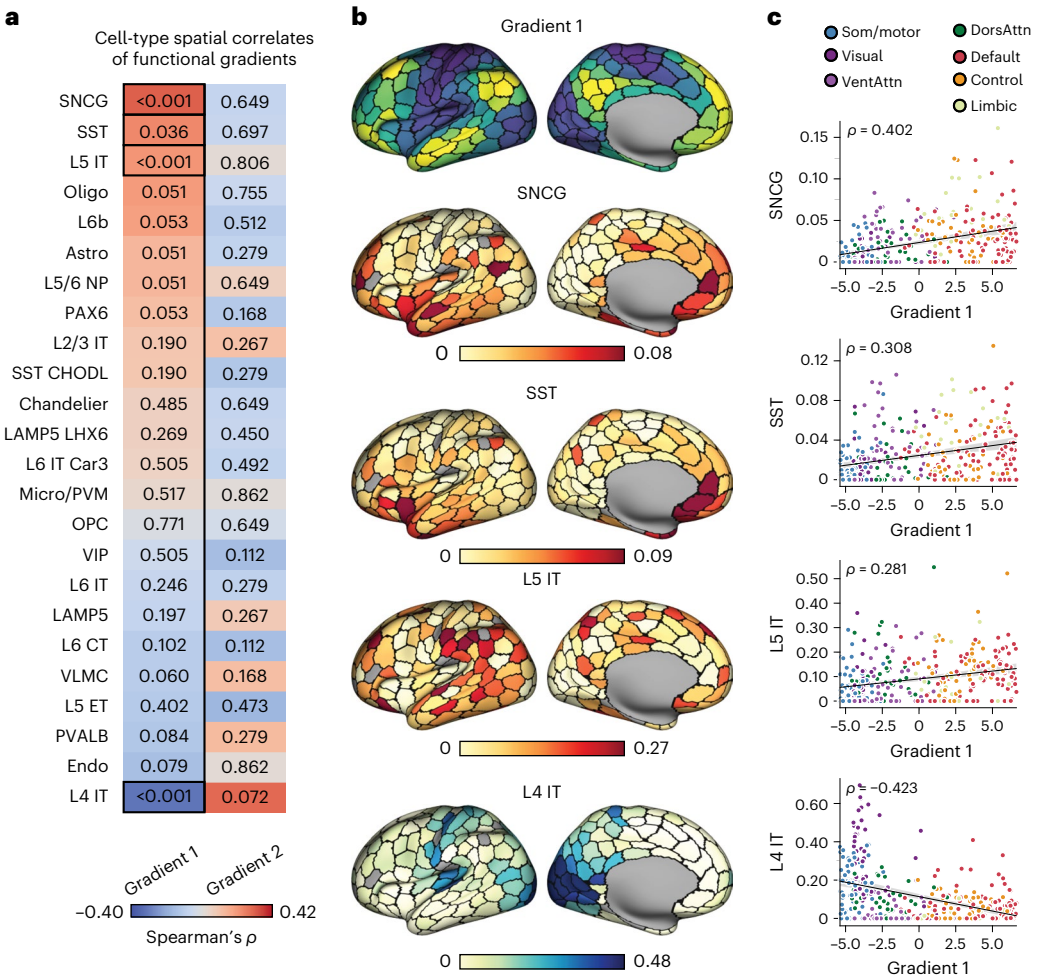


Fig. 2 | Univariate cell-type distributions align with functional gradient topographies. a, Cell types are imputed from gene expression in AHBA bulk-tissue samples. Single-cell signatures are constructed from independent tissue samples from eight cortical territories[41]. Resulting abundances of cell types were rank-ordered by spatial correlation to each principal functional gradient. Warm colors indicate a positive correlation, and numbers in each cell reflect the associated *P* value calculated from a two-sided spin test that controls for partial autocorrelation and was FDR-corrected for 24 comparisons. Correlations with $P < 0.05$ are denoted by black outlined boxes. Surviving from significance tests, the spatial pattern of gradient 1 was correlated with two interneuron subtypes (SNCG and SST), as well as two excitatory neuron subtypes (L5 IT and L4 IT). **b**, Imputed cell-type abundance distributed across the cortex suggests SNCG, SST and L5 IT are preferentially distributed around the transmodal end of gradient 1 (bright yellow), whereas L4 IT is preferentially distributed around the unimodal end (dark blue). Scale bars reflect estimated cell fractions. Parcels that are excluded from analyses and not covered by AHBA bulk samples are colored in gray. **c**, Scatter plots with each cortical parcel colored by the corresponding functional networks show that cell-type abundances gradually increase/decrease across the networks distributed along the gradients, with enrichment/absence evident within certain networks. Correlations were estimated by Spearman's ρ (as reflected in the scale bar in **a**), but for visual reference, the fitted linear regression lines with a 95% confidence interval (CI) predicting cellular abundance from gradient values are also displayed.

SST originates in the medial ganglionic eminence (MGE) and broadly corresponds to In8 (ref. 43) in Lake dataset (Extended Data Fig. 3b,c). SST is one of the biggest classes of interneurons[38]. The majority of these cells are localized in L2/3 across the seen cortical areas, whereas specialized subtypes are mainly present in L4–L5 of primary visual cortex[38]. In humans, the L2/3 SST (SST CALB1)[50] cells have morphologically distinct forms that reflect (1) Martinotti cells with dendrites mostly targeting L1, (2) double bouquet cells with dendrites targeting L1–L5 and (3) basket cells with mostly L3-targeting dendrites. These subtypes demonstrate diverse firing patterns ranging from fast to regular spiking.

Transmodal-associated L5 IT has a high abundance in the inferior parietal lobe, temporal lobe, dorsomedial prefrontal cortex and middle and anterior cingulate cortex. In humans, L5 IT has shorter basal dendrites (ends at L2/3) and simple apical dendrites compared with ITs in upper layers[51,52]. Rodent L5 IT homologs demonstrate broad axonal projections to the lateral hemisphere of cortex and striatum[36,53]. On the contrary, the unimodal-associated L4 IT is densely distributed in the primary visual, auditory, motor and somatosensory cortex. Mostly enriched in L4, which receives the sensory information from the thalamus, L4 IT has more diverse subtypes that are highly adapted to the corresponding sensory input[38]. Our observation that L4 IT is abundant in the primary motor cortex, which lacks a cell-packed granular L4, is well aligned with findings from rodent models[36,54]. While these analyses reveal isolated cell types that are preferentially distributed along functional gradients, in large part anchored at each gradient extrema, the large-scale organization of cortex is likely most apparent when simultaneously considering the distribution of multiple cell types. We further examined the combinatorial alignment of cell types to gradients in the next section.

Multivariate cellular profiles track cortical gradients

Above, we established spatial relationships between cell types studied in isolation and the first two functional gradients of connectivity[20]. However, the extent to which the macroscale organization of cortex

https://doi.org/10.1038/s41593-024-01812-2

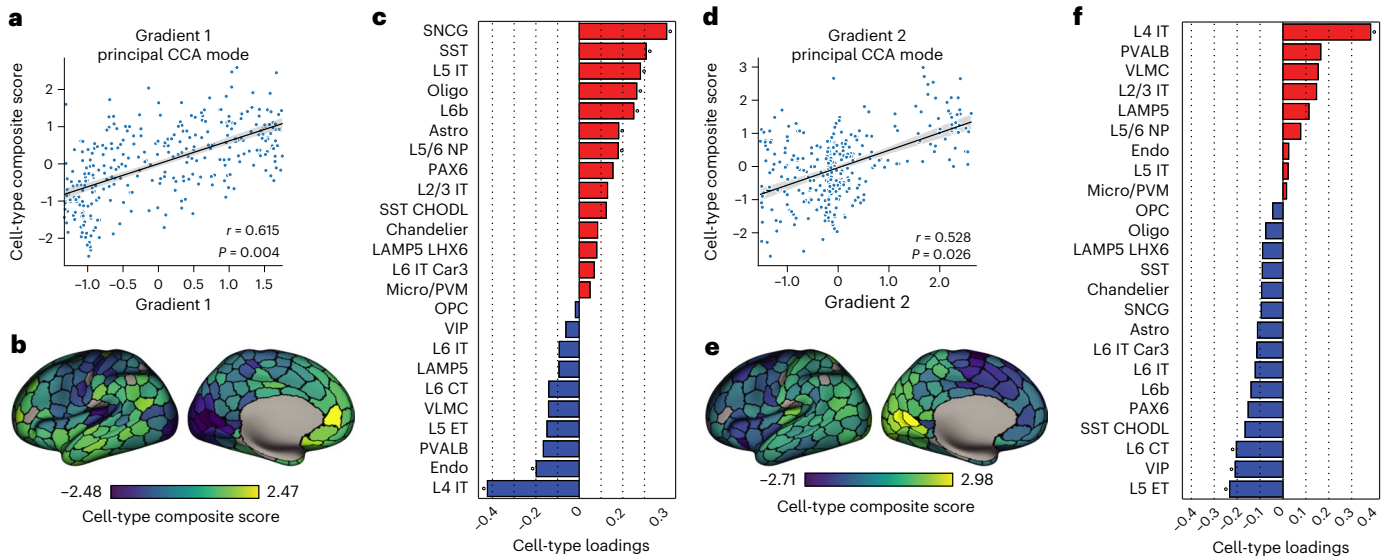


Fig. 3 | Multivariate cellular profiles follow the macroscale organization of cortex. a, The scatter plot displays the results of PermCCA, where the first functional gradient was positively associated with a composite score of cell-type abundances (95% CI). **b**, Cell-type composite score associated with the first functional gradient projected to the cortical surface. **c**, Canonical loadings of each cell type to the composite score implicate SNCG, SST, L5 IT, Oligo, layer 6b excitatory neuron (L6b), astrocyte (Astro), layer 5/6 near-projecting excitatory neuron (L5/6 NP), L4 IT and endothelial cells (Endo) (red, positive associations; blue, negative associations). **d**, The second functional gradient was positively associated with a cell-type composite score of cell-type abundance (95% CI). **e**, Cell-type composite score associated with the second functional gradient mapped to the cortical surface. **f**, Canonical loadings of each cell type to the composite score significantly implicated L4 IT, layer 5 extratelencephalic-projecting (L5 ET), VIP and L6 CT. *P* values in **a** and **d** are calculated from two-sided spin tests after correcting the two comparisons. Dots in **c** and **f** indicate indicate significant loadings after FDR corrections.

may be reflected in the spatial distribution of multiple cell types remains to be established. To examine the presence of these multivariate cellular profiles, we used permutational canonical correlation analysis (PermCCA)[55], which seeks the linear combination of cell-type distributions that are maximally correlated to each gradient. Inference on the resulting canonical variates was performed using spin-test permutations[44] to account for spatial autocorrelation.

Here the resulting cell-type composite scores reflect the linear combination of spatial distributions across all cell types that are maximally correlated to each functional gradient. As displayed in Fig. 3 (Lake dataset in Extended Data Fig. 7 and Supplementary Fig. 3), cell-type composite scores were correlated with both the first and second gradients across the three technical replications (Fig. 3a,d, Extended Data Fig. 7a,d and Supplementary Fig. 3a,d). In Fig. 3b,e (Lake dataset in Extended Data Fig. 7b,e and Supplementary Fig. 3b,e), we display the strength of involvement of each cell-type composite score associated with the first and second functional gradient projected to the cortical surface. At the univariate level, as reported above, a common spatial distribution for these gradient-associated cell types emerged where the profile of enrichment is anchored at one end of the gradient (Fig. 2). Consistent with this pattern, as reflected in our PermCCA cell-type loadings, gradient 1 ($r = 0.615$, $P = 0.004$) is most positively correlated with SNCG, SST and L5 IT and most negatively correlated with L4 IT. Here the order of the cells mirrors the order in their univariate correlation strength with gradient 1 (Fig. 2a). Gradient 2 ($r = 0.528$, $P = 0.026$) is most positively correlated with L4 IT that is matched up with Ex4 in Lake dataset (Extended Data Fig. 3). Ex4 is also the dominant contributor to the cellular multivariate correlation with gradient 2 (Extended Data Fig. 7 and Supplementary Fig. 3).

Although the cell types associated with gradients in isolation were also preferentially linked in the multivariate analyses, the variance explained by using a composite score now increases to 37.8% (gradient 1) and 27.8% (gradient 2; Fig. 3a,d) from 17.9% (gradient 1; Fig. 2c) in the univariate analyses (Fig. 2d). Accordingly, rather than being specific to isolated classes of cells, the observed topographic similarities may be most apparent when considering the combined spatial profiles of multiple cell types across cortex. To test this idea, we repeated the PermCCA iteratively, excluding different combinations of cells that reflect the major contributors for gradients 1 and 2 (Supplementary Tables 1–3). Critically, when simultaneously removing all the significant gradient-associated cell types revealed in the univariate analysis (gradient 1: SNCG, SST, L5 IT and L4 IT), the PermCCA results held for gradient 1 ($r = 0.503$, $P_{FDR} = 0.002$). Therefore, while certain cell types may preferentially follow the gradient architecture of cortex, the observed spatial relationships are robust and broadly conserved across a host of cell types (see cell-type abundance correlation matrix in Supplementary Figs. 4 and 5).

The cellular composition of large-scale functional networks

Certain cell types may possess a preferential relationship with specific functional networks, hinting at the presence of computation modules supporting a given network's specialized properties. To assess these network–cell relationships, we calculated an enrichment score for each cell type (Methods) across seven canonical functional networks[23]. For ease of interpretation, networks sharing similar enrichment profiles across cells and cell types with similar enrichment profiles across networks were grouped via hierarchical clustering[47] (Fig. 4a and Lake dataset in Extended Data Fig. 8). The primary divergence in cell-type hierarchical clusters followed the cell types that are associated with gradient 1 (Figs. 2a and 3c). Most interneurons and supporting cells are broadly abundant across networks at either unimodal or transmodal ends, as captured by clusters 1, 2, 6 and 8. Among them, SNCG, chandelier parvalbumin-expressing interneuron (Chandelier), oligodendrocytes (Oligo) and microglia/perivascular macrophage (Micro/PVM) are significantly enriched in the limbic network.

