



## King's Research Portal

DOI:

[10.1371/journal.pbio.3000976](https://doi.org/10.1371/journal.pbio.3000976)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Ball, G., Seidlitz, J., O'Muircheartaigh, J., Dimitrova, R., Fenchel, D., Makropoulos, A., Christiaens, D., Schuh, A., Passerat-Palmbach, J., Hutter, J., Cordero-Grande, L., Hughes, E., Price, A., Hajnal, J. V., Rueckert, D., Robinson, E. C., & Edwards, A. D. (2020). Cortical morphology at birth reflects spatiotemporal patterns of gene expression in the fetal human brain. *PLoS Biology*, *18*(11), [e3000976].  
<https://doi.org/10.1371/journal.pbio.3000976>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

1

2

3

4

5

6

7

8

9 Gareth Ball\*,<sup>1,2,3</sup> Jakob Seidlitz,<sup>4,5</sup> Jonathan O’Muircheartaigh,<sup>2</sup> Ralica Dimitrova,<sup>2</sup> Daphna Fenchel,<sup>2</sup>  
10 Antonios Makropoulos,<sup>2</sup> Daan Christiaens,<sup>2,6</sup> Andreas Schuh,<sup>7</sup> Jonathan Passerat-Palmbach,<sup>7</sup> Jana  
11 Hutter,<sup>2</sup> Lucilio Cordero-Grande,<sup>2</sup> Emer Hughes,<sup>2</sup> Anthony Price,<sup>2</sup> Jo V Hajnal,<sup>2</sup> Daniel Rueckert,<sup>7</sup>  
12 Emma C Robinson.<sup>2</sup> & A David Edwards.<sup>2</sup>

13

14

1. Developmental Imaging, Murdoch Children’s Research Institute, Melbourne, Australia

15

2. Centre for the Developing Brain, Department of Perinatal Imaging & Health, King’s College London, UK

16

3. Department of Paediatrics, University of Melbourne, Melbourne, Australia

17

4. Developmental Neurogenomics Unit, National Institute of Mental Health, Bethesda, USA

18

5. Department of Psychiatry, University of Cambridge, UK

19

6. Department of Electrical Engineering, ESAT/PSI, KU Leuven, Belgium

20

7. Biomedical Image Analysis Group, Department of Computing, Imperial College London, UK

21

22

\* gareth.ball@mcri.edu.au

23

24

25 **Abstract**

26 Interruption to gestation through preterm birth can significantly impact cortical development and have  
27 long-lasting adverse effects on neurodevelopmental outcome. We compared cortical morphology  
28 captured by high-resolution, multimodal MRI in n=292 healthy newborn infants (mean age at  
29 birth=39.9 weeks) with regional patterns of gene expression in the fetal cortex across gestation  
30 (n=156 samples from 16 brains, aged 12 to 37 post-conceptual weeks). We tested the hypothesis  
31 that noninvasive measures of cortical structure at birth mirror areal differences in cortical gene  
32 expression across gestation, and in a cohort of n=64 preterm infants (mean age at birth = 32.0  
33 weeks), we tested whether cortical alterations observed after preterm birth were associated with  
34 altered gene expression in specific developmental cell populations. Neonatal cortical structure was  
35 aligned to differential patterns of cell-specific gene expression in the fetal cortex. Principal component  
36 analysis of six measures of cortical morphology and microstructure showed that cortical regions were  
37 ordered along a principal axis, with primary cortex clearly separated from heteromodal cortex. This  
38 axis was correlated to estimated tissue maturity, indexed by differential expression of genes  
39 expressed by progenitor cells and neurons; and engaged in stem cell differentiation, neuron migration  
40 and forebrain development. Preterm birth was associated with altered regional MRI metrics and  
41 patterns of differential gene expression in glial cell populations. The spatial patterning of gene  
42 expression in the developing cortex was thus mirrored by regional variation in cortical morphology and  
43 microstructure at term, and this was disrupted by preterm birth. This work provides a framework to link  
44 molecular mechanisms to noninvasive measures of cortical development in early life and highlights  
45 novel pathways to injury in neonatal populations at increased risk of neurodevelopmental disorder.  
46

47 

## Introduction

48 The mammalian cortex is composed of functionally distinct regions organised along broad gradients  
49 that reflect spatially ordered and concerted variations of cortical structure and function.[1–6] While the  
50 mechanisms behind the emergence of this complex topography are not fully understood, cortical  
51 patterning is underwritten by dynamic regulation of gene transcription during gestation.[7,8] During  
52 embryonic development, early patterning of the neuroepithelium is established through intrinsic genetic  
53 mechanisms [9–12] that regulate early neurodevelopmental processes including neurogenesis and  
54 neuronal migrations from around 6 – 8 post-conceptual weeks (pcw) in humans.[11,13] During fetal  
55 development, this leads to the establishment and expansion of transient neural structures, including  
56 the subventricular zone, preplate, and subplate, and, eventually, formation of the cortex.[11,14,15]

57

58 The advent of modern transcriptomic technologies has allowed the precise mapping of cortical gene  
59 expression during the human fetal period.[16–18] Gene transcription is highly differentially expressed  
60 during prenatal development and varies significantly across cortical areas.[8,16,17,19] Interruption to  
61 the precisely timed dynamics of gene transcription during gestation is implicated in the onset of  
62 common developmental cognitive and neuropsychiatric disorders.[18,20,21]

63

64 Recently, the *postmortem* transcription of thousands of genes across the adult brain has been  
65 compiled to form brain-wide, gene expression atlases.[18,22,23] This allows precise comparison  
66 between spatial patterns of cortical gene expression and neuroanatomy quantified using Magnetic  
67 Resonance Imaging (MRI).[24] Neuroimaging studies have found patterns of gene expression in the  
68 adult cortex are mirrored by regional variation in cortical morphometry [25] and functional  
69 organisation,[26] and are associated with neuroimaging markers of developmental disorders.[27]  
70 Similar databases detailing cerebral gene transcription across the full human lifespan from early  
71 embryonic stages to adulthood are now available.[16,18] This has created an unprecedented  
72 opportunity to explore the molecular correlates of neuroimaging markers of early brain development.

73

74 Advances in neonatal neuroimaging now permit the quantification of developmental neuroanatomy *in*  
75 *vivo* at a higher-resolution than previously possible.[28,29] Imaging studies of the developing human  
76 brain shortly after birth have characterised a highly dynamic period of cerebral change defined by  
77 significant increases in brain volume,[30,31] cortical thickness and surface area,[31–33] progressive  
78 white matter myelination[34,35] and ongoing configuration and consolidation of functional brain  
79 networks.[36–41] Several studies have also used diffusion MRI models to study the microstructure of  
80 the cortex at around the time of birth, identifying areal patterns of development that may relate to  
81 ongoing cellular processes including dendritic arborisation and synaptic formation.[42–46] Further, the  
82 truncation of gestation due to preterm birth is associated with widespread alterations in cortical  
83 morphometry and microstructure indexed by MRI at the time of normal birth that highlight the  
84 sensitivity of noninvasive neuroimaging to detect disruptions in early developmental processes  
85 [32,42–49].

86

87 The combination of these technologies opens a new window to study early human brain development,  
88 facilitating a comparison between patterns of prenatal cortical gene expression and the development  
89 of the brain at around the time of birth, as well as providing a platform to test mechanistic hypotheses  
90 about the impact of early disruptions to brain development during gestation. The potential of this  
91 approach has been previously demonstrated using *post mortem* MRI to reveal a correspondence  
92 between genes linked to neural development and the microstructure of the fetal cortex. [50] In this  
93 study, we explore the association between *in vivo* measures of cortical morphometry at birth and  
94 regional patterns of fetal gene transcription in the human brain. We test the hypothesis that  
95 noninvasive markers of neonatal cortical structure mirror areal differences in the timing of cellular  
96 processes underlying cortical development, as indexed by differential spatiotemporal patterning of  
97 gene expression in the fetal cortex. Additionally, we test whether cortical alterations observed after  
98 preterm birth and quantified with MRI are linked with a selective vulnerability of developmental  
99 neuronal and glial cell populations in the developing cortex.

100

101 We define a principal mode of variation in neonatal cortical structure that is aligned to differential  
102 patterns of genes expression in the fetal cortex, enriched for foundational neurodevelopmental  
103 processes including neuronal differentiation and migration, and disrupted by preterm birth.

## 104 Results

### 105 A principal axis of the neonatal cortex

106 Using high-resolution structural and diffusion MRI data acquired from a large cohort of healthy  
107 neonates (n=292, 54% male, median [range] gestational age at scan = 40.86 [37.29-44.71]), we  
108 extracted six measures of cortical morphology (cortical thickness) and microstructure (T1w/T2w  
109 contrast; Fractional Anisotropy, FA; Mean Diffusivity, MD; Intracellular Volume Fraction, fICVF; and  
110 Orientation Dispersion Index, ODI) from eleven cortical regions-of-interest (ROI) with corresponding  
111 mRNA-sequencing in a prenatal transcriptomic dataset [18] (Fig 1A, S1 Fig).

112

113 Normalised regional metrics for all subjects are shown in Fig 1B. Similar regional profiles are evident  
114 across metrics with the pattern of inter-regional variation reflecting the full transcortical patterns  
115 shown in Fig 1A. Comparable regional patterns were observed in FA and cortical thickness, and  
116 between ODI, fICVF and the T1w/T2w contrast (Fig 1B), with higher FA, thicker cortex and a higher  
117 T1w/T2w contrast in primary somatomotor cortex (Fig 1A & 1B). MD displayed an opposing trend  
118 across regions, lowest in primary somatomotor regions and highest in fronto-parietal regions.

119

120 Based on the similarities in cortical patterning across metrics, we hypothesised that regional variation  
121 across metrics could be represented by a small number of latent factors. Using Principal Component  
122 Analysis (PCA), we projected the regional metrics onto a set of principal axes that maximally

123 explained variance in the full set of cortical measures (Fig 1C). Using the group-average region  $\times$   
124 metric matrix, we found that the first two components explained 91.6% of the total variance (PC1 =  
125 72.3% and PC2 = 19.3%, respectively).

126

127 The ordering of regions along the principal axis (PC1; Fig 1C) illustrates a clear separation between  
128 primary and higher order cortical regions based on neuroimaging metrics, with primary  
129 somatosensory and motor cortex (A1C, M1, S1) situated at opposite ends to prefrontal, inferior  
130 parietal and temporal cortex (DLPFC, IPC, ITC). This pattern is apparent in all cortical metrics, most  
131 strongly in T1/T2 contrast, fICVF and mean diffusivity (S2 Fig). The second principal axis (PC2)  
132 predominantly captured anatomical and microstructural differences in V1 compared to other primary  
133 cortex (Fig 1C & 1D).

### 134 PC1 is associated with regional patterns of gene expression in mid-gestation

135 Using a developmental transcriptomic dataset of bulk tissue mRNA data sampled from cortical tissue  
136 in 16 prenatal human specimens,[18] we compared regional variation in cortical MRI metrics,  
137 represented by PC1, with prenatal gene expression in anatomically-correspondent cortical regions.  
138 Through comparison to five independent single-cell RNA studies of the developing fetal cortex  
139 [18,51–54], we selected a set of 5287 marker genes shown to be differentially expressed in cortical  
140 cell populations during gestation. We used a nonlinear mixed-effects approach to model  
141 developmental changes in gene expression as a smooth function of age, accounting for inter-  
142 specimen variability. The nonlinear model provided a better fit of the expression data for all genes  
143 compared to a comparable linear model (range AIC difference: -16.7 to -87.7; range BIC difference: -  
144 2.1 to -58.9).

145

146 Using specimen- and age-corrected RPKM values provided by the residuals of the nonlinear mixed  
147 model for each gene (S3 Fig), we tested the association between spatial variation in gene expression  
148 during gestation and regional PC1 score using non-parametric correlation (Kendall's  $\tau$ ). Of 5287  
149 genes, 120 displayed a significant (positive or negative) correlation with PC1 after correction for  
150 multiple comparisons with False Discovery Rate ( $p < 0.05$ ). In total, 71 genes were positively correlated  
151 with PC1, with increasing gene expression in regions with a higher PC1 score (mean  $\pm$  S.D.  $\tau = 0.208$   
152  $\pm 0.023$ ) and 49 genes displayed the opposite relationship, with higher expression in regions with a  
153 negative PC1 score (mean  $\pm$  S.D.  $\tau = -0.208 \pm 0.022$ ).

154

155 We reasoned that genes associated with the patterning of cortical morphometry at birth may subserv  
156 important neurodevelopmental functions. To test this, we performed an over-representation analysis  
157 (ORA)[55] for ontological terms associated with specific biological processes in both gene lists. Of 71  
158 genes with spatial patterns of expression positively correlated with PC1 (denoted PC+), 61 (86%)  
159 were annotated to specific functional terms. Using all protein-coding genes transcribed in the bulk  
160 RNA dataset as the background reference set, we found significant enrichment of several  
161 neurodevelopmental terms including: stem cell differentiation (FDR=0.001, enrichment ratio=9.32),

162 neuron migration (FDR=0.03, enrichment=7.94) and forebrain development (FDR=0.004,  
 163 enrichment=5.65) (Fig 2A; S1 Table). Terms relating to stem cell and neuronal differentiation  
 164 remained significantly enriched when restricting the background reference set to only include genetic  
 165 markers of fetal cortical cells (n=5287; S1 Table). Performing weighted gene correlation network  
 166 analysis (WGCNA) on the PC+ gene set, we identified two co-expression modules (Fig 2B). The  
 167 largest, contained 53 genes including a tightly correlated set of developmental genes with roles in  
 168 regulating cell growth and differentiation including *EOMES*, *NEUROD4*, *SFRP1* and *TFAP2C*. The  
 169 smaller, second module (Module 2) contained 13 genes, with roles including neuronal signalling  
 170 (*ERBB4*, *CALB2*, *SCGN*) and neuronal differentiation (*ZNF536*, *DLX1*).

