REVISED MANUSCRIPT – NN-A52734A

Resource

Methylation quantitative trait loci in the developing brain and their enrichment in schizophrenia-associated genomic regions.

Eilis Hannon[1], Helen Spiers[2], Joana Viana[1], Ruth Pidsley[3], Joe Burrage[1], Therese M Murphy[1], Claire Troakes[2], Gustavo Turecki[5], Michael C. O'Donovan[4], Leonard C. Schalkwyk[6], Nicholas J. Bray[2,4*], Jonathan Mill[1,2*$]

[1] University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK.

[2] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, SE5 8AF, UK.

[3] Garvan Institute of Medical Research, Sydney 2010, NSW, Australia

[4] MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff CF24 4HQ, UK.

[5] Douglas Mental Health Institute, McGill University, Montreal H4H 1R3, QC, Canada

[6] School of Biological Sciences, University of Essex, Colchester CO4 3SQ, UK.

Eilis Hannon: e.j.hannon@exeter.ac.uk

Helen Spiers: helen.h.spiers@kcl.ac.uk

Ruth Pidsley: r.pidsley@garvan.org.au

Joana Viana: j.viana@exeter.ac.uk

Joe Burrage: J.Burrage@exeter.ac.uk

Therese Murphy: T.Murphy@exeter.ac.uk

Claire Troakes: Claire.troakes@kcl.ac.uk

Gustavo Turecki: gustavo.turecki@mcgill.ca

Michael O'Donovan: ODonovanMC@cardiff.ac.uk

Leonard Schalkwyk: lschal@essex.ac.uk

Nicholas Bray: BrayN3@cardiff.ac.uk

Jonathan Mill: j.mill@exeter.ac.uk

* These authors contributed equally

[$] Correspondence to: Jonathan Mill, University of Exeter Medical School, RILD Building, Royal Devon & Exeter Hospital, Barrack Road, Exeter. EX2 5DW. UK. E-mail: j.mill@exeter.ac.uk

Abstract

We characterized DNA methylation quantitative trait loci (mQTLs) in a large collection (n=166) of human fetal brain samples spanning 56-166 days post-conception, identifying >16,000 fetal brain mQTLs. Fetal brain mQTLs are primarily *cis*-acting, enriched in regulatory chromatin domains and transcription factor binding sites, and show significant overlap with genetic variants also associated with gene expression in the brain. Using tissue from three distinct regions of the adult brain (prefrontal cortex, striatum and cerebellum) we show that most fetal brain mQTLs are developmentally stable, although a subset is characterized by fetal-specific effects. We show that fetal brain mQTLs are enriched amongst risk loci identified in a recent large-scale genome-wide association study (GWAS) of schizophrenia, a severe psychiatric disorder with a hypothesized neurodevelopmental component. Finally, we demonstrate how mQTLs can be used to refine GWAS loci through the identification of discrete sites of variable fetal brain methylation associated with schizophrenia risk variants.

Human brain development is orchestrated by complex transcriptional programs[1], which are guided and reinforced by epigenetic modifications to DNA and histone proteins. DNA methylation is the most extensively studied epigenetic modification, playing a key role in many important genomic regulatory processes[2]. Of note, the establishment and maintenance of cell- and tissue-specific DNA methylation patterns is crucial for normal mammalian development[3]. Although traditionally regarded as a mechanism of transcriptional repression, DNA methylation can be associated with both increases and decreases in gene expression[4], and has recently been implicated in other genomic functions including alternative splicing and promoter usage[5].

We recently characterized widespread changes in DNA methylation across human fetal brain development[6], although the factors influencing inter-individual methylomic variation during the prenatal period are unknown. Studies in a variety of tissues, including adult human brain[7,8], have shown that DNA methylation can be influenced by DNA sequence variation. These methylation quantitative trait loci (mQTLs) have been found to overlap with DNA variants associated with levels of gene expression (expression quantitative trait loci; eQTLs)[4,9], and may serve as markers for these as well as other genetic influences on gene regulation. Although mQTLs have previously been assessed in the adult human brain using low-resolution DNA methylation arrays[7,8], mQTLs in the developing human brain have not previously been explored.

