

## ORIGINAL ARTICLE

## Epigenomic profiling of preterm infants reveals DNA methylation differences at sites associated with neural function

S Sparrow<sup>1</sup>, JR Manning<sup>2</sup>, J Cartier<sup>3</sup>, D Anblagan<sup>4</sup>, ME Bastin<sup>4</sup>, C Piyasena<sup>3</sup>, R Pataky<sup>1</sup>, EJ Moore<sup>1</sup>, SI Semple<sup>5</sup>, AG Wilkinson<sup>6</sup>, M Evans<sup>7</sup>, AJ Drake<sup>3,8</sup> and JP Boardman<sup>1,4,8</sup>

DNA methylation (DNAm) plays a determining role in neural cell fate and provides a molecular link between early-life stress and neuropsychiatric disease. Preterm birth is a profound environmental stressor that is closely associated with alterations in connectivity of neural systems and long-term neuropsychiatric impairment. The aims of this study were to examine the relationship between preterm birth and DNAm, and to investigate factors that contribute to variance in DNAm. DNA was collected from preterm infants (birth < 33 weeks gestation) and healthy controls (birth > 37 weeks), and a genome-wide analysis of DNAm was performed; diffusion magnetic resonance imaging (dMRI) data were acquired from the preterm group. The major fasciculi were segmented, and fractional anisotropy, mean diffusivity and tract shape were calculated. Principal components (PC) analysis was used to investigate the contribution of MRI features and clinical variables to variance in DNAm. Differential methylation was found within 25 gene bodies and 58 promoters of protein-coding genes in preterm infants compared with controls; 10 of these have neural functions. Differences detected in the array were validated with pyrosequencing. Ninety-five percent of the variance in DNAm in preterm infants was explained by 23 PCs; corticospinal tract shape associated with 6th PC, and gender and early nutritional exposure associated with the 7th PC. Preterm birth is associated with alterations in the methylome at sites that influence neural development and function. Differential methylation analysis has identified several promising candidate genes for understanding the genetic/epigenetic basis of preterm brain injury.

*Translational Psychiatry* (2016) 6, e716; doi:10.1038/tp.2015.210; published online 19 January 2016

## INTRODUCTION

Preterm birth affects 5–13% of newborns,<sup>1</sup> and is a profound early-life stressor that is closely associated with cerebral palsy, cognitive impairment, autism spectrum disorder and psychiatric disease.<sup>2–6</sup> The prevalence of impairment is related to gestational age at birth and to adverse exposures such as inflammation, ischaemia, respiratory morbidity and sub-optimal nutrition,<sup>7</sup> but the mechanisms underlying these associations are poorly understood.

Epigenetic modification has a fundamental role in regulating gene expression and determining neural cell fate, and DNA methylation (DNAm) is one such modification that is highly conserved across species.<sup>8</sup> DNAm is dynamic during development, including in the brain<sup>9</sup> and this could provide a mechanism by which environmental factors lead to disturbances of neural development that underpin later impairment.<sup>10</sup> DNAm mediates gene–environment interactions between early-life stress and several neuropsychiatric outcomes,<sup>11–14</sup> but little is known about DNAm in relation to brain development after preterm birth.

Although DNAm patterns are tissue specific, a number of recent observations suggest consistency between peripheral tissues and brain. First, DNAm profiles are altered consistently between prefrontal cortex and T cells in a rhesus macaque model of early-life stress.<sup>15</sup> Second, the top enriched biological processes

from peripheral blood cells of adults with post-traumatic stress disorder and early-life trauma concern central nervous system development,<sup>12</sup> which suggests considerable overlap between tissues. Third, inter-individual variation tends to be consistent across tissue types.<sup>16</sup> Furthermore, sampling DNA from saliva rather than blood is informative in brain DNAm studies because: methylation profiles obtained from saliva show greater correspondence with brain tissue extracts than those obtained from blood;<sup>17</sup> inherent properties of DNAm from buccal cells (greater enrichment of DNaseI hypersensitivity sites, histone modifications and disease-associated single nucleotide polymorphisms (SNPs)) may make them a more favourable proxy tissue than blood for epigenome-wide association studies of non-haematological disease.<sup>18</sup>

Structural and diffusion magnetic resonance imaging (dMRI) reveal a cerebral signature of preterm birth that includes reduced connectivity of white matter tracts, focal tissue volume reduction in deep grey matter nuclei and reduced cortical complexity.<sup>19–24</sup> Specifically, fractional anisotropy (FA) and mean diffusivity ( $\langle D \rangle$ ) derived from dMRI provide measures of tract integrity in the newborn brain that have a predictable pattern of alteration in preterm infants at term equivalent age (TEA).<sup>25–27</sup> These biomarkers are sensitive to genetic and environmental risk modulators for injury, and can detect neuroprotective treatment effects.<sup>28–31</sup>

<sup>1</sup>MRC Centre for Reproductive Health, University of Edinburgh, Queen's Medical Research Institute, Edinburgh, UK; <sup>2</sup>MRC Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK; <sup>3</sup>University/BHF Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK; <sup>4</sup>Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK; <sup>5</sup>Clinical Research Imaging Centre, University of Edinburgh, Edinburgh, UK; <sup>6</sup>Department of Radiology, NHS Lothian, Edinburgh, UK and <sup>7</sup>Department of Pathology, NHS Lothian, Edinburgh, UK. Correspondence: Dr JP Boardman, MRC Centre for Reproductive Health, University of Edinburgh, Queen's Medical Research Institute, 47 Little France Crescent, Room W1.26, Edinburgh EH16 4TJ, UK.

E-mail: james.boardman@ed.ac.uk

<sup>8</sup>These authors contributed equally to this work.

Received 1 September 2015; revised 18 November 2015; accepted 19 November 2015

Probabilistic neighbourhood tractography (PNT) is an automatic segmentation technique, based on single seed point tractography, that can identify the same fasciculus-of-interest across a group of subjects by modelling how individual tracts compare with a predefined reference tract in terms of length and shape.<sup>32,33</sup> This tract shape modelling is unique to PNT and allows not only measurement of tract integrity parameters, such as tract-averaged  $\langle D \rangle$  and FA, but also provides a metric, the absolute goodness-of-fit of the segmented tract to the reference ( $R$ ), which can be used to quantify differences in tract shape between individuals. The inclusion of anatomic information in dMRI models makes PNT ideally suited to studies of genetic and epigenetic effects because brain structure is heritable.<sup>34,35</sup>

In this study, we tested the hypotheses that the stress of preterm birth leads to alterations in the methylome that are apparent early in the newborn period, and variance in DNAm is associated with dMRI parameters in major white matter tracts and clinical risk factors for adverse outcome.