171

172 No biological terms were significantly enriched in genes with a spatial pattern of expression negatively  
 173 correlated to PC1 (denoted PC-). Using WGCNA, three small modules of 7 genes each were  
 174 identified (Modules 1N-3N; S4 Fig), including genes with high neuronal expression (Module 1N;  
 175 *CDKL5*, *ZBTB18*, *SORCS1*), and genes involved in cellular processes including adhesion and  
 176 signalling (Module 2N: *ACTN2*, *PTPN2*, *SSX2IP*) and metabolic activity (Module 3N: *DUSP7*,  
 177 *ST3GAL1*).

178

179 Using independent microarray data from laser microdissections (LMD) of the 21pcw fetal cortex, [16]  
 180 we verified that PC+ genes had higher expression in the cortical plate of higher order regions  
 181 (DLPFC, VLPFC, IPC) compared to primary cortex (M1, S1) (Fig 2C; mean fold change = 1.36,  
 182  $p < 0.001$ , 10,000 permutations). Using the top 100 differentially expressed genes identified in the LMD  
 183 dataset, we also confirmed that genes with higher expression in regions with a higher PC1 score  
 184 (DLPFC, VLPFC, IPC) in mid-gestation were enriched for important neurodevelopment functions  
 185 including neuron differentiation (GO:0021953, FDR=0.019, reference: all genes; S2 Table). Additional  
 186 validation experiments using independent single-cell RNA-seq data [51] confirmed an association  
 187 between regional PC1 score at birth and expression of PC1+ and PC- gene sets in mid-gestation (Fig  
 188 2D).

### 189 Imaging-gene associations are enriched for specific cell types in the fetal cortex

190 To explore these relationships further, we reconstructed cellular gene expression profiles by  
 191 stratifying the bulk tissue expression data using genetic markers of cell type derived from single-cell  
 192 RNA studies of the fetal cortex [18,51–54].

193

194 Sets of genetic markers for eleven cortical cell classes were initially compiled by combining lists of  
 195 genes that are differentially expressed in fetal cortical cell populations (S3 Table). To verify this  
 196 grouping, we calculated the average expression trajectories for all genetic markers within each cell  
 197 type across gestation and used them to calculate a 2D embedding using Uniform Manifold  
 198 Approximation and Projection (UMAP; Fig 3A). Proximity in the embedded space reflects similarity  
 199 between average trajectories of gene expression within cell type over time. In the embedded space,

200 cell types clustered by assigned class, and maturational timing (e.g.: precursor or mature), as well as  
201 within cellular subtype (eg: inhibitory and excitatory neurons; S5 Fig).

202

203 We tested the enrichment of genes expressed by each cell class within the PC+ and PC- gene sets.  
204 We found that PC+ genes were significantly enriched for genes expressed by precursor cells  
205 ( $p=0.0003$ , reference: fetal gene markers), specifically, for genes expressed by intermediate  
206 progenitor cells (enrichment ratio = 1.63,  $p=0.0002$ ; S4 Table), and inhibitory neurons (enrichment  
207 ratio = 3.2,  $p<0.0001$ ; Fig 3B; S4 Table). Posthoc analysis within cell class revealed specific inhibitory  
208 neuron subtypes present in the mid-fetal brain and enriched in the PC+ gene set included migrating  
209 cortical interneurons from the caudal ganglionic eminence (In\_5,[51] IN-CTX-CGE2,[52] both  
210  $p<0.0001$ ) and newborn interneurons originating in the medial ganglionic eminence (nIN1,[52]  
211  $p=0.0017$ ).

212

213 In contrast, PC- genes, with a spatial pattern of expression that was higher in primary somatomotor  
214 regions at mid-gestation, were enriched for genes expressed by mature cell types (enrichment=1.18,  
215  $p=0.002$ ). In terms of cell class, genes expressed by oligodendrocytes were enriched within PC-,  
216 though not significantly (enrichment=1.75,  $p=0.056$ ; Fig 3B; S4 Table). When considering only marker  
217 genes uniquely expressed by each cell class, PC- genes were enriched for excitatory neuronal genes  
218 (enrichment=2.14,  $p=0.008$ ; S5 Table). Posthoc analysis within this class revealed a single enriched  
219 early maturing excitatory neuronal subtype (Ex\_4,[51],  $p<0.0001$ ). Similar patterns of cell class-  
220 specific expression of PC+ and PC- genes were observed in the single-cell RNA dataset (S6 Fig).

## 221 Variation in tissue maturation during gestation predicts cortical development at birth

222 These data suggest that the spatial patterning of gene expression in the developing cortex is mirrored  
223 by regional variation in cortical morphology and microstructure measured using MRI at birth. To test  
224 this hypothesis, we created a model of cortical maturity to capture the relationship between the  
225 regional timing of gene expression and tissue maturation.

226

227 We modelled tissue sample age as a function of gene expression using support vector regression. To  
228 ensure full coverage across the prenatal period, and to maximise the number of samples contributing  
229 to the model, we included additional data from all tissue samples from brains aged 8pcw to 4 months  
230 postnatal age ( $n=21$  total). Using mean cortical gene expression of all 120 (PC+ and PC-) genes (Fig  
231 4A) our model accurately predicted sample age across the full prenatal window (Fig 4B), up to 4  
232 months of age. We validated our model in a separate dataset comprising microarray data from the  
233 prefrontal cortex in  $n=46$  brains aged 13pcw to 4 months [56] (BrainCloud; Fig 4B).

234

235 Using predictions from this model, we estimated the correlation between regional age predictions and  
236 PC1 in the prenatal sample (Fig 4C, left). We expected that for a given brain, regions with a more  
237 advanced gene expression profile (i.e. more similar to older tissue samples) would return an older age  
238 prediction. We observed a negative association develop over gestation between a cortical region's



239 predicted maturity and its position along the principal axis at birth (Fig 4C, right;  $R^2=0.36$ ,  $p<0.001$   
240 [5000 permutations]), such that in older samples, a lower PC1 score was associated with an older  
241 predicted age based on gene expression.

242

243 Using nonlinear models of gene expression over time, we estimated regional genetic maturity at  
244 several points across gestation (S7 Fig). We found that the relative maturity of regions compared to  
245 the rest of the cortex varied over time. Primary somatomotor regions remained relatively advanced  
246 throughout gestation compared to the rest of the cortex. In contrast, V1 remained relatively delayed  
247 across gestation. A divergence in maturity becomes apparent within higher-order regions by mid-  
248 gestation, with some cortical areas (IPC, ITC) falling behind other regions towards the time of birth.  
249 These patterns were largely repeated using the full fetal gene marker set ( $n=5287$  genes, S8 Fig).

## 250 Preterm birth leads to alterations along the principal imaging axis

251 Based on this evidence, we hypothesised that an interruption to the length of gestation would yield  
252 differences in cortical morphology indexed by variation along PC1. To test this, we compared cortical  
253 morphology in healthy neonates ( $n=292$ ) to a cohort of preterm-born infants scanned at term-  
254 equivalent age ( $n=64$ , 59% male; mean [S.D] gestational age at birth = 32.00 [3.88] weeks).

255

256 We extracted neuroimaging metrics from each cortical region and projected each individual's region  $\times$   
257 metric matrix onto the principal imaging axis (S9 Fig). After correcting for age at scan and sex,  
258 regional variation along PC1 explained significantly less variance in preterm individual's imaging data  
259 than those born at term (ANCOVA:  $F=7.9$ ,  $p=0.005$ ; Fig 5A). Across both groups, the mean variance  
260 explained by PC1 increased with age (Fig 5A;  $F=46.0$ ,  $p<0.001$ ), with a stronger association in the  
261 preterm cohort (interaction:  $F=6.63$ ,  $p=0.01$ ) suggesting that arrangement along the principal axis is  
262 ongoing around the time of birth and altered by events surrounding preterm birth. There was no  
263 significant difference between sexes ( $F=1.12$ ,  $p=0.28$ ).

264

265 As differences in the variance explained by PC1 are dictated by individual differences in cortical  
266 metrics, we sought to test the specific effects of preterm birth on all imaging measures. Using mixed  
267 effects linear models including effects of age, birth status and regional PC1 score, we confirmed a  
268 significant main effect of birth status on all cortical metrics except for ODI (S6-S8 Table). The largest  
269 effect was evident in cortical T1w/T2w contrast ( $F_{1,354}=135.53$ ,  $p<0.0001$ , Cohen's  $d=1.62$ ; S8 Table).  
270 On average, cortical T1w/T2w was significantly lower in preterm infants (marginal means [95% C.I.] =  
271 1.32 [1.31,1.33], 1.20 [1.18,1.22] for term and preterm infants, respectively). To a lesser extent, both  
272 intracellular volume fraction and FA were, on average, higher in term infants ( $d=0.32$ , 0.56  
273 respectively), although the direction of this effect was not consistent across cortical regions (S10 Fig).  
274 In contrast, average cortical mean diffusivity ( $d=-1.17$ ) and, to a lesser extent, cortical thickness ( $d=-$   
275 0.65) were higher in preterm infants across all regions.

276

277 The magnitude of regional group differences across all cortical metrics varied as a function of PC1  
278 (Fig 5B; S7 Table; S10 Fig). This effect was most apparent in T1w/T2w contrast where the differences  
279 between term and preterm groups formed a strong negative association with PC1 ( $r = -0.78$ ,  $p=0.023$   
280 after FDR correction). Similar trends were seen in the other metrics, although none reached  
281 significance ( $|r| = 0.32$  to  $0.68$ , all  $p>0.05$ ).

## 282 Vulnerability of specific cell populations to the timing of preterm birth

283 These data show that cortical differences in preterm infants occur along the principal imaging axis and  
284 are most apparent in T1/T2w contrast. We investigated the potential that the differences observed in  
285 preterm cortex may reflect a selective vulnerability in specific cell populations due to coincidental  
286 timing of extrauterine exposure following preterm birth and temporal variations in gene expression.  
287 Focusing on the cortical differences observed in T1/T2w contrast, we first estimated gene expression  
288 trajectories over the latter stages of gestation (160 to 260 post-conceptual days, approximately 25 to  
289 39 weeks gestational weeks). We then split this period into 10 age windows and within each, we  
290 identified genes with expression significantly correlated to the magnitude of group differences in  
291 T1/T2w contrast at term-equivalent age (FDR-corrected  $p<0.05$ , Fig 5C). Within each window, we  
292 tested for enrichment of gene expression by each of 10 fetal cell types. In the early preterm period,  
293 we found that mean regional differences in T1w/T2w contrast at term-equivalent age were  
294 significantly associated with genes expressed by both inhibitory and excitatory neurons (windows 1, 2,  
295 3 and 5, hypergeometric statistic:  $p<0.05$ ; Fig 5C, top; reference: fetal cell markers). However, later in  
296 gestation, T1w/T2w differences were correlated with the expression of genes enriched for glial cell  
297 populations, including microglia, endothelial cells (windows 8,9 and 10; all  $p<0.05$ ) and  
298 oligodendrocytes (windows 1,2 and 8,9 and 10, all  $p<0.05$ ; Fig 5C, middle).

299  
300 Using gene expression data from a mouse model of preterm brain injury [57], we confirmed the  
301 relationship between preterm brain injury and altered gene expression in glial cell populations at birth  
302 (Fig 5D). Gene expression in the murine cortex was measured at P1.5, after hypoxic-ischemic insult  
303 at E16.5. Differentially expressed genes (DEG;  $p<0.05$ ) were mapped to human homologs with 217  
304 DEGs matched to human genes included in the current study. We found that the set of DEGs was  
305 enriched for genes both expressed by glial populations in the human fetal cortex and associated with  
306 T1/T2 differences in the neonatal cortex (Fig 5D; S9 Table). Relative to a background set of mapped  
307 genes ( $n=15,052$ ) we observed a significant enrichment of genes associated with T1/T2 in at least  
308 one age window and expressed by glial populations including microglia (enrichment ratio=6.9,  
309  $p<0.001$ ), endothelial cells (5.4,  $p<0.001$ ), oligodendrocytes (4.9,  $p<0.001$ ) and radial glia (2.2,  
310  $p=0.02$ ). These relationships remained significant in microglia and endothelial cells when restricting  
311 the background set to only include matched fetal gene markers ( $n=4733$ ).

312

## 313 Potential cellular processes disrupted in the preterm brain

314 To identify potential molecular pathways associated with the neuroimaging differences we observed in  
315 the preterm cortex, we identified genes expressed by glial cell types associated with both  
316 neuroimaging differences in the human preterm brain and experimental models of preterm brain  
317 injury. Using genes expressed by oligodendrocytes (Fig 6A), microglia (Fig 7) and endothelial cells  
318 (S11 Fig) and associated with T1/T2 differences across multiple prenatal age windows, we identified  
319 protein-protein interaction networks using the STRING database [58]. Networks for oligodendrocytes  
320 and microglia are shown in Fig 6 and Fig 7. We performed a functional enrichment analysis of  
321 Reactome pathways [59] using the whole genome as a reference to identify specific molecular  
322 processes involving genes in each PPI network and identified significantly enriched pathways in each  
323 cell population (S10-S12 Table).

324

325 In oligodendrocytes, pathway enrichment analysis revealed significant gene associations across  
326 multiple time windows. Genes involved in NMDA signalling in the MAPK/ERK pathway (HSA-438066,  
327 HSA-442729, HSA-442982; *DLG1*, *GRIN2A*) were significantly correlated to T1w/T2w differences  
328 across the majority of the preterm period. In contrast, regional expression of genes associated with  
329 the MyD88 and TLR signalling cascades (HSA-975871; *S100B*, *RPS6KA2*) were most closely  
330 correlated to T1w/T2w differences in the latter stages of gestation (windows 5 to 9 and 7 to 10,  
331 respectively). Other pathways linked genes expression over multiple time periods. Neurotrophin  
332 signalling pathways included genes *OMG* (correlated between windows 1-8) and *ARHGEF10*  
333 (windows 2-5), and the Rho-GTPase signalling pathway (HSA-194840) included both *ARHGEF10* and  
334 *RHOB* (windows 5-10). Finally, sphingolipid metabolism pathways included genes expressed across  
335 both the full prenatal in humans and differentially-expressed after fetal ischaemic insult in mice  
336 (*ACER3*, windows 2-9).