In this study we combine high-density DNA methylation profiling with genome-wide SNP genotyping in a large (n = 166) collection of human brain samples from the first and second trimester of gestation. Given the growing evidence that many common variants associated with complex diseases act through effects on gene regulation[10,11], we subsequently test for enrichment of fetal mQTL amongst risk loci identified in a recent large-scale GWAS of schizophrenia[12], a neuropsychiatric disorder with a hypothesized neurodevelopmental component[13,14]. Finally, we show how mQTL data can be used to refine broad GWAS loci through the identification of discrete sites of variable fetal brain methylation associated with schizophrenia risk variants. As a resource to the wider community we have developed a searchable online database of fetal brain mQTLs that can be accessed at http://epigenetics.essex.ac.uk/mQTL/.

Results

Methylation quantitative trait loci (mQTL) in the developing human brain are widespread and predominantly characterized by cis effects:

We performed genome-wide single nucleotide polymorphism (SNP) genotyping and DNA methylation profiling in 166 human fetal brain samples ranging from 56 to 166 days post-conception (see **Online Methods** and **Supplementary Table 1**). After stringent quality-control we tested for an additive effect of allele dosage on DNA methylation across all potential pairings of 430,304 SNPs and 314,554 DNA methylation sites to identify fetal brain mQTLs (**Supplementary Table 1**). We identified 16,809 mQTLs at a conservative Bonferroni-corrected significance threshold of $P < 3.69 \times 10^{-13}$ (**Supplementary Tables 2** and **3**). The median DNA methylation change per allele across all identified mQTLs was 6.69% (interquartile range (IQR) = 3.17%-8.96%) for each mQTL SNP (**Supplementary Fig. 1**), slightly larger than reported in a previous analysis of genome-wide mQTLs in the adult brain (median effect size = 4.11%, IQR = 2.13-6.97%)[7]. The majority of mQTL SNPs (74.17%) are associated with DNA methylation at only a single probe; in contrast, most DNA methylation sites (69.83%) showing evidence for association do so with multiple mQTL SNPs, presumably as a result of linkage disequilibrium (LD) between SNPs (**Supplementary Fig. 2**). A searchable database of fetal brain mQTLs is available at: http://epigenetics.essex.ac.uk/mQTL/.

The majority of fetal brain mQTLs (96.3%) involve SNPs and DNA methylation sites on the same chromosome (**Fig. 1a** and **Supplementary Table 2**). We defined significant SNP-methylation relationships spanning <500kb as "*cis*-mQTLs" (n = 15,942, 94.8% of total), with those spanning >500kb or characterized by inter-chromosomal effects (n = 867, 5.16% of total) as representing *trans*-mQTLs. The strong enrichment of *cis*-mQTLs concurs with data from other tissues and cell-types[7,15-18]. Amongst the *cis*-mQTLs, both effect size (i.e. DNA methylation change per allele, **Fig. 1b**) and significance (i.e. *P*-value, **Supplementary Fig. 3**) are related to the distance between the Illumina 450K array probe and mQTL SNP.

Despite the preponderance of *cis*-mQTLs, there are some notable *trans*-mQTL effects (**Fig. 1c** and **Supplementary Table 4**), consistent with previous reports of long-range genetic regulation of epigenetic variation in multiple cell-types[19]. Although the average effect size for *trans*-mQTLs is significantly lower than that observed for *cis*-mQTLs (two-sided Wilcoxon rank sum test, $P = 6.74 \times 10^{-7}$) (**Supplementary Fig. 1**), there is a higher proportion of larger (DNA methylation change per allele > 25%) effects among *trans*-mQTLs than *cis*-mQTLs (1.04% vs 0.715%). Of the 178 DNA methylation sites identified as being associated with *trans*-acting genetic variation in the fetal brain, 50 (28.09%) and 108 (60.67%) were also identified in studies of pancreatic islet cells[16] and lymphocytes[19], respectively (**Supplementary Table 4** and **Supplementary Table 5**). These long-range associations

between genotype and DNA methylation complement data showing interactions between regulatory elements spanning several Mb[20], and even between chromosomes[21].

