Molecular signatures of cortical expansion in the human fetal brain

G. Ball[1,2] S. Oldham[1] V. Kyriakopoulou[3,4] L. Z. J. Williams[3,4] V. Karolis[3,5] A. Price[3,4] J. Hutter[3,4] M.L. Seal[1,2] A. Alexander-Bloch[6,7,8,9] J.V. Hajnal[3,4] A.D. Edwards[3,4] E.C. Robinson[3,4] J. Seidlitz[6,7,8,9]

1. Developmental Imaging, Murdoch Children's Research Institute, Melbourne, Australia
2. Department of Paediatrics, University of Melbourne, Melbourne, Australia
3. Centre for the Developing Brain, King's College London, London, UK
4. School of Biomedical Engineering & Imaging Science, King's College London, London, UK
5. Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
6. Department of Child and Adolescent Psychiatry and Behavioral Sciences, The Children's Hospital of Philadelphia, Philadelphia, PA
7. Department of Psychiatry, University of Pennsylvania, Philadelphia, PA
8. Lifespan Brain Institute, The Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, PA
9. Institute of Translational Medicine and Therapeutics, University of Pennsylvania, Philadelphia, PA

Abstract

The third trimester of human gestation is characterised by rapid increases in brain volume and cortical surface area. A growing catalogue of cells in the prenatal brain has revealed remarkable molecular diversity across cortical areas.[1,2] Despite this, little is known about how this translates into the patterns of differential cortical expansion observed in humans during the latter stages of gestation. Here we present a new resource, μBrain, to facilitate knowledge translation between molecular and anatomical descriptions of the prenatal developing brain. Built using generative artificial intelligence, μBrain is a three-dimensional cellular-resolution digital atlas combining publicly-available serial sections of the postmortem human brain at 21 weeks gestation[3] with bulk tissue microarray data, sampled across 29 cortical regions and 5 transient tissue zones.[4] Using μBrain, we evaluate the molecular signatures of preferentially-expanded cortical regions during human gestation, quantified *in utero* using magnetic resonance imaging (MRI). We find that differences in the rates of expansion across cortical areas during gestation respect anatomical and evolutionary boundaries between cortical types[5] and are founded upon extended periods of upper-layer cortical neuron migration that continue beyond mid-gestation. We identify a set of genes that are upregulated from mid-gestation and highly expressed in rapidly expanding neocortex, which are implicated in genetic disorders with cognitive sequelae. Our findings demonstrate a spatial coupling between areal differences in the timing of neurogenesis and rates of expansion across the neocortical sheet during the prenatal epoch. The μBrain atlas is available from: https://garedaba.github.io/micro-brain/ and provides a new tool to comprehensively map early brain development across domains, model systems and resolution scales.

The human cortex is a tapestry of specialised cortical areas supporting diverse and complex behaviours, each identifiable on the basis of distinct patterns of cyto-architecture, chemo-architecture, and axonal connectivity.[6-10] During gestation, waves of neurons are generated from progenitor cells lining the cerebral ventricles and migrate outwards along supporting radial glia to form the layers of the cortex.[11-13] Prior to the ingress of extrinsic connections via the thalamus,[14] the progressive differentiation of cortical areas is orchestrated by transcription factors expressed along concentration gradients and translated from the ventricular zone (VZ) to secondary progenitors of the subventricular zone (SVZ), then onto neurons in the cortical plate (CP), forming functional territories.[2,11,15-18] This process follows a precise spatiotemporal schema,[11,18-22] the traces of which extend far beyond the nascent stages of neurogenesis and are echoed in patterns of cytoarchitecture, axonal connectivity and function.[23-29]

Focused on uncovering the mechanisms that govern areal differentiation, studies have begun to catalogue the cellular diversity of the developing human cortex, and genes that encode it, with increasing granularity and scale.[1,2,30-32] Regional specialisation of cell types has been observed from early in gestation, with diversity of cortical gene transcription most evident in mid- to late-gestation but persisting into adulthood and aligning with structural and functional organisation of the brain.[4,26,33-36]

The third trimester of human gestation is characterised by rapid and sustained increases in brain volume and cortical surface area.[11,37,38] Differential rates of areal expansion during human development mirror evolutionary trends in cortical scaling and function[39-44] with preferential expansion in areas vulnerable to disruption in neurodevelopmental,[45] neurological[46], genetic[47] and psychiatric[48] disorders. Juxtaposed hypotheses implicate either the production of glia[49,50], or neurons[51-54] from specialised progenitor populations of the outer SVZ, in the expansion of the primate cortex. Thus, the distribution of distinct cell populations across the developing cortex may mediate areal differences in expansion and vulnerability to insult[18,33] but we currently do not have a clear understanding of how this molecular diversity is translated into cortical organisation in humans *in vivo*.

µBrain: A three-dimensional microscale atlas of the fetal brain

To bridge this gap, we sought to construct a 3D digital atlas of the developing brain at micrometre scale using a public resource of 81 serial histological 2D sections of a prenatal human brain at 21 postconceptional weeks (PCW).[3,4] Source data included serial coronal sections (20µm thickness) obtained from the right hemisphere of a single prenatal brain specimen (21 PCW; female), Nissl-stained, imaged at 1 micron resolution and labelled with detailed anatomical annotations, alongside interleaved coronal sections stained with *in situ* hybridisation (ISH) of n=41 developmental gene markers, as reported by Ding et al[3] (**Figure S1**; **Table S1-S3;** see **Methods**). In this and 3 other specimens (15, 16 and 21 PCW, 2 female), anatomical annotations had been used to guide a series of laser microdissections (LMD) across multiple cortical areas and layers of the cortical anlage (e.g.: cortical plate, subplate, intermediate zone, ventricular zone; **Table S4**) in the left hemisphere to measure regional gene expression via RNA microarrays, as described by Miller et al.[4] Nissl- and ISH-stained sections with corresponding anatomical labels and LMD arrays were made available as part of the BrainSpan Developing Brain Atlas [https://atlas.brain-map.org/atlas?atlas=3].

Artefacts due to tissue preparation, sectioning and staining procedures (including tearing and folding of sections) are common in histological data and can present difficulties for downstream processing pipelines.[55-58] To correct for tissue artefacts present in the histological data, we designed an automated detect-and-repair pipeline for Nissl-stained sections based on *pix2pix*, a Generative Adversarial Network (GAN)[59,60] (**Figure 1a-d; Supplemental Methods**). Using 256×256 pixel image patches drawn from 73/81 labeled histological sections (n=8 reserved for model testing) with paired anatomical labels, we trained a GAN model to produce Nissl-contrast images conditioned on a set of 20 tissue

labels (**Figure 1a; Table S2**). After training, the model was able to produce realistic, Nissl-stained image patches matched on colour hue and saturation to the original data using tissue annotations alone (**Figure 1c**). Model performance was robust to different parameter settings and model architectures (**Figure S2**). Using the trained model, we generated synthetic Nissl-contrast image predictions from anatomical annotations of each section and identified artefacts in the histological data based on deviations in pixel hue and saturation from the model prediction. Outlier pixels were replaced with model predictions using Poisson image editing[61] (**Figure 1d**) resulting in n=79 (2 excluded due to extensive missing tissue) complete histological sections (**Figure S1; Table S1**).

Histological atlases of the cerebral cortex[6,10] have proven invaluable for understanding human brain organisation but are limited by the loss of spatial information inherent to 2D representations of 3D structures. Reconstructions of 3D brain volumes from serial tissue sections of post mortem tissue allow the examination of intact brain anatomy at a scale inaccessible to current neuroimaging technologies.[62] We combined repaired tissue sections into a 3D volume of the right hemisphere using iterative affine image registration constrained by a tissue shape reference derived from fetal MRI (**Figure S3; Supplemental Methods**),[63] followed by nonlinear alignment to account for warping between adjacent sections. Using the aligned data, we generated a 3D volume resampled to voxel resolution $150 \times 150 \times 150 \mu m$ with dimension $189 \times 424 \times 483$ voxels ($28.35 \times 63.60 \times 72.45$mm) (μBrain; see **Methods**; **Figure 1e, Figure S4a-b**). Following reconstruction, we benchmarked the size of the reconstructed μBrain volume against standard fetal growth metrics for a 23 week (gestational age; GA, equivalent to 21 PCW) fetus (μBrain length = 62.7mm, 23 week GA occipital-frontal diameter median [5th, 95th centile] = 73.3 mm [68.2, 78.5]),[64] and compared tissue volume estimates based on reconstructed anatomical labels (parenchymal volume = 25.8ml, right hemisphere) to previously reported 3D MRI-derived fetal brain volumes (supratentorial volume [both hemispheres] at 23 week GA = 60.26ml).[65] Adapting protocols from neuroimaging analysis, we extracted the inner and outer surfaces of the cortical plate and projected a set of 29 cortical area labels derived from the histological tissue annotations (**Table S2**) onto the surface vertices to form the μBrain cortical atlas (**Figure 1f**). The μBrain cortical atlas represents a new parcellation of the developing brain defined according to the hierarchical ontology of the reference annotations and matched to corresponding LMD microarray data (**Table S2, S4; Figure 4c-d**).

In addition to the whole brain volume and cortical atlas, we created partial 3D reconstructions of ISH staining for 41 genes (see **Methods**; **Figure 1g**). Based on an average 41 tissue sections per gene (**Table S3**), semi-quantitative maps of gene expression revealed the tissue- and region-specific distributions of several genes, including caudal enrichment of the transcription factor *EOMES* in the subventricular zone,[15] and markers of neuronal migration (*DCX*[66]) and synaptic transmission (*GRIK2*[67]), in the cortical plate (**Figure 1g**; **Figure S5**)

Existing histological brain atlases, including those of the adult human,[62,68,69] mouse,[70,71] and macaque[72] brains facilitate integration with other data modalities, including neuroimaging, and are amenable to advanced computational image analysis methods to extract quantitative measures of neuroanatomy across multiple scales.[73,74] Building upon existing resources,[3,4] we have created the μBrain atlas (**Figure 1; Figure S4**), a new and freely-available 3D volumetric model of the 21 PCW fetal brain at $150 \mu m$ resolution, accompanied by a set of n=20 cerebral tissue labels (**Figure S4a-b**); surface models of the cortical plate surface and cortical plate/subplate interface with n=29 cortical area labels (**Figure 4c**) and n=41 partial reconstructions of ISH expression data (**Figure 5**). Cortical areas are matched to normalised gene expression data from corresponding LMD microarrays (**Table S4; Figure S4d**) across multiple tissue zones in three additional prenatal specimens (total n=4) providing a 3D anatomical coordinate space to facilitate integrated imaging-transcriptomic analyses of the developing brain. Below, we use the μBrain atlas to evaluate the molecular and cellular correlates of cortical expansion in the third trimester of human gestation.
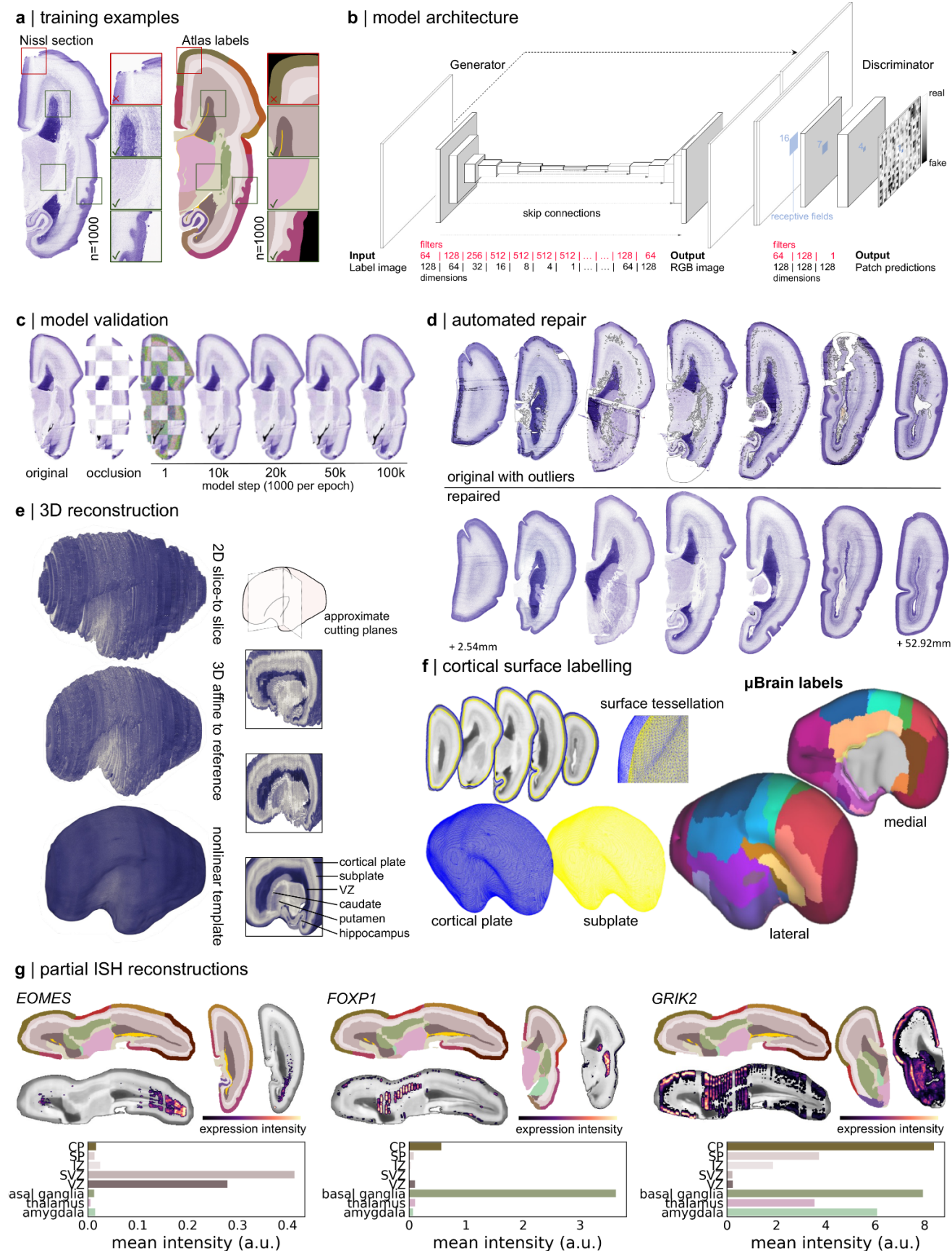
Figure 1: Generation of a 3D anatomical atlas of the mid-gestation fetal brain. **a**. Paired histological sections and simplified anatomical annotations were divided into 256 × 256 random patches (n=1000) for model training. Patches were quality checked prior to selection to ensure good overlap between labels and anatomy and no tissue damage. **b**. *Pix2pix* model architecture showing a U-Net generator coupled with a PatchGAN discriminator. Box sizes represent image width, height and number of filters/channels (depth) at each layer. Filters and dimensions of each layer are