337

338 In microglia, enriched pathways among genes associated with T1/T2 differences in the preterm cortex  
339 included signal transduction (HSA-162582), cytokine signalling (HSA-1280215) and stress response  
340 (Homology directed repair; HSA-5693538). The regional expression of genes associated with Rho  
341 GTPase signalling (HSA-194315; *CIT*, *RHOB*, *ARHGAP25*) spanned the prenatal period and were  
342 linked to brain injury in the mouse model (*CIT*). Similarly, *ITGAX*, associated with T1/T2 differences  
343 late in gestation and linked to cellular inflammatory responses, was differentially expressed in the  
344 mouse brain after fetal hypoxic-ischaemia.

345

346 In endothelial cells, a large PPI network enriched for genes associated with apoptotic mechanisms,  
347 including *CASP3*, *CDKN1A*, and *GADD45B*, each with patterns of expression correlated with T1/T2  
348 differences in the preterm cortex (S11 Fig).

349

350 This highlights metabolic signalling pathways associated with genes expressed in developmental glial  
351 populations during the period most at risk of interruption by preterm birth with a regional specificity

352 correlated to neuroimaging markers of preterm brain injury at birth and a functional role in  
353 experimental models of preterm brain injury.  
354

## 355 Discussion

356 In this study, we aimed to test the hypothesis that noninvasive neuroimaging measures of cortical  
357 structure at birth encode differential spatiotemporal patterning of genes underlying corticogenesis. We  
358 found that gene expression in the fetal cortex is mirrored by a principal mode of variation across  
359 multiple MRI metrics in the neonatal cortex. Specifically, regional variation in cortical morphometry  
360 and microstructure reflects differences in developmental maturity and tissue composition across  
361 cortical areas, indexed by the differential timing of gene expression across multiple cell types in the  
362 fetal cortex. Having established this relationship, we found that interruption to gestation through  
363 preterm birth resulted in significant disruptions to MRI-based measures of cortical development by the  
364 time of full-term birth. Further, the effects of preterm birth are temporally and spatially coincident to  
365 developmental processes involving cortical glial cell populations. This work provides an experimental  
366 framework to link molecular developmental mechanisms to macroscopic measures of cortical  
367 anatomy in early life, demonstrating the relationship between fetal gene expression and neonatal  
368 brain development and highlighting the specific impact of early exposure to the extrauterine  
369 environment due to preterm birth.

370  
371 Using advanced MRI acquired close to the time of birth in a large, healthy neonatal population, we  
372 mapped multiple measures of regional cortical morphometry onto a single mode of variation, defining  
373 a principal axis of the neonatal cortex. Ordering of cortical regions along this axis separated lower  
374 order sensory and motor regions from higher-order regions including parietal, frontal and superior  
375 temporal cortex situated at opposite ends. The shared spatial ordering of cortical properties is a  
376 common organisational feature of the mammalian brain,[1,2,60,61] reflected in regional variations in  
377 cell populations,[61] gene expression,[22,62] and connectivity[60] as well as MRI-based measures of  
378 functional topography[5] and cortical morphometry in both adults[63] and infants.[64] The optimal  
379 mapping of cortical properties onto one or two lower dimensions remains an area of active  
380 research,[1] however, studies have demonstrated that variation along one axis, or gradient, is largely  
381 reflected by concerted changes in others,[2,6,65] suggesting that lower-order representations of  
382 cortical organisation largely capture shared views of latent neurobiological variation. An important  
383 benefit of this framework is the reduction of multiple metrics into a single measure per subject. In our  
384 case, this takes advantage of the inherent redundancy across multiple structural and diffusion MRI  
385 measures of the same cortical regions, producing a latent representation of cortical structure across  
386 scales. Here, we applied a simple linear mapping, arranging cortical regions along a single axis using  
387 PCA. This was sufficient to explain a significant proportion of variation in regional MRI-based metrics,  
388 with areas with similar cortical profiles clustering together along the principal axis. Based on an  
389 observed differential expression of genes associated with specific progenitor cell populations, order

390 along this axis was correlated with spatial gene expression patterns that reflected a differential timing  
391 of cortical development across regions. Comparison to cell-specific gene expression profiles in late  
392 gestation suggested that MRI-based measures of cortical structure at birth correlated with gene  
393 expression by specific glial populations involving oligodendrocytes, microglia and endothelial cells.  
394 This correlation potentially reflects a spatial variation in the developmental timing of processes  
395 associated with myelination, neuronal guidance and the continued maturation of the brain's vascular  
396 networks at around the time of birth.[66–70]

397

398 The advent of modern transcriptomic technologies has enabled detailed analyses of the foundational  
399 molecular mechanisms underpinning corticogenesis in the human fetal brain.[16,52,53] Resolved to  
400 the level of individual cells, recent studies have performed systematic explorations of gene expression  
401 dynamics across cell-cycle progression, migration and differentiation of several major cell types in the  
402 fetal brain.[51] Combined with regional expression levels of bulk tissue mRNA measured across  
403 multiple cortical areas, this allows the spatio-temporal mapping of cell-specific gene expression  
404 profiles in the developing brain.[52] Here, we used a development atlas of gene expression,  
405 measured across 11 cortical regions from 12 to 37 post-conceptual weeks in 16 separate brain  
406 specimens.[18] This data resource provides unparalleled access to both the spatial and temporal  
407 dynamics of developmental mechanisms ongoing in the cortex during gestation. We found that a  
408 number of genes vary across cortical areas in line with a principal imaging axis. In particular, we  
409 found that genes with relatively higher expression in higher order regions during gestation were  
410 associated with developmentally earlier processes including neuronal differentiation and migration  
411 and were predominantly expressed by intermediate precursor cells and early-maturing inhibitory  
412 neurons. Using an alternative approach in four mid-gestation brain samples (aged 16-21 pcw), Miller  
413 et al. identified a generally rostro-caudal gradient of gene expression progressing along the contours  
414 of the developing brain and anchored in frontal and temporal cortex.[16] While some overlap was  
415 evident: 72/85 (85%) of frontally-enriched genes that were included in both studies were also  
416 positively correlated with the imaging axis, this indicates that variation along PC1 may reflect a  
417 combination of multiple overlapping intrinsic hierarchies or cellular gradients underlying cortical  
418 development.[18,61,62] Using a machine-learning approach designed to accommodate the large  
419 number of genes assayed, and validated in an independent sample, we established that the  
420 maturation of a given tissue sample could be accurately determined based on temporally-evolving  
421 profiles of gene expression. This approach takes advantage of the degree of variation in gene  
422 expression over development. Temporal variability in expression is present across most protein-  
423 coding genes and over 95% of genes that are differentially-expressed across cortical regions, are also  
424 differentially-expressed across gestation, with age explaining a large proportion of variance in gene  
425 expression.[18,19] Using the relative advancement or delay in predicted age across regional tissue  
426 samples, we observed a correlation between emerging differences in areal gene expression and  
427 cortical structure at birth, suggesting an interaction between the relative rate of development across  
428 regions and length of gestation. This was most notable in the protracted developmental trajectory of

429 the visual cortex in midgestation, as noted elsewhere.[16] Overall, our results lend evidential support  
430 to the presence of heterogeneous corticogenic timing over gestation.[71,72]

431

432 Based on these observations, we hypothesised that interruption to gestation would lead to cortical  
433 disruptions along the principal cortical axis, reflecting a deleterious interaction with genetically-  
434 determined developmental programs ongoing in the cortex in the latter stages of gestation. To test  
435 this, we compared cortical development in healthy newborns to a cohort of preterm-born infants  
436 scanned at the time of normal birth. In line with previous observations,[28,43,73] we found significant  
437 differences across most cortical metrics of macro- and microstructure in the preterm brain. The  
438 magnitude of differences between cohorts aligned with the principal imaging axis suggesting a  
439 differential impact of perinatal adversity on cortical development that is potentially encoded by a  
440 selective vulnerability across regions due to differential maturation rates. Adverse intrauterine  
441 environments can result in altered patterns of fetal gene expression and brain development [74–79]  
442 and we demonstrate overlapping genetic associations between alterations in preterm cortical  
443 structure and differential expression in an experimental model of fetal hypoxic-ischaemic brain  
444 injury.[57] However, the antecedents and impacts of preterm birth on brain development are  
445 multifactorial and we remain cautious on speculating about the causal mechanisms that may underlie  
446 the relationships observed in this study without further empirical evidence.

447

448 The largest effect was observed in the myelin-sensitive T1w/T2w contrast. In adults, regional variation  
449 in cortical T1w/T2w contrast is high correlated with quantitative MRI-measures of intracortical myelin  
450 and histological maps of cytoarchitecture.[80] Myelination in the neonatal cortex is minimal, however  
451 T1w and T2w signal vary as a function of position in the neonatal cortex and the transcortical pattern  
452 of T1w/T2w ratio observed in this study mirrors closely that reported in older cohorts, with high values  
453 predominant in primary sensory regions.[80] In addition, we find that genes with expression correlated  
454 to T1w/T2w contrast are enriched for genes expressed by glial cells, including microglia and  
455 oligodendrocytes, across the second half of gestation. This mirrors earlier reports, based on  
456 microarray data, of correlations between neonatal imaging phenotypes and glial gene expression  
457 during gestation.[81] Using a time-resolved analysis, we found several molecular pathways involving  
458 genes with spatial and temporal correlation to the potential timing of preterm birth. This method  
459 leveraged nonlinear models fit using the full prenatal sample, allowing the discrete mapping of varying  
460 gene expression associations across the mid- to late-fetal period. We identify co-expressed networks  
461 of genes expressed by microglia and oligodendrocytes in late gestation and associated with Rho-  
462 GTPase signalling pathways, critical for neuronal migration [82] and involved in oligodendrocyte  
463 maturation and myelination [83,84]; cytokine signalling and inflammatory response pathways involving  
464 NF- $\kappa$ B1 and associated with microglial activation after hypoxic insult;[85] the MAPK/ERK signaling  
465 pathway, associated with neuronal and oligodendrocyte proliferation, [86,87] as well as sphingolipid  
466 metabolic pathways and apoptotic pathways expressed in endothelial cells. These data provide  
467 supporting evidence to the important role of developmental glial populations in preterm brain  
468 injury.[88–90] We have previously identified risk alleles in preterm born infants in genes involved in

469 lipid metabolism and microglial activation in the developing brain and associated with altered patterns  
470 of brain development by term-equivalent age [91,92]. In this study, we validate our observations in an  
471 experimental model of fetal hypoxic-ischaemic injury highlighting a differential expression of glial-  
472 expressed genes after early brain injury.[57] We found several genes identified in both human and  
473 mouse studies and associated with T1/T2 differences in the preterm brain that were differentially-  
474 expressed after early brain injury suggested potential deleterious effects on glial cell populations that  
475 could lead to disrupted neuronal migration and formation of neural circuitry in the preterm  
476 brain.[57,57,93–96] While we recognise the need for further experimental research to elucidate the  
477 link to alterations in cortical structure, our findings highlight potential pathways by which preterm birth  
478 can result in altered cortical development due to coincidental timing with corticogenic processes in the  
479 fetal cortex.

480

481 We note several limitations to our study. While we have taken care to validate our observations  
482 across several datasets, we are careful to avoid causal language to describe the associations  
483 presented. Further experimental evidence is required to fully understand how spatial gene expression  
484 gradients lead to alterations in MRI-based measures of cortical structure in healthy and preterm  
485 brains. Due to the nature of the data, we compare *postmortem* gene expression data from the fetal  
486 cortex with *in vivo* measurements of cortical development in healthy and preterm neonates at the end  
487 of gestation. We recognise this approach depends upon a number of assumptions including that gene  
488 expression patterns can be generalised across preterm and healthy cohorts and that MRI-based  
489 metrics acquired at a single timepoint act as a surrogate measure for ongoing associations between  
490 cortical structure and gene expression during gestation. We mitigate some of these risks by  
491 performing validation experiments in independent datasets and comparing our findings to  
492 experimental models of preterm brain injury. To measure contemporaneous associations in imaging  
493 and gene expression, other studies [50] have employed *postmortem* MRI of fetal brains to acquire  
494 data in age-matched samples but this approach comes with the additional challenges of imaging *post*  
495 *mortem* tissue. With advancements in fetal MRI, we anticipate that future research will focus on  
496 examining imaging-transcriptomic associations across corresponding timepoints through mid- to late-  
497 gestation using healthy fetal MRI and preterm infants scanned shortly after birth to better capture the  
498 temporal evolution of the reported associations.

499

500 While the fetal dataset made available through PsychENCODE represents an unprecedented window  
501 into spatiotemporal gene expression in the developing brain, the relatively coarse spatial sampling  
502 limits our ability to map fine-grained spatial variation, or boundaries between primary and secondary  
503 areas. Our analyses only begins at 12pcw, after early gene expression gradients have begun to  
504 impose areal differentiation on the developing brain.[9] Similarly, the bulk tissue samples analysed  
505 contained fetal transient structures including the marginal zone and subplate, that differ from the  
506 cortex in terms of spatial and temporal development.[97] Although we verified our findings in the  
507 cortical plate using microarrays from layer-specific dissection in two 21pcw donor brains, this analysis  
508 was limited to a single timepoint. We anticipate, that with the increasing availability of spatially

509 resolved gene expression datasets, advances in fetal and neonatal imaging, as well as layer-specific  
510 imaging analysis [98], further exploration of this area will yield interesting insight into early cortical  
511 development. In addition, sexual dimorphism in the transient fetal structures of the brain have been  
512 reported.[99] While we felt the relatively small sample size precluded a direct assessment of sex  
513 differences in gene expression, we included sex as a factor in all models of gene expression over  
514 time, and in our analyses of cortical MRI measures.

515

516 To examine cell-specific gene expression, we performed a stratification of bulk tissue RNA using gene  
517 lists collated from several independent scRNA studies. This method assumes that areal differences in  
518 gene expression are due to differences in developmental timing represented by cellular differentiation  
519 as well as changes in the proportion of cell types in composite tissue. While we did not test directly  
520 whether the same cell types also show areal differences in gene expression, this was previously  
521 explored by Fan et al. [51] In 22-23pcw samples, they found that the dominant mode of variation  
522 across cells was cell type rather than regional location and that most areal differences were driven by  
523 significant variations in tissue composition, as well as differing patterns of maturation. The advent of  
524 high-resolution single-cell RNA maps [53] will hopefully lead to future studies where we can more  
525 directly test regional development of specific cell types, rather than broader cell classes.