## MATERIALS AND METHODS

### Participants

The study was conducted according to the principles of the Declaration of Helsinki, and ethical approval was obtained from the UK National Research Ethics Service. Written parental informed consent was obtained.

The cohort consisted of two groups of neonates who received care at the Royal Infirmary of Edinburgh between January 2012 and September 2014: (1) preterm neonates (defined as postmenstrual age (PMA) at birth < 32 completed weeks gestation); and control infants born at full term (> 37 weeks PMA). Infants were not eligible if they had dysmorphic features suggestive of a chromosomal abnormality that was confirmed by karyotype, a congenital malformation or a congenital infection.

### DNA extraction

The DNA OG-575 kit was used for sampling of saliva at TEA, defined as 38–42 weeks PMA (DNAGenotek, Ottawa, ON, Canada). DNA was extracted using an alcohol precipitation technique as per manufacturer's instructions, and was rehydrated in TE 0.5. Gel electrophoresis was used to qualify DNA extraction and Qubit 2.0 Fluorometer was utilised for quantification of DNA concentration (Invitrogen Life Sciences, Carlsbad, CA, USA).

### DNAm analysis

DNAm analysis was performed at the Genetics Core of the Edinburgh Clinical Research Facility (Edinburgh, UK). Bisulphite conversion of 500 ng input DNA was carried out using the EZ DNAm Kit (Zymo Research, Freiburg, Germany). Four microlitres of bisulphite-converted DNA was processed using the Infinium HD Assay for Methylation (Illumina Methylation 450k beadchip and Infinium chemistry (Illumina, San Diego, CA, USA)). Each sample was interrogated on the arrays against 485 000 methylation sites. The arrays were imaged on the Illumina HiScan platform and genotypes were called automatically using GenomeStudio Analysis software version 2011.1 (Illumina). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE72120 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72120>).

### MRI acquisition

The preterm infants underwent brain MRI at TEA. A Siemens MAGNETOM Verio 3T MRI clinical scanner (Siemens Healthcare, Erlangen, Germany) and 12-channel phased-array head coil were used to acquire: T1-weighted MPRAGE volume (~1 mm<sup>3</sup> resolution), T2-weighted STIR (~0.9 mm<sup>3</sup> resolution), T2-weighted FLAIR (~1 mm<sup>3</sup> resolution) and dMRI (11 T2- and 64 diffusion encoding direction ( $b = 750 \text{ s mm}^{-2}$ ) single-shot spin-echo echo planar imaging volumes with 2 mm isotropic voxels using a prototype sequence). All examinations were supervised by a paediatrician experienced in MRI procedures. Infants were examined in natural sleep and pulse oximetry, temperature and electrocardiography data were monitored throughout. Ear protection was used for each infant, comprising earplugs placed in the external ear and neonatal earmuffs (MiniMuffs, Natus Medical, San Carlos, CA, USA).

### Methylation analysis

Data were processed with the RnBeads tool<sup>36</sup> v 0.99.17, for processes including data loading, pre-processing, normalisation and differential methylation. The software was modified slightly to allow filtering of problematic CpG loci identified by Chen *et al.*<sup>37</sup>

During pre-processing and using the Chen annotation, probes were removed if their CpG loci overlapped with known SNPs from the 1000 Genomes Project ([www.1000genomes.org](http://www.1000genomes.org)), if a SNP occurred at the site of single-base extension, or if the probe had been shown to be non-specific. A further, smaller set of probes was removed where RnBeads' iterative 'GreedyCut' algorithm identified a large number of unreliable measurements across samples.

Normalisation was carried out using the beta mixture quantile dilation method of Teschendorff *et al.*<sup>38</sup> in which the differing distributions of Type I and Type II probes is taken into consideration. This method had been shown to perform well in comparison with other methods.<sup>39</sup> Following normalisation a batch correction was applied by use of ComBat<sup>40</sup> to account for the well-documented chip-wise batch effect of the Infinium platform. Before downstream analyses, probes were removed if their target sites occurred on sex chromosomes, or in non-CpG contexts.

Finally, differential methylation between term and preterm individuals was assessed in gene bodies and promoters. RnBeads includes gene-level annotations from Ensembl ([www.ensembl.org](http://www.ensembl.org); v77 for the version of RnBeads used), and assigns promoters as the regions from 1.5 kb upstream to 0.5 kb downstream of the transcription start site. Differentially methylated positions were assessed with Limma,<sup>41</sup> and aggregated for genes and promoters using a generalisation of Fisher's method. The false discovery rate (FDR)-corrected version of these aggregated region-level  $P$ -values was used to select genes with significantly differentially methylated regions (DMR) in bodies and/or promoters. Gene function annotation was determined from the National Center for Biotechnology Information Gene database (<http://www.ncbi.nlm.nih.gov/gene/about-generif>).

### Validation by pyrosequencing

Pyrosequencing was used to validate DNAm at five selected genes that showed differential methylation ( $P < 0.05$ , FDR corrected) between preterm infants at TEA and term controls: *SLC7A5*, *SLC1A2*, *NPBWR1* and *QPRT*. *APOL1* was included in validation studies because of its functional relevance and the significance value from the array was marginal ( $P = 0.05$ ). Bisulphite conversion was performed on 500 ng of genomic DNA with the EZ DNAm kit (Zymo Research, Freiburg, Germany). The converted DNA was amplified using the AmpliTaq Gold 360 kit (Applied Biosystems, Warrington, UK) with primers mapping to target regions containing CpGs assayed within the array. PCR primers were designed using PyroMark Assay Design Software 2.0 (Qiagen; <https://www.qiagen.com>). Pyrosequencing was performed using PyroMark Q24Gold reagents on a PyroMark Q24 Pyrosequencer (Qiagen) according to the manufacturer's instructions. Data were extracted and analysed using PyroMark Q24 1.0.10 software (Qiagen). Background non-conversion levels were ~1–3%.

