Methylomic trajectories across human fetal brain development

Helen Spiers,1 Ellis Hannon,2 Leonard C. Schalkwyk,3 Rebecca Smith,1 Chloe C.Y. Wong,1 Michael C. O’Donovan,4 Nicholas J. Bray,1 and Jonathan Mill1,2

1Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London SE5 8AF, United Kingdom; 2University of Exeter Medical School, University of Exeter, Exeter EX2 5DW, United Kingdom; 3School of Biological Sciences, University of Essex, Colchester CO4 3SQ, United Kingdom; 4MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff CF24 4HQ, United Kingdom

Epigenetic processes play a key role in orchestrating transcriptional regulation during development. The importance of DNA methylation in fetal brain development is highlighted by the dynamic expression of de novo DNA methyltransferases during the perinatal period and neurodevelopmental deficits associated with mutations in the methyl-CpG binding protein 2 (MECP2) gene. However, our knowledge about the temporal changes to the epigenome during fetal brain development has, to date, been limited. We quantified genome-wide patterns of DNA methylation at ~400,000 sites in 179 human fetal brain samples (100 male, 79 female) spanning 23 to 184 d post-conception. We identified highly significant changes in DNA methylation across fetal brain development at >7% of sites, with an enrichment of loci becoming hypomethylated with fetal age. Sites associated with developmental changes in DNA methylation during fetal brain development were significantly underrepresented in promoter regulatory regions but significantly overrepresented in regions flanking CpG islands (shores and shelves) and gene bodies. Highly significant differences in DNA methylation were observed between males and females at a number of autosomal sites, with a small number of regions showing sex-specific DNA methylation trajectories across brain development. Weighted gene coexpression network analysis (WGCNA) revealed discrete modules of comethylated loci associated with fetal age that are significantly enriched for genes involved in neurodevelopmental processes. This is, to our knowledge, the most extensive study of DNA methylation across human fetal brain development to date, confirming the prenatal period as a time of considerable epigenomic plasticity.

[Supplemental material is available for this article.]

Human brain development is an intricate process involving the dynamic orchestration of gene expression. Prenatal transcriptional changes in the brain occur more rapidly than at any other stage of life (Johnson et al. 2009; Colantuoni et al. 2011; Kang et al. 2011). The precise temporal regulation of transcriptional processes is necessary for the correct development of structural and functional complexity in the brain. Although cell-specific and temporally appropriate gene expression is primarily controlled through the direct action of transcription factors, there is a growing recognition of the role of epigenetic mechanisms in the dynamic regulation of gene function during cellular development and differentiation (Henikoff and Matzke 1997; Jaenisch and Bird 2003; Hirabayashi and Gotoh 2010).

DNA methylation is the most extensively studied epigenetic modification. It is known to play a role in many important genomic regulatory processes, including X-Chromosome inactivation, genomic imprinting, and the repression of tumor suppressor genes in cancer. DNA methylation refers to the addition of a single methyl group to DNA that is adjacent to cytosine residues. This process is catalyzed by DNA methyltransferases.

Chromatin remodeling proteins such as MECP2 remodel chromatin into a condensed heterochromatic state. Recent work has revealed a more nuanced relationship between DNA methylation and transcription that is dependent on genomic and cellular context (Jones 2012; although DNA methylation at promoter regulatory regions is typically associated with repressed expression, DNA methylation in the gene body is often positively correlated with expression (Ball et al. 2009; Maunakea et al. 2010) and is thought to play a role in alternative splicing (Maunakea et al. 2013). The establishment and maintenance of cell-specific DNA methylation patterns is crucial for normal mammalian development (Reik 2007; Geiman and Muegge 2010; Smith and Meissner 2013; Ziller et al. 2013). Recent evidence strongly implicates a role for dynamic epigenetic processes in the regulation of transcriptional plasticity in the developing brain (Numata et al. 2012; Lister et al. 2013). A critical role for DNA methylation in neurodevelopment is supported by the dynamic expression of the de novo DNA methyltransferases DNMT3A and DNMT3B during prenatal brain development (Feng et al. 2005), and by the occurrence of neurodevelopmental deficits in humans as a consequence of mutations in the methyl-CpG binding protein 2 (MECP2) gene, which interacts with methylated DNA to control neuronal gene expression (Guy et al. 2011; Jakovec and Akbarian 2012). Furthermore, the dynamic regulation of DNA methylation is known to influence key neurobiological and cognitive functions in the brain across the life course, including neuronal plasticity (Borrelli et al. 2008; Ma et al. 2010; Guo et al. 2011), memory formation and maintenance (Day
and Sweatt 2010; Zovkic et al. 2013), and circadian processes (Azzi et al. 2014). Knowledge about the specific temporal methylomic changes occurring during human fetal brain development has, however, been limited due to a lack of tissue samples, with previous studies focusing on a small number of samples obtained from a narrow range of fetal ages (Numata et al. 2012; Lister et al. 2013).

Here, we describe an analysis of neurodevelopmental trajectories in DNA methylation in human fetal brain samples, identifying changes to the epigenome during development across ~400,000 sites.

Results

Methodological overview

We assessed genome-wide patterns of DNA methylation in 179 human fetal brain samples (100 male, 79 female) spanning 23 to 184 d post-conception (DPC) (Fig. 1A). Fetal brain tissue was acquired frozen from the Human Developmental Biology Resource (HDBR) (http://www.hdbr.org) and MRC Brain Banks network (http://www.mrc.ac.uk/research/facilities/brain-banks/access-for-research) under strict ethical regulations and used to isolate genomic DNA.

DNA methylation was quantified using the Illumina Infinium HumanMethylation450 BeadChip, with preprocessing, normalization, and stringent quality control undertaken as previously described (Pidsley et al. 2013). Our analyses focused on identifying DNA methylation changes associated with brain development, and whether these were enriched in certain genomic regions and features or differed between males and females. We subsequently employed systems-level network-based approaches to identify modules of highly comethylated loci and their relationship to fetal brain development. See the Methods section for an in depth description of the samples and analytical approaches used in this study.

