OXFORD

GENERAL ARTICLE

Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic autism in post-mortem human brain tissue

Chloe C.Y. Wong[1], Rebecca G. Smith[2], Eilis Hannon[2], Gokul Ramaswami[3], Neelroop N. Parikshak[3], Elham Assary[4], Claire Troakes[1], Jeremie Poschmann[5], Leonard C. Schalkwyk[6], Wenjie Sun[7], Shyam Prabhakar[7], Daniel H. Geschwind[3,8,9] and Jonathan Mill[2,*]

[1]King's College London, Institute of Psychiatry, Psychology and Neuroscience, De Crespigny Park, London, SE5 8AF, UK, [2]University of Exeter Medical School, University of Exeter, Exeter, EX1 2LU, UK, [3]Center for Autism Research and Treatment, and Program in Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA, [4]Department of Biological and Experimental Psychology, School of Biological and Chemical Sciences, Queen Mary University of London, London, E1 4NS, UK, [5]Centre de Recherche en Transplantation et Immunologie, Institut de Transplantation Urologie Néphrologie (ITUN), CHU Nantes, Inserm, Université de Nantes, Nantes, France, [6]School of Biological Sciences, University of Essex, Colchester, CO4 3SQ, UK, [7]Computational and Systems Biology, Genome Institute of Singapore, 138672, Singapore, [8]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095-7088, USA and [9]Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095-6975, USA

*To whom correspondence should be addressed at: University of Exeter Medical School, RILD Building (Level 4), Royal Devon & Exeter Hospital, Barrack Rd, Exeter EX2 5DW, UK. Tel: 01392 408301; Email: j.mill@exeter.ac.uk

Abstract

Autism spectrum disorder (ASD) encompasses a collection of complex neuropsychiatric disorders characterized by deficits in social functioning, communication and repetitive behaviour. Building on recent studies supporting a role for developmentally moderated regulatory genomic variation in the molecular aetiology of ASD, we quantified genome-wide patterns of DNA methylation in 223 post-mortem tissues samples isolated from three brain regions [prefrontal cortex, temporal cortex and cerebellum (CB)] dissected from 43 ASD patients and 38 non-psychiatric control donors. We identified widespread differences in DNA methylation associated with idiopathic ASD (iASD), with consistent signals in both cortical regions that were distinct to those observed in the CB. Individuals carrying a duplication on chromosome 15q (dup15q), representing a genetically defined subtype of ASD, were characterized by striking differences in DNA methylation

across a discrete domain spanning an imprinted gene cluster within the duplicated region. In addition to the dramatic *cis*-effects on DNA methylation observed in dup15q carriers, we identified convergent methylomic signatures associated with both iASD and dup15q, reflecting the findings from previous studies of gene expression and H3K27ac. Cortical co-methylation network analysis identified a number of co-methylated modules significantly associated with ASD that are enriched for genomic regions annotated to genes involved in the immune system, synaptic signalling and neuronal regulation. Our study represents the first systematic analysis of DNA methylation associated with ASD across multiple brain regions, providing novel evidence for convergent molecular signatures associated with both idiopathic and syndromic autism.

Introduction

Autism spectrum disorder (ASD) encompasses a collection of complex neuropsychiatric disorders characterized by deficits in social interactions and understanding, repetitive behaviour and interests and impairments in language and communication development. ASD affects ~1% of the population and confers severe lifelong disability, contributing significantly to the global burden of disease (1,2). Evidence from neuroimaging, neuropathology, genetic and epidemiological studies has led to the conceptualization of ASD as a neurodevelopmental disorder, with etiological origins before birth (3,4). Quantitative genetic analyses have shown that ASD has a strong heritable component (5) with an emerging literature implicating rare single base-pair mutations, chromosomal rearrangements, *de novo* and inherited structural genomic variation and common (polygenic) risk variants in its pathogenesis (6–9). Despite the highly heterogeneous role of genetic variation in ASD, studies of transcriptional (10,11) and regulatory genomic variation (12) in post-mortem ASD brain provide evidence for a highly convergent molecular pathology, with individuals affected genetically defined subtypes of ASD sharing the core transcriptional signatures observed in idiopathic autism cases.

There is increasing evidence to support a role for non-sequence-based genomic variation in the aetiology of neurodevelopmental phenotypes including ASD (13). The epigenetic regulation of gene expression in the central nervous system is involved in modulating many core neurobiological and cognitive processes including neurogenesis (14,15), neuronal plasticity (16) and memory formation (17,18) and is known to be highly dynamic during human brain development (19). The dysregulation of epigenetic mechanisms underlies the symptoms of Rett syndrome and Fragile X syndrome, two disorders with considerable phenotypic overlap with ASD (20–22), and epigenetic variation has been recently associated with several neurodevelopmental phenotypes including ASD (12,23–31). Current epigenome-wide association studies of autism have focused primarily on DNA methylation, the best characterized and most stable epigenetic modification that acts to influence gene expression via physical disruption of transcription factor binding and through the attraction of methyl-binding proteins that initiate chromatin compaction and gene silencing (32). Despite finding evidence for ASD-associated methylomic variation, however, these analyses have been constrained by the analysis of small sample numbers and limited to the assessment of peripheral tissues or a single brain region (12,23–29).

In this study, we present results from the most systematic analysis of DNA methylation in ASD brain yet undertaken, quantifying methylomic variation in patients with idiopathic ASD (iASD) in addition to patients with a duplication of chromosome 15q11-13 ('dup15q'), which represents the most frequent cytogenetic abnormality associated with ASD occurring in ~1% of cases

(33,34). From each donor, we profiled matched post-mortem tissue from three brain regions—prefrontal cortex (PFC), temporal cortex (TC) and cerebellum (CB)—previously implicated in the pathophysiology of ASD. The frontal and temporal lobes, for example, play a role in social cognition, and animal models of ASD highlight cerebellar dysfunction (35–37). We find DNA methylation differences in both groups of iASD and dup15q patients, with consistent patterns of variation seen across the two cortical regions, distinct to those identified in CB. In addition to identifying dramatic *cis*-effects of the dup15q duplication in all three brain regions, we identify a significant overlap with the core methylomic differences observed in idiopathic autism iASD cases, reflecting findings from studies of transcriptional variation (11) and histone modifications (12).

Results

Methodological overview

We quantified DNA methylation across the genome using the Illumina Infinium HumanMethylation450 BeadChip ('450K array') in 223 post-mortem tissue samples comprising PFC, TC and CB dissected from 43 donors with ASD (including 7 patients with dup15q syndrome and 36 patients with iASD) and 38 non-psychiatric control subjects. After implementing a stringent quality control (QC) pipeline (see Materials and Methods), we obtained high-quality DNA methylation data from 76 PFC samples ($n = 36$ iASD patients, $n = 7$ dup15q patients, $n = 33$ controls), 77 TC samples ($n = 33$ iASD patients, $n = 6$ dup15q patients, $n = 38$ controls) and 70 CB samples ($n = 34$ iASD patients, $n = 7$ dup15q patients, $n = 29$ controls) (Supplementary Material, Table S1). Our primary analyses focused on identifying differentially methylated positions (DMPs) and differentially methylated regions (DMRs) associated with iASD and dup15q, controlling for cellular heterogeneity and other potential confounds, exploring the extent to which signals were shared across idiopathic and syndromic autism cases. Finally, we employed weighted gene co-methylation network analysis (WGCNA) to undertake a systems-level view of the DNA methylation differences associated with both iASD and dup15q across the three brain regions. An overview of our experimental approach is given in Supplementary Material, Figure S1.

