

Decoding the non-coding genome: elucidating genetic risk outside the coding genome

C. L. Barr^{†,‡,*} and V. L. Misener[†]

[†]Toronto Western Research Institute, University Health Network, Toronto, ON, Canada

[‡]Program in Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON, Canada

Abstract

Current evidence emerging from genome-wide association studies indicates that the genetic underpinnings of complex traits are likely attributable to genetic variation that changes gene expression, rather than (or in combination with) variation that changes protein-coding sequences. This is particularly compelling with respect to psychiatric disorders, as genetic changes in regulatory regions may result in differential transcriptional responses to developmental cues and environmental/psychosocial stressors. Until recently, however, the link between transcriptional regulation and psychiatric genetic risk has been understudied. Multiple obstacles have contributed to the paucity of research in this area, including challenges in identifying the positions of remote (distal from the promoter) regulatory elements (e.g. enhancers) and their target genes and the under-representation of neural cell types and brain tissues in epigenome projects – the availability of high-quality brain tissues for epigenetic and transcriptome profiling, particularly for the adolescent and developing brain, has been limited. Further challenges have arisen in the prediction and testing of the functional impact of DNA variation with respect to multiple aspects of transcriptional control, including regulatory-element interaction (e.g. between enhancers and promoters), transcription factor binding and DNA methylation. Further, the brain has uncommon DNA-methylation marks with unique genomic distributions not found in other tissues – current evidence suggests the involvement of non-CG methylation and 5-hydroxymethylation in neurodevelopmental processes but much remains unknown. We review here knowledge gaps as well as both technological and resource obstacles that will need to be overcome in order to elucidate the involvement of brain-relevant gene-regulatory variants in genetic risk for psychiatric disorders.

Keywords

5-Hydroxymethylation; adolescent brain; enhancers; epigenetics; gene expression; gene regulation; methylation; non-CG methylation; non-coding RNA; psychiatric disorders; transcriptome

*Corresponding author: Dr C. L. Barr, The Toronto Western Research Institute, 60 Leonard Avenue, Krembil Discovery Tower 8KD-412, Toronto, Ontario M5T 2S8, Canada. CBarr@uhnres.utoronto.ca.

Changes in gene expression are now acknowledged as a major contributor to the genetic risk for complex genetic traits. In this article, we review the key evidence supporting this, as well as mechanisms by which DNA variation influences gene expression and methodologies for studying these relationships. We note that a comprehensive review of the genetic findings for psychiatric disorders is beyond the scope of this review, as is a comprehensive review of epigenetic regulation. Instead, we focus this review on classes of transcripts and epigenetic marks that are prevalent in brain and that may make unique contributions to brain function. We also focus on methods that, in our opinion, are likely to make the greatest contributions to moving the field forward. Finally, we outline a number of obstacles that currently stand in the way of fully elucidating the relationships between gene expression and DNA variation underlying psychiatric disorders.

Transcriptional regulation of gene expression in genetic disease

Gene-regulatory elements controlling gene transcription include the promoter, enhancers, silencers, locus control regions and insulators. With the exception of the promoter that is located in the immediate vicinity of the transcription start site of the gene (Butler & Kadonaga 2002), the other elements can be located anywhere within a gene or even up to megabases away (Dean 2006; Nobrega *et al.* 2003; Spilianakis & Flavell 2004; Steidl *et al.* 2007), often within neighboring genes (Lettice *et al.* 2003). Genetic variation altering gene expression is a documented cause of genetic disease, and until recently, the majority of risk alleles studied were in the promoter, as this region is easily defined. However, genetic changes (duplications, deletions, single nucleotide changes and translocations) in remote regulatory elements – i.e. outside of the proximal promoter – resulting in altered gene expression have also been documented as disease mechanisms, both in Mendelian disorders (e.g. thalassemias, X-linked deafness and facioscapulohumeral muscular dystrophy) and in complex genetic disorders (e.g. cancer, diabetes, rheumatoid arthritis and systemic lupus erythematosus) (Dathe *et al.* 2009; De Gobbi *et al.* 2006; Driscoll *et al.* 1989; Furniss *et al.* 2008; Gabellini *et al.* 2002; Hatton *et al.* 1990; Lettice *et al.* 2003; Loots *et al.* 2005; Naranjo *et al.* 2010; Prokunina *et al.* 2002; Sharpe *et al.* 1992; Steidl *et al.* 2007; Sun *et al.* 2008; Tokuhira *et al.* 2003).

Recent results from genome-wide association studies (GWAS) provide further evidence that variation in gene expression contributes to genetic risk. Genome-wide association studies have recently been completed for a number of psychiatric disorders, including bipolar disorder (Ferreira *et al.* 2008; Green *et al.* 2013; Muhleisen *et al.* 2014; Psychiatric GWAS Consortium Bipolar Disorder Working Group 2011), attention-deficit hyperactivity disorder (ADHD) (Elia *et al.* 2011; Franke *et al.* 2009; Hinney *et al.* 2011; Lasky-Su *et al.* 2008, 2010; Lesch *et al.* 2008; Neale *et al.* 2008, 2010a; Stergiakouli *et al.* 2012), schizophrenia (Hamshere *et al.* 2013; O'Donovan *et al.* 2008, 2009; Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014), autism (Anney *et al.* 2012; Casey *et al.* 2012), major depression (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium 2013; The Psychiatric GWAS Consortium Steering Committee 2009), Tourette syndrome (Scharf *et al.* 2013) and obsessive-compulsive disorder (OCD) (Mattheisen *et al.* 2015; Stewart *et al.* 2013), as well as for adverse effects of therapeutic agents, including

antipsychotic-induced weight gain (Malhotra *et al.* 2012). While the majority of the early studies did not meet the significance threshold for genome-wide evidence of association, studies with larger sample sizes are now providing such evidence, with some of the findings replicating across samples and some showing evidence of shared signals across disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.* 2013).

Compared with GWAS analyses of other complex traits that have examined over 100 000 cases, investigations of psychiatric disease using GWAS are in their infancy. This is especially true for the majority of childhood-onset psychiatric disorders, for which GWAS sample sizes have been relatively small, numbering only a few thousand subjects (Anney *et al.* 2012; Neale *et al.* 2010b; Scharf *et al.* 2013). Nevertheless, with GWAS sample sizes for some psychiatric disorders now reaching the necessary numbers, replicated GWAS signals have emerged. Of note, the recent GWAS of 36 989 schizophrenia cases identified 108 GWAS-significant regions (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014).

Moving forward to identify the functional DNA variants (risk alleles) tagged by disorder-associated GWAS signals is now the focus of intense investigation, as we enter what has been coined as the post-GWAS era. In many instances, no change in the protein-coding region could be identified within the associated gene; in other instances, coding region changes were identified but could not fully explain the association (Dwyer *et al.* 2010; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Some associated markers are located at large distances (>10–100 kb) from an obvious candidate gene (Malhotra *et al.* 2012; Wang *et al.* 2010); others are located in gene ‘deserts’, with no identifiable genes within kilobases or even megabases of the positive markers (Franke *et al.* 2009; Moraes *et al.* 2012; Stewart *et al.* 2013; Zandi *et al.* 2007). Of note, only 11 of the 108 schizophrenia loci identified in the recent GWAS could be explained by a coding region single nucleotide polymorphism (SNP) (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Further, direct screening of a number of key, replicated candidate genes for psychiatric disorders could not identify coding region changes that would account for the association (e.g. *ZNF804A*; Dwyer *et al.* 2010). It appears likely that the vast majority of risk variants underlying psychiatric disorders will turn out to be located in non-coding regions of the genome.

That the landscape of genetic risk for psychiatric disorders will be largely non-coding aligns with overwhelming evidence that changes in gene-regulatory regions are a major contributor to risk for complex genetic traits: The majority (~93%) of disease-associated markers emerging from GWAS findings lie within the non-coding genome (Andersson *et al.* 2014; Farh *et al.* 2014; Maurano *et al.* 2012). The GWAS signals are enriched for expression quantitative trait loci (eQTL), i.e. loci that correlate with gene expression (Ellinghaus *et al.* 2012; Franke *et al.* 2010; Khor *et al.* 2011; Nicolae *et al.* 2010; Ramasamy *et al.* 2014; Zhong *et al.* 2010), and are enriched for markers located within gene-regulatory regions (Farh *et al.* 2014; International Genetics of Ankylosing Spondylitis Consortium *et al.* 2013; Maurano *et al.* 2012; Ricano-Ponce & Wijmenga 2013; Roadmap Epigenomics Consortium *et al.* 2015; Schaub *et al.* 2012).

Enhancers and genetic risk

Among the various types of regulatory elements in the genome, GWAS results for the majority of complex traits point increasingly to variants in one type in particular – namely, enhancers (Andersson *et al.* 2014; Ernst *et al.* 2011; Farh *et al.* 2014; Hnisz *et al.* 2013; Ward & Kellis 2012). Enhancers can reside anywhere within a gene and can even be millions of nucleotides distant from the genes they regulate, i.e. their target genes (Leveille *et al.* 2015). Enhancer activation, resulting in stimulation of target-gene transcription, involves the binding of transcription factors (TFs; proteins specialized for transcriptional regulation) to the enhancer sequence, modifying the epigenetic environment, with the enhancer looping to physically contact the target-gene promoter. Importantly, enhancers are the major genomic determinant of cell-type/tissue-specific gene expression (Jin *et al.* 2011). Further, DNA variants associated with a specific disease are more frequently found in enhancers specific to cell types relevant to that disease (Gerasimova *et al.* 2013), a finding that supports their contribution to disease risk.

Another important property of enhancers is that their target genes are typically not the nearest gene (only 27–40% target the nearest gene) (Andersson *et al.* 2014; Sanyal *et al.* 2012). This could explain GWAS signals in regions where the nearest gene(s) in the associated region are not likely candidates for involvement in the trait or disorder being investigated. Further, a single enhancer can regulate multiple genes (on average, 2.5) (Andersson *et al.* 2014; Sanyal *et al.* 2012). Thus, a genetic variant in an enhancer could have a profound impact on disease risk by dysregulating more than one gene (Verlaan *et al.* 2009). The identification of such multiple-gene targets will be critical to the subsequent modeling of disease mechanisms, as knockout or knockdown models of a single gene may not recapitulate the disease phenotype.

Despite the wealth of evidence that many GWAS signals could be pointing to remote regulatory elements in general, and enhancers in particular, a number of obstacles have stood in the way of easily moving forward to pinpoint the enhancers involved and identify the risk alleles. Systematic and comprehensive functional screening of the typically large genomic regions flagged by GWAS signals is prohibitively labor-intensive using standard reporter assays. On the other hand, functional testing approaches that assay only a randomly selected subset of DNA fragments spanning a large region of interest may miss the sought-after element(s). Multispecies comparisons have been used successfully to identify regulatory regions on the basis of sequence conservation, but principally for a restricted subset of genes (those expressed early in development with strict constraints on temporal and spatial expression) (Boffelli *et al.* 2004; Nobrega *et al.* 2003; Pennacchio *et al.* 2006); for other genes with only short stretches of conserved sequence, this approach may miss key regions. Further, conservation of regulatory region function can be maintained even without conservation of sequence, due to compensatory changes at other sites (Ludwig *et al.* 2000), and conversely, conservation of sequence may be maintained without conservation of function (Shen *et al.* 2012). DNA sequence differences between species may also result in new functions conferring evolutionary advantages (Dermitzakis & Clark 2002; Rockman *et al.* 2003) – including sequence variants that are human-specific (King & Wilson 1975; Prabhakar *et al.* 2006). For example, comparison of known functional elements in humans

indicated that only 60–68% are functional in rodents (Dermitzakis & Clark 2002). In addition, non-conserved elements may be relevant to genes contributing to human higher order complex cognitive traits (e.g. reading and language ability, verbal working memory and executive function).