### Diffusion MRI analysis

After conversion from DICOM to NIFTI-1 format, the dMRI data were preprocessed using FSL tools (<http://www.fmrib.ox.ac.uk/fsl>) to extract the brain and eliminate bulk patient motion and eddy current-induced artifacts by registering the diffusion-weighted to the first T2-weighted echo planar imaging volume of each subject. Using DTIFIT,  $\langle D \rangle$  and FA volumes were generated for every subject. From the underlying white matter connectivity data, eight major white matter fasciculi (genu and splenium of corpus callosum, left and right cingulum cingulate gyrus, left and right corticospinal tracts (CST), and left and right inferior longitudinal fasciculi) were identified using PNT optimised for neonatal dMRI. As described in detail in the study by Anblagan *et al.*,<sup>33</sup> this optimisation principally involved using reference tracts created from a group of 20 term controls.

### Principal components analysis

Dimension reduction using principal components (PC) analysis was used to inspect the dataset for signal in the methylation values that is related to clinical variables and imaging features that are associated with neurodevelopmental outcome (implemented in RnBeads). The clinical variables tested were: gender, PMA at birth, PMA at scan, chorioamnionitis, exposure to antenatal steroids, exposure to antenatal magnesium sulphate, number

of days requiring parenteral nutrition and one/more episodes of late-onset sepsis. The image features tested were tract-averaged FA, tract-averaged (D), and *R* for the eight major fasciculi. Properties of the dataset, which included coordinates in the PC space, clinical variables and image features were tested for association: if both properties contained categorical data, a two-sided Fisher's exact test was used; if both properties contain numerical data the correlation coefficient between the traits was computed, and a *P*-value was estimated using permutation tests with 10,000 permutations; and if property *A* was categorical and property *B* was numeric then the *P*-value for association was calculated by comparing the values of *B* for the different categories in *A* (two-sided Wilcoxon rank sum test when *A* defines two categories, or a Kruskal–Wallis one-way analysis of variance if *A* separates the samples into three or more categories. Because 33 variables were tested for each PC, *P*-values were corrected using FDR, and values < 0.05 were considered significant.

### Methylation as a function of clinical/imaging variables

Variables indicated as interesting via PCA were modelled directly using Limma (<http://bioinf.wehi.edu.au/limma/>), with methylation as a function of the variable in question.

## RESULTS

### Participants

We collected genomic DNA from 36 sex-matched preterm infants (mean PMA at birth 28<sup>+3</sup> weeks, range 23<sup>+2</sup>–32<sup>+6</sup>; mean birth weight 1057 g, range 568–1460) at TEA (mean PMA 39<sup>+5</sup> weeks, range 38–42<sup>+4</sup> weeks), and from 36 sex-matched controls born at term (mean PMA 40<sup>+0</sup>, range 38<sup>+1</sup>–42<sup>+0</sup>). Seventy out of 72 mothers (97%) reported taking folic acid supplements around the time of conception to at least 12 weeks gestation.

Of the preterm infants, 9 (25%) had intrauterine growth restriction, 35 (97%) had been exposed to antenatal steroids, 20 (56%) had been exposed to antenatal MgSO<sub>4</sub> and 11 (31%) had histological chorioamnionitis. The mean duration requiring parenteral nutrition after birth was 11 days (range 5–25).

Of the controls, none had intrauterine growth restriction, and none was exposed to MgSO<sub>4</sub> or steroids for threatened preterm labour at any stage in pregnancy. None received parenteral nutrition.

The mean (range) DNA yield was 45.9 ng μl<sup>-1</sup> (13.4–95.9) from preterm infants and 36.35 ng μl<sup>-1</sup> (8.12–80) from term infants.

### Association between DNAm and preterm birth

About 112,818 probes were removed after: first, pre-filtering (probes on SNPs (*n*=66,877); non-specific probes (*n*=26,505); sites with excess high detection *P*-values (*n*=8,852)); and second, post-filtering (non-CpG probes (*n*=1130); and probes on sex chromosomes (*n*=9,454)). The remaining probes were used to calculate aggregate *P*-values for DMRs in two categories: gene bodies and promoters. About 87 genes were assigned as differentially methylated by this approach (*P*<0.05, FDR corrected, Supplementary Table 1), of which 25 were protein coding. About 138 promoter regions were differentially methylated (Supplementary Table 2), of which 58 related to protein-coding genes. About 34 genes were present in both sets (partly due to the overlapping gene and promoter definitions), of which 11 were protein coding. Genes that encode proteins with neural function and/or those with neuropsychiatric disease associations are listed in Table 1.

### Array validation

To validate the array findings, pyrosequencing was performed at selected annotated CpG sites in five selected genes: *SLC7A5*, *SLC1A2*, *NPBWR1*, *APOL1* and *QPRT*. CpG sites in all five genes which were identified on the array were confirmed as being differentially methylated (Figure 1, Table 2). Because some assays covered additional neighbouring CpGs, which were not also

interrogated in the array, it was possible to assess methylation patterns in the nearby region. For *SLC7A5*, the assay covered three upstream CpGs that all showed similar differences in methylation; the *SLC7A2* and *APOL1* assays both captured a second neighbouring downstream CpG that was also differentially methylated; and the assay for *NPBWR1* cg26205771 covered 1 upstream and 1 downstream CpG, and both showed similar methylation patterns.

### Diffusion MRI analysis

Figure 2 shows illustrations of segmented tracts for a representative subject, while Table 3 presents descriptive statistics for (D), FA and *R* in the eight major fasciculi identified from the dMRI data in the preterm group using PNT.

### Principal components analysis

Ninety-five percent of the variance in the preterm methylome was explained by 23 components, with most variance explained by the first two PCs (31.8% and 20.1%, respectively), (Supplementary Table 3). In exploratory analyses, gender was associated with the first PC (*P*=0.0071); FA in the genu of the corpus callosum was associated with the 5th PC (3.3% variance; *P*=0.0061); right CST *R* and chorioamnionitis were associated with the 6th PC (2.9% variance; *P*=0.0011 and *P*=0.0053, respectively); and both gender (*P*=0.0016) duration of parenteral nutrition use (*P*=0.0017) were associated with the 7th PC (2.6% variance). After correction for multiple tests, three associations remained: right CST *R* with the 6th PC (*P*=0.036); and both gender (*P*=0.028) and duration of parenteral nutrition (*P*=0.028) with the 7th PC. No variable was significantly associated with any DMR when tested directly (adjusted *P*-value < 0.05).