The excitatory neurons displayed relatively specific network enrichment patterns (clusters 3, 4, 5 and 7), hinting at potential cellular motifs that may underpin associated network functions. For example, L5 IT and L2/3 IT in cluster 4 are preferentially distributed in the dorsal attention and frontoparietal networks involved in top–down attentional orienting and executive processes[56]. Dendrites of L5 IT terminate at

https://doi.org/10.1038/s41593-024-01812-2

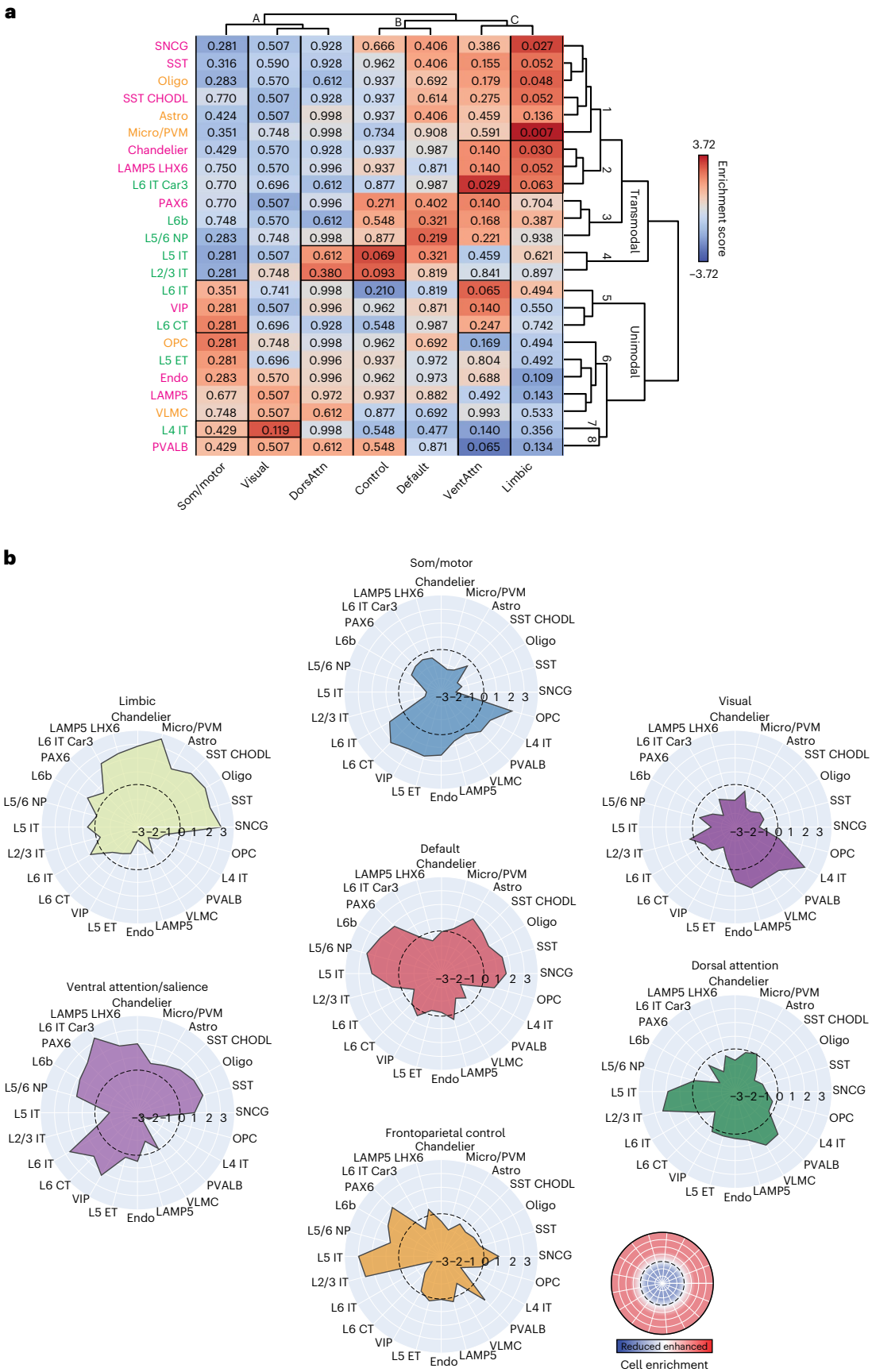


Fig. 4 | Large-scale functional networks demonstrate distinct cellular profiles. a, The table displays the relative cell-type enrichment, or absence, within each canonical functional network. For visualization, networks sharing similar enrichment profiles across cells and cell types with similar enrichment profiles across networks were grouped via hierarchical clustering. Letters (A, B and C) and numbers (1–8) denote network and cell-type clusters, respectively. Cell-type names are color-coded by broad class, with interneurons in pink, non-neuronal cells in orange and excitatory neurons in green. Empirical abundance for each cell type was aggregated within each of the seven large-scale functional networks. Corresponding null distributions were constructed from the parcel-level spin test, which accounts for spatial autocorrelation. Table fill colors reflect *z* scores, derived via subtracting the mean of the null from the observed empirical abundance and then dividing the difference by the s.d. of null distribution. Here *z* scores index empirical enrichment relative to the null. Warm colors indicate positive values (relatively enriched), and cold colors indicate negative values (relatively absent). Numbers in each cell represent the associated *P* value calculated from two-sided spin tests that control for spatial autocorrelation and were corrected for 24 comparisons. Reflecting the presence of a cell-type enrichment gradient spanning between somato/motor and limbic networks, each network shows a unique cell-type profile. Marked boxes reflect significant enrichment ($P < 0.05$). **b**, Polar plots of *z* score across 24 cell types for each network suggest the potential of cellular profiles that may serve as fingerprints that can distinguish each functional network. Score >0 lines (dashed) indicates when a cell type is enriched within a given network relative to the overall distribution across cortical parcels, whereas score <0 reflects the relative absence of a cell type.

the layer of L2/3 IT[51,52], the axons of which reversely target layers 5 and 6 (refs. 54,57). In cluster 5, vasoactive intestinal polypeptide-expressing interneuron (VIP), layer 6 intratelencephalic-projecting excitatory neuron (L6 IT) and layer 6 corticothalamic-projecting excitatory neurons (L6 CT) tend to be co-abundant in somato/motor and ventral attention networks involved in bottom–up sensory-driven attention/response[58]. L6 CT is innervated mostly by IT from higher-order cortical areas and partially by local and thalamic projections[57], whereas L6 IT tends to reciprocally connect with local neurons in layer 6 (refs. 54,57). In cluster 2, Chandelier, LAMP5- and LHX6-expressing interneuron (LAMP5 LHX6) and Car3-like layer 6 intratelencephalic-projecting excitatory neuron (L6 IT Car3) are (marginally) enriched in ventral attention and limbic networks. Unlike the fast-spiking PVALB that targets the perisomatic domain of pyramidal cells and broadly contributes to excitation–inhibition balance[53,59], the regular-spiking Chandelier preferentially innervates the initial axon segment, strengthening output control[60,61]. Although Chandelier can exist across layers 2–6, LAMP5 LHX6 is more specialized in the same layer[38] as L6 IT Car3.

The columns in Fig. 4a show enrichment profiles across cell types for each large-scale functional network. Consistent with their functional topography and/or spatial adjacency (Fig. 1a,d), somato/motor, visual and dorsal attention networks are grouped in cluster A, whereas default and frontoparietal networks are grouped in cluster B. Functionally and spatially distinct, ventral attention and limbic networks were grouped into cluster C. Of note, a hybrid cellular composition was evident in the ventral attention network, where properties of both unimodal and association territories were evident. While the ventral attention network primarily followed the cellular profile of the limbic system, it displayed L6 IT, VIP and L6 CT cells in a manner broadly consistent with the somato/motor network.

In addition to the gradient of cellular abundances from unimodal through association territories, as reflected in the polar plots in Fig. 4b, visually distinct network-level enrichment profiles are also apparent. Situated at distinct ends of the first functional gradient of connectivity (Fig. 1b), both the unimodal somato/motor and the limbic networks exhibit the most contrast between heightened or reduced enrichment profiles across cell types, with the ventral attention network also exhibiting pronounced variability in cellular enrichment. Broadly, the presence of distinct cellular fingerprints across functional networks suggests that cellular profiles derived from postmortem tissue samples may reflect, and be predictive of, the functional allegiance of a given cortical parcel.

Predicting network allegiances in postmortem tissue

Our findings raise an important question—can cellular profiles imputed from bulk-tissue samples be used to directly infer in vivo properties of brain organization? The above-mentioned analyses identify individual cell classes that preferentially follow cortical network topographies, suggesting the presence of network-specific cellular fingerprints. We use these results as a foundation to directly test the extent to which parcel-level cellular profiles can be used to predict their corresponding functional network assignments, as derived through fMRI. Support vector machines (SVMs) were trained to predict the network allegiance of postmortem tissue samples from parcel-level cellular abundances (Methods). The performance of models trained from empirical data was compared to a comprehensive set of increasingly stringent null models—(1) theoretical chance of predicting correctly given that the parcel is from the certain network (1/7, P_{chance}), (2) randomly permuted network labels (P_{perm}) or (3) shuffled labels while controlling for spatial autocorrelation (P_{spin}). Here we focus our interpretation on the most stringent significance null model (P_{spin}; see Fig. 5 for alternate significance thresholds and Lake dataset in Extended Data Fig. 9).

The SVM model successfully decoded parcel-level network assignments across cortex (empirical $F1_{median} = 0.377$, null $F1_{median} = 0.231$, $P_{spin} = 0.001$; see distribution plots in Supplementary Fig. 6), indicating that inferred cellular abundance from resected postmortem tissue reflects functionally relevant properties of brain organization. When considering individual networks, the SVM models trained on parcel-level imputed cell densities successfully predicted somato/motor ($P_{spin} = 0.023$), visual ($P_{spin} = 0.036$), ventral attention ($P_{spin} = 0.026$) and limbic ($P_{spin} = 0.038$) networks (all other $P_{spin} > 0.084$; Fig. 5a). When projecting parcel-level accuracies to the cortical surface (Fig. 5b), within-network variability was evident, indicating spatial heterogeneity of the cellular composition and associated network-level assignment accuracies (Supplementary Table 4). The confusion table, presented in Fig. 5c, highlights the assignment stability of the somato/motor, visual, ventral attention and limbic networks. Of note, while assignment accuracy is reduced in the three remaining networks, where misassignment occurs in parcels from frontoparietal, dorsal attention and default networks, they are likely to be assigned to other adjacent networks. For example, the dorsal attention network interdigitated between somato/motor and frontoparietal networks is more likely to be misassigned as the two. The SVM results emerging from the technical replication datasets and varying data exclusion criteria are displayed in Supplementary Figs. 6–12. Together, these data confirm the presence of distinct cellular fingerprints within some functional systems and support the need for additional research into the cytoarchitectonic determinants of network topography.

Discussion

Based on the first observation in ref. 2 regarding regional variations in the histological structure of gray matter across the cerebral cortex, the resection of postmortem tissue samples has revealed core insights into the cellular composition of the central nervous system. Recently, methodological advances have made it possible to map the macroscale organization of brain functions in vivo, providing the potential for deep biological insight into the genetic, molecular and cellular bases of cortical organization. Here integrating transcriptional and neuroimaging data, we demonstrate that imputed cell-type distributions follow the hierarchical functional architecture of the cortical sheet. Select cell types were found to spatially couple with aspects of cortical gradient and network organization, which was most evident in a distinction

https://doi.org/10.1038/s41593-024-01812-2

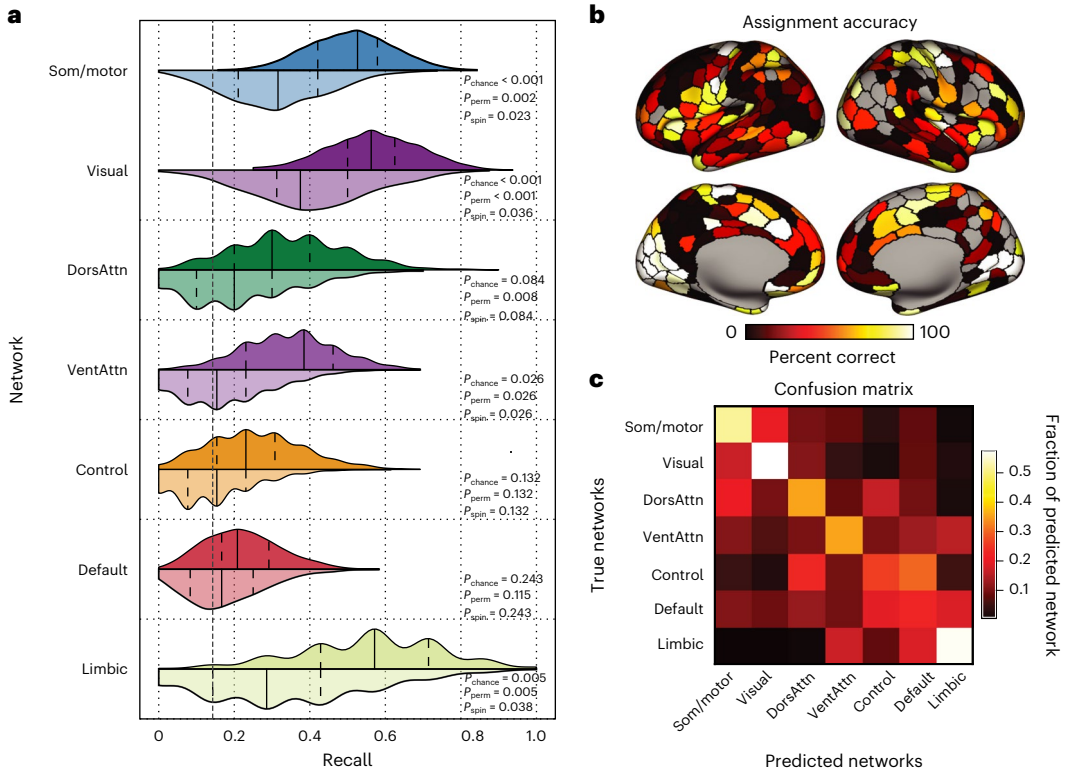


Fig. 5 | Large-scale functional network assignment can be predicted by cell-type abundance in postmortem tissue. a, Histograms display the SVM recall or the probability of correctly classifying a parcel to the associated network. These data suggest that the classifiers were able to predict somato/motor, ventral attention and limbic networks significantly above chance. Distributions in darker colors were constructed from 1,000 classifiers trained on real network labels, and the lighter-colored distribution represents classifiers trained on network labels shuffled by a spin test that controls for spatial autocorrelation. The solid lines indicate the median, and the dashed lines represent the quartiles of the distribution. *P* value of empirical recall was constructed from two-sided tests of increasing stringency—(1) theoretical chance (1/7, P_{chance}), (2) models trained from randomly permuted network labels (P_{perm}) or (3) labels shuffled by spin test (P_{spin}). **b**, Accuracy of network assignment across cortical parcels, calculated from all testing sets. **c**, Each row of the confusion matrix represents the fraction of parcels within the specific network that were predicted as belonging to each of the seven networks. The diagonal represents the percentage of correctly classified parcels within each network. Here the confusion matrix suggests a preferentially distinct cellular profile for somato/motor, visual and limbic networks. While classification accuracies were low for the remaining association cortex networks, dorsal attention, default and control networks display a higher rate of misclassifications among each other relative to unimodal networks.

between higher-order association and unimodal territories. Suggesting that regional variation in cellular profiles may reflect the layered aspects of cortical organization, multivariate cellular fingerprints captured a substantial portion of the spatial variability in both functional gradient topographies and parcel-level network assignments. Finally, imputed cell-type densities, derived from postmortem tissue, could be used to accurately predict parcel-level network assignments, demonstrating the presence of cellular markers of network-level brain functions as assessed through rs-fMRI. Together, these results indicate a close link between the functional organization of cortex and spatial variability of cell-type distributions, with important implications for the study of the cellular basis of brain functions across health and disease.

