Molecular Psychiatry (2012) 17, 728–740

www.nature.com/mp

ORIGINAL ARTICLE

A multi-tissue analysis identifies HLA complex group 9 gene methylation differences in bipolar disorder

Z Kaminsky[1,2,8], M Tochigi[1,3,8], P Jia[1,8], M Pal[1], J Mill[1,4], A Kwan[1], I Ioshikhes[5], JB Vincent[6], JL Kennedy[6], J Strauss[6], S Pai[1], S-C Wang[1,7] and A Petronis[1]

[1]*The Krembil Family Epigenetics Laboratory, Neuroscience Department, Centre for Addiction and Mental Health, Toronto, ON, Canada;* [2]*Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, USA;* [3]*Department of Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan;* [4]*King's College London, Social, Genetic and Developmental Psychiatry Department, Institute of Psychiatry, London, UK;* [5]*Ottawa Institute of Systems Biology (OISB) and Department of Biochemistry, Microbiology and Immunology (BMI), University of Ottawa, Ottawa, ON, Canada;* [6]*Neuroscience Department, Centre for Addiction and Mental Health, Toronto, ON, Canada and* [7]*Systems Biology and Bioinformatics Institute, National Central University, Jhongli, Taiwan*

Epigenetic studies of DNA and histone modifications represent a new and important activity in molecular investigations of human disease. Our previous epigenome-wide scan identified numerous DNA methylation differences in *post-mortem* brain samples from individuals affected with major psychosis. In this article, we present the results of fine mapping DNA methylation differences at the human leukocyte antigen (HLA) complex group 9 gene (*HCG9*) in bipolar disorder (BPD). Sodium bisulfite conversion coupled with pyrosequencing was used to interrogate 28 CpGs spanning ~700 bp region of *HCG9* in 1402 DNA samples from *post-mortem* brains, peripheral blood cells and germline (sperm) of bipolar disease patients and controls. The analysis of nearly 40 000 CpGs revealed complex relationships between DNA methylation and age, medication as well as DNA sequence variation (rs1128306). Two brain tissue cohorts exhibited lower DNA methylation in bipolar disease patients compared with controls at an extended *HCG9* region ($P = 0.026$). Logistic regression modeling of BPD as a function of rs1128306 genotype, age and DNA methylation uncovered an independent effect of DNA methylation in white blood cells (odds ratio (OR) = 1.08, $P = 0.0077$) and the overall sample (OR = 1.24, $P = 0.0011$). Receiver operating characteristic curve A prime statistics estimated a 69–72% probability of correct BPD prediction from a case vs control pool. Finally, sperm DNA demonstrated a significant association ($P = 0.018$) with BPD at one of the regions demonstrating epigenetic changes in the *post-mortem* brain and peripheral blood samples. The consistent multi-tissue epigenetic differences at *HCG9* argue for a causal association with BPD.
Molecular Psychiatry (2012) **17,** 728–740; doi:10.1038/mp.2011.64; published online 7 June 2011

Keywords: epigenetics; DNA methylation; pyrosequencing; bipolar disorder

Introduction

Complex non-Mendelian diseases exhibit numerous features that are difficult to explain by traditional mechanisms such as predisposing DNA sequence variation and hazardous environmental factors. The non-Mendelian irregularities include the discordance of monozygotic twins, parent of origin effects, gender dimorphism and relatively late age at onset.[1–3] Putative epigenetic misregulation that manifests via changes in modifications of DNA and histones can offer a more comprehensive and cohesive explanation for the various facets of complex disease.[4] Epigenetic factors such as DNA methylation control various genetic and genomic functions, including gene transcription.[5,6] DNA methylation is known to act via three mechanisms: critical site, general density of methylated sites in the DNA and multiple elements, which combines components of the first two models.[7] Critical site effects relate to the limitation of transcription factor binding affinity due to the methylation status of specific CpGs within their associated binding sites.[8,9] The overall density of methylcytosine in a gene regulatory region also contributes to gene activity with a large number of genes exhibiting an inverse correlation between the degree of methylation and the level of gene expression.[5,6] Epigenetic factors are relatively stably inherited across somatic cells; however, stochastic events and environmental influences can result

Correspondence: Dr A Petronis, The Krembil Family Epigenetics Laboratory, Neuroscience Department, Centre for Addiction and Mental Health, Toronto, ON, Canada M5T 1R8.
E-mail: arturas_petronis@camh.net
[8]Joint first authorship.
Received 23 January 2011; revised 17 March 2011; accepted 20 April 2011; published online 7 June 2011

in a drift of these regulatory molecular signals over time.[10,11] There is increasing evidence that some epigenetic signals can be transmitted from one generation to another,[12] which contradicts the idea that all epigenetic signals are wiped out during reprogramming in germ cells and points to a secondary molecular mechanism of heritability.

Over the last decade, it has been suggested that epigenetic misregulation is consistent with many epidemiological, clinical and molecular findings in bipolar disorder (BPD).[1,13,14] Partial epigenetic instability, or metastability, can account for both the discordance of monozygotic twins and clinical fluctuations over the course of BPD. The opposite disease phenotypes, depressive and manic phases, between which the patients can make a nearly complete recovery, are consistent with dynamic changes in epigenetics rather than alternations to the stable DNA sequence. Sex differences exist in BPD, such as the higher incidence of rapid cycling, mixed states and cyclothymia in women and a higher prevalence of early-onset BPD in men. These differences, as well as sex effects revealed in genetic linkage and association studies of BPD, could be rationalized by the known epigenetic effects of sex hormones.[15] Parent of origin-dependent clinical and molecular differences in BPD are likely to result from genomic imprinting,[16] a differential epigenetic modification of genes depending upon their parental origin.[17] Epigenetic misregulation of gene activity is also consistent with the numerous neurochemical differences detected in BPD such as altered densities/levels, of receptors, enzymes, peptides and neuromediators.[18] Increasing evidence for 'soft' inheritance (that is, partial epigenetic stability in meiosis)[12] offers a non-DNA sequence-based explanation for heritability of BPD. It is telling that valproic acid, a mood stabilizer and one of the most common medications to treat BPD, is a histone deacetylase inhibitor[19] and can also change DNA methylation.[20] Recent studies detected that lithium, another mood stabilizer, modulates histone acetylation and DNA methylation.[21,22]

BPD may result from a series of unfavorable epigenetic events beginning with a primary epigenetic defect, or pre-epimutation that hypothetically occurs in the germline during epigenetic reprogramming[23–25] and is inherited. Pre-epimutations are further altered during embryogenesis, childhood and adolescence by the multiple effects of tissue differentiation, stochastic factors, hormones and perhaps, external environmental factors.[23,26,27] Decades may elapse before epigenetic misregulation reaches the level of an epimutation, a defect of sufficient severity that the resultant cellular effects manifest clinically as BPD. The severity of an epimutation may fluctuate, as evidenced by clinical remission and relapse.

Recently, we completed the first DNA methylome study in major psychosis using the 12K CpG island microarray.[34] The study revealed a number of DNA methylation differences associated with BPD and schizophrenia (SCZ). This article reports data from our continuing effort to validate and finely map the identified epigenetic differences in the microarray project. One of the genes identified as differentially methylated in major psychosis was the human leukocyte antigen (HLA) complex group 9 gene (*HCG9*; Figure 1). In the microarray experiments, the human CpG island microarray probe UHNhscpg0007403 showed significantly lower DNA methylation in major psychosis patients relative to controls (fold change −0.42, false discovery rate (FDR) $P=0.03$). *HCG9* is located on chromosome 6p21.33 and is of particular interest as recent genome-wide association studies (GWAS) report an association between genetic variants at chromosome 6p21.3-22 with SCZ and BPD.[28–30] *HCG9* is a member of the PERB11 gene family, also known as MIC, many variants of which (such as MICA and MICB) function as ligands for natural killer cell receptors.[31] Three splice variants for *HCG9* have been identified so far (Figure 1). While it has been suggested based on experimental evidence that *HCG9* may be related to stress response and cell death,[32] its function is currently unknown.

The goals of our analysis of the *HCG9* gene were fivefold. Our primary objective was to validate the microarray findings in the Stanley Medical Research Institute (SMRI) *post-mortem* brain cohort, the sample which was used in our 12K CpG island microarray-based DNA methylome scan. Second, we attempted to replicate our primary findings in an independent *post-mortem* brain sample. In the event of validation, our third objective was to characterize *HCG9* methylation differences in the context of the available demographic and clinical information (for example, age-dependent dynamics of DNA methylation). Fourth, we were interested if pathological *HCG9* methylation differences can be detected in peripheral white blood cells (WBCs) and germline, tissues that are not directly involved in psychiatric disease but that are easily accessible and can be used in clinical applications. Finally, our fifth objective was to address the unknown function of the *HCG9* gene in order to better understand how it may fit into the etiopathogenesis of BPD and other psychiatric diseases.