## DISCUSSION

In a deeply phenotyped representative sample of newborns, preterm birth was associated with significant alterations in the methylome in 10 protein-coding genes whose products influence neural cell function and are associated with behavioural traits/neuropsychiatric disease. We found that specific risk modulators of neurodevelopmental outcome after preterm birth (gender, chorioamnionitis and early nutritional factors) explained a modest, but significant proportion of the variance in DNAm. Furthermore, there was an association between DNAm and white matter tract tissue integrity and shape inferred from dMRI, suggesting that epigenetic variation may contribute to the cerebral phenotype of preterm birth.

Epigenome-wide association studies have provided new insights into genes whose regulation pattern varies in the context of child abuse, post-traumatic stress disorder, schizophrenia and autism spectrum disorder,<sup>12,64,65</sup> but to our knowledge this is the first epigenome-wide association studies in preterm infants and healthy controls to identify differential methylation at loci that influence neural development. The magnitude of the DNAm differences between preterm and term infants varied between and within genes; however, in support of a potential biological role for these changes, differential methylation was identified at multiple CpGs on the array for most loci. Pyrosequencing analysis for all genes selected for validation confirmed the differences at individual CpGs seen on the array and also identified additional differentially methylated neighbouring CpGs, suggesting that preterm birth associates with widespread effects on DNAm at these loci. For the majority of these genes, differential DNAm was identified in the gene promoter, although for *LRG1* and *SLC7A5* differential methylation also extended into the gene body. In general, DNAm at DMRs has a negative correlation with gene expression, with recent studies reporting that this correlation is stronger not only for CpGs close to the transcription start site but also for intragenic DMRs, which do not necessarily mark intragenic

**Table 1.** Differential methylation between preterm infants at term equivalent age and healthy controls in protein-coding genes with neural functions and disease associations

Gene symbol	Gene name	Function/disease association	Position	Number of differentially methylated CpGs (P < 0.05 FDR corrected)	Direction of change in preterm infants relative to term infants
SLC7A5	Solute carrier family 7 (amino acid transporter light chain, L system), member 5	L-type amino acid transporter. Determining role in the permeation of branch chain amino acids and amino acid related drugs (L-Dopa) across the blood brain barrier; <sup>42</sup> and thyroid hormone uptake in foetal cortex <sup>43</sup>	Promoter and gene	26	↓
SLC1A2	Solute carrier family 1 (glial high affinity glutamate transporter), member 2	Principal membrane-bound transporter that clears the excitatory neurotransmitter glutamate from the extracellular space at synapses in the central nervous system. Associated with schizophrenia, bipolar disorder, and neurodegeneration <sup>44,45</sup>	Promoter	4	↓
NPBWR1	Neuropeptides B/W receptor 1	Neuropeptide and opioid receptor. Associated with memory function, <sup>46</sup> eating behaviours <sup>47</sup> and processing social information <sup>48</sup>	Promoter	1	↓
APOL1	Apolipoprotein L, 1	Secreted high density lipoprotein which binds to apolipoprotein A-I. Associated with schizophrenia susceptibility <sup>49,50</sup>	Gene	1	↓
QPRT	Quinolate phosphoribosyltransferase	Catabolises quinolate, a potent neuronal excitotoxin, and may inhibit apoptosis. <sup>51</sup> Associated with Alzheimer's disease <sup>52</sup> and epilepsy <sup>53</sup>	Promoter	1	↑
LRG1	Leucine-rich alpha-2-glycoprotein 1	Involved in protein-protein interaction, signal transduction, cell adhesion and development, is expressed in astrocytes of cerebral cortex and is linked with ageing and neurodegeneration <sup>54,55</sup>	Promoter and gene	2	↓
PRPH	Peripherin	Neuronal cytoskeletal protein. Associated with susceptibility to amyotrophic lateral sclerosis <sup>56</sup> and frontotemporal lobar degeneration <sup>57</sup>	Promoter	1	↑
GRIK5	Glutamate receptor, ionotropic, kainate 5	Member of the glutamate-gated ionic channel family. Associated with schizophrenia <sup>58,59</sup>	Promoter	1	↑
TREM2	Triggering receptor expressed on myeloid cells 2	Membrane protein that forms a receptor signalling complex with tyrosine kinase-binding protein. Mutations associated with pre-senile dementia and demyelination <sup>60,61</sup>	Promoter	2	↓
MCHR1	Melanin-concentrating hormone receptor 1	Member of the G protein-coupled receptor family 1, an integral plasma membrane protein that binds melanin-concentrating hormone. The encoded protein can inhibit cAMP accumulation and stimulate intracellular calcium flux. Differential methylation is associated with obesity <sup>62</sup> and receptor antagonists control obesity and influence mood <sup>63</sup>	Promoter	1	↑

Abbreviation: FDR, false discovery rate.



**Figure 1.** Differential methylation between preterm infants at term equivalent age and healthy infants born at term at CpG sites in protein-coding genes with neural function identified in the array ( $P < 0.05$ , corrected).

**Table 2.** Pyrosequencing results for 5 genes (13 CpG sites) that showed differential methylation between groups in the array