Select cell gene markers may spatially couple to regional differences in fMRI signal variability. Single-marker and polygenic cell deconvolution have established a spatially dependent relationship between heritable variance in in vivo fMRI signal amplitude and the topography of parvalbumin expression in postmortem brain tissue[12]. Extending upon this work, we demonstrate the presence of spatial alignment between regional cell densities and the functional gradient architecture of cortex. Most inhibitory interneurons broadly follow this macroscale architecture. SNCG (CCK/In1 in Lake dataset) and SST preferentially align with the transmodal end of the principal gradient (gradient 1), whereas PVALB are generally abundant around unimodal and adjacent areas. The presence of a dichotomous relationship between these groups of interneurons, embedded within a hierarchical somato/motor-association gradient, echoes their positioning during early embryonic development. Interneuron-destined cell types originate from distinct embryonic progenitor zones in the ventral telencephalon. Neural progenitor cells in CGE give rise to SNCG (CCK)[38], whereas the dorsal and ventral MGE progenitors generate SST and PVALB along anticorrelated spatial gradients[38,62,63]. After neurogenesis, these broad cell classes are governed by separate transcriptional cascades[62,64] that direct their tangential migration, layer-specific positioning and maturation[65,66] with distinct morpho-electrical properties[67]. The distinct computational properties of unimodal and heteromodal association territories are theorized to reflect the relative preponderance of associated cell types and their specialized roles in sensation and cognition[68].

SNCG, SST and PVALB differ across a host of morphological and firing properties, forming distinct local circuits that differentially bias network oscillations[69]. As a fast-spiking basket cell, PVALB has multipolar dendrites and radially arrayed axons[50] that broadly synapse on the perisomatic region of cortical projection neurons to regulate output[70]. Computational models in rodents suggest that a relative increase in PVALB, relative to SST, may result in stronger feedback inhibition on excitatory neurons within a given patch of cortex. This profile of excitability is thought to allow for short activation timescales that may be optimally suited for processing rapid-changing sensorimotor stimuli[12,68,71]. Conversely, SNCG cells encompass CCK interneurons that

display a broad range of spiking patterns, varying from synchronous transmission, which may enhance precise inhibition timing, through asynchronous repetitive activations that are thought to modulate inhibition strength and signal durations[72–74]. This wide spiking spectrum, well suited for the association cortex, enables them to integrate signals from various sources. Although future modeling work is warranted, these data suggest a link between the migration and maturation of local cellular circuits and the subsequent development and refinement of broader functional systems. Here we identify cell types that link to functional gradients in adults. Although these cell classes reach their cortical destinations in early life, the associated developmental and maturational trajectories extend through adolescence[40]. It remains to be determined how cell maturation, refinement and synaptic connectivity link to cortical organization across the lifespan.

Embedded within the sweeping cellular and transcriptional gradients that span cortex, select functional networks are co-enriched for distinct groups of excitatory and inhibitory neurons, which may reflect corresponding computational motifs. Enriched in unimodal sensory territories, especially the primary visual cortex, L4 IT cells are mainly distributed in layer 4 (or 3/5 if layer 4 is missing from the areal laminar structure)[38], which is highly specialized for processing incoming sensory information. In rodents, the dendrites of L4 IT are simple, not tufted[36] and suited to receive visual signals from thalamic ganglion cells that project at layer 4 (refs. 63,75). The local axons of L4 IT primarily project upwards to IT neurons in layer 2/3 (ref. 57) as part of the feedforward pathways, which ultimately project to layer 5 in association areas[36]. Here L2/3 IT and L5 IT are clustered into the same cell grouping due to their shared enrichment profiles across networks. This is particularly evident in the dorsal attention and frontoparietal networks involved in top–down attentional orienting and executive functioning[56]. Dendrites of L5 IT terminate at the layer of L2/3 IT[51,52], the axons of which reversely target layer 5 (refs. 54,57). While the cell–cell connectivity patterns between L5 IT and L2/3 IT require further study, the observed co-enrichment profile suggests the potential presence of neural circuits supporting the top–down processes across association networks. Similarly, L6 IT, L6 CT and VIP are co-abundant in somato/motor and ventral attention networks that support bottom–up sensory-driven attention/response[58]. In contrast to ITs in upper layers, L6 IT lacks tuft dendrites terminating at layers 1–4 (ref. 51), but possesses mutual connectivity with local neurons in layer 6 (refs. 51,57). Also situated in layer 6 (ref. 38), L6 CT cells receive signals from local and 'higher-order' cortical deep-layer ITs[36,57] and send efferent projections to thalamus. Distributed in layers 2 and 3, the regular-spiking VIP has dendrites and axons widely extending to layers 1–6 (ref. 50). Mostly regulating other interneurons, VIP might indirectly contribute to these processes. The salience/ventral attention network is linked to the accelerated maturation of adult-like cortical organization[76]. Here the co-abundance of mid-to-deep-layer cell types in the attention network may hint at potential mechanisms for cortical development, as transcriptional programs are coordinated within cell classes across developmental phases[77]. Future work should directly test the manner through which specialized cognitive processes supported by cortical networks may emerge from spatially distinct cellular connectivity motifs.

The large-scale network architecture of the cortex includes abrupt transitions that are embedded along continuous functional gradients[20]. Broadly, similar patterns are evident in cortical atlases defined through cell staining and morphological analysis, where homogeneous cell components are evident within local patches, while abrupt transitions can occur between some adjacent territories[3]. Emerging evidence suggests links between in vivo network parcellations of cortex and the presence of some cytoarchitectonic boundaries. Here functional network-defined boundaries and parcels have been observed to adhere to select histologically defined areas, including Broca's area (area 44) and aspects of the postcentral sulcus (areas 2 and 3)[32]. In the present analyses, using the imputed abundance across neuronal and non-neuronal cells in postmortem tissue, we demonstrate that the network allegiance across cortical parcels can be generally predicted. Although parcel-level prediction was broadly evident across cortex, substantial spatial heterogeneity was evident in the relations linking cellular profiles with network assignments. The somato/motor, visual, limbic and ventral attention networks exhibited the most distinct and predictable cellular profiles. This predictive profile emerges through the preferential enrichment or absence of cells in the somato/motor, visual and limbic networks relative to other networks, as well as the presence of a unique mixture of unimodal- and transmodal-associated cells in ventral attention-linked parcels. In contrast, control, dorsal attention and default networks displayed smoother and/or more homogeneous spatial transitions across cell types. This pattern may reflect the presence of smooth patterns of cellular composition near the boundaries between association networks, consistent with previous reports of transcriptional distinctiveness across cortex[31]. Highlighting the translational potential of multiscale neuroscience approaches, these data demonstrate predictive relationships connecting cellular profiles across the cortex with the in vivo functional organization of the human brain.

The present work should be interpreted in light of several limitations. First, the reported cell-type abundances are imputed from the bulk-tissue microarray data based on the gene expression signatures constructed from snRNA-seq. The microarray approach does not provide direct estimates of gene transcription; rather, we examine within-probe differences across samples. To obtain estimates robust across donors, bulk samples require aggregation into parcels, limiting the spatial resolution. As parcellation and cell-type definitions improve, the pattern observed will likely unfold in more detail. Second, we include a technical replication of our cell imputation approach by using additional single-cell data. Associated correspondences between imputation approaches are reported in the supplement. Of note, a lack of spatial granularity in the available tissue samples may have resulted in an underestimation of the true relationships linking the topographic distributions of cell types and brain functions. In mice, interneuron cell types are broadly conserved across cortical regions, while pyramidal cell diversity shows higher spatial variability[48], which is partially reflected in human data from ref. 38 across eight cortical areas. The imputed spatial distribution of cell types common in the available single-cell samples showed robust patterns, but it remains unclear how the cell-type diversity varies across the human cortex. As single-cell samples covering more cortical regions are gradually developed, future work should incorporate these spatially variable profiles when considering cellular abundances. Finally, in the present work, cells are defined from transcriptomics. Although initial efforts examining transcriptional profiles, morphology or firing patterns[70] have demonstrated reliable mapping between multimodal features and broad transcriptional cell types (the 24 cell classes)[38,46,50,52], finer clusters within these broad classes demonstrate nuanced morpho-electrical diversity[35]. Follow-up studies should consider how to best integrate the hierarchical structure of these diverse cell definitions in analyses while also examining alternate organizational models of in vivo brain functioning.

The present results demonstrate that the functional gradients and networks of the cerebral cortex are linked to spatial variability in cellular profiles. These data suggest that the imputed cell-type densities from postmortem tissue capture global patterns of FC as assessed through rs-fMRI, revealing the potential to bridge across in vivo and ex vivo methods in the study of human brain functions. Collectively, these discoveries highlight a connection between the functional organization of the cortex and its cellular underpinnings, with significant implications for understanding the cellular basis of brain functions across health and disease.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-024-01812-2.

References

1. Cajal, S. R. Y. *Recollections of My Life* (MIT Press, 1989).
2. Meynert, T. *Der Bau Der Gross-Hirnrinde Und Seine Örtlichen Verschiedenheiten, Nebst Einem Pathologisch-Anatomischen Corollarium: Separat-Abdruck Aus Der "Vierteljahrsschrift Für Psychiatrie Etc."* (Heuser, 1872).
3. Brodmann, K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, 1909).
4. Von Economo, C. F., Koskinas, G. N. & Triarhou, L. C. *Atlas of Cytoarchitectonics of the Adult Human Cerebral Cortex* 10 (Karger, 2008).
5. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
6. Mesulam, M. M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).
7. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
8. Anderson, K. M. et al. Gene expression links functional networks across cortex and striatum. *Nat. Commun.* **9**, 1428 (2018).
9. Hansen, J. Y. et al. Mapping gene transcription and neurocognition across human neocortex. *Nat. Hum. Behav.* **5**, 1240–1250 (2021).
10. Richiardi, J. et al. Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–1244 (2015).
11. 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. USA* **113**, E469–E478 (2016).
12. Anderson, K. M. et al. Transcriptional and imaging-genetic association of cortical interneurons, brain function, and schizophrenia risk. *Nat. Commun.* **11**, 2889 (2020).
13. Anderson, K. M. et al. Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. *Proc. Natl Acad. Sci. USA* **117**, 25138–25149 (2020).
14. Lashley, K. S. & Clark, G. The cytoarchitecture of the cerebral cortex of ateles: a critical examination of architectonic studies. *J. Comp. Neurol.* **85**, 223–305 (1946).
15. Gordon, E. M. et al. A somato-cognitive action network alternates with effector regions in motor cortex. *Nature* **617**, 351–359 (2023).
16. Lu, J. et al. Focal pontine lesions provide evidence that intrinsic functional connectivity reflects polysynaptic anatomical pathways. *J. Neurosci.* **31**, 15065–15071 (2011).
17. Bressler, S. L. & Menon, V. Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci.* **14**, 277–290 (2010).
18. Crossley, N. A. et al. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proc. Natl Acad. Sci. USA* **110**, 11583–11588 (2013).
19. Goldman-Rakic, P. S. Topography of cognition: parallel distributed networks in primate association cortex. *Annu. Rev. Neurosci.* **11**, 137–156 (1988).
20. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci. USA* **113**, 12574–12579 (2016).
21. Pang, J. C. et al. Geometric constraints on human brain function. *Nature* **618**, 566–574 (2023).
22. Sydnor, V. J. et al. Neurodevelopment of the association cortices: patterns, mechanisms, and implications for psychopathology. *Neuron* **109**, 2820–2846 (2021).
23. Yeo, Thomas et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
24. Anderson, K. M. et al. Heritability of individualized cortical network topography. *Proc. Natl Acad. Sci. USA* **118**, e2016271118 (2021).
25. Burger, B. et al. Disentangling cortical functional connectivity strength and topography reveals divergent roles of genes and environment. *NeuroImage* **247**, 118770 (2022).
26. Valk, S. L. et al. Genetic and phylogenetic uncoupling of structure and function in human transmodal cortex. *Nat. Commun.* **13**, 2341 (2022).
27. Finlay, B. L. & Uchiyama, R. Developmental mechanisms channeling cortical evolution. *Trends Neurosci.* **38**, 69–76 (2015).
28. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-scale gradients in human cortical organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
29. Froudist-Walsh, S. et al. Gradients of neurotransmitter receptor expression in the macaque cortex. *Nat. Neurosci.* **26**, 1281–1294 (2023).
30. Burt, J. B. et al. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259 (2018).
31. Wagstyl, K. et al. Transcriptional cartography integrates multiscale biology of the human cortex. *eLife* **12**, RP86933 (2024).
32. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
33. Haak, K. V., Marquand, A. F. & Beckmann, C. F. Connectopic mapping with resting-state fMRI. *NeuroImage* **170**, 83–94 (2018).
34. Zeng, H. What is a cell type and how to define it? *Cell* **185**, 2739–2755 (2022).
35. Berg, J. et al. Human neocortical expansion involves glutamatergic neuron diversification. *Nature* **598**, 151–158 (2021).
36. Callaway, E. M. et al. A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).
37. Boldog, E. et al. Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nat. Neurosci.* **21**, 1185–1195 (2018).
38. Jorstad, N. L. et al. Transcriptomic cytoarchitecture reveals principles of human neocortex organization. *Science* **382**, eadf6812 (2023).
39. Coifman, R. R. & Lafon, S. Diffusion maps. *Appl. Comput. Harmon. Anal.* **21**, 5–30 (2006).
40. Dong, H.-M., Margulies, D. S., Zuo, X.-N. & Holmes, A. J. Shifting gradients of macroscale cortical organization mark the transition from childhood to adolescence. *Proc. Natl Acad. Sci. USA* **118**, e2024448118 (2021).
41. Van Essen, D. C. et al. The WU-Minn Human Connectome Project: an overview. *NeuroImage* **80**, 62–79 (2013).
42. Jorstad, N. L. et al. Comparative transcriptomics reveals human-specific cortical features. *Science* **382**, eade9516 (2023).
43. Lake, B. B. et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* **36**, 70–80 (2018).
44. Alexander-Bloch, A. F. et al. On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage* **178**, 540–551 (2018).
45. Gordon, E. M. et al. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* **26**, 288–303 (2016).