Cytosine modifications were finely mapped using the sodium bisulfite conversion approach coupled with pyrosequencing. In total, we interrogated 1402 DNA samples across 28 CpG positions of *HCG9*.

Materials and methods

Samples

Post-mortem brain tissue of prefrontal, parietal and occipital cortices and corpus callosum from individuals with BPD ($N=34$), SCZ ($N=35$) and matched controls (Con) ($N=35$) was provided unblinded by the SMRI brain-array collection. A brain prefrontal cortex replication cohort from the McLean (McL) Hospital at the National Brain Databank ($N=34$ BPD, 30 SCZ and 50 Con) was tested along with a peripheral white

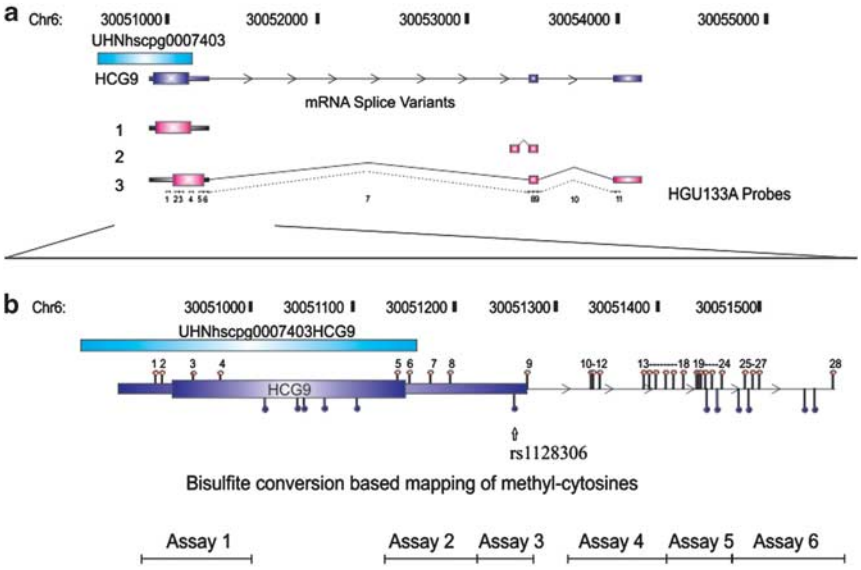

Figure 1 *HCG9* investigated region. A diagram of the *HCG9* gene located on Chr 6p21.33 (**a**) and magnified region (**b**) subjected to fine mapping of methylated cytosines. Representations are to scale based on the human genome build (NCBI 36) *hg18*. The position of the University Health Network (UHN) microarray probe (UHNhscpg0007403) that revealed DNA methylation differences in major psychiatric disease in our microarray analysis is shown in light blue. Reference sequence genes are depicted in dark blue while mRNA splice variants 1–3 for *HCG9* as determined by AceView are in pink. Relative gene expression probes from the *Affymetrix* HGU133A microarray relating to this gene are depicted in (**a**). Assays 1–6 in (**b**) represent the regions subjected to sodium bisulfite modification coupled with pyrosequencing. CpG dinucleotides subjected to DNA methylation analysis are represented by numbered red lollipops while un-interrogated CpGs and those omitted due to overlap with a database of single nucleotide polymorphisms (dbSNP) citation are shown as upside-down blue lollipops.

blood cell (WBC 1) sample provided by Dr James Kennedy ($N = 100$ BPD, 50 SCZ and 100 Con), a second WBC sample (WBC 2) from GlaxoSmithKline, London, UK ($N = 270$ BPD and 282 Con), and sperm samples collected by our laboratory at the Centre for Addiction and Mental Health ($N = 29$ BPD, 14 SCZ and 30 Con). Diagnoses had been made according to the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition). A summary of the demographic information of samples used in this study is shown in Table 1. Detailed information of the SMRI cohort was described elsewhere.[33] Extraction of all DNA was performed with a standard phenol–chloroform extraction method. The project has been approved by the Ethics Committee of the Centre for Addiction and Mental Health, Toronto.

Bisulfite treatment was performed by use of a standard protocol.[34] In brief, ~500 ng of genomic DNA was denatured in 0.3 M NaOH for 20 min at 42 °C. After the addition of freshly prepared saturated sodium metabisulfite (Sigma Aldrich, St Louis, MO, USA) and 10 mM hydroquinone (Sigma Aldrich) solution, samples were subjected to a 5-h incubation at 55 °C under exclusion of light. The samples were then purified and desulfonated with 0.3 M NaOH by using a Montage PCR96 96-well filtration plate (Millipore, Billerica, MA, USA).

Pyrosequencing

Four primer sets were designed to amplify the exon 1 and adjacent intron 1 region of the *HCG9* gene (Supplementary Table S1). Within these regions, six pyrosequencing assays were designed. In total, 28 CpGs in *HCG9* were covered. PCR amplifications were performed with a standard PCR protocol in 50 µl volume reactions containing 2 µl of sodium bisulfite-treated DNA, 0.15–0.5 µM primers, and master mix containing JumpStart or Platinum Taq DNA polymerase (Sigma Aldrich and Invitrogen, Carlsbad, CA, USA, respectively). After agarose gel electrophoresis to ensure successful amplification and specificity, PCR amplicons were processed for pyrosequencing analysis according to the manufacturer's standard protocol (Qiagen, Germantown, MD, USA). To validate the assay method, various proportions (0, 25, 50, 75 and 100%) of mixtures of CpGenome Universal Methylated and Unmethylated DNA (Millipore) were processed with the same protocol. Before proceeding with the clinical samples, it was confirmed that there are linear correlations between stepwise increments of methylated DNA and the measured methylation by pyrosequencing analysis in all CpG sites.

Genotyping

Genotyping of rs1128306 was performed by pyrosequencing. Selected genotypes were confirmed by resequencing in selected samples including the SMRI brain, McL brain and WBC 2 samples. PCR was performed using 10 ng of genomic DNA, 0.5 µM primers listed in Supplementary Table S1 and Taq DNA polymerase (New England BioLabs, Ipswich, MA, USA). PCR products were purified using

Table 1 Summary of the demographic variables of samples used in this study

	N	*Age (years)*	*Gender (M:F)*	*PMI (h)*	*AAO (years)*
Postmortem brains					
SMRI					
BPD	34	45.4 ± 10.7	16:18	37.9 ± 18.6	25.3 ± 9.23
SCZ	35	42.6 ± 8.47	26:9	31.4 ± 15.5	21.3 ± 6.07
CON	35	44.2 ± 7.58	26:9	29.4 ± 12.9	—
McL					
BPD	34	61.4 ± 18.9	14:20	22.2 ± 7.19	
SCZ	30	59.8 ± 13.3	20:10	22.8 ± 6.30	
CON	50	58.4 ± 15.6	30:20	20.9 ± 4.82	
Peripheral white blood cells					
Toronto					
BPD	100	42.1 ± 10	34:66	—	20.7 + 0.77
SCZ	50	43.3 ± 11.1	25:25	—	NA
CON	100	41.7 ± 10.9	34:66	—	—
GSK					
BPD	270	44.5 ± 12.6	121:149	—	20.3 + 0.55
CON	282	42.7 ± 0.81	136:146	—	—
Germline (sperm)					
BPD	29	40.2 ± 11.6	—	—	20.8 + 1.71
SCZ	14	37.4 ± 11.2	—	—	NA
CON	30	37.7 ± 10.3	—	—	—

Abbreviations: BPD, bipolar disorder; CON, control; GSK, GlaxoSmithKline; McL, brain samples from the McLean Brain Tissue bank including the prefrontal cortex; PMI, *post-mortem* interval in hours; SCZ, schizophrenia; SMRI, brain samples from the Stanley Medical Research Institute including prefrontal cortex, parietal cortex, occipital cortex and corpus callosum.

the QIAquick PCR purification kit (Qiagen) and processed for sequencing by the Center for Applied Genomics. Sequence alignment was performed by using ClustalW2.

Data analysis

Normality for the tested distributions was evaluated using an Anderson–Darling test. As a small percentage of CpG positions displayed non-Gaussian distributions, we evaluated significance using non-parametric tests. Mean values and s.e.m are reported in the text. $P \leqslant 0.05$ was deemed significant. The sliding window algorithm used as described in the text. As each test is not independent, we corrected for multiple testing using the Benjamini–Hochberg–Yekutieli-based method for calculating the FDR. The effect of potential confounding factors including brain pH, *post-mortem* interval and gender was assessed through non-parametric correlation. Correction for age and rs1128306 genotype was performed by taking the residuals of a linear model of DNA methylation as a function of age and rs1128306 genotype.

