GENES AND SCHIZOPHRENIA

Epigenetic Studies of Schizophrenia: Progress, Predicaments, and Promises for the Future

Emma Dempster1, Joana Viana2, Ruth Pidsley2, and Jonathan Mill*,1,2
1Exeter University Medical School, University of Exeter, UK; 2Institute of Psychiatry, King’s College London, UK
*To whom correspondence should be addressed; Exeter University Medical School, St Luke’s Campus, Exeter, EX1 2LU, UK; tel: +44(0)1392-726430, fax: +44(0)1392-722926, e-mail: j.mill@exeter.ac.uk

Increased understanding about the functional complexity of the genome has led to growing recognition about the role of epigenetic variation in the etiology of schizophrenia. Epigenetic processes act to dynamically control gene expression independently of DNA sequence variation and are known to regulate key neurobiological and cognitive processes in the brain. To date, our knowledge about the role of epigenetic processes in schizophrenia is limited and based on analyses of small numbers of samples obtained from a range of different cell and tissue types. Moving forward, it will be important to establish cause and effect in epigenetic studies of schizophrenia and broaden our horizons beyond DNA methylation. Rather than investigating genetic and epigenetic factors independently, an integrative etiological research paradigm based on the combination of genomic, transcriptomic, and epigenomic analyses is required.

Key words: schizophrenia/epigenetics/DNA methylation/genetics/epidemiology

Introduction

Although the last decade has witnessed tremendous advances in our understanding about the genetic basis of schizophrenia (SZ), a large amount of the variance in disease risk remains to be explained. Increased understanding about the functional complexity of the genome has led to growing recognition about the role of non-sequence-based variation in health and disease.1 Epigenetic processes, which developmentally regulate gene expression via modifications to DNA, histone proteins, and chromatin, have been widely heralded as the “missing piece” of the etiological puzzle for a spectrum of complex disease phenotypes including SZ.2,3 Caution is warranted, however, as empirical studies are in their infancy and the optimal strategies for epigenetic epidemiology are still being developed.4 This article will briefly introduce epigenetic mechanisms and their relevance to SZ before focusing on future directions for this emerging field of research.

The Epigenetic Regulation of Gene Expression

Sequencing the genome was only the first step in our quest to understand how genes are expressed and functionally regulated. Sitting above the DNA sequence is a second layer of information, the “epigenome,” which mediates mitotically heritable, but reversible, changes in gene expression.5 These processes are essential for normal cellular development and differentiation and allow the regulation of gene function through nonmutagenic mechanisms. In contrast to the DNA sequence, which is stable and strongly conserved, epigenetic processes can be both highly dynamic and potentially modified by external environmental factors.6 The primary focus of research in the context of health and disease has been on DNA methylation at CpG dinucleotides. DNA methylation is the most well characterized and stable epigenetic modification, influencing gene expression via the disruption of transcription factor binding and the attraction of methyl-binding proteins that initiate chromatin compaction and gene silencing.

Evidence for a Role of Epigenetic Processes in Schizophrenia

Epigenetic processes are known to regulate key neurobiological and cognitive processes in the brain. The importance of epigenetic mechanisms in “normal” brain development and function is exemplified by the neurodevelopmental deficits associated with mutations in methyl-CpG binding protein 2 (MECP2), whose product interacts with methylated CpGs to control neuronal gene expression.7 Several epidemiological, clinical, and molecular features associated with SZ
implicate an epigenetic contribution to etiology. For example, despite sharing the same DNA sequence the concordance rate between monozygotic (MZ) twins for SZ is consistently estimated to be <65%, potentially reflecting environmentally or stochastically mediated epigenetic variation. The most replicated environmental risks for SZ occur at critical periods early in development, a time of rapid cell replication when the epigenome is known to be particularly labile in response to external factors and the standard epigenic signals driving development and tissue differentiation are being established. For example, prenatal exposure to famine has been robustly associated with SZ and altered DNA methylation in epidemiological studies, suggesting that epigenetic pathways may mediate the link between the prenatal environment and the development of neuropsychiatric disease. Results of genetic studies also highlight a potential role for epigenetic mechanisms in SZ; a recent genome-wide association study (GWAS) concluded that variants in a cluster of core histone genes are associated with SZ. Finally, many of the currently prescribed treatments for SZ are associated with epigenomic changes, further implicating epigenetic processes in disease etiology.