Human fetal brain development is characterized by widespread changes in DNA methylation

A combined analysis of all 408,608 probes on the Illumina 450K array passing stringent quality control (QC) metrics showed that global levels of DNA methylation do not change significantly over the course of human fetal brain development ($r = -0.02, P = 0.76$) (Supplemental Fig. 1). In contrast, DNA methylation at individual autosomal sites was highly variable across neurodevelopment.
with levels at 28,718 sites (7.19% of the 399,364 autosomal probes assessed) differing according to fetal age at a significance level surpassing Bonferroni correction for all tested probes ($P < 1.25 \times 10^{-7}$) (Fig. 1B). We refer to these as developmentally differentially methylated positions (dDMPs). The 20 top autosomal dDMPs (ranked by $P$-value for age-association), characterized by a positive correlation between DNA methylation and fetal age (henceforth referred to as hypermethylated dDMPs), or a negative correlation (henceforth referred to as hypomethylated dDMPs), are listed in Table 1. A complete list of all 28,718 significant dDMPs is available from our laboratory website (http://epigenetics.iop.kcl.ac.uk/fetalbrain/dDMPs.xls) and in Supplemental File 1. Although the distribution of dDMPs is relatively consistent across most autosomal chromosomes (Supplemental Table 1), some chromosomes show a notable enrichment or depletion of significant sites; for example, 8.87% of CpG probes on Chromosome 13 were identified as Bonferroni-significant dDMPs (relative enrichment = 1.26, $P = 3.68 \times 10^{-10}$), whereas only 0.49% of CpG probes on Chromosome 19 reach this criterion (relative enrichment = 0.55, $P = 7.01 \times 10^{-7}$). The four top-ranked hyper- and hypomethylated fetal brain dDMPs are shown in Figure 1, C and D, respectively. Overall, there is a highly significant enrichment of hypomethylated autosomal dDMPs compared with hypermethylated dDMPs (Table 2) (hypermethylated sites: $n = 16,190$ [56.4%]; hypomethylated sites: $n = 12,528$ [43.6%]; $P = 6.74 \times 10^{-53}$), which is consistent with previous reports (Numata et al. 2012). We find a significant correlation ($r = 0.57, P = 2.82 \times 10^{-9}$) between developmental DNA methylation changes at the 100 top-ranked fetal cortex dDMPs identified in a previous, smaller-scale study (Numata et al. 2012) and changes observed at the same loci in our study (Supplemental Fig. 2). Although we have not directly generated transcriptomic data on the samples profiled in this study, fetal brain gene expression data for loci annotated to the top 20 hypermethylated and top 20 hypomethylated dDMPs (listed in Table 1) were extracted from the Brain Cloud resource (http://braincloud.jhmi.edu) (Colantuoni et al. 2011). This resource contains cortex transcriptomic data from 38

### Table 1. Top-ranked autosomal fetal brain dDMPs becoming hyper- and hypomethylated across development

<table>
<thead>
<tr>
<th>Probe</th>
<th>Chr</th>
<th>Position</th>
<th>Gene</th>
<th>Genic probe location</th>
<th>Relation to CpG island</th>
<th>Regression coefficient</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg02534854</td>
<td>11</td>
<td>147653072</td>
<td>C1QTNF3</td>
<td>3' UTR</td>
<td>Body</td>
<td>0.35</td>
<td>9.40 \times 10^{-22}</td>
</tr>
<tr>
<td>cg02527489</td>
<td>11</td>
<td>161033379</td>
<td>C1QTNF3</td>
<td>3' UTR</td>
<td>Body</td>
<td>0.17</td>
<td>2.07 \times 10^{-10}</td>
</tr>
<tr>
<td>cg19098189</td>
<td>12</td>
<td>1759721</td>
<td>CHSY1</td>
<td>TSS1500</td>
<td>Island</td>
<td>0.37</td>
<td>1.40 \times 10^{-22}</td>
</tr>
<tr>
<td>cg02534854</td>
<td>11</td>
<td>147653072</td>
<td>C1QTNF3</td>
<td>3' UTR</td>
<td>Body</td>
<td>0.35</td>
<td>9.40 \times 10^{-22}</td>
</tr>
<tr>
<td>cg02527489</td>
<td>11</td>
<td>161033379</td>
<td>C1QTNF3</td>
<td>3' UTR</td>
<td>Body</td>
<td>0.17</td>
<td>2.07 \times 10^{-10}</td>
</tr>
</tbody>
</table>

Chromosomal coordinates correspond to human genome build Feb. 2009 (GRCh37/hg19). Regression coefficient is the DNA methylation change (%) per day development. A full list of all 28,718 Bonferroni-significant dDMPs can be downloaded from http://epigenetics.iop.kcl.ac.uk/fetalbrain/dDMPs.xls and is also given in Supplemental File 1.
Table 2. Overview of the number of autosomal dDMPs significantly associated with human brain development split by genomic feature and direction of effect

<table>
<thead>
<tr>
<th>Total probes</th>
<th>Significant probes (%</th>
<th>Enrichment (95% CI)</th>
<th>P-value</th>
<th>Hypermethylated (%)</th>
<th>Hypomethylated (%)</th>
<th>Enrichment hypermethylated (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All probes</td>
<td>399,364</td>
<td>28,718 (7.19)</td>
<td>–</td>
<td>12,528 (43.6)</td>
<td>16,190 (56.4)</td>
<td>–</td>
<td>6.74 × 10⁻⁵³</td>
</tr>
<tr>
<td>CpG island feature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Island</td>
<td>132,511</td>
<td>2632 (1.99)</td>
<td>0.26 (0.25–0.27)</td>
<td>&lt;3.92 × 10⁻²⁹⁴</td>
<td>1582 (60.1)</td>
<td>1050 (39.9)</td>
<td>1.51 (1.35–1.68)</td>
</tr>
<tr>
<td>Shore</td>
<td>94,559</td>
<td>7806 (8.26)</td>
<td>1.16 (1.13–1.19)</td>
<td>1.23 × 10⁻²⁸</td>
<td>3824 (49.0)</td>
<td>3982 (51.0)</td>
<td>0.96 (0.90–1.02)</td>
</tr>
<tr>
<td>Shelf</td>
<td>36,072</td>
<td>3026 (8.39)</td>
<td>1.18 (1.14–1.23)</td>
<td>2.04 × 10⁻¹⁶</td>
<td>1043 (34.5)</td>
<td>1983 (65.5)</td>
<td>0.53 (0.47–0.58)</td>
</tr>
<tr>
<td>Not island/shore/shelf</td>
<td>136,222</td>
<td>15,254 (11.20)</td>
<td>1.63 (1.59–1.66)</td>
<td>&lt;3.92 × 10⁻²⁹⁴</td>
<td>6079 (39.9)</td>
<td>9175 (60.1)</td>
<td>0.66 (0.63–0.69)</td>
</tr>
<tr>
<td>Gene Feature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSS1500</td>
<td>71,560</td>
<td>3699 (5.17)</td>
<td>0.70 (0.68–0.73)</td>
<td>3.34 × 10⁻⁹²</td>
<td>1691 (45.72)</td>
<td>2008 (54.28)</td>
<td>0.84 (0.77–0.92)</td>
</tr>
<tr>
<td>TSS200</td>
<td>55,262</td>
<td>1419 (2.57)</td>
<td>0.34 (0.32–0.36)</td>
<td>&lt;3.92 × 10⁻²⁹⁴</td>
<td>554 (39.04)</td>
<td>865 (60.96)</td>
<td>0.64 (0.55–0.75)</td>
</tr>
<tr>
<td>5' UTR</td>
<td>56,383</td>
<td>2915 (5.17)</td>
<td>0.70 (0.68–0.73)</td>
<td>3.71 × 10⁻⁷⁵</td>
<td>1140 (39.11)</td>
<td>1775 (60.89)</td>
<td>0.64 (0.58–0.71)</td>
</tr>
<tr>
<td>1st exon</td>
<td>34,155</td>
<td>787 (2.30)</td>
<td>0.30 (0.28–0.33)</td>
<td>&lt;3.92 × 10⁻²⁹⁴</td>
<td>291 (36.98)</td>
<td>496 (63.02)</td>
<td>0.59 (0.48–0.72)</td>
</tr>
<tr>
<td>Gene body</td>
<td>146,435</td>
<td>12,432 (8.49)</td>
<td>1.20 (1.17–1.22)</td>
<td>4.71 × 10⁻⁵⁷</td>
<td>5181 (41.67)</td>
<td>7251 (58.33)</td>
<td>0.71 (0.66–0.75)</td>
</tr>
<tr>
<td>3' UTR</td>
<td>16,304</td>
<td>1225 (7.51)</td>
<td>1.05 (0.99–1.11)</td>
<td>0.1186</td>
<td>428 (34.94)</td>
<td>797 (65.06)</td>
<td>0.54 (0.46–0.63)</td>
</tr>
<tr>
<td>DNase I hypersensitivity sites</td>
<td>189,815</td>
<td>14,286 (7.53)</td>
<td>1.05 (1.03–1.07)</td>
<td>3.95 × 10⁻⁶</td>
<td>6109 (42.76)</td>
<td>8177 (57.24)</td>
<td>0.75 (0.71–0.78)</td>
</tr>
<tr>
<td>Transcription factor binding sites</td>
<td>194,306</td>
<td>11,173 (5.75)</td>
<td>0.79 (0.77–0.81)</td>
<td>1.65 × 10⁻⁹⁸</td>
<td>5044 (45.14)</td>
<td>6129 (54.86)</td>
<td>0.82 (0.78–0.87)</td>
</tr>
<tr>
<td>Alternative transcription events</td>
<td>148,061</td>
<td>9806 (6.62)</td>
<td>0.92 (0.89–0.94)</td>
<td>2.22 × 10⁻¹³</td>
<td>4127 (42.09)</td>
<td>5679 (57.91)</td>
<td>0.73 (0.69–0.77)</td>
</tr>
</tbody>
</table>