Fetal brain mQTLs are significantly enriched in functional regulatory domains:

We used data from ENCODE and the Roadmap Epigenomics Project[10,22-24] to assess whether sites characterized by genotype-associated DNA methylation co-localize with genomic regions associated with markers of transcriptional activity. We observed an enrichment of fetal brain mQTLs in genomic regions characterized by ChIP-seq peaks for repressive histone modifications in fetal brain – for example, H3K9me3 (relative enrichment = 1.43, P = 0.00107) and H3K27me3 (relative enrichment = 1.16, P = 0.000144) - and a significant depletion of fetal brain mQTLs in genomic regions defined by ChIP-seq peaks for histone modifications associated with active transcription – for example, H3K4me1 (relative enrichment = 0.828, $P = 6.55 \times 10^{-8}$) and H3K36me3 (relative enrichment = 0.543, $P = 1.39 \times 10^{-15}$ (**Supplementary Table 6**). Fetal brain mQTLs were found to be significantly enriched in regions of open chromatin indicated by DNase1 hypersensitivity sites (DHSs) identified in the adult human brain (**Supplementary Table 7**), consistent with the observation that intermediately methylated domains, one potential consequence of allele-specific DNA methylation, are enriched in DHSs[25]. We also identified a significant enrichment of genotype-associated DNA methylation sites overlapping annotated transcription factor binding sites identified by the ENCODE project[10,26] (relative enrichment = 1.26, $P = 2.96 \times 10^{-11}$) (**Supplementary Table 8**). Of note, there is a highly-significant enrichment (odds ratio = 1.35, $P = 1.66 \times 10^{-9}$) of fetal brain mQTLs influencing DNA methylation in CCCTC-binding factor (CTCF) motifs (**Supplementary Table 8**), confirming a finding from a previous study of heritable DNA methylation sites in the human brain[27]. CTCF is an 11 zinc-finger protein with insulator and chromatin barrier activity whose binding affinity is known to be strongly influenced by DNA methylation[28]. Given the important role of CTCF in core genomic processes including transcription, chromosomal interactions and chromatin structure[29], the enrichment of genetically-mediated DNA methylation at CTCF binding sites highlights an important potential mechanism linking genetic variation to genomic function. In addition to CTCF, a significant enrichment was also observed within binding sites for several other transcription factors, including IRF1, GABP, ELF1, Rad21 and CCNT2, with significant depletion in binding sites for others (e.g. SUZ12, CtBP2) (**Supplementary Table 8**).

Although the majority of fetal brain mQTLs are conserved in adult brain regions, there are fetal-specific genetic effects on DNA methylation at certain loci:

We next generated mQTL data from three adult human brain regions (prefrontal cortex (PFC), striatum (STR) and cerebellum (CER)) dissected from matched donors (n = 83; 21-96

years old, see **Online Methods** and **Supplementary Table 1**) to explore the extent to which fetal brain mQTLs are also present in the adult brain. Using a replication mQTL significance threshold of $P < 1x10^{-5}$, the majority (83.46%) of fetal brain mQTLs are present in at least one of the tested adult brain region (**Supplementary Table 9** and **Fig. 2a**), and there is a highly-significant overall correlation of individual mQTL effect sizes between fetal brain and each of the individual adult brain regions (PFC: r = 0.911, $P < 2.2x10^{-16}$; STR: r = 0.899, $P < 2.2x10^{-16}$; CER: r = 0.835, $P < 2.2x10^{-16}$) (**Supplementary Fig. 4**) across all Bonferroni-significant fetal brain mQTLs, even in mQTLs that do not meet our replication threshold (**Supplementary Fig. 5**). Of note, fetal brain mQTLs that do not replicate in adult brain are characterized by significantly lower effect sizes across all brain regions, including the fetal brain discovery sample ($P = 3.18x10^{-141}$) (**Supplementary Fig. 6**). Despite the overall strong concordance in the direction of mQTL effects between fetal and adult brain, there are notable examples of heterogeneity between fetal and adult brain tissue. We used a multilevel linear regression model to test the significance of an interaction term and identify differential mQTL effects across our datasets. Of the 10,663 fetal mQTL effects tested, 3,173 (29.76%) were found to be significantly heterogeneous (Bonferroni-corrected $P < 4.69 \times 10^{-6}$) across the fetal and adult datasets (**Supplementary Table 10**). These include mQTLs that had notably larger or notably smaller effects in the adult brain, and fetal-specific mQTLs showing no significant association with DNA methylation in any adult brain region (see **Fig. 2b**). We also identified a small number (n = 45) of fetal brain mQTLs that had opposite effects on DNA methylation in fetal and adult brain samples (**Fig. 2c** and **Supplementary Table 11**).