shown below. **c**. Model performance was evaluated on a set of sections that were not included in the training dataset. Checkerboard occlusions are shown with the original section, occluded patch predictions are shown using the trained model after a given number of iterations. **d**. The trained model was used to replace RGB values of outlying pixels with synthetic estimates. Top row: original sections spaced throughout the cerebral hemisphere with automatically identified outlier pixels outlined in grey. Bottom row, repaired sections. **e**. Repaired sections were aligned via linear, affine and iterative nonlinear registrations (see **Methods**) to create a 3D volume with final isotropic resolution of 150um. Right: Cut-planes illustrate internal structures after each stage of reconstruction. The reconstructed tissue label volume is shown in **Figure S4**. **f**. The outer (pial) and inner (subplate) cortical plate boundaries were extracted as surface tessellations. The μBrain cortical labels were projected onto the surface vertices to form the final cortical atlas (see **Figure S4**). Cortical areas correspond to matched LMD microarray data (**Table S4**; **Figure S4d**). **g**. Partial reconstructions of *EOMES*, *FOXP1* and *GRIK2* ISH data. ISH stained sections were registered to nearest Nissl-stained sections and aligned to the μBrain volume. Top row: selected axial and coronal sections of the μBrain volume and corresponding tissue labels with ISH expression of three developmental genes: *EOMES, FOXP1* and *GRIK2* overlaid. Expression intensity was derived from false-colour, semi-quantitative maps of gene expression. Bottom row: average expression intensity within each tissue or brain structure based on μBrain tissue labels. Averages were calculated only within sections where ISH was available for each gene.

Tissue- and region-specific gene expression in the mid-gestation brain

We sought to characterise patterns of gene expression in the mid-gestation brain and identify developmental and region-specific genes with putative roles in cortical expansion. To do so, we used publicly-available microarray data from four prenatal brain specimens aged 15 to 21 PCW.[4] Microarray probe annotations were updated and tissue samples matched to the μBrain atlas (**Table S4**) yielding expression data of 8771 genes sampled from between 18 and 27 brain regions and across 5 tissue zones for each specimen (see **Methods**; **Figure S4d**). Applying PCA to gene expression profiles, we found that tissue samples were primarily separated according to location in mitotic (VZ, SVZ) or post-mitotic tissue zones, rather than across regions (**Figure 2a**)[4] – a pattern that was replicated across all specimens when analysed separately (**Figure S6**). Focusing on expression profiles within each tissue zone, samples clustered according to maturity (**Figure 2a**; **Figure S7**) with developmental changes in gene expression most similar across adjacent mitotic (CP and SP, r= 0.67) and post-mitotic zones (SVZ and VZ, r=0.43; **Figure S8**). In line with evidence of a transition in VZ cell fate around mid-gestation,[75] we observed increased expression of genes enriched in post-mitotic excitatory neurons and interneurons (e.g.: *GRIK1-3*; *GLRA2*; *SCN3B*) between 15 and 21 PCW.[1] In the SP and CP, this transition coincided with an increase in genes expressed by radial glia (*BMP7*; *SOX3*) and oligodendrocyte precursor cells (OPCs; *CA10*) with a transitory decrease in microglia-enriched genes in the CP (*GPR34*; *TREM2*)[76] (**Figure 2b; Table S5**).

Across all tissue samples, we tested for differences in gene expression across zones (CP, SP, etc.), regions (motor, sensory, etc) and timepoints (early vs mid-gestation). This resulted in a subset of n=2145 (24.5%) genes with differential expression across all three factors, termed Zone-Region-Tissue (ZRT) genes (p<0.01 after FDR correction; **Figure 2c; Table S6)**. We reasoned that this subset, characterised by genes with dynamic regional and temporal expression in mid-gestation, would be associated with differential rates of cortical expansion during development. To support this line of reasoning, we found that the ZRT cluster was enriched for genes upregulated in the third trimester[77] (enrichment ratio = 1.89, hypergeometric test $p_{hypergeom}$<0.0001) and highly expressed in adolescent and adult brain tissue, compared to non-ZRT genes (**Figure S9**). ZRT genes included several human transcription factors (e.g.: *EGR1*, *JUNB*, *ZNF536*)[78] and were significantly enriched in radial glia (*SOX2*, *HES5,* $p_{hypergeom}$=0.03176), OPCs (*OLIG1*, *PDGFRA;* $p_{hypergeom}$=0.0424) and migrating interneurons (*CALB2*, *CNR1*; $p_{hypergeom}$=0.0009*;* **Figure 2d; Table S7**).[2] We observed highest ZRT expression in the subplate, with increasing expression of ZRT genes in postmitotic zones (CP, SP and IZ) compared to the SVZ and VZ, between 15 and 21 PCW (**Figure 2e**). Examining ZRT gene annotations revealed enrichment of critical neurodevelopmental functions including cell-cell adhesion (GO: 0098742;

5

CHD1, *EFNA5*, *NLGN1*, *NRXN1*; $p_{FDR}<0.0001$, background set = 8771 genes), forebrain development (GO: 0030900; *CASP3*, *CNTN2*, *DLX2*, *FOXP2*, *NEUROD6*; $p_{FDR}=0.026$) and neuron projection guidance (GO: 0097485; *EFNA2*, *EFNA5*; $p_{FDR} = 0.0034$) (**Table S8**). The ZRT geneset was additionally enriched for high-confidence ASD-linked genes (n = 43, $p_{hypergeom} = 0.034$)[79] including *SCL6A1*, *CACNA1C* and *CHD7* and pathogenic variants in 161 ZRT (7.5%) genes have been linked to neurodevelopmental and cognitive phenotypes and brain malformations[80] including *MAGEL2* (Schaaf-Yang syndrome[81]), *AFF2* (Fragile-X-E[82]) and *ADGRG1* (polymicrogyria[83]) (**Table S9**)
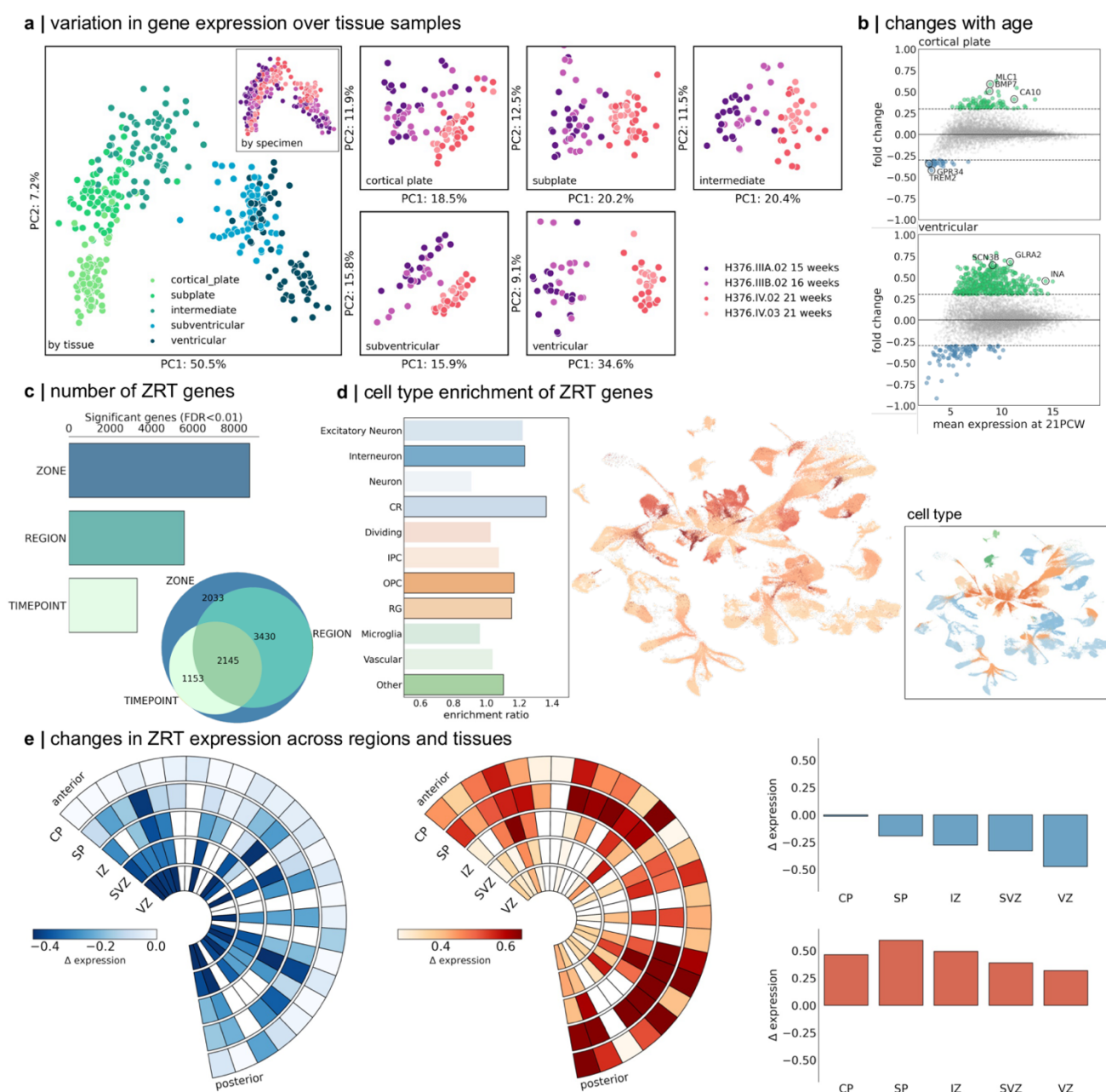

Figure 2: Regional gene expression in the mid-gestation fetal brain. a. PCA of LMD microarray data (n=8771 genes) in four prenatal brain specimens aged 15 PCW to 21 PCW. All tissue samples are shown (left) coloured by tissue zones (main) and specimen (inset). PCA was applied to all samples in each tissue zone separately (right). Samples are coloured by specimen and cluster by age. **b.** PC1 was associated with age-related change in all tissues and correlated between neighbouring zones. Plots show mean gene expression at 21 PCW (averaged over specimen and region) against fold change in gene expression between 15/16 PCW and 21 PCW for two tissue zones (cortical plate, top and ventricular zone, bottom). Genes with a log2(fold change) > 0.3 are shown in green (<-0.3 in blue). Representative genes are highlighted. **c.** Number of genes with differential expression over tissue zones (ZONE), cortical region (REGION) or timepoint (TIME). Venn diagram shows overlap of gene sets. In total, n=2145 were differentially

expressed across zone, region and time (ZRT genes). **d.** enrichment of ZRT genes in cell types previously identified in the mid-gestation fetal brain (left).[2] UMAP projection of cell types showing enriched clusters of OPCs and radial glia. Inset: UMAP projection coloured by cell type. **e.** ZRT gene expression over time and region. Wedge plots (left) show the pattern of expression of ZRT genes that decrease (left) or increase (right) between 15 and 21 PCW. Rows indicate tissue zones and columns indicate cortical regions ordered from anterior to posterior poles. Boxes are coloured by change in gene expression over time (Δ expression). Right: bar charts show mean change in gene expression for decreasing (top) and increasing (bottom) ZRT genes averaged within tissue zones.

Regional differences in the rate of cortical expansion *in utero* during the third trimester

We hypothesised that the dynamic temporal and regional patterning of ZRT genes across tissue zones support differential rates of areal expansion across the cortex. To test this, we acquired n=240 motion-corrected fetal brain MRI scans from 229 fetuses aged between 21^{+1} and 38^{+2} gestational weeks[+days] as part of the Developing Human Connectome Project (dHCP).[84] Volumetric T2-weighted scans were automatically reconstructed to 0.5mm isotropic voxel resolution[85,86]; then tissue segmentations were initially extracted using neonatal protocols[87], followed by extensive manual editing to ensure accuracy (**Methods**). Manually-corrected segmentations were then used to generate cortical surface reconstructions (**Figure 3a**).[88] For analysis, individual cortical surfaces were aligned to a fetal spatiotemporal atlas using a nonlinear, biomechanically-constrained surface registration (Multimodal Surface Matching [MSM]; **Figure 3a-c**).[89–92] At each stage, outputs were visually quality-checked and any failures removed. In total, data from n=195 scans in 190 fetuses (gestational age: 21^{+1} - 38^{+2} weeks; 88 female) were included in the analysis (**Figure S10**).