526

527 Finally, in this study, we focused on cortical structure rather than function. Future research may  
528 explore spatial associations between fetal gene expression and brain function as measured by MRI.  
529 In adults, a number of recent analyses have compared MRI metrics to patterns of gene expression  
530 reporting significant associations between correlated gene expression and both structural and  
531 functional measures [25,100–102] As discussed above, concerted variation in cortical properties is a  
532 common feature across MRI modalities [103,104] and spatial gene associations may be difficult to  
533 disentangle across correlated metrics. However, cortical measures that correlate at a single time point  
534 may not develop in tandem and future exploration of temporal development of cortical structure and  
535 function in relation to gene expression will yield interesting future research directions in this area.

536

537 In conclusion, we show that noninvasive imaging of the cortical structure in the neonatal brain is  
538 sensitive to differential spatiotemporal patterns of gene expression during gestation. In addition, we  
539 find that disruption to this developmental programming by preterm birth is associated with significant  
540 cortical alterations that appear to reflect the selective vulnerability of developing glial populations in  
541 the developing cortex.

542

543



544 **Figure Legends**

545 **Figure 1: A principal axis of the neonatal cortex indexed by multi-modal MRI.** A. Average cortical neuroimaging metrics in  
 546 a cohort of healthy, term-born neonates (n=292). Metrics derived from structural MRI (T1/T2 contrast, cortical thickness) and  
 547 diffusion MRI model parameters using DTI (FA and MD) and NODDI (fICVF and ODI). Right. Cortical regions-of-interest based  
 548 on anatomical references with corresponding developmental transcriptomic data (S1 Fig). B. Z-scored cortical metrics are  
 549 shown for each subject grouped within each cortical region-of-interest. C. Top. Cortical representations of the first two principal  
 550 components (PC1, PC2) derived from PCA of the regional MRI metric data in B. Bottom. Position of each cortical region-of-  
 551 interest in PCA state space, the position of each region is dictated by its component score (D) for the first two principal  
 552 components. Regions are labelled and coloured by PC1 score. D. PCA scores of each metric for the first two principal  
 553 components, coloured by PC1 score. See: <https://github.com/garedaba/baby-brains/tree/master/figures>.

554

555 **Figure 2: Genes associated with neuronal differentiation are differentially expressed along the principal imaging axis.**  
 556 A. Volcano plot showing enrichment of GO terms (Biological Processes) in genes with age-corrected expression levels  
 557 positively correlated with PC1. Significantly enriched terms (FDR < 0.05, reference: protein-coding genes) are labelled. B. Gene  
 558 co-expression analysis of all PC+ genes revealed two modules. Intra-modular connections are shown with node size and  
 559 colour indicating strength, and edge thickness and colour indicating weight. C. Differential expression of PC+ genes across  
 560 cortical regions (top) measured using laser microdissection (LMD) microarrays of the cortical plate in two 21pcw fetal samples  
 561 (<https://www.brainspan.org/lcm/>). Heatmap shows relative expression of all 71 PC+ genes in the inner and outer cortical plate  
 562 of each labelled region. Inset: mean fold change of PC+ genes compared to 10,000 random gene sets of the same size. D.  
 563 Total expression (in transcripts per million, TPM) of PC+ (top) and PC- (bottom) genes in single cells (n=572) extracted from  
 564 cortical regions in an independent single-cell RNA-seq survey of the mid-gestational fetal cortex [51]. Scatterplots show mean  
 565 TPM averaged over cells in each region, correlated with each region's PC1 score. See: [https://github.com/garedaba/baby-](https://github.com/garedaba/baby-brains/tree/master/figures)  
 566 [brains/tree/master/figures](https://github.com/garedaba/baby-brains/tree/master/figures) and <https://github.com/garedaba/baby-brains/tree/master/results/wgcna>.

567

568 **Figure 3: Cell-specific gene expression is associated with cortical morphology at birth.** A. UMAP embedding of 86 cell  
 569 types based on trajectories of relative gene expression over time recovers annotated cell classes. Subplots reflect enrichment  
 570 ratios of cell classes in PC+ and PC- gene sets (darker colour represent higher enrichment ratio). B. Enrichment ratio for fetal  
 571 marker genes expressed by each cell class is shown for PC+ (left) and PC- (right) gene sets. See:  
 572 <https://github.com/garedaba/baby-brains/tree/master/figures>.

573

574 **Figure 4: Tissue maturity correlates with regional variation in cortical morphometry at birth.** A. Developmental patterns  
 575 of mean cortical gene expression illustrated in each specimen for all 120 regionally variant genes (PC+ and PC-), ordered by  
 576 age. B. The relationship between predicted and true sample age for all regional samples (n=198 samples from n=21 brains) in  
 577 the PsychENCODE dataset aged between 50 and 400 postconceptional days (8pcw to 4 postnatal months), estimated using  
 578 support vector regression (SVR) and leave-one-out cross-validation. The SVR model was validated using additional samples  
 579 from the BrainCloud dataset (n=46 samples). Shaded area indicates 95% C.I. C. Left. The correlation between regional PC1  
 580 score and predicted tissue maturity is shown for each sample during gestation. Error bars show 95% C.I. for regional age  
 581 predictions over 1000 bootstrapped gene samples. Right. PC1 correlation is plotted against specimen age for each brain.  
 582 Shaded area indicates 95% C.I. for linear model fit over bootstrap samples. See: [https://github.com/garedaba/baby-](https://github.com/garedaba/baby-brains/tree/master/figures)  
 583 [brains/tree/master/figures](https://github.com/garedaba/baby-brains/tree/master/figures).

584

585 **Figure 5: Disruption of cortical development in preterm-born infants.** A. Left. Group difference in individual variance  
 586 across multiple neuroimaging metrics explained by the principal imaging axis in term (blue) and preterm (green) infants (left).  
 587 Right. The relationship between age at scan and variance explained by PC1 across all cortical metrics (right). Regression lines  
 588 are shown for term (blue) and preterm (green) infants with 95% C.I. B. Group differences in regional T1w/T2w contrast ordered  
 589 by position along PC1. (linear regression shown with 95% C.I.). C. Enrichment of gene sets from 10 fetal cortical cell classes  
 590 (top: neuronal; middle: non-neuronal; bottom: precursor) based on genes significantly associated (FDR p<0.05) with group

591 differences in T1w/T2w contrast at 10 timepoints in the preterm period. D. Enrichment of genes both expressed by each cell  
592 type and significantly correlated with T1/T2 in at least one age window in differentially-expressed genes measured in an  
593 experimental mouse model of preterm brain injury. [57] See: <https://github.com/garedaba/baby-brains/tree/master/figures>.

594

595 **Figure 6: Cellular pathways associated with genes expressed by oligodendrocytes and developmental alterations in**  
596 **the preterm cortex.** Left. Genes expressed by oligodendrocytes in the fetal cortex and significantly associated with group  
597 differences in T1w/T2w contrasts across at least 3 age windows are shown. Dark green indicates periods where gene  
598 expression and T1w/T2w contrast were significantly correlated for each gene (FDR  $p < 0.05$ ) across the preterm period. Right.  
599 Protein-protein interaction networks derived using STRING. Top functional enrichments of molecular pathways are shown  
600 where applicable. Genes associated with listed enriched pathway and genes differentially-expressed in an animal model of  
601 preterm brain injury are highlighted. See: [https://github.com/garedaba/baby-brains/tree/master/data/gene\\_lists](https://github.com/garedaba/baby-brains/tree/master/data/gene_lists).

602

603 **Figure 7: Cellular pathways associated with genes expressed by microglia and developmental alterations in the**  
604 **preterm cortex.** Left. Genes expressed by microglia in the fetal cortex and significantly associated with group differences in  
605 T1w/T2w contrasts across at least 3 age windows are shown. Dark blue indicates periods where gene expression and  
606 T1w/T2w contrast were significantly correlated for each gene (FDR  $p < 0.05$ ) across the preterm period. Right. Protein-protein  
607 interaction networks derived using STRING. Top functional enrichments of molecular pathways are shown where applicable .  
608 Genes associated with listed enriched pathway and genes differentially-expressed in an animal model of preterm brain injury  
609 are highlighted. See: [https://github.com/garedaba/baby-brains/tree/master/data/gene\\_lists](https://github.com/garedaba/baby-brains/tree/master/data/gene_lists).

610

611

## 612 Materials and Methods

### 613 Ethics statement

614 The study was approved by the UK Health Research Authority (Research Ethics Committee reference  
615 number: 14/LO/1169) and performed in accordance with the Declaration of Helsinki. Written informed  
616 parental consent was obtained for all participants.

### 617 Subjects

618 Infants were recruited and imaged at the Evelina Newborn Imaging Centre, St Thomas' Hospital,  
619 London, UK for the Developing Human Connectome Project (dHCP). Neuroimaging and basic  
620 demographic data from the dHCP are available to download from:  
621 <http://www.developingconnectome.org/second-data-release/>.

622

623 In total, 442 healthy, term-born infants (gestational age at birth  $> 37$  weeks) scanned between  
624 February 2015 and November 2018 as part of the dHCP were included in this study. From this cohort,  
625  $n=362$  were successfully processed via the dHCP structural processing pipeline (see *Image*  
626 *processing* below) and included after quality control. Of these, diffusion data from  $n=296$  was  
627 successfully processed using both DTI and NODDI pipelines (see *Image processing* below) A further  
628 four subjects were excluded following a final visual inspection due to cropped anatomical images. Of  
629 107 preterm infants (gestational age at birth  $< 37$  weeks) scanned at term-equivalent age during the  
630 same period, one was excluded due to incomplete demographic data,  $n=84$  completed structural MRI

631 processing and n=67 passed diffusion processing after quality control. A further n=3 were removed  
632 after final visual inspection.

633

634 After quality control and image processing, the final cohort comprised n=292 healthy term-born infants  
635 (54% male, mean [S.D] postmenstrual age at birth=39.96 [1.10] weeks, mean [S.D.] age at  
636 imaging=40.94 [1.56] weeks) and n=64 preterm infants scanned at term-equivalent age (59% male;  
637 born 32.00 [3.88] weeks and imaged at 40.57 [2.25] weeks).

## 638 Magnetic Resonance Imaging

639 MRI was performed on a 3T Philips Achieva (Philips, Netherlands) using a dedicated neonatal  
640 imaging system including a neonatal 32 channel phased array head coil.[29] Infants were imaged  
641 without sedation. T1- and T2-weighted anatomical images were acquired alongside diffusion and  
642 resting state functional MRI (total acquisition time: 63 minutes).

643

644 Inversion-recovery T1-weighted and T2-weighted images were acquired in sagittal and axial  
645 orientations (in-plane resolution:  $0.8 \times 0.8\text{mm}^2$ , slice thickness: 1.6mm with 0.8mm overlap) with  
646 TR=4795ms; TI=1740ms; TE=8.7ms; SENSE: 2.27 (axial) and 2.66 (sagittal) for T1-weighted images  
647 and TR=12000ms, TE=156ms; SENSE: 2.11 (axial), 2.60 (sagittal) for T2-weighted. T1- and T2-  
648 weighted image stacks were motion corrected and reconstructed using the multi-slice aligned  
649 sensitivity encoding method with integration into a 3D volume using a super-resolution scheme into  
650  $0.8 \times 0.8 \times 0.8\text{mm}$  resolution volumes.[105,106]

651

652 Diffusion MRI was acquired with a spherically-optimised set of directions over 4 b-shells ( $b=0\text{s/mm}^2$ :  
653 20 directions;  $b=400$ : 64 directions;  $b=1000$ : 88 directions;  $b=2600$ : 128 directions) with a multiband  
654 factor acceleration of 4, TR=3800ms; TE=90ms; SENSE: 1.2 and acquired resolution of  $1.5\text{mm} \times$   
655  $1.5\text{mm}$  with 3mm slices (1.5mm overlap) reconstructed using an extended SENSE technique into  $1.5$   
656  $\times 1.5 \times 1.5\text{mm}$  volumes.[107,108]

## 657 Image processing

658 T1- and T2-weighted images were processed using the dHCP structural pipeline  
659 (<https://github.com/BioMedIA/dhcp-structural-pipeline>).[28] Briefly, T2-weighted images were bias  
660 corrected (N4),[109], brain-extracted (BET)[110] and segmented into grey matter, white matter and  
661 cerebrospinal fluid using DRAW-EM.[111] Cortical surfaces of the right and left hemisphere were then  
662 extracted[112] and aligned to a population-specific cortical template[113] using spherical inflation and  
663 multimodal surface matching (MSM) with higher order constraints

664 ([https://github.com/ecr05/MSM\\_HOCR](https://github.com/ecr05/MSM_HOCR)).[114,115] This method ensures that all surfaces across  
665 participants have one-to-one vertex correspondence with the dHCP neonatal template. For each  
666 subject, we extracted the following metrics: cortical thickness (corrected for cortical curvature) and  
667 T1w/T2w contrast (calculated using rigidly aligned T1-weighted images).

668

669 Diffusion-weighted images were preprocessed by first denoising [116] and removing Gibbs ringing  
670 artefacts,[117] followed by a slice-to-volume motion and distortion correction with a slice-level outlier  
671 rejection using a multi-shell spherical harmonic signal representation (SHARD).[118] Visual inspection  
672 of the 4D images ensured motion correction and outlier rejection was successful and that images of  
673 poor quality were excluded from further analysis.

