Three transcriptional axes underpin anatomy, development, and disorders of the human cortex

Richard Dear, Jakob Seidlitz, Ross D. Markello, Aurina Arnatkevičiūtė, Kevin M. Anderson, Richard A.I. Bethlehem, Lifespan Brain Chart Consortium, Konrad Wagstyl, Edward T. Bullmore, Armin Raznahan, Petra E. Vértes

Abstract

The Allen Human Brain Atlas (AHBA) has catalysed many discoveries linking macroscale brain organisation to microscale neurobiology, including the existence of a transcriptional axis stretching from sensorimotor to association cortices. However, the complex spatial structure of the human cortex likely includes multiple gradients of gene expression that have yet to be characterised. Here, we reveal two additional transcriptional axes that are anatomically differentiated and have distinct profiles of metabolic and immune function, cytoarchitecture, and developmental trajectory. Moreover, we validate the three transcriptional axes in external data, show they explain more than half of spatial variation across key anatomical and functional neuroimaging maps, and demonstrate they are dissociably aligned with maps of autism, depression, and schizophrenia from one of the largest case-control imaging datasets. Together, these results reveal how human brain anatomy, development, and disorders are patterned along multiple cortical axes underpinned by the intrinsic structure of gene transcription.

Introduction

Brain function and dysfunction display complex spatial structure at the macroscale, yet the microscale underpinnings of these patterns remain unclear [1]. Landmark progress was made on this question with the release of the Allen Human Brain Atlas (AHBA), in which gene expression was measured at over 3000 locations throughout the human brain [2]. The AHBA has allowed diverse human neuroimaging maps – of anatomy, development, function, and disorders – to be interpreted in terms of their underlying biology, by identifying which genes have spatial expression patterns aligned to a given imaging map [3–19]. This approach has given rise to a new field of "imaging transcriptomics" [20–22].

A different approach for linking macroscale to microscale has been to study transcriptional axes – intrinsic structures of gene expression patterned along spatial gradients. In animal models, transcriptional brain axes have been shown to play important roles in cortical areal differentiation, axonal projection, laminar differentiation, and other neurodevelopmental processes [23,24]. For example, in chick and mice embryos, cell fate is patterned by a grid of multiple intersecting morphogen gradients organised along the anterior-posterior, medial-lateral, and dorsal-ventral axes, each of which are generated by distinct gene regulatory networks [25–28]. In adult mice, single-cell RNA-seq data has revealed multiple transcriptional brain axes [29], and in adult non-human primates, microarray transcriptomes revealed a rostral-caudal transcriptional axis as well as a gradient aligned to cortical layers [30]. In human embryos, recent evidence suggests that the chick and mouse embryonic transcriptional axes are largely conserved [31,32], but the extent to which transcriptional axes also pattern gene expression in adult human brains remains largely unexplored.

One relevant discovery is that the pattern which captures the most spatial variation in the AHBA – the first principal component (PC1) of human cortical transcription – is a rostral-caudal gradient running from sensorimotor to association cortices [2,33,34]. This PC1 gradient closely aligns to the ratio of T1- and T2-weighted MRI signals [34], a measure suggested to index both grey-matter myelination [35–37], as well as cortical specialisation as defined by feedforward-vs-feedback connectivity [34,38,39]. Genes with spatial patterns closely aligned to this transcriptional PC1 include key interneuron markers (*PVALB, CALB)*, synaptic subunits *(GRIN, GABRA)*, and risk genes for neurodegenerative disorders (*APOE*, *SNCA*). Moreover, these PC1 genes are over-represented in the gene annotation sets for both brain function and neuropsychiatric disorders [34]. It is thus clear that there is at least one transcriptional axis in the adult human cortex that likely plays a key role in healthy brain function.

Might more than one transcriptional axis pattern the spatial organisation of the adult human brain? To what extent does the intrinsic structure of gene expression underpin the macroscale complexity of brain function across health and disease? Recent work has proposed that a diverse range of neuroimaging maps which differentiate the sensorimotor and association brain cortices may all be associated with the PC1 transcriptional axis [39]. These include, among others, the principal gradient of functional connectivity [40], the map of brain metabolism as indexed by aerobic glycolysis [41], and the maps of evolutionary and developmental expansion [42]. If such diverse brain features all broadly relate to an underlying transcriptional axis, perhaps their differences are also underpinned by intrinsic structures of gene expression. This hypothesis that the transcriptional structure of the adult human cortex is intrinsically multidimensional is further supported by the multiple axes in both human embryos and adult animals [43]. However, thus far no additional transcriptional axes have been found and characterised in adult human brains.

Here we reveal two new gradients of human cortical gene expression that likely represent additional transcriptional axes beyond the already-described PC1 [34]. We also highlight the processing steps which are necessary to extract these generalisable higher-order patterns from the individually variable transcriptomic data in the AHBA. As with the first axis, the additional two transcriptional axes are not *a-priori* constrained to follow any spatial organisation of the brain (e.g. spatial coordinates, cytoarchitecture, or macroscale anatomy), but rather reveal how anatomical and axial alignment is embedded in the intrinsic structure of transcriptional variation. We then explore how the three axes are biologically and developmentally distinct, and demonstrate that they explain the majority of spatial variation across key anatomical and functional brain maps from neuroimaging. Finally, we show that phenotypes of autism spectrum disorder, major depressive disorder, and schizophrenia have specific associations to these three transcriptional axes that likely underpin the healthy development and function of the human brain.

Results

Revealing three generalisable axes of human cortical gene expression

We explored the intrinsic structure of cortical gene expression by processing the AHBA into a matrix representing the average expression of each gene within each of 180 macroscale brain regions defined by the HCP-MMP atlas [44], then applying dimension reduction to identify spatial patterns that are shared across a large number of genes [45]. First, we applied PCA, a widely applied dimension reduction method which was used by Burt *et al.* to characterise the first transcriptional axis (PC1) [34]. Second, we tested Diffusion Map Embedding (DME), an alternative method that in neuroimaging data has been shown to identify gradients representing key functional brain networks (e.g. the default mode) [40,46]. Both PCA and DME identify sets of spatial patterns representing distinct (i.e. uncorrelated) components of the spatial variation in brain gene expression, with the difference being that PCA (but not DME) also constrains the genetic profiles of the components to be uncorrelated.

A key consideration with seeking higher-order structures in the AHBA is whether patterns observed in only six healthy adult brains generalise to the broader population. Since the AHBA is the only data of its kind, we split the six-brain dataset into two disjoint triplets of three brains (e.g. donors 1, 2, 3 vs donors 4, 5, 6), processed the data independently in each triplet, and correlated the resulting principal components; the median of this correlation over all ten ways of splitting the six brains was taken as a proxy for the generalisability of the components (**Fig 1a**). Applying PCA to the cortical AHBA data processed with recommended parameter settings [47,48], the median triplet correlation was 0.72 for PC1, but only 0.12 and 0.11 for PC2 and PC3 (**Fig 1a**), suggesting that higher-order components computed in this way are not stable between subsets of the six AHBA brains, and therefore are unlikely to generalise.

We hypothesised that true (i.e. generalisable) higher-order gradients of cortical transcription may be obscured by both technical noise and individual variability in the six-brain AHBA. We therefore leveraged two methods to isolate the shared transcriptomic structure of interest. First, we filtered for more spatially consistent genes using the 'differential stability' metric [49] whereby genes are ranked by the mean correlation of their individual spatial expression patterns across pairs of donor brains. Second, we filtered for brain regions with more consistent sampling: when AHBA samples from individual donor brains are matched to regions in the HCP-MMP atlas, 43 of the 180 brain regions contain information from only one or two of the six donors (**Fig S1**). We therefore filtered for brain regions with information from at least 2 or at least 3 donors.