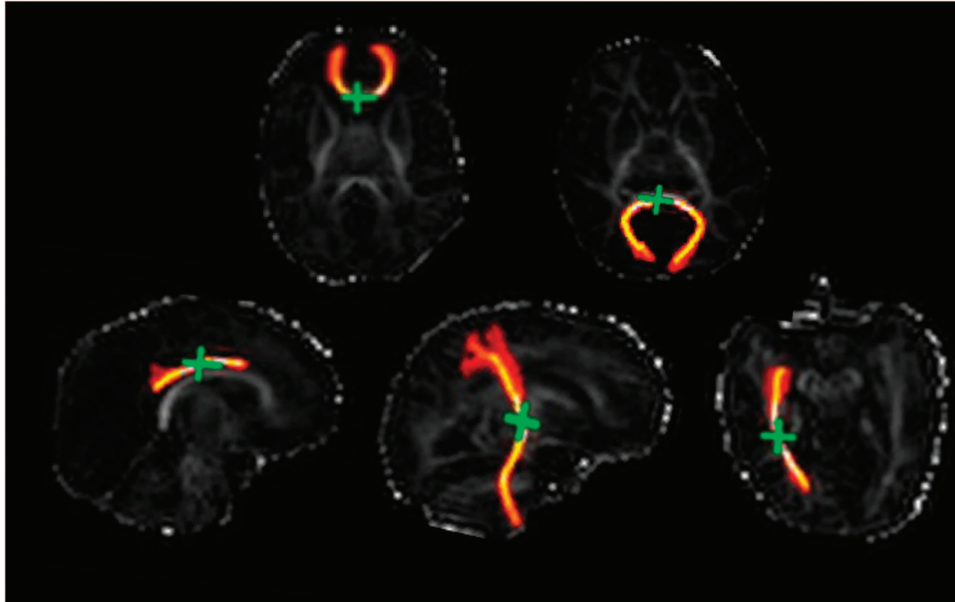
Gene symbol	CpG site	Term mean % methylation (s.d.)	Preterm mean % methylation (s.d.)	% Difference (term–preterm)	P-value
QPRT	cg06453916	11.4 (1.1)	12.5 (1.5)	–1.1	1.24E–03
SLC7A5	–45	77.2 (7.4)	68.9 (8.8)	8.3	8.30E–04
	–39	60.1 (8.2)	50.7 (8.9)	9.4	3.81E–04
	–23	80.0 (5.9)	71.0 (10.9)	9.1	7.68E–04
	cg05834639	36.6 (7.2)	28.6 (8.6)	8.1	8.74E–04
SLC1A2	cg25963980	31.2 (4.9)	25.7 (4.65)	5.5	5.94E–06
	+10	41.2 (5.0)	34.4 (6.0)	6.8	2.43E–06
	APOL1	cg36649144	29.5 (6.8)	24.3 (11.9)	5.2
	+9	14.6 (3.6)	11.5 (5.8)	3.14	0.0159
NPBWR1	cg07629017	6.0 (2.8)	4.74 (1.3)	1.26	0.0195
NPBWR1	–5	64.5 (7.5)	56.5 (10.3)	8.0	4.11E–04
	cg26205771	55.6 (7.1)	48.1 (8.1)	7.5	1.00E–04
	+4	58.4 (8.6)	50.0 (10.4)	8.4	4.57E–04

Positions of additional cytosines covered by pyrosequencing assays but which are not present on the array are given relative to the annotated CpG from the array.

CpG islands or CGI shores and may instead represent functional elements.<sup>66</sup>

A particularly large number of differentially methylated CpGs were identified in two members of the solute transporter family of membrane transport proteins, *SLC7A5* and *SLC1A2*. In preterm infants, a significant reduction (~10%) in DNAm was seen at multiple CpGs in the *SLC7A5* promoter and gene body on the array, and more were identified on pyrosequencing. *SLC7A5* also

known as *LAT1*) is a member of the solute transporter family of membrane transport proteins and is involved in the transport of large amino acids, including methionine, across the blood–brain barrier.<sup>67</sup> Methionine is a key component of S-adenosylmethionine, the major methyl donor and very recent data suggests that *SLC7A5* can act as an indirect regulator of the epigenome, at least in terms of histone modification through effects on the availability of methionine and the subsequent availability of



**Figure 2.** Illustration of segmented tracts overlaid on FA maps. Top row: genu (left) and splenium (right) of corpus callosum. Bottom row (right to left): left CCG, right CST and right ILF. CCG, cingulum cingulate gyrus; CST, corticospinal tract; FA, fractional anisotropy; ILF, inferior longitudinal fasciculus.

**Table 3.** Mean (s.d.) of tract-averaged FA and  $\langle D \rangle$ , and median (IQR/2) values of  $R$  for each major fasciculus

	<i>Genu</i>	<i>Splenium</i>	<i>Right CST</i>	<i>Left CST</i>	<i>Right CCG</i>	<i>Left CCG</i>	<i>Right ILF</i>	<i>Left ILF</i>
FA	0.20 (0.04)	0.26 (0.04)	0.27 (0.03)	0.28 (0.04)	0.20 (0.03)	0.19 (0.03)	0.22 (0.03)	0.19 (0.03)
$\langle D \rangle$ (s.d.)/ $\times 10^{-3} \text{ mm}^2 \text{ s}^{-1}$	1.497 (0.075)	1.586 (0.162)	1.192 (0.077)	1.240 (0.065)	1.375 (0.186)	1.365 (0.067)	1.664 (0.207)	1.685 (0.210)
$R$ (IQR/2)	-4.81 (2.27)	-7.73 (4.36)	-2.66 (1.41)	-3.30 (1.71)	-3.21 (2.44)	-2.78 (3.37)	-3.07 (2.67)	-0.97 (1.78)

Abbreviations: CCG, cingulum cingulate gyrus; CST, corticospinal tract; FA, fractional anisotropy; ILF, inferior longitudinal fasciculus; IQR, interquartile range.

S-adenosylmethionine.<sup>68</sup> *SLC1A2* (also known as *EAAT2*, *GLT-1*) is predominantly expressed in astrocytes but is also expressed by oligodendroglia and macrophages, and on neurons during development. It has a role for clearing glutamate throughout the neuroaxis and can be both transcriptionally and post-transcriptionally regulated.<sup>44</sup> The expression of *SLC1A2* is reported to be developmentally regulated, particularly at the window of peak vulnerability for the development of periventricular leukomalacia.<sup>69,70</sup>

The prevailing form of preterm brain injury is diffuse and involves multiple cell lines (reviewed by Back and Miller<sup>71</sup>). The pathogenesis includes death of pre-myelinating oligodendrocytes (pre-OLs) because of vulnerability to inflammatory mediators, reactive oxygen and nitrogen species, and glutamate excitotoxicity. This is followed by defective pre-OL regeneration and repair, leading to hypomyelination. In pre-clinical and human post-mortem studies the diffuse form of white matter injury coincides with enrichment of reactive glia (activated microglia/macrophages and reactive astrocytes) that inhibit the maturation of pre-OLs to myelin-forming oligodendrocytes. The neuronal population is not thought to degenerate under conditions that generate pre-OL loss (outside the context of tissue necrosis and cystic periventricular leukomalacia), but rather there is a dysmaturation response characterized by aberrant dendritic arborisation, disturbances in synaptic activity and reduced spine density. The functional profiles of the 10 genes that we found to be differentially methylated in preterm infants include neuronal and glial signalling,

neurotransmission, apoptosis and cellular energetics. Our findings focus attention on the role of these genes in mediating injury and regeneration/repair processes after preterm birth, and their candidacy is further strengthened by the neuropsychiatric disease associations in later life (Table 1).