As expected, total cortical surface area increased exponentially between 21- and 38-weeks gestation (β_{age}=0.054, p<0.001; **Figure 3c-d**).[93–95] While cortical surface area was moderately greater in males compared to females (β_{male}=0.011, p=0.002), this relationship did not change with age (p=0.946). At each vertex in the cortical surface mesh (n=30,248, excluding midline regions), we modelled areal expansion with respect to total surface area using log-log regression (see **Methods; Figure 3e**).[42] Hyperallometric expansion, occurring at a rate faster than the cortical surface as a whole, was observed across the lateral neocortical surface encompassing the fronto-parietal operculum and (granular) insula, primary motor and sensory cortex as well as dorsal parietal and frontal cortices, confirming previous observations based on fetal MRI and scans of preterm-born infants (**Figure 3c, e**).[95–98] In line with proposed models of cortical evolution and expansion,[5,9] slower rates of growth were observed in medial allocortex (including entorhinal, paleocortex and parahippocampal cortex) and the cingulate cortex (**Figure 3e**). The inclusion of sex and sex:age interaction effects in the scaling model did not affect estimated vertex scaling coefficients (r = 0.996). We confirmed that estimates of cortical expansion from cross-sectional analysis aligned closely to longitudinal observations from a single fetus scanned three times during gestation (**Figure S11**).

We calculated the non-parametric correlation (Kendall's τ) between regional estimates of ZRT gene expression in the cortical plate and subjacent tissue zones and average allometric scaling in each of cortical areas defined by the μBrain atlas (**Figure 3e**). In total, across both early and mid-gestation timepoints, expression of 433/2145 (20.1%) ZRT genes was spatially correlated with areal expansion during gestation in at least one tissue zone (ZRT$_{scaling}$; n=542 significant associations, p_{FDR} < 0.01) (**Table S9, S10**). Associations with areal scaling were significantly more common in ZRT genes than in non-ZRT genes (ZRT: 20.1%, non-ZRT: 8.3%; odds ratio=2.78, p<0.0001) with most significant ZRT$_{scaling}$ associations (414/542) localised to the CP (**Figure 3f; Table S10**). ZRT$_{scaling}$ genes in the CP included known molecular correlates of areal identity (*EFNA5*,[99] *GLI3*,[100] *FGFR2*[101]) and axonal guidance (*SLIT1*, *ROBO3*, *SRGAP1*).[102]

Differential expression of ZRT$_{scaling}$ genes largely captured differences between post-mitotic allocortex and neocortex, reflecting opposing allometric scaling across phylogenetic cortical types (**Figure 3e**). We found evidence at 15 PCW, but not at 21 PCW, that genes with higher expression in slower-expanding allocortex and peri-allocortex, were significantly enriched in early-born Cajal-Retzius neurons (e.g.: *CALB2*; overlap=17, enrichment=1.72, p$_{hypergeom}$=0.021),[2] cells that originate from the pallial-subpallial boundary and cortical hem and migrate tangentially across the developing neocortex in early gestation.[103,104] ZRT$_{scaling}$ genes involved in Notch signalling (*NOTCH2NLR*, *JAG1*)[105] and others critical for hippocampal dendritic development (*LRIG1*)[106] were also expressed highly in allocortical regions (**Table S10**). In contrast, ZRT$_{scaling}$ genes expressed in the preferentially expanded neocortex were enriched in progenitor cells at 15 PCW (*FBXO32, HES6*; IPC enrichment = 1.51, p$_{hypergeom}$ = 0.027), and general markers of deep layer neurons at both timepoints (*NEUROD6*, *SYT6;* 15 PCW: Neuron enrichment = 1.53, p$_{hypergeom}$ = 0.004; 21 PCW: enrichment = 1.51, p$_{hypergeom}$ = 0.007). While basic cell types are generally conserved across cortical areas,[107] previous evidence has shown that regional identity is imprinted during cell differentiation, with areal signatures most apparent in post-mitotic cell types but pervasive even at early stages of development across major brain structures.[2,15,31] In line with this, we found opposing enrichment of postmitotic neuronal markers specific to allocortex and neocortex in hypoallometric and hyperallometric ZRT$_{scaling}$ genes, respectively (**Figure 3g**).
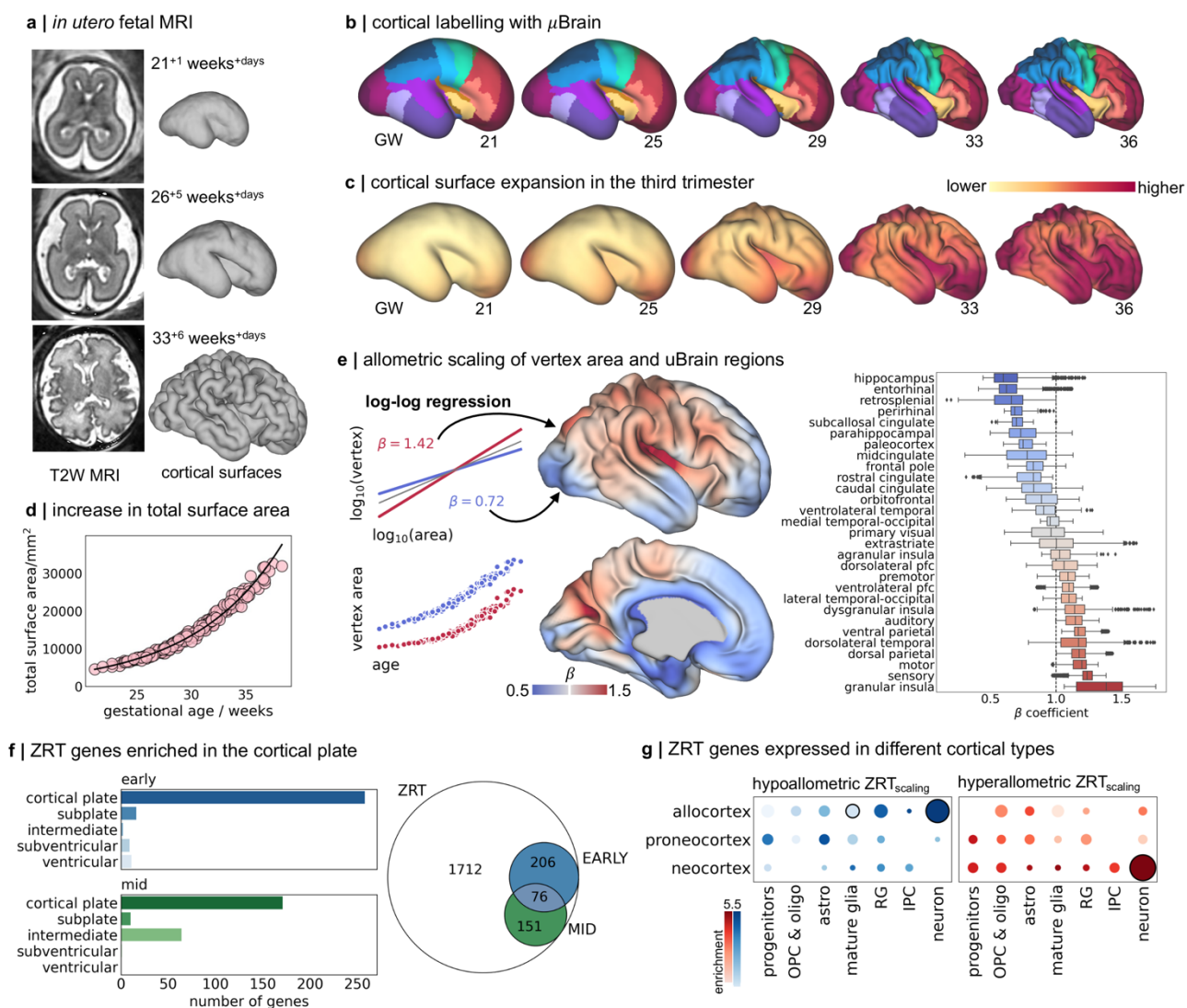

Figure 3: Preferential cortical expansion during the third trimester. a. n=195 fetal MRI scans were acquired during the third trimester of pregnancy. T2-weighted (T2W) scans were reconstructed using a motion-robust processing pipeline and used to generate tessellated cortical surface

representations that were aligned to the dHCP fetal surface template **b**. µBrain cortical labels projected onto dHCP fetal template surfaces from 21 to 36 weeks gestation using nonlinear surface registration. Surfaces are scaled to the same size for visualisation. **c.** For each timepoint, weighted average vertex area maps are displayed on the respective surface templates. Fetal cortical area maps were calculated from individual, co-registered and resampled fetal surfaces using a Gaussian kernel (sigma = 1 week). **d**. total cortical surface area calculated across all surface vertices (excluding the midline) as a function of gestational age at scan. **e.** Left**:** Models of allometric scaling were calculated for each vertex, modelling $\log_{10}$(vertex area) as a function of $\log_{10}$(total area)(top). In this framework, β>1 indicates hyperallometric growth (a relative expansion faster than the global rate). Note that a faster growth rate does not necessarily equate to greater total area at any given time (bottom). Middle: Hyperallometric scaling with respect to total cortical surface area ($\beta > 1$) plotted on the 36w template surface representing preferential cortical expansion during development. Right: Distribution of scaling coefficients for all vertices in each µBrain label in **a**, ordered by mean scaling. **f.** Right: In total, expression of 433 ZRT genes were correlated with areal scaling in gestation. Left: Significant associations (Kendall's τ, p_{FDR}<0.01) were observed across both early (15/16 PCW, n=2) and mid-gestation (21 PCW, n=2) timepoints and in all tissue zones. **g.** enrichment of hypoallometric (left) and hyperallometric (right) $ZRT_{scaling}$ genes in cortical-type specific cell markers.[2] Circle size denotes enrichment ratio, significant associations (p<0.05, hypergeometric test) are highlighted with black outline.

An expanded neocortex is a hallmark of the primate brain. A recent transcriptomic survey of the neocortex across primate species identified a set of genes differentially expressed in humans (hDEGS) and located near to genomic regions that are highly conserved across mammals but significantly altered along the human lineage, either through accelerated DNA substitution rates (human accelerated regions; HAR) or deletions (human conserved deletions; hCONDELS).[107-109] We tested whether these genes were associated with human neocortical expansion *in vivo*. We found that $ZRT_{scaling}$ genes were significantly enriched for hDEGs located near HARs (overlap = 37; enrichment=2.09, $p_{hypergeom}$<0.0001) and hCONDELS (overap=17; enrichment=2.0, $p_{hypergeom}$=0.008). Of these, 22 (56%) were expressed more highly in neocortical than allocortical regions, including several cell adhesion molecules (*DSCAM*, *PCDH7, PCDH9, LRFN2*), teneurins (*TENM3*) and ephrins (*EFNA5*), as well as genes with functional links to language acquisition (*FOXP2*) and neurodevelopmental disorders (*MEF2C, AFF2, ZEB2*) (**Table S9**).

Prolonged neural migration precedes faster expansion across the neocortex

Focusing further on neocortical expansion, we removed allo- and transitory periallo-cortical structures (hippocampus, retrosplenial cortex, entorhinal cortex and paleocortex) and repeated our regional correlation analysis over all ZRT genes. Within the neocortex, a subset of 116 ZRT genes (including 113 $ZRT_{scaling}$ genes) were significantly associated with differential rates of expansion across neocortical regions (ZRT_{neo}; p_{FDR}<0.01), with most associations localised to the intermediate zone (IZ; **Figure 4a; Table S10**). ZRT_{neo} genes were also enriched for hDEGS located near HARs (overlap= 10; enrichment=2.03, $p_{hypergeom}$=0.028) including *PCDH7*, *PCDH9*, *TENM3* and *AFF2* but not hCONDELS (**Table S9**).

We observed contrasting cell type enrichments of ZRT_{neo} genes at 15 and 21 PCW. Consistent with role of prolonged radial glial proliferation in proposed models of cortical expansion,[11,51,54] highly expressed ZRT_{neo} genes in areas with a higher rate of expansion over gestation were enriched in radial glia and intermediate progenitors at 15 PCW ($p_{hypergeom}$=0.045, 0.040 respectively; **Figure S12; Table S11**) with significant associations localised to the cortical plate, subplate and subventricular zone (**Figure 4a**). Early hyperallometric ZRT_{neo} genes are upregulated in both outer (*CDC42EP4*, *HS6ST1*) and ventricular (*FBXO32*) radial glial subpopulations[53] (**Figure S12**). In contrast, ZRT_{neo} genes expressed in neocortical areas with slower relative growth were localised to the cortical plate and subplate but not specifically enriched for any major cell types (all $p_{hypergeom}$>0.05; **Table S11**). However, hypoallometric ZRT_{neo} genes were expressed by neurons (*NFE2L*) and involved in dendritic (*ABGRB3*[110]) and synaptic (*NPTX2*[111]) plasticity, indicative of a population of maturing, not proliferative, cells with neuronal lineage in these regions.