DNA methylation differences between iASD cases and controls are consistent across cortical regions

No global differences in DNA methylation—estimated by averaging across all Illumina 450K array probes ($n = 417\,460$) included in our analysis—were identified between iASD patients and control subjects in any of the three brain regions (PFC: iASD = 48.4%, controls = 48.5%; TC: iASD = 48.4%, controls = 48.4%; CB:

iASD = 46.4%, controls = 46.4%). We observed a robust positive correlation between the estimated 'DNA methylation age'— calculated using an epigenetic clock based on DNA methylation values (38,39)—and recorded chronological age for each of the brain regions (PFC: $r = 0.98$, TC: $r = 0.97$, CB: $r = 0.94$) (Supplementary Material, Fig. S2), with no evidence for differential 'epigenetic aging' in iASD patients (PFC: $P = 0.10$, TC: $P = 0.24$, CB: $P = 0.80$). These findings indicate that ASD is not associated with any systemic differences in DNA methylation across the probes included on the Illumina 450K array in the brain regions tested, reflecting findings in studies of other complex neuropsychiatric phenotypes including Alzheimer's disease (40) and schizophrenia (41).

We next used a linear model including covariates for sex, age, brain bank and neuronal cell proportions derived from the DNA methylation data (except in the CB, as described in the Materials and Methods) to identify iASD-associated DMPs across the genome in each of the three brain regions. The top ranked iASD-associated DMPs in each brain region [PFC: cg08277486, which is located within *CCDC144NL* and hypermethylated in patients compared to controls ($P = 1.11e-06$); TC: cg08374799, which is located immediately upstream of *ITGB7* and hypomethylated in patients compared to controls ($P = 4.46e-07$); CB: cg01012394, which is located immediately upstream of *EYA3* and hypermethylated in patients compared to controls ($P = 1.01e-06$)] are shown in Figure 1A, with a list of all DMPs ($P < 5e-05$) detailed in Supplementary Materials, Tables S2–S4. Of note, iASD-associated DNA methylation differences are considerably more pronounced in both cortical regions (PFC: $n = 31$ DMPs; TC: $n = 52$ DMPs) than the CB ($n = 2$ DMPs). Hierarchical clustering of samples based on DNA methylation levels at these cortical DMPs distinguishes relatively well between iASD cases and controls in both PFC (Fig. 1B) and TC (Fig. 1C). As reported in previous analyses of epigenetic variation in the human brain (41,42), our data show that— at a global level—the patterns of DNA methylation in the CB are very distinct to the two cortical brain regions included in this study (Supplementary Material, Fig. S3). Effect sizes at iASD-associated DMPs are highly correlated between the two cortical regions but not between cortex and CB [top 100 PFC DMPs (PFC versus TC: $r = 0.77$, $P = 3.06e-21$; PFC versus CB: $r = 0.14$, $P = 0.18$), top 100 TC DMPs (TC versus PFC: $r = 0.81$, $P = 2.48e-24$; $r = 0.17$, $P = 0.09$) and top 100 CB DMPs (CB versus PFC: $r = 0.005$, $P = 0.96$; CB versus TC: $r = -0.03$, $P = 0.77$)] (Supplementary Material, Fig. S4). Given the striking consistency of effects across cortical regions, we used a multi-level linear mixed model (see Materials and Methods) to maximize our power to identify consistent iASD-associated differences across PFC and TC (Supplementary Materials, Fig. S5 and Table S5). We identified 157 DMPs ($P < 5e-05$), with the top-ranked cross-cortex iASD-associated difference (cg14392966, $P = 1.77E-08$) being located in the promoter region of *PUS3* (+6bp) and upstream of *DDX25* (+1136bp) on chromosome 11q24.2. Using data from an RNA-seq analysis of an overlapping set ($n = 40$, 25 iASD cases and 15 controls) of samples (11), we explored the extent to which genes annotated to these DMPs were also differentially expressed in iASD cortex. Of the 111 DMPs annotated to a gene, 20 (18.0%) were annotated to a transcript found to be differentially expressed in iASD cortex (False Discovery Rate (FDR) < 0.1) (Supplementary Material, Table S6), with an overall negative correlation ($r = -0.435$) between DNA methylation and gene expression effects sizes across these probe–gene pairs (Fig. 1D).

Dup15q, a genetically defined subtype of ASD, is associated with striking differences in DNA methylation across an imprinted gene cluster within the duplicated region

The duplication of human chromosome 15q11-13 ('dup15q') is the most frequent cytogenetic abnormality associated with ASD, occurring in ~1% of cases (33,34). In addition to the iASD cases profiled in this study, we quantified DNA methylation in PFC, TC and CB tissue from seven individuals with dup15q; using an established method to identify copy number variation (CNV) from Illumina 450K data (11), we confirmed dup15q status in each of the brain regions profiled in the seven carriers (Supplementary Materials, Figs S6–S8). As with the iASD patients, no global differences in DNA methylation were identified between dup15q carriers and controls in any of the three brain regions (PFC: dup15q carriers = 48.5%, controls = 48.5%; TC: dup15q carriers = 48.4%, controls = 48.4%; CB: dup15q carriers = 46.4%, controls = 46.4%). Again, we observed a strong positive correlation between the estimated 'DNA methylation age' with recorded chronological age for each of the brain regions (Supplementary Material, Fig. S2). Although there was no evidence for differential 'epigenetic aging' in dup15q patients compared to controls in either cortical region (PFC: $P = 0.06$, TC: $P = 0.45$), there was a nominally significant association in CB ($P = 0.012$) with dup15q carriers characterized by decelerated epigenetic aging compared to controls. Overall, these findings indicate that like iASD patients, dup15q carriers are not characterized by systemic differences in DNA methylation across the probes included on the Illumina 450K array in the brain regions tested.