Results

HCG9 *methylation analysis in the brain of BPD patients and controls*

We investigated the regions surrounding the 12K CpG island microarray probe of *HCG9*: the first exon and a GC-rich region located 3′ of the first exon that extends into the first intron. After treatment with sodium bisulfite, each genomic DNA sample was subjected to six pyrosequencing assays (Figure 1b). To maximize our chances of finding functionally relevant DNA methylation differences, we evaluated the mean DNA methylation density per individual for all adjacent CpG positions and for all possible window sizes of CpG 1–28 resulting in 406 independent tests. For each of the 406 tested windows, a non-parametric Wilcoxon rank sum test was performed on these mean *HCG9* methylation values between the BPD and control group; similar analysis was done to compare the SCZ and control group. All tests were subjected to Benjamini–Hochberg–Yekutieli-based FDR correction for multiple testing, which controls the FDR under dependency assumptions. This method was first used to evaluate DNA methylation status in the SMRI brain cohort (Con $N = 35$, SCZ $N = 35$, BPD $N = 34$; see Table 1 for details) in attempts to validate the microarray data. In SCZ studies testing the prefrontal cortex of *post-mortem* brains, most results were non-significant (a single CpG window at CpGs 13–15 that displayed higher methylation in SCZ as compared with controls; SCZ = 17.4 ± 2.67, Con = 10.4 ± 1.34, $P = 0.031$). SCZ samples in WBC 2 and germline were also tested with few significant results; and therefore, we decided to focus our efforts solely on BPD.

We tested BPD in multiple brain regions in two samples: the SMRI (multiple regions: prefrontal

cortex, occipital cortex and corpus callosum) and McL (prefrontal cortex). No significant differences were observed in parietal, occipital or corpus callosum brain regions. Prefrontal cortex in both the SMRI and McL cohorts revealed the largest number of significant *P* values (although none survived correction for multiple testing) but the identified CpG regions were discrepant between the two cohorts (Supplementary Table S2). We observed generally a lower degree of methylation in BPD of SMRI sample, but higher in the McL sample. For example in SMRI CpG 3 (BPD = 35.6 ± 5.22, Con 50.4 ± 4.48, $P = 0.018$) but in McL brain sample (BPD = 37.1 ± 1.36, Con 31.2 ± 1.72, $P = 0.011$). To understand the reasons for the discrepancies, we investigated the effects of age, medication and DNA sequence differences on *HCG9* methylation in the two brain collections.

Analysis of the discrepancies between the SMRI and McL brain cohorts

HCG9 *methylation increases with age in controls.* The McL cohort represents significantly older individuals than the SMRI cohort (SMRI = 44.7. ± 1.12 years, McL = 60.1 ± 1.70 years; *t*-test, $P = 2.4 \times 10^{-11}$). For control individuals in each cohort, we correlated the age at death with the average DNA methylation in each of the 406 CpG window combinations. In both cohorts, we detected an increase in DNA methylation with age. This phenomenon appears to be fairly generalized over all 28 CpGs in the SMRI cohort (Spearman's $\rho = 0.30$, $P = 0.037$; Supplementary Table S4). McL control DNA methylation was positively correlated over CpGs 10–28 (Spearman's $\rho = 0.30$, $P = 0.037$). In a combined analysis of both brain cohorts, CpGs ranging from 3 to 28 displayed significant positive correlations with age (Figure 2a). In the McL cohort, we plotted the average DNA methylation difference in the most age discrepant groups (30–40 vs 70–80 years) for all CpG windows (Figure 2b). DNA methylation in younger BPD patients was less than controls, resembling the SMRI results. Together, these results corroborate the hypothesis that age differences are at least partially responsible for the discrepant initial findings between the two brain cohorts.

Effect of medication on HCG9 *methylation.* In the SMRI cohort (medication information for the McL sample was not available), correlation of lifetime antipsychotic measures (defined as fluphenazine equivalents in milligrams) with average age corrected CpG methylation across the 28 represented CpGs revealed a marginally significant association at a single CpG 14 (Spearmans' $\rho = 0.34$, $P = 0.05$), which became insignificant after correction for multiple testing. These results suggest that antipsychotic medications do not affect DNA methylation at *HCG9*.

Single-nucleotide polymorphism rs1128306 is associated with DNA methylation differences. The third factor that may have contributed to discrepancies between the SMRI and McL cohorts is related to the effect of DNA sequence variation on DNA methylation status. Using sequencing data obtained for the *HCG9* region of the SMRI data set, we evaluated 12 single-nucleotide polymorphisms (SNPs) in the region for evidence of allele-specific methylation. We detected no DNA methylation difference between alleles of SNPs rs17180353, rs2071568, rs373472, rs58031868, rs6903753, rs690402, rs9260832 and rs9278524, while SNPs rs400488, rs422640, rs9278523 and rs1128306 demonstrated significant association with mean DNA methylation (Kruskal-Wallis $\chi^2 = 13.26$, Bonferroni $P = 0.0024$ in all cases). These four SNPs were in perfect linkage disequilibrium, and so only rs1128306 (G/A) was used as a marker of genetic polymorphism in the region affecting DNA methylation, as it could be genotyped using pyrosequencing assay 3 in the remaining samples. We genotyped the SNP (rs1128306; G/A) located between CpGs 8 and 9 (Figure 1) and investigated it for association with *HCG9* methylation at each CpG position. DNA methylation between CpGs 5 and 9 was consistently lower in GG compared with GA plus AA genotypes (numbers of AA homozygotes were too small to investigate separately and therefore such were included in the GA group) reaching the maximal difference at CpG 5 (Wilcoxon rank sum test, GG minus GA = 16.8 ± 2.3, $P = 1.7 \times 10^{-9}$) (Supplementary Figure S1; Supplementary Table S5). Similar analyses of DNA samples extracted from WBC and germline showed similar rs1128306 effects and will be discussed in the sections below.

As the rs1128306 allele A in heterozygotes GA is associated with a higher density of methylated cytosines, we tested to see if the SMRI and McL cohorts had differing case and control allele frequencies for this allele. The BPD group in the SMRI cohort has less allele A possessing individuals as compared with that of the McL cohort (allele A frequency: SMRI Con = 0.4, SMRI BPD = 0.16, McL Con = 0.4, McL BPD = 0.26, $\chi^2 = 31.2$, $P = 0.075$). The higher proportion of BPD cases containing the allele A in the McL disease group could be another reason for the observed higher DNA methylation in BPD patients in this cohort.

Combined analysis of SMRI and McL brain cohorts. We eliminated the effects of age and the rs1128306 genotype through linear model transformation and performed non-parametric Wilcoxon rank sum tests on the 406 windows of the overall brain sample consisting of both SMRI and McL cohorts. This analysis was equivalent to evaluating significance with a linear model with age and genotype as covariates (with the exception that it was a more conservative non-parametric test). Despite the differences previously identified, now these two brain cohorts displayed numerous consistent results in the exon 1 and exon 1 3′ UTR region ranging between CpGs 1 and 10 (BPD = −2.0 ± 1.22, Con 1.68 ± 0.86, $P = 0.026$) with CpG 5 displaying the largest effect (BPD = −2.3 ± 1.52, Con = 1.86 ± 1.13,