46. Chartrand, T. et al. Morphoelectric and transcriptomic divergence of the layer 1 interneuron repertoire in human versus mouse neocortex. *Science* **382**, eadf0805 (2023).
47. Nery, S., Fishell, G. & Corbin, J. G. The caudal ganglionic eminence is a source of distinct cortical and subcortical cell populations. *Nat. Neurosci.* **5**, 1279–1287 (2002).
48. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
49. Dudok, B. et al. Alternating sources of perisomatic inhibition during behavior. *Neuron* **109**, 997–1012 (2021).
50. Lee, B. R. et al. Signature morphoelectric properties of diverse GABAergic interneurons in the human neocortex. *Science* **382**, eadf6484 (2023).
51. Mohan, H. et al. Dendritic and axonal architecture of individual pyramidal neurons across layers of adult human neocortex. *Cereb. Cortex* **25**, 4839–4853 (2015).
52. Kalmbach, B. E. et al. Signature morpho-electric, transcriptomic, and dendritic properties of human layer 5 neocortical pyramidal neurons. *Neuron* **109**, 2914–2927 (2021).
53. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
54. Mao, X. & Staiger, J. F. Multimodal cortical neuronal cell type classification. *Pflügers Arch.* **476**, 721–733 (2024).
55. Winkler, A. M., Renaud, O., Smith, S. M. & Nichols, T. E. Permutation inference for canonical correlation analysis. *NeuroImage* **220**, 117065 (2020).
56. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
57. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
58. Vossel, S., Geng, J. J. & Fink, G. R. Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* **20**, 150–159 (2014).
59. Kong, X. et al. Sensory-motor cortices shape functional connectivity dynamics in the human brain. *Nat. Commun.* **12**, 6373 (2021).
60. Scala, F. et al. Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature* **598**, 144–150 (2021).
61. Gouwens, N. W. et al. Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. *Cell* **183**, 935–953 (2020).
62. Velmeshev, D. et al. Single-cell analysis of prenatal and postnatal human cortical development. *Science* **382**, eadf0834 (2023).
63. Hu, J. S., Vogt, D., Sandberg, M. & Rubenstein, J. L. Cortical interneuron development: a tale of time and space. *Development* **144**, 3867–3878 (2017).
64. Zhu, K. et al. Multi-omic profiling of the developing human cerebral cortex at the single-cell level. *Sci. Adv.* **9**, eadg3754 (2023).
65. Miyoshi, G. & Fishell, G. GABAergic interneuron lineages selectively sort into specific cortical layers during early postnatal development. *Cereb. Cortex* **21**, 845–852 (2011).
66. Miyoshi, G. Elucidating the developmental trajectories of GABAergic cortical interneuron subtypes. *Neurosci. Res.* **138**, 26–32 (2019).
67. Wamsley, B. & Fishell, G. Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nat. Rev. Neurosci.* **18**, 299–309 (2017).
68. Kim, Y. et al. Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* **171**, 456–469 (2017).
69. Whissell, P., Cajanding, J., Fogel, N. & Kim, J. C. Comparative density of CCK- and PV-GABA cells within the cortex and hippocampus. *Front. Neuroanat.* **9**, 124 (2015).
70. Kepecs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014).
71. 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**, 419–431 (2015).
72. Armstrong, C. & Soltesz, I. Basket cell dichotomy in microcircuit function. *J. Physiol.* **590**, 683–694 (2012).
73. Hefft, S. & Jonas, P. Asynchronous GABA release generates long-lasting inhibition at a hippocampal interneuron–principal neuron synapse. *Nat. Neurosci.* **8**, 1319–1328 (2005).
74. Kohus, Z. et al. Properties and dynamics of inhibitory synaptic communication within the CA3 microcircuits of pyramidal cells and interneurons expressing parvalbumin or cholecystokinin. *J. Physiol.* **594**, 3745–3774 (2016).
75. Nassi, J. J. & Callaway, E. M. Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* **10**, 360–372 (2009).
76. Dong, H.-M. et al. Ventral attention network connectivity is linked to cortical maturation and cognitive ability in childhood. *Nat. Neurosci.* **27**, 2009–2020 (2024).
77. Dear, R. et al. Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. *Nat. Neurosci.* **27**, 1075–1086 (2024).

Methods

FC gradient analysis

Gradients, or components with similar FC patterns, were derived in a manner consistent with ref. 20. Briefly, FC matrices averaged from 820 participants in the HCP dataset[41] coregistered via Multimodal Surface Matching, All modalities (MSMAll) were downloaded from ConnectomeDB[78] (https://www.humanconnectome.org/storage/app/media/documentation/s900/820_Group-average_rfMRI_Connectivity_December2015.pdf). The 10,242 × 10,242 per hemisphere cortical FC matrices (represented in fsaverage5 surface space[79]) for each participant was calculated from 1-h rs-fMRI concatenated from four minimally preprocessed[80–83], spatially normalized 15-min scans. From these group-averaged FC matrices, correlation coefficients were Fisher *z*-transformed via a hyperbolic tangent function to scale the value between −1 and 1. The top 10% connections of each vertex were preserved, and all other values were set to 0 to enforce sparsity. The cosine distance between any two rows of the FCz matrix was calculated and then subtracted from 1 to generate a symmetrical similarity matrix. Gradients were derived from the similarity matrix by diffusion map embedding, as validated and detailed in refs. 40,20 (https://neuroanatomyandconnectivity.github.io/gradient_analysis/). This approach nonlinearly projects high-dimensional FC into a low-dimensional space. Here a gradient reflects an axis of FC variance along which cortical vertices fall in a spatially continuous order, with adjacent vertices sharing similar geographically short- and long-range correlations to the rest of the cortex. The two gradients explaining the highest variance were selected for subsequent analysis. Vertex-level gradients were averaged across the 400 cortical parcels in the Schaefer functional atlas[32].

Functional parcellation analysis

To characterize the functional network structure of the cortical sheet, we used 400 roughly symmetric ROIs from seven specific brain networks[33] in the left and right hemispheres as derived through the cortical parcellation of ref. 32 (https://github.com/ThomasYeoLab/CBIG). The functional networks used here were previously derived and validated using data from 1000 adults in the Genomics Superstruct Project[32,33,84]. In short, each network is a cluster of vertices that shares homogenous rs-fMRI FC with the rest of the cortex.

Brain gene expression data processing

Human microarray gene expression data obtained from bulk samples of six postmortem brains were downloaded from the AHBA dataset (http://human.brain-map.org/)[7]. Raw data were processed using the abagen toolbox (https://abagen.readthedocs.io/en/stable)[85,86] at the sample level, following the practice recommended in ref. 87 and implemented by other studies[9]. Probes were reannotated using data provided in ref. 85, and those without Entrez IDs were excluded. Probes that exceed background noise in 30% of all tissue samples were included, among which the probe with the highest differential stability for each gene was selected. In total, 16,383 genes were retained after the processing. For each donor, tissue sample expression values were first *z*-scored across genes, and then these gene expression values were then *z*-scored across samples. Consistent with ref. 12, individual cortical tissue samples were mapped to each AHBA donor's Freesurfer-derived cortical surfaces (FreeSurfer v6.0.0), downloaded from ref. 88. Native space mid-thickness surfaces were transformed to a common fsLR32k group space while maintaining the native cortical geometry of each individual donor. The native voxel coordinate of each tissue sample was mapped to the closest surface vertex using tools from the HCP workbench. Tissue samples were included if they were collected from less than 4 mm from the nearest surface vertex, resulting in 1,676 analyzable cortical samples.

Cell-type deconvolution

Cortical cell-type abundance distributions were inferred following the procedures detailed in ref. 12. In brief, single-nucleus droplet-based sequencing (snDrop-seq) data obtained from ref. 38 were downloaded from cellxgene-census (https://cellxgene.cziscience.com/collections/d17249d2-0e6e-4500-abb8-e6c93fa1ac6f). Data from ref. 43 were downloaded from the Gene Expression Omnibus website (GSE97930; https://www.ncbi.nlm.nih.gov/geo). Count matrices derived from unique molecular identifiers (UMIs) were preprocessed via Seurat[89], where outlier cells and minimally expressed genes were filtered, and then the data were log-normalized. Genes were referred to by Entrez IDs, among which only the IDs shared by both Jorstad/Lake and AHBA datasets were included. The marker genes that were found reliable across all eight cortical areas by ref. 38 were used to further narrow down the gene sets to reduce computational load. The groupings across area subclasses defined in ref. 38 or superordinate cell identities defined in ref. 43 were applied for categorizing transcriptionally similar cell types to reduce collinearity. After processing, the snDrop-seq data were de-log-transformed before feeding into CIBERSORTx (https://cibersortx.stanford.edu/)[90] as a reference for cell-type abundance imputation on each AHBA bulk-tissue sample. For the Jorstad dataset, given that the marker genes were selected by robustness across the eight cortical areas, all the cortical samples were combined together to derive the gene signature matrices of the 24 cell types. In the Lake dataset, gene signature matrices for 18 cell types were derived from visual (Lake dataset VIS) and frontal (Lake dataset DFC) samples separately. Cell-type abundances were consequently imputed from across AHBA samples, taking each signature matrix as a reference per donor. Cell-type matchings between the Jorstad and Lake datasets were conducted via the label-transferring algorithms in Seurat v5 (https://satijalab.org/seurat/)[91] based on the UMI count matrices (R v4.1.1), following procedures in ref. 92. Single cells from Jorstad dataset were transferred to the reference Lake dataset. The matching scores between cell types from the two datasets were estimated as the percentage of cells from each Jorstad dataset cluster aligned with the specific Lake dataset cell cluster. The correlation between Lake dataset VIS- and Lake dataset DFC-derived gene signatures was validated in ref. 12. To further minimize the effects of spatial heterogeneity of single-cell transcriptomic signatures, later analysis only used common cell types between Lake dataset DFC (excluding Ex2) and Lake dataset VIS (excluding In2), which resulted in 17 cell types in total. The cell-type abundances for each AHBA cortical sample were mapped to the cortical vertices represented in fsaverage6 surface space and then parceled into the Schaefer 400 atlas[32]. Here samples were first averaged at the individual donor level within parcels and then averaged across donors.

Identification of cell types spatially correlated to functional gradients

The spatial pattern of each of the 24/17 cell-type abundances parceled in the Schaefer atlas was correlated (Spearman's ρ) with the primary and secondary functional gradients parceled using the same atlas. The statistical significance of the correlation was assessed through spin test[44] (MATLAB R2021a), which permuted the gradient at the vertex level (represented in fsaverage5) 1,000 times while reserving the spatial autocorrelation (https://github.com/spin-test/spin-test). The permuted gradient and cell-type correlations were used to construct a null distribution of correlation values. The cell-type abundances inferred from each technical replication (Jorstad and Lake datasets) were tested separately and were FDR-corrected for 24/34 multiple comparisons.

Examination of cell types' combinational correlation to functional gradients

PermCCA was used to investigate the multivariate relationship between the spatial distribution of cell types and each gradient. CCA allows us to examine the linear combination of all the cell-type abundances that maximally correlate with each gradient. The statistical significance of the canonical variates was tested via a permutation method

that controls for cortical spatial autocorrelation (https://github.com/andersonwinkler/PermCCA)[55], where the null distributions of gradients were generated from the spin test described in the previous section. Then FDR was applied to correct for multiple comparisons (two gradients). The cell types' linear combinational correlation to the first two gradients was measured by cell-type composite score. Each cell type's contribution to this correlation was measured by loadings, the correlation between cell-type abundance distribution and gradient canonical variate.

PermCCA was repeated with the removal of different combinations of the cell types that were each univariately correlated with gradients 1 and 2. Each group of PermCCA results was FDR-corrected for the number of removal combinations (12 for gradient 1 with Lake dataset DFC imputed cell types, 10 for gradient 1 with Lake dataset VIS imputed cell types; 3 for gradient 2 with Lake dataset DFC imputed cell types, 1 for gradient 2 with Lake dataset VIS imputed cell types).

Cell-type enrichment in functional networks

Each one of the 400 cortical parcels was assigned to a functional network in a validated seven-network solution derived in ref. 23. Across each cell type, the empirical abundances were averaged across parcels within a given functional network. The empirical abundances were then permuted across the cortex 1,000 times, controlling for spatial autocorrelation, yielding 1,000 null models for each cell type. Given that the cell-type abundances were aggregated at the parcel level, the Cornblath version of spin test[93] was used as recommended by a study discussed in ref. 94, which projects parcel abundances to vertices, rotates and takes the mean of vertices in each parcel. The same within-network cell-type abundance averaging process was repeated on these null models, generating a null distribution of mean abundance for each cell type within each network. For each type of cell, the enrichment score of a network was calculated by taking the difference between the empirical abundance and the mean of the null abundance distribution and then dividing by the s.d. of the null abundance distribution. *P* values were first calculated from two-tailed tests and then FDR-corrected for multiple corrections across all cell types for each network.

Cell types predicting functional networks

SVMs (https://scikit-learn.org/stable/modules/svm.html#svm-classification) were trained to predict the functional network each cortical parcel belongs to based on abundances of 24/17 cell types within that parcel (Python v3.8.8). Because two of the six donors in AHBA had samples from the right hemisphere, models were trained on parcels from only the left hemisphere and from both hemispheres as two parallel groups. For the Lake dataset, within each group, three subgroups of models were trained separately, based on the cell-type abundance imputed from independent single-cell samples, Lake dataset DFC and Lake dataset VIS, as two replicates. Two subgroups were trained from the two replicates, respectively, and one subgroup ensembled the information from the two replicates. The Jorstad dataset only has one group that has encompassed samples from all eight cortical areas at the gene marker level, as detailed in the previous section. For each model, parcels were randomly shuffled and split into 1,000 distinct train (70%) and test (30%) sets without replacement. Given that the number of parcels within each functional network is not balanced, the train-test split was stratified within each network category. Nested threefold cross-validation was implemented to select and validate the hyperparameters in the training set. Kernels and regularization parameters were first selected and tuned in the inner twofold cross-validation; the models' performance was then evaluated in the outer threefold cross-validation, and the final model was the parameter combination with the highest score. The regularization parameter was set to adjust weights inversely proportional to class frequencies in the training data to control for the unbalanced class size within each split. The *F*1 score was used to evaluate the models' overall performance. Recall, which indexes the probability of correctly classifying a parcel given it is from a certain class, was used to evaluate the predicting performance within each class. Accuracy for each parcel was calculated by the total number of times it was correctly classified over the total number of times it was included in the test set. These metrics were evaluated from the 1,000 test sets.