Using a linear model including covariates for sex, age, brain bank and neuronal cell proportions derived from the DNA methylation data (except in the CB as described in the Materials and Methods), we identified numerous DMPs in dup15q carriers in each of the three brain regions (Supplementary Materials, Tables S7–S9 and Figs 9–11). As in our analysis of iASD, we used a multi-level model (see Materials and Methods) to identify consistent dup15q-associated differences across PFC and TC (Fig. 2; Supplementary Material, Table S10). Using *comb-p* (43), we also identified spatially correlated regions of differential DNA methylation associated with dup15q status (Sidak corrected $P < 0.05$) (Supplementary Materials, Tables S11–S14). Our analyses revealed striking *cis*-effects on DNA methylation, with the majority of significant DMPs (Fig. 2C; Supplementary Materials, Figs S6–S8) and DMRs located within a ~7 Mb cluster in the 15q11.1-13.2 duplication region. Despite these strong *cis*-effects, however, sites within the 15q duplicated region are not ubiquitously differentially methylated in CNV carriers. Dup15q-associated DMPs were found to be focused in a specific region within the duplication, with this discrete differentially methylated domain including clusters of probes that are both hyper- and hypo-methylated in carriers (Fig. 2C; Supplementary Materials, Figs S6–S8). Interestingly, these DMPs overlap a genomically imprinted gene cluster within the duplicated region containing transcripts monoallelically expressed from either the paternal (*SNRPN*, *snoRNAs*) or maternal (*UBE3A*, *ATP10A*) alleles. Although DMPs located in the dup15q region are highly consistent across each of the three brain regions, the overall pattern of dup15q-associated variation is more similar between the two cortical regions than between cortex and CB (Supplementary Material, Fig. S12), reflecting the patterns observed for iASD. Interestingly, despite the large effects observed within the dup15q region, a number
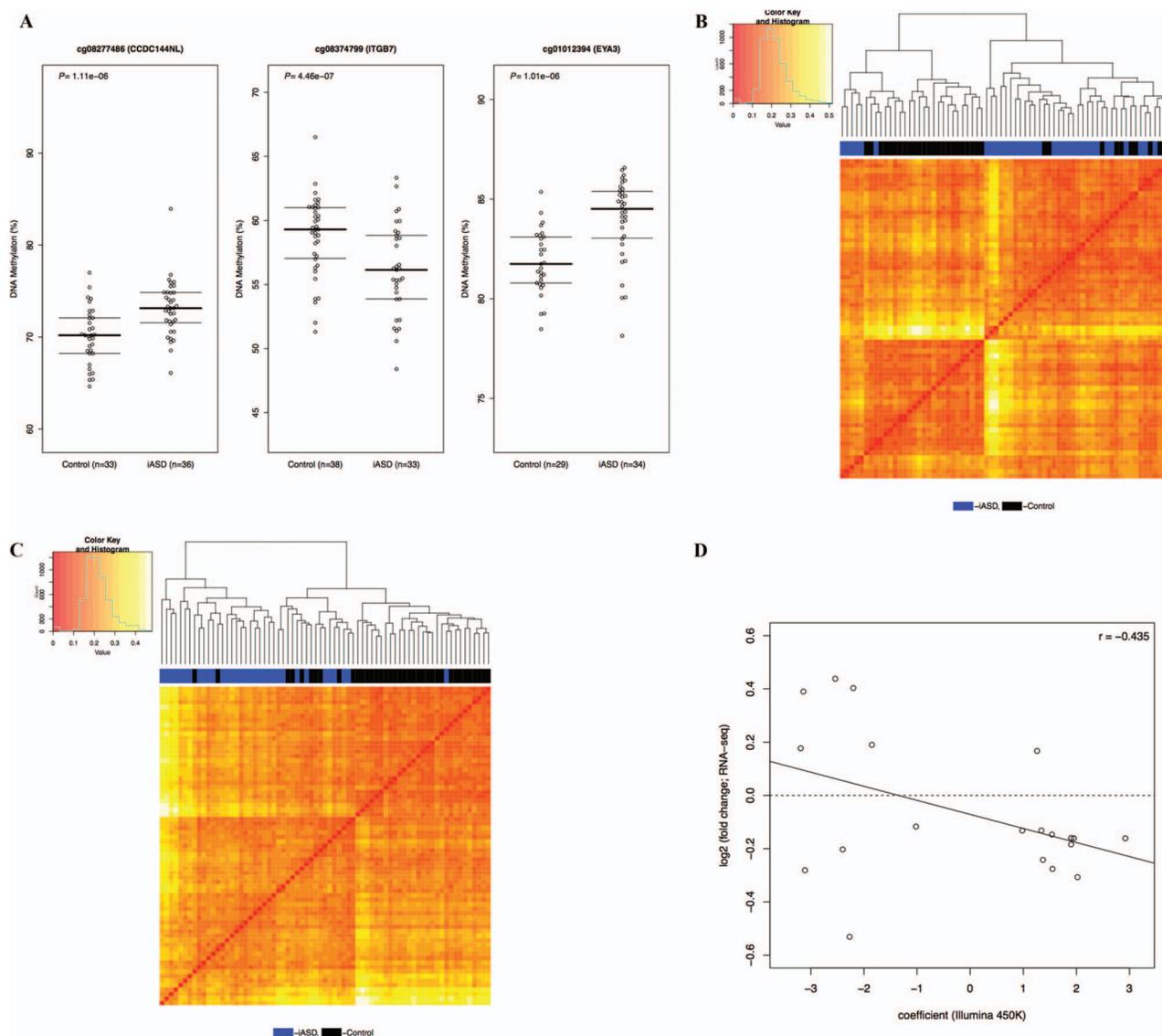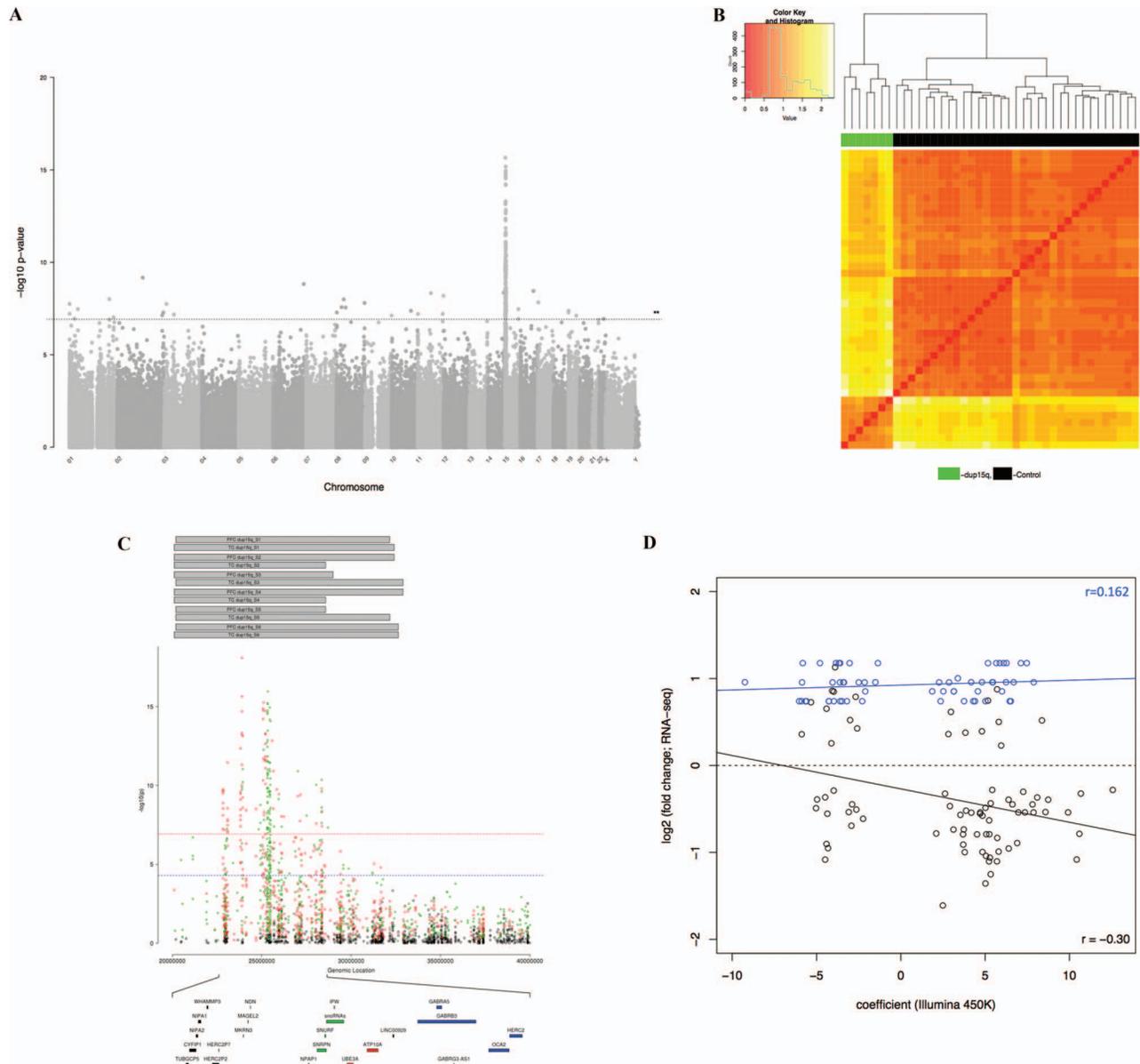
Figure 1. DNA methylation differences at sites associated with idiopathic autism cluster cases and controls and are correlated with gene expression differences. Site-specific changes in DNA methylation associated with idiopathic autism. (**A**) Shown are the top-ranked iASD-associated DMPs in PFC (cg08277486), TC (cg08374799) and CB (cg01012394). A complete list of all iASD-associated DMPs (P < 5e-05) is provided in Supplementary Materials, Tables S2–S4. (**B–C**) Hierarchical clustering of samples based on DNA methylation at iASD-associated DMPs in the cortex. Shown is the clustering of samples based on DNA methylation levels (red = low, yellow = high) at iASD-associated DMPs (P < 5e-05) in (b) PFC and (c) TC. In both cortical regions, the two primary clusters are clearly aligned to disease status (PFC: cluster 1 = 77% iASD, cluster 2 = 20% iASD; TC: cluster 1 = 92% iASD, cluster 2 = 22% iASD). (**D**) Effect sizes at iASD-associated DMPs are negatively correlated with gene expression level. Shown is the relationship between DNA methylation difference (X-axis) and gene expression difference (Y-axis) for genes annotated to cortical DMPs identified in a RNA-seq study performed on an overlapping set of samples.

of DMPs (Supplementary Materials, Tables S7–S10) and DMRs (Supplementary Materials, Tables S11–S14) outside the vicinity of the duplication were also identified in each of the three brain regions, suggesting that structural variation on chromosome 15 may influence regulatory genomic variation at other chromosomal locations in *trans*. Using data from an RNA-seq analysis of an overlapping set (*n* = 21; 6 dup15q cases and 15 controls) of samples (44), we explored the extent to which genes annotated to these DMPs were also differentially expressed in dup15q ASD cortex. Of the 699 DMPs annotated to a gene, 139 (19.9%) were annotated to a transcript found to be differentially expressed in dup15q cortex (FDR < 0.1) (Supplementary Material, Table S15), with an overall negative correlation (*r* = −0.313) between DNA methylation and gene

expression effects sizes across these probe–gene pairs (Fig. 2D). Of note, there was a significant difference in the correlation (Fisher's Z test, P = 0.007) between DNA methylation and gene expression for loci within the dup15q region (*r* = 0.162) compared to those elsewhere in the genome (*r* = −0.300).