Mapping gene regulatory elements

Histone modification, movement and/or replacement play a dynamic role in gene regulation (Eberharter & Becker 2002; Higgs *et al.* 2006; Peters & Schubeler 2005), and specific histone modifications (e.g. H3 and H4 acetylation, methylation) mark regulatory elements (Bernstein *et al.* 2005; Schubeler *et al.* 2004). This property of histones can be leveraged to locate regulatory regions, focusing the search for risk alleles within these regions (Couto *et al.* 2010; Darabi *et al.* 2015; Elbert *et al.* 2011; Heintzman *et al.* 2007; Ni *et al.* 2008). By chemically cross-linking DNA to histones, antibodies specific for histone modifications can be used to capture the modified-histone-associated DNA regions; coupling this chromatin immunoprecipitation (ChIP) approach with DNA identification methods, such as genomic tiling arrays (ChIP-chip) (Ren *et al.* 2000) or next-generation sequencing (ChIP-seq) (Barski *et al.* 2007; Johnson *et al.* 2007; Mikkelsen *et al.* 2007; Robertson *et al.* 2007), flags modified chromatin across genomes, creating genome-wide maps of regulatory regions. Comparative analyses found that the locations of sites marked by methylated histones are conserved between mouse and human genomes even though most of the genomic sequences at those sites are not (Bernstein *et al.* 2005). Thus, modified chromatin identifies regulatory regions where function, but not necessarily sequence, is conserved.

The ChIP-seq data, most notably from the ENCODE project and Epigenomic Roadmap (Bernstein *et al.* 2010, 2012; Kim *et al.* 2007; Roadmap Epigenomics Consortium *et al.* 2015), have recently resulted in an explosion of information on the genomic positions of modified histones and other epigenomic features, as markers for regulatory elements. Much of this information is now publicly available on websites such as UCSC (<https://genome.ucsc.edu/cgi-bin/hgGateway>) and through the Epigenome Road-map (<http://www.roadmapepigenomics.org>).

The available ChIP-seq maps of regulatory regions allow us to begin studies of these regions in relation to disease; even so, the maps are – for the study of psychiatric disease – still quite limited in two major respects. First, as noted earlier, there is considerable tissue/cell-type specificity inherent in transcriptional regulation. Thus, the applicability of the ChIP-seq information obtained is limited by the cell/tissue types examined, and unfortunately, little information is available on cell types relevant to the nervous system. Early on, there was little data in ENCODE for human neural cell types or human brain tissue, and this tissue is still underrepresented, given that brain tissue has high levels of gene expression with the majority of annotated genes expressed in the brain (Datson *et al.* 2001; Miller *et al.* 2014). For example, on the basis of microarray data, it was estimated that 84% of Reference Sequence (RefSeq) genes are expressed in adult neocortex (Hawrylycz *et al.* 2012) and 95% in the developing neocortex (Miller *et al.* 2014). Second, brain-relevant maps that have been produced are far from complete. The Epigenomic Roadmap has mapped histone modifications (both acetylations and methylations) in brain tissues from only one adult

individual (eight brain regions) and from fetal brain at only a few gestational time points (Roadmap Epigenomics Consortium *et al.* 2015). Until recently, the majority of histone-modification data available were for H3K4me3, a mark of promoters, with little information on remote regulatory regions. However, the number of epigenomic marks has now expanded to include six additional histone marks, including those that mark enhancers (Roadmap Epigenomics Consortium *et al.* 2015), a start to providing more information on the epigenome of brain.

Functional studies of enhancers

Impact of DNA variants on enhancer function

The mapping of gene-regulatory regions allows us to correlate the positions of genetic association signals with the positions of regulatory elements. However, the data do not provide information on the actual function of specific alleles within these elements. One cannot assume simply on the basis of location that a SNP within a regulatory region changes gene expression – functional studies will still be required. Here again, the cell/tissue-type specificity of transcriptional regulation is an important consideration. Function can only be tested reliably in cells/tissues that have the ‘correct’ components/configuration of transcriptional machinery. High-throughput functional screening of putative regulatory regions can be performed using animal models that test expression in all tissues (Pennacchio *et al.* 2006), but this is expensive. Further, regulatory elements functional in humans may not be functional in rodents (Dermitzakis & Clark 2002). Thus, functional studies with appropriate disease-relevant human cell types are required to show the impact of DNA variation on transcriptional regulatory function.

High-throughput enhancer testing

Standard methods using *in vitro* transfection assays can be used to test individual regulatory regions and the impact of DNA variation on transcription. However, even when the locale has been narrowed down (through overlay of the ChIP-seq and GWAS maps), both the construction of vectors and the assays themselves are time-consuming and tedious to perform for each individual gene/regulatory region/DNA-variant combination. A recent breakthrough in molecular biology has been the development of high-throughput methods to screen for enhancer function using next-generation sequencing (Arnold *et al.* 2013; Melnikov *et al.* 2012; Patwardhan *et al.* 2012; Vanhille *et al.* 2015). These protocols make use of ‘bar-coded’ transfection vectors that allow for the transfection of cells with hundreds to thousands of enhancers in separate reporter vectors – the different bar codes are used to tag the individual enhancers and quantify the expression levels generated from each of the corresponding vectors.

Super-enhancers in genetic disease

Despite the methodological breakthrough described above, there is a recently identified class of enhancers, termed stretch or super-enhancers (Hnisz *et al.* 2013; Parker *et al.* 2013; Whyte *et al.* 2013), that poses additional challenges for study. Super-enhancers are clusters of enhancers larger (>3 kb) than typical enhancers, some spanning as much as 50 kb (Whyte *et al.* 2013), and they are associated with key genes that control cell identity. Disease-

associated alleles are enriched in super-enhancers (Hnisz *et al.* 2013) and tend to occur in the super-enhancers of disease-relevant cell types, more so than in the typical enhancers (Hnisz *et al.* 2013). Unfortunately, however, unlike typical enhancers, the super-enhancers cannot be tested for function in their entirety – generally, they are too big to be efficiently cloned into traditional test vectors and transfected into cells. Given their key role in cell identity and their association with risk alleles, alternative methods to test for functional effects of genetic variation in super-enhancers will be required such as genome editing methods (CRISPR/Cas9) (Charpentier & Doudna 2013; Jinek *et al.* 2012; Ran *et al.* 2013). Genome editing is a powerful tool for the study of putative functional variants; however, the availability of neural cell types for these studies will be a limitation for psychiatric disease.

DNA variation and TF binding

The transcriptional activity brought about by enhancers depends, fundamentally, on the concerted activities of TFs –DNA-binding proteins that orchestrate the appropriate selection and activation of cell-type-specific and/or signal-dependent enhancers to spawn new enhancer-promoter interaction loops or modulate existing ones, ultimately stimulating target-gene transcription (Heinz *et al.* 2015). Bioinformatics approaches can be used to predict the binding of a particular TF to a particular DNA sequence (Whitaker *et al.* 2015) and to predict the impact of DNA variation on that binding, but because of the short length and degenerate nature of consensus TF-binding sequences, such predictions are not precise (Schug 2008; Weirauch *et al.* 2013; Whitaker *et al.* 2015). Prediction methods that take the three-dimensional structures of the DNA motifs into account (in addition to their simple linear sequences) may help to improve the situation, but such methods have only begun to emerge and it remains to be seen whether they will be useful in predicting the impact of DNA variation on TF binding (Gao & Ruan 2015; Maienschein-Cline *et al.* 2012; Zhou *et al.* 2015).

A further complication when it comes to predicting the impact of motif variation on TF-binding activity is that TFs bind collaboratively, and the larger genomic context influences the probability of binding (Heinz *et al.* 2013, 2015). The presence of multiple homotypic TF-binding sites can buffer the loss of a single site (Kilpinen *et al.* 2013). Further, there can be interdependency of collaborating heterotypic sites: direct loss of TF binding (due to sequence variation) at one site can indirectly abolish binding of another TF to its site, without any direct change of sequence there (Heinz *et al.* 2013).

The limitations described above can be overcome, to some extent, using methods that assay TF binding directly. For example, the *in vitro* method of electrophoretic mobility shift assay (EMSA) has been used for decades; it is, however, cumbersome and limited by low throughput. Addressing this problem, more-recent, high-throughput *in vitro* methods that assay TF binding to immobilized oligonucleotides have been developed (e.g. using ELISA-based, bead-based and protein-binding microarray technologies) to automate screening for hundreds of DNA regions simultaneously (Berger & Bulyk 2009; Bhattacharya *et al.* 2010; Chorley *et al.* 2008; Weirauch *et al.* 2013). Further, the *in vivo* methods of ChIP – using TF-specific antibody – can be used to capture and map the genomic sequences to which a specified TF is bound in the cell type or tissue type of interest. Some of this *in vivo* TF-

binding information is available from the ENCODE and Epigenomic Roadmap projects. However, because of the issue of cell-type specificity described earlier and because of the possibility that the genotype(s) of the samples used may be influencing TF binding at some site(s), evidence of non-binding in TF-ChIP is not necessarily conclusive. Even when efforts to overcome these shortcomings are implemented, however (e.g. using appropriate cell/tissue types and genotypes), there may remain a technical obstacle: the availability of ChIP-grade TF-specific antibodies is currently limited.

Despite the drawbacks encountered in ascertaining the impacts of DNA variation on TF binding, the need to further develop and pursue this line of research in relation to psychiatric disease cannot be overstated – identifying disease-relevant transcriptional networks predicated on the coordinate actions of TFs at their cognate DNA binding sites will likely be the key to ultimately designing more-specific and effective therapeutic interventions. Highlighting this scenario is the recent example where the analysis of motifs predicted to be changed by GWAS alleles in regulatory regions across diverse autoimmune and inflammatory disorders pointed to a network of TFs in the JAK/STAT-mediated type I interferon response (Maurano *et al.* 2012); JAK inhibitors are currently showing promise in clinical trials for immune-mediated disorders (Garber 2011; Sandborn *et al.* 2012; Tanaka *et al.* 2012). Information regarding the networks of TFs and second messengers regulating risk genes may indicate direct targets for interventions. Thus, we will need to understand the TF networks that regulate risk genes underlying psychiatric disorders and how risk alleles perturb these networks.