Fetal brain mQTLs overlap with genetic variants associated with RNA transcript abundance in the brain:

We used eQTL data from ten adult brain regions[30] to test whether identified fetal brain mQTLs overlap with genetic variants associated with RNA transcript abundance in *cis*. We compared the distribution of the minimum brain eQTL *P*-value of all interrogated SNPs split into the subsets of those identified as fetal brain mQTLs and those not (**Supplementary Fig. 7**), finding that variants associated with DNA methylation are indeed more likely to be associated with gene expression in *cis* (Wilcoxon rank sum test $P < 2.2x10^{-16}$). Of the 414,172 SNPs tested in both the mQTL and eQTL datasets, 9,869 were identified as being Bonferroni-significant *cis* mQTLs and 2,674 as Bonferroni-significant ($P < 5.99x10^{-9}$) eQTLs, with an overlap of 750 variants associated with 227 DNA methylation probes and 127 transcript probes (**Supplementary Table 12**). At a more relaxed eQTL threshold ($P < 1.00x10^{-7}$), there is an overlap of 1,042 variants associated with 315 DNA methylation probes and 183 transcript probes. A list of all variants associated with both DNA methylation and gene expression in *cis* is given in **Supplementary Table 13**. This overlapping set of variants likely

includes multiple SNPs in LD that are associated with one gene expression transcript and DNA methylation site. Because the extent and magnitude of LD varies across the genome, we established an LD-independent set of SNPs associated with DNA methylation and tested the overlap of these with the sentinelized subsignals – i.e. the most associated marker from a set in high LD ($r^2 > 0.8$) - from the brain eQTL dataset[30]. Compared to 1 million simulated mQTL SNP sets matched for allele frequency, this overlap was significantly greater than expected (relative enrichment = 4.23, $P < 1.00\text{x}10^{-6}$ after 1 million simulations; **Supplementary Table 14** and **Supplementary Fig. 8**).

There is a significant enrichment of schizophrenia-associated GWAS variants in fetal brain mQTLs:

Our catalogue of fetal brain mQTLs provides a unique resource for investigating putative functional consequences of genetic variation associated with postulated neurodevelopmental disorders such as schizophrenia. A recent large-scale GWAS identified 108 independent genomic loci exhibiting genome-wide significant association with the disorder ($P < 5\text{x}10^{-8}$), with evidence for a substantial polygenic component within signals that fall below this stringent level of significance[12]. Because the majority of these variants reside in regions of strong LD and do not index coding variants affecting protein structure, there remains considerable uncertainty about the causal genes involved in pathogenesis and the way in which they are functionally regulated by schizophrenia risk variants. We used PLINK[31] to 'clump' our list of significant ($P < 3.69\text{x}10^{-13}$) fetal brain mQTL variants into a set of quasi-independent SNPs (SNP pairwise $r^2 < 0.25$ within 250kb (non-major histocompatibility complex (MHC)) or 10000kb (MHC)), see **Online Methods**) and tested for enrichment of schizophrenia-associated variants across a range of GWAS significance-thresholds, using up to 1,000,000 simulated SNP sets to generate empirical *P* values (see **Online Methods**). We observed a highly-significant enrichment (relative enrichment = 4.11, $P = 3.0\text{x}10^{-6}$) of genome-wide significant schizophrenia risk variants amongst fetal brain mQTLs, with a trend for stronger enrichment at more stringent levels of GWAS significance (**Supplementary Fig. 9** and **Table 1**). To examine the specificity of any enrichment, these analyses were repeated using large GWAS datasets from i) a non-neurodevelopmental brain disorder (Alzheimer's disease (AD)[32]) and ii) two non-neurological phenotypes (body mass index (BMI)[33] and type 2 diabetes (T2D)[34]. Although our confidence in the enrichment of fetal brain mQTL in these datasets is limited by the smaller number of semi-independent GWAS SNPs, levels of enrichment were found to be notably lower for all other tested phenotypes (**Table 1**). Variants associated with AD are, however, nominally significantly enriched at the most relaxed GWAS threshold (GWAS threshold $P < 5\text{x}10^{-5}$: relative enrichment = 3.18, $P = 0.022$) and