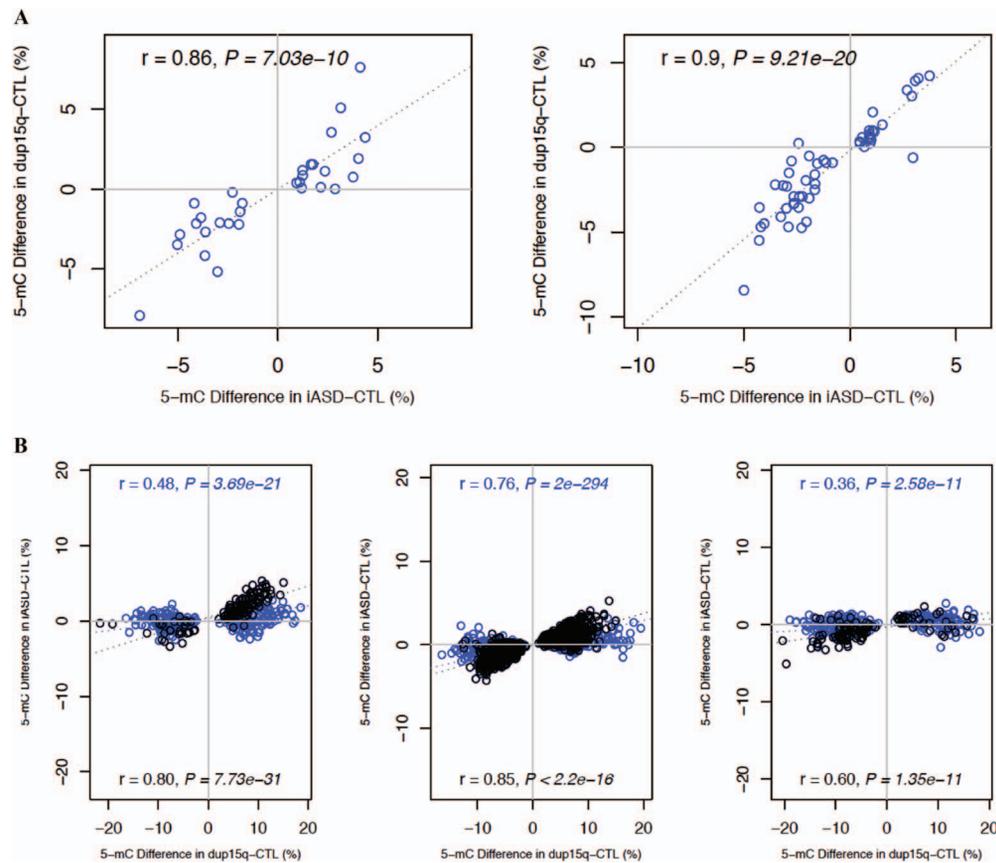## Methylomic differences are shared between iASD and dup15q carriers

Building on a recent analysis of gene expression (11) that revealed a core pattern of cortical transcriptional dysregulation observed in both iASD and dup15q carriers, we next examined the extent to which disease associated DNA methylation differences are shared between these two distinct subgroups

Figure 2. The duplication of chromosome 15q is associated with distinct DNA methylation profiles in all three brain regions. (**A**) Manhattan plot of *P*-values from a multi-level model used to identify consistent dup15q-associated differences across both cortical regions (PFC and TC). The majority of dup15q-associated DMPs are located within the duplicated region. The line highlights a bonferroni significance threshold of $P < 1.198 \times 10^{-7}$. (**B**) Hierarchical clustering of samples based on DNA methylation at dup15q-associated DMPs in the cortex. Shown is the clustering of samples based on PFC DNA methylation levels (red = low, yellow = high) at dup15q-associated DMPs ($P < 5e\text{-}05$). (**C**) Cortical dup15q-associated DMPs are localized within a narrow region within the duplicated region. Shown is the distribution of *P*-values across the dup15q region from our cross-cortex (FC and TC) model. DMPs are stratified by direction of effect (red = hypermethylated in dup15q ASD, green = hypomethylated in dup15q ASD). Shown at the top are the estimated break-points for individual dup15q samples derived from PFC and TC DNA methylation data for each individual donor. The dup15q differentially methylated domain includes clusters of probes that are both hyper- and hypo-methylated overlapping a known imprinted gene cluster containing paternally expressed (green), maternally expressed (red) and biallelically expressed (blue) genes. (**D**) Effect sizes at dup15q-associated DMPs are inversely correlated with gene expression level. Shown is the relationship between DNA methylation difference (X-axis) and gene expression difference (Y-axis) for genes annotated to cortical DMPs identified in a RNA-seq study performed on an overlapping set of samples. The overall correlation between DNA methylation is gene expression was −0.313, although there were striking differences between loci within the dup15q region ($r = 0.162$; coloured in blue) and elsewhere in the genome ($r = −0.300$; coloured in black).

of autism patients. Effect sizes at iADS DMPs ($P < 5 \times 10^{-5}$) were significantly correlated between iASD and dup15q patients (PFC: $r = 0.86$, $P = 7.03e\text{-}10$; TC: $r = 0.9$, $P = 9.21e\text{-}20$; CB: not tested because of the small number of significant DMPs) (Fig. 3A). Likewise, effect sizes at dup15q DMPs ($P < 5e\text{-}05$) were found to be highly correlated between dup15q and iASD

patients (PFC: $r = 0.48$, $P = 3.69e\text{-}21$; TC: $r = 0.76$, $P = 2e\text{-}294$; CB: $r = 0.36$, $P = 2.58e\text{-}11$) (Fig. 3B); of note, although these correlations were particularly strong for probes outside of dup15q region (PFC: $r = 0.80$, $P = 7.73e\text{-}31$; TC: $r = 0.85$, $P < 2.2e\text{-}16$; CB: $r = 0.60$, $P = 1.35e\text{-}11$), significant correlations were also seen for dup15q-associated DMPs located in the duplicated

Figure 3. Consistent DNA methylation differences are seen in iASD and dup15q patients. (**A**) Effect sizes at cortical iASD-associated DMPs are significantly correlated between iASD and dup15q patients. Shown is the correlation in effect sizes for PFC (r = 0.86, P = 7.03e-10, left panel) and TC (r = 0.9, P = 9.21e-20, right panel). (**B**) Effect sizes at dup15q-associated DMPs are highly correlated between dup15q and iASD patients. Shown is the correlation in effect sizes for PFC (across all probes r = 0.48, P = 3.69e-21; probes outside of the dup15q region r = 0.80, P = 7.73e-31; left panel), TC (across all probes r = 0.76, P = 2e-294; probes outside of the dup15q region r = 0.85, P < 2.2e-16; middle panel) and CB (across all probes r = 0.36, P = 2.58e-11; probes outside of the dup15q region r = 0.60, P = 1.35e-11; right panel).

region (PFC: r = 0.22, P = 8.4e-4; TC: r = 0.53, P = 8.59e-18; CB: r = 0.22, P = 0.001). Hierarchical clustering of samples based on DNA methylation values at iASD-associated DMPs ($P < 5 \times 10^{-5}$) shows that dup15q carriers cluster together with iASD cases (Supplementary Materials, Figs S13 and S14), highlighting convergent methylomic signatures associated with both idiopathic and syndromic forms of autism.