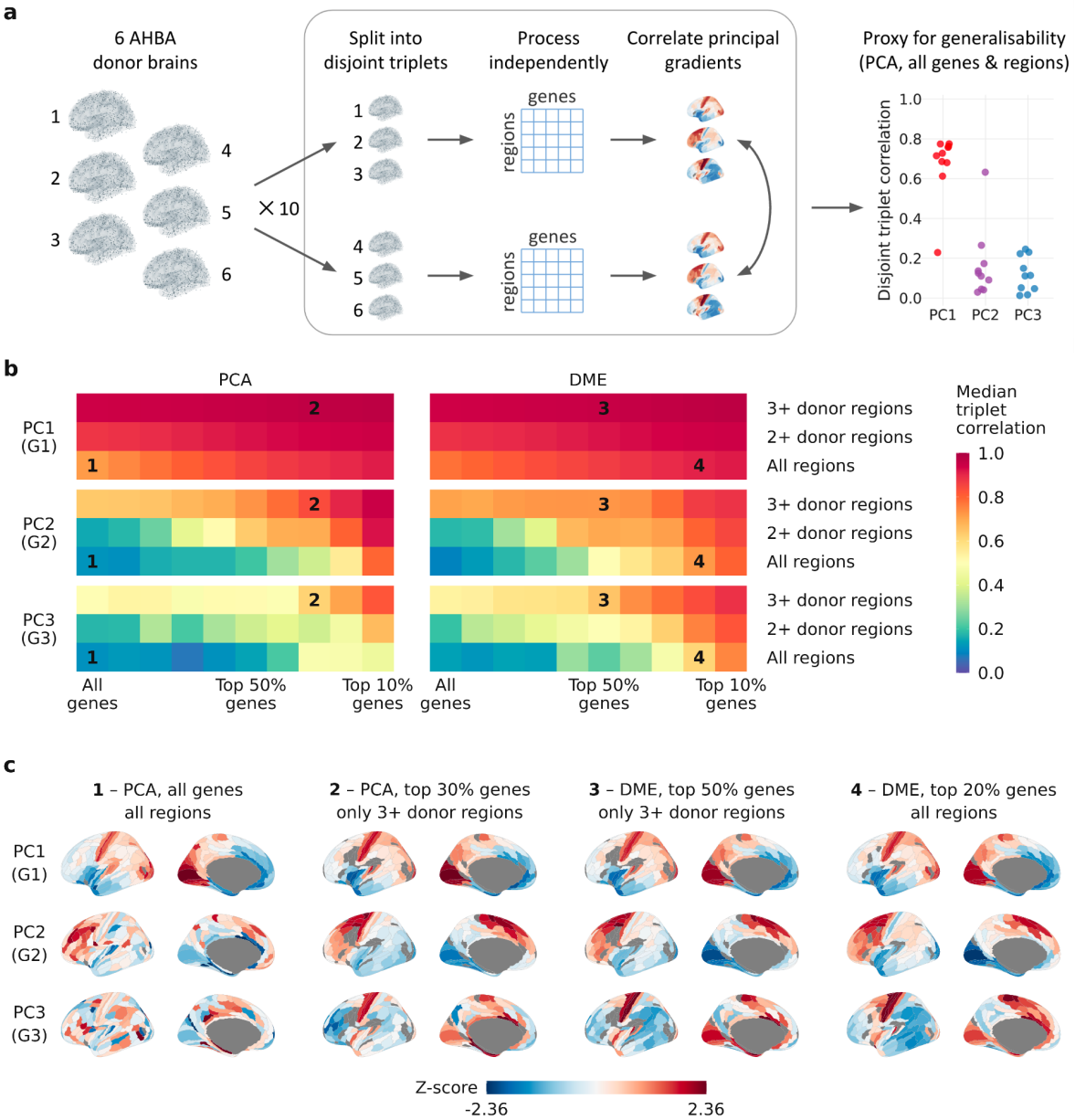

Figure 1: *Filtering genes and regions reveals three generalisable axes of gene expression.* ***a****, Splitting the six AHBA donor brains into triplets and comparing independently processed principal gradients is a proxy for generalisability. When using PCA with all regions and genes, PC2 and PC3 are unlikely to generalise.* ***b****, Generalisability proxy (i.e. median correlation of gradients from disjoint triplets) computed across all combinations of: (i, left/right panels) choice of PCA or DME, (ii, x-axis) filtering genes for X% most stable genes as computed by differential stability [49], and (iii, y-axis) filtering for only regions with samples matched from either 2+ or 3+ donors. All three parameters increase the generalisability of the principal gradients.* ***c****, Brain maps of the principal gradients from the four options of parameter combinations marked* ***1,2,3,4*** *in* ***b****. Options* ***2-4*** *are on the boundary where the third component has median triplet correlation > 0.6. These options are also equivalent (****Fig S3****). Brain maps in* ***c*** *are computed using all six AHBA brains together, which closely matches maps from any individual triplet (R>0.8). PCA, principal component analysis; DME, diffusion map embedding.*

The three parameters of gene filter, region filter, and choice of PCA or DME were independently varied, and we computed the proxy for generalisability of components (i.e. median correlation between components for disjoint sets of three donor brains) for each combination of parameters (**Fig 1b**). We observed that the three parameters all increased generalisability relative to the base case of PCA with all genes and regions (option **1**). For example, when using PCA, filtering for only the top 10% most stable genes increased the median triplet correlation of PC2 to 0.79, while filtering to only regions with samples from 3+ donors increased it to 0.64. However, to capture the fullest picture of cortical transcription it is valuable to retain the maximum extent of both genes and regions. To find an optimal balance between the competing constraints of generalisability and including more genes and regions, we compared three combinations of parameters at the boundary where PC1-3 (termed G1-3 when using DME) had median correlation > 0.6: PCA with the top 30% of genes and only regions with data from 3+ donors (option **2**); DME with top 50% of genes and only regions with data from 3+ donors (option **3**); and DME with only the top 20% of genes but all regions (option **4**). There was no combination of parameters where PC4+/G4+ had > 0.6 generalisability while retaining more than 10% of genes (**Fig S2**), so we limited further analysis to the first three components. We then used all six AHBA donor brains together to calculate the spatial maps corresponding to each parameter combination (**Fig 1c**). We found that these three sets of brain maps are equivalent (e.g. between option **3** and option **4**, G1-3 are respectively correlated at 1.0, 0.96, and 0.92; **Fig S3**). This equivalency suggests that identifying generalisable higher-order transcriptional structures requires a tradeoff of filtering more genes or more regions – for example, using DME, the same three cortical gradients can be identified in the AHBA with either 137/180 regions and half (~8000) the genes (option **3**), or all the regions but only 20% (~3000) of the most stable genes (option **4**).

After any of these filtering approaches the three intrinsic gradients displayed clear and distinct axial alignments in three-dimensional (i.e. anatomical) space (**Fig 1c**). G1 was negatively aligned to the rostral-caudal axis (R=-0.64) but positively aligned to the dorsal-ventral axis (R=+0.54), whilst G2 was positively aligned to both rostral-caudal and dorsal-ventral axes (R=+0.49 and +0.58). Neither G1 or G2 were correlated to the medial-lateral axis (R<0.1), whereas this was the only axis to which G3 was aligned (R=-0.39). Furthermore, the three gradients appeared spatially autocorrelated ('smooth'), which was not the case for PC2 and PC3 as originally computed without the filtering strategies described above (**Fig 1c**). We quantified spatial autocorrelation with the Moran's I metric for which I=1 is perfect autocorrelation, I=-1 is perfect anticorrelation, and I=0 is the pattern expected from random noise [50]. Without filters (option **1**), PC1-3 respectively have Moran's I of 0.50, 0.23, and 0.00, whereas after filtering to increase generalisability, Moran's I is 0.51, 0.49, and 0.29 (for option **4**).

Crucially, the first three intrinsic gradients are not sensitive to the choice of 0.6 as a threshold for generalisability: relaxing or tightening either gene or region filter from one of the above options results in equivalent gradients (correlated at > 0.9, **Fig S3**). Similarly, the results are not sensitive to a wide range of other processing choices for the AHBA (**Fig S3**), such as whether right hemisphere samples are mirrored to the left hemisphere or dropped, how the microarray probes are aggregated to genes and normalised, choice of cortical parcellation template (**Fig S4**), etc. With any of these options, PC1 (referred to as G1 using DME) closely aligns to the original PC1 gradient characterised by Burt *et al.,* which explained 22% of variance in the region-by-gene expression matrix. Since the gene and region filters remove much of the unexplained variance, the three new axes G1, G2, and G3 respectively explain 51%, 7%, and 4% of variance (for option **3**). Hereafter we focus on G1-3 computed using option **3** (DME, top 50% of genes, 137/180 regions) so as to retain more genes and gain clearer insight into the biology of these gradients that likely represent three transcriptional axes of the human cortex.

Three transcriptional axes have distinct neurobiological profiles

We first asked whether the three transcriptional axes display distinct neurobiological profiles. For each axis we derived weights quantifying the relative contribution of individual genes in defining these axes (see Methods). As expected given the spatial alignment, the gene weights for the G1 axis reproduce the results of Burt *et al.*, having strong positive and negative weights (z-score > 1) for specific interneuron marker genes *SST*, *PVALB*, *VIP*, *CCK*, as well as glutamatergic neuronal genes *GRIN* and *GABRA* [34] (**Fig S6**). Exploring Gene Ontology Biological Process [51] enrichments for the two newly identified transcriptomic axes, we found that G2 positive weighted genes were significantly (*FDR* < 0.001) enriched for metabolic biological processes such as generation of energy, mitochondrion organisation, and metabolic process, while negative weighted genes are enriched for RNA regulatory terms (**Fig 2b**, full enrichment list for G1-3 in **Table S1**). For G3, we observed a highly significant (*FDR* < 0.0001) enrichment in positive weighted genes for immune function, including inflammatory process, antigen processing, adaptive immune process, T-cell activation, and leukocyte migration, as well as for vasculature processes such as angiogenesis. Conversely, the negative G3 weights were enriched for synaptic development.