Analysis of the transcriptome

Expression quantitative trait locus mapping

Within linkage disequilibrium (LD) regions, post-GWAS searches for functional risk alleles can be expedited by leveraging eQTL data. As noted earlier, eQTL are genetic markers that correlate with gene expression – thus presumably tagging functional variants altering gene expression. The production of genome-wide eQTL maps follows the same principle as GWAS, except that rather than assessing panels of genome-wide markers for associations with a single trait or disease/disorder, the markers are assessed for associations with the expression levels of RNA species in the transcriptomes of cells/tissues. Thus, eQTL markers identified in a disease-relevant cell/tissue type that are found in high LD with disease-associated markers from a GWAS can be extremely useful – they can help to focus on functional (gene-regulatory) risk alleles for the disease in question and can help to shed light on the identity of the target-gene ‘culprits’ whose dysregulation is causal to the disease (Darabi *et al.* 2015; Ramasamy *et al.* 2014; Zou *et al.* 2012).

Recently, there has been a great deal of progress in the availability of public databases containing transcriptome and genotype information (eQTL databases) (Colantuoni *et al.* 2011; Kang *et al.* 2011; Mele *et al.* 2015; Ramasamy *et al.* 2014; Zou *et al.* 2012); however, these resources are currently limited in terms of either the number of individuals represented or the developmental periods or brain regions covered. Early resources used microarrays [BrainCloud (Colantuoni *et al.* 2011), Brain eQTL Almanac (<http://ukbec.wordpress.com/>) (Ramasamy *et al.* 2014; Zou *et al.* 2012)]; thus, the data were limited to annotated genes

and, as a result, were missing unannotated exons and transcripts, particularly non-coding transcripts [e.g. long non-coding RNAs, small RNAs including microRNAs (miRNAs)] (Batista & Chang 2013; De Santa *et al.* 2010; Derrien *et al.* 2012; Kocerha *et al.* 2015; Orom & Shiekhhattar 2013; Orom *et al.* 2010; van de Leemput *et al.* 2014).

Non-coding RNA

Over the past several years, it has become increasingly apparent that the genome produces an abundance of non-coding transcripts (De Santa *et al.* 2010; Derrien *et al.* 2012; Kim *et al.* 2010; Mele *et al.* 2015; van de Leemput *et al.* 2014); non-coding RNAs now outnumber annotated coding RNAs (Derrien *et al.* 2012). Non-coding RNAs have been documented to downregulate (Batista & Chang 2013; Kocerha *et al.* 2015) as well as upregulate gene expression (Batista & Chang 2013; De Santa *et al.* 2010; Lai *et al.* 2013; Orom & Shiekhhattar 2013), with some species recently shown to function in an ‘enhancer-like’ fashion (Lai *et al.* 2013; Orom & Shiekhhattar 2013; Orom *et al.* 2010). Non-coding RNAs play key roles in neurogenesis, neurodevelopment and activity-dependent brain plasticity (Barry 2014; Kocerha *et al.* 2015; Lipovich *et al.* 2012). Further, specific miRNA genes have been documented as genetic risk factors in psychiatric disorders (Kocerha *et al.* 2015) and others will likely be implicated as genetic studies progress.

Despite these important findings, a relative paucity of detailed information on the non-coding transcriptome in the brain continues to hinder our ability to interpret GWAS signals in psychiatric disorders. In some instances, genetic association signals may be pointing to an unannotated non-coding transcript or to a regulatory region controlling the expression of a non-coding transcript locus as opposed to a coding gene (Farh *et al.* 2014). Recent eQTL resources such as the GTEx project (GTEx Consortium 2013, 2015; Mele *et al.* 2015), BrainSpan (<http://brain-map.org/>) and other databases (Jaffe *et al.* 2015) are beginning to make some headway in this area by using RNA-seq data. However, the RNA libraries for some of these were polyadenylation-selected and non-polyadenylated transcripts are not represented (Derrien *et al.* 2012; Kim *et al.* 2010). It is not clear how this omission will impact our understanding of gene regulation in brain, but non-polyadenylated transcripts are abundant in brain. One category of RNA, the long non-coding RNAs (lncRNAs), are enriched for non-polyadenylated transcripts compared with mRNA (Derrien *et al.* 2012), and the highest number of tissue-specific lncRNAs are found in brain (Derrien *et al.* 2012; Francescato *et al.* 2014). Further, there is a typically non-polyadenylated species of RNA transcribed at enhancers (eRNAs) (Kim *et al.* 2010; Koch *et al.* 2011) and these eRNAs themselves function as key transcriptional regulators (Leveille *et al.* 2015).

A further complication when it comes to the detection of non-coding RNA species in RNA-seq datasets is that many non-coding RNAs exhibit lower expression (~10-fold lower for lncRNAs) and more tissue specificity than the mRNAs of coding genes (Cabili *et al.* 2011; Djebali *et al.* 2012). Also, some are expressed only transiently (Cabili *et al.* 2011). Thus, annotating these transcripts will require combining deep sequencing with the extensive brain regional representation that was achieved previously using microarrays (Hawrylycz *et al.* 2012). Current RNA-seq methods require libraries to be constructed differently for long and short transcripts (e.g. microRNAs) and this need for two separate libraries increases the

sequencing expense. Because of costs, it has been argued that it currently may be impractical to obtain full annotation of rare transcripts; more-sensitive and/or less-expensive sequencing methods will be required.

Despite the advances afforded by SNP-expression (eQTL) datasets in narrowing the field of candidates in post-GWAS searches for functional risk alleles (Darabi *et al.* 2015; Ramasamy *et al.* 2014), these studies still cannot distinguish the functional alleles from those in strong LD with them. That is, the datasets identify correlations between particular genetic variants and the expression of particular transcripts, but do not identify the actual functional allele(s) among all the DNA variants in strong LD. Given that LD can span megabases from the associated allele(s) identified in a GWAS, many DNA variants may lie within the region of LD and correlate with expression. Functional studies will still be required to distinguish – among these correlated markers – the risk allele(s) that are actually influencing expression and that, in turn, are likely to be causal in the disease/disorder. The most illuminating studies will be those that not only identify eQTL located in regulatory elements but also shed light on the relevant target gene(s) – even when the target lies outside the LD region.

Direct determination of enhancer targets through looping analyses

Enhancers work by looping to physically contact the promoter(s) of their target gene(s) and stimulate transcription. Chromosome conformation capture (3C) methods (Miele & Dekker 2008; Ni *et al.* 2008; Sajan & Hawkins 2012) are looping assays that detect these physical interactions between enhancers and their target(s) and thus can be used to identify the disease-relevant gene(s) influenced by risk alleles within the enhancers.

In 3C studies, cells are first fixed to stabilize – by cross-linking – the physical contact points between regulatory elements within the nucleus. The cells are then lysed and the chromatin digested with a restriction enzyme, leaving the ends of what were once linearly distant but conformationally close DNA sequences available for ligation to one another; religation under dilute conditions promotes the desired ligation events, positioning the originally distant DNA sequences adjacent to each other in the ligation products (Lieberman-Aiden *et al.* 2009; Miele & Dekker 2008; Rao *et al.* 2014; Shen *et al.* 2012). The cross-links on the chromatin are then reversed, and the ligation products analyzed either by quantitative PCR or by DNA sequencing (depending on the application) (Dixon *et al.* 2012; Hughes *et al.* 2014; Rao *et al.* 2014).

In its most basic form, 3C is combined with quantitative PCR for targeted interrogation of individual prespecified candidate promoter–enhancer interactions, and is the least expensive method available (Abou El Hassan & Bremner 2009; Miele & Dekker 2008). However, this method requires *a priori* prediction as to which promoters and enhancers pair, and hence, which DNA sequence pairs are most likely to interact. Given that promoter targets may be hundreds of kilobases away from the enhancer, even outside the associated LD region found in a GWAS, either all known promoters within a large area must be tested or the promoter targets must be predicted. Thus, this method is limited. Nevertheless, the recently released FANTOM5 dataset provides some basis for predictions (Andersson *et al.* 2014). Leveraging the observation that active enhancers are, typically, themselves transcribed, an analysis of

the FANTOM5 dataset used pairwise expression correlations of the transcriptional activity of putative enhancers with the transcriptional activity of transcription start sites to predict the targets of enhancers across a diverse set of cell types.

Given the current numbers of GWAS signals for psychiatric disorders and the expected increase in these numbers as sample sizes increase, high-throughput genomic methods for the parallel interrogation of multiple potential promoter–enhancer interactions are required. The state of the art is Hi-C, a method that combines 3C with deep sequencing of all points of chromatin interaction, genome-wide (Dixon *et al.* 2012; Lieberman-Aiden *et al.* 2009). The resolution to identify interactions is dependent on sequencing depth. Unfortunately, because the costs of reaching the sequencing depths required for whole-genome coverage are currently high, the typical resolution achievable does not allow for the identification of specific promoter–enhancer interactions for most datasets (Zuin *et al.* 2014). However, the results can provide information on chromatin domains that can be used to estimate the extent of potential interactions (Dixon *et al.* 2012).

Bridging the gap between the limited ‘one-to-one’ scope of basic 3C and the limited resolution of the whole-genome (‘all-to-all’) Hi-C alternative, there are a number of other 3C-related methods that combine next-generation sequencing with targeted investigation of elements or regions of interest. These include methods such as 5C, a ‘many-to-many’ approach that identifies all interactions occurring within a particular region (as opposed to genome-wide as in Hi-C) (Dostie *et al.* 2006), ChIA-PET, a hybrid method that incorporates ChIP to interrogate the subset of all chromatin looping points that are bound by a protein of interest (Fullwood *et al.* 2009; Heidari *et al.* 2014), and 4C, a ‘one-to-all’ approach that investigates all (genome-wide) interactions with one particular element (e.g. an enhancer or a promoter) – the element of interest is designated the bait or viewpoint (Simonis *et al.* 2006; Splinter *et al.* 2012; van de Werken *et al.* 2012a,b; Zhao *et al.* 2006). Among these methods, the 4C strategy is particularly well-suited to the task of finding the target(s) of an enhancer that has been singled out for post-GWAS investigation. There are also some adaptations of 4C that can be further advantageous. For example, enhanced 4C (e4C) includes a bait/viewpoint-enrichment step (using biotin-coupled bait-specific primer extension) that results in at least 100-fold enrichment of bait-sequence representation in the library – improving the sensitivity to detect weaker (yet important) interactions that might otherwise go undetected (Sexton *et al.* 2012). Other adaptations, Capture-C and Capture Hi-C (CHi-C), use a hybridization capture approach to enrich for many different bait sequences (viewpoints) simultaneously. These multiplex enrichment strategies followed by high-throughput sequencing allow for interrogation of the interaction points of hundreds to many thousands of different bait elements (e.g. enhancers or promoters) simultaneously (Hughes *et al.* 2014; Mifsud *et al.* 2015).

Ultimately, one must show that the putative risk alleles identified in gene-regulatory elements alter such interactions – either destroy them or create new interactions (Visser *et al.* 2012). When appropriate cell types are available, specific nucleotides can be altered using genome editing (Charpentier & Doudna 2013; Jinek *et al.* 2012) and the impact on these interactions observed. This approach will not be possible, however, when the interactions are found only in human brain tissues. In that case, a large multi-individual panel of brain

tissues screened by genotype would be necessary to correlate the presence of the putative risk allele with altered interactions. Such studies would, however, require careful attention to tissue quality – particularly post-mortem (or post-surgical) intervals, as some interactions may be prone to degradation (Mitchell *et al.* 2014).