The predictive metrics for every model were evaluated against models fitted from permuted network labels[95,96]. A Hungarian version of the spin test was applied, which uniquely reassigns each parcel's network label for every rotation that controls spatial autocorrelation[94,97]. Each permutation was used to train and test a null model using a randomly selected hyperparameter combination from the set of 1,000 optimal hyperparameter combinations for the original model[98]. Prediction performance metrics from each of the original model's 1,000 train-test splits were then compared to the median prediction accuracy from the null distribution. Consistent with previous works[95,96], the *P* value for each metric's significance is defined as the proportion of 1,000 original models with a performance score less than or equal to the median performance of the null model. Performance metrics were considered to be significant if they performed better than the median null performance for more than 950 of the 1,000 original models.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

A summary of data availability for each dataset used in this study is described below. FC matrices averaged from 820 participants in HCP dataset coregistered via MSMAll were downloaded from ConnectomeDB (https://www.humanconnectome.org/storage/app/media/documentation/s900/820_Group-average_rfMRI_Connectivity_December2015.pdf). Schaefer parcellation was downloaded from Computational Brain Imaging Group (CBIG) GitHub (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal). Human microarray gene expression data obtained from bulk samples of six postmortem brains were downloaded from the AHBA dataset (http://human.brain-map.org/). snDrop-seq data obtained from refs. 38,43 were downloaded from cellxgene-census (https://cellxgene.cziscience.com/collections/d17249d2-0e6e-4500-abb8-e6c93fa1ac6f) and Gene Expression Omnibus website (GSE97930, https://www.ncbi.nlm.nih.gov/geo), respectively.

Code availability

Code is accessible via GitHub—(1) analysis-ready cell-type abundance maps and cell-type analysis code (https://github.com/XihanZhang/human-cellular-func-con), (2) FC gradient analysis (https://neuroanatomyandconnectivity.github.io/gradient_analysis/), (3) spin test (https://github.com/spin-test/spin-test), (4) PermCCA (https://github.com/andersonwinkler/PermCCA) and (5) abagen toolbox for AHBA preprocessing (https://abagen.readthedocs.io/en/stable).

Brain visualization tools were as follows: neuromaps v0.0.5 (https://pypi.org/project/neuromaps/), surfplot (https://surfplot.readthedocs.io/), brainspace v0.1.10 (https://brainspace.readthedocs.io/), enigma toolbox (https://enigma-toolbox.readthedocs.io/) and netneurotools v0.2.4 (https://netneurotools.readthedocs.io/).

Cell-type imputation from mixtures via CIBERSORTx (https://cibersortx.stanford.edu/).

Cell types mapping between single-cell datasets via Seurat v5 (https://satijalab.org/seurat/).

References

78. Marcus, D. et al. Informatics and data mining tools and strategies for the Human Connectome Project. *Front. Neuroinform.* **5**, 4 (2011).

79. Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J. & Coalson, T. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb. Cortex* **22**, 2241–2262 (2012).
80. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
81. Fischl, B. FreeSurfer. *NeuroImage* **62**, 774–781 (2012).
82. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* **17**, 825–841 (2002).
83. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).
84. Holmes, A. J. et al. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Sci. Data* **2**, 150031 (2015).
85. Arnatkevičiūtė, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage* **189**, 353–367 (2019).
86. Markello, R. D. et al. Standardizing workflows in imaging transcriptomics with the abagen toolbox. *eLife* **10**, e72129 (2021).
87. Arnatkeviciute, A., Markello, R. D., Fulcher, B. D., Misic, B. & Fornito, A. Toward best practices for imaging transcriptomics of the human brain. *Biol. Psychiatry* **93**, 391–404 (2023).
88. Romero-Garcia, R. et al. Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. *NeuroImage* **171**, 256–267 (2018).
89. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
90. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
91. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
92. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
93. Cornblath, E. J. et al. Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Commun. Biol.* **3**, 261 (2020).
94. Markello, R. D. & Misic, B. Comparing spatial null models for brain maps. *NeuroImage* **236**, 118052 (2021).
95. Parkes, L. et al. Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Transl. Psychiatry* **11**, 232 (2021).
96. Dhamala, E. et al. Proportional intracranial volume correction differentially biases behavioral predictions across neuroanatomical features, sexes, and development. *NeuroImage* **260**, 119485 (2022).
97. Kuhn, H. W. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* **2**, 83–97 (1955).
98. Dhamala, E. et al. Brain-based predictions of psychiatric illness-linked behaviors across the sexes. *Biol. Psychiatry* **94**, 479–491 (2023).

Acknowledgements

This work was supported by the National Institute of Mental Health (grants R01MH120080 and R01MH123245 to A.J.H.). This work was also supported by the following awards to E.D.: the Northwell Health/Feinstein Institutes for Medical Research Advancing Women in Science and Medicine Career Development Award and the Feinstein Institutes for Medical Research Barbara Zucker Emerging Scientist Award. Analyses were made possible by the high-performance computing facilities provided through the Yale Center for Research Computing. This work used data from the Allen Institute for Brain Science. Additional data were provided, in part, by the HCP, WU-Minn Consortium (principal investigators: D.V. Essen and K. Ugurbil; 1U54MH091657), funded by the 16 National Institutes of Health (NIH) Institutes and Centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. This manuscript reflects the views of the authors and may not reflect the opinions or views of the Allen Institute, NIH or the HCP consortia investigators.

Author contributions

X.-H.Z. and A.J.H. designed the research, made figures and wrote the paper, which all authors commented on and edited. X.-H.Z. conducted the research, analyzed the data and published the code. X.H.Z., K.M.A., H.-M.D., S.C., E.D., P.S.E., M.B.G., D.M. and A.J.H. analyzed and interpreted the results. All authors provided analytic support.

Competing interests

K.M.A. is an employee of Neumora Therapeutics. The remaining authors declare no competing interests.

Additional information

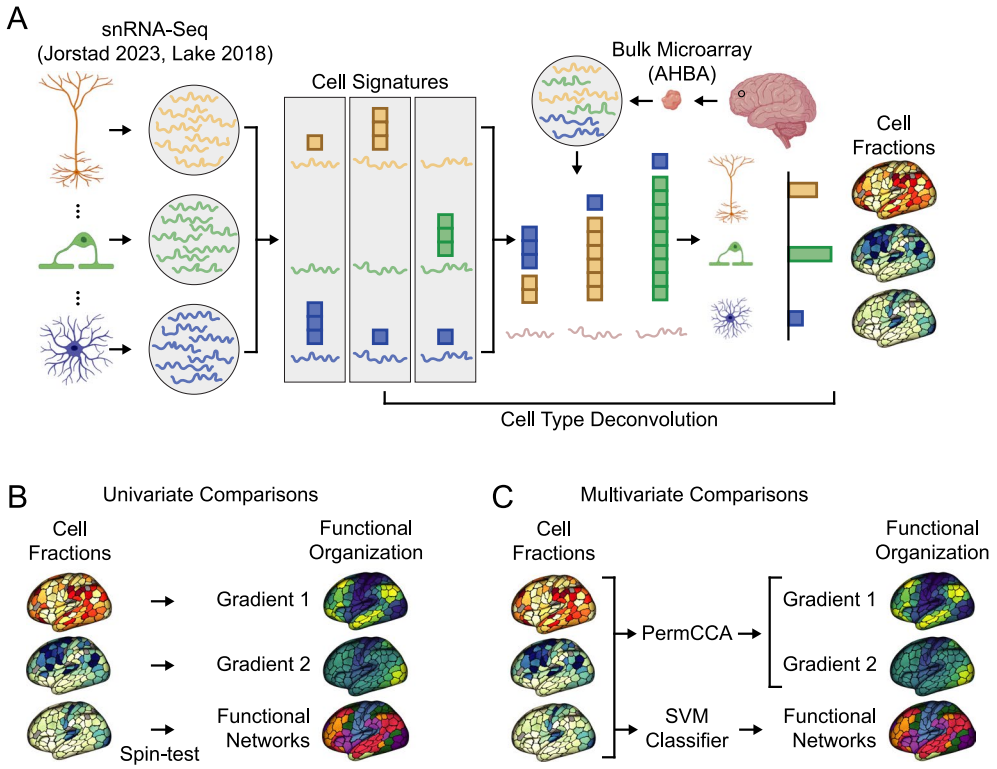
Extended data is available for this paper at https://doi.org/10.1038/s41593-024-01812-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-024-01812-2.

Correspondence and requests for materials should be addressed to Xi-Han Zhang or Avram J. Holmes.

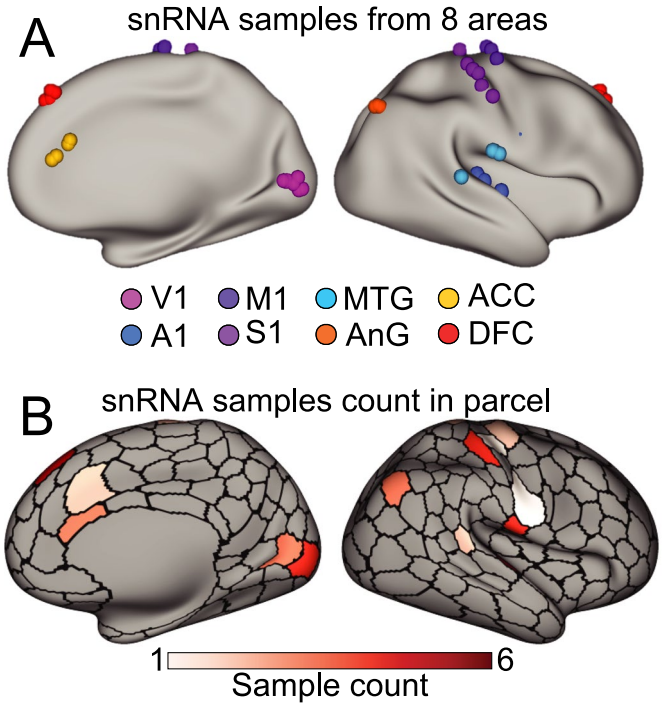
Peer review information *Nature Neuroscience* thanks Michael Hawrylycz and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41593-024-01812-2



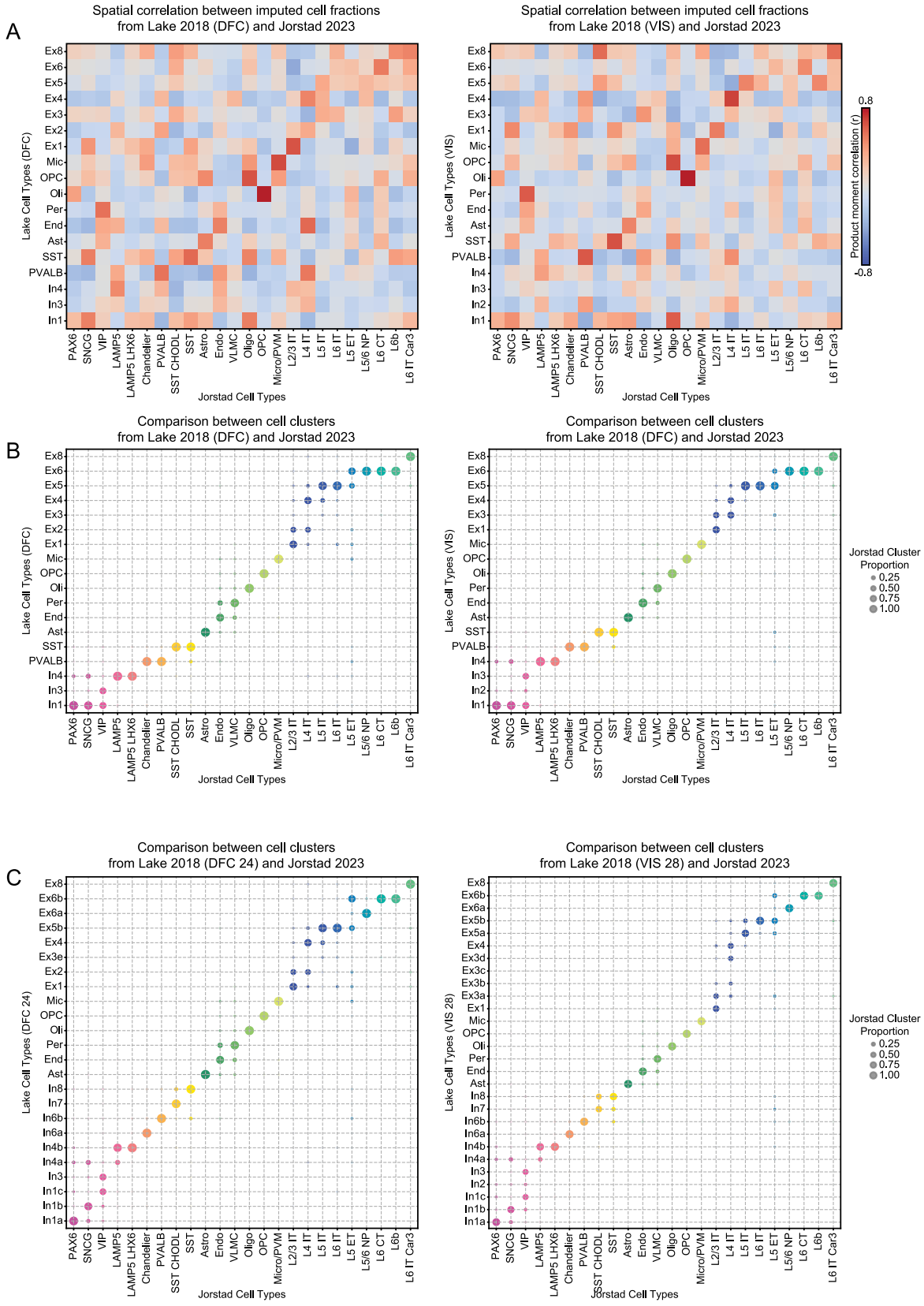
Extended Data Fig. 1 | Workflow for the analysis pipeline. a, Procedures of cell-type deconvolution via CIBERSORTx. Jorstad dataset and Lake dataset[38,42,43] are used as two technical replications for the main analyses. Within each dataset, transcriptomes of single nuclei with labeled cell types were compared to select genes that were differentially expressed across the cell types to constitute an expression signature matrix. Then the signature matrix was used to impute fractions of each cell type within the total transcriptions of each AHBA bulk sample estimated by microarray. **b**, The univariate relationship between cell-type abundance and functional architectures was calculated separately for gradients and networks. (1) The imputed fractions for each cell type are compared with the two gradients across cortical parcels. The resulting empirical correlation is compared against the correlations between cell-type fractions and null gradients permuted by spin test to construct the p-values that control for spatial autocorrelation. (2) The imputed cell-type fractions were averaged within each functional network to calculate the empirical network-level fractions for each cell type. Cell-type fractions were permuted via spin test and then averaged within the same functional atlas for each network to constitute the null distributions. Cell-type enrichment was calculated from the comparisons between empirical cell-type fractions and the nulls. **c**, (1) Multivariate relationships between cell types and gradients were examined through CCA where the null distribution was constructed from the spin test described in **b**. (2) The fractions of all cell types within each parcel were used to predict its corresponding functional network via SVM.