Epigenetic Studies of Schizophrenia: The Current State of Play

Despite considerable speculation about the role of epigenetics in SZ, this is a relatively nascent area of research. Compared to other disorders such as cancer, where an epigenetic contribution to disease is well established, few empirical data are available to support the epigenetic theory of SZ. Although publications in the field are growing at an exponential rate (figure 1), a considerable proportion of these represent non-research articles. An overview of the current findings from studies of DNA methylation in SZ is given in table 1; these findings are described in more detail in several recent review articles. Here, we focus more broadly on issues pertaining to the interpretation of epigenetic studies in SZ, exploring the caveats, problems, and promises that characterize this rapidly evolving field.

Our knowledge about the role of epigenetic processes in SZ is garnered from analyses of small numbers of samples obtained from a range of different cell and tissue types. These studies have primarily focused on only one epigenetic modification (ie, DNA methylation) across very specific genomic regions (ie, promoter CpG islands associated with a priori candidate genes). A plethora of different methodological approaches have been used to interrogate epigenetic variation, although there is little validation of disease-associated differences using alternative approaches or replication in additional samples. The field is somewhat reminiscent of the early days of psychiatric genetics, although given the added biological, technological, and methodological issues pertaining to epigenetic epidemiology, many of the current results may be even harder to interpret than the often-contradictory, underpowered candidate gene–association studies that preceded GWAS.

Recent advances in technology mean that genome-scale studies of the epigenome across much larger sample collections are now feasible, particularly for DNA methylation, and a number of genome-wide association studies (EWAS) for SZ have been performed or are currently underway. For instance, a study using frontal cortex brain tissue to assess DNA methylation across ~12 000 regulatory regions of the genome uncovered DNA methylation differences in the vicinity of multiple loci, including several involved in glutamatergic and gamma aminobutyric acid–releasing (GABAergic) neurotransmission, brain development, and other processes functionally linked to disease etiology. Recently, we published the first GWAS of DNA methylation differences between MZ twins discordant for SZ, observing considerable disease-associated variation and a significant enrichment of epigenetic disruption in biological networks and pathways relevant to psychiatric disease and neurodevelopment. Although these findings provide some support for the role of epigenetic dysfunction in SZ, they represent a fairly crude initial snapshot of the regulatory changes involved in pathogenesis and indicate that a more systematic investigation across much larger numbers of samples is warranted.