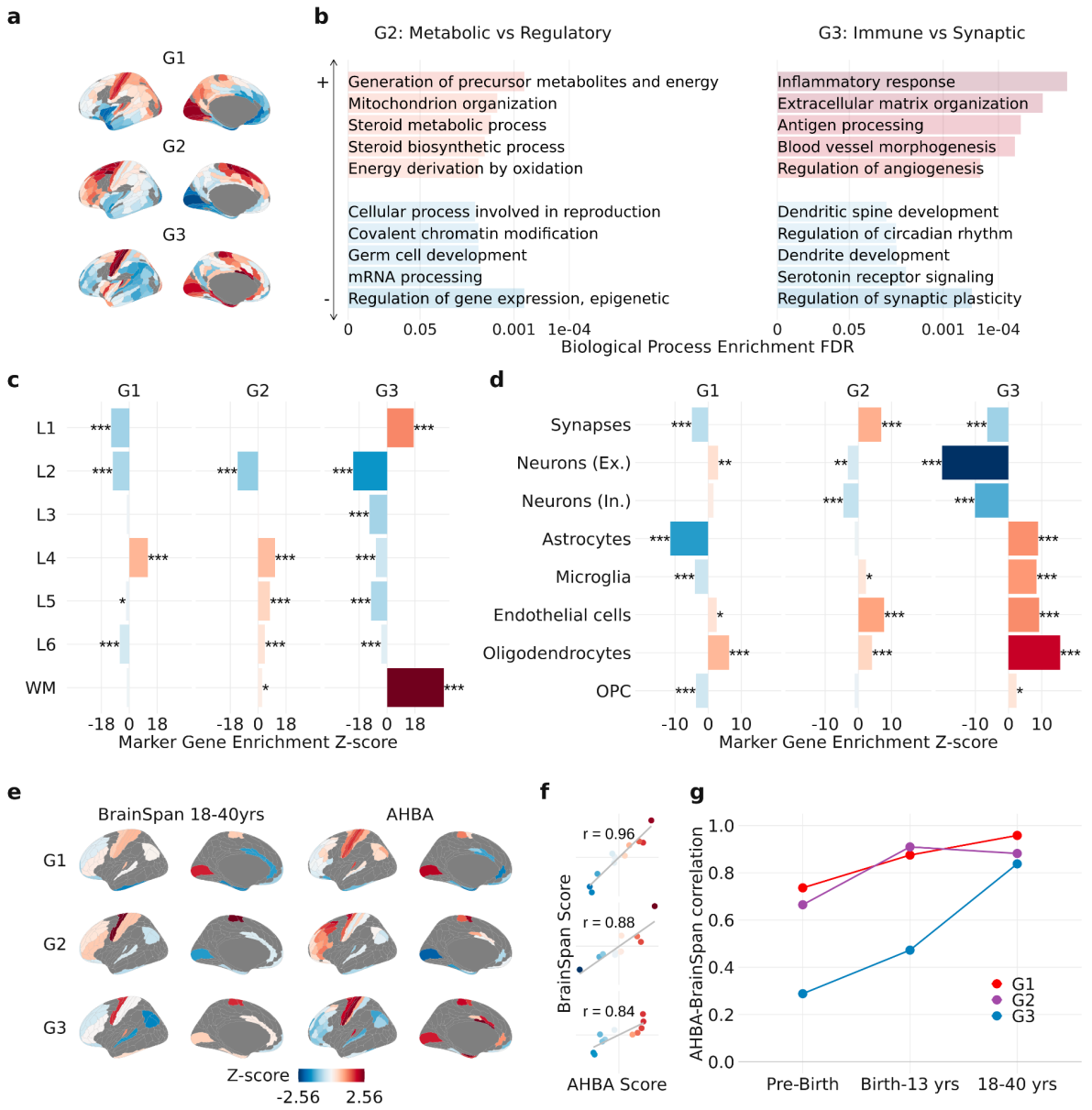

Figure 2: *Three axes of cortical gene expression have distinct biological profiles.* ***a****, AHBA axes from which gene weights are computed.* ***b****, Gene Ontology Biological Process enrichments for G2 and G3, top 5 positive and negative by enrichment FDR (full table in supplement).* ***c****, Enrichments of G1-3 gene weights for marker genes of six cortical layers and white matter (WM) [52]. Significance computed as percentile of mean weight of marker gene set relative to randomly permuted gene weights, and FDR-corrected across all 24 tests. *, **, *** indicate FDR < 0.05, 0.01, 0.001 respectively.* ***d****, Enrichments of G1-3 gene weights for cell type gene sets [7], and for synaptic marker genes [53]. Significance as in* ***c****.* ***e****, Scores of AHBA axis gene weights in BrainSpan adult brains compared to AHBA regions.* ***f****, Scatter plots of matched region scores (from* ***e****) show that the transcriptional axes genes have the same pattern in AHBA and BrainSpan adult brains.* ***g****, BrainSpan vs AHBA axis correlations (computed as in* ***f****) at different developmental periods.*

We next explored whether the three axes are enriched for cortical layers and cell-types. Using lists of marker genes for each cortical layer [52], we reproduced the result from Burt *et al.* that the G1 axis was positively enriched for layer L4 and negatively enriched for all other layers (**Fig 2c**). We further observed that all three axes had distinct enrichments for layer marker genes: G2 was negatively enriched for L2 and positively enriched for infragranular layers L4-6, while G3 was positively enriched for L1 and WM and negatively enriched for L2-6. Diversity of enrichment profiles was also observed for marker genes of synapses and cell-types derived from single-cell RNAseq studies [7,53]. G3 was positively enriched for glial cells and negatively enriched for both neurons and synapses (**Fig 2d**), aligning with its enrichment for immune vs synaptic Biological Processes, as well as for glia-enriched layers L1 and WM. For G2, cell type enrichments were more specific, with positive enrichments for synapses, endothelial cells, and oligodendrocytes, and negative enrichments for neurons. G1 was negatively enriched for synapses, astrocytes, microglia, and oligodendrocyte progenitor cells (OPC), but positively enriched for oligodendrocytes. A more fine-grained cell-type enrichment analysis using the neuronal clusters identified in Lake *et al.* [54] revealed stronger enrichments for specific neuronal types in G1 and G2, while G3 remained significantly negatively enriched for all neuron types (**Fig S6**). We also analysed the 19 genes enriched in the human supragranular cortex relative to mice and that are expanded in association cortex with dense cortico-cortical connections [55]. We found these genes were not enriched in G1 (FDR_{G1}=0.05), but were significantly positively enriched in G2 (FDR_{G2}=0.008) and negatively in G3 (FDR_{G3}=0.01).

Lastly, we hypothesised that the three axes may be related to development, which we investigated using BrainSpan – a separate human gene expression dataset where gene expression was measured in between 4 and 14 cortical regions from 35 donor brains at different developmental stages [31]. As BrainSpan lacks sufficient spatial granularity to independently reproduce the axes, we instead asked if the gene weights representing the biological profiles for each AHBA axis (G1-3) would have a consistent spatial pattern in both the AHBA and BrainSpan. When projecting the G1-3 gene weights from AHBA onto the gene expression of only the adult brains in BrainSpan (**Fig 2e**, see Methods), we found that the resulting spatial patterns were indeed highly correlated with the corresponding regions in the AHBA axes (R=0.96, 0.88, and 0.84 respectively) (**Fig 2f**), validating that the three axes are likely generalisable to individuals outside the six AHBA donors (e.g. the 35 brains sampled in BrainSpan). We then similarly compared the adult AHBA axes with BrainSpan brains from earlier developmental stages. We observed that for G1 and G2, BrainSpan and AHBA were highly consistent even in preadolescent and prenatal donor brains (Birth-13yr: R_{G1}=0.87, R_{G2}=0.91; Prenatal R_{G1}=0.74, R_{G2}=0.66), whereas G3 was consistent only in the adult brains and not stable at earlier developmental

timepoints (Birth-13yr: R_{G3}=0.47; Prenatal R_{G3}=0.29) (**Fig 2g**). Taken together, the above analyses show how the three transcriptional axes have different anatomical, biological, and developmental profiles: G1, previously described as the sensorimotor-association axis, is present from or before birth and is aligned with specific inhibitory and excitatory marker genes; G2 is enriched for metabolic processes and also emerges early in development; whilst G3 is enriched for immune vs synaptic processes and for markers of glia (and glial layers) vs neurons & synapses, and may only emerge during adolescence.

Neuroimaging maps are underpinned by multiple gene expression axes

Having demonstrated that the three AHBA components are biologically distinct, we next asked how they relate to macroscale brain features from neuroimaging. We explored this by spatially correlating the axes with ten previously published neuroimaging maps of brain structure and function (**Fig 3c**). These include maps of macroscale structure (cortical thickness [56] and allometric scaling [57]), MRI proxies for microscale structure (T1w/T2w index of myelination [35], and externopyramidisation, i.e. soma size [58]), metabolism (aerobic glycolysis [59] and cerebral blood flow [60]), function (principal variation in fMRI connectivity [40] and Neurosynth cognitive terms [61]), and maps of cortical expansion both over development and in humans relative to non-human primates [42]. These maps were chosen as they have all been suggested to be patterned along the same sensorimotor-association axis as the transcriptional PC1[39]. We reproduced prior results from Burt *et al.* that the G1 sensorimotor/association axis has strong spatial correlations with cortical myelination (R_{G1}=0.74, p_s<10^{-4}) and cortical thickness (R_{G1}=-0.66, p_s<10^{-4}) [34] (**Fig 3d**). For the metabolically-enriched G2 axis we observed a strong correlation with the map of aerobic glycolysis as measured by glycolytic index (R_{G2}=0.66, p_s=0.004), which was much stronger than the aerobic glycolysis map correlation with G1 (R_{G1}=0.17, p_s=0.7). Across the ten maps, six had stronger correlations with G2 than with G1, in particular the maps of developmental expansion (R_{G1}=-0.08, p_s=0.75; R_{G2}=0.44, p_s=0.01), evolutionary expansion (R_{G1}=-0.20, p_s=0.5; R_{G2}=0.44, p_s=0.012), and externopyramidisation (R_{G1}=0.08, p_s=0.9; R_{G2}=-0.57, p_s=0.09). None of the ten maps were most aligned with G3.