At 21 PCW, after the peak period of neurogenesis, we observed the opposite pattern of cell type enrichments. ZRT_{neo} genes expressed in the IZ subjacent to preferentially expanded cortical areas were enriched in neuronal populations (enrichment = 2.19, $p_{hypergeom}$ = 0.00011) (**Figure 4a-c**) whereas, hypoallometric ZRT_{neo} genes were enriched in proliferative glial cell types (IPC: enrichment=2.96, $p_{hypergeom}$ <0.0001; RG: enrichment=2.36, $p_{hypergeom}$ <0.0001; **Figure 4b,c; Table S11**). The presence of post-mitotic neuronal markers in the IZ at 21 PCW suggested that neuronal migration was ongoing in cortical areas with the fastest rate of expansion in the third trimester of gestation. This is consistent with a conserved mechanism of mammalian cortical expansion whereby longer neurogenic periods lead to an expanded neocortex.[11,112–115] In this context, on both phylogenetic and ontogenetic scales, later developing cortical regions would exhibit faster rates of expansion.[51,115,116] A prominent hypothesis of neocortical expansion has suggested that, in primates, this process is realised through the continued production of upper layer neurons from outer radial glia (oRG) populations situated in the outer SVZ, a cell population greatly expanded in the primate brain.[51,54]
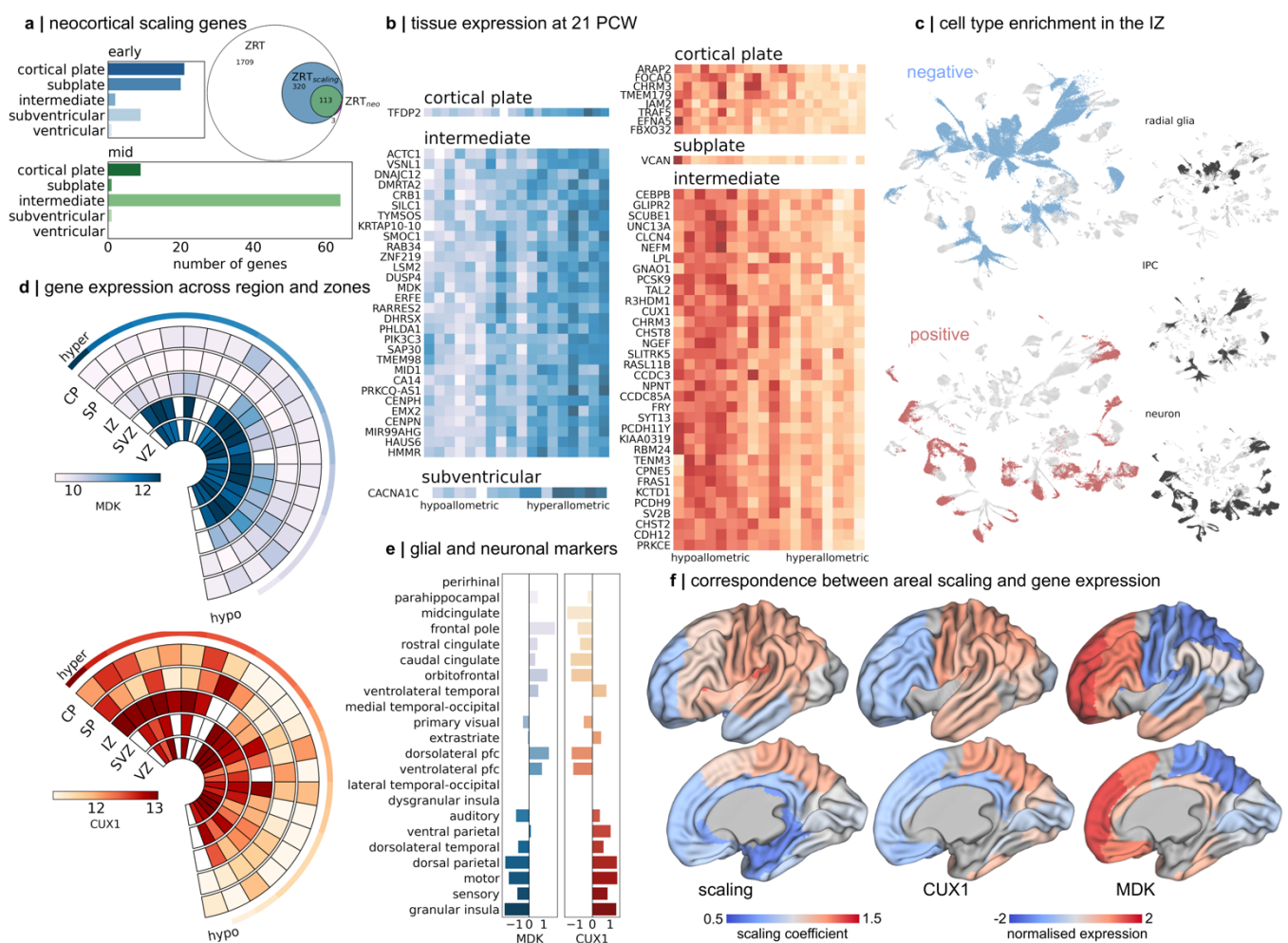

Figure 4: Preferential neocortical expansion is associated with differential timing of neurogenesis and gliogenesis. a. 133 ZRT genes were associated (p_{FDR}<0.01) with areal scaling of the neocortex (after excluding paleo- and archi-cortex; ZRT_{neo}). Most significant associations were localised to the IZ. **b.** normalised (Z-score) expression profiles for genes correlated with areal scaling in each tissue zone at 21 PCW. Associations at 15 PCW are shown in **Figure S12**. Negative associations (higher relative expression in hypoallometric regions) shown in blue, positive associations are in red. Lighter colours indicate higher relative expression. Most significant associations are in the IZ. **c.** Mid-gestation cell clusters[2] significantly enriched (p<0.01) for genes associated with areal scaling in the IZ at 21 PCW. Territories of three cell types are shown. Negative and positive ZRT_{neo} genes are enriched in progenitor cells and neurons, respectively **d.** wedge plots are shown for two ZRT_{neo} genes expressed by specific cell types: *MDK* (glial) and *CUX1* (upper layer neurons). Rows indicate tissue zones and columns indicate cortical regions ordered according to allometric scaling from

10

hyper to hypoallometric. Colour bar indicates normalised expression levels (a.u.). **e**. expression (Z-score) of *MDK* and *CUX1* in all regions sampled in the IZ, ordered from hypo (top) to hyperallometric (bottom) scaling. **f**. IZ expression of *CUX1* (middle) and *MDK* (right) projected onto corresponding μBrain surface atlas labels and displayed on the 36w dHCP template surface. Regions where expression for a given gene was not available are shown in grey. For comparison, average allometric scaling in each region is displayed (left).

To examine this proposed mechanism in humans, we focused on *CUX1*, a marker of layer III/IV neurons that regulates dendritic morphology[117] and is expressed highly in preferentially-expanded cortical regions (**Figure 4d-f**). *CUX1* is located downstream of HAR426 and pathogenic mutations in *CUX1* are associated with ASD, intellectual disability and epilepsy.[118,119] We find that, in the IZ at 21 PCW, *CUX1* exhibits expression that varies along a hypo-to-hyperallometric gradient (**Figure 4d,e**; τ=0.52, p_{FDR}=0.002). To validate these observations, we examine ISH staining of a second upper layer marker, *SATB2*, in five regions with differential allometric scaling, finding examples of upper layer SATB2+ neurons within the IZ of regions with a faster rate of expansion in mid- to late-gestation (**Figure S13**). The prolonged migration of these cell populations in expanding neocortical regions is a potential consequence of differential neurogenic timing across the neocortical sheet that, at least in part, supports the accelerated expansion of hyperallometric cortical regions during gestation.

Several mechanisms exist to regulate gene transcription during early brain development.[120,121] To identify potential regulators of ZRT gene expression in the developing fetal cortex, we used a recent chromatin accessibility atlas[122] to examine the position of open chromatin regions (OCR) in the mid-gestation brain relative to ZRT genes. We found that ZRT genes were more likely than non-ZRT genes to be located near to predicted regulatory elements (pREs), a subset of OCRs that are likely to function as neurodevelopmental enhancers in mid-gestation[122] (OR: 1.38 p < 0.0001; **Figure S14**; **Table S12**). Moreover, $ZRT_{scaling}$ and ZRT_{neo} genes were significantly enriched for genes located near to pREs (enrichment = 1.25, 1.31 $p_{hypergeom}$<0.0001, <0.005 respectively; **Table S12**). Focusing on laminar specificity of ZRT gene expression, we found that over 25% of ZRT_{neo} genes were located immediately up- or downstream of OCRs specific to the upper layers of the cortical plate, compared to 9% located near to deep layer OCRs (**Figure S14B**). Transcription factor motifs contained within OCRs specific to upper cortical layers and proximal to ZRT_{neo} (n=20) included bHLH, LIM and POU homeobox and HMG-box motif families (**Figure S14C**) that bind to transcription factors which regulate superficial neuronal identify (e.g.: *E2A*, *BRN1*, *LHX2*).[123–125] Thus, the differential accessibility of specific regulatory elements can resolve the laminar identity of maturing upper-layer excitatory neurons migrating through the IZ at 21 PCW.

Based on this evidence, we reasoned that neuronal migration, and thus neural proliferation, would be complete or near complete at 21 PCW in neocortical areas with slower expansion rates in the third trimester. In this case, expression of proliferative cell markers (**Figure 4**) would reflect gliogenesis rather than neurogenesis. To test, we compared ZRT_{neo} genes associated with cortical scaling at 21 PCW in the IZ to region-specific cell type signatures in the mid-fetal brain.[2] Reflecting the proximity to medial allocortex and periallocortical regions, we identified several midline identity genes (*MID1*, *DMRT5*, *EMX2*) with high expression in hypoallometric cortex as well as markers of cell proliferation (*HMMR*, *HAUS6*, *CENPN*, *CENPH*, *PIK3C3*) (**Figure 4b; Table S10**). In support of our hypothesis,, we found that hypoallometric ZRT_{neo} genes were specifically enriched in (peri)allocortical glial cell populations (*MDK*, *SAP30*, *TMEM98;* astroglia, $p_{hypergeom}$=0.01; OPC, $p_{hypergeom}$=0.07).

TMEM98 is a MYRF-interacting protein specifically expressed in newly-differentiated oligodendrocytes in the developing central nervous system[126] whereas MDK is a growth factor expressed in pre-OPCs that can induce differentiation in oligodendrocyte-lineage OL1+ cells *in vitro*.[127–129] *SAP30* forms a co-repressor complex with *HDAC1*

11

and *HDAC2*, class I histone deacetylases that regulate gene transcription and are essential for oligodendrocyte maturation.[130–133] Similar negative correlations with cortical expansion were recorded in OPC cell population markers S100B (**Table S10**; τ=-0.33, p_{FDR}=0.070), NKX2-2 (τ=-0.41, p_{FDR}=0.019) and the glial progenitor marker *EGFR* (**Figure S15**), which has been validated previously as a critical gene related to brain size.[134] In an independent dataset,[33] we observed similar trends in *OLIG1* expression in mid-gestation across cortical regions with differential developmental expansion (**Figure S16**). Overall, these data suggest that the developmental timing of the neuro- to gliogenic switch varies across the neocortical sheet, with the length of the neurogenic period supporting differential rates of neocortical expansion during the third trimester of gestation.

Neocortical scaling genes are critical for typical neurodevelopment

Given their likely importance in shaping early normative neurodevelopment, we hypothesized that the ZRT_{neo} genes would be susceptible to severely disruptive mutations (i.e., loss-of-function variants). We found significant enrichment of hyperallometric (median loss of function observed/expected upper bound fraction (LOEUF) score = 0.26, permutation p = 0.0003 using random gene sets of similar size: $p_{permutation}$) but not hypoallometric (median LOEUF score = 0.40, $p_{permutation}$ = 1) ZRT_{neo} genes, suggesting a disproportionate level of vulnerability to loss-of-function variation in genes whose expression is greater in areas that expand fastest in the third trimester. Within these, we identified a set of constrained genes expressed highly in the subventricular zone at 15 PCW in hyperallometric regions. These genes are involved in extracellular matrix formation and interaction (*EFEMP2*, LOEUF=0.56, *PTPRM*, LOUEF=0.33), and epithelial-to-mesenchymal transition (*FBXO32*,[135] LOEUF=0.64), pathways crucial to outer radial glia specification and differentiation in germinal zones of the developing brain.[53] Follow-up analyses using genome-wide metrics for dosage sensitivity[136] confirmed the enrichment of hyperallometric ZRT_{neo} genes as haploinsufficient (62% of genes, $p_{permutation}$ < 0.0001 using random gene sets of similar size) and not triplosensitive (19%, $p_{permutation}$ = 0.9418) – a highly pathogenic mechanism for loss-of-function mutations.

To assess the clinical relevance of these distinct ZRT_{neo} gene sets (i.e., hypoallometric and hyperallometric), we performed enrichment analyses using MAGMA[137] across an array of previously published genome-wide association studies (GWAS). We found that ZRT_{neo} gene sets were not enriched for birth outcomes (gestational duration) or cognition (educational attainment), but hypoallometric ZRT_{neo} genes were enriched for externalizing behavior (β=0.17 , p= 0.007) and hyperallometric ZRT_{neo} genes were enriched for schizophrenia (SCZ; β=0.17, p=0.004). Further analysis using postmortem gene expression data from patients with neurodevelopmental disorders revealed significant enrichment of ZRT_{neo} gene sets within multiple co-expression modules.[138] Both hypoallometric and hyperallometric ZRT_{neo} genes were enriched in cross-disorder module *CD1* (both $p_{permutation}$ < 0.05) – downregulated in autism spectrum disorder (ASD), SCZ, and bipolar disorder, and containing neuron-enriched genes and genes with ASD- and SCZ-associated nonsynonymous de novo variants from whole-exome sequencing; and hyperallometric ZRT_{neo} genes were enriched in module *CD13* ($p_{permutation}$ < 0.05) – also downregulated in ASD, SCZ, and bipolar disorder, and containing neuron-enriched genes.