several individual GWAS variants identified for this and the other tested phenotypes are also significant mQTLs in fetal brain mQTLs (**Supplementary Table 15**).

The identified mQTLs residing in schizophrenia-associated GWAS regions do not necessarily represent the actual causal risk variants; in many instances we are likely to be capturing 'passenger' effects whereby the variant influencing DNA methylation and the schizophrenia-associated SNP are instead co-segregating in the same LD block. Therefore we sought to identify instances where a likely causal risk variant for schizophrenia was an mQTL SNP. We used 1000 Genomes Project data (http://www.1000genomes.org/) to identify all variants in strong LD ($r^2 > 0.8$) with the 125 autosomal index SNPs associated with schizophrenia[12]. Of note, two of the actual schizophrenia GWAS index SNPs represent Bonferroni-significant fetal brain mQTLs: rs2535627 (associated with DNA methylation at cg11645453, $P = 3.05x10^{-13}$, **Supplementary Fig. 10**) and rs4648845 (associated with DNA methylation at cg02275930, $P = 4.54x10^{-15}$, **Supplementary Fig. 11**). 46 additional mQTL variants that are in strong LD with another six index SNPs are part of 86 highly-significant fetal brain mQTL pairs (**Supplementary Table 16**).

mQTLs can be used to localize putative causal loci within large genomic regions associated with schizophrenia:

To generate a more comprehensive database of mQTLs in the fetal brain and identify more examples where the same SNP is associated with both DNA methylation and disease, we imputed our genotype data using the most recent panel downloaded from the 1000 Genome Project (see **Online Methods**). Using an imputed set of 5,177,320 variants we identified an additional 256,040 mQTLs, which reflected the non-imputed dataset in terms of genomic distribution and observed effect sizes (**Supplementary Table 2** and **Supplementary Fig. 12**). The full list of fetal brain mQTLs after imputation can be downloaded from: http://epigenetics.essex.ac.uk/mQTL//All_Imputed_BonfSignificant_mQTLs.csv.gz. Our imputed data enabled us to identify 1,067 instances where the same SNP is associated with both DNA methylation and schizophrenia, with a comprehensive list available for download from:
http://epigenetics.essex.ac.uk/mQTL//PGC_IndexSNPs_QTLs_Inc2PCs_AllTissues_MatchSNPPosition.csv). Because they could be biased by the LD structure at associated loci the imputed mQTL data were not used for subsequent enrichment analyses, but they enabled us to further refine schizophrenia candidate regions and undertake co-localization analyses in order to identify variants associated with both DNA methylation and schizophrenia. We performed a Bayesian co-localization analysis[35] across the 105 autosomal regions associated with schizophrenia[12], spanning 19,378 DNA methylation sites included in our analysis. Instead of

focusing only on the intersection of significant variants associated with two phenotypes independently, this approach compares the pattern of association results from the schizophrenia GWAS and mQTL analyses across a region, combining the summary statistics into posterior probabilities for five hypotheses (see **Online Methods**). As this methodology assumes that the causal variant is present, or at least very well tagged, in the dataset, these co-localization analyses were performed using our imputed fetal brain mQTL dataset. The posterior probabilities for 65 regions, involving 296 DNA methylation sites in 306 pairs, were supportive of a co-localized association signal for both schizophrenia and DNA methylation in that region (PP3+PP4 > 0.99; **Supplementary Table 17**). Twenty-six of these pairs (covering 15 regions associated with schizophrenia) had a higher posterior probability for both schizophrenia and DNA methylation being associated with the same causal variant ($PP_4/PP_3 > 1$), with 16 (10 regions) of these having sufficient support for them to be considered as 'convincing' ($PP_4/PP_3 > 5$) according to the criteria of Giambartolomei and colleagues[35]. Of note, three of these top-ranked pairs are annotated to the *AS3MT* locus within a robust schizophrenia-associated region on chromosome 10 (**Supplementary Table 17**). The utility of mQTL mapping for localizing putative causal loci associated with disease within this region is shown in **Fig. 3**, with additional examples for other schizophrenia-associated regions available on our website (http://epigenetics.essex.ac.uk/mQTL/).