674

675 We fit each subject's diffusion data with both a diffusion tensor model, fitted to the  $b=1000s^2/mm$  shell  
676 and implemented in MRtrix,[119] and the NODDI (Neurite Orientation Dispersion and Density  
677 Imaging) model [120]. For the diffusion data, NODDI was implemented with the NODDI MATLAB  
678 toolbox using the *invivopreterm* tissue type options with the default parameters of intrinsic diffusivity  
679 fixed to  $1.7 \times 10^{-3} mm^2/s$  and the starting point for values considered as the fraction of intra-neurite  
680 space lowered to 0-0.3 (instead of 0-1 in the adult brain) to better fit higher water content in the  
681 newborn compared to the mature adult brain.[49,121]

682

683 From these models, we derived parametric maps of fractional anisotropy (FA) and mean diffusivity  
684 (MD) from DTI, as well as maps of orientation dispersion index (ODI), quantifying the angular variation  
685 of neurite orientation within a voxel and intra-cellular volume fraction (fICVF), indexing the tissue  
686 volume fraction restricted within neurites. Cortical diffusion maps were projected to the cortical  
687 surface after co-registration with the corresponding anatomical data.

688

689 Images were visually inspected after acquisition and after reconstruction, and following each  
690 processing pipeline. Any images that failed to successfully complete the processing pipelines or failed  
691 visual inspection at any stage were removed from further analysis. As an additional step, we  
692 quantified in-scanner movement and image quality using a summary metric of the total head  
693 translation, rotation and the ratio of detected outlier slices. These three metrics were combined into  
694 one aggregated quality assurance measure [118,122] This measure did not significantly differ  
695 between groups (term [mean +/- S.D] = 1.61 +/- 1.79, preterm = 1.55 +/- 1.03;  $t=0.24$ ,  $p=0.81$ ).  
696 Including QA as a covariate in our analyses of group differences didn't impact our reported  
697 observations.

## 698 Bulk tissue gene expression data

699 Preprocessed, bulk tissue cortical gene expression data were made available as part of the  
700 PsychENCODE project (available to download at: <http://development.psychencode.org/>).[18] Tissue  
701 was collected after obtaining parental or next of kin consent and with approval by the institutional

702 review boards at the Yale University School of Medicine, the National Institutes of Health, and at each  
703 institution from which tissue specimens were obtained.

704

705 Tissue processing is detailed elsewhere. [18] In brief, mRNA data were available for post-mortem  
706 human brain tissue collected from n=41 specimens aged between 8 post-conceptual weeks (pcw)  
707 and 40 postnatal years. For each brain, regional dissection of up to 16 cerebral regions was  
708 performed, including 11 neocortical regions (dorsolateral frontal cortex, DLPFC; ventrolateral frontal  
709 cortex, VLPFC; orbitofrontal cortex, OFC; medial frontal cortex, MFC; primary motor cortex, M1;  
710 primary sensory cortex, S1; inferior parietal cortex, IPC; primary auditory cortex, A1C; superior  
711 temporal cortex, STC; inferior temporal cortex, ITC; primary visual cortex, V1), and five sub-cortical  
712 regions (hippocampus, amygdala, striatum, thalamus and cerebellar cortex). Detailed anatomical  
713 boundaries for each cortical region at each stage of development are provided elsewhere.[17,18]  
714 Regional tissue samples were subject to mRNA-sequencing using an Illumina Genome Analyzer IIx  
715 (Illumina, San Diego, CA) and mRNA-seq data processed using RSEQtools (v0.5).[123] Gene  
716 expression was measured as reads per kilobase of transcript per million mapped reads (RPKM).  
717 Conditional quantile normalisation was performed to remove GC-content bias and ComBat used to  
718 remove technical variance due to processing site (Yale or USC).[18,124,125]

719

720 In this study, we included RPKM data from neocortical samples of prenatal specimens aged 12 post-  
721 conceptual weeks onwards (n=16, age range = 12-37 pcw, mean [S.D.] age = 18.4 [7.7] pcw, 50%  
722 male, mean [S.D.] number of cortical regions sampled = 9.75 [1.6], mean [S.D.] post-mortem interval  
723 = 7.1 [12.6] hours, mean [S.D.] RNA integrity number[RIN][126] = 9.26 [0.73]). Prenatal specimens  
724 from the earliest developmental window (8-9 postconceptional weeks) were excluded as some cortical  
725 regions (e.g.: M1 and S1) were combined together to account for immature cortical anatomy.[17,18]

726

727 The prenatal gene expression data was initially filtered to only include protein-coding genes (NCBI  
728 GRCh38.p12, n=18,524 out of a possible 20,720). In order to restrict our analysis to focus on genes  
729 expressed in the developing cortex, we further filtered this list to only contain genes expressed by  
730 cells in the fetal cortex based on the composite list of prenatal cell markers from five independent  
731 single-cell RNA studies of the developing fetal cortex (see '*Genetic markers of cell type*' below). This  
732 resulted in expression data from a final set of 5287 genes.

### 733 Additional gene expression datasets

#### 734 *BrainCloud*

735 Preprocessed microarray data from n=46 human prefrontal cortex tissue samples aged approximately  
736 95 to 390 postconceptional-days (14 pcw to 4 months postnatal age) were downloaded from GEO  
737 (accession: GSE30272). Prior to analysis, individual gene expression was modelled using nonlinear  
738 splines. Surrogate Variable Analysis was performed to remove technical variation and batch effects  
739 (31 surrogate variables) while retaining variation due to age. For further details please see Colantuoni

740 et al. [56] For each gene, expression was Z-transformed prior to modelling. In total, gene expression  
741 for 4986/5287 fetal gene markers was available.

742

#### 743 *Single cell RNA*

744 Regional single cell RNA gene expression data was made available via GEO (accession:  
745 GSE103723).[51] Briefly, 4213 single cells were isolated from 20 anatomical regions of the 22 and  
746 23pcw fetal cortex and subject to single cell RNA-sequencing. Normalised expression data in  
747 transcripts per million (TPM) were available for 96 cells per tissue sample. We selected data from  
748 single cells extracted from matching cortical regions to those described above and classified to one of  
749 10 classes based on clusters identified in [51], as detailed '*Genetic markers of cell type*' below (n=572  
750 cells).

751

#### 752 *Laser microdissection (LMD) microarray data*

753 LMD microarray data was accessed via the BrainSpan data portal (brainspan.org/lcm). This provides  
754 access to DNA microarray data from four midgestational brains, dissected into around 300 anatomical  
755 samples. Detailed information is provided elsewhere.[16] We performed a differential search to  
756 identify microarray probes with differential expression in the cortical plate of the DLPFC (regional ID:  
757 fCPdl), VLPFC (fCPvl) and IPC (pCPpv) compared to M1 (fCPm1) and S1 (fCPs1) measured in two  
758 21pcw donor brains. Differential expression data for 23,000 probes was downloaded and  
759 corresponding data mapped to the set of preselected fetal marker genes for comparison.

760

#### 761 *Experimental mouse model of preterm brain injury*

762 Gene expression levels in P1.5 mouse cortex was measured for control or ischaemic pups, where  
763 ischaemia was induced by maternal uterine artery occlusion at E16.5. Four mice in each group were  
764 included. [57] Data were made available via GEO (accession: GSE89998). Analysis was performed  
765 using *GEO2R* and the *limma* package. Expression data were first log<sub>2</sub>-transformed before fold-  
766 change was estimated across groups. Mouse genes were mapped to human homologs using  
767 Ensembl and matched to the list of human fetal gene markers.

#### 768 **Cortical regions-of-interest**

769 To facilitate comparison between developmental RNA and MRI data, we created a set of cortical  
770 regions-of-interest (ROI) labels corresponding to the anatomical dissections used for mRNA analysis  
771 and aligned to the dHCP imaging data.

772

773 Reference post-mortem MRI data were acquired as part of the Allen Institute BrainSpan Atlas of the  
774 Developing Human Brain. Details of tissue processing and MRI acquisition are available at:  
775 [https://help.brain-map.org/download/attachments/3506181/BrainSpan\\_MR\\_DW\\_DT\\_Imaging.pdf](https://help.brain-map.org/download/attachments/3506181/BrainSpan_MR_DW_DT_Imaging.pdf). In  
776 brief, MRI was acquired at 3T (Siemens, Germany) in a post-mortem, whole-brain specimens aged 22  
777 pcw. In addition, anatomical annotations corresponding to the regional dissections in Miller et al.,[16]  
778 Kang et al.[17] and Li et al.[18] were provided on a reconstructed cortical surface from a 19pcw

779 prenatal specimen.[50] Cortical ROI data were available to download in VTK file format, separately for  
780 left and right cortical hemispheres (S1 Fig).

781

782 To generate a set of dHCP-compatible cortical labels, we reconstructed the cortical surface of a 3T  
783 post-mortem MRI image from a 22pcw brain. First, manually creating a brain mask to remove non-  
784 brain tissue, then smoothing using a mean filter of 3mm width. We performed automated tissue  
785 segmentation on the smoothed image using the dHCP structural pipeline, manually correcting tissue  
786 segmentations on a slice-by-slice basis for accuracy prior to cortical surface reconstruction. Using  
787 dHCP tools, the fetal cortical surface was extracted and cortical labels manually transferred onto it  
788 based on the reference labels provided by Huang et al.[50] and anatomical descriptions in Li et al.[18]  
789 Finally, the fetal surface was inflated to a sphere and co-registered to the earliest timepoint (36 weeks  
790 gestational age) of the dHCP cortical surface atlas using multimodal surface matching  
791 (MSM).[113,115]

792

793 This resulted in a set of 11 cortical ROI, each associated with regional bulk tissue mRNA data  
794 sampled across gestation and co-registered with dHCP neuroimaging data to allow correspondent  
795 sampling of cortical imaging metrics in the neonatal brain (S1 Fig).

## 796 Cortical imaging metric analysis

797 For every subject, mean values of each imaging metric (thickness, T1w/T2w contrast, FA, MD, fICVF,  
798 ODI) were calculated within each cortical ROI. Metric values were averaged across hemispheres and  
799 outlier values identified and removed using a median absolute deviation (MAD) of  $> 3.5$ .

800

801 For all healthy term-born infants, regional metrics were Z-transformed and averaged across subjects  
802 to produce a group average region  $\times$  metric matrix representing the relative variation of each imaging  
803 metric across cortical regions.

804

805 We projected the group average data onto two axes using Principal Component Analysis (PCA) via  
806 eigendecomposition of the data covariance matrix. This results in a set of  $L$  eigenvectors,  $W_L$ , that  
807 map the original  $n \times p$  data matrix,  $X$  onto a set of orthogonal axes as:  $T_L = XW_L$ . As generally  $L < p$ ,  
808 the truncated  $n \times L$  matrix,  $T_L$ , forms a low-dimensional representation of the original data. We can  
809 then project each subject's region  $\times$  metric matrix,  $X_s$ , onto a common set of axes as  $T_L^s = X_s W_L$ ,  
810 where  $T_L^s$  represents the  $L$  component scores for each subject,  $s$ .

811

812 All analysis was performed in Python (3.7.3) using Scipy (1.3.0)[127] and Scikit-Learn (0.21.2).[128]

813 **Modelling gene expression trajectories**

814 For each gene, we modelled the relationship between gene expression and specimen age using  
 815 mixed-effects models. Using bulk tissue RPKM data described above, each gene's expression data  
 816 were first Winsorised to set very small or large outlying values to the 5th and 95th centile values,  
 817 respectively, to stabilise against extreme values before log2-transformation.

818

819 We compared two models, modelling regional gene expression as either a linear or nonlinear function  
 820 of age with fixed effects of sex and RNA integrity number. We accounted for sample-specific variation  
 821 by including in the model a random intercept for each specimen, such that:

822

$$823 \quad y \sim f(v) + X\beta + Zb$$

824

825 Where  $f(\cdot)$  is a nonlinear function of predictor  $v$ ,  $X$  is an  $m$ -observation  $\times$   $p$  design matrix modelling  $p$   
 826 linear, fixed effects and  $Z$  is an  $m \times (n \cdot r)$  design matrix modelling  $r$  random effects across  $n$   
 827 specimens. In this case, age was included as either a nonlinear predictor,  $f(v)$ , or as a fixed linear  
 828 effect alongside sex and RIN. We specified a relatively smooth nonlinear function of age using a  
 829 natural cubic spline with four knots evenly spaced across the age span. To estimate region-specific  
 830 trajectories, we calculated a second nonlinear model, additionally including separate smooth functions  
 831 for each cortical region. Models were compared using AIC and BIC.

832

833 We calculated age-corrected RPKM values for each gene in all cortical samples using the residuals of  
 834 the best-fit nonlinear mixed model (S3 Fig) to test the spatial association between gene expression  
 835 and the principal imaging gradient using non-parametric correlation (Kendall's  $\tau$ ).

836

837 Modelling was performed in R (3.6.1) using *nlme*[129] and *mgcv*[130] packages.

838 **Genetic markers of cell type**

839 Genetic markers of cortical cell types were collated from five independent single-cell RNA studies of  
 840 the fetal cortex.[18,51–54] Using single-cell RNA-seq, each study identified sets of genes differentially  
 841 expressed across cell clusters or types. Cell types were independently defined in each study and a list  
 842 of all cell types included in this study ( $n=87$ ) are shown in S3 Table. Where applicable, for a given cell  
 843 type, differentially-expressed genes were included as cell type markers if they were found to be  
 844 expressed in at least 50% all cells surveyed.[18,51,52] Across all five studies, each cell type was  
 845 manually assigned to one of 11 cell classes based on text descriptions from each study (astrocyte,  
 846 endothelial cell, microglia, neuron:excitatory, neuron:inhibitory, neuron:unclassified, oligodendrocyte,  
 847 oligodendrocyte precursor cell [OPC], pericyte, intermediate progenitor cell, radial glia) and classified  
 848 as either a precursor or mature cell type (S3 Table). For each cell class, omnibus gene lists were  
 849 created by collating identified gene markers for all cell types within a class. Unique gene lists were  
 850 created by excluding any genes identified as a marker of more than one cell class.