Bonferroni correction (P < 1.25 × 10⁻⁷) was used as a threshold to determine significance. For further details regarding specific transcription factor binding sites and alternative transcription events, see Supplemental Tables 10 and 11 (Wang et al. 2008; Koscielny et al. 2009; The ENCODE Project Consortium 2012; Slikker et al. 2013; Flicek et al. 2014; Ensembl, http://www.ensembl.org).
human fetal samples spanning 14 to 20 gestational weeks (~84–126 DPC). Despite the much narrower range of ages in this cohort, nine (29.03%) of the 31 transcripts for which data were available displayed a nominally significant \((P < 0.05)\) association between transcript abundance and brain development (Supplemental Table 2), with several transcripts showing highly significant associations with developmental age (Supplemental Fig. 3). Many dDMRs do not represent isolated changes at specific sites but rather occur in clusters (Fig. 1E; Supplemental Fig. 4), representing developmentally differentially methylated regions (dDMRs). We used \(\text{comb-p} \) (Pedersen et al. 2012) to identify spatially correlated regions of differential DNA methylation significantly associated with fetal brain development (Sidak-corrected \(P < 0.05\), number of probes \(\geq 3\)). In total, 4825 dDMRs were identified with a mean size of 391 base pairs (bp) \((SD = 250 \text{bp})\) and spanning an average of 4.77 probes (Supplemental Table 3). The top-ranked dDMRs (ranked by Sidak-corrected \(P\)-value) are shown in Supplemental Table 4 and Supplemental Figure 4, and a complete list of all significant dDMRs is available for download from our laboratory website (http://epigenetics.iop.kcl.ac.uk/fetalbrain/4825_age_DMRs.csv) and in Supplemental File 2. Of note, several dDMRs are located in the vicinity of loci known to play a key role in neurodevelopment, including the Wnt antagonist \(\text{SFRP1}\) that regulates cell differentiation and proliferation in the developing brain (Augustine et al. 2001), the nuclear receptor gene \(\text{NR4A2}\), which is important for the differentiation and maintenance of dopaminergic neurons during neurogenesis (Jankovic et al. 2005; Vuillermot et al. 2012), and \(\text{SHANK2}\), encoding a synaptic protein that functions as a molecular scaffold in the developing brain which has been implicated in neurodevelopmental disorders, including autism (Schmeisser et al. 2012; Guilmatre et al. 2014).

There is a significant reduction in non-CpG DNA methylation during human fetal brain development

Of the 408,608 probes passing QC, a small subset \((n = 743)\) interrogated DNA methylation at non-CpG sites. The level of DNA methylation at these sites is generally low (mean DNA methylation across all non-CpG sites \(= 14.93\%\), \(SD = 6.83\%\)), with the majority of probes being relatively unmethylated throughout brain development, confirming the results of a previous study reporting negligible levels of non-CpG methylation in fetal cortex (Lister et al. 2013). There is, however, a small but significant overall decrease in average non-CpG DNA methylation across the period of fetal brain development assessed in this study \((r = –0.24, P\text{-value} = 1.14 \times 10^{-3})\) (Supplemental Fig. 5), with 698 (93.9\%) of the interrogated non-CpG probes becoming less methylated with developmental age, supporting previous data showing that non-CpG methylation is reduced following differentiation of embryonic stem cells (Lister et al. 2009). In total, 14 non-CpG sites \((1.88\%)\) demonstrated Bonferroni-significant \((P < 6.73 \times 10^{-5})\) changes in DNA methylation across brain development (Supplemental Fig. 6; Supplemental Table 5).

The distribution of fetal brain dDMRs differs depending on CG density

Although fetal brain dDMRs are seen throughout the genome, they are not equally distributed with respect to annotated genic features (Table 2). There is a significant underrepresentation of dDMRs in promoter regulatory regions, including CpG islands (CGIs) (percentage of significant probes \(= 1.99\%\), relative enrichment \(= 2.30, P < 3.92 \times 10^{-294}\)), genic first exons (percentage of significant probes \(= 2.57\%\), relative enrichment \(= 3.34, P < 3.92 \times 10^{-294}\)), and 5’ untranslated regions (5’ UTRs) (percentage of significant probes \(= 5.17\%\), relative enrichment \(= 0.70, P = 3.71 \times 10^{-75}\)), and in the vicinity of annotated transcription start sites (TSSs) (TSS-200; percentage of significant probes \(= 2.57\%\), relative enrichment \(= 0.34, P < 3.92 \times 10^{-294}\); TSS-1500; percentage of significant probes \(= 5.17\%\), relative enrichment \(= 0.70, P = 3.34 \times 10^{-292}\)) (Fig. 2A). Conversely, fetal brain dDMRs are significantly enriched in CGI shores (percentage of significant probes \(= 8.26\%\), relative enrichment \(= 1.16, P = 1.23 \times 10^{-28}\)), CGI shelves (percentage of significant probes \(= 8.39\%\), relative enrichment \(= 1.18, P = 2.04 \times 10^{-16}\)), and gene bodies (percentage of significant probes \(= 8.49\%\), relative enrichment \(= 1.20, P = 4.71 \times 10^{-77}\)). Interestingly, dDMRs within

\[
\begin{align*}
A & \quad \text{All probes} & \text{CpG island feature} & \text{Gene feature} \\
\text{Significant probes (\%)} & \quad 50 & 0 & 100 \\
\text{Genomic Feature} & \quad \text{TSS-200} & \text{TSS-1500} & \text{TSS-3000} \\
\text{B} & \quad \% \text{DMR hypermethylated} & \% \text{DMR hypomethylated} \\
\text{Percentage significant probes (\%)} & \quad 100 & 0 \\
\text{Genomic Feature} & \quad \text{All probes} & \text{CpG island feature} & \text{Gene feature} \\
\end{align*}
\]