The Conundrum of Causality

The simple brute-force “science by numbers” approach that has been successfully employed in SZ GWAS analyses is not directly translatable to epigenetic epidemiology.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Tissue Source</th>
<th>Sample Number</th>
<th>Method</th>
<th>Loci Assessed</th>
<th>Summary of Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Frontal cortex [BA 9 and 10]</td>
<td>5 SZ, 5 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>RELN</td>
<td>SZ-associated hypermethylation at RELN promoter in frontal cortex</td>
</tr>
<tr>
<td>14</td>
<td>Frontal cortex [BA 46]</td>
<td>35 SZ, 35 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>COMT</td>
<td>SZ-associated hypomethylation at COMT promoter in tissue samples from left side of brain</td>
</tr>
<tr>
<td></td>
<td>Frontal lobe [BA 9 and 10]</td>
<td>10 SZ, 10 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>HTR2A</td>
<td>SZ-associated hypermethylation at HTR2A promoter</td>
</tr>
<tr>
<td>15</td>
<td>Frontal cortex [BA 46]</td>
<td>35 SZ, 35 controls</td>
<td>High-resolution melting assay</td>
<td>5HTR1A</td>
<td>Hypermethylation in SZ</td>
</tr>
<tr>
<td>16</td>
<td>Peripheral blood leukocytes</td>
<td>40 SZ, 67 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>MAOA</td>
<td>Significant hypermethylation at individual CpG sites in MAOA promoter in male but not female patients</td>
</tr>
<tr>
<td>17</td>
<td>Peripheral blood cells</td>
<td>371 SZ, 288 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>COMT</td>
<td>No significant difference in COMT promoter methylation between SZ and controls</td>
</tr>
<tr>
<td>18</td>
<td>Cerebellum</td>
<td>15 SZ, 15 controls</td>
<td>Pyrosequencing of bisulfite-PCR amplicons</td>
<td>COMT</td>
<td>No significant difference in COMT promoter methylation between SZ and controls</td>
</tr>
<tr>
<td>19</td>
<td>Peripheral blood</td>
<td>22 twin pairs discordant for SZ</td>
<td>Illumina 27K HumanMethylation microarray</td>
<td>~27 000 CpG genome-wide sites</td>
<td>Numerous loci with DNA methylation differences between discordant twins</td>
</tr>
<tr>
<td>20</td>
<td>Saliva</td>
<td>63 SZ, 76 controls; first-degree relatives: 15 SZ</td>
<td>MSP and bisulfite sequencing</td>
<td>HTR2A</td>
<td>T102C polymorphic site significantly hypomethylated in patients and their first-degree relatives compared to controls</td>
</tr>
<tr>
<td>21</td>
<td>Occipital cortex</td>
<td>10 SZ, 10 controls</td>
<td>Direct sequencing of bisulfite-PCR amplicons</td>
<td>RELN</td>
<td>Significant hypermethylation of 4 CpG sites in the RELN promoter in SZ</td>
</tr>
<tr>
<td>22</td>
<td>Prefrontal cortex [BA 9 and 10]</td>
<td>5 SZ, 5 controls</td>
<td>Clonal sequencing of bisulfite-PCR amplicons</td>
<td>SOX10</td>
<td>SZ-associated hypermethylation and reduced expression of SOX10</td>
</tr>
<tr>
<td>23</td>
<td>Postmortem prefrontal cortex (BA10)</td>
<td>11 SZ, 12 controls</td>
<td>Pyrosequencing of bisulfite-PCR amplicons</td>
<td>S-COMT</td>
<td>Significant relationship between COMT promoter-region methylation, physical activity, and metabolic syndrome in 158Val/Met patients</td>
</tr>
<tr>
<td>24</td>
<td>Whole blood</td>
<td>85 SZ</td>
<td>Luminometric methylation assay (LUMA) Global DNA methylation</td>
<td>S-COMT 5-HTT</td>
<td>Significant SZ-associated global hypomethylation in SZ with medication and age effects</td>
</tr>
<tr>
<td>25</td>
<td>Peripheral blood leukocytes</td>
<td>177 SZ, 171 controls</td>
<td>Pyrosequencing of bisulfite-PCR amplicons</td>
<td>~12 000 CpG genome-wide islands</td>
<td>S-COMT hypermethylated in SZ</td>
</tr>
<tr>
<td>26</td>
<td>Frontal cortex</td>
<td>35 SZ, 35 controls</td>
<td>Enrichment of unmethylated DNA and hybridization to CpG island microarrays</td>
<td>~12 000 CpG genome-wide islands</td>
<td>Significant differences in DNA methylation at multiple loci</td>
</tr>
<tr>
<td>27</td>
<td>Saliva</td>
<td>20 SZ, 20 unaffected family members, 25 controls</td>
<td>MSP and bisulfite sequencing</td>
<td>MB-COMT</td>
<td>SZ-associated hypomethylation of MB-COMT promoter</td>
</tr>
<tr>
<td>28</td>
<td>Peripheral blood lymphocytes</td>
<td>1 MZ pair concordant, 1 MZ pair discordant</td>
<td>Clonal sequencing of bisulfite-PCR amplicons</td>
<td>DRD2</td>
<td>SZ twin in pair discordant for SZ had DRD2 methylation levels “closer” to the affected concordant twin pair than to its unaffected co-twin</td>
</tr>
<tr>
<td>29</td>
<td>Prefrontal cortex [BA 10]</td>
<td>15 SZ, 15 controls</td>
<td>Pyrosequencing of bisulfite-PCR amplicons</td>
<td>RELN</td>
<td>No significant differences</td>
</tr>
<tr>
<td>30</td>
<td>Both hemispheres of the superior temporal gyrus, parahippocampus gyrus, and cingulate gyrus</td>
<td>293 SZ, 340 controls</td>
<td>Clonal sequencing of bisulfite-PCR amplicons</td>
<td>FOXP2</td>
<td>Hypermethylation in the left parahippocampus gyrus in SZ patients, opposite patterns in controls</td>
</tr>
<tr>
<td>31</td>
<td>Peripheral blood leukocytes</td>
<td>48 same-sex discordant sibling pairs</td>
<td>MSP and direct bisulfite sequencing</td>
<td>DRD2</td>
<td>No significant differences</td>
</tr>
</tbody>
</table>