Cortical co-methylation modules associated with ASD are enriched for immune, synaptic and neuronal processes

We next used WGCNA (45) to characterize systems-level differences in DNA methylation associated with ASD. We built co-methylation networks using all 'variable' DNA methylation sites (defined as those where the range of DNA methylation values for the middle 80% of individuals was greater than 5%; N = 251 311) using cross-cortex (PFC and TC) data from all donors (see Materials and Methods). WGCNA identified 61 co-methylation modules (Supplementary Material, Table S16), and we used the 'module eigengene' (i.e. the first principal component) for each module to explore differences between controls (n = 29), individuals with iASD (n = 30), individuals with dup15q (n = 6) and a combined ASD group (n = 36). We identified several co-methylation modules robustly asso-

ciated (FDR < 0.05) with at least one diagnostic category (Table 1, Supplementary Material, Fig. S15). We tested whether the genes annotated to probes in each ASD-associated co-methylation module were enriched for specific gene ontology (GO) pathways using a method that groups related pathways to control for the hierarchical structure of the ontological annotations (Table 1 and Supplementary Material, Table S17), identifying a number of pathways relevant to the known aetiology of ASD. For example, the most enriched pathways among genes annotated to probes in the 'skyblue3' module, which was associated with both iASD (FDR = 0.0063) and the combined ASD group (FDR = 0.0033), are related to immune function (e.g. interleukin-1 beta production, P = 8.55E-20), consistent with findings from genetic (46), transcriptomic (47) and epidemiological data (48,49). Among genes annotated to probes in the 'darkorange' module, which was also associated with both iASD (FDR = 0.014) and the combined ASD group (FDR = 0.01), the top-ranked pathways were related to synaptic signalling and regulation, in particular phosphatidylinositol 3-kinase (PI3K) activity (P = 1.95E-08), which plays an important role in synaptic formation and plasticity (50). Interestingly, dysregulation of the PI3K signalling pathway has been associated with neuropsychiatric disorders including schizophrenia (51,52) and autism (53,54). Also of note, pathways enriched among genes annotated to probes in the 'pink' module, which

Table 1. Cortical co-methylation modules associated with ASD

Module	Module size (probes)	Association	GO ID	Top-ranked GO terms	GO term enrichment (P-value)
skyblue3	398	iASD versus CTL *(FDR = 0.0063)* ASD (iASD + dup15q) versus CTL *(FDR = 0.0033)*	GO:0032611 GO:0045785	Interleukin-1 beta production Positive regulation of cell adhesion	8.55E-20 1.82E-19
darkorange	977	iASD versus CTL *(FDR = 0.014)* ASD (iASD + dup15q) versus CTL *(FDR = 0.01)*	GO:0016307 GO:0007156	Phosphatidylinositol phosphate kinase activity Homophilic cell adhesion	1.95E-08 2.49E-08
Tan	2697	iASD versus CTL *(FDR = 0.032)* ASD (iASD + dup15q) versus CTL *(FDR = 0.029)*	GO:0005694 GO:0007156	Chromosome Homophilic cell adhesion	2.40E-11 1.17E-10
honeydew1	123	iASD versus CTL *(FDR = 0.034)* ASD (iASD + dup15q) versus CTL *(FDR = 0.03)*	GO:0051639 GO:0010893	Actin filament network formation Positive regulation of steroid biosynthetic process	1.97E-20 7.41E-15
green	6305	ASD (iASD + dup15q) versus CTL *(FDR = 0.047)*	GO:0043565 GO:0003700	Sequence-specific DNA binding Sequence-specific DNA-binding transcription factor activity	7.24E-38 1.33E-29
blue	19109	ASD (iASD + dup15q) versus CTL *(FDR = 0.025)*	GO:0050919 GO:0016301	Negative chemotaxis Kinase activity	8.20E-06 8.55E-06
pink	4595	dup15q versus CTL *(FDR = 0.00046)*	GO:0097458 GO:0014069	Neuron part Postsynaptic density	3.41E-15 1.87E-10
lightsteelblue1	297	dup15q versus CT *(FDR = 0.012)*	GO:0007156 GO:0005509	Homophilic cell adhesion Calcium ion binding	<2.79E-99 2.79E-99

A full list of GO pathways enriched among genes annotated to the Illumina 450K probes in each module is given in Supplementary Material, Table S17.

was associated with dup15q carriers (FDR = 0.00046) were related to pathways important in neurons (*P* = 3.41E-15) and postsynaptic density (*P* = 1.87E-10). A full list of all significant pathways for each of the co-methylation modules is given in Supplementary Material, Table S17.

Discussion

In this study, we quantified DNA methylation in 223 post-mortem tissue samples isolated from the PFC, TC and CB dissected from 43 donors with ASD and 38 non-psychiatric control subjects. To our knowledge, this represents the most systematic analysis of DNA methylation in ASD using disease-relevant tissue and the first to compare variation identified in idiopathic and syndromic forms of ASD. We report ASD-associated DNA methylation differences at numerous CpG sites with more pronounced effects in both cortical regions compared to the CB. This finding is consistent with previous gene expression studies illustrating that ASD-related molecular changes are substantially smaller in the CB compared to the cortex (11).

Although structural variation on chromosome 15 was found to be associated with striking *cis*-effects on DNA methylation, with a discrete differentially methylated domain spanning an imprinted gene cluster within the duplicated region, variation in DNA methylation associated with autism in the cerebral cortex was highly correlated between iASD and dup15q patients. These results suggest that there are convergent molecular signatures in the cortex associated with different forms of ASD, reinforcing the findings from a recent study of gene expression (11) and H3K27ac (12) undertaken in an overlapping set of samples. Interestingly, a recent study reported partially overlapping peripheral blood

DNA methylation signatures across 14 different Mendelian paediatric neurodevelopmental syndromes (55).

Co-methylation network analyses highlighted systems-level changes in cortical DNA methylation associated with both iASD and dup15q, with associated modules being enriched for sites annotated to genes involved in the immune system, synaptic signalling and neuronal regulation. Our results corroborate findings from other DNA methylation (26,30,31) and gene expression analyses (10,44) that have concluded that ASD-related co-methylation and co-expression modules are significantly enriched for synaptic, neuronal and immune dysfunction genes. Finally, we used existing RNA-seq data on an overlapping set of samples to show a remarkable overlap of cortical ASD-associated DMPs with differential gene expression. Future studies should focus on further understanding the transcriptional consequences of the observed associations and testing whether these associations are causal or a consequence of disease and/or medication.

This study has several strengths. Our epigenome-wide analysis of ASD is, to our knowledge, the largest post-mortem cohort so far and included tissue from three brain regions that have been previously implicated in the pathophysiology of ASD. This contrasts with previous studies that have been undertaken on much smaller numbers of samples and focused on only one or two brain regions. The inclusion of both iASD patients and dup15q carriers in our analyses enabled us to explore evidence for convergent molecular signatures associated with both idiopathic and syndromic forms of autism.

Despite this being the first study to quantify DNA methylation across three different brain regions from both idiopathic and syndromic ASD patients and controls, this study has a number of important limitations that should be considered when interpreting the results. First, DNA methylation was quantified using the

Illumina 450K array; although this is a robust and highly reliable platform with content spanning regulatory regions associated with the majority of known annotated genes, it interrogates DNA methylation at a relatively small proportion of sites across the whole genome. Second, because epigenetic processes play an important role in defining cell-type-specific patterns of gene expression (56,57), the use of bulk tissue from each brain region is a potential confounder in DNA methylation studies (58). Despite our efforts to control for the effect of cell-type diversity in DNA methylation quantification in our analyses using *in silico* approaches, this approach is not suitable to estimate the neuronal proportion in the CB and cannot inform us about disease-relevant DNA methylation changes specific to individual brain cell types. Of note, our general findings are in line with those reported by Nardone and colleagues (26) that interrogated DNA methylation in cell-sorted cortical neurons (from 16 ASD cases and 15 controls). Third, we were limited by the availability of post-mortem brain tissue, and the average age of the cases—especially the seven dup15q carriers—was lower compared to the controls. We, therefore, included 'age' as a covariate in all analyses to minimize potential confounding. Fourth, there is increasing awareness of the importance of 5-hydroxymethyl cytosine (5-hmC) in the human brain (59), although this modification cannot be distinguished from DNA methylation using standard bisulfite-based approaches. It is plausible that many of the ASD-associated differences identified in this study are confounded by modifications other than DNA methylation. To date, no study has evaluated the role of genome-wide 5-hmC in ASD, although recent studies from our group quantified levels of 5-hmC across the genome in human cortex and CB (60) and across neurodevelopment (61).