Figure 2. The distribution and direction of fetal brain dDMRs differs markedly across genomic features. (A) Compared to the genome average, dDMRs are significantly underrepresented in CpG islands, transcription start-sites, 5’ UTRs, and first exons, but significantly enriched in CpG island shores, CpG island shelves, and the gene body (see also Table 2). (B) There is an overall enrichment of hypomethylated fetal brain dDMRs \((P = 6.74 \times 10^{-12})\), but CpG islands are unique in showing an enrichment for sites becoming hypermethylated with fetal brain development \((P = 2.00 \times 10^{-12})\) (see also Table 2).
CGIs are more likely to become hypermethylated with fetal age (hypermethylated \( n = 1582 \) [60.1%] vs. hypomethylated \( n = 1050 \) [39.9%]), relative enrichment \( = 1.61, P = 2.00 \times 10^{-31} \) in contrast to dDMRs annotated to other genic features, which tend to become hypomethylated (see Table 2; Fig. 2B). Stratifying genomic regions by CpG density confirms a highly significant enrichment of dDMRs in all non-CGI genic features, with the highest percentage of significant probes found in non-CGI intergenic sites (percentage of significant probes = 13.16%, relative enrichment = 1.96, \( P = 3.92 \times 10^{-294} \)). In contrast, dDMRs were underrepresented in CGI-rich regions, especially in CGI proximal promoter sites (percentage of significant probes = 0.60%, relative enrichment = 0.08, \( P = 3.92 \times 10^{-294} \)) (Supplemental Table 6; Supplemental Fig. 7).

**Fetal brain dDMRs are enriched in certain functional elements and depleted in others**

A modest enrichment of dDMR compared to the genome average was seen in DNase I hypersensitivity sites (percentage of significant probes = 7.53%, relative enrichment = 1.05, \( P = 3.95 \times 10^{-37} \)) (Table 2; The ENCODE Project Consortium 2012; Slikker et al. 2013) marking regulatory DNA regions (Maurano et al. 2012). In contrast, probes associated with transcription factor binding sites (TFBSs) (The ENCODE Project Consortium 2012; Slikker et al. 2013) were under-enriched for dDMRs compared to the genome average (percentage of significant probes = 5.75%, relative enrichment = 0.79, \( P = 1.65 \times 10^{-50} \)). In total, 117 of the 148 (79.05%) classes of TFBS motifs assessed were significantly (\( P < 0.05 \)) under-enriched for dDMRs, with 12 (8.11%) being significantly enriched for dDMRs (see Supplemental Table 7 for a list of under- and over-enriched TFBSs). Taken together, regions characterized by alternative transcription events (Wang et al. 2008; Koscielny et al. 2009; Flicek et al. 2014; Ensembl, http://www.ensembl.org) were under-enriched for dDMRs (percentage of significant probes = 6.62%, relative enrichment = 0.92, \( P = 2.22 \times 10^{-13} \)) (Table 2), although there was considerable heterogeneity between different domains (Supplemental Table 8; Supplemental Fig. 8). There is a significant over-enrichment of dDMRs in alternative last exons (percentage of significant probes = 9.82%, relative enrichment = 1.41, \( P = 1.05 \times 10^{-21} \)), cassette exons (percentage of significant probes = 8.24%, relative enrichment = 1.16, \( P = 8.09 \times 10^{-21} \)), and mutually exclusive exons (percentage of significant probes = 8.11%, relative enrichment = 1.14, \( P = 4.62 \times 10^{-5} \)), with all other events displaying a significant under-enrichment of dDMRs.

There are sex differences and distinct sex-specific developmental trajectories in the human fetal brain methylome, including at autosomal locations

A total of 8059 (1.99%) of the 408,594 autosomal and X-linked probes assessed were identified as significantly (Bonferroni-corrected threshold: \( P < 1.22 \times 10^{-7} \)) differentially methylated between all male and female samples (Supplemental Fig. 9). As expected, the majority (\( n = 7538, 93.53\% \)) of these differences are located on the X Chromosome and are likely associated with X Chromosome dosage compensation mechanisms in females. However, an appreciable proportion of sex differences (\( n = 521, 6.47\% \)) are autosomal (see Fig. 3A and Table 3 for examples). Of these, approximately half are hypermethylated in females (\( n = 261, 50.10\% \)). A full list of all 521 Bonferroni-significant autosomal sex DMRs can be downloaded from our website (http://epigenetics.iop.kcl.ac.uk/fetalbrain/sDMRs.xlsx) and is also given in Supplemental File 3. Of the 521 sites showing sex differences, 119 overlap with significant autosomal DNA methylation differences identified between males and females in a recent analysis of the postnatal adult cortex (Xu et al. 2014), with 99.16% (\( n = 119 \)) of these displaying the same direction of effect (Supplemental Table 9). Furthermore, there is a highly significant correlation (\( r = 0.88, P < 1 \times 10^{-205} \)) between the magnitude and direction of sex differences in DNA methylation at autosomal probes reported as significantly different in the Xu et al. (2014) analysis of adult brain (\( n = 544 \)) and sex differences at the same CGI sites in fetal brain (Supplemental Fig. 10), indicating that most gender differences in the brain methylome manifest early in fetal development and are stable across the life course. A total of 1099 DMRs between males and females in the fetal brain were identified (Sidak-corrected \( P < 0.05 \), number of probes \( \geq 3 \)) (Supplemental Table 3), of which the majority (\( n = 908, 82.62\% \)) are located on the X Chromosome, with 191 (17.38%) located on the autosomes (see Supplemental Table 10; Supplemental Fig. 11). A list of all significant sex-associated DMRs is available from our laboratory website (http://epigenetics.iop.kcl.ac.uk/fetalbrain/1099_sex_DMRs.csv) and in Supplemental File 4. We next explored sex-specific changes in DNA methylation across fetal brain development by analyzing interactions between sex and developmental age. We identified 61 sites (\( n = 59 \) autosomal, \( n = 2 \) X Chromosomes) characterized by significant (Bonferroni-corrected threshold: \( P < 1.22 \times 10^{-7} \)) sex-specific DNA methylation trajectories across brain development (Table 4; Supplemental Fig. 12). The top-ranked autosomal sex-specific dDMR (cg23916284, \( P = 1.72 \times 10^{-15} \)) (Fig. 3B) is located in the gene body of RBPF, which encodes a transcriptional regulator in the Notch signaling pathway that is involved in neurogenesis and neuronal maturation. The pattern of DNA methylation observed at this site is typical of other autosomal sex-specific dDMRs; a high level of DNA methylation is observed in both males and females until –100 DPC, when males diverge from males, becoming progressively hypomethylated. These findings should be treated with some caution given the low number of samples at these later gestational ages. The two significant X Chromosome sex-specific dDMRs (cg09167861 [HHL1], \( P = 3.50 \times 10^{-9} \); cg00718858, \( P = 9.22 \times 10^{-8} \)) show more distinct male and female trajectories across brain development (Fig. 3C) and potentially reflect dynamic changes in X Chromosome dosage compensation mechanisms. Using comb-p, we identified 66 DMRs displaying sex-specific DNA methylation trajectories across brain development (Supplemental Fig. 13; Supplemental Tables 3, 11). A complete list of regions characterized by sex-specific DNA methylation trajectories can be downloaded from http://epigenetics.iop.kcl.ac.uk/66_age_sex_DMRs.csv and is also given in Supplemental File 5.