*Note:* 5-HTT, 5-hydroxy-tryptamine transporter; BA, Brodmann area; COMT, catechol-O-methyltransferase; DRD2, dopamine D2 receptor; FOXP2, Forkhead box protein P2; HTR1A, 5-hydroxytryptamine (serotonin) receptor 1A; HTR2A, 5-hydroxytryptamine (serotonin) receptor 2A; LUMA, luminometric methylation assay; MAOA, monoamine oxidase A; MSP, methylation-specific PCR; MZ, monozygotic twins; PCR, polymerase chain reaction; RELN, reelin; SOX10, SRY (sex determining region Y)-box 10; SZ, schizophrenia.
Instead, more conventional epidemiological approaches that consider a spectrum of confounding factors, including tissue/cell type, age, sex, medication exposure, and reverse causation, have to be used.

One of the major challenges in studying epigenetic changes in neuropsychiatric disorders is the availability of samples of the primary target tissue, i.e., the brain. Because it is not yet possible to perform in vivo epigenetic studies in the brain, only retrospective study designs using postmortem brain samples are viable. These experiments are limited by access to high-quality, well-phenotyped samples. There are a number of potential confounders that could influence epigenetic analyses of such samples, including postmortem interval, storage time, pH, and cellular heterogeneity. The small number of brain samples available for SZ epigenetic research means that many investigations are limited in terms of their power to detect significant associations, especially given the small absolute changes that are likely to be uncovered and the heterogeneous cellular composition of the cerebral cortex. Ultimately, given the cell-specific differences identified in epigenetic gene regulation, it may be optimal to isolate specific cell types (e.g., neurons, glia, and astrocytes) from brain tissue via processes such as laser capture microdissection.

In reality, it may not be feasible to undertake adequately powered studies of SZ using postmortem brain material, especially if numbers approaching those used in GWAS analyses are required. A key question in epigenetic epidemiology, particularly in regard to the study of neuropsychiatric disease, therefore concerns the extent to which easily accessible peripheral tissues such as blood can be used to ask questions about interindividual phenotypic variation manifest in inaccessible tissues such as the brain. We recently undertook the first within-individual cross-tissue (multiple brain regions plus blood) characterization of the DNA methylome. Although, as expected, distinct tissue-specific patterns of DNA methylation were identified and between-tissue variation was found to greatly exceed between-individual differences within any one tissue, it was striking that some interindividual variation was significantly correlated between blood and brain tissue, indicating that peripheral sources of DNA may have utility in epidemiological studies of complex neurobiological phenotypes. As shown in table 1, several studies have successfully used peripheral samples to identify significant disease-associated epigenetic variation in blood, although it remains important to establish whether these differences are also reflected in SZ-relevant regions of the brain.