https://doi.org/10.1038/s41593-024-01812-2



Extended Data Fig. 2 | Jorstad snDrop-seq sample distribution. a, Locations of the 44 Jorstad snRNA samples[38] from 8 cortical areas: primary visual cortex (V1), primary auditory cortex (A1), primary motor cortex (M1), primary sensory cortex (S1), middle temporal gyrus (MTG), angular gyrus (AnG), dorsolateral prefrontal cortex (DFC). **b**, Counts of samples falling within each parcel from the 400-parcel functional atlas of ref. 32. Parcels that are not covered by snRNA samples are colored in gray.

https://doi.org/10.1038/s41593-024-01812-2

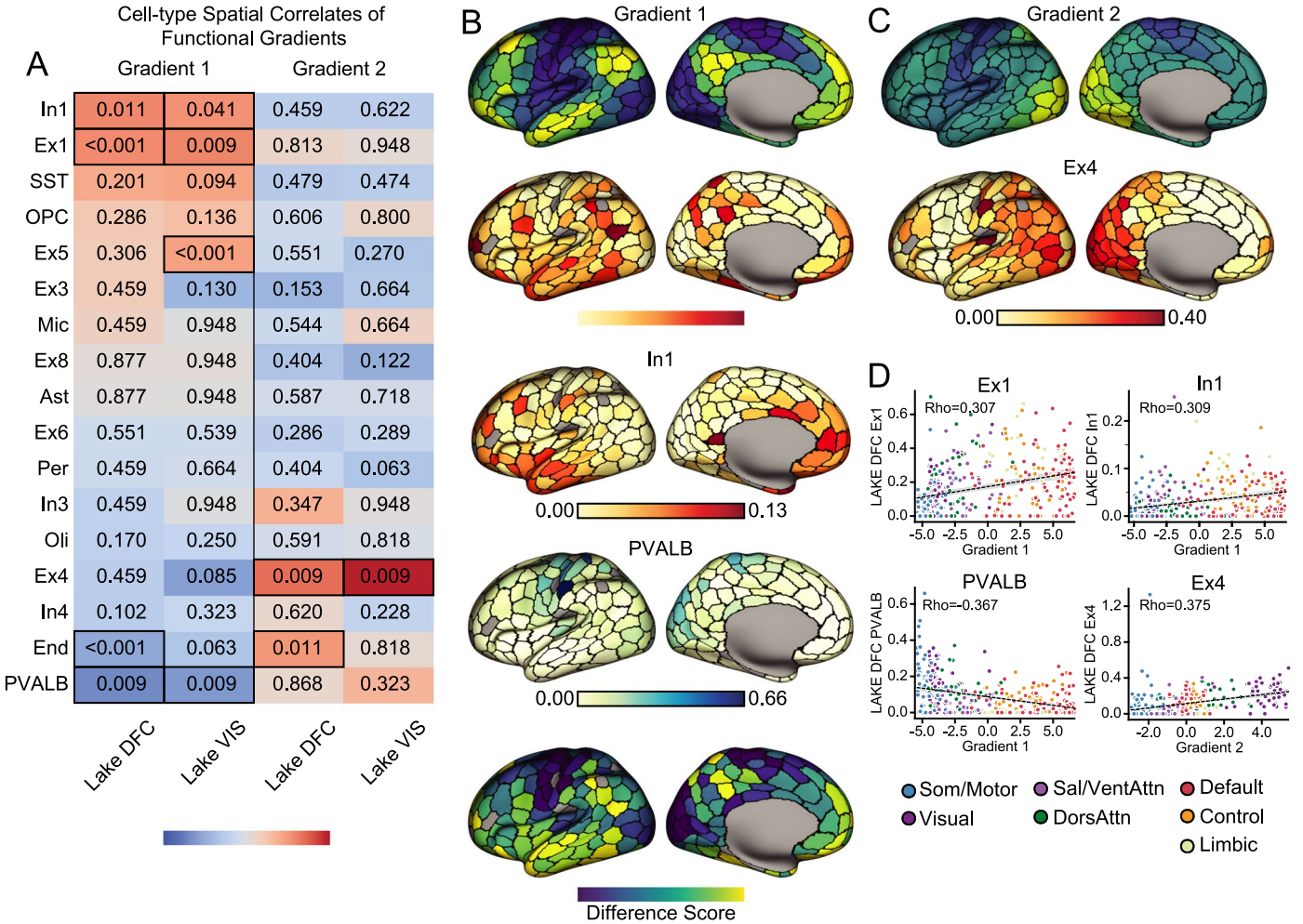


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41593-024-01812-2

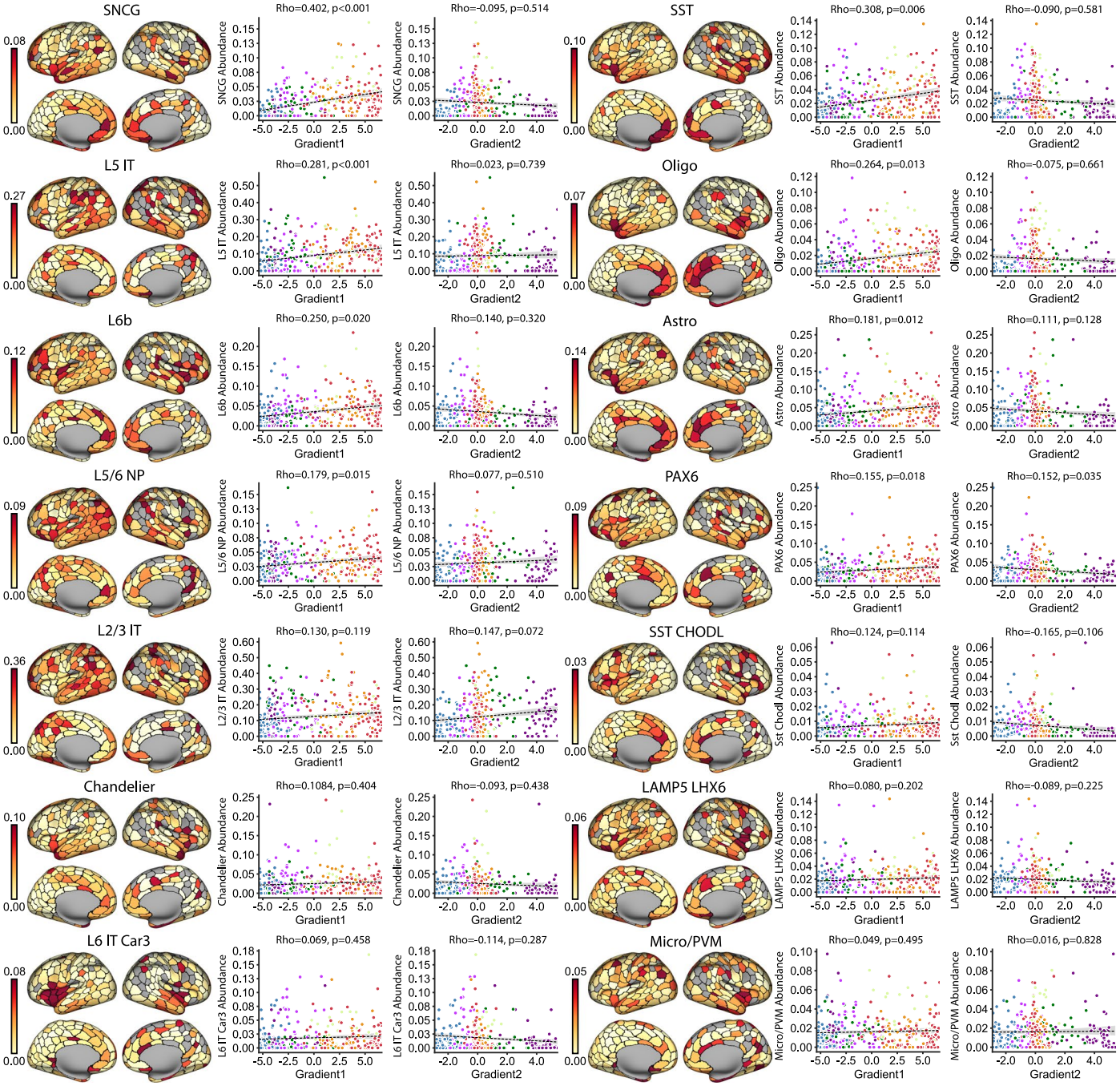
Extended Data Fig. 3 | Cell-type comparisons between Jorstad and Lake snDrop-seq datasets. a, The cell fractions imputed from Lake[43] and Jorstad[38] samples are compared across cortical parcels. Left, comparison between Jorstad 24 cell types derived from 8 cortical regions and Lake 18 cell types derived from prefrontal cortex (DFC). Right, comparison between Jorstad cell types and Lake 18 cell types derived from the visual cortex (VIS). **b** and **c** are the transcriptional identity comparisons between Jorstad and Lake cell types. Dot plots showing the proportion of each Jorstad nuclei cluster (cell-type) that matches Lake clusters based on the gene marker unique molecular identifier (UMI) counts for each nucleus. Lake clusters from DFC are on the left and VIS are on the right. Within each matrix, GABAergic inhibitory interneurons are in warm colors in the bottom left corner. glutamatergic excitatory neurons are in cold colors in the top right corner. Non-neuronal cells in the middle are colored in green. **b**, The sub-clusters of the same cluster in Lake data were combined into one (for example, In1a-c is combined as In1, In6a-b are combined as PVALB and In7-8 are combined as SST). The majority of the Lake interneuron clusters can be broken down into two to three clusters in Jorstad. Most non-neuronal clusters between the two datasets have one-to-one mapping. L6 IT Car3 and Ex8 are the only one-to-one matched excitatory pairs between Jorstad and Lake. Deeper-layer excitatory clusters (Ex5 and Ex6) in Lake can be broken down into multiple deeper-layer clusters in Jorstad (L5 IT–L6b). Shallower layer excitatory neuron clusters between the two datasets have many-to-many mapping. **c**, When considering sub-clusters in Lake, additional one-to-one mapping in the GABAergic clusters was evident: In1a and PAX6, In1b and SNCG, In6a and chandelier, In6b and PVALB. An additional glutamatergic one-to-one pair is Ex6a and L5/L6 NP.