To conclude, we identified widespread differences in DNA methylation associated with iASD, with consistent signals in both cortical regions that were distinct to those observed in the CB. Individuals carrying a duplication on chromosome 15q (dup15q), representing a genetically defined subtype of ASD, were characterized by striking differences in DNA methylation across a discrete domain spanning an imprinted gene cluster within the duplicated region. In addition to the striking *cis*-effects on DNA methylation observed in dup15q carriers, we identified convergent methylomic signatures associated with both iASD and dup15q, reflecting the findings from previous studies of gene expression and H3K27ac. Our study represents the first systematic analysis of DNA methylation associated with ASD across multiple brain regions, providing novel evidence for convergent molecular signatures associated with both idiopathic and syndromic autism and highlighting potential disease-associated pathways that warrant further investigation.

Materials and Methods

Post-mortem brain tissue from autism cases and controls

Tissue samples for this study were acquired from the Autism Tissue Program brain bank at the Oxford UK Brain Bank for Autism (www.brainbankforautism.org.uk), Harvard Brain and Tissue Bank (https://hbtrc.mclean.harvard.edu) and the National Institute for Child Health and Human Development Eunice Kennedy Shriver Brain and Tissue Bank for Developmental Disorders. (http://www.medschool.umaryland.edu/btbank/). All subjects were de-identified prior to acquisition, and all samples were dissected by trained neuropathologists, snap-frozen and

stored at −80°C. A total of 223 samples obtained from 95 individuals were included in this study with up to three brain regions from each individual donor: dorsolateral or medial PFC [corresponding to Brodmann area 9 and denoted as 'prefrontal cortex' (PFC)], superior temporal gyrus [corresponding to BA41, BA42 or BA22 and denoted as 'temporal cortex' (TC)] and cerebellar vermis (CB) (Supplementary Material, Fig. S1). Further information about the samples is given in Supplementary Material, Table S1. Genomic DNA was isolated from ∼100 mg of each dissected brain region using a standard phenol–chloroform extraction method and tested for degradation and purity before analysis.

DNA methylation profiling and data QC

All samples were randomized with respect to phenotypic status, age, sex and brain bank 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™ to minimize potential bisulfite-related biases (Zymo Research, Irvine, CA, USA). Genome-wide DNA methylation was quantified using the pooled bisulfite-converted DNA samples and the Illumina Infinium HumanMethylation450 BeadChip ('450K array') (Illumina, San Diego, CA, USA), scanned on an Illumina HiScan System. Illumina GenomeStudio software was used to extract signal intensities for each probe, generating a final report that was imported in to the R statistical environment 3.0.2 (www.r-project.org) (62) using the *methylumi* (63) package. Data QC and pre-processing were performed using the *wateRmelon* package as described previously (64). Multidimensional scaling plots of sex chromosome probes were used to check that the predicted sex corresponded with the reported sex for each individual, and comparison of 65 single nucleotide polymorphism (SNP) probes on the array confirmed that matched tissues were sourced from the same individual. Stringent filtering of the pre-normalized Illumina 450K data was performed. First, we removed the 65 SNP probes, cross-reactive probes and probes overlapping polymorphic CpGs containing an SNP with minor allele frequency of >5% within 10 bp of the single base extension position as detailed in the Illumina annotation file and identified in recent publications (65,66). Second, CpG sites with a detection P-value of >0.05 in 1% of samples identified by the *pfilter* function within the *wateRmelon* R package were removed. Third, polymorphic SNP control probes ($n = 65$) located on the array were used to confirm that matched cortex and CB tissues were sourced from the same individual. The final data set consists of a total of 417 460 probes from 76 PFC samples ($n = 36$ iASD patients, $n = 7$ dup15q patients, $n = 33$ controls), 77 TC samples ($n = 33$ iASD patients, $n = 6$ dup15q patients, $n = 38$ controls) and 70 CB samples ($n = 34$ iASD patients, $n = 7$ dup15q patients, $n = 29$ controls) (Supplementary Material, Table S1) was normalized with the *dasen* function of the *wateRmelon* R package and then batch-corrected with the *ComBat* function of the *ComBat* R package (67). For the dup15q samples, their duplication status and breakpoints were confirmed by genotyping (11) and further validated by CNV calling from the 450K DNA methylation data using R package ChAMP (68).

Identification of autism-associated differential methylation

All statistical analyses were conducted using R statistical package (version 3.1.1). Analyses were performed to test for DMPs and

DMRs associated with disease status for each brain region. The R package *Cell EpigenoType Specific* (CETS) mapper (69) designed for the quantification and normalization of differing neuronal proportions in genome-wide DNA methylation data sets was used to estimate brain cellular heterogeneity in cortex (both PFC and TC) samples. CETS-based neuronal cell composition estimate was not applied on the CB samples given the known high proportion of non-NeuN-expressing neurons in this brain region (41,69). To model the effect of sample-specific variables, we performed linear regression for each probe using age, gender, brain bank, CETS (for PFC and TC but not CB) and diagnosis as independent variables. We also applied a linear mixed effect (LME) model framework for samples with data from multiple cortical regions to identify consistent differential ASD-associated DNA methylation markers across the cortex. The individual donor identified was treated as a random effect, and age, gender, brain bank, CETS, brain region and diagnosis were treated as fixed effects. All disease-associated DMPs identified and reported have to pass the 'discovery' threshold of $P < 5e\text{-}05$ and/or the stringent experiment-wide significance threshold of $P < 1.198 \times 10^{-7}$. ASD-associated DMRs were identified using the *Python* module *comb-p* (43) to group spatially correlated DMPs (seed $P < 1.00\text{E-}03$, minimum of two probes) at a maximum distance of 300 bp in each analytical group. Significant DMRs were identified as those with at least two probes and a corrected $P < 0.05$ using Sidak correction (70).

Weighted gene correlation network analysis

For the co-methylation network analyses, modules of co-methylated probes were identified using the WGCNA package in R (45), and a network was constructed for cross-cortex samples. For each network, the input probe set was pruned to remove probes with minimal variability across samples. We required that each probe have a minimum range of methylation values of 5% within the middle 80% of samples, resulting in 251 311 probes for cortex. In the cross-cortex network, to enrich for modules relating to diagnosis, we regressed out the effect of CET score and age after fitting a linear model for each probe with diagnosis, age, sex, cortical region, CET score and brain bank as independent variables. A signed network was constructed using the biweight mid-correlations between probes with a soft power of 7. This network was generated blockwise, using the WGCNA function blockwiseModules with the following parameters: maxBlockSize = 15 000, mergeCutHeight = 0.1, deepSplit = 4 and minModuleSize = 100. For each module, a linear mixed effects model was fit between its eigengene and diagnosis, age, sex, cortical region, batch, CET score and brain bank as fixed effects with individual ID as a random effect. This LME model was fit for three subsets of samples: (iASD + dup15q) versus control, iASD versus control and dup15q versus control.