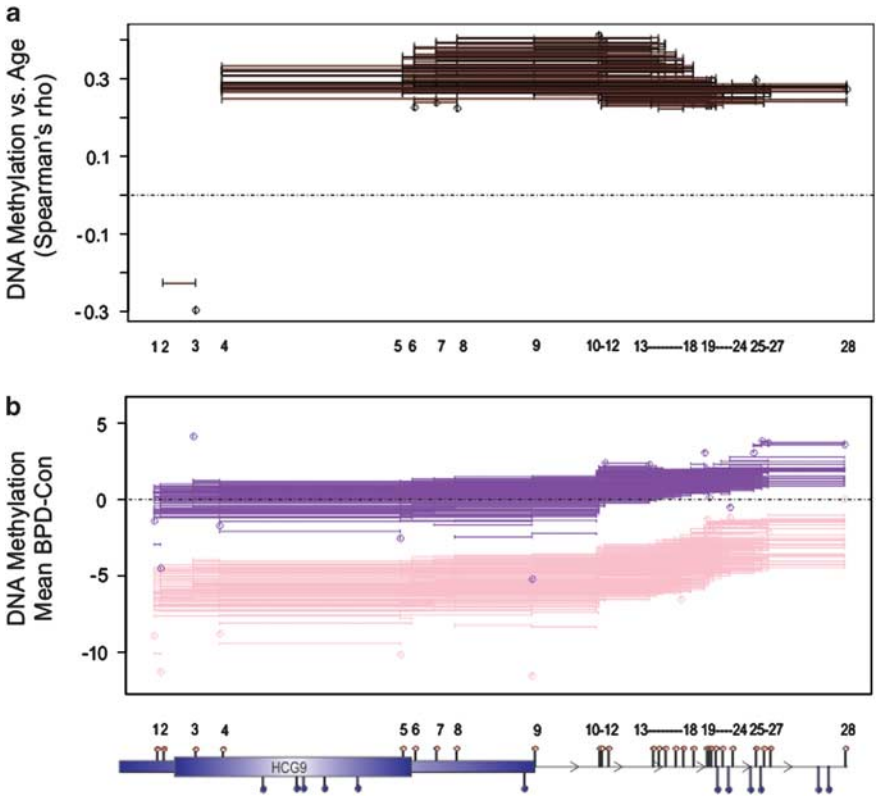

Figure 2 Age-related increase in *HCG9* methylation. (**a**) Spearman's correlation coefficients (ρ) for age at death and average DNA methylation in control individuals from Stanley Medical Research Institute (SMRI) and McLean (McL) brain cohorts combined ($N=84$); each line shows the corresponding window of CpGs. Only windows with $P \leqslant 0.05$ are depicted. (**b**) Average difference in DNA methylation density between younger (purple: 30–40 years) and older (pink: 70–80 years) bipolar disorder (BPD) and control group individuals (McL cohort only). Data shown for all possible CpG window variants.

$P=0.05$) (Supplementary Table S3). Negative mean DNA methylation values are the result of data correction.

HCG9 *methylation studies in WBCs*

In addition to the brain studies, we investigated the *HCG9* methylation profiles in peripheral WBCs (total $N=370$ BPD patients and $N=382$ controls) using the same approach. Before disease analysis, we evaluated the effect of age on DNA methylation in controls from the WBC cohort. Similar to the brain studies, DNA methylation increased with age over a majority of CpGs in the WBC cohort with significant correlations observed in windows between CpGs 8 and 28 (Spearman's $\rho=0.3$, $P=0.00012$). Another similarity with the brain findings was that GG homozygotes at rs1128306 displayed significantly lower methylation compared with GA heterozygotes in the control population (Supplementary Figure S1; Supplementary Table S5). After correction for age and rs1128306 genotype, we evaluated DNA methylation associations with BPD. CpGs 5–8 displayed lower methylation in BPD (BPD $=-0.39\pm0.14$, Con $=0.35\pm0.21$, $P=0.0006$) (Supplementary Figure S2). All of these identified differences are located in the same region and same direction as those identified in the brain studies.

Information available in the WBC cohort gave us the opportunity to evaluate the effect of mood stabilizers on DNA methylation in the region. We determined daily mood stabilizer dosage information in the WBC cohort (mood stabilizers included lithium and valproate) and correlated our 406 possible CpG windows against these values. No significant correlations were observed for lithium dose, while for valproate, we obtained significant correlations ranging from CpGs 2 to 9 (Spearman's $\rho=0.13$, $P=0.05$) and CpG 28 (Spearman's $\rho=0.14$, $P=0.043$). The direction of correlation suggests that increased mood stabilizer dose increases DNA methylation toward control levels. While the strength of the correlations is not strong enough to conclude that mood stabilizers help to alleviate *HCG9* DNA methylation abnormalities in BPD patients, these results suggest our disease-specific associations are not the spurious result of a class of medications commonly prescribed to BPD patients.

To understand the BPD predictive value of DNA methylation changes in peripheral WBCs, we performed a logistic regression analysis, modeling

the effect of DNA methylation on disease, controlling for age and rs1128306 genotype. The mean density of modified cytosines at CpGs 6–9 displayed the most significant effect (odds ratio (OR) = 1.07, $P = 0.0046$).

Next, we evaluated the logistic regression model based on age, rs1128306 genotype and DNA methylation at CpGs 5–8 for its ability to predict a diagnosis of BPD in two ways. We initially generated the model using WBC 2 ($N = 270$ BPD patients and 282 controls) to predict diagnosis in WBC 1 ($N = 100$ BPD patients and 100 controls) (see Samples; Figure 3a) and subsequently randomly selected 90% of the combined WBC sample to predict diagnosis based on the remaining 10% consisting of 38 BPD patients and 37 controls (Figure 3b). We compared the predicted probabilities with the true diagnoses over a range of probability thresholds to generate a receiver operating characteristic curve and generated area under the receiver operating characteristic curve (A prime) values of 0.69 and 0.72, respectively (Figure 3). This value means that, given one BPD and one control sample, these models have 69 and 72% probability, respectively, of assigning a higher value to the BPD case.

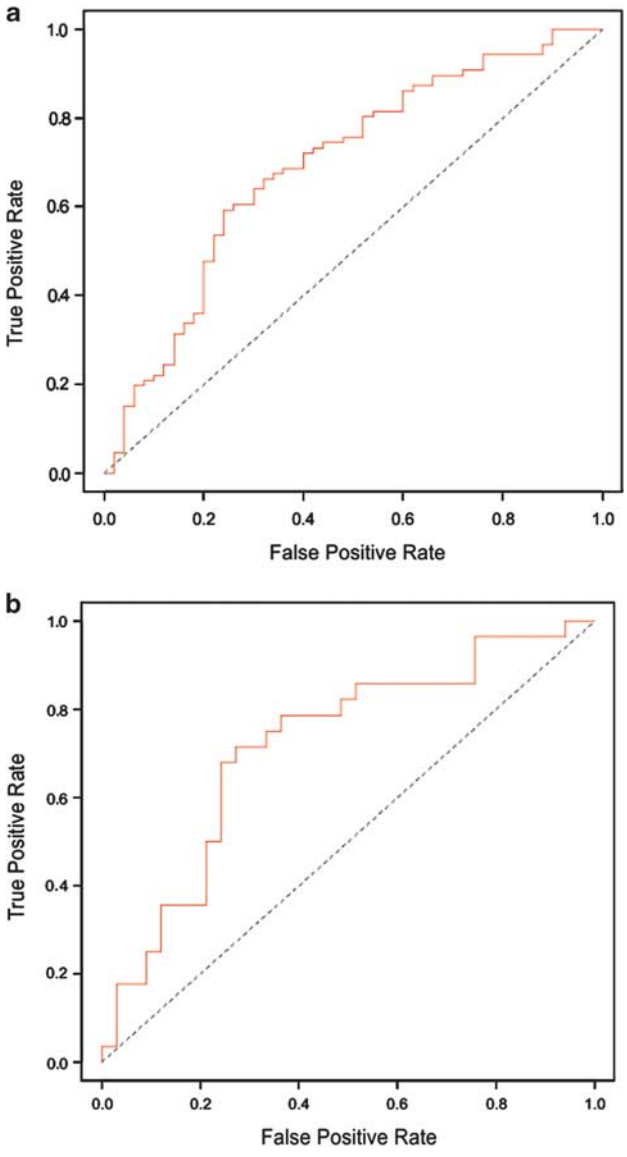

Figure 3 Bipolar disorder (BPD) predictive model characteristics. Receiver operating characteristic curves generated for the prediction of the WBC 1 cohort and a randomly selected 10% of the combined WBC cohorts using logistic regression models generated using the mean DNA methylation density at CpGs 5–8, rs1128306 genotype and age in the WBC 2 (**a**) cohorts and 90% of the WBC sample (**b**). WBC, white blood cell.

HCG9 *methylation studies in the germline*

In the germline samples, the epigenetic effects of the rs1128306 genotype were consistent with the other tissues (Supplementary Figure S1; Supplementary Table S5). After correction for rs1128306 genotype and age, a significantly lower DNA methylation density was detected at CpG 5 (BPD = -0.68 ± 0.44, Con = 0.65 ± 0.54, $P = 0.028$), which is also consistent with brain and WBC findings. Interestingly, in the germline, unlike the brain and WBC, older age was associated with lower *HCG9* methylation. For example, raw Spearman's correlation values were $\rho = -0.65$ ($P = 0.021$) for CpGs 4–20, $\rho = -0.4$ ($P = 0.033$) for CpGs 1, $\rho = -0.41$ ($P = 0.035$) for CpGs 11–12 and $\rho = -0.39$ ($P = 0.036$) for CpG 16. The opposite age effects in germline and WBC suggest that germline methylation differences in BPD patients were genuine and not artifacts of WBC contamination, which may reach 5% of the total cell count in semen samples. Also, GG and GA methylation differences in the germline have a peak at CpGs 5–8 but a dip in WBC (Supplementary Figure S1), which argues for the same.