https://doi.org/10.1038/s41593-024-01812-2



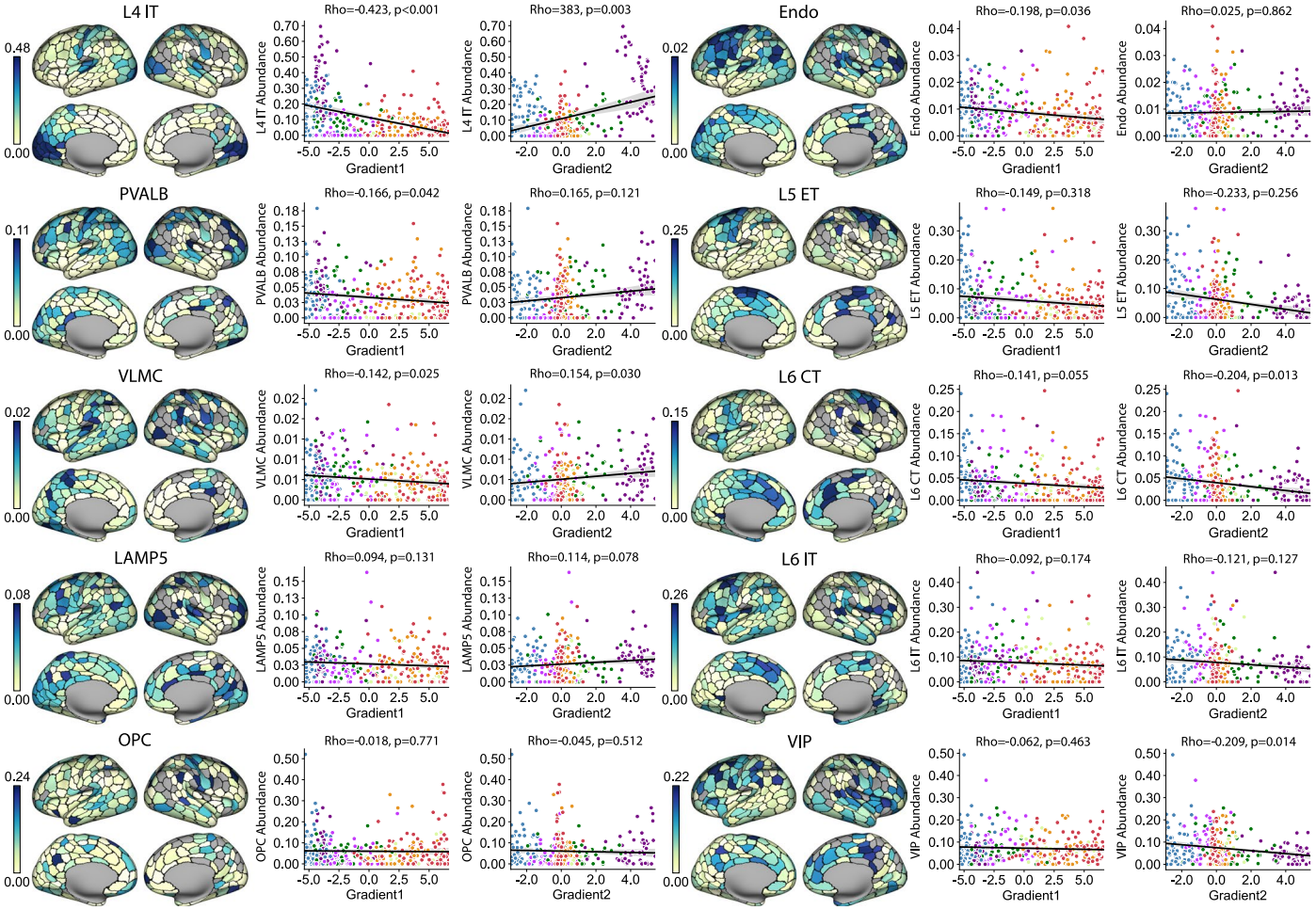
Extended Data Fig. 4 | Univariate cell-type distributions (imputed from Lake snRNA-seq[43]) align with functional gradient topographies. a, Cell types are imputed from gene expression in AHBA bulk-tissue samples. Single-cell signatures are constructed from independent tissue samples in the frontal (Lake DFC) and visual (Lake VIS) lobes, allowing for a technical replication of the cell-type imputation scheme. Resulting abundance of cell types was rank-ordered by spatial correlation to each principal functional gradient. Warm colors indicate positive correlation, and numbers in each cell reflect associated p-value calculated from a two-sided spin test that controls for partial autocorrelation and were FDR-corrected for 17 comparisons. Correlations with $p < 0.05$ are denoted by black outlined boxes. Surviving from significance tests in both Lake DFC and Lake VIS, the spatial pattern of gradient 1 was correlated with two interneuron subtypes: In1 and PVALB, as well as one excitatory neuron subtype: Ex1. Gradient 2 was significantly correlated to Ex4 excitatory neurons. **b**, Imputed cell-type abundance distributed across cortex suggests Ex1 and In1 are preferably distributed around the transmodal end of gradient 1 (bright yellow), whereas PVALB is preferably distributed around the unimodal end (dark blue). Scale bars reflect estimated cell abundances. A difference score between Ex1 and PVALB distributions generates a pattern spatially consistent with the first functional gradient. **c**, Ex4 follows a spatial pattern aligning to the second gradient, peaking in visual pole (bright yellow) then gradually decreasing as it approaches the somato/motor and auditory cortices (dark blue). Parcels that are excluded from analyses, and not covered by AHBA bulk samples, are colored in gray. **d**, Scatter plots with each cortical parcel colored by the corresponding functional networks show that cell-type abundance gradually increases/decreases across the networks distributed along the gradients, with enrichment/absence evident within certain networks. Correlations were estimated by Spearman's rho (as reflected in the scale bar in **a**), but for visual reference, the fitted linear regression lines (95% CI) predicting cellular abundance from gradient values are also displayed.

https://doi.org/10.1038/s41593-024-01812-2



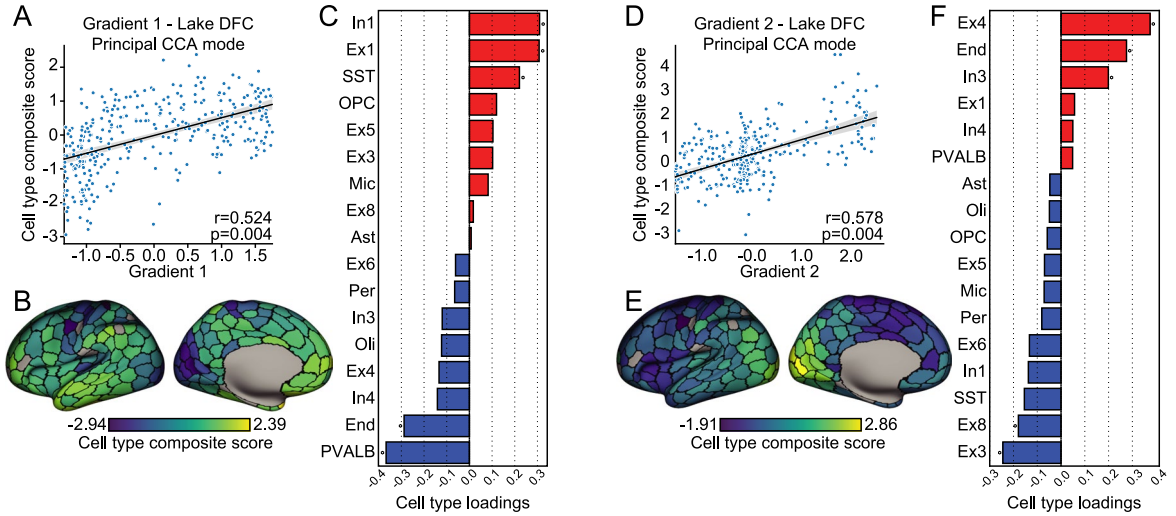
Extended Data Fig. 5 | Spatial distribution of cell types (imputed from Jorstad snDrop-seq[38]) preferentially distributed in association cortex. The imputed cell-type abundances across cortex in both hemispheres are aggregated in 400 Schaffer parcels. The scatter plot on the left shows the correlation between cell abundance and the first principle gradient (gradient 1) across cortical parcels, and the correlation with the second gradient (gradient 2) is displayed on the right. Correlations were estimated by Spearman's rho, but for visual reference, the fitted linear regression lines (95% CI) predicting cellular abundance from gradient values are also displayed. P-values were constructed from two-sided spin tests that control spatial autocorrelation. Dots are color-coded by the functional network allegiance of each parcel. A positive correlation with gradient 1 indicates the cell is preferentially spaced on the transmodal association cortex. Such distribution preference reflected on the scatter plot of gradient 2 is a peak near 0. Although only SNCG, SST and L5 IT survived the multiple corrections, the patterns are visually identifiable in the rest of the top five cells (Oligo, L6b).

https://doi.org/10.1038/s41593-024-01812-2



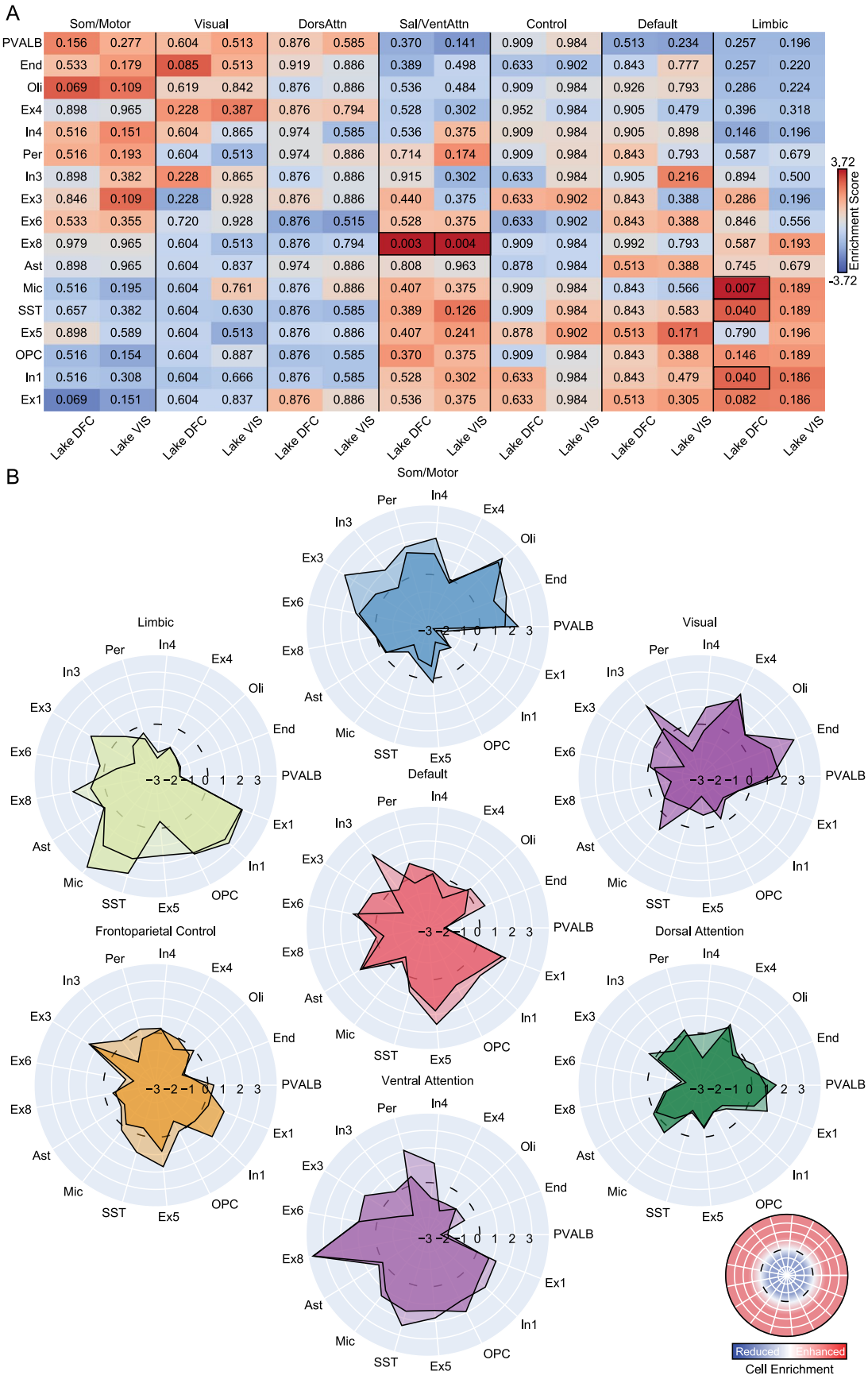
Extended Data Fig. 6 | Spatial distribution of cell types (imputed from Jorstad snRNA-seq[38]) preferentially distributed in unimodal areas. The imputed cell-type abundances across cortex in both hemispheres are aggregated in 400 Schaffer parcels. The scatter plot on the left shows the correlation between cell abundance and the first principle gradient (gradient 1) across cortical parcels, and the correlation with the second gradient (gradient 2) is displayed on the right. Correlations were estimated by Spearman's rho, but for visual reference, the fitted linear regression lines (95% CI) predicting cellular abundance from gradient values are also displayed. P-values were constructed from two-sided spin tests that control spatial autocorrelation. Dots are color-coded by the functional network allegiance of each parcel. A negative correlation with gradient 1 indicates the cell is preferentially spaced on the unimodal areas. Although only L4 IT survived the multiple corrections, the patterns are visually identifiable in Endo and PVALB. Their matched cell types in Lake are End and PVALB/In6b, both of which survived multiple corrections for the negative correlation with gradient 1 (Supplementary Fig. 4; end in Lake DFC: rho = −0.288, $p_{FDR} < 0.001$; PVALB in Lake DFC: rho = −0.367, $p_{FDR} = 0.009$; PVALB in Lake VIS: rho = −0.325, $p_{FDR} = 0.009$).

https://doi.org/10.1038/s41593-024-01812-2



Extended Data Fig. 7 | Multivariate cellular (imputed from Lake DFC[43]) profiles follow the macroscale organization of cortex. a, The scatter plot displays results of PermCCA where the first functional gradient was positively associated with a composite score of cell-type abundance (95% CI). **b**, Cell-type composite score associated with the first functional gradient projected to the cortical surface. **c**, Canonical loadings of each cell type to the composite score implicate In1, Ex1, PVALB and the SST interneurons and endothelial cells (End; red indicates positive associations; blue, negative associations). **d**, The second functional gradient was positively associated with a cell-type composite score of cell-type abundance (95% CI). **e**, Cell-type composite score associated with the second functional gradient mapped to the cortical surface. **f**, Canonical loadings of each cell type to the composite score significantly implicated Ex4, Ex3 (marked by gene NEFM) and Ex8 (MCTP2, NR4A2) excitatory neurons, and In3 (TSHZ2, SHISA8) interneuron and End. P-values in **a** and **d** are calculated from two-sided spin tests after correcting the two comparisons.

https://doi.org/10.1038/s41593-024-01812-2

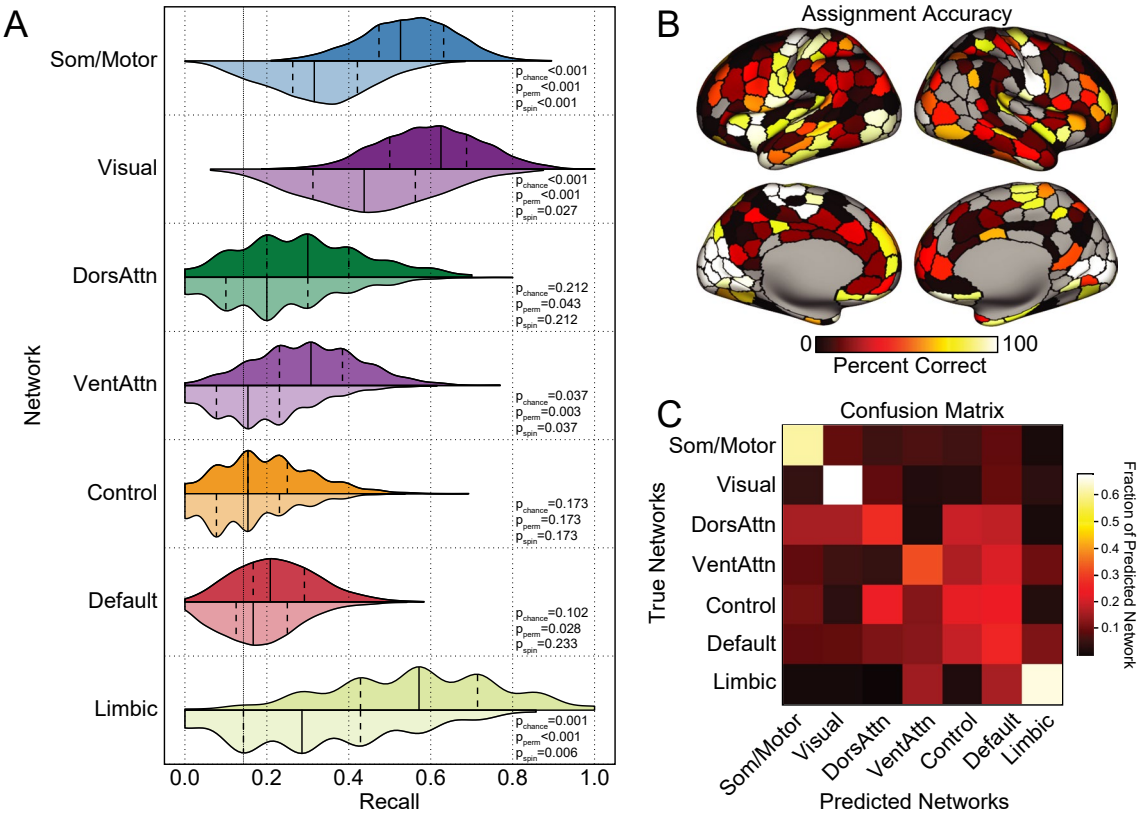


Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41593-024-01812-2

Extended Data Fig. 8 | Large-scale functional networks demonstrate distinct cellular profiles (imputed from Lake snRNA-seq[43]). a, The table displays the relative cell-type enrichment, or absence, within each canonical functional network. Networks are ordered as their estimated position along the first principal gradient. Empirical abundance for each cell type was aggregated within each of the 7 large-scale functional networks. Corresponding null distributions were constructed from parcel-level spin test, which accounts for spatial autocorrelation. Table fill colors reflect z-scores, derived via subtracting the mean of the null from the observed empirical abundance and then dividing the difference by the standard deviation of the null distribution. Here z-scores index empirical enrichment relative to the null. Warm colors indicate positive values (relatively enriched), and cold colors indicate negative values (relatively absent). Numbers in each cell represent the associated p-value calculated from two-sided spin tests that control for spatial autocorrelation and were corrected for 17 comparisons. Reflecting the presence of a cell-type enrichment gradient spanning between somato/motor and limbic networks, each network shows a unique cell-type profile. Marked boxes reflect significant enrichment ($p < 0.05$). **b**, Polar plots of z-score across 17 cell types for each network suggest the potential of cellular profiles that may serve as fingerprints that can distinguish each functional network. Score above zero lines (dashed) indicates when a cell type is enriched within a given network relative to the overall distribution across cortical parcels, whereas score below zero reflects the relative absence of a cell type. Polar plot corresponds to imputed cell densities from Lake DFC and Lake VIS stacked together, with the overlapping area in darker color.