DNA variation and methylation

DNA methylation is known to play a critical role in brain function and behavior, and deregulated methylation results in disease (Davies *et al.* 2007; Labrie *et al.* 2012; Lyst & Bird 2015; Shin *et al.* 2014); however, the impact of specific DNA variants on methylation and the impact of the resultant methylation changes on gene function have been studied for relatively few loci (Dempster *et al.* 2013; Labrie *et al.* 2012). Genetic changes influencing methylation may result in the creation of a repressive or permissive chromatin environment and/or may impact the binding sites of specific TFs (Dantas Machado *et al.* 2015; Lyst & Bird 2015; Sun *et al.* 2014), thereby altering transcriptional regulation.

CG methylation

Until recently, the focus of DNA methylation studies was on methylated C (5mC) in CG dinucleotides (mCG). DNA variation may directly, through removal or addition of a cytosine nucleotide, or indirectly, through removal or addition of a guanine nucleotide, cause a loss or gain of methylation that could alter gene expression. Further, nucleotides surrounding the CG may also influence the probability that the cytosine is methylated. Allele-specific methylation analysis in mouse frontal cortex indicated that methylation of CG is biased by the sequence of adjacent nucleotides on either side of the CG, with some motifs preferentially hypermethylated and others hypomethylated (Xie *et al.* 2012). For example, the motif CTCGCG was found to be hypermethylated, whereas the motif CACGTG was found to be hypomethylated. Analysis of human methylomes also showed bias in the methylation of these motifs (Xie *et al.* 2012), indicating cross-species conservation of sequence preference of the methylation machinery; DNA variation of these motifs may affect methylation and impact gene expression.

Non-CG methylation

Non-CG methylation was first reported in mouse frontal cortex (Xie *et al.* 2012), challenging dogma that in differentiated mammalian tissues, C is only methylated (5mC) in the context of an adjacent G (CG methylation). Up until this discovery, non-CG methylation had been documented only in embryonic stem (ES) cells, oocytes and preimplantation embryos (He & Ecker 2015; Lister *et al.* 2009). Non-CG methylation disappeared upon differentiation of ES cells and reappeared in the creation of induced pluripotent stem (iPS) cells (Lister *et al.* 2009), and thus was thought to be a marker of pluripotent cell types. The finding of non-CG methylation in adult brain challenged this view and paved the way to better understanding the role of this epigenetic modification.

Additional studies confirmed the prevalence of non-CG methylation in mouse brain (Guo *et al.* 2014; Lister *et al.* 2013) and showed that it is also present in human brain (Guo *et al.* 2014; Lister *et al.* 2013). In adult brain tissue, 2.9% of cytosines genome-wide were found in

methylated CGs and 1.3–1.5% were found in methylated non-CGs (He & Ecker 2015; Lister *et al.* 2013). Differentiating between brain cell types, non-CG methylation appears to be prevalent in neurons (~6% of cytosines methylated in the non-CG context) but low in glial cells (<1%) (He & Ecker 2015; Lister *et al.* 2013; Schultz *et al.* 2015).

Non-CG and CG methylation differ in their developmental timing: CG methylation is established during fetal development, whereas non-CG methylation is low in fetal brain (Lister *et al.* 2013; Spiers *et al.* 2015) and accumulates in neurons in the course of postnatal development, gradually increasing during brain maturation to late adolescence before declining slightly during aging (Guo *et al.* 2014; Lister *et al.* 2013). The most rapid neuronal accumulation of non-CG methylation occurs from birth to 2 years of age, with slower accumulation continuing through adolescence till it becomes the dominant form of methylation in adult neurons (Lister *et al.* 2013). This trajectory may be related to the dramatic rates of gene expression changes seen in frontal cortex during development: the highest rates are in fetal brain, and then during infancy, the rates of changes in gene expression become much slower, slowing further still during childhood (1–10 years of age) to adolescence, after which low rates of change persist into mid-adulthood (Colantuoni *et al.* 2011).

The genomic distribution of non-CG methylation is non-random and differs from CG methylation, indicating unique functions; there are, for example, large (in the order of megabases) regions that are largely devoid of non-GC methylation but retain mCG (He & Ecker 2015). Elsewhere, non-CG methylation is enriched in repeat elements, inactive enhancers and gene bodies (He & Ecker 2015). Like CG methylation, non-CG methylation in brain and neurons is influenced by surrounding sequence: A preference for methylation of CA motifs (Lister *et al.* 2013; Mo *et al.* 2015; Xie *et al.* 2012) compared with other nucleotides is similar to that seen in oocytes (Tomizawa *et al.* 2011) and ES cells (Lister *et al.* 2009). But when motif analysis was expanded to include additional flanking nucleotides, differences in sequence preference between tissues were evident – the motif TACAC is preferentially methylated in neurons and glia, whereas TACAG is preferentially methylated in pluripotent cells (Schultz *et al.* 2015). This difference suggests that the function of non-CG methylation in neurons may be distinct from that in pluripotent cells, and analysis of gene expression supports this view (He & Ecker 2015). Methylation of CAC across the gene body is negatively correlated with expression in neurons, whereas methylation of CAG across the gene body is positively correlated with expression in pluripotent cells (He & Ecker 2015). Genetic variation in the surrounding motif could influence the probability of methylation, and hence gene expression, in a tissue-specific manner, and possibly the resultant phenotype.

Recently, non-CG methylation was reported in several other differentiated tissues, but at much lower levels than in brain – less than 0.4% of cytosines methylated in the non-CG context (Schultz *et al.* 2015) compared with 1–2% in frontal cortex (He & Ecker 2015). The mark seems to be in a subpopulation of cells within the tissues, as replicate samples excised from the same tissue differed in their percentage of non-CG methylation, indicating non-homogenous distribution (Schultz *et al.* 2015). A motif, TNCAC, similar to the TACAC motif methylated in neurons was observed, as well as a NNCAN motif (where N is any

nucleotide). The authors speculate that these non-CG methylation patterns may represent adult stem cells present in the tissues (Schultz *et al.* 2015). Another possibility to consider is that neural cells within these tissues could be contributing to the observed non-CG methylation.

The specific role(s) of non-CG methylation in the brain and whether this epigenetic mark functions in disease are currently unknown (He & Ecker 2015). One intriguing possibility is that non-CG methylation is directly related to Rett syndrome, a neurodevelopmental disorder resulting from deleterious mutations in the gene encoding methyl-CpG-binding protein 2 (MeCP2) (Lyst & Bird 2015). Non-CG methylation is correlated with repressed transcription, and MeCP2, in addition to binding mCG, binds methylated non-CG with both high affinity and a preference for methylated CA (mCA) (Chen *et al.* 2015; Gabel *et al.* 2015; Guo *et al.* 2014). It is speculated that the timing of onset of Rett syndrome symptoms – typically not until 6–18 months after birth – is directly related to the acquisition of increased levels of non-CG methylation in a key subset of neuronal genes that begins after birth in the maturing brain (Chen *et al.* 2015; Gabel *et al.* 2015). The genes implicated are typically long (>100 kb) and mCA-rich and appear to be subject to MeCP2-binding-mediated transcriptional repression. In the absence of functional MeCP2, however (i.e. in Rett syndrome), derepression ensues, resulting in the pathological overexpression of these heavily non-CG-methylated genes (Chen *et al.* 2015; Gabel *et al.* 2015). In cultured neurons lacking MeCP2, long-gene derepression could be reversed pharmacologically, using a topoisomerase inhibitor, and this rebalancing of gene expression was associated with an improvement in cellular health (Gabel *et al.* 2015). The authors speculate that the rebalancing of long-gene expression in neurons could attenuate neuronal dysfunction in Rett syndrome and may be a strategy for treatment (Gabel *et al.* 2015).

5-Hydroxymethylation

First identified in the early 1970s, 5-hydroxymethylation (5-hydroxymethylcytosine, 5hmC) was thought to be an artifact until its presence in brain and ES cells was confirmed in 2009 (Kriaucionis & Heintz 2009; Kriukiene *et al.* 2012; Tahiliani *et al.* 2009). It was then thought that 5hmC was a transient intermediary in the demethylation of 5mC to demethylated C (Kinde *et al.* 2015). However, 5hmC has since been shown to be a *bona fide* epigenetic mark, with some loci stable across the life span and others dynamic, changing during development (Kinde *et al.* 2015; Sun *et al.* 2014; Szulwach *et al.* 2011). In brain tissue, 5hmC is abundant, accounting for 25–40% of modified cytosines (Hahn *et al.* 2013; Lister *et al.* 2013; Munzel *et al.* 2010; Nestor *et al.* 2012; Wen & Tang 2014) and increasing during neuronal differentiation (Hahn *et al.* 2013). In mouse and human brain, 5hmC was found to markedly increase from the early postnatal stage to adulthood, suggesting a role in neurodevelopment (Lister *et al.* 2013; Song *et al.* 2011; Szulwach *et al.* 2011; Wang *et al.* 2012). The vast majority (97–99.98%) of 5hmC is found in the CG context (Lister *et al.* 2013; Wen & Tang 2014; Wen *et al.* 2014a; Yu *et al.* 2012); however, studies of 5hmC during fear extinction have also pointed to changes in non-CG methylation sites (Li *et al.* 2014). 5hmC is enriched in enhancers and gene bodies and is positively correlated with gene expression in brain (Wen & Tang 2014). Further, 5hmC is more enriched on the sense strand

than on the antisense strand, whereas the antisense strand is more enriched for 5mC (Wen & Tang 2014).

As with non-CG methylation, the role of 5hmC modification in brain is not entirely clear (Kinde *et al.* 2015; Wen & Tang 2014). This modification has been implicated in multiple functions, including exon splicing, learning and memory, hippocampal neurogenesis and DNA-methylation-related plasticity (Kinde *et al.* 2015; Li *et al.* 2014; Santiago *et al.* 2014; Sun *et al.* 2014; Wen & Tang 2014). Some of these functions have been indicated indirectly, through manipulation of the ten-eleven translocation (Tet) family of dioxygenases (Kinde *et al.* 2015; Santiago *et al.* 2014; Sun *et al.* 2014; Wen & Tang 2014) – the enzymes responsible for the conversion of 5mC to 5hmC, and also for the conversion of 5mC and/or 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (He *et al.* 2011; Ito *et al.* 2011). 5-Formylcytosine and 5caC are found at much lower levels in the brain than 5hmC (Liu *et al.* 2013a), thus the assumption that these functions are likely attributable to 5hmC. However, the potential role of either 5fC or 5caC vs. the specific function of 5hmC in these processes is unknown. 5-Formylcytosine shows a different distribution in brain compared with 5mC and 5hmC (Wagner *et al.* 2015). Further, 5fC levels are higher in non-neuronal populations than in neurons and 5fC shows a different pattern of age-related distribution, its levels decreasing with age starting at early developmental stages (Wagner *et al.* 2015). These findings suggest different functions, but further studies are required.