Gene ontology pathway analysis

Illumina UCSC gene annotation, which is derived from the genomic overlap of probes with RefSeq genes or up to 1500 bp of the transcription start site of a gene, was used to create a test gene list from the probes identified in the disease-associated modules for pathway analysis. Where probes were not annotated to any gene (i.e. in the case of intergenic locations), they were omitted from this analysis; where probes were annotated to multiple genes, all were included. A logistic regression approach was used to test if genes in this list predicted pathway mem-

bership, while controlling for the number of probes that passed QC (i.e. were tested) annotated to each gene. Pathways were downloaded from the GO website (http://geneontology.org/) and mapped to genes, including all parent ontology terms. All genes with at least one 450 K probe annotated and mapped to at least one GO pathway were considered. Pathways were filtered to those containing between 10 and 2000 genes. After applying this method to all pathways, the list of significant pathways ($P < 0.05$) was refined by grouping to control for the effect of overlapping genes. This was achieved by taking the most significant pathway and retesting all remaining significant pathways while controlling additionally for the best term. If the test genes no longer predicted the pathway, the term was said to be explained by the more significant pathway, and hence these pathways were grouped together. This algorithm was repeated, taking the next most significant term, until all pathways were considered as the most significant or found to be explained by a more significant term.

Supplementary Material

Supplementary Material is available at *HMG* online.

References

1. Elsabbagh, M., Divan, G., Koh, Y.J., Kim, Y.S., Kauchali, S., Marcin, C., Montiel-Nava, C., Patel, V., Paula, C.S., Wang, C. et al. (2012) Global prevalence of autism and other pervasive developmental disorders. *Autism Res.*, **5**, 160–179.
2. Kim, Y.S., Leventhal, B.L., Koh, Y.J., Fombonne, E., Laska, E., Lim, E.C., Cheon, K.A., Kim, S.J., Kim, Y.K., Lee, H. et al. (2011) Prevalence of autism spectrum disorders in a total population sample. *Am. J. Psychiatry*, **168**, 904–912.
3. Gliga, T., Jones, E.J., Bedford, R., Charman, T. and Johnson, M.H. (2014) From early markers to neuro-developmental mechanisms of autism. *Developmental Rev.*, **34**, 189–207.
4. Parikshak, N.N., Gandal, M.J. and Geschwind, D.H. (2015) Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat. Rev. Genet.*, **16**, 441.
5. Gaugler, T., Klei, L., Sanders, S.J., Bodea, C.A., Goldberg, A.P., Lee, A.B., Mahajan, M., Manaa, D., Pawitan, Y. and Reichert, J. (2014) Most genetic risk for autism resides with common variation. *Nat. Genet.*, **46**, 881.

6. Abrahams, B.S. and Geschwind, D.H. (2008) Advances in autism genetics: on the threshold of a new neurobiology. *Nat. Rev. Genet.*, **9**, 341–355.

7. Berg, J.M. and Geschwind, D.H. (2012) Autism genetics: searching for specificity and convergence. *Genome Biol.*, **13**, 247.

8. de la Torre-Ubieta, L., Won, H., Stein, J.L. and Geschwind, D.H. (2016) Advancing the understanding of autism disease mechanisms through genetics. *Nat. Med.*, **22**, 345.

9. Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R.K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O.A. and Anney, R. (2019) Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.*, **41**, 431–444.

10. Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R., Blencowe, B.J. and Geschwind, D.H. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, **474**, 380.

11. Parikshak, N.N., Swarup, V., Belgard, T.G., Irimia, M., Ramaswami, G., Gandal, M.J., Hartl, C., Leppa, V., de la Torre Ubieta, L. and Huang, J. (2016) Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature*, **540**, 423–427.

12. Sun, W., Poschmann, J., del Rosario, R.C.-H., Parikshak, N.N., Hajan, H.S., Kumar, V., Ramasamy, R., Belgard, T.G., Elanggovan, B. and Wong, C.C.Y. (2016) Histone acetylome-wide association study of autism spectrum disorder. *Cell*, **167**, 1385–1397. e1311.

13. Ciernia, A.V. and LaSalle, J. (2016) The landscape of DNA methylation amid a perfect storm of autism aetiologies. *Nat. Rev. Neurosci.*, **17**, 411.

14. Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., Li, E., Zhang, Y. and Sun, Y.E. (2010) Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science*, **329**, 444–448.

15. Yao, B., Christian, K.M., He, C., Jin, P., Ming, G.-L. and Song, H. (2016) Epigenetic mechanisms in neurogenesis. *Nat. Rev. Neurosci.*, **17**, 537.

16. Tognini, P., Napoli, D. and Pizzorusso, T. (2015) Dynamic DNA methylation in the brain: a new epigenetic mark for experience-dependent plasticity. *Front. Cell. Neurosci.*, **9**, 331.

17. Heyward, F.D. and Sweatt, J.D. (2015) DNA methylation in memory formation: emerging insights. *Neuroscientist*, **21**, 475–489.

18. Zovkic, I.B., Guzman-Karlsson, M.C. and Sweatt, J.D. (2013) Epigenetic regulation of memory formation and maintenance. *Learn. Mem.*, **20**, 61–74.

19. Spiers, H., Hannon, E., Schalkwyk, L.C., Smith, R., Wong, C.C., O'Donovan, M.C., Bray, N.J. and Mill, J. (2015) Methylomic trajectories across human fetal brain development. *Genome Res.*, **25**, 338–352.

20. Persico, A.M. and Bourgeron, T. (2006) Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends Neurosci.*, **29**, 349–358.

21. Grafodatskaya, D., Chung, B., Szatmari, P. and Weksberg, R. (2010) Autism spectrum disorders and epigenetics. *J. Am. Acad. Child Adolesc. Psychiatry*, **49**, 794–809.

22. Loke, Y.J., Hannan, A.J. and Craig, J.M. (2015) The role of epigenetic change in autism spectrum disorders. *Front. Neurol.*, **6**, 107. doi: 10.3389/fneur.2015.00107.

23. Ladd-Acosta, C., Hansen, K.D., Briem, E., Fallin, M.D., Kaufmann, W.E. and Feinberg, A.P. (2014) Common DNA methylation alterations in multiple brain regions in autism. *Mol. Psychiatry*, **19**, 862–871.

24. Ginsberg, M.R., Rubin, R.A., Falcone, T., Ting, A.H. and Natowicz, M.R. (2012) Brain transcriptional and epigenetic associations with autism. *PLoS One*, **7**, e44736. doi: 10.1371/journal.pone.0044736.

25. Nguyen, A., Rauch, T.A., Pfeifer, G.P. and Hu, V.W. (2010) Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. *FASEB J*, **24**, 3036–3051.

26. Nardone, S., Sams, D.S., Zito, A., Reuveni, E. and Elliott, E. (2017) Dysregulation of cortical neuron DNA methylation profile in autism spectrum disorder. *Cerebral Cortex*, **12**, 5739–5754.

27. Nardone, S., Sams, D.S., Reuveni, E., Getselter, D., Oron, O., Karpuj, M. and Elliott, E. (2014) DNA methylation analysis of the autistic brain reveals multiple dysregulated biological pathways. *Transl. Psychiatry*, **4**, e433. doi: 10.1038/tp.2014.70.

28. Wong, C.C., Meaburn, E.L., Ronald, A., Price, T.S., Jeffries, A.R., Schalkwyk, L.C., Plomin, R. and Mill, J. (2014) Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Mol. Psychiatry*, **19**, 495–503.

29. Hannon, E., Schendel, D., Ladd-Acosta, C., Grove, J., Hansen, C.S., Andrews, S.V., Hougaard, D.M., Bresnahan, M., Mors, O. and Hollegaard, M.V. (2018) Elevated polygenic burden for autism is associated with differential DNA methylation at birth. *Genome Med.*, **10**, 19.

30. Dunaway, K.W., Islam, M.S., Coulson, R.L., Lopez, S.J., Ciernia, A.V., Chu, R.G., Yasui, D.H., Pessah, I.N., Lott, P. and Mordaunt, C. (2016) Cumulative impact of polychlorinated biphenyl and large chromosomal duplications on DNA methylation, chromatin, and expression of autism candidate genes. *Cell Rep.*, **17**, 3035–3048.

31. Vogel Ciernia, A., Laufer, B.I., Dunaway, K.W., Mordaunt, C.E., Coulson, R.L., Yasui, D.H. and LaSalle, J.M. (2018) Epigenomic convergence of genetic and immune risk factors in autism brain. *bioRxiv*, 270827. doi: https://doi.org/10.1101/270827.