## 851 Cell type embedding

852 Using the region-specific, nonlinear model specified above, expression trajectories for every gene  
 853 were estimated for each region at 50 evenly spaced points across the full observation window (12pcw  
 854 - 37pcw). For each cell type identified in the fetal cortex (see above), expression trajectories for all  
 855 cell-type gene markers were normalised to unit length, concatenated over regions and averaged to  
 856 capture both temporal and spatial variation in average gene expression across cell types. Similarity  
 857 between cell-type gene expression trajectories were then visualised by embedding into a two-  
 858 dimensional space using Uniform Manifold and Approximation Projection (UMAP) based on Euclidean  
 859 distance.[131]

## 860 Enrichment analyses

861 We performed over-representation analysis (ORA) of each list of gene markers for each of 10 cell  
 862 classes (excluding neuron:unclassified), calculating the hypergeometric statistic:

863

$$864 \quad p = 1 - \sum_{i=0}^x \frac{\binom{K}{i} \binom{M-K}{N-i}}{\binom{M}{N}}$$

865

866 Where  $p$  is the probability of finding  $x$  or more genes from a cell-class-specific gene list  $K$  in a set of  
 867 randomly selected genes,  $N$  drawn from a background set,  $M$ . We calculated enrichment ratios as the  
 868 proportion of cell-class-specific genes in the gene list of interest, compared to the proportion in the full  
 869 background set. The background gene set was defined as the full list of protein-coding genes  
 870 included in the analysis (n=5287) unless otherwise specified. We corrected for multiple comparisons  
 871 across cell classes using False Discovery Rate (FDR).

872

873 We additionally performed ORA for Gene Ontology terms using WebGestalt.[55]

## 874 Weighted Gene Correlation Network Analysis

875 We used WGCNA [132] to identify co-expression modules within PC+ and PC- gene sets. We  
 876 performed topology analysis using a gene  $\times$  gene adjacency matrix constructed from the residualised  
 877 log2-transformed RPKM data, after accounting for variance due to age, sex and sample effects (see  
 878 *Modelling gene expression trajectories*, above). A soft threshold was chosen to approximate scale  
 879 free topology in the adjacency matrix (PC+: power=5,  $r^2=0.77$ ; PC-: power=10,  $r^2=0.78$ ), [133] before  
 880 transformation into a topological overlap matrix. Hierarchical clustering was used to assign genes to  
 881 modules based on the dynamic tree-cutting method.[134] Analysis was performed in R (3.6.1) with the  
 882 *WGCNA* package.[132]

### 883 Predicting tissue maturity

884 We used gene expression over time to construct a predictive model of genetic maturity using support  
 885 vector regression. To maximise coverage across the prenatal period, we included additional samples  
 886 aged 8pcw up to 4 months of postnatal age (n=21 total). Using the n=120 regionally-varying genes  
 887 (PC+ and PC-), we first calculated regional gene expression profiles, corrected for variance due to  
 888 sex, RIN and specimen ID while retaining variance due to age, using nonlinear mixed-effects models.  
 889 We then averaged gene expression across cortical regions in each specimen to create a specimen  $\times$   
 890 gene (21  $\times$  120) mean gene expression matrix, where each row represents the normalised  
 891 log<sub>2</sub>(RPKM) of each gene for a given specimen, averaged across cortical regions.

892

893 In machine learning, kernels can be applied high-dimensional data sets to improve model fitting where  
 894  $n \ll p$ . To calculate regional variation in genetic maturity, we implemented a leave-one-out (LOO)  
 895 model using Support Vector Regression with a linear kernel (Scikit-Learn; regularisation parameter  
 896 set to  $C=10.0$ ) and modelling the association between specimen age (in post-conceptual days) and  
 897 mean cortical gene expression data in 20 out of 21 specimens. We then used this model to predict  
 898 age using the *regional* gene expression profiles of the remaining, left-out specimen, resulting in  
 899 eleven age predictions, one per cortical region. We repeated this process, leaving out a different  
 900 specimen each time.

901

902 In order to estimate a stable prediction of tissue maturity, we repeated the modelling using a  
 903 bootstrapped selection of genes, repeating gene sampling with replacement 1000 times. We also  
 904 repeated the modelling using all 5287 genes. We calculated the correlation between regional genetic  
 905 maturity (averaged over 1000 bootstraps) and PC1 score for each specimen and tested the  
 906 significance of this relationship by permuting mean gene expression profiles with respect to specimen  
 907 age 5000 times during model training.

### 908 Group comparison of cortical morphology

909 We compared regional cortical metrics in term and preterm cohorts using a linear mixed effects  
 910 modelling approach. For each of six metrics, we modelled metric value as a combination of age, sex,  
 911 regional PC1 score and birth group status (term or preterm). We included an interaction term for PC1  
 912 and birth status to test the hypothesis that preterm birth incurs differential effects across cortical  
 913 regions in line with PC1. We also included subject ID as a random effect to account for correlated  
 914 within-subject observations across regions. We fit nested models by Maximum Likelihood, comparing  
 915 model fits with and without the inclusion of birth status using AIC and BIC (S6 Table).

### 916 Developmental gene enrichment

917 In order to test cell class enrichment over time, we split the preterm period (approximately 160 to 260  
 918 post-conceptual days) into 10 age windows. Using nonlinear gene expression trajectories,  
 919 calculated across cortical regions (see *Modelling gene expression trajectories* above), we averaged

920 modelled gene expression within each window for every cortical region. Then, in each window, we  
921 calculated the non-parametric association (Kendall's  $\tau$ ) between gene expression and the mean  
922 difference between term and preterm groups in T1w/T2w contrast in each cortical region and  
923 recorded significantly associated genes (FDR-corrected at  $p < 0.05$ ). Finally, we performed cell-class  
924 enrichment (see 'Enrichment Analyses' above), in each of the 10, time-resolved gene sets.

#### 925 PPI networks

926 Protein-protein interactions networks were visualised in Cytoscape (3.7.2) using the StringDB protein  
927 query. Pathway enrichment of Reactome pathways was performed for subnetworks.  
928

929 **References**

- 930 1. Huntenburg JM, Bazin P-L, Margulies DS. Large-Scale Gradients in Human Cortical  
931 Organization. *Trends in Cognitive Sciences*. 2018;22: 21–31. doi:10.1016/j.tics.2017.11.002
- 932 2. Hilgetag CC, Goulas A. ‘Hierarchy’ in the organization of brain networks. *Philosophical*  
933 *Transactions of the Royal Society B: Biological Sciences*. 2020;375: 20190319.  
934 doi:10.1098/rstb.2019.0319
- 935 3. Raut RV, Snyder AZ, Raichle ME. Hierarchical dynamics as a macroscopic organizing principle of  
936 the human brain. *PNAS*. 2020 [cited 13 Aug 2020]. doi:10.1073/pnas.2003383117
- 937 4. Harris JA, Mihalas S, Hirokawa KE, Whitesell JD, Choi H, Bernard A, et al. Hierarchical  
938 organization of cortical and thalamic connectivity. *Nature*. 2019;575: 195–202.  
939 doi:10.1038/s41586-019-1716-z
- 940 5. Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G, et al. Situating the  
941 default-mode network along a principal gradient of macroscale cortical organization. *PNAS*.  
942 2016;113: 12574–12579. doi:10.1073/pnas.1608282113
- 943 6. Burt JB, Demirtaş M, Eckner WJ, Navejar NM, Ji JL, Martin WJ, et al. Hierarchy of transcriptomic  
944 specialization across human cortex captured by structural neuroimaging topography. *Nat*  
945 *Neurosci*. 2018;21: 1251–1259. doi:10.1038/s41593-018-0195-0
- 946 7. Lein ES, Belgard TG, Hawrylycz M, Molnár Z. Transcriptomic Perspectives on Neocortical  
947 Structure, Development, Evolution, and Disease. *Annu Rev Neurosci*. 2017;40: 629–652.  
948 doi:10.1146/annurev-neuro-070815-013858
- 949 8. Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. The Cellular and Molecular Landscapes of  
950 the Developing Human Central Nervous System. *Neuron*. 2016;89: 248–268.  
951 doi:10.1016/j.neuron.2015.12.008
- 952 9. Bishop KM, Rubenstein JLR, O’Leary DDM. Distinct Actions of Emx1, Emx2, and Pax6 in  
953 Regulating the Specification of Areas in the Developing Neocortex. *J Neurosci*. 2002;22: 7627–  
954 7638. doi:10.1523/JNEUROSCI.22-17-07627.2002
- 955 10. Kiecker C, Lumsden A. Hedgehog signaling from the ZLI regulates diencephalic regional identity.  
956 *Nat Neurosci*. 2004;7: 1242–1249. doi:10.1038/nn1338
- 957 11. Bystron I, Blakemore C, Rakic P. Development of the human cerebral cortex: Boulder Committee  
958 revisited. *Nat Rev Neurosci*. 2008;9: 110–122. doi:10.1038/nrn2252
- 959 12. Cadwell CR, Bhaduri A, Mostajo-Radji MA, Keefe MG, Nowakowski TJ. Development and  
960 Arealization of the Cerebral Cortex. *Neuron*. 2019;103: 980–1004.  
961 doi:10.1016/j.neuron.2019.07.009
- 962 13. Meyer G, Schaaps JP, Moreau L, Goffinet AM. Embryonic and Early Fetal Development of the  
963 Human Neocortex. *J Neurosci*. 2000;20: 1858–1868. doi:10.1523/JNEUROSCI.20-05-  
964 01858.2000
- 965 14. Kostovic I, Jovanov-Milosevic N. The development of cerebral connections during the first 20-45  
966 weeks’ gestation. *SeminFetalNeonatal Med*. 2006;11: 415–422. doi:10.1016/j.siny.2006.07.001
- 967 15. Kostović I, Jovanov-Milošević N, Radoš M, Sedmak G, Benjak V, Kostović-Srzić M, et al.  
968 Perinatal and early postnatal reorganization of the subplate and related cellular compartments in  
969 the human cerebral wall as revealed by histological and MRI approaches. *Brain Struct Funct*.  
970 2014;219: 231–253. doi:10.1007/s00429-012-0496-0

- 971 16. Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional landscape of the  
972 prenatal human brain. *Nature*. 2014;508: 199–206. doi:10.1038/nature13185
- 973 17. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the  
974 human brain. *Nature*. 2011;478: 483–489. doi:10.1038/nature10523
- 975 18. Li M, Santpere G, Imamura Kawasawa Y, Evgrafov OV, Gulden FO, Pochareddy S, et al.  
976 Integrative functional genomic analysis of human brain development and neuropsychiatric risks.  
977 *Science*. 2018;362. doi:10.1126/science.aat7615
- 978 19. Pletikos M, Sousa AMM, Sedmak G, Meyer KA, Zhu Y, Cheng F, et al. Temporal Specification  
979 and Bilaterality of Human Neocortical Topographic Gene Expression. *Neuron*. 2014;81: 321–  
980 332. doi:10.1016/j.neuron.2013.11.018
- 981 20. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. Shared molecular  
982 neuropathology across major psychiatric disorders parallels polygenic overlap. *Science*.  
983 2018;359: 693–697. doi:10.1126/science.aad6469
- 984 21. Velmeshev D, Schirmer L, Jung D, Haeussler M, Perez Y, Mayer S, et al. Single-cell genomics  
985 identifies cell type-specific molecular changes in autism. *Science*. 2019;364: 685–689.  
986 doi:10.1126/science.aav8130
- 987 22. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. An anatomically  
988 comprehensive atlas of the adult human brain transcriptome. *Nature*. 2012;489: 391–399.  
989 doi:10.1038/nature11405
- 990 23. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, et al. Comprehensive functional genomic  
991 resource and integrative model for the human brain. *Science*. 2018;362.  
992 doi:10.1126/science.aat8464
- 993 24. Fornito A, Arnatkevičiūtė A, Fulcher BD. Bridging the Gap between Connectome and  
994 Transcriptome. *Trends Cogn Sci (Regul Ed)*. 2019;23: 34–50. doi:10.1016/j.tics.2018.10.005
- 995 25. Shin J, French L, Xu T, Leonard G, Perron M, Pike GB, et al. Cell-Specific Gene-Expression  
996 Profiles and Cortical Thickness in the Human Brain. *Cereb Cortex*. 2018;28: 3267–3277.  
997 doi:10.1093/cercor/bhx197
- 998 26. Richiardi J, Altmann A, Milazzo A-C, Chang C, Chakravarty MM, Banaschewski T, et al.  
999 Correlated gene expression supports synchronous activity in brain networks. *Science*.  
1000 2015;348: 1241–1244. doi:10.1126/science.1255905
- 1001 27. Romero-Garcia R, Warrier V, Bullmore ET, Baron-Cohen S, Bethlehem RAI. Synaptic and  
1002 transcriptionally downregulated genes are associated with cortical thickness differences in  
1003 autism. *Molecular Psychiatry*. 2018; 1. doi:10.1038/s41380-018-0023-7
- 1004 28. Makropoulos A, Robinson EC, Schuh A, Wright R, Fitzgibbon S, Bozek J, et al. The developing  
1005 human connectome project: A minimal processing pipeline for neonatal cortical surface  
1006 reconstruction. *NeuroImage*. 2018;173: 88–112. doi:10.1016/j.neuroimage.2018.01.054
- 1007 29. Hughes EJ, Winchman T, Padormo F, Teixeira R, Wurie J, Sharma M, et al. A dedicated neonatal  
1008 brain imaging system. *Magnetic Resonance in Medicine*. 2017;78: 794–804.  
1009 doi:10.1002/mrm.26462
- 1010 30. Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, et al. A Structural MRI Study  
1011 of Human Brain Development from Birth to 2 Years. *J Neurosci*. 2008;28: 12176–12182.  
1012 doi:10.1523/JNEUROSCI.3479-08.2008
- 1013 31. Gilmore JH, Knickmeyer RC, Gao W. Imaging structural and functional brain development in early  
1014 childhood. *Nat Rev Neurosci*. 2018;19: 123–137. doi:10.1038/nrn.2018.1