HCG9 *methylation across tissues of BPD patients and controls*

In the process of this project, we noticed that despite variable density of modified cytosines at each specific position *HCG9* methylation profiles were similar across the brain, WBC and germline samples, which suggests that germline *HCG9* epigenetic pattern was partially retained in the somatic tissues. As expected, the average *HCG9* methylation profile across controls in each tissue correlated with the germline (Figure 4a; Supplementary Table S6).

We next investigated the relative risk to disease of genotype- and age-independent DNA methylation effects using the entire multi-tissue sample. We evaluated the cumulative effects for brain, WBC and germline samples using a random effects meta-analysis to combine logistic regression model DNA methylation ORs for each tissue across the 406 CpG windows. Lower DNA methylation at CpGs 5–9 represented the most significant effect on disease risk (summary OR = 1.24, $P = 0.0011$; Figure 4b). The results of this analysis suggest that, independent of age and the rs1128306 genotype, lower DNA methylation at *HCG9* at CpGs 5–9 is associated with BPD across the brain, WBC and germline.

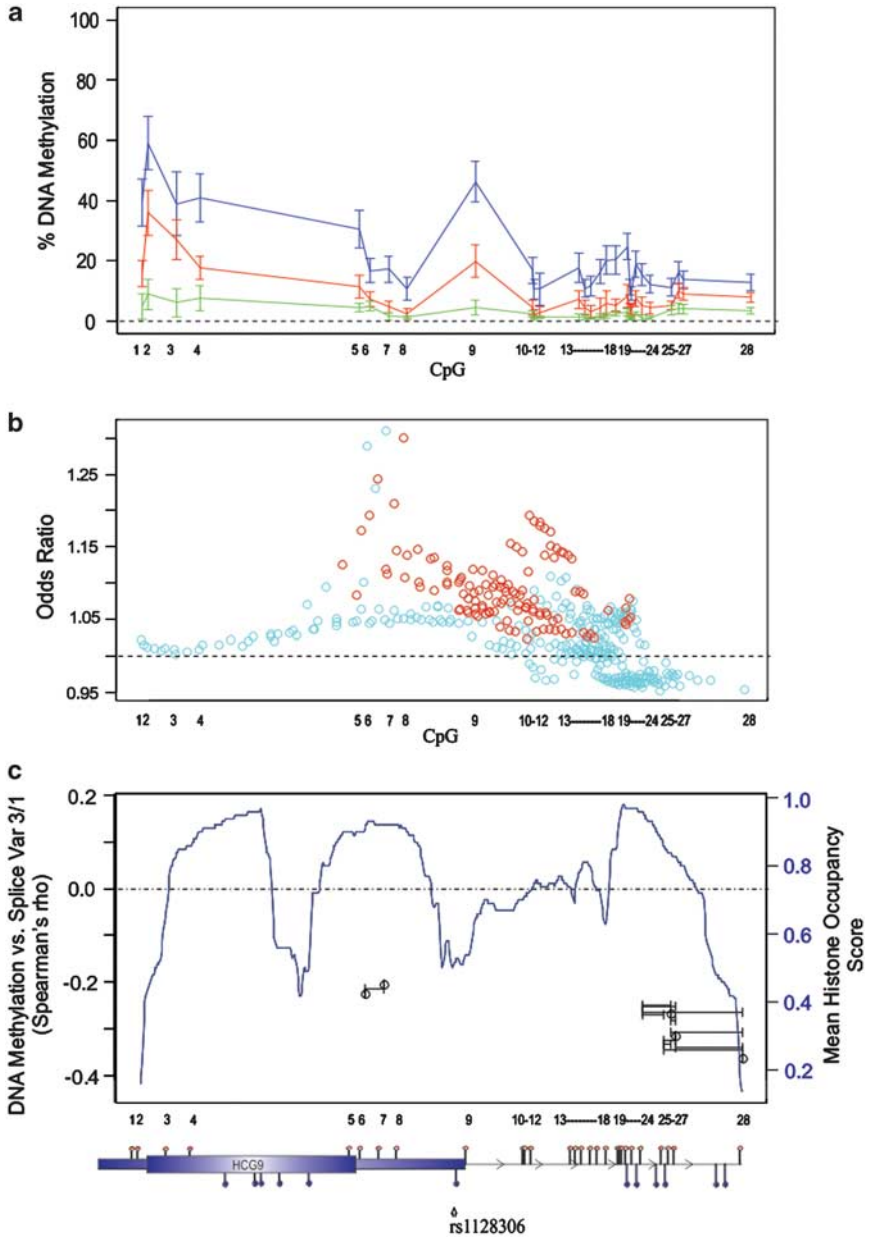

Figure 4 (**a**) *HCG9* DNA methylation profiles by tissue. Mean DNA methylation for the control groups for each of the 28 CpGs for the brain (blue), white blood cell (WBC) (red) and germline (green) cohorts is plotted, demonstrating the conserved DNA methylation patterns across tissues. (**b**) Combined sample analysis of bipolar disorder (BPD) risk by genomic location. The odds ratio (OR) for BPD risk due solely to DNA methylation differences in BPD patients and controls is plotted relative to the middle genomic position of each of the 406 tested CpG windows. The rs1128306 single-nucleotide polymorphism (SNP) and age were modeled as covariates in a logistic regression model and ORs due to DNA methylation were combined with a random effects meta-analysis for all tissues. The middle genomic positions of CpG windows where DNA methylation significantly contributes to risk are plotted in red. Of interest, when the rs1128306 polymorphism alone was evaluated in BPD evidence for association was observed for the combined brain sample and WBC sample; however, allele G exerted an increased risk in the combined brain (OR = 2.45, $P = 0.017$) but protective effect in WBC (OR = 0.72, $P = 0.044$, respectively). No association was observed in the germline sample or an analysis of all samples combined. (**c**) DNA methylation vs *HCG9* steady-state mRNA levels. Non-parametric Spearman correlations (ρ) between DNA methylation and the ratio of *HCG9* splice variants 3 and 1 (windows with $P \leqslant 0.05$) (black, *Y* axis left), and mean histone occupancy scores as modeled by a sequence-based histone prediction program (blue, *Y* axis right).

Putative pathological roles of HCG9 *in BPD*

As the pathological role of intronic methylation differences at *HCG9* in BPD is not straightforward, we sought to elucidate any relationship between DNA methylation differences and *HCG9* steady-state mRNA levels. We analyzed three gene expression data sets (studies 2, 3 and 7) from the SMRI genomics database that were performed on the *Affymetrix*

HGU133A gene expression arrays, which contain 11 probes covering *HCG9* (Figure 1a). A limited number of McL cohort individuals ($N=34$) also contained steady-state mRNA data from the same array platform, and we performed *HCG9* mRNA analyses on the combined SMRI and McL brain data set. As there are indications that intronic chromatin modifications may be involved in alternative splicing, we were interested to ascertain steady-state mRNA levels for individual *HCG9* splice variants (Figure 1a). In order to do this, we subjected all data to background correction and quantile normalization and took the mean $\log_2$ probe value for each of the 11 probes spanning *HCG9*. Probes 10 and 11 (probe group 1) uniquely bind mRNA from splice variant 3, while probes 1–4 (probe group 2) bind mRNA from variants 1 and 3 (Figure 1a). After averaging these two probe groups, we determined the levels of *HCG9* splice variant 1 by taking the residuals of a linear model between probe groups 1 and 2, statistically subtracting splice variant 3 from the combined 1 and 3 measures. A significantly higher log ratio of *HCG9* mRNA splice variant 3/1 was detected in BPD compared with controls (Wilcoxon rank sum test, mean ± s.e.m.: BPD = 1.45 ± 2.17, Con = −1.77 ± 15.6, $P=0.034$).

Since the maximal DNA methylation differences in BPD were detected around CpG 5 in the two brain cohorts, and CpGs 5–9 represented the most significant effect on disease risk in the total sample of the three tissues, we hypothesized that epigenetic modifications may be contributing to splicing decisions in this region which corresponds to the overlapping splice variants 1 and 3. The log ratio of splice variant 3 over splice variant 1 was tested using the DNA methylation sliding window. CpG 6 (that is, next to CpG 5) displayed a significant negative correlation with splice variant ratios (Spearman's $\rho=-0.23$, $P=0.029$; Figure 4c) although did not survive correction for multiple testing. The largest correlation was detected for CpG 25 (Spearman's $\rho=-0.44$, $P=5.33\times10^{-5}$, FDR $P=0.022$; Figure 4c). Interestingly, CpG 25 maps next to CpGs 23–24 which in the SMRI brain cohort displayed the highest significance for differential methylation between the BPD and control groups (BPD = 10.7 ± 3.28, Con 15.2 ± 1.60, $P=0.0021$) (Supplementary Table S1). Other significant correlations are shown in Supplementary Table S7. Both CpG 6 and CpG 25 corresponded to peak mean histone occupancy scores 7–14 bp downstream of CpG 6 and ~60 bp upstream of CpG 25 as predicted by a nucleosome prediction algorithm[35] (Figure 4c). Another algorithm[36] mapped tentative nucleosome centers 20–50 bp upstream of CpG 6 and ranging from ~45 bp upstream to 17 bp downstream of CpG 25, while a third program, FineStr 1.0,[37] mapped nucleosome centers ~60 bp downstream of CpG 6 and 2 bp downstream of CpG 25. As the amount of DNA that coils around a single histone octamer is ~150 bp, CpGs 5–8 and 24–27 can be reasonably associated with a single nucleosome peak. Consistently with our initial observation of lower methylation at *HCG9* in BPD patients, these data suggest that as *HCG9* DNA methylation decreases in BPD, there is an increase in the ratio of splice variant 3 to splice variant 1 in the brain.