There are distinct modules of comethylated loci in the developing human brain

We next employed weighted gene coexpression network analysis (WGCNA) (Langfelder and Horvath 2008) to undertake a systems-level characterization of the DNA methylation changes associated with human brain development. WGCNA identified 22 discrete modules of comethylated positions (Supplemental Table 12), and the first principal component of each individual module (the “module eigengene”) was used to assess their relationship with fetal brain development (Fig. 4A). Nine modules were found to be significantly correlated with DPC, most notably the “yellow” (\( n = 21,999 \) probes, \( r = -0.59, P = 3.94 \times 10^{-18} \)) and “brown” modules (\( n = 24,002 \) probes, \( r = 0.54, P = 5.63 \times 10^{-15} \)) (Fig. 4B; Supplemental Fig. 14). Module membership within these modules is
highly correlated with probe significance (yellow module: \( r = 0.79, P < 1 \times 10^{-13} \); brown module: \( r = 0.72, P < 1 \times 10^{-39} \)) (Fig. 4C), showing a clear relationship between DNA methylation at the core members of each module and fetal brain development (Fig. 4D). To test whether the identified modules were biologically meaningful, over-enrichment analyses (see Methods) were performed using gene ontology (GO) pathways (Ashburner et al. 2000). Both modules were highly enriched for specific functional pathways directly related to development of the nervous system and neurobiological function (Supplemental Tables 13, 14). In particular, the brown module was highly enriched for pathways related to development of the brain including epithalamus development (odds ratio = 78.18, \( P = 3.79 \times 10^{-28} \)), central nervous system neuron axonogenesis (odds ratio = 8.26, \( P = 9.70 \times 10^{-21} \)), neuron maturation (odds ratio = 6.97, \( P = 9.78 \times 10^{-15} \)), and nervous system development (odds ratio = 1.61, \( P = 1.25 \times 10^{-13} \)). In contrast, modules not associated with fetal brain development showed no significant enrichment of neurodevelopmental functions (for example, see Supplemental Table 15 for pathways associated with the light green WGCNA module). The two modules strongly associated with sex across all fetal brain samples (light cyan, Supplemental Table 16; black, Supplemental Table 17) were, as expected, primarily comprised of X-linked loci and enriched for pathways relating to epigenetic processes (e.g., negative regulation of histone H3K36 methylation, odds ratio = 63.83, \( P = 7.34 \times 10^{-30} \); light cyan module), sex-specific genomic regulation (e.g., sex determination, odds ratio = 6.97, \( P = 7.45 \times 10^{-31} \); black module), and brain development (e.g., establishment or maintenance of neuroblast polarity, odds ratio = 43.97, \( P = 7.26 \times 10^{-33} \); black module). These analyses suggest that the large coordinated changes in DNA methylation identified across our samples relate specifically to development of the brain. It is, however, noteworthy that a recent “epigenetic clock” algorithm developed by Steve Horvath (Horvath 2013), based on DNA methylation values at 353 CpG sites shown to be associated with postpartum age, accurately predicts that each of our samples is prenatal, with a significant positive correlation (\( r = 0.15, P = 0.038 \)) between predicted and actual age at the resolution of days (Fig. 4E). This supports the notion that a core set of loci are likely to be more generally associated with aging across the life course and that these basic aging processes start very early in fetal life and occur across tissues.

dDMPs are common in several genomic regions associated with two neurodevelopmental disorders (schizophrenia and autism)

Schizophrenia and autism spectrum disorder are highly heritable neuropsychiatric disorders that are believed to have an early neurodevelopmental component (Weinberger 1995; Geschwind and

Figure 3. Sex-specific DNA methylation differences in the developing fetal brain. A total of 8059 (1.99% of the 408,594 autosomal and X-linked probes assessed) were identified as significantly differentially methylated between males and females in the fetal brain, with most differences occurring on the X Chromosome as expected. However, a number of autosomal loci also show significant differences between males and females, as illustrated by the four probes shown in A. Of note, DNA methylation at a number of (B) autosomal and (C) X Chromosome probes showed significant interactions between sex and developmental stage, indicating differential trajectories of DNA methylation in the developing male and female brain.
Levitt 2007; Fatemi and Folsom 2009). A recent large-scale genome-wide association study (GWAS) of schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014) identified 108 conservatively defined loci meeting genome-wide significance. To aid the functional characterization of these loci and identify genes potentially involved in the neurodevelopmental origins of this disorder, we provide annotation of the fetal brain dDMPs within the 89 genomic regions containing at least one informative CpG site (Supplemental Table 18). Of note, several specific schizophrenia-associated genomic regions contain multiple dDMPs (for example, see Fig. 5A,B), highlighting regions of potential interest for understanding the neurodevelopmental origins of this disorder. Likewise, for autism, high-confidence candidate gene regions were selected from the Simons Foundation Autism Research Initiative (SFARI) gene database (Basu et al. 2009; Abrahams et al. 2013) and complemented by loci identified in recent exome sequencing and selected from the Simons Foundation Autism Research Initiative gene database (Basu et al. 2009; Abrahams et al. 2013) and has been hypothesized to enhance transcription elongation efficiency, suppress the spurious initiation of transcription within active genes, contribute to splicing mechanisms, and regulate alternative promoter use. These changes in gene body DNA methylation during brain development is therefore interesting given the dramatic transcriptional plasticity occurring in the brain at this time (Colantuoni et al. 2011).