- 1015 32. Kapellou O, Counsell SJ, Kennea N, Dyet L, Saeed N, Stark J, et al. Abnormal cortical  
1016 development after premature birth shown by altered allometric scaling of brain growth. *PLoS*  
1017 *Med.* 2006;3: e265. doi:10.1371/journal.pmed.0030265
- 1018 33. Wang F, Lian C, Wu Z, Zhang H, Li T, Meng Y, et al. Developmental topography of cortical  
1019 thickness during infancy. *PNAS.* 2019;116: 15855–15860. doi:10.1073/pnas.1821523116
- 1020 34. Deoni SCL, Dean DC, O’Muircheartaigh J, Dirks H, Jerskey BA. Investigating white matter  
1021 development in infancy and early childhood using myelin water fraction and relaxation time  
1022 mapping. *Neuroimage.* 2012;63: 1038–1053. doi:10.1016/j.neuroimage.2012.07.037
- 1023 35. Counsell SJ, Maalouf EF, Fletcher AM, Duggan P, Battin M, Lewis HJ, et al. MR Imaging  
1024 Assessment of Myelination in the Very Preterm Brain. *AJNR Am J Neuroradiol.* 2002;23: 872–  
1025 881.
- 1026 36. Doria V, Beckmann CF, Arichi T, Merchant N, Groppo M, Turkheimer FE, et al. Emergence of  
1027 resting state networks in the preterm human brain. *Proc Natl Acad Sci U S A.* 2010;107: 20015–  
1028 20020. doi:10.1073/pnas.1007921107
- 1029 37. Fransson P, Åden U, Blennow M, Lagercrantz H. The Functional Architecture of the Infant Brain  
1030 as Revealed by Resting-State fMRI. *Cereb Cortex.* 2011;21: 145–154.  
1031 doi:10.1093/cercor/bhq071
- 1032 38. Allievi AG, Arichi T, Tusor N, Kimpton J, Arulkumaran S, Counsell SJ, et al. Maturation of  
1033 Sensori-Motor Functional Responses in the Preterm Brain. *Cereb Cortex.* 2016;26: 402–413.  
1034 doi:10.1093/cercor/bhv203
- 1035 39. Wilcox T, Haslup JA, Boas DA. Dissociation of processing of featural and spatiotemporal  
1036 information in the infant cortex. *Neuroimage.* 2010;53: 1256–1263.  
1037 doi:10.1016/j.neuroimage.2010.06.064
- 1038 40. Cao M, He Y, Dai Z, Liao X, Jeon T, Ouyang M, et al. Early Development of Functional Network  
1039 Segregation Revealed by Connectomic Analysis of the Preterm Human Brain. *Cereb Cortex.*  
1040 2017;27: 1949–1963. doi:10.1093/cercor/bhw038
- 1041 41. Karen T, Morren G, Haensse D, Bauschatz AS, Bucher HU, Wolf M. Hemodynamic response to  
1042 visual stimulation in newborn infants using functional near- infrared spectroscopy. *Hum Brain*  
1043 *Mapp.* 2007;29: 453–460. doi:10.1002/hbm.20411
- 1044 42. Ouyang M, Jeon T, Sotiras A, Peng Q, Mishra V, Halovanic C, et al. Differential cortical  
1045 microstructural maturation in the preterm human brain with diffusion kurtosis and tensor  
1046 imaging. *PNAS.* 2019;116: 4681–4688. doi:10.1073/pnas.1812156116
- 1047 43. Ball G, Srinivasan L, Aljabar P, Counsell SJ, Durighel G, Hajnal JV, et al. Development of cortical  
1048 microstructure in the preterm human brain. *Proc Natl Acad Sci USA.* 2013;110: 9541–9546.  
1049 doi:10.1073/pnas.1301652110
- 1050 44. Batalle D, O’Muircheartaigh J, Makropoulos A, Kelly CJ, Dimitrova R, Hughes EJ, et al. Different  
1051 patterns of cortical maturation before and after 38 weeks gestational age demonstrated by  
1052 diffusion MRI in vivo. *NeuroImage.* 2019;185: 764–775. doi:10.1016/j.neuroimage.2018.05.046
- 1053 45. Ouyang M, Dubois J, Yu Q, Mukherjee P, Huang H. Delineation of early brain development from  
1054 fetuses to infants with diffusion MRI and beyond. *Neuroimage.* 2019;185: 836–850.  
1055 doi:10.1016/j.neuroimage.2018.04.017
- 1056 46. Yu Q, Ouyang A, Chalak L, Jeon T, Chia J, Mishra V, et al. Structural Development of Human  
1057 Fetal and Preterm Brain Cortical Plate Based on Population-Averaged Templates. *Cereb*  
1058 *Cortex.* 2016;26: 4381–4391. doi:10.1093/cercor/bhv201

- 1059 47. Ball G, Boardman JP, Rueckert D, Aljabar P, Arichi T, Merchant N, et al. The effect of preterm  
1060 birth on thalamic and cortical development. *Cereb Cortex*. 2012;22: 1016–1024.  
1061 doi:10.1093/cercor/bhr176
- 1062 48. Bouyssi-Kobar M, Brossard-Racine M, Jacobs M, Murnick J, Chang T, Limperopoulos C.  
1063 Regional microstructural organization of the cerebral cortex is affected by preterm birth.  
1064 *NeuroImage: Clinical*. 2018;18: 871–880. doi:10.1016/j.nicl.2018.03.020
- 1065 49. Eaton-Rosen Z, Melbourne A, Orasanu E, Cardoso MJ, Modat M, Bainbridge A, et al.  
1066 Longitudinal measurement of the developing grey matter in preterm subjects using multi-modal  
1067 MRI. *Neuroimage*. 2015;111: 580–589. doi:10.1016/j.neuroimage.2015.02.010
- 1068 50. Huang H, Jeon T, Sedmak G, Pletikos M, Vasung L, Xu X, et al. Coupling Diffusion Imaging with  
1069 Histological and Gene Expression Analysis to Examine the Dynamics of Cortical Areas across  
1070 the Fetal Period of Human Brain Development. *Cereb Cortex*. 2013;23: 2620–2631.  
1071 doi:10.1093/cercor/bhs241
- 1072 51. Fan X, Dong J, Zhong S, Wei Y, Wu Q, Yan L, et al. Spatial transcriptomic survey of human  
1073 embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Res*. 2018;28: 730–745.  
1074 doi:10.1038/s41422-018-0053-3
- 1075 52. Nowakowski TJ, Bhaduri A, Pollen AA, Alvarado B, Mostajo-Radji MA, Di Lullo E, et al.  
1076 Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human  
1077 cortex. *Science*. 2017;358: 1318–1323. doi:10.1126/science.aap8809
- 1078 53. Polioudakis D, de la Torre-Ubieta L, Langerman J, Elkins AG, Shi X, Stein JL, et al. A Single-Cell  
1079 Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron*.  
1080 2019;103: 785-801.e8. doi:10.1016/j.neuron.2019.06.011
- 1081 54. Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH, et al. Low-coverage single-cell  
1082 mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in  
1083 developing cerebral cortex. *Nat Biotechnol*. 2014;32: 1053–1058. doi:10.1038/nbt.2967
- 1084 55. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive,  
1085 powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res*.  
1086 2017;45: W130–W137. doi:10.1093/nar/gkx356
- 1087 56. Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, et al. Temporal dynamics and genetic  
1088 control of transcription in the human prefrontal cortex. *Nature*. 2011;478: 519–523.  
1089 doi:10.1038/nature10524
- 1090 57. Kubo K-I, Deguchi K, Nagai T, Ito Y, Yoshida K, Endo T, et al. Association of impaired neuronal  
1091 migration with cognitive deficits in extremely preterm infants. *JCI Insight*. 2017;2.  
1092 doi:10.1172/jci.insight.88609
- 1093 58. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-  
1094 protein association networks with increased coverage, supporting functional discovery in  
1095 genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47: D607–D613.  
1096 doi:10.1093/nar/gky1131
- 1097 59. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome  
1098 Pathway Knowledgebase. *Nucleic Acids Res*. 2018;46: D649–D655. doi:10.1093/nar/gkx1132
- 1099 60. Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex.  
1100 *CerebCortex*. 1991;1: 1–47.
- 1101 61. Cahalane DJ, Charvet CJ, Finlay BL. Systematic, balancing gradients in neuron density and  
1102 number across the primate isocortex. *Front Neuroanat*. 2012;6. doi:10.3389/fnana.2012.00028

- 1103 62. Gomez J, Zhen Z, Weiner KS. Human visual cortex is organized along two genetically opposed  
1104 hierarchical gradients with unique developmental and evolutionary origins. *PLOS Biology*.  
1105 2019;17: e3000362. doi:10.1371/journal.pbio.3000362
- 1106 63. Wagstyl K, Ronan L, Goodyer IM, Fletcher PC. Cortical thickness gradients in structural  
1107 hierarchies. *Neuroimage*. 2015;111: 241–250. doi:10.1016/j.neuroimage.2015.02.036
- 1108 64. Larivière S, Vos de Wael R, Hong S-J, Paquola C, Tavakol S, Lowe AJ, et al. Multiscale  
1109 Structure–Function Gradients in the Neonatal Connectome. *Cereb Cortex*. 2019 [cited 29 Nov  
1110 2019]. doi:10.1093/cercor/bhz069
- 1111 65. Huntenburg JM, Bazin P-L, Goulas A, Tardif CL, Villringer A, Margulies DS. A Systematic  
1112 Relationship Between Functional Connectivity and Intracortical Myelin in the Human Cerebral  
1113 Cortex. *Cereb Cortex*. 2017;27: 981–997. doi:10.1093/cercor/bhx030
- 1114 66. Fonta C, Imbert M. Vascularization in the Primate Visual Cortex during Development. *Cereb*  
1115 *Cortex*. 2002;12: 199–211. doi:10.1093/cercor/12.2.199
- 1116 67. Travis K, Ford K, Jacobs B. Regional dendritic variation in neonatal human cortex: a quantitative  
1117 Golgi study. *Dev Neurosci*. 2005;27: 277–287. doi:10.1159/000086707
- 1118 68. Jakovcevski I, Zecevic N. Sequence of oligodendrocyte development in the human fetal  
1119 telencephalon. *Glia*. 2005;49: 480–491. doi:10.1002/glia.20134
- 1120 69. Verney C, Monier A, Fallet-Bianco C, Gressens P. Early microglial colonization of the human  
1121 forebrain and possible involvement in periventricular white-matter injury of preterm infants. *J*  
1122 *Anat*. 2010;217: 436–448. doi:10.1111/j.1469-7580.2010.01245.x
- 1123 70. Reemst K, Noctor SC, Lucassen PJ, Hol EM. The Indispensable Roles of Microglia and  
1124 Astrocytes during Brain Development. *Front Hum Neurosci*. 2016;10.  
1125 doi:10.3389/fnhum.2016.00566
- 1126 71. Huttenlocher PR, Dabholkar AS. Regional differences in synaptogenesis in human cerebral  
1127 cortex. *Journal of Comparative Neurology*. 1997;387: 167–178. doi:10.1002/(SICI)1096-  
1128 9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z
- 1129 72. Guillery RW. Is postnatal neocortical maturation hierarchical? *Trends in Neurosciences*. 2005;28:  
1130 512–517. doi:10.1016/j.tins.2005.08.006
- 1131 73. Dubois J, Benders M, Cachia A, Lazeyras F, Ha-Vinh Leuchter R, Sizonenko SV, et al. Mapping  
1132 the early cortical folding process in the preterm newborn brain. *Cereb Cortex*. 2008;18: 1444–  
1133 1454. doi:10.1093/cercor/bhm180
- 1134 74. Garbett KA, Hsiao EY, Kálmán S, Patterson PH, Mirnics K. Effects of maternal immune activation  
1135 on gene expression patterns in the fetal brain. *Transl Psychiatry*. 2012;2: e98.  
1136 doi:10.1038/tp.2012.24
- 1137 75. Tilley SK, Joseph RM, Kuban KCK, Dammann OU, O’Shea TM, Fry RC. Genomic biomarkers of  
1138 prenatal intrauterine inflammation in umbilical cord tissue predict later life neurological  
1139 outcomes. *PLoS One*. 2017;12. doi:10.1371/journal.pone.0176953
- 1140 76. Edlow AG, Guedj F, Sverdlov D, Pennings JLA, Bianchi DW. Significant Effects of Maternal Diet  
1141 During Pregnancy on the Murine Fetal Brain Transcriptome and Offspring Behavior. *Front*  
1142 *Neurosci*. 2019;13. doi:10.3389/fnins.2019.01335
- 1143 77. Miguel PM, Pereira LO, Silveira PP, Meaney MJ. Early environmental influences on the  
1144 development of children’s brain structure and function. *Developmental Medicine & Child*  
1145 *Neurology*. 2019;61: 1127–1133. doi:10.1111/dmcn.14182