Schizophrenia-associated genomic regions are enriched for fetal-specific mQTLs:

Given the hypothesized neurodevelopmental component to schizophrenia, we examined the extent to which the mQTLs overlapping with schizophrenia-associated variants are characterized by fetal-specific effects (**Supplementary Table 16**). Across the 78 mQTLs also tested in adult brain samples, overall effect sizes were significantly larger in fetal brain than all adult brain regions tested (Wilcoxon rank-sum test: PFC $P = 0.0420$, STR $P = 0.00226$, CER $P = 0.00998$) (**Fig. 4**). Our heterogeneity analysis highlighted 16 (20.5%) instances where significantly different relationships between genotype and DNA methylation are found across the adult and fetal datasets, with eight classed as fetal-specific variants (i.e. those not reaching our replication threshold ($P < 1.00 \times 10^{-5}$) in any adult brain region) and the remaining eight demonstrating smaller effects across the adult brain.

Discussion

To explore the functional consequences of genetic variation in the developing human brain, we characterized mQTLs in human fetal brain samples (spanning 56-166 days post-conception), identifying over 16,000 associated pairs of SNPs and DNA methylation sites. We find that fetal brain mQTLs are significantly enriched in functional regulatory domains including DHSs, regions of repressive histone modifications, and specific transcription factor

binding sites across the genome, and show significant overlap with genetic variants influencing gene expression in the brain. Although the majority of fetal brain mQTLs appear to be conserved across adult brain regions, we find evidence for fetal-specific genetic effects at certain loci. Our data concur with findings from an independent study of cortical mQTLs across development [Jaffe et al, Nat Neuro, in press]; mQTL effects are highly consistent across both analyses (**Supplementary Fig. 13**) and largely conserved between fetal and adult brain.

There is growing evidence that the majority of common variants associated with complex traits act through effects on gene regulation[10,11]. Our data add to a growing literature showing that DNA methylation is genetically influenced[27], with mQTLs representing a potential mechanism linking genetic variation to complex phenotypes[4,9,36,37]. We find a significant enrichment of schizophrenia-associated GWAS variants in fetal brain mQTLs, indicating that common genetic variants conferring risk for schizophrenia are associated with altered DNA methylation in fetal human brain. The hypothesis that schizophrenia has an early neurodevelopmental component is supported by several lines of epidemiological and neuropathological evidence[13,14]. However, direct molecular evidence of schizophrenia risk factors operating in the fetal brain is scarce[38,39]. We have recently found that genomic loci that are differentially methylated between schizophrenia patients and unaffected controls in the adult brain are enriched at those undergoing dynamic changes in DNA methylation during human fetal brain development[6,40]. In the present study, we find that genetic variants exhibiting genome-wide significant association with schizophrenia[12] show a four-fold enrichment amongst fetal brain mQTLs, directly implicating altered gene regulation during fetal brain development in the etiology of the disorder.

To conclude, we report the first systematic analysis of genetically-mediated DNA methylation in the developing human brain. Our data support the hypothesis that a significant proportion of the genetic variants conferring schizophrenia risk have regulatory effects that become manifest early in the prenatal period, and demonstrate the utility of mQTL mapping for localizing putative causal loci associated with complex disease phenotypes within large genomic regions. As a resource to the wider community, we have developed a searchable online database of fetal brain mQTLs that can be accessed at http://epigenetics.essex.ac.uk/mQTL/.