Discussion

Despite the altriciality of the human brain at birth, areal expansion of the cortex during the second and third trimester of gestation is critical for later neurodevelopmental function. Cortical surface area increases exponentially during the third trimester of gestation, permitted by rapid cortical folding over the same period. Powered by a new 3D atlas of the developing brain, our results provide a multiscale understanding of fetal cortical expansion in the second half of

pregnancy. We find that differential expansion of cortical areas in gestation respects anatomical and evolutionary boundaries between cortical types[5,9] and is supported by an extended period of neural migration through mid-gestation.[51,52,54]

Neurogenesis exhibits a conserved order but nonlinear scaling across species.[139] Longer neurogenic periods in larger-brained species, supported by a larger pool of progenitors in proliferative zones, result in the preferential expansion of later developing structures.[11,112–114,139] In mammals, differences in the timing and rate of neuron production vary across cortical areas with evidence to suggest that progressive termination of cortical neurogenesis occurs along a rostral-caudal axis.[140–142] In this case, earlier termination of neuronal production in the anterior cortex could create a potential affordance for increased neuronal size and arborisation, leading to increased areal expansion during development.[25,139,143] However, further evaluation of areal differences in neurogenic timing in the primate cortex presents a more complex picture, with neurogenesis terminating first in limbic and allocortical structures but continuing in the prefrontal cortex beyond mid-gestation.[139,144,145] Coupled with the nonlinear progression of human gyrification over gestation,[146] this suggests that areal differences in cortical scaling are likely founded upon an alternative schema.[49,51]

Alternative hypotheses have been put forward on the role of oRG proliferation, and the prolonged production of neurons or glia, in cortical expansion.[49–52] Our findings demonstrate that prior to gyrification, but after the peak period of neurogenesis, supragranular neurons continue to migrate to neocortical areas with the fastest rate of expansion in the third trimester. In the primate brain, oRGs produce large numbers of upper layer neurons, provide a scaffold for neural migration and, upon completion of neurogenesis, act as a source of glial cells in mid- to late-gestation.[49,52,54,147] Thus, regulation of neuro-to-gliogenic timing in the oRG subpopulation may represent a plausible candidate for differential rates of neocortical expansion[49,147] Though present in other mammals, the oSVZ is expanded in primate species[51,53,54] and proliferation in the oSVZ, marked by mitotic activity, is highest in regions that expand most in later development.[148] While our data suggest rapid areal expansion is preceded by an extended neurogenic period, we lack the data to confirm a similarly extended period of gliogenesis. In humans, neurogenesis precedes cortical folding with the subsequent gliogenic period more closely aligned to the timing of cortical expansion.[49,50] An extended neurogenic period coupled with a longer migration time due to the expanding volume of the brain may necessitate an extended gliogenic period to populate the expanding neuropil.[50] Evaluating the temporal and spatial regulation of glial fate transition and proliferation in the oSVZ during the second half of gestation represents a critical next step in understanding this process.

In gyrencephalic species, the buckling and folding of the cerebral cortex allows for increased surface area of the cortical grey matter. Greater tangential expansion of superficial cortical layers relative to subcortical tissue represents a core feature of biomechanical models of cortical growth and folding.[149–152] However, uniform rates of tangential expansion can not fully account for the consistency in location of cortical folds across individuals, with additional genetic contributions to gyral patterning clearly demonstrated in twin studies.[153,154] In contrast, genetically determined areal differences in expansion rate may give rise to the consistent patterns of folding observed across the neocortical sheet.[148,155] Recently, large-scale neuroimaging studies have identified patterns of altered cortical morphometry that are shared across common neuropsychiatric conditions and human genetics studies have begun to converge on putative mechanisms underlying cortical abnormalities in developmental genetic disorders.[156–160] Here, we identify significant enrichment of pathogenic loss-of-function variants in genes that are expressed in mid-gestation, linked to specification of outer radial glia and associated with differential rates of cortical expansion. Taken together, these findings suggest that there are temporal windows of susceptibility in the early stages of brain development where areal

differences in the timing of fundamental neurogenic processes could underlie observable cortical abnormalities and postnatal functional pathologies in neurogenetic disorders.[47]

Spatially-embedded gene expression atlases of the adult human[35,68] and mouse[70,161] have proven exceptionally powerful in recent years, bridging resolution gaps to common neuroimaging modalities[162,163] and providing insight into the molecular correlates of structural[26,42,164] and functional neuroanatomy,[36,165,166] brain development,[167–169] disease and disorder.[47,170,171] In such studies, comparisons with *in vivo* neuroanatomy can only be fully realised through three-dimensional localisation of tissue samples within a common coordinate space.[62,68,70] To date, a limitation of this approach has been either the sampling of a narrow age range outside of key developmental periods[35,69] or, in developmental datasets, a lack of 3D spatial information[3] and relatively coarse anatomical sampling.[34] To fill this gap, we provide a new resource, μBrain, built upon existing open-source data, to allow researchers to map developmental neuroanatomy of the human fetal brain onto early histogenic processes using contemporaneous *post mortem* data. The reconstructed 3D μBrain atlas brings detailed tissue microarray and *in situ* hybridisation data into alignment with a developmental anatomical atlas of the fetal brain.[89] The μBrain atlas will enable future studies to examine tissue or region-specific expression signatures in relation to aspects of structural or functional brain development *in utero* or identify spatial or temporal windows of vulnerability for genetic or neurodevelopmental disorders.

Developmental MRI studies provide unique insight into early human brain development. Due to large differences in size, shape and tissue contrast, specialised tools are required for the analysis of infant and neonatal MRI. Similarly, we cannot rely on common cortical atlases that are based on adult neuroanatomy.[87,172] Here, we used annotations derived from cytoarchitecture of the mid-fetal brain to generate a new cortical atlas to facilitate further research in early brain development. A key area for future research in this field is the development and validation of improved methods to align early MRI to common template spaces. The geometry of the fetal cortex is smooth, making alignment of cortical morphometry an ill-posed problem. Newer, anatomically-constrained registration techniques and larger longitudinal cohorts with multiple scans during mid- to late-trimester will enable more precise estimates of cortical expansion in the future.[90,95]

With increasingly granular surveys of the developing brain at a single-cell level[2,107] the advent of spatial transcriptomic technologies,[173] and a series of large-scale and open-access perinatal neuroimaging studies,[84,174,175] we anticipate μBrain will provide a foundation for developmental and comparative neuroscience to integrate and transfer knowledge of early brain development across domains, model systems and resolution scales.

Acknowledgements

Author contributions

G.B., S.O. V.K., L.Z.J.W., V.K., A.P., J.V.H., J.H., E.C.R. & J.S. performed data acquisition and data processing. G.B., S.O. & J.S performed data analysis. G.B, J.S., V.K., L.Z.J.W & E.C.R. contributed to methodology. G.B., M.L.S., A.A-B., J.V.H., A.D.E., E.C.R. & J.S. provided resources and supervision. E.C.R. provided software. Project conceptualisation: G.B & J.S. Writing drafts, revisions and editing: all authors.

Conflicts of interest

JS and AFA-B are co-founders of Centile Bioscience.

Material and Methods

Public data sources

Source data underlying the μBrain atlas were made available as part of the BrainSpan Developing Brain Atlas [https://atlas.brain-map.org/atlas?atlas=3] with detailed tissue processing protocols available from Ding et al.[3] In brief, a single prenatal brain specimen (21 PCW; female) was bisected and the right hemisphere used for serial sectioning. The brain specimen was cut into four coronal slabs and frozen in isopentane. Serial coronal sectioning at 20μm thickness was performed slab-by-slab with sequential sections submitted to Nissl, AChE or ISH staining with 43 gene probes and stained sections digitally scanned at 1μm / pixel resolution. In total, 81 out of 174 Nissl-stained sections with varying sampling densities (~0.5mm to 1.2mm between sections) were selected for annotation.[3] Expert anatomical annotations were conducted manually on each section. Nissl- and ISH-stained sections with corresponding anatomical labels were made available for download. Anatomical annotations were also used to guide laser microdissections for DNA microarray analysis across the developing cerebral tissue in the left hemisphere of 4 separate mid-gestation specimens (see **Microarray Data** below). The section numbers and approximate coronal positions of sections used in the construction of the 3D μBrain atlas are listed in **Table S1**.

Image processing

We downloaded each high-resolution Nissl-stained section (n=81; downsampled to 2μm/pixel) as RGB images in JPG format with corresponding anatomical labels as SVG files.

After converting SVG to RGB PNG format, we manually combined anatomical labels according to the hierarchical ontology of the reference atlas[3] to create two compact annotations, one for image repair comprising 20 tissue structure labels (*brain-labels*) and one for statistical analysis containing only cortical labels (*cortex-labels*, n=30, including one generic 'brain tissue' label for non-cortical structures; see **Table S2**). Due to the small size and degree of missing data precluding reconstruction, marginal zone and subpial granular zones were not considered in this analysis. Nissl-stained sections and corresponding label images were then downsampled to 20μm /pixel resolution.

Histological reconstruction

Pix2pix is a conditional generative adversarial network (GAN) trained to perform image-to-image translation between pairs of image examples.[59] We used the *pix2pix* architecture (**Figure 1b**) to synthesise Nissl-stained images from label annotations in order to replace artefacts within tissue sections (**Figure 1c,d**). Following conventional GAN structure, the model combines a generator network, G, with a classifier (or discriminator, D) whose objective is to determine if images are real or fake (**Figure 1b**). GAN training is performed adversarially with the generator network competing to generate more and more realistic synthetic image from label annotations, and the discriminator working to discriminate between real and fake examples. An L_1 regularisation term is added to enforce that generated images are as close as possible to the ground truth. Full model architecture and training details are included in the **Supplemental Methods**.

Image repair

To perform repair of whole sections, we split each label image into patches of 256×256 pixels with an 8 pixel overlap and passed them through the trained generator. The resulting, synthetic Nissl contrast patches were stitched together into a full section matching the dimensions of the original image (**Figure 1d**). Patch prediction and image reconstruction was performed using MightyMosaic [https://pypi.org/project/MightyMosaic/].

To detect regions of the original Nissl-stained section that needed repair, we designed an automated outlier detection method based on the Median Absolute Deviation (MAD) of pixel hue and saturation. The original Nissl-stained sections and corresponding GAN-generated predictions were transformed to HSV format and blurred with a box filter (width = height = 5 pixels). We identified outliers with median absolute differences in hue and saturation between pixels in the ground truth image and its synthetic equivalent in hue and saturation greater than threshold, θ, set to 2.5, whereby lowering θ would increase the number of pixels marked as outliers.

For each section, a binary mask was created containing all pixels identified as outliers in both hue and saturation. A final opening operation was applied to the outlier mask using an elliptical filter (iterations = 3, width = 3 pixels) to remove speckles in the mask. Identified outlier pixels were then replaced with the corresponding, intensity-matched pixels from the synthetic image using Poisson image editing to effect image repair (Figure 1d).[61] Outlier detection and repair was performed in Python using OpenCV (4.5.2) [https://opencv.org/].

µBrain volume construction

Following automated repair of major tissue artefacts present in the histological data, we aimed to develop a 3-dimensional reconstruction of the fetal brain to facilitate comparison with *in vivo* MR imaging data. Image alignment and reconstruction steps are summarised below. Full details are included in **Supplemental Methods**.

Slice-to-slice alignment

Using the middle section as a reference, repaired Nissl-stained sections were aligned using a graph-based, slice-to-slice registration.[176,177] Pairwise rigid transforms were estimated between each section and its neighbouring sections in the direction of the reference. Dijkstra's shortest-path algorithm was then used to calculate the set of transforms with lowest cost to align a given section to the reference.[176,177] The selected transforms were composed and applied to both the image and its corresponding labels to bring all sections into approximate alignment (**Figure 1e**; **Figure S17a**).

Affine registration to a fetal brain shape reference

Reconstructing 3D volumes from the consecutive alignment of 2D sections commonly produces an artefact termed 'z-shift' caused by the propagation of registration errors between adjacent slices and resulting in a distorted three-dimensional structure in the final volume.[178] To overcome this effect, it is common to use a shape prior to guide

registration and preserve 3D shape.[62,178,179] In lieu of a ground-truth volume for the sectioned data, we employed a population-based average anatomical image: specifically the 22-week timepoint of the Gholipour et al. spatio-temporal fetal MRI atlas (**Figure S3**).[63]

After matching MRI-based tissue labels to the μBrain tissue labels, we upsampled the MRI template to 50μm isotropic resolution and converted the MRI labels into an image Nissl-like contrast using the trained GAN model (**Figure S3c-d**). Nissl-contrast images were re-stacked into a 3D volume to act as an anatomical prior for registration.

We performed an iterative affine registration procedure between the MRI-based shape prior and the 3D stack of histological sections.[176] This process was repeated for a total 5 iterations, producing a final 3D volume with aligned coronal slices and a global shape approximately matched to the *in utero* fetal brain (**Figure 1e; Figure S17a**).

Final template construction

To create the final 3D volume, we employed a data augmentation technique, generating n=50 unique representations of the affinely-aligned data by applying nonlinear distortions along all three image axes. For each volume, we performed a weighted nonlinear registration between neighbouring sections to account for residual misalignments. Finally, to create a smooth 3D reconstructed volume, we co-registered all 50 augmented and aligned volumes into a single probabilistic anatomical template with voxel resolution $150 \times 150 \times 150$μm using an iterative, whole-brain nonlinear registration (**Figure 1e**; **Figure S17a**; **Supplemental Methods**). All image registration was performed in Python 3.7 using *antspyx* (0.2.7).[14]

Cortical reconstruction

To reconstruct the fetal cortical surface, we adapted existing protocols for *ex vivo* [https://freesurfer.net/fswiki/ExVivo] and non-human primate [https://prime-re.github.io/] surface reconstruction with Freesurfer.[15] We used the μBrain tissue labels to generate a 'white matter' mask (all subcortical structures and tissue zones, excluding the cortical plate). We used this mask to generate inner and outer surfaces for the μBrain volume (**Figure 1f**). Surfaces were smoothed and inspected for topological errors. All processing was performed with Freesurfer (7.3.2).