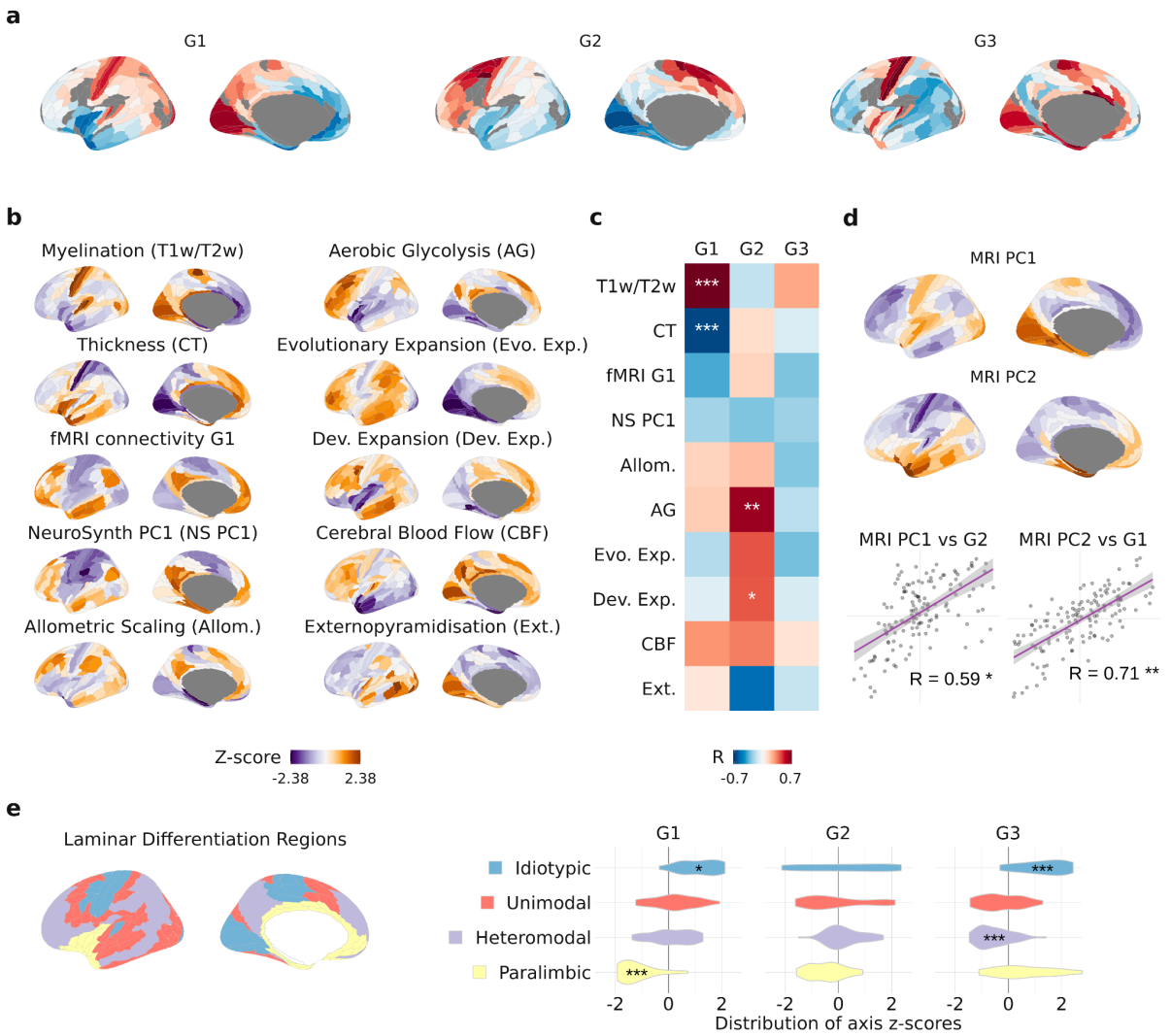

Figure 3: *Neuroimaging maps are underpinned by multiple axes of gene expression.* ***a,*** *Three transcriptional axes.* ***b,*** *Ten neuroimaging maps that have been proposed to align to a sensorimotor-association axis [39], references in main text.* ***c,*** *Spatial correlations of neuroimaging maps with transcriptional axes. Significance assessed against spatially autocorrelated spin permutations: *, **, *** indicate* p_s *< 0.05, 0.01, 0.001 respectively.* ***d,*** *The common spatial variation among key neuroimaging maps identifies the first two transcriptional axes: PC1 of the set of ten MRI maps aligns with the metabolically-enriched G2 axis, while PC2 of the MRI maps aligns with the sensorimotor-association G1 axis. These associations are robust to bootstrapped sampling of the ten maps (**Fig S6**).* ***e,*** *Distributions of G1-3 scores in the Mesulam atlas of laminar differentiation [62] showing that G3 distinguishes the regions where cortical layers are most (idiotypic) and least (heteromodal) distinct. Significance assessed as the percentile of the true mean score relative to the mean score from spatially autocorrelated spin permutations.*

We further hypothesised that spatial patterns which are shared across the neuroimaging maps themselves may relate to the underlying transcriptional axes described here. We investigated this by applying PCA to the set of ten neuroimaging maps, independent from any gene expression data (**Fig**

3e). We found that 51% of the variance in these ten imaging maps was explained by their top two PCs, and that PC1 of the neuroimaging maps was significantly aligned only to the metabolically-enriched G2 transcriptional axis from the AHBA (R_{G2}=0.59, p_s=0.01), and not to the G1 axis (R_{G1}=-0.29, p_s=0.38). This PC1 of neuroimaging maps explained 31% of variance and was driven by the primary fMRI connectivity gradient, aerobic glycolysis, evolutionary expansion, and developmental expansion. Conversely, PC2 of the neuroimaging maps aligned only with the sensorimotor-association G1 axis (R_{G1}=0.71, p_s=0.003), and not with G2 (R_{G2}=0.36, p_s=0.48). It explained 20% of variance and was driven by the maps of cortical thickness, the Neurosynth cognitive axis, cerebral blood flow, and the T1w/T2w ratio. These results were robust to both bootstrapping and subsampling the ten maps before computing PCA (**Fig S7**). In other words, the same basic axes of brain organisation emerge independently, both from the common patterns that explain over half the spatial variance in neuroimaging maps of diverse macroscale features, and from the spatial expression patterns of the ~8,000 genes whose expression patterns are most stable across individuals.

Lastly, we investigated whether the three transcriptional axes may be related to atlases of cortical cytoarchitecture [63], functional networks [64], and laminar differentiation [62]. G1 had significant associations with regions of all three atlases (**Fig S8**), while G3 was specifically associated with the histologically-identified regions of greatest and least laminar differentiation [62] (**Fig 3e,** idiotypic vs heteromodal). The G3-positive brain regions enriched for immune processes, glia, and L1/WM (**Fig 2b,c,d**) were co-located with the idiotypic brain regions identifying primary sensory-motor cortex where the cortical layers are most distinct – these regions were also associated with G1. Conversely, the synaptically-enriched G3-negative regions were co-located with the heteromodal association cortex defined by neuronal inputs and responses representing multiple sensory modalities, and where cortical layers are least distinct [62]. Thus, although G3 was not clearly associated with any of the neuroimaging features analysed above, its strong biological enrichments may reflect laminar differences in cortical microstructure.

Neurodevelopmental disorders are underpinned by multiple gene expression axes

Cortical transcriptional axes may underpin not only brain function, but also brain dysfunction. We therefore explored associations with the recently published BrainChart data, in which normative models derived from >125,000 MRI scans were used to characterise regional deviation in cortical volume of psychiatric patients relative to median growth charts of brain development [65] (**Fig 4a**). Using the 34-region Desikan-Killiany parcellation imposed by this dataset (**Fig 4b**, see Methods), we

found that even under spin tests corrected for false discovery rate across all nine comparisons, autism spectrum disorder (ASD) was significantly associated with both G1 and G2 (R_{G1}=0.49, p_s=0.0002, FDR_s=0.002; R_{G2}=-0.28, p_s=0.0006, FDR_s=0.003), while major depressive disorder (MDD) was associated only with G1 (R_{G1}=0.42, p_s=0.01, FDR_s=0.02), and schizophrenia (SCZ) was significantly associated only with G3 (R_{G3}=0.44, p_s=0.008, FDR_s=0.02) (**Fig 4c**).