Detection of DNA variation effects on DNA methylation

A number of sequence-based methods can be used to interrogate DNA variation and methylation (Harris *et al.* 2010; Plongthongkum *et al.* 2014; Shull *et al.* 2015; Ulahannan & Greally 2015). The state of the art is whole-genome single-nucleotide-resolution methylation analysis (methylC-seq) (Lister *et al.* 2009, 2013; Xie *et al.* 2012), which relies on bisulfite conversion of the DNA. Currently, it is cost-prohibitive to perform methylC-seq for more than a few individuals because of the depth of sequencing required to accurately call methylation events in the bisulfite-treated DNA. Enrichment methods paired with sequencing (without bisulfite conversion), such as methylated DNA immunoprecipitation sequencing (MeDIP-seq), allow genome-wide coverage at relatively low cost, but have lower ability to identify regions with intermediate levels of methylation and cannot distinguish allele-specific methylation (Harris *et al.* 2010; Plongthongkum *et al.* 2014; Shull *et al.* 2015). It is possible, however, to achieve single-nucleotide resolution for detection of allele-specific methylation within the methylation-enriched DNA by subsequently interrogating the captured (methylation-enriched) DNA fragments with bisulfite sequencing (Choi *et al.* 2010; Shull *et al.* 2015). Another well-established genome-wide approach with single-nucleotide resolution for subgenome-scale methylation study is reduced representation bisulfite sequencing (RRBS) (Meissner *et al.* 2005); however, RRBS relies on restriction enzyme sites and is estimated to cover only 8–14% of CpGs (Harris *et al.* 2010; Shull *et al.* 2015). In contrast to these genome-wide approaches, targeted bisulfite-sequencing protocols that provide single-nucleotide resolution for methylation analysis of selected regions of the genome have also been developed. Such methods are dependent on targeted oligonucleotide capture either before or after bisulfite conversion of the DNA and can be used to investigate hundreds of individuals for thousands of selected 5mCs (Plongthongkum *et al.* 2014; Shull

et al. 2015). For example, a method that uses padlock probes (Deng *et al.* 2009; Diep *et al.* 2012) (annealed to the bisulfite-treated genomic DNA, circularized and amplified by PCR for subsequent high-throughput sequencing) has been optimized such that a genome-scale set of over 330 000 padlock probes can be used in a single-tube reaction to cover >500 000 selected CG sites (Diep *et al.* 2012) (http://genome-tech.ucsd.edu/public/Gen2_BSPP/); it also allows for the development of additional probes for more in-depth coverage of regions of particular interest. This method is advantageous because DNA variants are called from the sequence and the impact on methylation identified.

The analysis of 5hmC presents further challenges: 5hmC cannot be distinguished from 5mC by standard bisulfite sequencing but can be distinguished by other methods, such as Tet-assisted bisulfite sequencing (TAB-seq) (Yu *et al.* 2012). 5-Formylcytosine and 5caC also cannot be distinguished using bisulfite sequencing and are too rare to be identified using immunoprecipitation-enrichment techniques. A number of methods for the identification of 5fC have been published including 5fC-selective chemical labeling (fC-Seal) for genome-wide profiling of 5fC and a 5fC chemically assisted bisulfite sequencing (fCAB-seq) method for the base-resolution detection of 5fC (Song *et al.* 2013).

DNA methylome maps

The quantitation of DNA methylation across the genome, through the types of approaches described above, produces methylome maps that can be further investigated through multiple types of association analysis, including methylome-wide association study (MWAS) and methylation QTL (meQTL) and expression methylation-site (eMS) analyses. The MWAS approach analyzes DNA methylation sites in relation to a complex trait or disorder (Sun 2014), the meQTL approach analyzes the methylation sites in relation to genome-wide genetic markers (Luijk *et al.* 2015) and the eMS approach analyzes them in relation to genome-wide gene/RNA expression (Liu *et al.* 2013b). Post-GWAS studies of complex disorders are increasingly integrating these methylomics analyses, not only with the GWAS signals themselves but also with eQTL and regulatory-region (e.g. enhancer) maps, in order to uncover functional risk variants – including ones that may be involved in brain function and/or psychiatric disorders (Clark *et al.* 2015; Gamazon *et al.* 2013; Gibbs *et al.* 2010; Kumar *et al.* 2015; van Eijk *et al.* 2015; Wockner *et al.* 2014; Zhao *et al.* 2014).

Obstacles to the study of gene regulation in psychiatric disorders

The study of gene regulation involved in psychiatric disorders has unique limitations not encountered in the study of many other genetic disorders. Specifically, the field is limited by the availability of neural cell types that can be cultured as well as by the availability of high-quality post-mortem brain tissues from both patients and controls. These limitations are particularly problematic for the study of sensitive developmental time periods (prenatal, early childhood and adolescence), sex differences and/or for the study of specific genotypes. We review these obstacles below.

Representative neural cell types for molecular studies

A major obstacle in moving genetic association findings for brain disorders forward to understanding altered gene function has been the limited ability to culture neurons and glial cells for molecular studies. Human iPS cells (Ho *et al.* 2015; Maury *et al.* 2015; Wright *et al.* 2014), differentiated human ES cells (Espuny-Camacho *et al.* 2013; Maury *et al.* 2015; Nicholas *et al.* 2013) and olfactory neurosphere-derived cells (Matigian *et al.* 2010) can be used as a source of neural cell types. Among these, iPS cells from patients are particularly valuable in the examination of cellular phenotypes (Brennand *et al.* 2011; Chen *et al.* 2014; Madison *et al.* 2015; Robicsek *et al.* 2013; Wang *et al.* 2014), especially when derived from patients with known risk genotypes (Pedrosa *et al.* 2011; Wen *et al.* 2014b; Yoon *et al.* 2014). Currently, these cellular models can capture a number of neural cell types, but as yet cannot fully capture the diversity of cells found in brain (Ho *et al.* 2015; Maury *et al.* 2015; Merkle *et al.* 2015; Nicholas *et al.* 2013). There are also multiple sources of variability, such as cell origin, cell-derivation methods, cell-type heterogeneity within the culture and cell-culture conditions, that limit the interpretation of cellular characteristics and outcomes in cell cultures derived from patient samples (Ho *et al.* 2015; Nityanandam & Baldwin 2015; Sandoe & Eggan 2013). Appropriate representation in terms of neuronal maturity is also an issue (Ho *et al.* 2015; Nicholas *et al.* 2013; Sandoe & Eggan 2013). For example, human iPS neural precursor cells and 6-week-old neurons exhibit a gene expression pattern similar to that of first-trimester fetal forebrain (Brennand *et al.* 2015). These cells could serve as a model for neurodevelopmental disorders that are primarily prenatally influenced but may not model disorders where the critical period is later in life. Because of these limitations, there is a crucial need for the development of methods to create mature neurons (Nityanandam & Baldwin 2015).

Representation of key developmental risk periods

Currently, key developmental risk periods for psychiatric disorders are not well represented in transcriptome and epigenome datasets with the majority of data obtained from adult tissues. This represents a major limitation because a number of the psychiatric disorders are considered neurodevelopmental (Estes *et al.* 2015; Owen *et al.* 2011; Weinberger 1987; Zwaigenbaum *et al.* 2005), with symptoms seen very early in development. A critical time point for some of these may be prenatal. Fetal brain is represented in the Epigenomic Roadmap, but for only a very limited number of gestational ages (<http://www.ncbi.nlm.nih.gov/geo/roadmap/epigenomics/?search=fetal+brain&display=50>). Gene expression dramatically and continually changes over the course of fetal development (Jaffe *et al.* 2015; Kang *et al.* 2011) and such a limited sample set may not capture the totality of the changes (Kang *et al.* 2011).

The most important developmental time period for the emergence of many psychiatric disorders is late childhood through adolescence (Kessler *et al.* 2005). Adolescence is a time of dramatic physiological, cognitive and emotional changes and is associated with an increased risk of psychiatric disorders, suicide and aggression among adolescent youth. This increased risk of disturbance is most likely attributable to both psychosocial and hormonal factors – with learning and experience interacting with biologically programmed developmental influences to modify structure and function in the brain (Sisk & Zehr 2005).

First documented in post-mortem studies, recent neuroimaging studies confirm major structural changes occurring in the brain (e.g. decreases in gray matter and increases in white matter, myelination, synaptogenesis and synaptic pruning) from late childhood to early adolescence, and functional imaging studies have also revealed differences in regional activation in the brain during this period (Blakemore 2012). Analysis of brain transcriptome data reflects these changes, with adolescent brain showing the second highest number of differentially expressed transcripts after fetal brain (Jaffe *et al.* 2015).

Maps of regulatory regions functioning in brain during this critical period are not currently represented in the Epigenomic Roadmap (Roadmap Epigenomics Consortium *et al.* 2015) and are underrepresented in current transcriptome databases (http://help.brain-map.org/download/attachments/3506181/Human_Brain_Seq_Stages.pdf?version=1&modificationDate=1433436980032) (Colantuoni *et al.* 2011; Jaffe *et al.* 2015; Kang *et al.* 2011). This is because of the paucity of brain tissues from these age groups available in brain banks.

Representation of male and female brain

There is a sex bias in the prevalence of most psychiatric disorders (Baron-Cohen *et al.* 2011; Kigar & Auger 2013; Ngun *et al.* 2011). Many of the early childhood-onset disorders have a male predominance (e.g. ADHD, Tourette syndrome and autism) (Zahn-Waxler *et al.* 2008). Depression in children with onset at less than 14 years of age is found in almost equal proportions in boys and girls, with a slight predominance of male cases (Hankin *et al.* 1998; Kovacs *et al.* 2003). During midpuberty, at around Tanner Stage III (~age 13–14), this ratio changes, with a dramatic increase in incidence among girls and a shift to female predominance (Angold *et al.* 1998).

In brain, sex differences are evident by structural neuroimaging as well as by neural activation patterns in functional imaging paradigms (Blakemore 2012; Ingalhalikar *et al.* 2014). Estrogen is present in the adult, postnatal and, possibly, prenatal brain, influencing brain organization and assisting in the feminization of both the brain and behavior (see Collaer & Hines 1995). Testosterone is present in the fetus as early as 8 weeks of gestation, with a surge reaching almost pubertal levels (Hines 2008). There is another surge after birth, before dropping to low levels at 4–6 months (Forest *et al.* 1974). Gonadal hormones influence histone modifications as well as DNA methylation, and levels of mRNA for methyltransferases differ in specific regions of the male and female brain (Nugent *et al.* 2012), likely contributing to differential gene expression. Sex-chromosome-linked genes also display sex-biased gene expression and these may contribute to sex differences in brain (Deng *et al.* 2014; Loke *et al.* 2015; Mele *et al.* 2015; Ngun *et al.* 2011; Trabzuni *et al.* 2013; Vawter *et al.* 2004; Weickert *et al.* 2009). Notable is the Y chromosome gene SRY (Sex-determining Region on the Y chromosome), the male sex determination and differentiation gene that plays a key role in regulating the catecholaminergic system (Loke *et al.* 2015; Ngun *et al.* 2011). A number of autosomal genes exhibit sex-specific or sex-biased gene expression and/or exon usage as borne out in comparative analyses of exon-level transcriptome data from human male and female brain samples taken at various stages – from fetal to adult (Kang *et al.* 2011; Trabzuni *et al.* 2013). Thus, epigenetic changes in the

male and female developing and adult brain are of extreme interest for the understanding of psychiatric disorders and resources should be increased for more comprehensive study in this area.