PCs analysis was used to explore whether dMRI measures in the major white matter fasciculi or clinical risk factors contributed to structure in the methylation data of the preterm group. After correction for multiple tests and exclusion of probes on sex chromosomes, a small proportion of the variance was explained by the shape of CST, which was associated with the 6th principal component (3.3% of variance); and of the clinical factors tested, gender and number of days requiring parenteral nutrition, both associated with the 7th principal component (2.9% of variance). However, none of these three variables was significantly associated with any DMR when tested directly, which indicates that if an effect is present, it is subtle and distributed over many loci.

We sampled the methylome at a single time point chosen to reflect the allostatic load of preterm birth and neonatal intensive care among children who survive to hospital discharge, but this leaves uncertainty about the temporal cues for epigenetic modification in the perinatal period. A recent study of DNAm in umbilical cord blood of 11 preterm infants and 11 term controls demonstrated 20 DMRs between the groups, including loci in 3 genes that are involved with neuronal development: *PPT2*, *GABBR1*, *PLEKHB1*.<sup>72</sup> We did not identify DMRs in these genes,

which may be explained by differences in study population, timing of sampling or tissue-type sampled. A multiple sampling design that includes parental samples, placental tissue, cord blood and extends across the life-course would be required to investigate the relative contributions of *in utero* and postnatal exposures to changes in DNAm, and the extent to which preterm birth leaves a legacy on the methylome.<sup>73</sup>

In conclusion, these novel data show that the profound early-life stress of preterm birth is associated with differential methylation at sites in several protein-coding genes. The analysis of differential methylation has identified promising candidate genes for understanding genetic influences on brain development after preterm birth.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

The study was funded by Theirworld (<http://www.theirworld.org>), and was sponsored by the University of Edinburgh. The study was supported by NHS Research Scotland and NHS Lothian Research and Development. We are grateful to the families who consented to take part in the study and to the nursing and radiography staff at the Clinical Research Imaging Centre, the University of Edinburgh (<http://www.cric.ed.ac.uk>) who scanned the infants. We thank Thorsten Feiweier at Siemens Healthcare for collaborating with dMRI acquisitions (Works-in-Progress Package for Advanced EPI Diffusion Imaging). AJD receives funding from the Scottish Funding Council (SCD/09).