**Table 3.** The 20 top-ranked autosomal sex differences in DNA methylation in the fetal brain

<table>
<thead>
<tr>
<th>Probe</th>
<th>P-value</th>
<th>Chr</th>
<th>Position</th>
<th>Gene(s)</th>
<th>Genic probe location</th>
<th>Relation to CpG island</th>
<th>Male mean</th>
<th>Male SD</th>
<th>Female mean</th>
<th>Female SD</th>
<th>Difference (Female - Male)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg12691488</td>
<td>$1.66 \times 10^{-8}$</td>
<td>1</td>
<td>243053673</td>
<td>KRT77</td>
<td>Body</td>
<td>Island</td>
<td>36.11</td>
<td>4.08</td>
<td>15.28</td>
<td>3.87</td>
<td>$-20.83$</td>
</tr>
<tr>
<td>cg03691818</td>
<td>$7.59 \times 10^{-6}$</td>
<td>12</td>
<td>53085038</td>
<td>RTN1</td>
<td>Body</td>
<td>Island</td>
<td>9.06</td>
<td>1.00</td>
<td>16.08</td>
<td>2.15</td>
<td>$-7.01$</td>
</tr>
<tr>
<td>cg11643285</td>
<td>$8.78 \times 10^{-6}$</td>
<td>3</td>
<td>16116667</td>
<td>KRT77</td>
<td>Body</td>
<td>Island</td>
<td>72.38</td>
<td>2.79</td>
<td>83.76</td>
<td>2.95</td>
<td>$-11.38$</td>
</tr>
<tr>
<td>cg02325951</td>
<td>$1.99 \times 10^{-5}$</td>
<td>14</td>
<td>8978619</td>
<td>FOXN3</td>
<td>Body</td>
<td>Shelf</td>
<td>78.28</td>
<td>4.48</td>
<td>65.08</td>
<td>3.17</td>
<td>$-13.20$</td>
</tr>
<tr>
<td>cg03687700</td>
<td>$3.97 \times 10^{-5}$</td>
<td>2</td>
<td>24271844</td>
<td>FKBP18</td>
<td>TSS1500</td>
<td>Shore</td>
<td>20.34</td>
<td>3.18</td>
<td>32.55</td>
<td>4.28</td>
<td>$-12.12$</td>
</tr>
<tr>
<td>cg04946709</td>
<td>$2.93 \times 10^{-5}$</td>
<td>16</td>
<td>59789030</td>
<td>Body</td>
<td>Island</td>
<td>Shore</td>
<td>84.30</td>
<td>2.49</td>
<td>76.47</td>
<td>2.56</td>
<td>$-7.83$</td>
</tr>
<tr>
<td>cg15817705</td>
<td>$5.81 \times 10^{-4}$</td>
<td>1</td>
<td>209406063</td>
<td>Body</td>
<td>Island</td>
<td>Shore</td>
<td>84.38</td>
<td>2.60</td>
<td>75.70</td>
<td>3.22</td>
<td>$-8.67$</td>
</tr>
<tr>
<td>cg22345911</td>
<td>$4.08 \times 10^{-4}$</td>
<td>17</td>
<td>80321263</td>
<td>CSNK1D</td>
<td>5' UTR, 1st exon</td>
<td>Island</td>
<td>10.24</td>
<td>0.74</td>
<td>12.67</td>
<td>0.95</td>
<td>$-2.43$</td>
</tr>
<tr>
<td>cg23719534</td>
<td>$2.17 \times 10^{-4}$</td>
<td>15</td>
<td>101099284</td>
<td>Body</td>
<td>Island</td>
<td>Island</td>
<td>82.86</td>
<td>2.76</td>
<td>90.64</td>
<td>2.79</td>
<td>$-7.77$</td>
</tr>
<tr>
<td>cg11841231</td>
<td>$2.20 \times 10^{-3}$</td>
<td>2</td>
<td>20554339</td>
<td>PARD3B</td>
<td>Body</td>
<td>Island</td>
<td>78.89</td>
<td>3.29</td>
<td>68.67</td>
<td>4.83</td>
<td>$-10.22$</td>
</tr>
<tr>
<td>cg06711937</td>
<td>$6.65 \times 10^{-3}$</td>
<td>13</td>
<td>23489940</td>
<td>Body</td>
<td>Island</td>
<td>Island</td>
<td>8.63</td>
<td>0.70</td>
<td>12.23</td>
<td>2.03</td>
<td>$-3.59$</td>
</tr>
<tr>
<td>cg23001456</td>
<td>$2.45 \times 10^{-3}$</td>
<td>17</td>
<td>2615074</td>
<td>CLUH</td>
<td>TSS200</td>
<td>Island</td>
<td>8.51</td>
<td>0.60</td>
<td>10.00</td>
<td>0.66</td>
<td>$-1.49$</td>
</tr>
<tr>
<td>cg03218192</td>
<td>$1.23 \times 10^{-3}$</td>
<td>17</td>
<td>33914403</td>
<td>AP2B1</td>
<td>1st exon;</td>
<td>Island</td>
<td>21.11</td>
<td>2.25</td>
<td>26.58</td>
<td>2.56</td>
<td>$-5.46$</td>
</tr>
<tr>
<td>cg06642617</td>
<td>$5.05 \times 10^{-3}$</td>
<td>2</td>
<td>73461528</td>
<td>CCT7</td>
<td>Body; TSS1500; 5' UTR</td>
<td>Island</td>
<td>8.37</td>
<td>1.17</td>
<td>6.30</td>
<td>0.41</td>
<td>$-2.07$</td>
</tr>
<tr>
<td>cg04874129</td>
<td>$5.48 \times 10^{-3}$</td>
<td>16</td>
<td>55690873</td>
<td>SLC6A2</td>
<td>1st exon</td>
<td>Island</td>
<td>11.87</td>
<td>1.98</td>
<td>16.22</td>
<td>1.94</td>
<td>$4.36$</td>
</tr>
<tr>
<td>cg24016844</td>
<td>$3.97 \times 10^{-3}$</td>
<td>11</td>
<td>111506641</td>
<td>LR1F1</td>
<td>TSS1500</td>
<td>Island</td>
<td>16.04</td>
<td>2.18</td>
<td>12.08</td>
<td>1.38</td>
<td>$-3.97$</td>
</tr>
<tr>
<td>cg22921760</td>
<td>$2.63 \times 10^{-3}$</td>
<td>14</td>
<td>22005368</td>
<td>SALL2</td>
<td>TSS200</td>
<td>Island</td>
<td>24.40</td>
<td>2.76</td>
<td>30.36</td>
<td>2.95</td>
<td>$-5.96$</td>
</tr>
<tr>
<td>cg21148594</td>
<td>$3.50 \times 10^{-3}$</td>
<td>14</td>
<td>65704156</td>
<td>Body</td>
<td>Island</td>
<td>68.33</td>
<td>4.05</td>
<td>60.69</td>
<td>3.79</td>
<td>$-7.64$</td>
<td></td>
</tr>
<tr>
<td>cg12614789</td>
<td>$3.72 \times 10^{-3}$</td>
<td>15</td>
<td>76638473</td>
<td>Body</td>
<td>Island</td>
<td>Island</td>
<td>15.81</td>
<td>2.51</td>
<td>20.02</td>
<td>2.79</td>
<td>$-4.22$</td>
</tr>
<tr>
<td>cg09977154</td>
<td>$4.15 \times 10^{-3}$</td>
<td>16</td>
<td>89557657</td>
<td>ANKRD11</td>
<td>TSS1500</td>
<td>Island</td>
<td>11.15</td>
<td>1.70</td>
<td>8.22</td>
<td>1.12</td>
<td>$-2.93$</td>
</tr>
</tbody>
</table>