In situ hybridisation

In addition to serial Nissl staining, interleaved coronal sections were used for *in situ hybridisation* (ISH) of a series of neurodevelopmental marker genes (**Table S3**).[3] High-throughput ISH staining was performed for each gene, with stained sections digitised at 1μm resolution. Quantification of the intensity of expression detection was performed using an automated procedure that pseudo-colour coded levels of expression for visualisation, with low-to-high expression represented as blue-to-red.[161]

Compared to Nissl-stained sections (n=79 after quality control), fewer ISH stained sections were available for each gene (mean n = 41 after quality control), precluding a full 3D reconstruction of each. We downloaded each set of ISH-stained sections and removed any with large artefacts (tearing, folding, missing tissue). From each false-colour expression map, we extracted the red channel to focus only on higher expressing cells. Each section was registered to the nearest, repaired Nissl-stained section using affine registration. Registrations were visually inspected and any failures removed. Aligned sections were then stacked together, with blank slices in place of missing sections and reconstructed into a 3D volume using the previously calculated slice-to-volume alignments for each section (see '**μBrain volume construction**').

Microarray data

We downloaded prenatal LMD microarray data from the BrainSpan database [https://www.brainspan.org/]. For details on tissue processing and dissection see Miller et al.[4] and the technical white paper available at: [https://help.brain-map.org/download/attachments/3506181/Prenatal_LMD_Microarray.pdf]. In total, normalised microarray data from 58,692 probes in 1206 tissue samples were available to download, obtained from the left hemisphere of four post-mortem fetal brain specimens (age 15-21 PCW, 3 female).[4] Each probe was assigned a 'present' or 'absent' annotation based on strength of average probe expression over corresponding background signal. Through comparison with the BrainSpan reference atlas, we matched each tissue sample's anatomical label to i) corresponding cortical labels included in the µBrain atlas and ii) one of five tissue zones (cortical plate, subplate, intermediate zone, subventricular zone, ventricular zone) (**Table S4; Figure S4**). Samples that could not be matched to labeled regions in the cortical plate or corresponding subjacent tissue zones were removed, including samples from subcortical nuclei, midbrain structures and brainstem.

Microarray processing

We updated gene assignments for the Allen microarray probes using Re-Annotator[180] and removed any probes assigned to more than one gene, resulting in a probe set (n=46,156) mapped to 20,262 unique genes. Low signal probes designated 'absent' were removed (34.67% of probes), as were tissue samples from the marginal zone, subpial granular zone and subcortical and midbrain structures (54.46% of samples). Where multiple probes mapped to a single gene, the probe with the highest differential stability (DS),[181] the average pairwise correlation between tissue sample expression over all specimens, was assigned. Probes with DS<0.2 were removed.

Where more than one sample was available for a given region or zone, e.g.: samples from the outer and inner cortical plate in the same region, gene expression was averaged across samples. Finally, any probes with missing data in more than 10% of tissue samples were removed (n=1253). This resulted in expression data from 8771 genes across 27 regions and 5 tissue zones for analysis (**Figure S4**).

Fetal MRI

To measure cortical expansion *in utero* during the third trimester, we analysed high-resolution MRI from a large cohort of fetuses.

MRI acquisition

Fetal MRI datasets (n=240 scans from 229 fetuses aged between 21^{+1} and 38^{+2} gestational weeks^{+days}) were acquired as part of the Developing Human Connectome Project (dHCP) using a Philips Achieva 3T system, with a 32-channel cardiac coil in maternal supine position. Structural T1-weighted (T1w), T2w, functional MRI and diffusion MRI data were acquired for a total scan time of approximately 45 minutes.[85] T2-weighted SSTSE volumes were acquired with TE=250ms, acquisition resolution 1.1 x 1.1mm, slice thickness 2.2mm, -1.1mm gap and 6 stacks. All 3D brain images were reconstructed using a fully automated slice-to-volume reconstruction (SVR) pipeline[86] to 0.5mm resolution and reoriented to the standard radiological space.

The study was approved by the UK Health Research Authority (Research Ethics Committee reference 452 number: 14/LO/1169) and written parental consent was obtained in every case for imaging and open data release of the anonymized data. All data was acquired at St Thomas Hospital, London, United Kingdom.

After image processing and quality control, the final dataset comprised n=195 fetal MRI datasets acquired from n=190 fetuses aged 21^{+1} to 38^{+2} gestational weeks (88 female). Repeated scans were acquired from four fetuses.

MRI processing

While neonatal protocols for automated MRI tissue segmentation exist,[87,182] due to the differences in size, tissue contrast and signal-to-noise ratio, segmentations derived from fetal MRI often require extensive manual editing to ensure accuracy.[183]

Here, we used an optimised neonatal tissue segmentation pipeline (Draw-EM)[87] with tissue priors adapted to a fetal MRI template to create a 'first-pass' tissue segmentation for each fetal MRI volume. Tissue segmentations were then visually checked and extensive manual corrections performed where needed to correct gross segmentation errors and ensure accuracy of tissue boundaries (CSF/cortex/white matter). Manually-corrected tissue segmentations were then used to generate anatomically and topologically correct inner and outer cortical surfaces using *Deformable*.[88] Note that all intensity-based correction terms were turned off during surface reconstruction and each surface was generated using just the corrected tissue segmentations. At each stage, images and derived outputs were visually inspected for accuracy.

Alignment to fetal template

We aligned individual cortical surfaces to the dHCP fetal atlas, a spatiotemporal surface atlas, spanning 21-36 weeks of gestation with weekly timepoints.[89,91] Using MSM with higher-order clique reduction, we calculated non-linear transforms of individual surfaces to their closest fetal timepoint based on spherical registration of sulcal depth features.[90,92] The MSM transform was used to resample individual surface topology (pial, midthickness, and white) onto the template surface vertices, ensuring that all surfaces across individuals had the same vertex correspondence. Resampled surfaces were manually checked to ensure the quality of the registration.

Alignment to μBrain

We aligned the μBrain cortical surface to the earliest timepoint of the dHCP fetal template surface using a two-step nonlinear surface registration guided by a set of anatomical priors (**Figure S17b,c**). We used MSM to perform an initial nonlinear spherical registration between μBrain and dHCP surfaces based on alignment of sulcal depth. After this, we created a set of coarse cortical labels on the dHCP surface matched to corresponding μBrain labels by combining a) dHCP cortical atlas labels,[87] b) manual labels guided by sulcal anatomy on the 36 week fetal surface and c) combining μBrain labels in the same lobes (e.g.: ventrolateral frontal, dorsolateral frontal, orbitofrontal) were into single anatomical labels. The full list of 11 matched cortical regions included: auditory cortex; cingulate cortex; frontal cortex; insular cortex; primary motor; primary sensory; occipital cortex; parahippocampal cortex; parietal cortex; superior temporal cortex; ventrolateral temporal cortex. A secondary multivariate spherical registration between μBrain and fetal surfaces was initialised using the previously calculated sulcal alignment and driven by alignment of cortical ROIs across surfaces.[90] This approach leverages anatomical labels (defined based on cytoarchitecture, or using older fetal anatomy in μBrain and dHCP atlases, respectively), to inform cortical alignment in the absence of geometric features. A similar approach has proven successful accommodating large deformations across primate species.[184]

μBrain labels were propagated to each timepoint of the dHCP fetal atlas (**Figure 3b**) and onto the surface topology of each fetal scan. Cortical labelling was visually quality checked for alignment.

Statistical analysis

Allometric scaling of cortical surface area

Each subject's outer cortical topology was resampled onto the dHCP template surface (32,492 vertices) and vertex-wise estimates of cortical surface area were corrected for folding bias by regressing out cortical curvature[185,186] and smoothed with a Gaussian kernel (FWHM = 10mm). Total cortical surface area was calculated as the sum of all vertices in the cortical mesh, excluding the medial wall. At each vertex, v, we modeled scaling relationships with brain size by estimating the log-log regression coefficient for total surface area as a predictor of vertex area, a_v:[42]

$$log_{10}(a_v) = 1 + \beta log_{10}(\sum_{v=1}^{V} a) + \varepsilon$$

Such that the scaling coefficient, β, can be directly interpreted relative to 1 (representing linear scaling between vertex area and total area) with $\beta > 1$ and $\beta < 1$ representing hyper- and hypoallometric scaling of vertices with respect to total area, respectively. Models were fit using Ordinary Least Squares (OLS) regression. We tested alternative models including sex and age:sex interactions. Analyses were repeated after removing repeated scans to satisfy i.i.d. assumptions of OLS regression (n=190; **Figure S18**).

Prior to analysis, vertexwise outliers were identified and removed (**Figure S19**). To account for age-related increases in area, outliers were identified using a sliding window over age (outliers >2.5 S.D. from the mean within a given window, maximum window size=25 scans, sorted by age). Data from five scans were removed prior to analysis due to the presence of outliers in more than 5% of vertices.

Vertexwise maps of areal scaling (β coefficients) were parcellated using the μBrain cortical labels, calculating average scaling within each parcel for further analysis.

Modelling changes in gene expression over zone (Z), region (R) and time (T)

For each gene (n=8771), we modelled the main effects of cortical tissue zone, region and timepoint on expression using a general linear model. Significant effects (p<0.01) were identified after False Discovery Rate correction for multiple comparisons over genes. Statistical analysis was performed in *statsmodels* (0.13.5)

Enrichment analyses

For all enrichment analyses, we calculated the enrichment ratio as the ratio of the proportion of genes-of-interest within each geneset/marker list to the proportion of background genes within each geneset. Unless otherwise stated, the background set was defined as the full list of genes included in the study (n=8771). Significance was determined using the hypergeometric statistic:

$$p = 1 - \sum_{i=0}^{x} \frac{\binom{K}{i}\binom{M-K}{N-i}}{\binom{M}{N}}$$

Where p is the probability of finding x or more genes from a specific geneset K in a set of randomly selected genes, N drawn from a background set, M. Where stated, False Discovery Rate (FDR) correction was applied to multiple comparisons.

Code and data availability

The μBrain digital template with corresponding cortical surfaces and atlas labels is available from https://garedaba.github.io/micro-brain alongside code supporting data processing and analysis for this manuscript.

All dHCP data, fetal brain reconstructions, brain region segmentations and cortical surfaces are available for download from the NDA https://nda.nih.gov/edit_collection.html?id=3955

Source histological and microarray data are available from the Allen Brain Institute https://www.brainspan.org/

References

1. Polioudakis, D. *et al.* A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron* **103**, 785-801.e8 (2019).

2. Bhaduri, A. *et al.* An atlas of cortical arealization identifies dynamic molecular signatures. *Nature* **598**, 200–204 (2021).

3. Ding, S.-L. *et al.* Cellular resolution anatomical and molecular atlases for prenatal human brains. *J. Comp. Neurol.* **530**, 6–503 (2022).

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

5. García-Cabezas, M. Á., Zikopoulos, B. & Barbas, H. The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Struct. Funct.* **224**, 985–1008 (2019).

6. Brodmann, K. *Vergleichende Lokalisationslehre der Grosshirnrinde*. (Barth, 1909).

7. Sanides, F. The cyto-myeloarchitecture of the human frontal lobe and its relation to the phylogenetic differentiation of the cerebral cortex. *J. Hirnforsch.* **7**, 269–282 (1964).

8. Petrides, M. & Pandya, D. N. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* **11**, 1011–1036 (1999).

9. Barbas, H. & Rempel-Clower, N. Cortical structure predicts the pattern of corticocortical connections. *Cereb. Cortex* **7**, 635–646 (1997).

10. Economo, C. von & Koskinas, G. N. *Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen*. (J. Springer, 1925).

11. Rakic, P. Specification of cerebral cortical areas. *Science* **241**, 170–176 (1988).

12. Letinic, K., Zoncu, R. & Rakic, P. Origin of GABAergic neurons in the human neocortex. *Nature* **417**, 645–649 (2002).

13. Kriegstein, A. R. & Noctor, S. C. Patterns of neuronal migration in the embryonic cortex. *Trends Neurosci.* **27**, 392–399 (2004).

14. O'Leary, D. D., Schlaggar, B. L. & Tuttle, R. Specification of neocortical areas and thalamocortical connections. *Annu. Rev. Neurosci.* **17**, 419–439 (1994).

15. Elsen, G. E. *et al.* The protomap is propagated to cortical plate neurons through an Eomes-dependent intermediate map. *Proc. Natl. Acad. Sci.* **110**, 4081–4086 (2013).

16. Ypsilanti, A. R. *et al.* Transcriptional network orchestrating regional patterning of cortical progenitors. *Proc. Natl.*

Acad. Sci. **118**, e2024795118 (2021).

17. Greig, L. C., Woodworth, M. B., Galazo, M. J., Padmanabhan, H. & Macklis, J. D. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* **14**, 755–769 (2013).

18. Cadwell, C. R., Bhaduri, A., Mostajo-Radji, M. A., Keefe, M. G. & Nowakowski, T. J. Development and Arealization of the Cerebral Cortex. *Neuron* **103**, 980–1004 (2019).

19. Sansom, S. N. & Livesey, F. J. Gradients in the Brain: The Control of the Development of Form and Function in the Cerebral Cortex. *Cold Spring Harb. Perspect. Biol.* **1**, a002519 (2009).

20. Rubenstein, J. L. *et al.* Genetic control of cortical regionalization and connectivity. *Cereb. Cortex N. Y. N 1991* **9**, 524–532 (1999).