HCG9 has three potential splice variants, generating 77, 37 and 121 amino-acid (aa) proteins (Figure 1). Using Protein BLAST, we determined the 37 and 77 aa peptides lack conserved domains and do not share homology with any proteins in the non-redundant protein database. The 121 aa peptide also lacked conserved domains, however, did produce a single hit in a known locus with the exported acetyl esterase (YP_001901494) in *Xanthomonas campestris*. These class of enzymes have been implicated in regulating B-cell immunological tolerance, and the sialic acid acetylesterase was recently associated with autoimmune disorders including rheumatoid arthritis and type 1 diabetes.[38]

We next attempted to elucidate the function of *HCG9* by performing gene ontology (GO) analysis on those genes whose expression from the previously mentioned SMRI data sets is most tightly correlated with that of *HCG9*. For the BPD and control individuals in the SMRI studies 2, 3 and 7 performed on the HGU133A gene expression platform, we correlated the expression of all probes against that of *HCG9* using Spearman method followed by FDR correction. We selected only those genes significantly correlated with *HCG9* in all three studies using an FDR $P<0.05$ and subsequently performed GO analysis. GO categories associated with genes in this selected group were tested against their frequency of occurrence on the *Affymetrix* HGU133A array. Significance was assessed using the binomial test in R and corrected for multiple testing using the Bonferroni method. Significant GO categories (Bonferroni $P<0.01$) appear in Supplementary Table S8. The results suggest that when expression of *HCG9* changes, so does that of genes involved in signaling and immune system-related functions such as inflammation and regulation of B cell-based immunological tolerance, among others.

Discussion

Microarray-based DNA methylation profiling provides an unbiased approach for the identification of epigenetic differences that may have an etiological role in human disease. Our earlier study of ~1% of the methylome revealed a number of candidate regions displaying altered DNA methylation signatures associated with major psychosis.[39] In this project, we explored methylation patterns of an HLA gene, *HCG9*, in *post-mortem* brain tissues, WBC, and the germline, which all consistently revealed lower methylation of *HCG9* in BPD patients compared with the controls.

HCG9 methylation patterns in the brain are complex and depend on age as well as DNA sequence variation. Several reports implicated age as a

contributing factor to DNA methylation change.[40–42] *HCG9* methylation increases in the aging brain and seems to be approaching the normal state observed in controls, which is consistent with the common observation that clinical symptoms of BPD in aging individuals become less severe.[43] Alternatively, DNA methylation levels decreased with age in the germline sample, toward the risk epigenotype. This finding suggests that with increasing paternal age, the probability of contributing low DNA methylation sperm upon fertilization increases and may have relevance to epidemiological observations of higher incidence of BPD among children of older fathers.[44–46]

DNA methylation differences associated with genetic polymorphisms have been documented and are known as allele-specific methylation patterns.[39,47] The rs1128306 SNP at *HCG9* also contributes to DNA methylation differences across individuals. More specifically, allele A exhibited a higher density of methylated cytosines in ~65 bp in each direction of CpG 8 in comparison with the alternative allele G. This was observed across all three investigated tissues. While we identified associations of this SNP with BPD between the combined brain sample and WBC sample, alone, the sample sizes of these individual cohorts are likely too small to identify true SNP associations with BPD and may explain why allele G appeared to exert a risk inducing and protective effect in the respective cohorts. These results highlight the possibility that DNA polymorphisms associated with disease, such as those identified in GWAS studies, may be acting synergistically with epigenetic misregulation.[48,49] The relative contribution of DNA sequence variation vs epigenetic variation to the disease phenotype may vary significantly. The extreme case is when only epigenetic differences have a pathological role while SNPs are only the reporters of such epimutations. It is also interesting to note that the G/A polymorphism (flanking nucleotides are C and A) may have originated from a C→T mutation in the complementary strand (T[C/T]G). The dinucleotide CpG is a 'hotspot' for mutation in the human genome, most likely due to high frequency of spontaneous deamination of 5-methylcytosine to thymidine.[50] The ancestral allele of rs1128306 SNP is unknown but based on a primarily Caucasian population, allele G(C) is three times as common as allele A(T) (0.75 and 0.25, respectively), and the later allele exhibits higher methylation compared with the former, consistent with the epigenetic origin of rs1128306 SNP. Therefore, in addition to a direct pathogenic effect, DNA methylation can contribute to human genetic disease indirectly, that is, via increased mutation rates.

The observation of a DNA sequence–epigenetic association is particularly interesting in the light of recent findings suggesting that 6p21.3-6p22 (*HCG9* maps to NCBI 36; chr6:30050871-30054156) has been implicated by several large GWAS in major psychosis.[28–30] The rs1128306 SNP was a strong marker for a portion of the epigenetic variation in the region. Unfortunately, no linkage disequilibrium data exists in HapMap for this SNP; however, this and possibly other SNPs could be linked to the markers tested by the GWAS, suggesting that a portion of the disease association signal coming out of the major histocompatibility complex (MHC) could represent this and possibly other epigenetically misregulated regions.

Of particular interest and added value were the two non-brain tissues, WBC and sperm, which 'mirrored' most of the brain findings and provided additional confidence that the *HCG9* methylation changes in the *post-mortem* brains of BPD patients are genuine. The identification of an epigenetic difference in BPD detectable in WBCs holds promise for the identification of predictive biomarkers for the disease. Although area under the receiver operating characteristic curves demonstrated only a mildly predictive value for disease, given that *HCG9* was selected from an interrogation of only ~1% of the epigenome,[39] the *HCG9* epigenetic biomarker for BPD stands as a proof of principle that the epigenomic studies of multiple tissues may result in clinical applications. Identification of biomarkers should be far more successful if the candidate genes or regions are selected from the most significant changes observed in epigenome-wide analyses.

In epigenetic studies, cellular heterogeneity is a potential confounder. Peripheral WBC DNA comes from a wide variety of blood cells, of which the proportion of cells retaining the inherited or acquired pre-epimutation may be quite small. For example, the pre-epimutation may be lost in differentiation from a precursor stem cell to some mature cell types and not others. Similarly, in the brain tissues, *HCG9* methylation differences could be resultant from having only a particular subset of neurons or glial cells affected and ultimately conferring the observed epigenetic difference. The epigenetic noise caused by cellular heterogeneity could be affecting our predictive model of BPD diagnosis. One of the next steps of *HCG9* methylation studies in BPD or other psychiatric disease could be dedicated to investigation of particular WBC subtypes or homogenous brain cell populations. Cellular heterogeneity issues call for the development and optimization of technologies capable of isolating and investigating small numbers of cells, even to the level of the single cell to adequately address these issues.

One major caveat of epigenetic studies of complex disease is that the etiological significance of identified DNA methylation changes is not immediately evident as epigenetic patterns in the diseased tissue may be influenced by the disease state, treatment or other events related to the pathological process. We sought to evaluate the cause vs effect relationship between the identified DNA methylation differences and disease at the *HCG9* gene by analyzing non-brain tissues, WBC and germline. The lower *HCG9* methylation density in peripheral WBCs mapping to the same region of *HCG9* nominated from the combined brain cohort argues (although does not prove) for an etiological role of *HCG9* methylation differences in BPD. The causal link

between *HCG9* epimutation and BPD is further supported by our analysis of sperm DNA, which also revealed *HCG9* methylation differences (for example, CpG 5) in BPD patients. In general, DNA methylation is known to be subject to reprogramming during both gametogenesis and after fertilization in humans and mice; however, examples exist where epigenetic alterations in the parental generation are passed to the offspring.[51–53] The results of germline studies allow us to infer the possibility that *HCG9* epimutation may represent one of the heritable epigenetic risk factors in BPD. It is important to keep this finding in perspective, however, as the epigenetic difference was identified in the sperm of men affected with BPD and not the fathers of BPD patients. We can only suggest that the identified *HCG9* methylation profiles are to some extent meiotically stable and thus may have the potential to survive successive epigenetic resetting and be passed to the next generation. The mechanisms of such a passage of epigenetic information are beyond the scope of this study but may involve a communication between inherited DNA methylation and histone protein modifications, non-coding RNAs or interaction with DNA sequence variation in the region.