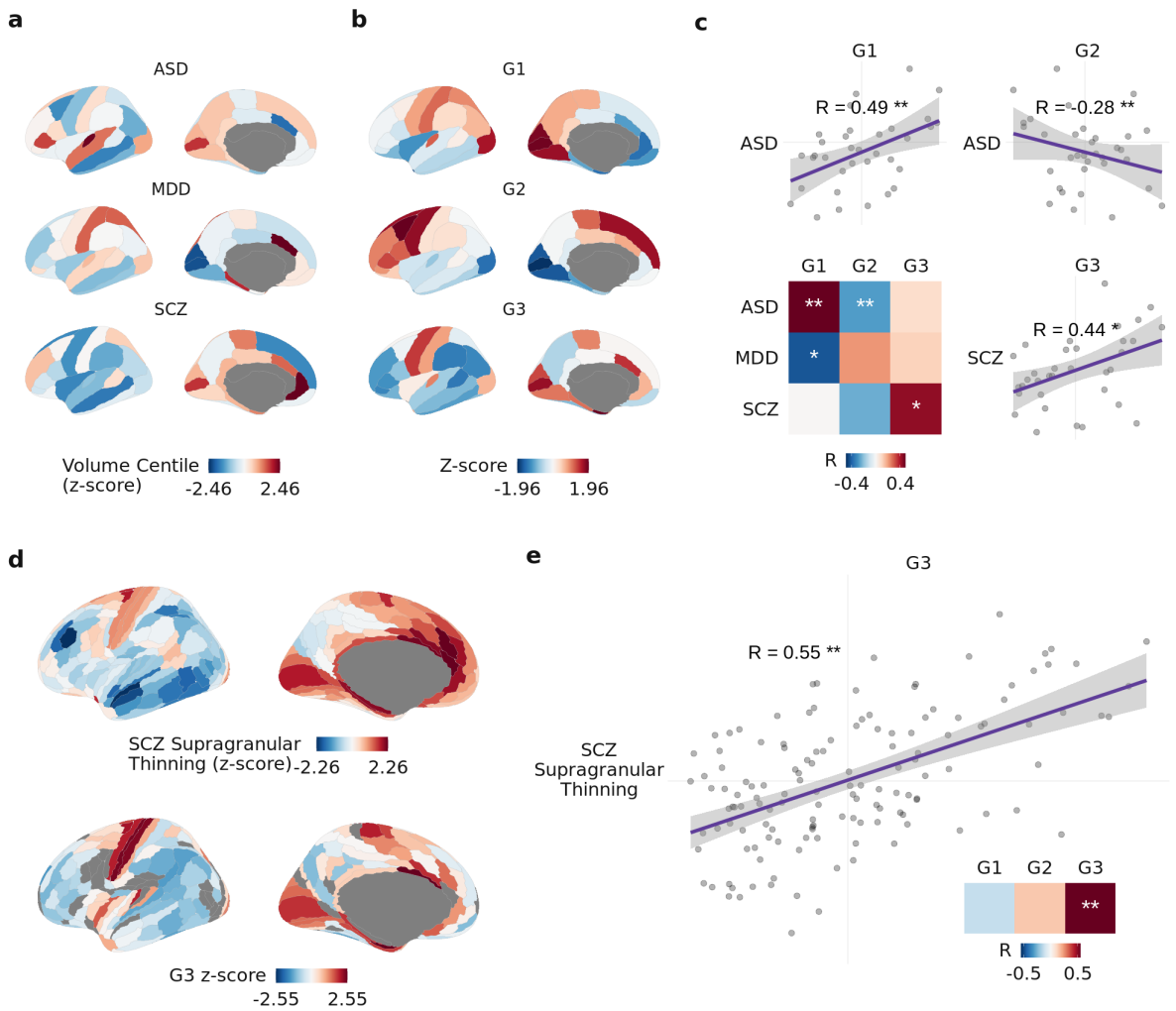

Figure 4: *Transcriptional axes may underpin psychiatric disorders.* ***a****, Cortical volume differences in autism spectrum disorder (ASD), major depressive disorder (MDD), and schizophrenia (SCZ) cases from normative modelling.* ***b****, AHBA axes projected into Desikan-Killiany parcellation.* ***c****, Spatial correlations between maps of cortical volume differences in ASD, SCZ, MDD (as in* ***a****) with the G1-3 transcriptional axes (as in* ***b****). Scatter plots shown for the strongest associations of each of G1-3. Significance assessed against spatially autocorrelated spin permutations, and corrected for multiple comparisons: *, **, *** indicate FDR_s < 0.05, 0.01, 0.001 respectively.* ***d****, Map of supragranular-specific cortical thinning in SCZ [66] in HCP-MMP parcellation, and G3.* ***e****, Scatter plot of region z-scores in* ***d****, significance assessed as in* ***b****.* ***Inset****, spatial associations of SCZ supragranular thinning map with G1-3.*

Our result that G1 is spatially associated to both ASD and MDD aligns with prior suggestions that genes aligned to the first transcriptional axis are implicated in multiple psychiatric disorders, including autism and depression [34]. For G2 – the axis we found to be genetically enriched for metabolic processes (**Fig 2b**) and spatially associated with aerobic glycolysis (**Fig 3c**) – the spatial association to ASD also makes sense in light of prior research suggesting that autism may be related to metabolic and mitochondrial disorders [67–69]. Similarly, the specific spatial association between the immunity-enriched G3 transcriptional axis and SCZ aligns with prior research suggesting a role for neuroinflammation in schizophrenia [70–75]. The strong and significant association between G3 and the brain regions of most and least laminar differentiation (**Fig 3e**) also suggests that the G3-SCZ association may relate to laminar architecture. To investigate this laminar hypothesis, we compared G3 with a prior map of specifically supragranular cortical thinning in SCZ [66] (**Fig 4d**). We found this layer-specific SCZ phenotype to be strongly and specifically associated with G3 (R_{G3}=0.55, p_s=0.001, FDR_s=0.004) (**Fig 4e**), suggesting that the G3 association with overall SCZ cortical volume reduction may be driven by the co-location of supragranular-specific thinning with regions of lower laminar differentiation. Taken together, these findings suggest that intrinsic axes of brain transcription may also underpin aspects of healthy brain organisation that are disrupted in neurodevelopmental psychiatric disorders.

Discussion

In this study we hypothesised that gene expression in the healthy adult human neocortex may exhibit multiple dimensions of intrinsic spatial structure, akin to the multiple transcriptional brain axes observed in both embryonic human, chick, and mice brains, and the adult brains of mice and non-human primates [23–32]. The existence of multiple transcriptional axes is particularly relevant in the context of neuroimaging, as the ubiquity of associations between neuroimaging maps and the first principal component of gene expression [34] has been taken as evidence that diverse brain features may all be organised along a common sensorimotor-association axis [39]. However, the search for additional transcriptional axes has faced obstacles due to gene expression variations across donor brains and potential limited generalisability of the six-brain AHBA - the only high-resolution spatial transcriptome for the adult human brain.

In this study we sought to overcome these obstacles by defining an explicit proxy for generalisability, namely consistency within independent subsets of the AHBA dataset. By optimising data processing for this proxy measure, we identified two previously undiscovered transcriptional axes. We validated the three axes by showing that their genetic profiles reflect the same spatial pattern in both the AHBA and the lower-resolution BrainSpan RNA-seq dataset based on a larger, independent cohort of 35 brains. We showed that the three axes have distinct enrichment profiles for biological functions, cortical layers, cell-types, and developmental trajectories. We then asked how maroscale brain structure and function relate to the three transcriptional axes. Here we referenced a set of neuroimaging maps previously associated with a single sensorimotor-association axis linked to transcriptional PC1 [39]. Although these diverse maps are all somewhat aligned to a sensorimotor-association axis, they are still quite distinct, and we therefore hypothesised that their differences may be explained by other axes of transcriptional structure. We found that half of this set of neuroimaging maps are more aligned to the metabolically-enriched G2 axis than to the G1 sensorimotor-association axis (**Fig 3c**). Moreover, half of the variance in the same set of maps are captured by spatial patterns aligned to transcriptional G1 and G2, which respectively explain 20% and 31% of variance across the neuroimaging maps (**Fig 3d**). This suggests that the common method of interpreting brain maps via spatial correlations with previously reported maps may be misleading, in that a significant correlation may not imply specificity of association, but rather reflect aspects of a shared underlying transcriptional architecture. Furthermore, using the AHBA to 'decode' the biology of a neuroimaging map by identifying gene sets with similar spatial patterns often results in convergent gene enrichments even from independent and apparently unrelated neuroimaging maps [76]. Fulcher *et al.* showed that these convergences likely result from not taking into account the spatial

co-expression and autocorrelation of brain transcription – this intrinsic spatial structure of gene expression is what our study of transcriptional axes seeks to characterise biologically. Future studies in imaging transcriptomics may therefore benefit from situating neuroimaging maps with respect to the three transcriptional axes presented here.

In addition to underpinning the organisation of macroscale brain features measured by MRI, the three transcriptomic axes may represent important aspects of brain organisation in their own right. For example, the sensorimotor-association G1 axis has been suggested to reflect the relative feedforward-vs-feedback specialisation of cortical regions [34], which may be associated with grey-matter myelination. Consistent with this, we found that genes associated with myelination such as *MBP* were most highly weighted in G1, which is notable given that the broader set of oligodendrocyte marker genes were more enriched in G3 than G1. For the G2 axis, the metabolic process enrichments (**Fig 2b**) and spatial co-location with the map of aerobic glycolysis (an index of brain metabolism) (**Fig 3c**) align with prior analysis by Goyal *et al.*, who applied dimension reduction to 900 samples from just one of the six AHBA brains and similarly found that regions of high and low aerobic glycolysis cluster along the second transcriptomic dimension [41]. Goyal *et al.* also related aerobic glycolysis to development and specifically to developmental processes that are protracted in humans relative to other species (neoteny), which accords with our results that the G2 axis is spatially associated with maps of developmental and evolutionary expansion [42], and genetically associated with genes enriched in the human supragranular cortex relative to mice [55]. Thus, if the sensorimotor-association G1 axis represents feedforward-feedback cortical specialisation [34], it may be the case that G2 captures a second dimension of cortical differentiation distinguishing metabolically expensive and evolutionarily recent frontal cortices whose expansion throughout development is driven by the expression of G2 genes. Finally for the G3 axis, it is notable that the axis distinguishes brain regions transcriptionally enriched for immune processes, glia, and white matter, from those enriched for synaptic processes, neurons, and layers 2 through 5 (**Fig 2c,d**). These poles of the G3 axis are co-located with the regions of greatest and least laminar differentiation as determined by histology [62] (**Fig 3e**), which have in turn been proposed as a basis for structure-function coupling [77,78]. As both structure-function coupling and laminar differentiation continue to change throughout adolescence [17,79], the G3 transcriptional axis may similarly not be consolidated until adulthood, in accordance with our result that the expression pattern of G3 genes is not conserved in pre-adult brains (**Fig 2g**).