Outlook

The GWAS findings clearly point to variants in the non-coding genome as being significant contributors to genetic risk. While large sequencing studies are underway for several of the psychiatric disorders, the interpretation of the findings for variants identified in the non-coding genome will be hampered by our ignorance of the function of these regions and the impact of genetic variants within them. Although resources providing transcriptome and epigenomic maps for brain tissue are increasing, there are still gaps in the developmental stages covered by these resources as well as gaps in the available epigenetic marks and transcriptome coverage. As these resources continue to expand, integrating the multiple layers of ‘omics’ data produced will yield a more complete understanding of the complex relationships that exist among DNA variants and the diversity of epigenetic features involved in transcriptional regulation (Banovich *et al.* 2014; Baubec & Schubeler 2014; Zhi *et al.* 2013) – this will require the ongoing development of analytic methods that can handle the increasingly multilayered datasets (Huang 2014; Pineda *et al.* 2015). Even with this advanced profiling of the epigenetic landscape, however, the links between specific variants and function will require experimental validation. The development of neural cell types from pluripotent cells to closely approximate cellular functions in brain will be critical for these studies.

Acknowledgments

This study was supported by Krembil Foundation, Ontario Mental Health Foundation and the Canadian Institutes of Health Research.

References

- Abou El Hassan M, Bremner R. A rapid simple approach to quantify chromosome conformation capture. *Nucleic Acids Res.* 2009; 37:e35. [PubMed: 19181703]
- Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature.* 2014; 507:455–461. [PubMed: 24670763]
- Angold A, Costello EJ, Worthman CM. Puberty and depression: the roles of age, pubertal status and pubertal timing. *Psychol Med.* 1998; 28:51–61. [PubMed: 9483683]
- Anney R, Klei L, Pinto D, et al. Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet.* 2012; 21:4781–4792. [PubMed: 22843504]
- Arnold CD, Gerlach D, Stelzer C, Boryn LM, Rath M, Stark A. Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science.* 2013; 339:1074–1077. [PubMed: 23328393]
- Banovich NE, Lan X, McVicker G, van de Geijn B, Degner JF, Blischak JD, Roux J, Pritchard JK, Gilad Y. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. *PLoS Genet.* 2014; 10:e1004663. [PubMed: 25233095]
- Baron-Cohen S, Lombardo MV, Auyeung B, Ashwin E, Chakrabarti B, Knickmeyer R. Why are autism spectrum conditions more prevalent in males? *PLoS Biol.* 2011; 9:e1001081. [PubMed: 21695109]
- Barry G. Integrating the roles of long and small non-coding RNA in brain function and disease. *Mol Psychiatry.* 2014; 19:410–416. [PubMed: 24468823]

- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007; 129:823–837. [PubMed: 17512414]
- Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell*. 2013; 152:1298–1307. [PubMed: 23498938]
- Baubec T, Schubeler D. Genomic patterns and context specific interpretation of DNA methylation. *Curr Opin Genet Dev*. 2014; 25:85–92. [PubMed: 24614011]
- Berger MF, Bulyk ML. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat Protoc*. 2009; 4:393–411. [PubMed: 19265799]
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ III, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005; 120:169–181. [PubMed: 15680324]
- Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol*. 2010; 28:1045–1048. [PubMed: 20944595]
- Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]
- Bhattacharya N, Sarno A, Idler IS, Fuhrer M, Zenz T, Dohner H, Stilgenbauer S, Mertens D. High-throughput detection of nuclear factor-kappaB activity using a sensitive oligo-based chemiluminescent enzyme-linked immunosorbent assay. *Int J Cancer*. 2010; 127:404–411. [PubMed: 19924814]
- Blakemore SJ. Imaging brain development: the adolescent brain. *Neuroimage*. 2012; 61:397–406. [PubMed: 22178817]
- Boffelli D, Nobrega MA, Rubin EM. Comparative genomics at the vertebrate extremes. *Nat Rev Genet*. 2004; 5:456–465. [PubMed: 15153998]
- Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, Li Y, Mu Y, Chen G, Yu D, McCarthy S, Sebat J, Gage FH. Modelling schizophrenia using human induced pluripotent stem cells. *Nature*. 2011; 473:221–225. [PubMed: 21490598]
- Brennand K, Savas JN, Kim Y, Tran N, Simone A, Hashimoto-Torii K, Beaumont KG, Kim HJ, Topol A, Ladrán I, Abdelrahim M, Matikainen-Ankney B, Chao SH, Mrksich M, Rakic P, Fang G, Zhang B, Yates JR III, Gage FH. Phenotypic differences in hiPSC NPCs derived from patients with schizophrenia. *Mol Psychiatry*. 2015; 20:361–368. [PubMed: 24686136]
- Butler JE, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*. 2002; 16:2583–2592. [PubMed: 12381658]
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011; 25:1915–1927. [PubMed: 21890647]
- Casey JP, Magalhaes T, Conroy JM, et al. A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Hum Genet*. 2012; 131:565–579. [PubMed: 21996756]
- Charpentier E, Doudna JA. Biotechnology: rewriting a genome. *Nature*. 2013; 495:50–51. [PubMed: 23467164]
- Chen HM, DeLong CJ, Bame M, Rajapakse I, Herron TJ, McInnis MG, O’Shea KS. Transcripts involved in calcium signaling and telencephalic neuronal fate are altered in induced pluripotent stem cells from bipolar disorder patients. *Transl Psychiatry*. 2014; 4:e375. [PubMed: 25116795]
- Chen L, Chen K, Lavery LA, Baker SA, Shaw CA, Li W, Zoghbi HY. MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proc Natl Acad Sci USA*. 2015; 112:5509–5514. [PubMed: 25870282]
- Choi JH, Li Y, Guo J, Pei L, Rauch TA, Kramer RS, Macmill SL, Wiley GB, Bennett LB, Schnabel JL, Taylor KH, Kim S, Xu D, Sreekumar A, Pfeifer GP, Roe BA, Caldwell CW, Bhalla KN, Shi H.

- Genome-wide DNA methylation maps in follicular lymphoma cells determined by methylation-enriched bisulfite sequencing. *PLoS One*. 2010; 5doi: 10.1371/journal.pone.0013020
- Chorley BN, Wang X, Campbell MR, Pittman GS, Noureddine MA, Bell DA. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res*. 2008; 659:147–157. [PubMed: 18565787]
- Clark SL, Aberg KA, Nerella S, Kumar G, McClay JL, Chen W, Xie LY, Harada A, Shabalin AA, Gao G, Bergen SE, Hultman CM, Magnusson PK, Sullivan PF, van den Oord EJ. Combined whole methylome and genomewide association study implicates CNTN4 in alcohol use. *Alcohol Clin Exp Res*. 2015; 39:1396–1405. [PubMed: 26146898]
- Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, Colantuoni EA, Elkahlon AG, Herman MM, Weinberger DR, Kleinman JE. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*. 2011; 478:519–523. [PubMed: 22031444]
- Collaer ML, Hines M. Human behavioral sex differences: a role for gonadal hormones during early development? *Psychol Bull*. 1995; 118:55–107. [PubMed: 7644606]
- Couto JM, Livne-Bar I, Huang K, Xu Z, Cate-Carter T, Feng Y, Wigg K, Humphries T, Tannock R, Kerr EN, Lovett MW, Bremner R, Barr CL. Association of reading disabilities with regions marked by acetylated H3 histones in KIAA0319. *Am J Med Genet*. 2010; 153B:447–462. [PubMed: 19588467]
- Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI, Ripke S, Santangelo S, Sullivan PF. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013; 381:1371–1379. [PubMed: 23453885]
- Dantas Machado AC, Zhou T, Rao S, Goel P, Rastogi C, Lazarovici A, Bussemaker HJ, Rohs R. Evolving insights on how cytosine methylation affects protein-DNA binding. *Brief Funct Genomics*. 2015; 14:61–73. [PubMed: 25319759]
- Darabi H, McCue K, Beesley J, et al. Polymorphisms in a putative enhancer at the 10q21.2 breast cancer risk locus regulate NRBF2 expression. *Am J Hum Genet*. 2015; 97:22–34. [PubMed: 26073781]
- Dathe K, Kjaer KW, Brehm A, Meinecke P, Nurnberg P, Neto JC, Brunoni D, Tommerup N, Ott CE, Klopocki E, Seemann P, Mundlos S. Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *Am J Hum Genet*. 2009; 84:483–492. [PubMed: 19327734]
- Datson NA, van der Perk J, de Kloet ER, Vreugdenhil E. Expression profile of 30,000 genes in rat hippocampus using SAGE. *Hippocampus*. 2001; 11:430–444. [PubMed: 11530848]
- Davies W, Isles AR, Humby T, Wilkinson LS. What are imprinted genes doing in the brain? *Epigenetics*. 2007; 2:201–206. [PubMed: 18094620]
- De Gobbi M, Viprasak V, Hughes JR, Fisher C, Buckle VJ, Ayyub H, Gibbons RJ, Vernimmen D, Yoshinaga Y, de Jong P, Cheng JF, Rubin EM, Wood WG, Bowden D, Higgs DR. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science*. 2006; 312:1215–1217. [PubMed: 16728641]
- De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol*. 2010; 8:e1000384. [PubMed: 20485488]
- Dean A. On a chromosome far, far away: LCRs and gene expression. *Trends Genet*. 2006; 22:38–45. [PubMed: 16309780]
- Dempster E, Viana J, Pidsley R, Mill J. Epigenetic studies of schizophrenia: progress, predicaments, and promises for the future. *Schizophr Bull*. 2013; 39:11–16. [PubMed: 23211398]
- Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, Daley GQ, Eggan K, Hochedlinger K, Thomson J, Wang W, Gao Y, Zhang K. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat Biotechnol*. 2009; 27:353–360. [PubMed: 19330000]
- Deng X, Berletch JB, Nguyen DK, Distèche CM. X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet*. 2014; 15:367–378. [PubMed: 24733023]