Functional implications of the detected epigenetic difference at CpGs 5–9, just downstream of the first coding exon, on *HCG9* activity and its role in BPD require further study. Although traditionally the regulatory function of DNA methylation is most characteristically recognized in its role in the gene promoter,[54–57] in the case of *HCG9* and BPD, the disease-related mechanism may be different. One possibility is that DNA methylation at CpGs 5–9 may be contributing to alternative splicing scenarios of the *HCG9* mRNA transcript. According to AceView, *HCG9* has three mRNA splice variants, only two of which (77 and 121 aa) may be encoding protein products; however, it remains possible that one or both of these splice variants represent non-coding transcripts. There is a growing body of evidence that chromatin conformation and histone modifications help direct aspects of co-transcriptional splicing.[58] In particular, reduction of histone 3 lysine 4 tri-methylation (H3K4me3) in conjunction with CHD1 chromatin-remodeling ATPase knockdown was found to alter splicing efficiency,[59] while increased levels of histone 3 lysine 9 (H3K9) acetylation triggered by neuronal depolarization were associated with exon skipping.[60] In our experiments, we found correlations between DNA methylation and *HCG9* splice variant ratios at peaks of predicted histone occupancy, suggesting that our DNA methylation levels may be a reflection of associated histone modification status and that epigenetic alterations in these positions are important for alternative splicing mechanisms. The decreased *HCG9* DNA methylation in BPD patients may account for the increased ratio of splice variant 3 to splice variant 1 in the brain of affected individuals. The observed increased ratio of the 121 aa over 77 aa protein in BPD may represent a gain of deleterious neural inflammatory function or a loss of protective anti-inflammatory function in BPD and may contribute to disease risk.

Using GO analysis, we observed that *HCG9* is co-expressed with genes involved in inflammation while Protein BLAST identified homology of the 121 aa peptide with a class of genes negatively associated with inflammatory response.[61] CNS inflammation has been hypothesized as a contributing factor for BPD,[62] while pro-inflammatory biomarkers have been found to be predictive of acute episodes in BPD.[63] Furthermore, antipsychotic and antidepressant medications have been known for their ability to reduce inflammation.[64–68] A downregulation of *HCG9* was observed in a recent transcriptome analysis after treatment with pro-inflammatory factors, which is consistent with the interpretation that one of the splice variants of *HCG9* confers an anti-inflammatory signal.[32] Loss of HCG9 protective function as a result of alternative splicing could account for the deleterious effects of the epimutation and its association with disease. Recently, one study identified a change in *HCG9* expression levels with the addition of a histone deacetylase inhibitor suberoylanilide hydroxamic acid and taxanes to human breast cancer cells,[69] further corroborating an epigenetic component to *HCG9* expression and suggesting that epigenetic modifiers may have therapeutic benefit in BPD.

Another important activity in the further characterization of the role of *HCG9* in BPD is identification of the role of 5-hydroxymethylcytosine, which is known to be present at high levels in the brain.[70,71] It has been recently shown that 5-hydroxymethylcytosine, like 5-methylcytosine, does not undergo C-to-T transitions after bisulfite treatment, indicating that these two modified cytosine species are indistinguishable by the bisulfite sequencing technique.[72] Therefore, all our estimates of methylated cytosines in *HCG9*, in fact, represent both types of cytosine modification. Recently, some new approaches have been developed to differentiate the two types of cytosine modifications[73–75] and these can be used in future studies to address the roles of 5-hydroxymethylcytosine in health and disease.

The results of this study highlight the importance of controlling for confounding influences in brain tissue-based studies of complex disease. We identified significant DNA methylation associations with age and genetic influences that, if ignored, have the potential to result in false-positive or false-negative findings. Additionally, we used the analytical strategy that favors replication across different tissues rather than direct statistical evidence for isolated analysis of each tissue. By using a sliding window-based method and performing statistical tests on all possible adjacent CpG combinations, we identified overlaps across tissues and replication cohorts that we would have otherwise missed. This approach may become exemplary in the primary studies of DNA modification differences when neither the number of cytosines involved, nor the multiple ways of their interaction is known. This study highlights the need for methodological principles for epigenomic and epigenetic

studies of complex disease that are quite different from the ones used in DNA sequence-based approaches.

Note added in proof: While the manuscript was reviewed and processed, we performed a pilot analysis of 5-hydroxymethylation (5-hmC) at CpG 6 which happens to be at the target site for MspI restriction enzyme. Glucosylation of the 5-hmC at CCGG renders such targets resistant to MspI digestion, and the differential enrichment of MspI-target amplicons with and without glucosylation estimates 5-hmC. In this experiment, a lower level of 5-hmC was identified in a subset of BPD prefrontal cortex samples compared to controls, suggesting that a portion of the identified DNA methylation differences in this study may be related to 5-hmC and warranting further follow-up analyses.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We would like to thank the Stanley Medical Research Institute (SMRI) and Dr Francine Benes of the McLean Hospital Brain Bank for the provided brain tissues. ZK and JM were CIHR fellows. MT was supported by the Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad. AP is Senior Fellow, Ontario Mental Health Foundation, and Tapscott Chair in Schizophrenia Studies, University of Toronto. This work was supported by the Canadian Institutes for Health and Research (186007) and the National Institutes of Health (MH074127; MH088413) to AP.

Web Resources

AceView, http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/
Clustal W2, http://www.ebi.ac.uk/Tools/clustalw2/index.html
FineStr, http://www.cs.bgu.ac.il/~nucleom/
Nucleosome Prediction Algorithm, http://genie.weizmann.ac.il/software/nucleo_prediction.html/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/
Stanley Medical Research Institute Genomics Database, http://www.stanleygenomics.org/

References

1 Petronis A. Human morbid genetics revisited: relevance of epigenetics. *Trends Genet* 2001; **17**: 142–146.
2 Robertson KD. DNA methylation and human disease. *Nat Rev Genet* 2005; **6**: 597–610.
3 Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007; **447**: 433–440.
4 Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 2010; **465**: 721–727.
5 Yeivin A, Razin A. Gene methylation patterns and expression. In: Jost J, Saluz, H (eds). *DNA Methylation: Molecular Biology and Biological Significance*. Birkhauser Verlag: Basel, 1993 pp. 523–568.
6 Holliday R, Ho T, Paulin R. Gene silencing in mammalian cells. In: VEA Russo RM, Riggs AD (eds). *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 1996, pp 47–59.
7 Riggs AD, Xiong Z, Wang L, Lebon JM. Methylation dynamics, epigenetic fidelity and X chromosome structure. *Novartis Found Symp* 1998; **214**: 214–225; discussion 225–32.
8 Ehrlich M, Ehrlich, K. Effect of DNA methylation and the binding of vertebrate and plant proteins to DNA. In: Jost J, Saluz P (eds). *DNA Methylation: Molecular Biology and Biological Significance*. Birkhauser Verlag: Basel, Switzerland, 1993, pp 145–168.
9 Riggs A, Xiong Z, Wang L, Jm L. Methylation dynamics, epigenetic fidelity and X chromosome structure. In: Wolffe, A (ed). *Epigenetics*. John Wiley & Sons: Chistester, 1998, pp 214–227.
10 Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML *et al*. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 2005; **102**: 10604–10609.
11 Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, Wong AH *et al*. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet* 2009; **41**: 240–245.
12 Richards EJ. Inherited epigenetic variation—revisiting soft inheritance. *Nat Rev Genet* 2006; **7**: 395–401.
13 Petronis A. Epigenetics and bipolar disorder: new opportunities and challenges. *Am J Med Genet* 2003; **123C**: 65–75.
14 Petronis A. The genes for major psychosis: aberrant sequence or regulation? *Neuropsychopharmacology* 2000; **23**: 1–12.
15 Kaminsky Z, Wang SC, Petronis A. Complex disease, gender and epigenetics. *Ann Med* 2006; **38**: 530–544.
16 Hall JG. Genomic imprinting: review and relevance to human diseases. *Am J Hum Genet* 1990; **46**: 857–873.
17 Barlow DP. Gametic imprinting in mammals. *Science* 1995; **270**: 1610–1613.
18 Newberg AR, Catapano LA, Zarate CA, Manji HK. Neurobiology of bipolar disorder. *Expert Rev Neurother* 2008; **8**: 93–110.
19 Guidotti A, Dong E, Kundakovic M, Satta R, Grayson DR, Costa E. Characterization of the action of antipsychotic subtypes on valproate-induced chromatin remodeling. *Trends Pharmacol Sci* 2009; **30**: 55–60.
20 Dong E, Chen Y, Gavin DP, Grayson DR, Guidotti A. Valproate induces DNA demethylation in nuclear extracts from adult mouse brain. *Epigenetics* 2010; **5**: 730–735.
21 Kwon B, Houpt TA. Phospho-acetylation of histone H3 in the amygdala after acute lithium chloride. *Brain Res* 2010; **1333**: 36–47.
22 Mccoll G, Killilea DW, Hubbard AE, Vantipalli MC, Melov S, Lithgow GJ. Pharmacogenetic analysis of lithium-induced delayed aging in Caenorhabditis elegans. *J Biol Chem* 2008; **283**: 350–357.
23 Petronis A. The origin of schizophrenia: genetic thesis, epigenetic antithesis, and resolving synthesis. *Biol Psychiatry* 2004; **55**: 965–970.
24 Schumacher A, Petronis A. Epigenetics of complex diseases: from general theory to laboratory experiments. *Curr Top Microbiol Immunol* 2006; **310**: 81–115.
25 Flanagan JM, Popendikyte V, Pozdniakovaite N, Sobolev M, Assadzadeh A, Schumacher A *et al*. Intra- and interindividual epigenetic variation in human germ cells. *Am J Hum Genet* 2006; **79**: 67–84.
26 Sutherland JE, Costa M. Epigenetics and the environment. *Ann NY Acad Sci* 2003; **983**: 151–160.
27 Jaenisch, R, Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003; **33**(Suppl): 245–54.
28 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I *et al*. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–757.
29 Purcell SM, Wray NR, Stone JL, Visscher PM, O'donovan MC, Sullivan PF *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
30 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al*. Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.
31 Gaudieri, S, Longman-Jacobsen N, Tay GK, Dawkins RL. Sequence analysis of the MHC class I region reveals the basis of the genomic matching technique. *Hum Immunol* 2001; **62**: 279–285.
32 Seidl, C, Port M, Apostolidis C, Bruchertseifer F, Schwaiger M, Senekowitsch-Schmidtke R *et al*. Differential gene expression