Acknowledgements

This work was supported by grants from the UK Medical Research Council (MRC) grant numbers MR/K013807/1 (to J.M.) and MR/L010674/1 (to N.J.B.), and US National Institutes

of Health (grant number AG036039) to J.M. R.P. and H.S. were funded by MRC PhD studentships. The human embryonic and fetal material was provided by the Joint MRC (grant number G0700089)/Wellcome Trust (grant number GR082557) Human Developmental Biology Resource. We thank Dr Mike Weale for providing eQTL data from the BRAINEAC database.

Author contributions

J.M. and N.J.B. conceived and supervised the study, and obtained funding. E.H. undertook primary data analysis and bioinformatics. L.S. provided analytical support. H.S., J.V., R.P., T.M. and J.B. performed laboratory work. C.T. and G.T. provided samples for analysis. M.O'D. provided support fot GWAS enrichment analyses. E.H., N.J.B. and J.M. drafted the manuscript. All of the authors read and approved the final submission.

Competing financial interests

The authors declare no competing financial interests.

References for main text

1. Kang, H.J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483-9 (2011).
2. Jones, P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**, 484-92 (2012).
3. Ziller, M.J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477-81 (2013).
4. Wagner, J.R. *et al.* The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biol* **15**, R37 (2014).
5. Maunakea, A.K. *et al.* Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253-7 (2010).
6. Spiers, H. *et al.* Methylomic trajectories across human fetal brain development. *Genome Res* **25**, 338-52 (2015).
7. Gibbs, J.R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* **6**, e1000952 (2010).
8. Gamazon, E.R. *et al.* Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry* **18**, 340-6 (2013).
9. Gutierrez-Arcelus, M. *et al.* Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *Elife* **2**, e00523 (2013).
10. Maurano, M.T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190-5 (2012).
11. Nicolae, D.L. *et al.* Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* **6**, e1000888 (2010).
12. Schizophrenia Working Group of the PGC *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-+ (2014).
13. Fatemi, S.H. & Folsom, T.D. The neurodevelopmental hypothesis of schizophrenia, revisited. *Schizophr Bull* **35**, 528-48 (2009).

14. Weinberger, D.R. From neuropathology to neurodevelopment. *Lancet* **346**, 552-7 (1995).
15. Teh, A.L. *et al.* The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res* **24**, 1064-74 (2014).
16. Olsson, A.H. *et al.* Genome-wide associations between genetic and epigenetic variation influence mRNA expression and insulin secretion in human pancreatic islets. *PLoS Genet* **10**, e1004735 (2014).
17. Drong, A.W. *et al.* The presence of methylation quantitative trait loci indicates a direct genetic influence on the level of DNA methylation in adipose tissue. *PLoS One* **8**, e55923 (2013).
18. Gutierrez-Arcelus, M. *et al.* Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. *PLoS Genet* **11**, e1004958 (2015).
19. Lemire, M. *et al.* Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci. *Nat Commun* **6**, 6326 (2015).
20. Dekker, J., Marti-Renom, M.A. & Mirny, L.A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* **14**, 390-403 (2013).
21. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R. & Flavell, R.A. Interchromosomal associations between alternatively expressed loci. *Nature* **435**, 637-45 (2005).
22. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
23. Slieker, R.C. *et al.* Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin* **6**, 26 (2013).
24. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
25. Elliott, G. *et al.* Intermediate DNA methylation is a conserved signature of genome regulation. *Nat Commun* **6**, 6363 (2015).
26. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
27. McRae, A.F. *et al.* Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol* **15**, R73 (2014).
28. Wang, H. *et al.* Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* **22**, 1680-8 (2012).
29. Ong, C.T. & Corces, V.G. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet* **15**, 234-46 (2014).
30. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* **17**, 1418-28 (2014).
31. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
32. Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-8 (2013).
33. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
34. Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* **44**, 981-90 (2012).
35. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).
36. Meaburn, E.L., Schalkwyk, L.C. & Mill, J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics* **5**, 578-82 (2010).
37. van Eijk, K.R. *et al.* Identification of schizophrenia-associated loci by combining DNA methylation and gene expression data from whole blood. *Eur J Hum Genet* (2014).