## REFERENCES

- Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R *et al*. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012; **379**: 2162–2172.
- Moore T, Hennessy EM, Myles J, Johnson SJ, Draper ES, Costeloe KL *et al*. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ* 2012; **345**: e7961.
- Mackay DF, Smith GC, Dobbie R, Pell JP. Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. *PLoS Med* 2010; **7**: e1000289.
- Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. *J Am Acad Child Adolesc Psychiatry* 2010; **49**: 453–463.
- Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Autism spectrum disorders in extremely preterm children. *J Pediatr* 2010; **156**: 525–531.
- Nosarti C, Reichenberg A, Murray RM, Cnattingius S, Lambe MP, Yin L *et al*. Pre-term birth and psychiatric disorders in young adult life. *Arch Gen Psychiatry* 2012; **69**: E1–E8.
- Volpe JJ. Systemic inflammation, oligodendroglial maturation, and the encephalopathy of prematurity. *Ann Neurol* 2011; **70**: 525–529.
- Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. *Science* 2010; **330**: 622–627.
- Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CC, O'Donovan MC *et al*. Methylomic trajectories across human fetal brain development. *Genome Res* 2015; **25**: 338–352.
- Ozanne SE, Constanca M. Mechanisms of disease: the developmental origins of disease and the role of the epigenotype. *Nat Clin Pract Endocrinol Metabol* 2007; **3**: 539–546.
- Labonte B, Yerko V, Gross J, Mechawar N, Meaney MJ, Szyf M *et al*. Differential glucocorticoid receptor exon 1(B), 1(C), and 1(H) expression and methylation in suicide completers with a history of childhood abuse. *Biol Psychiatry* 2012; **72**: 41–48.
- Mehta D, Klengel T, Conneely KN, Smith AK, Altmann A, Pace TW *et al*. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proc Natl Acad Sci USA* 2013; **110**: 8302–8307.
- Perroud N, Dayer A, Piguat C, Nallet A, Favre S, Malafosse A *et al*. Childhood maltreatment and methylation of the glucocorticoid receptor gene NR3C1 in bipolar disorder. *Br J Psychiatry* 2014; **204**: 30–35.
- Khulan B, Manning JR, Dunbar DR, Seckl JR, Raikonen K, Eriksson JG *et al*. Epigenomic profiling of men exposed to early-life stress reveals DNA methylation differences in association with current mental state. *Transl Psychiatry* 2014; **4**: e448.
- Provencal N, Suderman MJ, Guillemin C, Massart R, Ruggiero A, Wang D *et al*. The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells. *J Neurosci* 2012; **32**: 15626–15642.
- Sommershof A, Aichinger H, Engler H, Adenauer H, Catani C, Boneberg EM *et al*. Substantial reduction of naive and regulatory T cells following traumatic stress. *Brain Behav Immun* 2009; **23**: 1117–1124.
- Smith AK, Kilaru V, Klengel T, Mercer KB, Bradley B, Conneely KN *et al*. DNA extracted from saliva for methylation studies of psychiatric traits: evidence tissue specificity and relatedness to brain. *Am J Med Genet B* 2015; **168B**: 36–44.
- Lowe R, Gemma C, Beyan H, Hawa MI, Bazeos A, Leslie RD *et al*. Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics* 2013; **8**: 445–454.
- Boardman JP, Counsell SJ, Rueckert D, Kapellou O, Bhatia KK, Aljabar P *et al*. Abnormal deep grey matter development following preterm birth detected using deformation-based morphometry. *Neuroimage* 2006; **32**: 70–78.
- Kapellou O, Counsell SJ, Kennea N, Dyet L, Saeed N, Stark J *et al*. Abnormal cortical development after premature birth shown by altered allometric scaling of brain growth. *PLoS Med* 2006; **3**: e265.
- Woodward LJ, Anderson PJ, Austin NC, Howard K, Inder TE. Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. *N Engl J Med* 2006; **355**: 685–694.
- Counsell SJ, Edwards AD, Chew AT, Anjari M, Dyet LE, Srinivasan L *et al*. Specific relations between neurodevelopmental abilities and white matter microstructure in children born preterm. *Brain* 2008; **131**: 3201–3208.
- van Kooij BJ, de Vries LS, Ball G, van Hl, Benders MJ, Groenendaal F *et al*. Neonatal tract-based spatial statistics findings and outcome in preterm infants. *AJNR Am J Neuroradiol* 2012; **33**: 188–194.
- Ball G, Pazderova L, Chew A, Tusor N, Merchant N, Arichi T *et al*. Thalamocortical connectivity predicts cognition in children born preterm. *Cereb Cortex* 2015; **25**: 4310–4318.
- Huppi PS, Maier SE, Peled S, Zientara GP, Barnes PD, Jolesz FA *et al*. Microstructural development of human newborn cerebral white matter assessed in vivo by diffusion tensor magnetic resonance imaging. *Pediatr Res* 1998; **44**: 584–590.
- Huppi PS, Murphy B, Maier SE, Zientara GP, Inder TE, Barnes PD *et al*. Microstructural brain development after perinatal cerebral white matter injury assessed by diffusion tensor magnetic resonance imaging. *Pediatrics* 2001; **107**: 455–460.
- Miller SP, Vigneron DB, Henry RG, Bohland MA, Ceppi-Cozzio C, Hoffman C *et al*. Serial quantitative diffusion tensor MRI of the premature brain: development in newborns with and without injury. *J Magn Reson Imaging* 2002; **16**: 621–632.
- Ball G, Counsell SJ, Anjari M, Merchant N, Arichi T, Doria V *et al*. An optimised tract-based spatial statistics protocol for neonates: applications to prematurity and chronic lung disease. *Neuroimage* 2010; **53**: 94–102.
- Porter EJ, Counsell SJ, Edwards AD, Allsop J, Azzopardi D. Tract-based spatial statistics of magnetic resonance images to assess disease and treatment effects in perinatal asphyxial encephalopathy. *Pediatr Res* 2010; **68**: 205–209.
- Boardman JP, Walley A, Ball G, Takousis P, Krishnan ML, Hughes-Carre L *et al*. Common genetic variants and risk of brain injury after preterm birth. *Pediatrics* 2014; **133**: e1655–e1663.
- O'Gorman RL, Bucher HU, Held U, Koller BM, Huppi PS, Hagmann CF. Tract-based spatial statistics to assess the neuroprotective effect of early erythropoietin on white matter development in preterm infants. *Brain* 2015; **138**: 388–397.
- Clayden JD, Maniega SM, King MD, Bastin ME, Clark CA. TractoR: magnetic resonance imaging and tractography with *R*. *J Stat Softw* 2011; **44**: 1–18.
- Anblagan D, Bastin ME, Sparrow S, Piyasena C, Pataky R, Moore EJ *et al*. Tract shape modeling detects changes associated with preterm birth and neuroprotective treatment effects. *Neuroimage* 2015; **8**: 51–58.
- Toga AW, Thompson PM. Genetics of brain structure and intelligence. *Annu Rev Neurosci* 2005; **28**: 1–23.
- Thompson PM, Cannon TD, Narr KL, van ET, Poutanen VP, Huttunen M *et al*. Genetic influences on brain structure. *Nat Neurosci* 2001; **4**: 1253–1258.
- Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods* 2014; **11**: 1138–1140.
- Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW *et al*. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013; **8**: 203–209.
- Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D *et al*. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2013; **29**: 189–196.
- Marabita F, Almgren M, Lindholm ME, Ruhmann S, Fagerstrom-Billai F, Jagodic M *et al*. An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. *Epigenetics* 2013; **8**: 333–346.
- Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118–127.