32. Schübeler, D. (2015) Function and information content of DNA methylation. *Nature*, **517**, 321.

33. Schroer, R.J., Phelan, M.C., Michaelis, R.C., Crawford, E.C., Skinner, S.A., Cuccaro, M., Simensen, R.J., Bishop, J., Skinner, C. and Fender, D. (1998) Autism and maternally derived aberrations of chromosome 15q. *Am. J. Med. Genet.*, **76**, 327–336.

34. Depienne, C., Moreno-De-Luca, D., Heron, D., Bouteiller, D., Gennetier, A., Delorme, R., Chaste, P., Siffroi, J.-P., Chantot-Bastaraud, S. and Benyahia, B. (2009) Screening for genomic rearrangements and methylation abnormalities of the 15q11-q13 region in autism spectrum disorders. *Biol. Psychiatry*, **66**, 349–359.

35. Tsai, P.T., Hull, C., Chu, Y., Greene-Colozzi, E., Sadowski, A.R., Leech, J.M., Steinberg, J., Crawley, J.N., Regehr, W.G. and Sahin, M. (2012) Autistic-like behavior and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. *Nature*, **488**, 647.

36. Hadjikhani, N., Joseph, R.M., Snyder, J. and Tager-Flusberg, H. (2005) Anatomical differences in the mirror neuron system and social cognition network in autism. *Cereb. Cortex*, **16**, 1276–1282.

37. Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P. and Morgan, J. (2007) Mapping early brain development in autism. *Neuron*, **56**, 399–413.

38. Horvath, S. (2013) DNA methylation age of human tissues and cell types. *Genome Biol.*, **14**, 3156.

39. Horvath, S. (2015) Erratum to: DNA methylation age of human tissues and cell types. *Genome Biol.*, **16**, 96.

40. Lunnon, K., Smith, R., Hannon, E., De Jager, P.L., Srivastava, G., Volta, M., Troakes, C., Al-Sarraj, S., Burrage, J. and Macdonald, R. (2014) Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nature Neurosci.*, **17**, 1164–1170.

41. Viana, J., Hannon, E., Dempster, E., Pidsley, R., Macdonald, R., Knox, O., Spiers, H., Troakes, C., Al-Saraj, S. and Turecki, G. (2016) Schizophrenia-associated methylomic variation: molecular signatures of disease and polygenic risk burden across multiple brain regions. *Hum. Mol. Genet.*, **26**, 210–225.

42. Davies, M.N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R.A., Milosavljevic, A. and Troakes, C. (2012) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol.*, **13**, R43.

43. Pedersen, B.S., Schwartz, D.A., Yang, I.V. and Kechris, K.J. (2012) Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics*, **28**, 2986–2988.

44. Parikshak, N.N., Swarup, V., Belgard, T.G., Irimia, M., Ramaswami, G., Gandal, M.J., Hartl, C., Leppa, V., de la Torre Ubieta, L. and Huang, J. (2016) Genome-wide changes in lncRNA, alternative splicing, and cortical patterning in autism. bioRxiv (in press), 077057.

45. Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

46. Nazeen, S., Palmer, N.P., Berger, B. and Kohane, I.S. (2016) Integrative analysis of genetic data sets reveals a shared innate immune component in autism spectrum disorder and its co-morbidities. *Genome Biol.*, **17**, 228.

47. Gupta, S., Ellis, S.E., Ashar, F.N., Moes, A., Bader, J.S., Zhan, J., West, A.B. and Arking, D.E. (2014) Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nature Commun.*, **5**, 5748. doi: 10.1038/ncomms6748.

48. Estes, M.L. and McAllister, A.K. (2015) Immune mediators in the brain and peripheral tissues in autism spectrum disorder. *Nat. Rev. Neurosci.*, **16**, 469.

49. Needleman, L.A. and McAllister, A.K. (2012) The major histocompatibility complex and autism spectrum disorder. *Dev. Neurobiol.*, **72**, 1288–1301.

50. Brunet, A., Datta, S.R. and Greenberg, M.E. (2001) Transcription-dependent and -independent control of neuronal survival by the PI3K–Akt signaling pathway. *Curr. Opin. Neurobiol.*, **11**, 297–305.

51. Kalkman, H.O. (2012) A review of the evidence for the canonical Wnt pathway in autism spectrum disorders. *Mol. Autism*, **3**, 10.

52. Law, A.J., Wang, Y., Sei, Y., O'Donnell, P., Piantadosi, P., Papaleo, F., Straub, R.E., Huang, W., Thomas, C.J. and Vakkalanka, R. (2012) Neuregulin 1-ErbB4-PI3K signaling in schizophrenia and phosphoinositide 3-kinase-p110δ inhibition as a potential therapeutic strategy. *Proc. Natl. Acad. Sci. USA.*, **109**, 12165–12170.

53. Enriquez-Barreto, L. and Morales, M. (2016) The PI3K signaling pathway as a pharmacological target in Autism related disorders and Schizophrenia. *Mol. Cell. Ther.*, **4**, 2. doi: 10.1186/s40591-40016-40047-40599.

54. Bourgeron, T. (2009) A synaptic trek to autism. *Curr. Opin. Neurobiol.*, **19**, 231–234.

55. Aref-Eshghi, E., Rodenhiser, D.I., Schenkel, L.C., Lin, H., Skinner, C., Ainsworth, P., Paré, G., Hood, R.L., Bulman, D.E. and Kernohan, K.D. (2018) Genomic DNA methylation signatures enable concurrent diagnosis and clinical genetic variant classification in neurodevelopmental syndromes. *Am. J. Hum. Genet.*, **102**, 156–174.

56. Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. and Ziller, M.J. (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317.

57. Akbarian, S., Liu, C., Knowles, J.A., Vaccarino, F.M., Farnham, P.J., Crawford, G.E., Jaffe, A.E., Pinto, D., Dracheva, S. and Geschwind, D.H. (2015) The psychencode project. *Nat. Neurosci.*, **18**, 1707.

58. Mill, J. and Heijmans, B.T. (2013) From promises to practical strategies in epigenetic epidemiology. *Nat. Rev. Genet.*, **14**, 585.

59. Branco, M.R., Ficz, G. and Reik, W. (2012) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.*, **13**, 7.

60. Lunnon, K., Hannon, E., Smith, R.G., Dempster, E., Wong, C., Burrage, J., Troakes, C., Al-Sarraj, S., Kepa, A. and Schalkwyk, L. (2016) Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol.*, **17**, 27.

61. Spiers, H., Hannon, E., Schalkwyk, L.C., Bray, N.J. and Mill, J. (2017) 5-hydroxymethylcytosine is highly dynamic across human fetal brain development. *BMC Genomics*, **18**, 738.

62. Team, R.D.C. (2010) R Foundation for Statistical Computing, Vienna, Austria, Vol. Retrieved from http://www.R-project.org.

63. Davis, S., Du, P., Bilke, S., Triche, T., & Bootwalla, M. (2018) Methylumi: handle Illumina methylation data. In R package, Vol. version 2.28.0.

64. Pidsley, R., Wong, C.C.Y., Volta, M., Lunnon, K., Mill, J. and Schalkwyk, L.C. (2013) A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, **14**, 293.

65. Chen, Y.A., Lemire, M., Choufani, S., Butcher, D.T., Grafodatskaya, D., Zanke, B.W., Gallinger, S., Hudson, T.J. and Weksberg, R. (2013) Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium Human Methylation 450 microarray. *Epigenetics*, **8**, 203–209.

66. Price, M.E., Cotton, A.M., Lam, L.L., Farre, P., Emberly, E., Brown, C.J., Robinson, W.P. and Kobor, M.S. (2013) Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin*, **6**, 4.

67. Johnson, W.E., Li, C. and Rabinovic, A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.

68. Morris, T.J., Butcher, L.M., Feber, A., Teschendorff, A.E., Chakravarthy, A.R., Wojdacz, T.K. and Beck, S. (2014) ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics*, **1**, 428–430.

69. Guintivano, J., Aryee, M.J. and Kaminsky, Z.A. (2013) A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*, **8**, 290–302.

70. Šidák, Z. (1967) Rectangular confidence regions for the means of multivariate normal distributions. *J. Am. Stat. Assoc.*, **62**, 626–633.