38. Hill, M.J. & Bray, N.J. Evidence that schizophrenia risk variation in the ZNF804A gene exerts its effects during fetal brain development. *Am J Psychiatry* **169**, 1301-8 (2012).
39. Tao, R. *et al.* Expression of ZNF804A in human brain and alterations in schizophrenia, bipolar disorder, and major depressive disorder: a novel transcript fetally regulated by the psychosis risk variant rs1344706. *JAMA Psychiatry* **71**, 1112-20 (2014).
40. Pidsley, R. *et al.* Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *Genome Biol* **15**, 483 (2014).

Figure Legends

Figure 1: mQTLs in the developing human brain are predominantly *cis*-acting with effect size related to distance, although notable *trans*-effects are also present. (a) The genomic distribution of Bonferroni significant (P = 3.69 x10^{-13}) mQTLs in fetal brain, where the position on the x-axis indicates the location of Illumina 450K HumanMethylation array probes and the position on the y-axis indicates the location of informative SNPs. The color of the point corresponds to the difference in DNA methylation per allele compared to the reference allele, with the largest effects plotted in dark red. A clear positive diagonal can be observed demonstrating that the majority of mQTLs in fetal brain are associated with genotype in *cis*. (b) The relationship between mQTL effect size (DNA methylation change per allele) and distance between the Illumina 450K HumanMethylation array probe and informative SNPs in fetal brain, confirming the predominantly *cis* nature of mQTLs. A similar relationship is seen between mQTL significance and distance (see **Supplementary Fig. 3**). (c) *Trans*-mQTLs in the developing human brain. Shown are all Bonferroni significant (P = 3.69 x10-13) *trans*-mQTLs identified in fetal brain samples. The thickness of each line depicts association effect size, with color reflecting the chromosomal location of the mQTL SNP.

Figure 2: Despite a highly-significant overall correlation of individual mQTL effects between fetal brain and adult brain regions, a subset of loci are characterized by fetal-specific mQTLs. (a) Heatmap showing effect sizes in adult brain for all fetal brain mQTLs tested. Using a replication mQTL significance threshold of P < 1.00E-5, the majority (83.46%) of fetal brain mQTLs are present in at least one of the tested adult brain region. Despite the overall strong concordance in the direction of mQTL between fetal and adult brain, there are notable examples of heterogeneity in mQTL effects between fetal and adult brain tissue. Shown are examples of (b) a fetal-specific mQTL between rs10447470 and cg07900658 (heterogeneity P = 7.23E-40) and (c) an mQTL between rs2108854 and cg21577356 showing an opposite direction of effect in fetal brain and cerebellum (heterogeneity P = 8.39E-36).

Figure 3: mQTL mapping can localize putative causal loci associated with disease. Co-localization analyses yielded strong support for variants annotated to the *AS3MT* gene being associated with both schizophrenia and DNA methylation within a broad genomic region (chr10:104535135-105006335) identified in a recent GWAS analysis of schizophrenia[12]. All potential causal variants (r^2 > 0.8 with index variant) present in the imputed mQTL dataset are indicated by vertical blue lines and all DNA methylation probes within each region by

vertical red lines. Bonferroni-significant mQTLs are indicated by black lines between the respective variant and DNA methylation probe, where the line width reflects the magnitude of the effect. Additional examples of fetal brain mQTLs in genomic regions showing genome-wide significant association with schizophrenia are available on our website (http://epigenetics.essex.ac.uk/mQTL/).

Figure 4: Fetal mQTLs in schizophrenia-associated regions have larger effects on DNA methylation during neurodevelopment than the in adult brain. For robustly-associated schizophrenia GWAS variants characterized by human fetal brain mQTLs (**Supplementary Table 16**), we compared effect sizes between fetal and adult brain. Effect sizes (red to blue) for the corresponding mQTLs in adult brain are significantly lower across all three adult brain regions tested (Wilcoxon rank-sum test: PFC $P = 0.0420$, STR $P = 0.00226$, CER $P = 0.00998$). The heterogeneity *P*-value for each mQTL is depicted by the green (high heterogeneity) to yellow (low heterogeneity) column. White indicates that the mQTL was not tested in the adult brain samples.