Mean and standard deviation (SD) represent DNA methylation (%). A full list of all 521 Bonferroni-significant autosomal sex DMPs can be downloaded from http://epigenetics.iop.kcl.ac.uk/fetalbrain/dDMPs.xlsx and is also given in Supplemental File 3.
DNA methylation at 61 CpG sites was characterized by Bonferroni-significant interaction between developmental stage (DPC) and sex. Chromosomal coordinates correspond to human genome build Feb. 2009 (GRCh37/hg19). Examples of specific interactions are shown in Figure 3, B and C.
Figure 4. Gene comethylation modules in the developing human brain. (A) Heatmap representing the correlation between module eigenvalues (ME) and the samples traits of DPC and sex. Each row represents a module, as indicated on the y-axis, and each column a trait. As shown in the color scale bar, strong positive correlation is indicated by dark red, strong negative correlation is indicated by dark green, and white indicates no correlation. Each cell contains the corresponding correlation and \( P \)-value given in parentheses. The yellow (\( P = 3.94 \times 10^{-18} \)) and brown (\( P = 5.63 \times 10^{-15} \)) WGCNA modules are most significantly associated with human fetal brain development. (B) The module eigengene of the yellow and brown modules is significantly associated with brain development, and (C) module membership in both modules is strongly correlated with probe significance with brain development (yellow module: \( r = 0.79, P < 1 \times 10^{-19} \); brown module: \( r = 0.72, P < 1 \times 10^{-19} \)). Probe significance is the absolute value of the correlation between DNA methylation at a CpG site and developmental age, and module membership is defined as the correlation between the module eigengene and the probe methylation profile. (D) Heatmap of probes in the yellow and brown modules showing changes across brain development. Color corresponds to DNA methylation level at probes with a module membership >0.85. (E) There is a significant correlation (\( r = 0.15, P = 0.038 \)) between predicted and actual age at the resolution of days.
Figure 5. Fetal brain dDMPs in the vicinity of genetic loci associated with neurodevelopmental disorders. (A) Fetal brain dDMPs within high-confidence regions identified in a recent large GWAS of schizophrenia may provide insight regarding causal variants, e.g., region Chr 20: 37361494–37485994 contains two genes, of which one overlaps with five highly significant dDMPs: PPP1R16B. The position of the most significant SNP (rs6065094) is indicated by a dashed red line. (B) The top-ranked dDMP (cg26680502) in the schizophrenia-associated region displays significant hypermethylation across fetal brain development \((P = 9.51 \times 10^{-16})\). (C) Examples of dDMPs in the vicinity of autism candidate gene NRXN1, with (D) the top-ranked NRXN1 probe (cg23851515) becoming significantly hypomethylated across fetal brain development \((P = 3.23 \times 10^{-14})\). (E) Examples of dDMPs in the vicinity of autism candidate gene SHANK1, with (F) the top-ranked probe (cg01427575) becoming significantly hypomethylated across fetal brain development \((P = 5.91 \times 10^{-9})\).
2006; Fujimoto et al. 2009; Imayoshi et al. 2010). Of note, recent evidence suggests there is considerable interaction between NRG signaling and sex hormones including estradiol (Soares et al. 2004; Arevalo et al. 2011; Ruiz-Palmero et al. 2011). Sex hormones are believed to contribute to the sexual differentiation of the brain that occurs in the second half of pregnancy (Bakker and Brock 2010) and are known to be potent modifiers of epigenetic status and transcription in the brain (Ghahramani et al. 2014). In this regard, it is interesting that male-female divergence at many loci occurs at ~100 DPC (see Fig. 3B), coinciding with a testosterone peak occurring between weeks 12 and 18 of pregnancy that is believed to mediate sexual differentiation of the brain (Finegan et al. 1989).

Schizophrenia and autism spectrum disorders are common neuropsychiatric disorders that are believed to have an early neurodevelopmental component (Weinberger 1995; Geschwind and Levitt 2007; Fatemi and Folsom 2009). Understanding the epigenomic changes occurring during fetal brain development could therefore provide useful insight about the molecular etiology of these conditions, especially given recent evidence for disease-associated epigenetic dysfunction in the brain (Mill et al. 2008; Ladd-Acosta et al. 2013; Pidsley et al. 2014). We include the location of dDMRs in relation to high-confidence schizophrenia- and autism-associated loci, which could provide clues to the location of functional genetic variants contributing to the etiology of these disorders.

There are several limitations to this study. First, legal restrictions on later-term abortions precluded the assaying of brain samples from later stages of fetal development. However, in comparison with previous studies, we have assessed a much larger number of brain samples spanning a wider range of fetal ages (Numata et al. 2012; Lister et al. 2013). For example, the largest previous study (Numata et al. 2012) assayed only 30 prenatal samples spanning ~100 to 130 DPC. Another limitation is that the brain tissue used in the present study was acquired frozen and undissected, and was therefore processed as a composite of brain regions and cell types. Although this precludes us from specifying a particular brain region or cell type from which our data are derived, our results have the potential advantage of being representative of the predominant changes in the developing human brain as a whole. Third, although the Illumina 450K platform can accurately quantify DNA methylation at single-base resolution with probes associated with 99% of genes and 96% of CGIs, the array targets ~2% of the CpG sites in the human genome, and probes are not equally distributed across all genomic features. As costs diminish, future studies should employ sequencing-based genomic-profiling technologies to more thoroughly interrogate the epigenome across development in large numbers of samples. Fourth, it is likely that the Bonferroni-corrected P-value cutoff applied in this study is overly conservative due to the nonindependence of CpG sites (Shoemaker et al. 2010). Although we report that ~7% of probes represent true fetal brain dDMRs, the reality is likely to be higher. Fifth, no demographic or phenotypic information other than developmental age was available for the samples assayed, although we were able to determine sex using an XY Chromosome PCR assay and the DNA methylation data. Sixth, we do not have gene expression data from these samples and are therefore unable to make direct conclusions about the transcriptional consequences of the developmental changes in DNA methylation we observe. Using publicly available microarray gene expression data from an independent set of fetal brain samples, however, we find evidence that a number of genes in the vicinity of our top-ranked dDMRs are differentially expressed across brain development. Given the low number of samples and limited developmental ages available in the resource utilized (Colantuoni et al. 2011), as well as the lack of information on alternative splicing and isoform abundance, these data suggest that many of the fetal brain dDMRs will have transcriptional consequences in the developing brain. Finally, we are not able to distinguish between DNA methylation and its oxidized derivative, DNA hydroxymethylation, as standard bisulfite sequencing does not distinguish between these modifications (Huang et al. 2010). This is an important confounder given that the latter is known to be abundant and to have functional relevance in the brain (Globisch et al. 2010; Khare et al. 2012; Lister et al. 2013). It will therefore be important to further characterize these modifications across fetal brain development using techniques capable of discriminating DNA methylation and DNA hydroxymethylation (Booth et al. 2012), as well as additional epigenetic modifications, including the more highly oxidized DNA methylation derivatives, 5-formylcytosine and 5-carboxycytosine, and histone modifications (Pongthongkum et al. 2014).