triggered by highly cytotoxic alpha-emitter-immunoconjugates in gastric cancer cells. *Invest New Drugs* 2010; **28**: 49–60.
33 Torrey EF, Webster M, Knable M, Johnston N, Yolken RH. The Stanley foundation brain collection and neuropathology consortium. *Schizophr Res* 2000; **44**: 151–155.
34 Kaminsky Z, Petronis A. Methylation SNaPshot: a method for the quantification of site-specific DNA methylation levels. *Methods Mol Biol* 2009; **507**: 241–255.
35 Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y *et al* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 2009; **458**: 362–366.
36 Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J *et al.* A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* 2008; **18**: 1073–1083.
37 Gabdank I, Barash D, Trifonov EN. FineStr: a web server for single-base-resolution nucleosome positioning. *Bioinformatics* 2010; **26**: 845–846.
38 Surolia I, Pirnie SP, Chellappa V, Taylor KN, Cariappa A, Moya J *et al.* Functionally defective germline variants of sialic acid acetylesterase in autoimmunity. *Nature* 2010; **466**: 243–247.
39 Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, Bouchard L *et al.* Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 2008; **82**: 696–711.
40 Fuke C, Shimabukuro M, Petronis A, Sugimoto J, Oda T, Miura K *et al.* Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Ann Hum Genet* 2004; **68**(Part 3): 196–204.
41 Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H *et al.* Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* 2010; **20**: 434–439.
42 Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H *et al.* Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 2010; **20**: 440–446.
43 Depp CA, Jeste DV. Bipolar disorder in older adults: a critical review. *Bipolar Disord* 2004; **6**: 343–367.
44 Frans EM, Sandin S, Reichenberg A, Lichtenstein P, Langstrom N, Hultman CM. Advancing paternal age and bipolar disorder. *Arch Gen Psychiatry* 2008; **65**: 1034–1040.
45 Menezes PR, Lewis G, Rasmussen F, Zammit S, Sipos A, Harrison GL *et al.* Paternal and maternal ages at conception and risk of bipolar affective disorder in their offspring. *Psychol Med* 2010; **40**: 477–485.
46 Crow TJ. Mutation and psychosis: a suggested explanation of seasonality of birth. *Psychol Med* 1987; **17**: 821–828.
47 Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E *et al.* Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet* 2008; **40**: 904–908.
48 Tycko, B. Allele-specific DNA methylation: beyond imprinting. *Hum Mol Genet* 2010; **19**(R2): R210–R220.
49 Meaburn, EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome Implications for genetic studies of complex disease. *Epigenetics* 2010; **5**: 578–582.
50 Cooper DN, Krawczak M. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet* 1989; **83**: 181–188.
51 Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999; **23**: 314–318.
52 Oates NA, Van Vliet J, Duffy DL, Kroes HY, Martin NG, Boomsma DI *et al.* Increased DNA methylation at the AXIN1 gene in a monozygotic twin from a pair discordant for a caudal duplication anomaly. *Am J Hum Genet* 2006; **79**: 155–162.
53 Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV *et al.* Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci USA* 2003; **100**: 2538–2543.
54 Dulac C. Brain function and chromatin plasticity. *Nature* 2010; **465**: 728–735.
55 Ehrlich M, Wang RY. 5-Methylcytosine in eukaryotic DNA. *Science* 1981; **212**: 1350–1357.
56 Brown DD. Gene expression in eukaryotes. *Science* 1981; **211**: 667–674.
57 Razin A, Riggs AD. DNA methylation and gene function. *Science* 1980; **210**: 604–610.
58 Schwartz S, Ast G. Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. *EMBO J* 2010; **5**: 578–582.
59 Sims 3rd RJ, Millhouse S, Chen CF, Lewis BA, Erdjument-Bromage H, Tempst P *et al.* Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* 2007; **28**: 665–676.
60 Schor IE, Rascovan N, Pelisch F, Allo M, Kornblihtt AR. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci USA* 2009; **106**: 4325–4330.
61 Hubbard RE, O'mahony MS, Calver BL, Woodhouse KW. Plasma esterases and inflammation in ageing and frailty. *Eur J Clin Pharmacol* 2008; **64**: 895–900.
62 Matigian N, Windus L, Smith H, Filippich C, Pantelis C, Mcgrath J *et al.* Expression profiling in monozygotic twins discordant for bipolar disorder reveals dysregulation of the WNT signalling pathway. *Mol Psychiatry* 2007; **12**: 815–825.
63 Kapczinski F, Dal-Pizzol F, Teixeira AL, Magalhaes PV, Kauer-Sant'anna M, Klamt F *et al.* Peripheral biomarkers and illness activity in bipolar disorder. *J Psychiatr Res* 2010; **45**: 156–161.
64 Laan W, Grobbee DE, Selten JP, Heijnen CJ, Kahn RS, Burger H. Adjuvant aspirin therapy reduces symptoms of schizophrenia spectrum disorders: results from a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry* 2010; **71**: 520–527.
65 Monji A, Kato T, Kanba S. Cytokines and schizophrenia: microglia hypothesis of schizophrenia. *Psychiatry Clin Neurosci* 2009; **63**: 257–265.
66 Goldstein BI, Kemp DE, Soczynska JK, Mcintyre RS. Inflammation and the phenomenology, pathophysiology, comorbidity, and treatment of bipolar disorder: a systematic review of the literature. *J Clin Psychiatry* 2009; **70**: 1078–1090.
67 Sugino H, Futamura T, Mitsumoto Y, Maeda K, Marunaka Y. Atypical antipsychotics suppress production of proinflammatory cytokines and up-regulate interleukin-10 in lipopolysaccharide-treated mice. *Prog Neuropsychopharmacol Biol Psychiatry* 2009; **33**: 303–307.
68 Muller N. Inflammation and the glutamate system in schizophrenia: implications for therapeutic targets and drug development. *Expert Opin Ther Targets* 2008; **12**: 1497–1507.
69 Chang H, Jeung HC, Jung JJ, Kim TS, Rha SY, Chung HC. Identification of genes associated with chemosensitivity to SAHA/taxane combination treatment in taxane-resistant breast cancer cells. *Breast Cancer Res Treat* 2011; **125**: 55–63.
70 Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; **324**: 930–935.
71 Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009; **324**: 929–930.
72 Huang, Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* 2010; **5**: e8888.
73 Zheng Y, Cohen-Karni D, Xu D, Chin HG, Wilson G, Pradhan S *et al.* A unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Res* 2010; **38**: 5527–5534.
74 Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 2010; **38**: e181.
75 Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* 2010; **38**: e125.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)