- 1146 78. Schneider C, Krischke G, Rascher W, Gassmann M, Trollmann R. Systemic hypoxia differentially  
1147 affects neurogenesis during early mouse brain maturation. *Brain Dev.* 2012;34: 261–273.  
1148 doi:10.1016/j.braindev.2011.07.006
- 1149 79. Schwartz ML, Vaccarino F, Chacon M, Yan WL, Ment LR, Stewart WB. Chronic neonatal hypoxia  
1150 leads to long term decreases in the volume and cell number of the rat cerebral cortex. *Semin*  
1151 *Perinatol.* 2004;28: 379–388. doi:10.1053/j.semperi.2004.10.009
- 1152 80. Glasser MF, Van Essen DC. Mapping Human Cortical Areas In Vivo Based on Myelin Content as  
1153 Revealed by T1- and T2-Weighted MRI. *J Neurosci.* 2011;31: 11597–11616.  
1154 doi:10.1523/JNEUROSCI.2180-11.2011
- 1155 81. Pecheva D, Lee A, Poh JS, Chong Y-S, Shek LP, Gluckman PD, et al. Neural Transcription  
1156 Correlates of Multimodal Cortical Phenotypes during Development. *Cereb Cortex.* [cited 15 Jan  
1157 2020]. doi:10.1093/cercor/bhz271
- 1158 82. Govek E-E, Hatten ME, Van Aelst L. The Role of Rho GTPase Proteins in CNS Neuronal  
1159 Migration. *Dev Neurobiol.* 2011;71: 528–553. doi:10.1002/dneu.20850
- 1160 83. Barres BA, Raff MC, Gaese F, Bartke I, Dechant G, Barde YA. A crucial role for neurotrophin-3 in  
1161 oligodendrocyte development. *Nature.* 1994;367: 371–375. doi:10.1038/367371a0
- 1162 84. Feltri ML, Suter U, Relvas JB. The function of RhoGTPases in axon ensheathment and  
1163 myelination. *Glia.* 2008;56: 1508–1517. doi:10.1002/glia.20752
- 1164 85. Yao L, Kan EM, Kaur C, Dheen ST, Hao A, Lu J, et al. Notch-1 Signaling Regulates Microglia  
1165 Activation via NF- $\kappa$ B Pathway after Hypoxic Exposure In Vivo and In Vitro. *PLOS ONE.* 2013;8:  
1166 e78439. doi:10.1371/journal.pone.0078439
- 1167 86. Ishii A, Furusho M, Bansal R. Sustained Activation of ERK1/2 MAPK in Oligodendrocytes and  
1168 Schwann Cells Enhances Myelin Growth and Stimulates Oligodendrocyte Progenitor Expansion.  
1169 *J Neurosci.* 2013;33: 175–186. doi:10.1523/JNEUROSCI.4403-12.2013
- 1170 87. Samuels IS, Karlo JC, Faruzzi AN, Pickering K, Herrup K, Sweatt JD, et al. Deletion of ERK2  
1171 Mitogen-Activated Protein Kinase Identifies Its Key Roles in Cortical Neurogenesis and  
1172 Cognitive Function. *J Neurosci.* 2008;28: 6983–6995. doi:10.1523/JNEUROSCI.0679-08.2008
- 1173 88. Dommergues M-A, Plaisant F, Verney C, Gressens P. Early microglial activation following  
1174 neonatal excitotoxic brain damage in mice: a potential target for neuroprotection. *Neuroscience.*  
1175 2003;121: 619–628. doi:10.1016/s0306-4522(03)00558-x
- 1176 89. Back SA, Luo NL, Borenstein NS, Levine JM, Volpe JJ, Kinney HC. Late oligodendrocyte  
1177 progenitors coincide with the developmental window of vulnerability for human perinatal white  
1178 matter injury. *JNeurosci.* 2001;21: 1302–1312.
- 1179 90. Baburamani AA, Supramaniam VG, Hagberg H, Mallard C. Microglia toxicity in preterm brain  
1180 injury. *Reproductive Toxicology.* 2014;48: 106–112. doi:10.1016/j.reprotox.2014.04.002
- 1181 91. Krishnan ML, Wang Z, Aljabar P, Ball G, Mirza G, Saxena A, et al. Machine learning shows  
1182 association between genetic variability in PPAR $\gamma$  and cerebral connectivity in preterm infants.  
1183 *PNAS.* 2017;114: 13744–13749. doi:10.1073/pnas.1704907114
- 1184 92. Van Steenwinckel J, Schang A-L, Krishnan ML, Degos V, Delahaye-Duriez A, Bokobza C, et al.  
1185 Decreased microglial Wnt/ $\beta$ -catenin signalling drives microglial pro-inflammatory activation in  
1186 the developing brain. *Brain.* 2019;142: 3806–3833. doi:10.1093/brain/awz319
- 1187 93. Squarzoni P, Oller G, Hoeffel G, Pont-Lezica L, Rostaing P, Low D, et al. Microglia Modulate  
1188 Wiring of the Embryonic Forebrain. *Cell Reports.* 2014;8: 1271–1279.  
1189 doi:10.1016/j.celrep.2014.07.042

- 1190 94. Stolp HB, Fleiss B, Arai Y, Supramaniam V, Vontell R, Birtles S, et al. Interneuron Development  
1191 Is Disrupted in Preterm Brains With Diffuse White Matter Injury: Observations in Mouse and  
1192 Human. *Front Physiol.* 2019;10. doi:10.3389/fphys.2019.00955
- 1193 95. Fleiss B, Gressens P, Stolp HB. Cortical Gray Matter Injury in Encephalopathy of Prematurity:  
1194 Link to Neurodevelopmental Disorders. *Front Neurol.* 2020;11: 575.  
1195 doi:10.3389/fneur.2020.00575
- 1196 96. Moretti R, Pansiot J, Bettati D, Strazielle N, Ghersi-Egea J-F, Damante G, et al. Blood-brain  
1197 barrier dysfunction in disorders of the developing brain. *Front Neurosci.* 2015;9.  
1198 doi:10.3389/fnins.2015.00040
- 1199 97. Vasung L, Rollins CK, Velasco-Annis C, Yun HJ, Zhang J, Warfield SK, et al. Spatiotemporal  
1200 Differences in the Regional Cortical Plate and Subplate Volume Growth during Fetal  
1201 Development. *Cereb Cortex.* 2020;30: 4438–4453. doi:10.1093/cercor/bhaa033
- 1202 98. Wagstyl K, Larocque S, Cucurull G, Lepage C, Cohen JP, Bludau S, et al. BigBrain 3D atlas of  
1203 cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices.  
1204 *PLOS Biology.* 2020;18: e3000678. doi:10.1371/journal.pbio.3000678
- 1205 99. Vasung L, Rollins CK, Yun HJ, Velasco-Annis C, Zhang J, Wagstyl K, et al. Quantitative In vivo  
1206 MRI Assessment of Structural Asymmetries and Sexual Dimorphism of Transient Fetal  
1207 Compartments in the Human Brain. *Cereb Cortex.* 2020;30: 1752–1767.  
1208 doi:10.1093/cercor/bhz200
- 1209 100. Ritchie J, Pantazatos SP, French L. Transcriptomic characterization of MRI contrast with focus  
1210 on the T1-w/T2-w ratio in the cerebral cortex. *Neuroimage.* 2018;174: 504–517.  
1211 doi:10.1016/j.neuroimage.2018.03.027
- 1212 101. Wang G-Z, Belgard TG, Mao D, Chen L, Berto S, Preuss TM, et al. Correspondence between  
1213 resting state activity and brain gene expression. *Neuron.* 2015;88: 659–666.  
1214 doi:10.1016/j.neuron.2015.10.022
- 1215 102. Vértes PE, Rittman T, Whitaker KJ, Romero-Garcia R, Váša F, Kitzbichler MG, et al. Gene  
1216 transcription profiles associated with inter-modular hubs and connection distance in human  
1217 functional magnetic resonance imaging networks. *Philos Trans R Soc Lond B Biol Sci.*  
1218 2016;371. doi:10.1098/rstb.2015.0362
- 1219 103. Baum GL, Cui Z, Roalf DR, Ciric R, Betzel RF, Larsen B, et al. Development of structure–  
1220 function coupling in human brain networks during youth. *PNAS.* 2020;117: 771–778.  
1221 doi:10.1073/pnas.1912034117
- 1222 104. Vázquez-Rodríguez B, Suárez LE, Markello RD, Shafiei G, Paquola C, Hagmann P, et al.  
1223 Gradients of structure–function tethering across neocortex. *PNAS.* 2019;116: 21219–21227.  
1224 doi:10.1073/pnas.1903403116
- 1225 105. Cordero-Grande L, Hughes EJ, Hutter J, Price AN, Hajnal JV. Three-dimensional motion  
1226 corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to  
1227 neonatal brain imaging. *Magn Reson Med.* 2018;79: 1365–1376. doi:10.1002/mrm.26796
- 1228 106. Kuklisova-Murgasova M, Quaghebeur G, Rutherford MA, Hajnal JV, Schnabel JA.  
1229 Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Med  
1230 Image Anal.* 2012;16: 1550–1564. doi:10.1016/j.media.2012.07.004
- 1231 107. Hutter J, Tournier JD, Price AN, Cordero-Grande L, Hughes EJ, Malik S, et al. Time-efficient  
1232 and flexible design of optimized multishell HARDI diffusion. *Magn Reson Med.* 2018;79: 1276–  
1233 1292. doi:10.1002/mrm.26765

- 1234 108. Zhu K, Dougherty RF, Wu H, Middione MJ, Takahashi AM, Zhang T, et al. Hybrid-Space  
1235 SENSE Reconstruction for Simultaneous Multi-Slice MRI. *IEEE Trans Med Imaging*. 2016;35:  
1236 1824–1836. doi:10.1109/TMI.2016.2531635
- 1237 109. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3  
1238 Bias Correction. *IEEE Trans Med Imaging*. 2010;29: 1310–1320.  
1239 doi:10.1109/TMI.2010.2046908
- 1240 110. Smith SM. Fast robust automated brain extraction. *Human Brain Mapping*. 2002;17: 143–155.  
1241 doi:10.1002/hbm.10062
- 1242 111. Makropoulos A, Gousias IS, Ledig C, Aljabar P, Serag A, Hajnal JV, et al. Automatic whole brain  
1243 MRI segmentation of the developing neonatal brain. *IEEE Trans Med Imaging*. 2014;33: 1818–  
1244 1831. doi:10.1109/TMI.2014.2322280
- 1245 112. Schuh A, Makropoulos A, Wright R, Robinson EC, Tusor N, Steinweg J, et al. A deformable  
1246 model for the reconstruction of the neonatal cortex. 2017 IEEE 14th International Symposium on  
1247 Biomedical Imaging (ISBI 2017). 2017. pp. 800–803. doi:10.1109/ISBI.2017.7950639
- 1248 113. Bozek J, Makropoulos A, Schuh A, Fitzgibbon S, Wright R, Glasser MF, et al. Construction of a  
1249 neonatal cortical surface atlas using Multimodal Surface Matching in the Developing Human  
1250 Connectome Project. *Neuroimage*. 2018;179: 11–29. doi:10.1016/j.neuroimage.2018.06.018
- 1251 114. Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, et al. MSM: a new  
1252 flexible framework for Multimodal Surface Matching. *Neuroimage*. 2014;100: 414–426.  
1253 doi:10.1016/j.neuroimage.2014.05.069
- 1254 115. Robinson EC, Garcia K, Glasser MF, Chen Z, Coalson TS, Makropoulos A, et al. Multimodal  
1255 surface matching with higher-order smoothness constraints. *Neuroimage*. 2018;167: 453–465.  
1256 doi:10.1016/j.neuroimage.2017.10.037
- 1257 116. Veraart J, Fieremans E, Novikov DS. Diffusion MRI noise mapping using random matrix theory.  
1258 *Magnetic Resonance in Medicine*. 2016;76: 1582–1593. doi:10.1002/mrm.26059
- 1259 117. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local  
1260 subvoxel-shifts. *Magnetic Resonance in Medicine*. 2016;76: 1574–1581.  
1261 doi:10.1002/mrm.26054
- 1262 118. Christiaens D, Cordero-Grande L, Pietsch M, Hutter J, Price AN, Hughes EJ, et al. Scattered  
1263 slice SHARD reconstruction for motion correction in multi-shell diffusion MRI of the neonatal  
1264 brain. arXiv:190502996 [physics]. 2019 [cited 14 Nov 2019]. Available:  
1265 <http://arxiv.org/abs/1905.02996>
- 1266 119. Tournier J-D, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, et al. MRtrix3: A fast,  
1267 flexible and open software framework for medical image processing and visualisation.  
1268 *NeuroImage*. 2019;202: 116137. doi:10.1016/j.neuroimage.2019.116137
- 1269 120. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite  
1270 orientation dispersion and density imaging of the human brain. *Neuroimage*. 2012;61: 1000–  
1271 1016. doi:10.1016/j.neuroimage.2012.03.072
- 1272 121. Kunz N, Zhang H, Vasung L, O'Brien KR, Assaf Y, Lazeyras F, et al. Assessing white matter  
1273 microstructure of the newborn with multi-shell diffusion MRI and biophysical compartment  
1274 models. *NeuroImage*. 2014;96: 288–299. doi:10.1016/j.neuroimage.2014.03.057
- 1275 122. Dimitrova R, Pietsch M, Christiaens D, Ciarrusta J, Wolfers T, Batalle D, et al. Heterogeneity in  
1276 Brain Microstructural Development Following Preterm Birth. *Cereb Cortex*. 2020;30: 4800–4810.  
1277 doi:10.1093/cercor/bhaa069

- 1278 123. Habegger L, Sboner A, Gianoulis TA, Rozowsky J, Agarwal A, Snyder M, et al. RSEQtools: a  
1279 modular framework to analyze RNA-Seq data using compact, anonymized data summaries.  
1280 *Bioinformatics*. 2011;27: 281–283. doi:10.1093/bioinformatics/btq643
- 1281 124. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional  
1282 quantile normalization. *Biostatistics*. 2012;13: 204–216. doi:10.1093/biostatistics/kxr054
- 1283 125. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using  
1284 empirical Bayes methods. *Biostatistics*. 2007;8: 118–127. doi:10.1093/biostatistics/kxj037
- 1285 126. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an  
1286 RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular  
1287 Biology*. 2006;7: 3. doi:10.1186/1471-2199-7-3
- 1288 127. SciPy.org — SciPy.org. [cited 2 Oct 2019]. Available: <https://www.scipy.org/>
- 1289 128. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn:  
1290 Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12: 2825–2830.
- 1291 129. nlme: Nonlinear Mixed-Effects Models in nlme: Linear and Nonlinear Mixed Effects Models.  
1292 [cited 2 Oct 2019]. Available: <https://rdrr.io/cran/nlme/man/nlme.html>
- 1293 130. Wood S. *Generalized Additive Models: An Introduction with R*. 2nd ed. CRC Press; 2017.  
1294 Available: [https://www.crcpress.com/Generalized-Additive-Models-An-Introduction-with-R-  
1295 Second-Edition/Wood/p/book/9781498728331](https://www.crcpress.com/Generalized-Additive-Models-An-Introduction-with-R-Second-Edition/Wood/p/book/9781498728331)
- 1296 131. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for  
1297 Dimension Reduction. arXiv:180203426 [cs, stat]. 2018 [cited 2 Oct 2019]. Available:  
1298 <http://arxiv.org/abs/1802.03426>
- 1299 132. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis.  
1300 *BMC Bioinformatics*. 2008;9: 559. doi:10.1186/1471-2105-9-559
- 1301 133. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis.  
1302 *Stat Appl Genet Mol Biol*. 2005;4: Article17. doi:10.2202/1544-6115.1128
- 1303 134. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the  
1304 Dynamic Tree Cut package for R. *Bioinformatics*. 2008;24: 719–720.  
1305 doi:10.1093/bioinformatics/btm563
- 1306
- 1307