The three transcriptional axes are likely to have translational relevance for psychiatric disorders. Using normative modelling of cortical volume derived from the BrainChart dataset that includes over 125,000 fully harmonised MRI scans [65], we observed autism spectrum disorder (ASD) to be significantly associated with both G1 and G2, while schizophrenia (SCZ) was significantly associated with G3 (**Fig 4c**). The ASD association aligns with recent work by Wagstyl *et al.*, whose gene co-expression network analysis of the AHBA revealed two maps that are closely aligned to G1 and G2 and both enriched for disjoint sets of ASD risk genes [80]. Moreover, both the ASD and SCZ associations also accord with our developmental analysis (**Fig 2g**): ASD is diagnosable early in development and G1 and G2 are conserved in preadolescent and prenatal brains, whereas SCZ is only diagnosable in adolescence and G3 was similarly only visible in post-adolescent brains. Moreover, G3 was also closely aligned with the map of specifically supragranular cortical thinning in schizophrenia [66] (**Fig 4e**). Together with the fact that G3 differentiates regions of greatest and least laminar differentiation (**Fig 4d**), this suggests that the mechanisms underlying cortical thinning in schizophrenia may be related to genes that are (i) most highly expressed in supragranular layers, (ii) enriched for synaptic and neuronal processes and cell-types, and (iii) highly expressed in one extreme of the G3 transcriptional axis, namely the brain regions devoted to heteromodal associative brain function [62]. Finally, major depressive disorder (MDD) was weakly associated with G1, despite G1 existing in early development yet MDD (as with SCZ) typically emerging in adolescence. A possible reason for this discrepancy is Burt *et al.*'s proposal that G1 represents the most foundational aspects of brain organisation and as such is associated with all disorders [34]. If so, the lack of an association between G1 and SCZ may be due to limited signal-to-noise, and remains a question for future research. Overall, our results complement recent findings pointing to shared molecular and structural underpinnings across psychiatric disorders [81–85].

Our study has several limitations. Although we were able to show consistency of the spatial gene expression axes between the AHBA and BrainSpan, this was not a full replication of the three axes, which would require a second transcriptomic brain atlas with high spatial resolution. Critically, postmortem gene expression is a snapshot of a dynamic process that varies throughout the lifespan, as well as being dependent on individual genetics and environmental factors [86]. However, the six-brain AHBA only allows the study of those components of the brain's transcriptional architecture that are least dynamic and susceptible to perturbation, potentially limiting the relevance of these findings. This concern is most relevant for genes prominent in G3, as our analysis suggests that G1 and G2 are relatively stable over development (**Fig 2g**). Our study was also limited to the left cortical hemisphere, so investigations of hemispheric asymmetry and subcortical associations are left for

future work. Lastly, the AHBA comprises only bulk microarray data, meaning that we cannot distinguish spatial variation due to varying cell-type composition from spatial variation of gene expression within the same cell-types. The data may also be prone to the biases of microarray measurement, particularly for genes with the highest and lowest overall expression due to nonlinearity in how mRNA concentration relates to intensity of fluorescence. With the continual advancement of single-cell RNA-seq technology [87], future studies may be able to further validate these results in distinct cell-types, and more broadly explore how the brain's transcriptional axes relate to the myriad dimensions of human and animal transcriptomes.

The ability to link neuroimaging to genetics has catalysed a new era of bridging macroscale and microscale perspectives on the brain. Ten years since the release of the AHBA, questions have arisen regarding the generalisability of results from only six brains, the lack of higher-order principal components beyond PC1, and the convergence in transcriptomic correlates of diverse neuroimaging maps. Future studies using the AHBA may therefore benefit from the methodological advances proposed here, such as defining a proxy for generalisability, and filtering for genes and regions to optimise it. On the quest to bridge between macroscale and microscale, the three transcriptional axes revealed here may support further research into how human cortical anatomy, development, and disorders are underpinned by the multidimensional intrinsic structure of gene expression in the brain.

Methods

AHBA data and parcellation images

Probe-level gene expression data with associated spatial coordinates was obtained from the Allen Institute website (https://human.brain-map.org/static/download). HCP-MMP1.0 parcellation images matched to the individual native MRI space of each donor brain were obtained from Arnatkeviciute *et al.* (https://figshare.com/articles/dataset/AHBAdata/6852911) [47].

AHBA processing parameters

To correctly match AHBA samples to regions in native donor space parcellation images using published processing pipelines, it is recommended to use either (i) abagen version 0.1.3 or greater (for python) [48], or (ii) the version of the AHBAprocessing pipeline updated in June 2021 or later (for matlab) [47]. Here we processed the AHBA with the abagen package using the below parameters, which in unless otherwise specified follow published recommendations [47]:

- *Hemisphere*: The right hemisphere samples that are present for two of the six donors were mirrored onto the left hemisphere of those donor datasets.
- *Intensity-based filter*: Probes were filtered to retain only those exceeding background noise in at least 50% of the samples, using the t-test provided by the Allen Institute.
- *Probe aggregation*: Probes were aggregated to genes by differential stability, meaning that for each gene, the probe with the highest mean correlation across donor pairs was used.
- *Distance threshold*: Samples were matched to regions with a tolerance threshold of 2mm, using the voxel-mass algorithm in the abagen package.
- *Sample normalisation*: Prior to aggregating over donors, samples were normalised to have the same mean expression level across all genes using a scaled robust sigmoid.
- *Gene normalisation*: Prior to aggregating over donors, genes were normalised to have the same mean expression level across all samples.

To ensure robustness the primary analysis was repeated in a series of sensitivity analyses varying all of the parameters above (e.g. not mirroring right hemisphere samples to the left hemisphere, different or no intensity filter for genes, different methods for aggregating and normalising probes) (**Fig S4**). We also performed the processing using alternative parcellation templates (**Fig S5**).

Gene filtering by differential stability

Genes were filtered for those that show more similar spatial patterns of expression across the six donors using the metric of differential stability (DS) [49]. For each gene, DS is calculated as the average correlation of that gene's regional expression vector between each donor pair (15 pairs with all six

brains, or three pairs in the triplets analysis (Fig 1a)). Genes are ranked by DS and then only the top X percentile of genes are retained, where X was varied between 10% and 100% (see Results).

Filtering regions by donors represented

Regions were filtered for those that include samples from at least three of the six AHBA donor brains, which in the HCP-MMP1.0 parcellation retains 137/180 regions. In the triplets analysis (Fig 1a), this process means only brain regions with samples from all three donors in the triplet were retained.

Dimension reduction

Dimension reduction was performed using both Principal Component Analysis (PCA) and Diffusion Map Embedding (DME), the latter having been extensively described for use in spatial brain analysis by Margulies *et al.* [40]. For DME, normalised cosine was used as the kernel for the affinity matrix. No sparsity was added, and the alpha parameter was set at 1. These parameters were chosen as they optimised the inter-triplet correlation metric for generalisability. Both methods were implemented using the BrainSpace package [88].

PCA generates gene weights (i.e. loadings) for each PC directly. With DME, we computed equivalent gene weights using partial least squares (PLS) with the gene expression matrix as the input matrix and the DME spatial map for each component as the response vector [89]; this was applied independently for each component (so the gene weights for each component are not constrained to be orthogonal as in PCA). As partial least squares finds coefficients that maximise covariance, with normalised gene expression matrices the weights are equivalent to the Pearson correlations between the spatial expression vector of each gene and the component map. We verified that in the case of PCA, this method for computing gene weights reproduces the same weights as the normal PCA coefficients.

Disjoint triplet correlation as a proxy for generalisability

To test for generalizability we separated the six AHBA brains into pairs of disjoint triplets (for example donor brains 1,2,3 in one triplet and 4,5,6 in another). We applied our full analysis pipeline (including all processing steps e.g. probe aggregation, normalisation, and filters) independently to each of the twenty possible combinations of triplets and correlated the gene weights and region scores between each of the ten disjoint pairs. All correlations were computed over only the intersection of the regions and genes retained in each triplet of the pair. As the order of principal components can vary across different triplets, we employed a matching algorithm in which the full correlation matrix was computed for the top five principal components of each triplet. The highest absolute correlation

value in the matrix was then identified as representing two matched components, and the process repeated until all components were matched.

Biological process gene enrichment analysis

Biological process enrichments of the gene weights for each component were computed using the 'proteins with values/ranks' function of the STRING online software [90], which tests whether the mean weight of each annotated gene list is significantly higher or lower than random permutations of the same gene weights. This method therefore does not require choosing a cutoff threshold for target/background lists.