https://doi.org/10.1038/s41593-024-01812-2



Extended Data Fig. 9 | Large-scale functional network assignment can be predicted by cell-type abundance (imputed from Lake snRNA-seq[43]) in post-mortem tissue. a, Histograms display the SVM recall or the probability of correctly classifying a parcel to the associated network. These data suggest that the classifiers were able to predict somato/motor, visual, ventral attention and limbic networks significantly above chance. Distributions in darker color were constructed from 1000 classifiers trained on real network labels, and the lighter-colored distribution represents classifiers trained on network labels shuffled by spin test that controls for spatial autocorrelation. The solid lines indicate median, and the dashed lines represent quartiles of the distribution. P-value of empirical recall was constructed from two-sided tests of increasing stringency: (1) theoretical chance (1/7, p_{chance}); (2) models trained from randomly permuted network labels (p_{perm}); or (3) labels shuffled by spin test (p_{spin}). **b**, Accuracy of network assignment across cortical parcels, calculated from all testing sets. **c**, Each row of the confusion matrix represents the fraction of parcels within the specific network that were predicted as belonging to each of the 7 networks. The diagonal represents the percentage of correctly classified parcels within each network. Here the confusion matrix suggests a preferentially distinct cellular profile for somato/motor, visual and limbic networks. While classification accuracies were low for the remaining association cortex networks, dorsal attention, default and control networks display a higher rate of misclassifications among each other relative to unimodal networks.

https://doi.org/10.1038/s41593-024-01812-2

Extended Data Table 1 | Cell-type annotations from the Jorstad dataset

Cell-type	Description	Neighborhood	Class
LAMP5	LAMP5-expressing interneuron	CGE	GABAergic
LAMP5 LHX6	LAMP5- and LHX6-expressing interneuron	CGE	GABAergic
PAX6	PAX6-expressing interneuron	CGE	GABAergic
SNCG	SNCG-like interneuron	CGE	GABAergic
VIP	VIP-expressing interneuron	CGE	GABAergic
Chandelier	Chandelier parvalbumin-expressing interneuron	MGE	GABAergic
PVALB	Parvalbumin-expressing interneuron	MGE	GABAergic
SST	Somatostatin-expressing interneuron	MGE	GABAergic
SST CHODL	Somatostatin- and Chodl-expressing interneuron	MGE	GABAergic
L2/3 IT	Layer 2/3 intratelencephalic-projecting	IT	Glutamatergic
L4 IT	Layer 4 intratelencephalic-projecting	IT	Glutamatergic
L5 IT	Layer 5 intratelencephalic-projecting	IT	Glutamatergic
L6 IT	Layer 6 intratelencephalic-projecting	IT	Glutamatergic
L6 IT Car3	Layer 6 intratelencephalic-projecting (Car3-like)	IT	Glutamatergic
L5 ET	Layer 5 extratelencephalic-projecting	Non-IT	Glutamatergic
L5/6 NP	Layer 5/6 near-projecting	Non-IT	Glutamatergic
L6 CT	Layer 6 corticothalamic-projecting	Non-IT	Glutamatergic
L6b	Layer 6b	Non-IT	Glutamatergic
Astro	Astrocyte	NN	Non-neuronal
Endo	Endothelial cell	NN	Non-neuronal
Micro/PVM	Microglia/Perivascular macrophage	NN	Non-neuronal
Oligo	Oligodendrocyte	NN	Non-neuronal
OPC	Oligodendrocyte precursor cell	NN	Non-neuronal
VLMC	Vascular leptomeningeal cell	NN	Non-neuronal

Transcriptional taxonomy of nuclei revealed hierarchical structures, where 'class' refers to three higher level branches as follows: GABAergic inhibitory interneurons, glutamatergic excitatory neurons and non-neuronal cells. Within GABAergic class, the interneuron clusters were grouped into two major 'neighborhoods' that reflect the development origins—caudal ganglionic eminence (CGE) and medial ganglionic eminence (MGE). The two major neighborhoods in glutamatergic excitatory neurons demonstrated different projection targets—intratelencephalic-projecting (IT) vs non-IT. The 24 clusters grouped into these neighborhoods are the cell types considered in the present study. Broadly, GABAergic cell types reflect the signature genes that typically play key roles in cell maturation and migration during development. The glutamatergic cell types mainly reflect the layer-specificity and projection targets. The non-neuronal cell types capture the major glial cells and others supporting distinct functions. These annotations were originally produced from the patch-seq[35] and many other multimodal characterization techniques for single cells collected together with the snRNA-seq in mice and then mapped to human samples based on transcriptional homology and validated through the multimodal cellular characterization[36].

nature portfolio

Corresponding author(s): Avram J. Holmes & Xi-Han Zhang

Last updated by author(s): Sep 11, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software were used for data collection
Data analysis	Custom code for our analysis is accessible via GitHub (https://github.com/XihanZhang/human-cellular-func-con) MATLAB R2021a scripts were used for the following major analysis: Functional gradient calculation code downloaded from Margulies et al., 2016 (https://neuroanatomyandconnectivity.github.io/gradient_analysis/) Spin-test permutating gradients via code downloaded from Alexander-Bloch et al., 2018 (https://github.com/spin-test/spin-test) Permutational canonical correlation analysis via code downloaded from Winkler et al., 2020 (https://github.com/andersonwinkler/PermCCA) Vertex-parcel convertion via ENIGMA-1.1.3 (https://github.com/MICA-MNI/ENIGMA) Python v3.8.8 scripts were used for the following major analysis: AHBA sample preprocessing via abagen v0.1.3 (https://abagen.readthedocs.io/en/stable) MNI to vertex mapping for the AHBA samples via FreeSurfer v6.0.0 (https://surfer.nmr.mgh.harvard.edu/) Cortical parcellation via code downloaded from Schaefer et al., 2018 (https://github.com/ThomasYeoLab/CBIG) Brain visualization via neuromaps v0.0.5 (https://pypi.org/project/neuromaps/), surfplot (https://surfplot.readthedocs.io/), brainspace v0.1.10 (https://brainspace.readthedocs.io/), enigma toolbox (https://enigma-toolbox.readthedocs.io/), netneurotools v0.2.4 (https://netneurotools.readthedocs.io/) SVM classification via sklearn (https://scikit-learn.org/stable/modules/svm.html#svm-classification) Cell type matching between Lake and Jorstad datasets via Seurat v5 (https://satijalab.org/seurat/) in R v4.1.1

Cell type imputation from mixtures via CIBERSORTx (https://cibersortx.stanford.edu/)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Functional connectivity matrices averaged from 820 subjects in HCP dataset coregistered via MSMAll were downloaded from ConnectomeDB (https://www.humanconnectome.org/storage/app/media/documentation/s900/820_Group-average_rfMRI_Connectivity_December2015.pdf).
Schaefer parcellation were downloaded from CBIG github (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal).
Human microarray gene expression data obtained from bulk samples of 6 postmortem brains were downloaded from AHBA dataset (http://human.brain-map.org/).
Single-nucleus droplet-based sequencing (snDrop-seq) data obtained by Jorstad et al. were downloaded from cellxgene-census (https://cellxgene.cziscience.com/collections/d17249d2-0e6e-4500-abb8-e6c93fa1ac6f)
snDrop-seq data from Lake et al. were downloaded from Gene Expression Omnibus website ("GSE97930" [https://www.ncbi.nlm.nih.gov/geo]).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	AHBA include 6 donors, 5 were male and 1 was female (in sex). snRNA-seq data from Jorstad et al. has 4 males and 2 females. HCP 820R include 366 male and 454 female (in gender).
Reporting on race, ethnicity, or other socially relevant groupings	Grouping was not relevant to our analyses because they were always conducted on the full set of available data.
Population characteristics	AHBA mean age 42.5 years, range 24-57 years. snRNA-seq from Jorstad et al. mean age 45.7, range 29-60 HCP 820R mean age 28.8 years, range 22-37 years.
Recruitment	Anonymized open access data were used for the present analysis. Recruitment details of AHBA donors can be seen at: Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489, 391–399. 10.1038/nature11405. For snRNA-seq provided by Jorstad et al.: Jorstad, N. L., Close, J., Johansen, N., Yanny, A. M., Barkan, E. R., Travaglini, K. J., ... & Lein, E. S. (2023). Transcriptomic cytoarchitecture reveals principles of human neocortex organization. Science, 382(6667), eadf6812. For HCP 820R: Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., and Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. NeuroImage 80, 62–79. 10.1016/j.neuroimage.2013.05.041.
Ethics oversight	Not applicable to the current study, where anonymized open access data were used.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Our analyses utilized open-access genetic and neuroimaging consortia data, so no sample size calculations were performed. Rather, we used all the data provided (all the 6 donors from AHBA and rest-state functional connectivity matrices of all the 820 subjects from HCP 820R).

Data exclusions	No participants were excluded from the present analysis
Replication	The main analysis was perform based on the cell types imputed from Jorstad dataset and replicated with the cell types imputed from Lake dataset in the current study. Though the cell clusters defining the cell types from the two datasets were not perfectly one-on-one mapping, the major cell classes, such as PVALB, SST, IN1 in Lake and SNCG in Jorstad etc (as discussed with more details in maintext) shared similar univariate alignment with functional gradients. The effect size and robustness of multivariate relationship between cell types and gradients are comparable between the two datasets. The cell types enrichment across seven functional networks between the two datasets shared greater taxonomy class. The predicting accuracy on functional networks were similar between the cell types imputed from the two datasets.
Randomization	Randomization was not relevant to our analyses because they were always conducted on the full set of available data and did not include "case/control" experimental design.
Blinding	Blinding was not relevant since analyses were conducted on publicly available archival data and did not use an experimental/control group design.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

Magnetic resonance imaging

Experimental design

Design type	resting state functional magnetic resonance imaging
Design specifications	Four 15-minute rest-run per subject
Behavioral performance measures	We analyzed resting-state imaging data, which does not have an associated behavioral metric.

Acquisition

Imaging type(s)	resting state functional magnetic resonance imaging
Field strength	3T
Sequence & imaging parameters	TR=720ms,TE=33.1ms, spatial resolution=2*2*2mm^3, more details can be found in the S900 release manual available at db.humanconnectome.org
Area of acquisition	whole-brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Functional connectivity matrices averaged from 820 subjects in HCP dataset coregistered via MSMAll were downloaded from ConnectomeDB (https://www.humanconnectome.org/storage/app/media/documentation/s900/820_Group-average_rfMRI_Connectivity_December2015.pdf). Gradients, or components with similar functional connectivity patterns, were derived by the published code (https://neuroanatomyandconnectivity.github.io/gradient_analysis/)
Normalization	Functional connectivity matrices averaged from 820 subjects in HCP dataset coregistered via MSMAll were downloaded from ConnectomeDB (https://www.humanconnectome.org/storage/app/media/documentation/s900/820_Group-average_rfMRI_Connectivity_December2015.pdf).
Normalization template	Standard HCP pre-processed data were downloaded for the current analysis.
Noise and artifact removal	Standard HCP pre-processed data were downloaded for the current analysis.
Volume censoring	Standard HCP pre-processed data were downloaded for the current analysis.

Statistical modeling & inference

Model type and settings	spin-test, permutation-based canonical correlation analysis, support vector machines
Effect(s) tested	we tested whether cell type spatial distributions correlated with functional gradients and network topographies.

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

Anatomical location(s)	ROIs were defined using the Schaefer 400 cortical functional atlas. Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., and Yeo, B.T.T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb. Cortex 28, 3095–3114.
Statistic type for inference (See Eklund et al. 2016)	analysis were conducted at network and individual parcel level
Correction	where relevant, statistical tests were tested for multiple comparisons.

Models & analysis

n/a	Involved in the study
☐	☒ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity	Spearman's Rho between cell type abundance and functional gradient value across ROIs parceled in Schaefer 400.
Multivariate modeling and predictive analysis	Permutational based canonical correlation analysis was used to calculate the maximum correlation between each functional gradient and 17 cell type abundances. Support vector machines were trained to predict the functional network each cortical parcel belongs to based on abundances of 17 cell types within that parcel. SVMs were trained from 70% parcels and tested on the left-out 30% parcels. Nested three-fold cross-validation was implemented to select and validate the hyperparameters in the training set. Kernels and regularization parameters were first selected and tuned in the inner 2-fold cross-validation, the models' performance were then evaluated in the outer 3-fold cross-validation, the final model was the parameter combination with the highest score