Another important issue concerns the ability to distinguish between cause and effect; the disease-associated modifications described in table 1 may arise prior to illness and contribute directly to the disease phenotype or could alternatively be a secondary effect of the disease process or the medications used in treatment. It is clear that a range of factors, including genotype, age, and sex, are correlated with epigenetic variation, and so including these as covariates in analyses is important. The analysis of SZ-discordant MZ twins represents a powerful strategy for overcoming many of these issues as both twins have their genomes, parents, birth date, and gender in common and are likely to have been exposed to a highly similar pre- and perinatal environment. An optimal study design would involve the longitudinal assessment of epigenetic changes in MZ twins, examining within-pair epigenomic variation as they become discordant for disease. Finally, because many psychotropic drugs have been shown to alter the epigenome, such longitudinal study designs would allow us to determine whether epigenetic changes are independent of, or are mediated by, medication.

Experimental Approaches and Limitations

Most studies of epigenetic processes in SZ have focused on DNA methylation. There is, however, a diverse range of enrichment strategies and mapping platforms that have been employed to assess DNA methylation in SZ. This has made comparisons between studies difficult, although the recently released Illumina 450K Methylation Beadchip that is being widely adopted for EWAS analyses offers an economical balance between coverage and precision. The majority of probes on this array, however, are still located in CpG-rich promoters and may not be optimal for identifying the most phenotypically relevant epigenetic variation. Recent studies highlight the importance of DNA methylation outside of promoter CpG islands, finding functionally relevant epigenomic variation primarily at nonpromoter CpG islands, low CG-content promoters, and the gene body. With the cost of whole-genome bisulfite sequencing on a rapid downward trajectory, the era of whole-methylome analyses of disorders such as SZ is not far away. A number of additional DNA modifications (5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxylcytosine) have recently received considerable attention. For example, 5-hydroxymethylcytosine is believed to result from the active demethylation of methylated cytosine and appears to be abundant in brain tissue, warranting further investigation in the context of neuropsychiatric phenotypes. Posttranslational histone modifications are another major source of epigenetic regulation, which have been largely neglected in epidemiologically informative study designs of SZ.

Finally, rather than investigating genetic and epigenetic factors independently, it is clear that an integrative etiological research paradigm is required. It has been demonstrated, e.g., that DNA sequence variation can directly influence DNA methylation in cis and evidence has been found for an enrichment of methylation quantitative trait loci (mQTLs) amongst psychosis-nominated GWAS loci. The interpretation of GWAS data could be improved by incorporating such “epiallelic” information into analyses; although mQTLs provide a functional mechanism for apparently noncoding genetic variation,
other epigenetic patterns may complicate the direct identification of disease-associated loci, contributing toward the “missing heritability” of complex diseases by masking direct associations between genotype and phenotype. Because epigenetic processes are influenced by a spectrum of external environmental factors, including diet, toxins, drugs, and stress, the observation that polymorphisms can also exert an effect on gene function via epigenetic processes occurring in cis suggests a common pathway behind both genetic and environmental effects and a potential mechanism for gene–environment interaction.

### Into the Future

Technological advances in epigenomic profiling methodologies mean that it will soon be feasible to economically map the epigenome at single base-pair resolution in large cohorts of samples. Moving forward, it will be important to establish cause and effect in epigenetic studies of SZ and broaden our horizons beyond DNA methylation. Furthermore, maximum information will be obtained from studies integrating epigenomic information with genomic, transcriptomic, and proteomic data obtained from the same samples. Ultimately, the dynamic and reversible nature of epigenetic processes holds huge promise for the development of novel diagnostic and therapeutic measures for SZ.

### References