Layer and cell-type enrichment analyses

The gene lists for layer marker genes were obtained from published spatial transcriptomic enrichment analyses [52]. For each study we selected the genes with significantly greater expression in each cortical layer relative to all other layers. The synaptic gene enrichment uses the full gene list from SynaptomeDB [53]. All other cell-type gene lists were from Seidlitz *et al.* [7], who compiled lists of significantly differentially expressed genes from five independent single-cell RNA-seq studies [54,91–94]. Enrichments for cell-types were computed as above, by calculating the z-score of the true mean gene weight of each gene list against mean weights from 10,000 random permutations of the same gene weights. P-values were calculated as the percentile of the above, and corrected for multiple comparisons with the Benjamini-Hochberg false discovery rate adjustment.

BrainSpan developmental gene expression processing

BrainSpan data was obtained directly from the Allen Institute website [31] and processed as follows:

1. The 11 cortical regions in the BrainSpan data were matched to the HCP-MMP1.0 parcellation according to the descriptions in the BrainSpan documentation.
2. Exon-level expression data were filtered for only the matched BrainSpan regions.
3. Donor brains from which fewer than 4 regions were sampled were dropped.
4. Within each donor, expression of each gene was z-normalised over regions.
5. Donors were aggregated into three age buckets (prenatal, preadolescent, and adult) and expression was averaged for each gene.

AHBA-BrainSpan developmental consistency analysis

Consistency between the AHBA axes and BrainSpan was evaluated as follows:

1. Processed BrainSpan data were filtered for only the stable genes retained in the AHBA axes.

2. The dot product of the gene weights for each AHBA axis was taken against the BrainSpan data, resulting in three scores for each BrainSpan region at each age bucket.
3. The AHBA spatial maps were filtered and aggregated into the same 11 BrainSpan regions, and correlated against the BrainSpan scores for each age bucket (Pearson R).

Neuroimaging brain maps: bivariate analysis

The ten neuroimaging maps are publicly available in the Neuromaps package [95], as well as in Sydnor *et al.* [39]. Pearson correlations were computed between each imaging map and each of the three transcriptional axes, and significance tested using spin tests, leveraging tooling from Neuromaps [95].

Neuroimaging brain maps: multivariate analysis

PCA was applied to the set of ten neuroimaging maps and the first principal components compared against the AHBA axes. To test robustness, this analysis was repeated over 100 subsets of the five of the ten maps (**Fig S7**).

Psychiatric disorders: spatial associations

Maps of the regional centile score differences in cortical volume for ASD, SCZ, and MDD were obtained from the BrainCharts project by Bethlehem, Seidlitz, White *et al.* [65], in which normative models were computed for multiple brain phenotypes across the human lifespan from a harmonised dataset of >125,000 MRI scans. As these data were in the Desikan-Killiany parcellation, the AHBA axes in the HCP-MMP parcellation were mapped to fsaverage vertices then re-averaged into Desikan-Killiany. Pearson correlations were computed, significance assessed by spin tests, and corrected for FDR across all nine tests. Maps of MRI markers of supragranular cortical thinning in SCZ were taken from Wagstyl *et al.* [66], and parcellated in using the HCP-MMP1.0 parcellation. Pearson correlations were computed, significance assessed by spin tests with FDR corrections.

Author information

Richard Dear, Jakob Seidlitz, Ross D. Markello, Aurina Arnatkevičiūtė, Kevin M. Anderson, Richard A.I. Bethlehem, Lifespan Brain Chart Consortium, Konrad Wagstyl, Edward T. Bullmore, Armin Raznahan, Petra E. Vértes

Department of Psychiatry, University of Cambridge, Cambridge, UK

Richard Dear, Richard A.I. Bethlehem, Edward T. Bullmore, Petra E. Vértes

Lifespan Brain Institute, The Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, PA, USA

Jakob Seidlitz

Department of Child and Adolescent Psychiatry and Behavioral Science, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

Jakob Seidlitz

Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Jakob Seidlitz

McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

Ross D. Markello

Turner Institute for Brain and Mental Health, Monash University, Melbourne, Australia

Aurina Arnatkevičiūtė

Department of Psychology, Yale University, New Haven, CT, USA

Kevin M. Anderson

Wellcome Centre for Human Neuroimaging, London, UK

Konrad Wagstyl

Section On Developmental Neurogenomics, National Institute of Mental Health, Bethesda, MD, USA

Armin Raznahan

Lifspan Brain Chart Consortium

List of members https://github.com/brainchart/Lifespan

Author contributions

P.E.V., R.D., J.S., K.W., and A.R. designed research. J.S., A.A., K.W., and R.B. contributed data. R.D. and P.E.V. performed research, and K.A., E.T.B., K.W., A.R., J.S., R.M., and A.A. helped interpret results. R.D. and P.E.V. wrote the manuscript. All authors reviewed the manuscript.

Disclosures

KMA is an employee of Neumora Therapeutics. RDM is an employee of Octave Biosciences.

Acknowledgements

The authors acknowledge and thank Sofie Valk and Varun Warrier for their helpful comments on the manuscript.

R.D. was supported by the Gates Cambridge Trust. J.S. was supported by NIMH T32MH019112-29 and K08MH120564. A.A. was funded by a grant from the Australian Research Council (ARC) under its Linkage Project scheme (LP160101592). R.A.I.B. was supported by the Autism Research Trust. K.S.W. was supported by the Wellcome Trust (215901/Z/19/Z). E.T.B. was supported by an NIHR Senior Investigator award and the Wellcome Trust collaborative award for the Neuroscience in Psychiatry Network. A.R. was supported by the National Institute of Mental Health Intramural Research Program (NIH Annual Report Number, 1ZIAMH002949-04). P.E.V. is a Fellow of MQ: Transforming Mental Health (MQF17_24).

Data were curated and analysed using a computational facility funded by an MRC research infrastructure award (MR/M009041/1) to the School of Clinical Medicine, University of Cambridge and supported by the mental health theme of the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the NIH, NHS, the NIHR or the Department of Health and Social Care. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

References

1. van den Heuvel, M. P. & Yeo, B. T. T. A Spotlight on Bridging Microscale and Macroscale Human Brain Architecture. *Neuron* **93**, 1248–1251 (2017).
2. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
3. Vértes, P. E. *et al.* Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, (2016).
4. Morgan, S. E. *et al.* Cortical patterning of abnormal morphometric similarity in psychosis is associated with brain expression of schizophrenia-related genes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 9604–9609 (2019).
5. Hansen, J. Y. *et al.* Mapping gene transcription and neurocognition across human neocortex. *Nat Hum Behav* (2021) doi:10.1038/s41562-021-01082-z.
6. Romero-Garcia, R. *et al.* Schizotypy-Related Magnetization of Cortex in Healthy Adolescence Is Colocated With Expression of Schizophrenia-Related Genes. *Biol. Psychiatry* **88**, 248–259 (2020).
7. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat. Commun.* **11**, 3358 (2020).
8. Anderson, K. M. *et al.* Gene expression links functional networks across cortex and striatum. *Nat. Commun.* **9**, 1428 (2018).
9. Whitaker, K. J. *et al.* Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 9105–9110 (2016).
10. 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).
11. Shafiei, G. *et al.* Topographic gradients of intrinsic dynamics across neocortex. *Elife* **9**, (2020).
12. Anderson, K. M. *et al.* Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 25138–25149 (2020).

13. Anderson, K. M. *et al.* Transcriptional and imaging-genetic association of cortical interneurons, brain function, and schizophrenia risk. *Nat. Commun.* **11**, 2889 (2020).

14. Váša, F. *et al.* Conservative and disruptive modes of adolescent change in human brain functional connectivity. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 3248–3253 (2020).

15. Seidlitz, J. *et al.* Morphometric Similarity Networks Detect Microscale Cortical Organization and Predict Inter-Individual Cognitive Variation. *Neuron* **97**, 231–247.e7 (2018).

16. Romero-Garcia, R., Warrier, V., Bullmore, E. T., Baron-Cohen, S. & Bethlehem, R. A. I. Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. *Mol. Psychiatry* **24**, 1053–1064 (2019).

17. Paquola, C. *et al.* Shifts in myeloarchitecture characterise adolescent development of cortical gradients. *Elife* **8**, (2019).

18. Park, B.-Y. *et al.* An expanding manifold in transmodal regions characterizes adolescent reconfiguration of structural connectome organization. *Elife* **10**, (2021).

19. Wang, Y. *et al.* Long-range functional connections mirror and link microarchitectural and cognitive hierarchies in the human brain. *Cereb. Cortex* (2022) doi:10.1093/cercor/bhac172.

20. Fornito, A., Arnatkevičiūtė, A. & Fulcher, B. D. Bridging the Gap between Connectome and Transcriptome. *Trends Cogn. Sci.* **23**, 34–50 (2019).

21. Arnatkeviciute, A., Fulcher, B., Bellgrove, M. & Fornito, A. Imaging transcriptomics of brain disorders. *PsyArXiv* (2021) doi:10.31234/osf.io/4exug.

22. Martins, D. *et al.* Imaging transcriptomics: Convergent cellular, transcriptomic, and molecular neuroimaging signatures in the healthy adult human brain. *Cell Rep.* **37**, 110173 (2021).

23. Polleux, F., Ince-Dunn, G. & Ghosh, A. Transcriptional regulation of vertebrate axon guidance and synapse formation. *Nat. Rev. Neurosci.* **8**, 331–340 (2007).