21. Bishop, K. M., Goudreau, G. & O'Leary, D. D. M. Regulation of Area Identity in the Mammalian Neocortex by Emx2 and Pax6. *Science* **288**, 344–349 (2000).

22. O'Leary, D. D. M., Chou, S.-J. & Sahara, S. Area Patterning of the Mammalian Cortex. *Neuron* **56**, 252–269 (2007).

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

24. Goulas, A., Majka, P., Rosa, M. G. P. & Hilgetag, C. C. A blueprint of mammalian cortical connectomes. *PLOS Biol.* **17**, e2005346 (2019).

25. Charvet, C. J., Cahalane, D. J. & Finlay, B. L. Systematic, Cross-Cortex Variation in Neuron Numbers in Rodents and Primates. *Cereb. Cortex N. Y. NY* **25**, 147–160 (2015).

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

27. Amir, Y., Harel, M. & Malach, R. Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *J. Comp. Neurol.* **334**, 19–46 (1993).

28. Markov, N. T. *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex N. Y. N 1991* **24**, 17–36 (2014).

29. Froudist-Walsh, S. *et al.* Gradients of neurotransmitter receptor expression in the macaque cortex. *Nat. Neurosci.* **26**, 1281–1294 (2023).

30. Fan, X. *et al.* Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Res.* **28**, 730–745 (2018).

31. Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).

32. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated

signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).

33. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, (2018).

34. Kang, H. J. *et al.* Spatiotemporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).

35. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).

36. Richiardi, J. *et al.* Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–1244 (2015).

37. Sakai, T. *et al.* Fetal brain development in chimpanzees versus humans. *Curr. Biol.* **22**, R791–R792 (2012).

38. Bethlehem, R. a. I. *et al.* Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).

39. Molnár, Z. & Pollen, A. How unique is the human neocortex? *Dev. Camb. Engl.* **141**, 11–16 (2014).

40. Uylings, H. B. & van Eden, C. G. Qualitative and quantitative comparison of the prefrontal cortex in rat and in primates, including humans. *Prog. Brain Res.* **85**, 31–62 (1990).

41. 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).

42. Reardon, P. K. *et al.* Normative brain size variation and brain shape diversity in humans. *Science* **360**, 1222–1227 (2018).

43. Rilling, J. K. Comparative primate neuroimaging: insights into human brain evolution. *Trends Cogn. Sci.* **18**, 46–55 (2014).

44. Donahue, C. J., Glasser, M. F., Preuss, T. M., Rilling, J. K. & Van Essen, D. C. Quantitative assessment of prefrontal cortex in humans relative to nonhuman primates. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5183–E5192 (2018).

45. Hazlett, H. C. *et al.* Early brain development in infants at high risk for autism spectrum disorder. *Nature* **542**, 348–351 (2017).

46. Wagstyl, K. *et al.* Atlas of lesion locations and postsurgical seizure freedom in focal cortical dysplasia: A MELD study. *Epilepsia* **63**, 61–74 (2022).

47. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat. Commun.* **11**, 3358 (2020).

48. Hettwer, M. D. *et al.* Coordinated cortical thickness alterations across six neurodevelopmental and psychiatric disorders. *Nat. Commun.* **13**, 6851 (2022).

49. Rash, B. G. *et al.* Gliogenesis in the outer subventricular zone promotes enlargement and gyrification of the primate cerebrum. *Proc. Natl. Acad. Sci.* **116**, 7089–7094 (2019).

50. Rash, B. G., Arellano, J. I., Duque, A. & Rakic, P. Role of intracortical neuropil growth in the gyrification of the primate cerebral cortex. *Proc. Natl. Acad. Sci.* **120**, e2210967120 (2023).

51. Kriegstein, A., Noctor, S. & Martínez-Cerdeño, V. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat. Rev. Neurosci.* **7**, 883–890 (2006).

52. Hansen, D. V., Lui, J. H., Parker, P. R. L. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554–561 (2010).

53. Pollen, A. A. *et al.* Molecular Identity of Human Outer Radial Glia during Cortical Development. *Cell* **163**, 55–67 (2015).

54. Nowakowski, T. J., Pollen, A. A., Sandoval-Espinosa, C. & Kriegstein, A. R. Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral Cortex Development. *Neuron* **91**, 1219–1227 (2016).

55. Pichat, J., Iglesias, J. E., Yousry, T., Ourselin, S. & Modat, M. A Survey of Methods for 3D Histology Reconstruction. *Med. Image Anal.* **46**, 73–105 (2018).

56. Dauguet, J. *et al.* Three-dimensional reconstruction of stained histological slices and 3D non-linear registration with in-vivo MRI for whole baboon brain. *J. Neurosci. Methods* **164**, 191–204 (2007).

57. Choe, A. S. *et al.* Accuracy of image registration between MRI and light microscopy in the ex-vivo brain. *Magn. Reson. Imaging* **29**, 683–692 (2011).

58. Agarwal, N., Xu, X. & Gopi, M. Geometry Processing of Conventionally Produced Mouse Brain Slice Images. *J. Neurosci. Methods* **306**, 45–56 (2018).

59. Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. *ArXiv161107004 Cs* (2018).

60. Goodfellow, I. J. *et al.* Generative Adversarial Networks. *ArXiv14062661 Cs Stat* (2014).

61. Pérez, P., Gangnet, M. & Blake, A. Poisson image editing. *ACM Trans. Graph.* **22**, 313–318 (2003).

62. Amunts, K. *et al.* BigBrain: An Ultrahigh-Resolution 3D Human Brain Model. *Science* (2013) doi:10.1126/science.1235381.

63. Gholipour, A. *et al.* A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. *Sci. Rep.* **7**, 476 (2017).

64. Papageorghiou, A. T. *et al.* The INTERGROWTH-21st fetal growth standards: toward the global integration of pregnancy and pediatric care. *Am. J. Obstet. Gynecol.* **218**, S630–S640 (2018).

65. Scott, J. A. *et al.* Growth trajectories of the human fetal brain tissues estimated from 3D reconstructed in utero MRI. *Int. J. Dev. Neurosci.* **29**, 529–536 (2011).

66. Gleeson, J. G., Lin, P. T., Flanagan, L. A. & Walsh, C. A. Doublecortin Is a Microtubule-Associated Protein and Is

Expressed Widely by Migrating Neurons. *Neuron* **23**, 257–271 (1999).

67. Bahn, S., Volk, B. & Wisden, W. Kainate receptor gene expression in the developing rat brain. *J. Neurosci.* **14**, 5525–5547 (1994).

68. Ding, S.-L. *et al.* Comprehensive cellular-resolution atlas of the adult human brain. *J. Comp. Neurol.* **524**, 3127–3481 (2016).

69. Adria Casamitjana *et al.* A next-generation, histological atlas of the human brain and its application to automated brain MRI segmentation. *bioRxiv* 2024.02.05.579016 (2024) doi:10.1101/2024.02.05.579016.

70. Wang, Q. *et al.* The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936-953.e20 (2020).

71. Li, A. *et al.* Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain. *Science* **330**, 1404–1408 (2010).

72. Howard, A. F. D. *et al.* An open resource combining multi-contrast MRI and microscopy in the macaque brain. *Nat. Commun.* **14**, 4320 (2023).

73. Wagstyl, K. *et al.* BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices. *PLOS Biol.* **18**, e3000678 (2020).

74. Paquola, C. *et al.* The BigBrainWarp toolbox for integration of BigBrain 3D histology with multimodal neuroimaging. *eLife* **10**, e70119 (2021).

75. Lee, D. R. *et al.* Transcriptional heterogeneity of ventricular zone cells in the ganglionic eminences of the mouse forebrain. *eLife* **11**, e71864.

76. Menassa, D. A. *et al.* The spatiotemporal dynamics of microglia across the human lifespan. *Dev. Cell* **57**, 2127-2139.e6 (2022).

77. Werling, D. M. *et al.* Whole-Genome and RNA Sequencing Reveal Variation and Transcriptomic Coordination in the Developing Human Prefrontal Cortex. *Cell Rep.* **31**, 107489 (2020).

78. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172**, 650–665 (2018).

79. Abrahams, B. S. *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol. Autism* **4**, 36 (2013).

80. Thormann, A. *et al.* Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat. Commun.* **10**, 2373 (2019).

81. Fountain, M. D. *et al.* The phenotypic spectrum of Schaaf-Yang syndrome: 18 new affected individuals from 14 families. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **19**, 45–52 (2017).

82. Knight, S. J. L. *et al.* Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental

retardation. *Cell* **74**, 127–134 (1993).

83. Piao, X. *et al.* G protein-coupled receptor-dependent development of human frontal cortex. *Science* **303**, 2033–2036 (2004).

84. Edwards, A. D. *et al.* The Developing Human Connectome Project Neonatal Data Release. *Front. Neurosci.* **16**, 886772 (2022).

85. Price, A. N. *et al.* The developing Human Connectome Project (dHCP): fetal acquisition protocol. in *Proceedings of the 27th ISMRM Annual Meeting & Exhibition* vol. 27 (Abstract no. 244) (2019).

86. Cordero-Grande, L., Hughes, E. J., Hutter, J., Price, A. N. & Hajnal, J. V. Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to neonatal brain imaging. *Magn. Reson. Med.* **79**, 1365–1376 (2018).

87. Makropoulos, A. *et al.* The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction. *NeuroImage* **173**, 88–112 (2018).

88. Schuh, A. *et al.* A deformable model for the reconstruction of the neonatal cortex. in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* 800–803 (2017). doi:10.1109/ISBI.2017.7950639.

89. Karolis, V. *et al.* Developing Human Connectome Project spatio-temporal surface atlas of the fetal brain. 300 MiB (2023) doi:10.12751/G-NODE.QJ5HS7.

90. Robinson, E. C. *et al.* Multimodal surface matching with higher-order smoothness constraints. *NeuroImage* **167**, 453–465 (2018).

91. Bozek, J. *et al.* Construction of a neonatal cortical surface atlas using Multimodal Surface Matching in the Developing Human Connectome Project. *NeuroImage* **179**, 11–29 (2018).

92. Robinson, E. C. *et al.* MSM: a new flexible framework for Multimodal Surface Matching. *NeuroImage* **100**, 414–426 (2014).

93. Kyriakopoulou, V. *et al.* Normative biometry of the fetal brain using magnetic resonance imaging. *Brain Struct. Funct.* **222**, 2295–2307 (2017).

94. Kapellou, O. *et al.* Abnormal cortical development after premature birth shown by altered allometric scaling of brain growth. *PLoS Med.* **3**, e265 (2006).

95. Garcia, K. E. *et al.* Dynamic patterns of cortical expansion during folding of the preterm human brain. *Proc. Natl. Acad. Sci.* **115**, 3156–3161 (2018).

96. Vasung, L. *et al.* Spatiotemporal Differences in the Regional Cortical Plate and Subplate Volume Growth during Fetal Development. *Cereb. Cortex* **30**, 4438–4453 (2020).

97. Moeskops, P. *et al.* Development of Cortical Morphology Evaluated with Longitudinal MR Brain Images of Preterm

Infants. *PLOS ONE* **10**, e0131552 (2015).

98. Xia, J. *et al.* Fetal cortical surface atlas parcellation based on growth patterns. *Hum. Brain Mapp.* **40**, 3881–3899 (2019).

99. Mackarehtschian, K., Lau, C. K., Caras, I. & McConnell, S. K. Regional differences in the developing cerebral cortex revealed by ephrin-A5 expression. *Cereb. Cortex N. Y. N 1991* **9**, 601–610 (1999).

100. Theil, T., Alvarez-Bolado, G., Walter, A. & Rüther, U. Gli3 is required for Emx gene expression during dorsal telencephalon development. *Dev. Camb. Engl.* **126**, 3561–3571 (1999).

101. Stevens, H. E. *et al.* Fgfr2 Is Required for the Development of the Medial Prefrontal Cortex and Its Connections with Limbic Circuits. *J. Neurosci.* **30**, 5590–5602 (2010).

102. Simpson, J. H., Bland, K. S., Fetter, R. D. & Goodman, C. S. Short-range and long-range guidance by Slit and its Robo receptors: a combinatorial code of Robo receptors controls lateral position. *Cell* **103**, 1019–1032 (2000).

103. Bielle, F. *et al.* Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. *Nat. Neurosci.* **8**, 1002–1012 (2005).

104. Abraham, H., Pérez-García, C. G. & Meyer, G. p73 and Reelin in Cajal-Retzius cells of the developing human hippocampal formation. *Cereb. Cortex N. Y. N 1991* **14**, 484–495 (2004).

105. Irvin, D. K., Zurcher, S. D., Nguyen, T., Weinmaster, G. & Kornblum, H. I. Expression patterns of Notch1, Notch2, and Notch3 suggest multiple functional roles for the Notch-DSL signaling system during brain development. *J. Comp. Neurol.* **436**, 167–181 (2001).

106. Alsina, F. C. *et al.* Lrig1 is a cell-intrinsic modulator of hippocampal dendrite complexity and BDNF signaling. *EMBO Rep.* **17**, 601–616 (2016).

107. Jorstad, N. L. *et al.* Transcriptomic cytoarchitecture reveals principles of human neocortex organization. *Science* **382**, eadf6812 (2023).

108. Girskis, K. M. *et al.* Rewiring of human neurodevelopmental gene regulatory programs by human accelerated regions. *Neuron* **109**, 3239-3251.e7 (2021).

109. Kronenberg, Z. N. *et al.* High-resolution comparative analysis of great ape genomes. *Science* **360**, eaar6343 (2018).