In conclusion, we identify highly dynamic hypo- and hypermethylation occurring throughout the genome during human fetal brain development, with notable sex-specific methylation trajectories observed for several dDMRs. There are discrete modules of comethylated loci associated with brain development that are significantly enriched for genes involved in neurodevelopmental processes. Our data indicate that epigenetic changes during human brain development are highly dynamic, sometimes sex-specific, and may provide clues to the origins of common neurodevelopmental disorders.

### Methods

#### Samples

Human fetal brain tissue was acquired from the Human Developmental Biology Resource (HDBR) (http://www.hdbr.org) and the MRC Brain Banks network (http://www.mrc.ac.uk/research/facilities/brain-banks/access-for-research). Ethical approval for the HDBR was granted by the Royal Free Hospital research ethics committee under reference 08/H0712/34 and Human Tissue Authority (HTA) material storage license 12220; ethical approval for the MRC Brain Bank was granted under reference 08/ME09/38. The age of these samples ranged from 23 to 184 DPC, determined by Carnegie staging in the case of embryonic samples (defined as ≥56 DPC) and foot and knee to heel length measurements for fetal samples (defined as ≤57 DPC). No additional phenotypic or demographic information was available. An overview of the samples is given in Figure 1A and Supplemental Table 20. Brain tissue was obtained frozen and had not been dissected into regions. We therefore homogenized half of the brain tissue from each individual fetus for subsequent genomic DNA extraction, which was performed by standard phenol-chloroform procedures. Sample sex was determined via PCR amplification of a region in the vicinity of the amelogenin gene (AMELY) that produces different sized PCR products for the X and Y Chromosomes (977 and 788 bp, respectively) (Eng et al. 1994). Following methylation array processing, sex-calls were confirmed using multidimensional scaling plots of the variable XY Chromosome probes (see below).

#### Genome-wide DNA methylation quantification

All samples were randomized with respect to age and sex to avoid batch effects throughout all experimental procedures. Genomic DNA (500 ng) from each sample was treated in duplicate with sodium bisulfite using the Zymo EZ DNA Methylation-Lightning Kit (Zymo Research). Genome-wide DNA methylation was quantified using the Illumina Infinium HumanMethylation450 BeadChip (Illumina) and scanned on the HiScan System (Illumina). Illumina
GenomeStudio software (Illumina) was used to extract signal intensities for each probe, generating a final report that was imported into the R statistical environment 3.0.2 (http://www.r-project.org) (R Development Core Team 2005) using the methylumi and minfi packages (http://www.bioconductor.org/packages/release/bioc/html/methylumi.html). Data quality control and preprocessing were performed using the watermelon package as described previously (Pidsley et al. 2013). Stringent filtering of the prenormalized Illumina 450K data was performed using the pfilter function; CpG sites with a beadcount of <3 in 5% of samples were removed (n = 146) as were CpG sites with a detection P-value > 0.05 in 1% of samples (n = 22,379). Finally, cross-reactive probes and polymorphic CpGs, as detailed in the Illumina annotation file and identified in recent publications (Chen et al. 2013; Price et al. 2013), were removed (Supplemental Table 21), leaving 408,608 probes for further analysis (including n = 9230 X Chromosome and n = 14 Y Chromosome probes). Data were normalized with the dasen function of the watermelon package. Demographic information for the final cohort of the 179 samples remaining following quality control (n = 100 males [age range = 23 to 184 DPC]; n = 79 females [age range = 37 to 161 DPC]) is shown in Supplemental Table 20.

Statistical analysis
A multiple linear regression model with DPC and sex was used to identify methylation changes associated with brain development or sex. Autosomal probes (n = 399,364) were considered significantly associated with DPC if they passed a Bonferroni-corrected significance threshold of P < 1.25 × 10⁻⁷ (α = 0.05), while autosomal and X Chromosome probes (n = 408,594) were considered to be significantly associated with sex if they passed a Bonferroni-corrected significance threshold of P < 1.22 × 10⁻⁷ (α = 0.05). The model was then extended to include an interaction term between sex and DPC and fitted for all autosomal and X Chromosome probes (n = 408,594); a significance threshold of P < 1.22 × 10⁻⁷ (α = 0.05) was used to identify probes with different developmental methylation trajectories between the sexes.

WGCNA analysis
Weighted gene co-expression network analysis (WGCNA) was used to identify modules of highly coexpressed probes (Langfelder and Horvath 2008). Modules were identified from the pairwise correlation, using a signed block-wise network construction at power 4. Modules identified were assigned a color name, and a weighted average-methylation profile, known as a module eigengene (ME), was calculated for each. Correlations between the MEs and the phenotypic traits available (DPC, sex) were used to identify modules associated with these traits. The relationship between each probe and each module was assessed through calculation of module membership (MM), the absolute correlation between a probe’s DNA methylation status and the ME, allowing the identification of the subset of probes with the strongest membership to each module. Probe significance (PS), the absolute value of correlation between each probe and phenotypic trait, was calculated to quantify the association of probes with age. To assess the biological meaning of modules associated with age and sex, genes associated with probes possessing an absolute MM value of >0.85 were extracted and assessed using pathway and gene ontology analysis.

Gene ontology and pathway analysis
Gene annotations provided by Illumina were used to map probes to genes. Gene ontology terms were downloaded and annotated to all probes based on these gene annotations where all parent terms for each child term were also included. Probes without genetic annotation or without any gene ontology annotation were excluded from the subsequent analysis. Fisher’s exact tests were used to test for an over-enrichment of WGCNA module probes in each pathway compared to all probes passing QC and inputted to the WGCNA algorithm. Only pathways containing at least 50 probes were considered. By running the analysis at the probe level, we were able to control for the uneven distribution of probes per gene, minimizing any bias introduced by genes with an above average number of probes.

Enrichment analyses in genomic regions associated with autism
Candidate loci were selected from the SFARI gene database (selecting genes having an evidence level of 1 [high confidence] or 2 [strong candidate]) (Basu et al. 2009; Abrahams et al. 2013), supplemented with genes implicated by exome sequencing studies (Neale et al. 2012; O’Roak et al. 2012a,b; Sanders et al. 2012), as well as those identified as high-confidence ASD genes in a recent study (Willsey et al. 2013). Probes annotated to these genes were tested for enrichment of dDMPs using with a two-tailed Fisher’s exact test, compared to the frequency of dDMPs in all gene-annotated probes on the Illumina 450K array (38,845 [11.03%] of 352,191 probes).

Data access
Raw and normalized illumina 450K methylation data has been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE58885. A searchable database of our fetal brain DNA methylation data is available at http://epigenetics.iop.kcl.ac.uk/fetalbrain.

Acknowledgments
This project was in part supported by an MRC project grant to J.M. and research funds provided by the University of Exeter Medical School. H.S. is supported by an MRC studentship. The human embryonic and fetal material was provided by the Joint MRC (grant G0700089/Welcome Trust (grant GR082557) Human Developmental Biology Resource (http://www.hdbr.org).

References


Genes Dev **25**:1010–1023.


Received July 14, 2014; accepted in revised form December 15, 2014.