24. Kudo, L. C., Karsten, S. L., Chen, J., Levitt, P. & Geschwind, D. H. Genetic analysis of anterior posterior expression gradients in the developing mammalian forebrain. *Cereb. Cortex* **17**, 2108–2122 (2007).

25. O'Leary, D. D. M., Chou, S.-J. & Sahara, S. Area patterning of the mammalian cortex. *Neuron* **56**, 252–269 (2007).

26. Wonders, C. P. *et al.* A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. *Dev. Biol.* **314**, 127–136 (2008).

27. Rakic, P., Ayoub, A. E., Breunig, J. J. & Dominguez, M. H. Decision by division: making cortical maps. *Trends Neurosci.* **32**, 291–301 (2009).

28. Kicheva, A. *et al.* Coordination of progenitor specification and growth in mouse and chick spinal cord. *Science* **345**, 1254927 (2014).

29. Saunders, A. *et al.* Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* **174**, 1015–1030.e16 (2018).

30. Bernard, A. *et al.* Transcriptional architecture of the primate neocortex. *Neuron* **73**, 1083–1099 (2012).

31. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).

32. Rayon, T., Maizels, R. J., Barrington, C. & Briscoe, J. Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development* **148**, (2021).

33. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-Scale Gradients in Human Cortical Organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).

34. Burt, J. B. *et al.* Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259 (2018).

35. Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011).

36. Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E. & Van Essen, D. C. Trends and properties of human cerebral cortex: correlations with cortical myelin content. *Neuroimage* **93 Pt 2**, 165–175 (2014).

37. Huntenburg, J. M. *et al.* A Systematic Relationship Between Functional Connectivity and Intracortical Myelin in the Human Cerebral Cortex. *Cereb. Cortex* **27**, 981–997 (2017).

38. Fulcher, B. D., Murray, J. D., Zerbi, V. & Wang, X.-J. Multimodal gradients across mouse cortex. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 4689–4695 (2019).

39. Sydnor, V. J. *et al.* Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* **109**, 2820–2846 (2021).

40. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12574–12579 (2016).

41. Goyal, M. S., Hawrylycz, M., Miller, J. A., Snyder, A. Z. & Raichle, M. E. Aerobic glycolysis in the human brain is associated with development and neotenous gene expression. *Cell Metab.* **19**, 49–57 (2014).

42. Hill, J. *et al.* Similar patterns of cortical expansion during human development and evolution. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13135–13140 (2010).

43. Goulas, A., Margulies, D. S., Bezgin, G. & Hilgetag, C. C. The architecture of mammalian cortical connectomes in light of the theory of the dual origin of the cerebral cortex. *Cortex* **118**, 244–261 (2019).

44. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).

45. Bernhardt, B. C., Smallwood, J., Keilholz, S. & Margulies, D. S. Gradients in brain organization. *Neuroimage* **251**, 118987 (2022).

46. Margulies, D. S. *et al.* Cortical Gradients and Their Role in Cognition. in *Encyclopedia of Behavioral Neuroscience, 2nd edition (Second Edition)* (ed. Della Sala, S.) 242–250 (Elsevier, 2022). doi:10.1016/B978-0-12-819641-0.00010-4.

47. Arnatkeviciute, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353–367 (2019).

48. Markello, R. D. *et al.* Standardizing workflows in imaging transcriptomics with the abagen

toolbox. *bioRxiv* 2021.07.08.451635 (2021) doi:10.1101/2021.07.08.451635.

49. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* **18**, 1832–1844 (2015).
50. Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).
51. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
52. Maynard, K. R. *et al.* Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).
53. Pirooznia, M. *et al.* SynaptomeDB: an ontology-based knowledgebase for synaptic genes. *Bioinformatics* **28**, 897–899 (2012).
54. 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).
55. Krienen, F. M., Yeo, B. T. T., Ge, T., Buckner, R. L. & Sherwood, C. C. Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E469–78 (2016).
56. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80**, 62–79 (2013).
57. Reardon, P. K. *et al.* Normative brain size variation and brain shape diversity in humans. *Science* **360**, 1222–1227 (2018).
58. Paquola, C. *et al.* A multi-scale cortical wiring space links cellular architecture and functional dynamics in the human brain. *PLoS Biol.* **18**, e3000979 (2020).
59. Vaishnavi, S. N. *et al.* Regional aerobic glycolysis in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17757–17762 (2010).
60. Satterthwaite, T. D. *et al.* Impact of puberty on the evolution of cerebral perfusion during adolescence. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8643–8648 (2014).
61. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated

synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).

62. Mesulam, M.-M. *Principles of behavioral neurology*. (Oxford University Press, USA, 1985).
63. von Economo, C. *Atlas of cytoarchitectonics of the adult human cerebral cortex*. (Karger, 1925).
64. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
65. Bethlehem, R. A. I. *et al.* Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).
66. Wagstyl, K. *et al.* Multiple markers of cortical morphology reveal evidence of supragranular thinning in schizophrenia. *Transl. Psychiatry* **6**, e780 (2016).
67. Zecavati, N. & Spence, S. J. Neurometabolic disorders and dysfunction in autism spectrum disorders. *Curr. Neurol. Neurosci. Rep.* **9**, 129–136 (2009).
68. Valenti, D., de Bari, L., De Filippis, B., Henrion-Caude, A. & Vacca, R. A. Mitochondrial dysfunction as a central actor in intellectual disability-related diseases: an overview of Down syndrome, autism, Fragile X and Rett syndrome. *Neurosci. Biobehav. Rev.* **46 Pt 2**, 202–217 (2014).
69. Kurochkin, I. *et al.* Metabolome signature of autism in the human prefrontal cortex. *Commun Biol* **2**, 234 (2019).
70. Müller, N. Inflammation in Schizophrenia: Pathogenetic Aspects and Therapeutic Considerations. *Schizophr. Bull.* **44**, 973–982 (2018).
71. Khandaker, G. M. *et al.* Inflammation and immunity in schizophrenia: implications for pathophysiology and treatment. *Lancet Psychiatry* **2**, 258–270 (2015).
72. Kirkpatrick, B. & Miller, B. J. Inflammation and schizophrenia. *Schizophr. Bull.* **39**, 1174–1179 (2013).
73. Mongan, D., Ramesar, M., Föcking, M., Cannon, M. & Cotter, D. Role of inflammation in the pathogenesis of schizophrenia: A review of the evidence, proposed mechanisms and implications for treatment. *Early Interv. Psychiatry* **14**, 385–397 (2020).
74. Fan, X., Goff, D. C. & Henderson, D. C. Inflammation and schizophrenia. *Expert Rev. Neurother.* **7**,

789–796 (2007).

75. García-Bueno, B. *et al.* Pro-/anti-inflammatory dysregulation in patients with first episode of psychosis: toward an integrative inflammatory hypothesis of schizophrenia. *Schizophr. Bull.* **40**, 376–387 (2014).
76. Fulcher, B. D., Arnatkeviciute, A. & Fornito, A. Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data. *Nat. Commun.* **12**, 2669 (2021).
77. Paquola, C. *et al.* Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS Biol.* **17**, e3000284 (2019).
78. Valk, S. L. *et al.* Genetic and phylogenetic uncoupling of structure and function in human transmodal cortex. *Nat. Commun.* **13**, 2341 (2022).
79. Park, B.-Y. *et al.* Adolescent development of multiscale structural wiring and functional interactions in the human connectome. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2116673119 (2022).
80. Wagstyl, K. *et al.* Transcriptional Cartography Integrates Multiscale Biology of the Human Cortex. *bioRxiv* 2022.06.13.495984 (2022) doi:10.1101/2022.06.13.495984.
81. Hansen, J. Y. *et al.* Local molecular and global connectomic contributions to cross-disorder cortical abnormalities. *Nat. Commun.* **13**, 4682 (2022).
82. Park, B.-Y. *et al.* Multiscale neural gradients reflect transdiagnostic effects of major psychiatric conditions on cortical morphology. *Commun Biol* **5**, 1024 (2022).
83. Hettwer, M. D. *et al.* Coordinated cortical thickness alterations across psychiatric conditions: A transdiagnostic ENIGMA study. *bioRxiv* (2022) doi:10.1101/2022.02.03.22270326.
84. Hansen, J. Y. *et al.* Molecular and connectomic vulnerability shape cross-disorder cortical abnormalities. *bioRxiv* 2022.01.21.476409 (2022) doi:10.1101/2022.01.21.476409.
85. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
86. Viñuela, A. *et al.* Age-dependent changes in mean and variance of gene expression across

tissues in a twin cohort. *Hum. Mol. Genet.* **27**, 732–741 (2018).

87. Langseth, C. M. *et al.* Comprehensive in situ mapping of human cortical transcriptomic cell types. *Commun Biol* **4**, 998 (2021).
88. Vos de Wael, R. *et al.* BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Commun Biol* **3**, 103 (2020).
89. Abdi. Partial least square regression (PLS regression). *Encyclopedia for research methods for the social*.
90. Szklarczyk, D. *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2018).
91. Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37–53 (2016).
92. Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
93. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7285–7290 (2015).
94. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, (2018).
95. Markello, R. D. *et al.* neuromaps: structural and functional interpretation of brain maps. *bioRxiv* 2022.01.06.475081 (2022) doi:10.1101/2022.01.06.475081.