110. Lanoue, V. *et al.* The adhesion-GPCR BAI3, a gene linked to psychiatric disorders, regulates dendrite morphogenesis in neurons. *Mol. Psychiatry* **18**, 943–950 (2013).

111. Lee, S.-J. *et al.* Presynaptic Neuronal Pentraxin Receptor Organizes Excitatory and Inhibitory Synapses. *J. Neurosci.* **37**, 1062–1080 (2017).

112. Stepien, B. K. *et al.* Lengthening Neurogenic Period during Neocortical Development Causes a Hallmark of

Neocortex Expansion. *Curr. Biol. CB* **30**, 4227-4237.e5 (2020).

113. Stepien, B. K., Vaid, S. & Huttner, W. B. Length of the Neurogenic Period—A Key Determinant for the Generation of Upper-Layer Neurons During Neocortex Development and Evolution. *Front. Cell Dev. Biol.* **9**, (2021).

114. Lewitus, E., Kelava, I., Kalinka, A. T., Tomancak, P. & Huttner, W. B. An Adaptive Threshold in Mammalian Neocortical Evolution. *PLOS Biol.* **12**, e1002000 (2014).

115. Rakic, P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci.* **18**, 383–388 (1995).

116. Fenlon, L. R. Timing as a Mechanism of Development and Evolution in the Cerebral Cortex. *Brain. Behav. Evol.* **97**, 8–32 (2021).

117. Cubelos, B., Briz, C. G., Esteban-Ortega, G. M. & Nieto, M. Cux1 and Cux2 selectively target basal and apical dendritic compartments of layer II-III cortical neurons. *Dev. Neurobiol.* **75**, 163–172 (2015).

118. Oppermann, H. *et al.* CUX1-related neurodevelopmental disorder: deep insights into phenotype-genotype spectrum and underlying pathology. *Eur. J. Hum. Genet.* 1–10 (2023) doi:10.1038/s41431-023-01445-2.

119. Doan, R. N. *et al.* Mutations in Human Accelerated Regions (HARs) Disrupt Cognition and Social Behavior. *Cell* **167**, 341 (2016).

120. Won, H. *et al.* Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016).

121. de la Torre Ubieta, L. *et al.* The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* **172**, 289-304.e18 (2018).

122. Markenscoff-Papadimitriou, E. *et al.* A Chromatin Accessibility Atlas of the Developing Human Telencephalon. *Cell* **182**, 754-769.e18 (2020).

123. Dominguez, M. H., Ayoub, A. E. & Rakic, P. POU-III Transcription Factors (Brn1, Brn2, and Oct6) Influence Neurogenesis, Molecular Identity, and Migratory Destination of Upper-Layer Cells of the Cerebral Cortex. *Cereb. Cortex N. Y. NY* **23**, 2632–2643 (2013).

124. Muralidharan, B. *et al.* LHX2 Interacts with the NuRD Complex and Regulates Cortical Neuron Subtype Determinants Fezf2 and Sox11. *J. Neurosci.* **37**, 194–203 (2017).

125. Pfurr, S. *et al.* The E2A splice variant E47 regulates the differentiation of projection neurons via p57(KIP2) during cortical development. *Dev. Camb. Engl.* **144**, 3917–3931 (2017).

126. Huang, H. *et al.* Interactive Repression of MYRF Self-Cleavage and Activity in Oligodendrocyte Differentiation by TMEM98 Protein. *J. Neurosci.* **38**, 9829–9839 (2018).

127. Fujikawa, A. & Noda, M. Role of pleiotrophin-protein tyrosine phosphatase receptor type Z signaling in

myelination. *Neural Regen. Res.* **11**, 549–551 (2016).

128. Ichihara-Tanaka, K., Oohira, A., Rumsby, M. & Muramatsu, T. Neuroglycan C is a novel midkine receptor involved in process elongation of oligodendroglial precursor-like cells. *J. Biol. Chem.* **281**, 30857–30864 (2006).

129. Pooyan, P. *et al.* The Dynamic Proteome of Oligodendrocyte Lineage Differentiation Features Planar Cell Polarity and Macroautophagy Pathways. *GigaScience* **9**, giaa116 (2020).

130. Shen, S., Li, J. & Casaccia-Bonnefil, P. Histone modifications affect timing of oligodendrocyte progenitor differentiation in the developing rat brain. *J. Cell Biol.* **169**, 577–589 (2005).

131. Marin-Husstege, M., Muggironi, M., Liu, A. & Casaccia-Bonnefil, P. Histone deacetylase activity is necessary for oligodendrocyte lineage progression. *J. Neurosci. Off. J. Soc. Neurosci.* **22**, 10333–10345 (2002).

132. Zhang, Y. *et al.* SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* **1**, 1021–1031 (1998).

133. Ye, F. *et al.* HDAC1 and HDAC2 Regulate Oligodendrocyte Differentiation By Disrupting β-Catenin-TCF Interaction. *Nat. Neurosci.* **12**, 829–838 (2009).

134. Seidlitz, J. *et al.* The molecular genetic landscape of human brain size variation. *Cell Rep.* **42**, 113439 (2023).

135. Sahu, S. K. *et al.* FBXO32 promotes microenvironment underlying epithelial-mesenchymal transition via CtBP1 during tumour metastasis and brain development. *Nat. Commun.* **8**, 1523 (2017).

136. Collins, R. L. *et al.* A cross-disorder dosage sensitivity map of the human genome. *Cell* **185**, 3041-3055.e25 (2022).

137. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).

138. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).

139. Finlay, B. L. & Darlington, R. B. Linked regularities in the development and evolution of mammalian brains. *Science* **268**, 1578–1584 (1995).

140. Rakic, P. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science* **183**, 425–427 (1974).

141. Gardette, R., Courtois, M. & Bisconte, J. C. Prenatal development of mouse central nervous structures: time of neuron origin and gradients of neuronal production. A radioautographic study. *J. Hirnforsch.* **23**, 415–431 (1982).

142. Sanderson, K. J. & Weller, W. L. Gradients of neurogenesis in possum neocortex. *Brain Res. Dev. Brain Res.* **55**, 269–274 (1990).

143. Elston, G. N., Oga, T. & Fujita, I. Spinogenesis and Pruning Scales across Functional Hierarchies. *J. Neurosci.* **29**, 3271–3275 (2009).

144. Rakic, P. Pre- and post-developmental neurogenesis in primates. *Clin. Neurosci. Res.* **2**, 29–39 (2002).

145. Granger, B., Tekaia, F., Le Sourd, A. M., Rakic, P. & Bourgeois, J.-P. Tempo of neurogenesis and synaptogenesis in the primate cingulate mesocortex: Comparison with the neocortex. *J. Comp. Neurol.* **360**, 363–376 (1995).

146. Chi, J. G., Dooling, E. C. & Gilles, F. H. Gyral development of the human brain. *Ann. Neurol.* **1**, 86–93 (1977).

147. Huang, W. *et al.* Origins and Proliferative States of Human Oligodendrocyte Precursor Cells. *Cell* **182**, 594-608.e11 (2020).

148. Reillo, I., de Juan Romero, C., García-Cabezas, M. Á. & Borrell, V. A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb. Cortex N. Y. N 1991* **21**, 1674–1694 (2011).

149. Richman, D. P., Stewart, R. M., Hutchinson, J. W. & Caviness, V. S. Mechanical model of brain convolutional development. *Science* **189**, 18–21 (1975).

150. Bayly, P. V., Okamoto, R. J., Xu, G., Shi, Y. & Taber, L. A. A cortical folding model incorporating stress-dependent growth explains gyral wavelengths and stress patterns in the developing brain. *Phys. Biol.* **10**, 016005 (2013).

151. Budday, S., Raybaud, C. & Kuhl, E. A mechanical model predicts morphological abnormalities in the developing human brain. *Sci. Rep.* **4**, 5644 (2014).

152. Tallinen, T., Chung, J. Y., Biggins, J. S. & Mahadevan, L. Gyrification from constrained cortical expansion. *Proc. Natl. Acad. Sci.* **111**, 12667–12672 (2014).

153. Lohmann, G., von Cramon, D. Y. & Steinmetz, H. Sulcal variability of twins. *Cereb. Cortex N. Y. N 1991* **9**, 754–763 (1999).

154. Lohmann, G., von Cramon, D. Y. & Colchester, A. C. F. Deep sulcal landmarks provide an organizing framework for human cortical folding. *Cereb. Cortex N. Y. N 1991* **18**, 1415–1420 (2008).

155. Ronan, L. *et al.* Differential Tangential Expansion as a Mechanism for Cortical Gyrification. *Cereb. Cortex N. Y. NY* **24**, 2219–2228 (2014).

156. Modenato, C. *et al.* Effects of eight neuropsychiatric copy number variants on human brain structure. *Transl. Psychiatry* **11**, 1–10 (2021).

157. Sønderby, I. E. *et al.* Effects of copy number variations on brain structure and risk for psychiatric illness: Large-scale studies from the ENIGMA working groups on CNVs. *Hum. Brain Mapp.* **43**, 300–328 (2022).

158. Opel, N. *et al.* Cross-Disorder Analysis of Brain Structural Abnormalities in Six Major Psychiatric Disorders: A Secondary Analysis of Mega- and Meta-analytical Findings From the ENIGMA Consortium. *Biol. Psychiatry* **88**, 678–686 (2020).

159. Writing Committee for the ENIGMA-CNV Working Group. Association of Copy Number Variation of the 15q11.2 BP1-BP2 Region With Cortical and Subcortical Morphology and Cognition. *JAMA Psychiatry* **77**, 420–430 (2020).

160. Boedhoe, P. S. W. *et al.* Subcortical Brain Volume, Regional Cortical Thickness, and Cortical Surface Area Across Disorders: Findings From the ENIGMA ADHD, ASD, and OCD Working Groups. *Am. J. Psychiatry* **177**, 834–843 (2020).

161. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).

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

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

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

165. Shafiei, G. *et al.* Topographic gradients of intrinsic dynamics across neocortex. *eLife* **9**, e62116 (2020).

166. Hansen, J. Y. *et al.* Mapping gene transcription and neurocognition across human neocortex. *Nat. Hum. Behav.* **5**, 1240–1250 (2021).

167. Ball, G., Seidlitz, J., Beare, R. & Seal, M. L. Cortical remodelling in childhood is associated with genes enriched for neurodevelopmental disorders. *NeuroImage* **215**, 116803 (2020).

168. Parker, N. *et al.* Assessment of Neurobiological Mechanisms of Cortical Thinning During Childhood and Adolescence and Their Implications for Psychiatric Disorders. *JAMA Psychiatry* **77**, 1127–1136 (2020).

169. Lau, H. Y. G., Fornito, A. & Fulcher, B. D. Scaling of gene transcriptional gradients with brain size across mouse development. *NeuroImage* **224**, 117395 (2021).

170. Romero-Garcia, R. *et al.* Transcriptomic and connectomic correlates of differential spatial patterning among gliomas. *Brain* **146**, 1200–1211 (2022).

171. Lee, H. M. *et al.* Multimodal mapping of regional brain vulnerability to focal cortical dysplasia. *Brain J. Neurol.* **146**, 3404–3415 (2023).

172. Dufford, A. J. *et al.* (Un)common space in infant neuroimaging studies: A systematic review of infant templates. *Hum. Brain Mapp.* **43**, 3007–3016 (2022).

173. Li, Y. *et al.* Spatiotemporal transcriptome atlas reveals the regional specification of the developing human brain. *Cell* **0**, (2023).

174. Howell, B. R. *et al.* The UNC/UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. *NeuroImage* **185**, 891–905 (2019).

175. Volkow, N. D., Gordon, J. A. & Freund, M. P. The Healthy Brain and Child Development Study—Shedding Light on Opioid Exposure, COVID-19, and Health Disparities. *JAMA Psychiatry* **78**, 471–472 (2021).

176. Majka, P. & Wójcik, D. K. Possum—A Framework for Three-Dimensional Reconstruction of Brain Images from Serial Sections. *Neuroinformatics* **14**, 265–278 (2016).

177. Yushkevich, P. A. *et al.* 3D mouse brain reconstruction from histology using a coarse-to-fine approach. in *Proceedings of the Third international conference on Biomedical Image Registration* 230–237 (Springer-Verlag, 2006). doi:10.1007/11784012_28.

178. Malandain, G., Bardinet, E., Nelissen, K. & Vanduffel, W. Fusion of autoradiographs with an MR volume using 2-D and 3-D linear transformations. *NeuroImage* **23**, 111–127 (2004).

179. Beare, R., Richards, K., Murphy, S., Petrou, S. & Reutens, D. An assessment of methods for aligning two-dimensional microscope sections to create image volumes. *J. Neurosci. Methods* **170**, 332–344 (2008).

180. Arloth, J., Bader, D. M., Röh, S. & Altmann, A. Re-Annotator: Annotation Pipeline for Microarray Probe Sequences. *PLOS ONE* **10**, e0139516 (2015).

181. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* **18**, 1832–1844 (2015).

182. Adamson, C. L. *et al.* Parcellation of the neonatal cortex using Surface-based Melbourne Children's Regional Infant Brain atlases (M-CRIB-S). *Sci. Rep.* **10**, 4359 (2020).

183. Uus, A. U. *et al.* BOUNTI: Brain vOlumetry and aUtomated parcellatioN for 3D feTal MRI. *eLife* **12**, (2023).

184. Eichert, N. *et al.* Cross-species cortical alignment identifies different types of anatomical reorganization in the primate temporal lobe. *eLife* **9**, e53232 (2020).

185. Glasser, M. F. & Essen, D. C. V. Mapping Human Cortical Areas In Vivo Based on Myelin Content as Revealed by T1- and T2-Weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011).

186. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).