- 41 Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (eds). *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer: New York, USA, 2005, pp 397–420.
- 42 Matsuo H, Tsukada S, Nakata T, Chairoungdua A, Kim DK, Cha SH *et al*. Expression of a system L neutral amino acid transporter at the blood-brain barrier. *Neuroreport* 2000; **11**: 3507–3511.
- 43 Chan SY, Martin-Santos A, Loubiere LS, Gonzalez AM, Stieger B, Logan A *et al*. The expression of thyroid hormone transporters in the human fetal cerebral cortex during early development and in N-Tera-2 neurodifferentiation. *J Physiol* 2011; **589**: 2827–2845.
- 44 Sheldon AL, Robinson MB. The role of glutamate transporters in neurodegenerative diseases and potential opportunities for intervention. *Neurochem Int* 2007; **51**: 333–355.
- 45 Shao L, Vawter MP. Shared gene expression alterations in schizophrenia and bipolar disorder. *Biol Psychiatry* 2008; **64**: 89–97.
- 46 Ksiazek-Winiarek DJ, Kacperska MJ, Glabinski A. MicroRNAs as novel regulators of neuroinflammation. *Mediat Inflamm* 2013; **2013**: 172351.
- 47 Ishii M, Fei H, Friedman JM. Targeted disruption of GPR7, the endogenous receptor for neuropeptides B and W, leads to metabolic defects and adult-onset obesity. *Proc Natl Acad Sci USA* 2003; **100**: 10540–10545.
- 48 Watanabe N, Wada M, Irukayama-Tomobe Y, Ogata Y, Tsujino N, Suzuki M *et al*. A single nucleotide polymorphism of the neuropeptide B/W receptor-1 gene influences the evaluation of facial expressions. *PLoS One* 2012; **7**: e35390.
- 49 Tomita Y, Ikeda M, Mutoh H, Inada T, Iwata N, Ozaki N *et al*. Association study between Apolipoprotein L and schizophrenia by exhaustive and rule-based combination analysis for identification of multilocus interactions. *J Biosci Bioeng* 2007; **103**: 303–310.
- 50 Mimmack ML, Ryan M, Baba H, Navarro-Ruiz J, Iritani S, Faull RL *et al*. Gene expression analysis in schizophrenia: reproducible up-regulation of several members of the apolipoprotein L family located in a high-susceptibility locus for schizophrenia on chromosome 22. *Proc Natl Acad Sci USA* 2002; **99**: 4680–4685.
- 51 Ishidoh K, Kamemura N, Imagawa T, Oda M, Sakurai J, Katunuma N. Quinolinic phosphoribosyl transferase, a key enzyme in de novo NAD(+) synthesis, suppresses spontaneous cell death by inhibiting overproduction of active-caspase-3. *Biochim Biophys Acta* 2010; **1803**: 527–533.
- 52 Rahman A, Ting K, Cullen KM, Braidy N, Brew BJ, Guillemin GJ. The excitotoxic quinolinic acid induces tau phosphorylation in human neurons. *PLoS One* 2009; **4**: e6344.
- 53 Feldblum S, Rougier A, Loiseau H, Loiseau P, Cohadon F, Morselli PL *et al*. Quinolinic-phosphoribosyl transferase activity is decreased in epileptic human brain tissue. *Epilepsia* 1988; **29**: 523–529.
- 54 Nakajima M, Miyajima M, Ogino I, Watanabe M, Hagiwara Y, Segawa T *et al*. Brain localization of leucine-rich alpha2-glycoprotein and its role. *Acta Neurochir Suppl* 2012; **113**: 97–101.
- 55 Miyajima M, Nakajima M, Motoi Y, Moriya M, Sugano H, Ogino I *et al*. Leucine-rich alpha2-glycoprotein is a novel biomarker of neurodegenerative disease in human cerebrospinal fluid and causes neurodegeneration in mouse cerebral cortex. *PLoS One* 2013; **8**: e74453.
- 56 Corrado L, Carlomagno Y, Falasco L, Mellone S, Godi M, Cova E *et al*. A novel peripherin gene (PRPH) mutation identified in one sporadic amyotrophic lateral sclerosis patient. *Neurobiol Aging* 2011; **552**: e1–e6.
- 57 Martins-de-Souza D, Guest PC, Mann DM, Roeber S, Rahmoune H, Bauder C *et al*. Proteomic analysis identifies dysfunction in cellular transport, energy, and protein metabolism in different brain regions of atypical frontotemporal lobar degeneration. *J Proteome Res* 2012; **11**: 2533–2543.
- 58 Shibata H, Aramaki T, Sakai M, Ninomiya H, Tashiro N, Iwata N *et al*. Association study of polymorphisms in the GluR7, KA1 and KA2 kainate receptor genes (GRIK3, GRIK4, GRIK5) with schizophrenia. *Psychiatry Res* 2006; **141**: 39–51.
- 59 Porter RH, Eastwood SL, Harrison PJ. Distribution of kainate receptor subunit mRNAs in human hippocampus, neocortex and cerebellum, and bilateral reduction of hippocampal GluR6 and KA2 transcripts in schizophrenia. *Brain Res* 1997; **751**: 217–231.
- 60 Kiialainen A, Veckman V, Saharinen J, Paloneva J, Gentile M, Hakola P *et al*. Transcript profiles of dendritic cells of PLOSL patients link demyelinating CNS disorders with abnormalities in pathways of actin bundling and immune response. *J Mol Med* 2007; **85**: 971–983.
- 61 Klunemann HH, Ridha BH, Magy L, Wherrett JR, Hemelsoet DM, Keen RW *et al*. The genetic causes of basal ganglia calcification, dementia, and bone cysts: DAP12 and TREM2. *Neurology* 2005; **64**: 1502–1507.
- 62 Stepanow S, Reichwald K, Huse K, Gausmann U, Nebel A, Rosenstiel P *et al*. Allele-specific, age-dependent and BMI-associated DNA methylation of human MCHR1. *PLoS One* 2011; **6**: e17711.
- 63 Borowsky B, Durkin MM, Ogozalek K, Marzabadi MR, DeLeon J, Lagu B *et al*. Antidepressant, anxiolytic and anorectic effects of a melanin-concentrating hormone-1 receptor antagonist. *Nat Med* 2002; **8**: 825–830.
- 64 Ladd-Acosta C, Hansen KD, Briem E, Fallin MD, Kaufmann WE, Feinberg AP. Common DNA methylation alterations in multiple brain regions in autism. *Mol Psychiatry* 2014; **19**: 862–871.
- 65 Aberg KA, McClay JL, Nerella S, Clark S, Kumar G, Chen W *et al*. Methylome-wide association study of schizophrenia: identifying blood biomarker signatures of environmental insults. *JAMA Psychiatry* 2014; **71**: 255–264.
- 66 Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D *et al*. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* 2015; **523**: 212–216.
- 67 O’Kane RL, Hawkins RA. Na<sup>+</sup>-dependent transport of large neutral amino acids occurs at the abluminal membrane of the blood-brain barrier. *Am J Physiol Endocrinol Metab* 2003; **285**: E1167–E1173.
- 68 Dann SG, Ryskin M, Barsotti AM, Golas J, Shi C, Miranda M *et al*. Reciprocal regulation of amino acid import and epigenetic state through Lat1 and EZH2. *EMBO J* 2015; **34**: 1773–1785.
- 69 Desilva TM, Billiards SS, Borenstein NS, Trachtenberg FL, Volpe JJ, Kinney HC *et al*. Glutamate transporter EAAT2 expression is up-regulated in reactive astrocytes in human periventricular leukomalacia. *J Comp Neurol* 2008; **508**: 238–248.
- 70 Desilva TM, Kinney HC, Borenstein NS, Trachtenberg FL, Irwin N, Volpe JJ *et al*. The glutamate transporter EAAT2 is transiently expressed in developing human cerebral white matter. *J Comp Neurol* 2007; **501**: 879–890.
- 71 Back SA, Miller SP. Brain injury in premature neonates: a primary cerebral dysmaturation disorder? *Ann Neurol* 2014; **75**: 469–486.
- 72 Fernando F, Keijser R, Henneman P, van der Kevie-Kersemaekers AM, Mannens MM, van der Post JA *et al*. The idiopathic preterm delivery methylation profile in umbilical cord blood DNA. *BMC Genomics* 2015; **16**: 736.
- 73 Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, Sheehan P *et al*. Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. *Genome Med* 2013; **5**: 96.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)