- Dermitzakis ET, Clark AG. Evolution of transcription factor binding sites in mammalian gene regulatory regions: conservation and turnover. *Mol Biol Evol.* 2002; 19:1114–1121. [PubMed: 12082130]
- Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012; 22:1775–1789. [PubMed: 22955988]
- Diep D, Plongthongkum N, Gore A, Fung HL, Shoemaker R, Zhang K. Library-free methylation sequencing with bisulfite padlock probes. *Nat Methods.* 2012; 9:270–272. [PubMed: 22306810]
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012; 485:376–380. [PubMed: 22495300]
- Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature.* 2012; 489:101–108. [PubMed: 22955620]
- Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 2006; 16:1299–1309. [PubMed: 16954542]
- Driscoll MC, Dobkin CS, Alter BP. Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. *Proc Natl Acad Sci USA.* 1989; 86:7470–7474. [PubMed: 2798417]
- Dwyer S, Williams H, Holmans P, Moskvina V, Craddock N, Owen MJ, O'Donovan MC. No evidence that rare coding variants in ZNF804A confer risk of schizophrenia. *Am J Med Genet B Neuropsychiatr Genet.* 2010; 153B:1411–1416. [PubMed: 20862696]
- Eberharder A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep.* 2002; 3:224–229. [PubMed: 11882541]
- van Eijk KR, de Jong S, Strengman E, Buizer-Voskamp JE, Kahn RS, Boks MP, Horvath S, Ophoff RA. Identification of schizophrenia-associated loci by combining DNA methylation and gene expression data from whole blood. *Eur J Hum Genet.* 2015; 23:1106–1110. [PubMed: 25424713]
- Elbert A, Lovett MW, Cate-Carter T, Pitch A, Kerr EN, Barr CL. Genetic variation in the KIAA0319 5' region as a possible contributor to dyslexia. *Behav Genet.* 2011; 41:77–89. [PubMed: 21207242]
- Elia J, Glessner JT, Wang K, et al. Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder. *Nat Genet.* 2011; 44:78–84. [PubMed: 22138692]
- Ellinghaus D, Ellinghaus E, Nair RP, et al. Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Hum Genet.* 2012; 90:636–647. [PubMed: 22482804]
- Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature.* 2011; 473:43–49. [PubMed: 21441907]
- Espuny-Camacho I, Michelsen KA, Gall D, Linaro D, Hasche A, Bonnefont J, Bali C, Orduz D, Bilheu A, Herpoel A, Lambert N, Gaspard N, Peron S, Schiffmann SN, Giugliano M, Gaillard A, Vanderhaeghen P. Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron.* 2013; 77:440–456. [PubMed: 23395372]
- Estes A, Zwaigenbaum L, Gu H, St John T, Paterson S, Elison JT, Hazlett H, Botteron K, Dager SR, Schultz RT, Kostopoulos P, Evans A, Dawson G, Eliason J, Alvarez S, Piven J. IBIS Network. Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *J Neurodev Disord.* 2015; 7:24. [PubMed: 26203305]
- Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature.* 2014; 518:337–343. [PubMed: 25363779]
- Ferreira MA, O'Donovan MC, Meng YA, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet.* 2008; 40:1056–1058. [PubMed: 18711365]

- Forest MG, Sizonenko PC, Cathiard AM, Bertrand J. Hypophyso-gonadal function in humans during the first year of life. 1. Evidence for testicular activity in early infancy. *J Clin Invest.* 1974; 53:819–828. [PubMed: 4812441]
- Francescato M, Vitezic M, Heutink P, Saxena A. Brain-specific noncoding RNAs are likely to originate in repeats and may play a role in up-regulating genes in cis. *Int J Biochem Cell Biol.* 2014; 54:331–337. [PubMed: 24993078]
- Franke B, Neale BM, Faraone SV. Genome-wide association studies in ADHD. *Hum Genet.* 2009; 126:13–50. [PubMed: 19384554]
- Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet.* 2010; 42:1118–1125. [PubMed: 21102463]
- Fullwood MJ, Liu MH, Pan YF, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature.* 2009; 462:58–64. [PubMed: 19890323]
- Furniss D, Lettice LA, Taylor IB, Critchley PS, Giele H, Hill RE, Wilkie AO. A variant in the sonic hedgehog regulatory sequence (ZRS) is associated with triphalangeal thumb and deregulates expression in the developing limb. *Hum Mol Genet.* 2008; 17:2417–2423. [PubMed: 18463159]
- Gabel HW, Kinde B, Stroud H, Gilbert CS, Harmin DA, Kastan NR, Hemberg M, Ebert DH, Greenberg ME. Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature.* 2015; 522:89–93. [PubMed: 25762136]
- Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell.* 2002; 110:339–348. [PubMed: 12176321]
- Gamazon ER, Badner JA, Cheng L, et al. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry.* 2013; 18:340–346. [PubMed: 22212596]
- Gao Z, Ruan J. A structure-based multiple-instance learning approach to predicting in vitro transcription factor-DNA interaction. *BMC Genomics.* 2015; 16(Suppl 4):S3.
- Garber K. Psoriasis: from bed to bench and back. *Nat Biotechnol.* 2011; 29:563–566. [PubMed: 21747375]
- Gerasimova A, Chavez L, Li B, Seumois G, Greenbaum J, Rao A, Vijayanand P, Peters B. Predicting cell types and genetic variations contributing to disease by combining GWAS and epigenetic data. *PLoS One.* 2013; 8:e54359. [PubMed: 23382893]
- Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, Arepalli S, Dillman A, Rafferty IP, Troncoso J, Johnson R, Zielke HR, Ferrucci L, Longo DL, Cookson MR, Singleton AB. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 2010; 6:e1000952. [PubMed: 20485568]
- Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E, Grozeva D, Kirov G, Holmans P, Moran JL, Purcell S, Sklar P, Owen MJ, O'Donovan MC, Jones L, Jones IR, Craddock N. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol Psychiatry.* 2013; 18:1302–1307. [PubMed: 23070075]
- GTEX Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet.* 2013; 45:580–585. [PubMed: 23715323]
- GTEX Consortium. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015; 348:648–660. [PubMed: 25954001]
- Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, Zhu H, Chang Q, Gao Y, Ming GL, Song H. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci.* 2014; 17:215–222. [PubMed: 24362762]
- Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J, Jui J, Jin SG, Jiang Y, Pfeifer GP, Lu Q. Dynamics of 5-hydroxymethylcytosine and chromatin marks in mammalian neurogenesis. *Cell Rep.* 2013; 3:291–300. [PubMed: 23403289]
- Hamshere ML, Walters JT, Smith R, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the schizophrenia PGC. *Mol Psychiatry.* 2013; 18:708–712. [PubMed: 22614287]

- Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *J Abnorm Psychol.* 1998; 107:128–140. [PubMed: 9505045]
- Harris RA, Wang T, Coarfa C, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol.* 2010; 28:1097–1105. [PubMed: 20852635]
- Hatton CS, Wilkie AO, Drysdale HC, Wood WG, Vickers MA, Sharpe J, Ayyub H, Pretorius IM, Buckle VJ, Higgs DR. Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human alpha globin gene cluster. *Blood.* 1990; 76:221–227. [PubMed: 2364173]
- Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature.* 2012; 489:391–399. [PubMed: 22996553]
- He Y, Ecker JR. Non-CG methylation in the human genome. *Annu Rev Genomics Hum Genet.* 2015; 16:55–77. [PubMed: 26077819]
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011; 333:1303–1307. [PubMed: 21817016]
- Heidari N, Phanstiel DH, He C, Grubert F, Jahanbani F, Kasowski M, Zhang MQ, Snyder MP. Genome-wide map of regulatory interactions in the human genome. *Genome Res.* 2014; 24:1905–1917. [PubMed: 25228660]
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* 2007; 39:311–318. [PubMed: 17277777]
- Heinz S, Romanoski CE, Benner C, Allison KA, Kaikkonen MU, Orozco LD, Glass CK. Effect of natural genetic variation on enhancer selection and function. *Nature.* 2013; 503:487–492. [PubMed: 24121437]
- Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol.* 2015; 16:144–154. [PubMed: 25650801]
- Higgs DR, Vernimmen D, Hughes J, Gibbons R. Using genomics to study how chromatin influences gene expression. *Annu Rev Genomics Hum Genet.* 2006; 14:14.
- Hines M. Early androgen influences on human neural and behavioural development. *Early Hum Dev.* 2008; 84:805–807. [PubMed: 18938049]
- Hinney A, Scherag A, Jarick I, et al. Genome-wide association study in German patients with attention deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet.* 2011; 156B:888–897. [PubMed: 22012869]
- Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. *Cell.* 2013; 155:934–947. [PubMed: 24119843]
- Ho SM, Topol A, Brennand KJ. From “directed differentiation” to “neuronal induction”: modeling neuropsychiatric disease. *Biomark Insights.* 2015; 10:31–41. [PubMed: 26045654]
- Huang YT. Integrative modeling of multiple genomic data from different types of genetic association studies. *Biostatistics.* 2014; 15:587–602. [PubMed: 24705142]
- Hughes JR, Roberts N, McGowan S, Hay D, Giannoulatou E, Lynch M, De Gobbi M, Taylor S, Gibbons R, Higgs DR. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet.* 2014; 46:205–212. [PubMed: 24413732]
- Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R. Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci USA.* 2014; 111:823–828. [PubMed: 24297904]
- Cortes A, Hadler J, et al. International Genetics of Ankylosing Spondylitis Consortium. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet.* 2013; 45:730–738. [PubMed: 23749187]
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* 2011; 333:1300–1303. [PubMed: 21778364]

- Jaffe AE, Shin J, Collado-Torres L, Leek JT, Tao R, Li C, Gao Y, Jia Y, Maher BJ, Hyde TM, Kleinman JE, Weinberger DR. Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat Neurosci.* 2015; 18:154–161. [PubMed: 25501035]
- Jin F, Li Y, Ren B, Natarajan R. Enhancers: multi-dimensional signal integrators. *Transcription.* 2011; 2:226–230. [PubMed: 22231119]
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012; 337:816–821. [PubMed: 22745249]
- Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. *Science.* 2007; 316:1497–1502. [PubMed: 17540862]
- Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature.* 2011; 478:483–489. [PubMed: 22031440]
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry.* 2005; 62:593–602. [PubMed: 15939837]
- Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature.* 2011; 474:307–317. [PubMed: 21677747]
- Kigar SL, Auger AP. Epigenetic mechanisms may underlie the aetiology of sex differences in mental health risk and resilience. *J Neuroendocrinol.* 2013; 25:1141–1150. [PubMed: 23841484]
- Kilpinen H, Waszak SM, Gschwind AR, et al. Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science.* 2013; 342:744–747. [PubMed: 24136355]
- Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell.* 2007; 128:1231–1245. [PubMed: 17382889]
- Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. *Nature.* 2010; 465:182–187. [PubMed: 20393465]
- Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME. Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2. *Proc Natl Acad Sci USA.* 2015; 112:6800–6806. [PubMed: 25739960]
- King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science.* 1975; 188:107–116. [PubMed: 1090005]
- Kocerha J, Dwivedi Y, Brennand KJ. Noncoding RNAs and neurobehavioral mechanisms in psychiatric disease. *Mol Psychiatry.* 2015; 20:677–684. [PubMed: 25824307]
- Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C, Eick D, Gut I, Ferrier P, Andrau JC. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol.* 2011; 18:956–963. [PubMed: 21765417]
- Kovacs M, Obrosky DS, Sherrill J. Developmental changes in the phenomenology of depression in girls compared to boys from childhood onward. *J Affect Disord.* 2003; 74:33–48. [PubMed: 12646297]
- Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science.* 2009; 324:929–930. [PubMed: 19372393]
- Kriukiene E, Liutkeviciute Z, Klimasauskas S. 5-Hydroxymethylcytosine – the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev.* 2012; 41:6916–6930. [PubMed: 22842880]
- Kumar G, Clark SL, McClay JL, Shabalin AA, Adkins DE, Xie L, Chan R, Nerella S, Kim Y, Sullivan PF, Hultman CM, Magnusson PK, Aberg KA, van den Oord EJ. Refinement of schizophrenia GWAS loci using methylome-wide association data. *Hum Genet.* 2015; 134:77–87. [PubMed: 25284466]
- Labrie V, Pai S, Petronis A. Epigenetics of major psychosis: progress, problems and perspectives. *Trends Genet.* 2012; 28:427–435. [PubMed: 22622229]

- Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhatter R. Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature*. 2013; 494:497–501. [PubMed: 23417068]
- Lasky-Su J, Neale BM, Franke B, et al. Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *Am J Med Genet B Neuropsychiatr Genet*. 2008; 147B:1345–1354. [PubMed: 18821565]
- Lasky-Su J, Won S, Mick E, Anney RJ, Franke B, Neale B, Biederman J, Smalley SL, Loo SK, Todorov A, Faraone SV, Weiss ST, Lange C. On genome-wide association studies for family-based designs: an integrative analysis approach combining ascertained family samples with unselected controls. *Am J Hum Genet*. 2010; 86:573–580. [PubMed: 20346434]
- van de Leemput J, Boles NC, Kiehl TR, Corneo B, Lederman P, Menon V, Lee C, Martinez RA, Levi BP, Thompson CL, Yao S, Kaykas A, Temple S, Fasano CA. CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. *Neuron*. 2014; 83:51–68. [PubMed: 24991954]
- Lesch KP, Timmesfeld N, Renner TJ, Halperin R, Roser C, Nguyen TT, Craig DW, Romanos J, Heine M, Meyer J, Freitag C, Warnke A, Romanos M, Schafer H, Walitza S, Reif A, Stephan DA, Jacob C. Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J Neural Transm*. 2008; 115:1573–1585. [PubMed: 18839057]
- Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E. A long-range *Shh* enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet*. 2003; 12:1725–1735. [PubMed: 12837695]
- Leveille N, Melo CA, Agami R. Enhancer-associated RNAs as therapeutic targets. *Expert Opin Biol Ther*. 2015; 15:723–734. [PubMed: 25819025]
- Li X, Wei W, Zhao QY, Widagdo J, Baker-Andresen D, Flavell CR, D'Alessio A, Zhang Y, Bredy TW. Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation. *Proc Natl Acad Sci USA*. 2014; 111:7120–7125. [PubMed: 24757058]
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozcy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009; 326:289–293. [PubMed: 19815776]
- Lipovich L, Dacht F, Cai J, Bagla S, Balan K, Jia H, Loeb JA. Activity-dependent human brain coding/noncoding gene regulatory networks. *Genetics*. 2012; 192:1133–1148. [PubMed: 22960213]
- Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462:315–322. [PubMed: 19829295]
- Lister R, Mukamel EA, Nery JR, et al. Global epigenomic reconfiguration during mammalian brain development. *Science*. 2013; 341:1237905. [PubMed: 23828890]
- Liu S, Wang J, Su Y, Guerrero C, Zeng Y, Mitra D, Brooks PJ, Fisher DE, Song H, Wang Y. Quantitative assessment of Tet-induced oxidation products of 5-methylcytosine in cellular and tissue DNA. *Nucleic Acids Res*. 2013a; 41:6421–6429. [PubMed: 23658232]
- Liu Y, Ding J, Reynolds LM, et al. Methylomics of gene expression in human monocytes. *Hum Mol Genet*. 2013b; 22:5065–5074. [PubMed: 23900078]
- Loke H, Harley V, Lee J. Biological factors underlying sex differences in neurological disorders. *Int J Biochem Cell Biol*. 2015; 65:139–150. [PubMed: 26028290]
- Loots GG, Kneissel M, Keller H, Baptist M, Chang J, Collette NM, Ovcharenko D, Plajzer-Frick I, Rubin EM. Genomic deletion of a long-range bone enhancer misregulates sclerostin in Van Buchem disease. *Genome Res*. 2005; 15:928–935. [PubMed: 15965026]
- Ludwig MZ, Bergman C, Patel NH, Kreitman M. Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature*. 2000; 403:564–567. [PubMed: 10676967]

- Luijk R, Goeman JJ, Slagboom EP, Heijmans BT, van Zwet EW. An alternative approach to multiple testing for methylation QTL mapping reduces the proportion of falsely identified CpGs. *Bioinformatics*. 2015; 31:340–345. [PubMed: 25282645]
- Lyst MJ, Bird A. Rett syndrome: a complex disorder with simple roots. *Nat Rev Genet*. 2015; 16:261–275. [PubMed: 25732612]
- Madison JM, Zhou F, Nigam A, et al. Characterization of bipolar disorder patient-specific induced pluripotent stem cells from a family reveals neurodevelopmental and mRNA expression abnormalities. *Mol Psychiatry*. 2015; 20:703–717. [PubMed: 25733313]
- Maienschein-Cline M, Dinner AR, Hlavacek WS, Mu F. Improved predictions of transcription factor binding sites using physicochemical features of DNA. *Nucleic Acids Res*. 2012; 40:e175. [PubMed: 22923524]
- Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium I. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 2013; 18:497–511. [PubMed: 22472876]
- Malhotra AK, Correll CU, Chowdhury NI, Muller DJ, Gregersen PK, Lee AT, Tiwari AK, Kane JM, Fleischhacker WW, Kahn RS, Ophoff RA, Lieberman JA, Meltzer HY, Lencz T, Kennedy JL. Association between common variants near the melanocortin 4 receptor gene and severe antipsychotic drug-induced weight gain. *Arch Gen Psychiatry*. 2012; 69:904–912. [PubMed: 22566560]
- Matigian N, Abrahamsen G, Sutharsan R, et al. Disease-specific, neurosphere-derived cells as models for brain disorders. *Dis Model Mech*. 2010; 3:785–798. [PubMed: 20699480]
- Mattheisen M, Samuels JF, Wang Y, et al. Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol Psychiatry*. 2015; 20:337–344. [PubMed: 24821223]
- Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012; 337:1190–1195. [PubMed: 22955828]
- Maury Y, Come J, Piskorski RA, Salah-Mohellibi N, Chevalyere V, Peschanski M, Martinat C, Nedelec S. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol*. 2015; 33:89–96. [PubMed: 25383599]
- Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. 2005; 33:5868–5877. [PubMed: 16224102]
- Mele M, Ferreira PG, Reverter F, et al. The human transcriptome across tissues and individuals. *Science*. 2015; 348:660–665. [PubMed: 25954002]
- Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG Jr, Kinney JB, Kellis M, Lander ES, Mikkelsen TS. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol*. 2012; 30:271–277. [PubMed: 22371084]
- Merkle FT, Maroof A, Wataya T, Sasai Y, Studer L, Eggan K, Schier AF. Generation of neuropeptidergic hypothalamic neurons from human pluripotent stem cells. *Development*. 2015; 142:633–643. [PubMed: 25670790]
- Miele, A., Dekker, J. Mapping cis- and trans-chromatin interaction networks using chromosome conformation capture (3C). In: Hancock, R., editor. *The Nucleus: Volume 2: Chromatin, Transcription, Envelope, Proteins, Dynamics, and Imaging*. Humana Press; Totowa, NJ, USA: 2008. p. 105-121.
- Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA, Herman B, Happe S, Higgs A, LeProust E, Follows GA, Fraser P, Luscombe NM, Osborne CS. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet*. 2015; 47:598–606. [PubMed: 25938943]
- Mikkelsen TS, Ku M, Jaffe DB, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007; 448:553–560. [PubMed: 17603471]
- Miller JA, Ding SL, Sunkin SM, et al. Transcriptional landscape of the prenatal human brain. *Nature*. 2014; 508:199–206. [PubMed: 24695229]

- Mitchell AC, Bharadwaj R, Whittle C, Krueger W, Mirmics K, Hurd Y, Rasmussen T, Akbarian S. The genome in three dimensions: a new frontier in human brain research. *Biol Psychiatry*. 2014; 75:961–969. [PubMed: 23958183]
- Mo A, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Ulrich MA, Nery JR, Sejnowski TJ, Lister R, Eddy SR, Ecker JR, Nathans J. Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron*. 2015; 86:1369–1384. [PubMed: 26087164]
- Moraes CF, Lins TC, Carmargos EF, Naves JO, Pereira RW, Nobrega OT. Lessons from genome-wide association studies findings in Alzheimer's disease. *Psychogeriatrics*. 2012; 12:62–73. [PubMed: 22416831]
- Muhleisen TW, Leber M, Schulze TG, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun*. 2014; 5:3339. [PubMed: 24618891]
- Munzel M, Globisch D, Bruckl T, Wagner M, Welzmler V, Michalakis S, Muller M, Biel M, Carell T. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem*. 2010; 49:5375–5377. [PubMed: 20583021]
- Naranjo S, Voeselek K, de la Calle-Mustienes E, Robert-Moreno A, Kokotas H, Grigoriadou M, Economides J, Van Camp G, Hilgert N, Moreno F, Alsina B, Petersen MB, Kremer H, Gomez-Skarmeta JL. Multiple enhancers located in a 1-Mb region upstream of POU3F4 promote expression during inner ear development and may be required for hearing. *Hum Genet*. 2010; 128:411–419. [PubMed: 20668882]
- Neale BM, Lasky-Su J, Anney R, et al. Genome-wide association scan of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet*. 2008; 147B:1337–1344. [PubMed: 18980221]
- Neale BM, Medland S, Ripke S, et al. Case-control genome-wide association study of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2010a; 49:906–920. [PubMed: 20732627]
- Neale BM, Medland SE, Ripke S, et al. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2010b; 49:884–897. [PubMed: 20732625]
- Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, Katz E, Dixon JM, Harrison DJ, Meehan RR. Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res*. 2012; 22:467–477. [PubMed: 22106369]
- Ngun TC, Ghahramani N, Sanchez FJ, Bocklandt S, Vilain E. The genetics of sex differences in brain and behavior. *Front Neuroendocrinol*. 2011; 32:227–246. [PubMed: 20951723]
- Ni Z, Abou El Hassan M, Xu Z, Yu T, Bremner R. The chromatin-remodeling enzyme BRG1 coordinates CIITA induction through many interdependent distal enhancers. *Nat Immunol*. 2008; 9:785–793. [PubMed: 18500344]
- Nicholas CR, Chen J, Tang Y, Southwell DG, Chalmers N, Vogt D, Arnold CM, Chen YJ, Stanley EG, Elefanty AG, Sasai Y, Alvarez-Buylla A, Rubenstein JL, Kriegstein AR. Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. *Cell Stem Cell*. 2013; 12:573–586. [PubMed: 23642366]
- Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010; 6:e1000888. [PubMed: 20369019]
- Nityanandam A, Baldwin KK. Advances in reprogramming-based study of neurologic disorders. *Stem Cells Dev*. 2015; 24:1265–1283. [PubMed: 25749371]
- Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. Scanning human gene deserts for long-range enhancers. *Science*. 2003; 302:413. [PubMed: 14563999]
- Nugent BM, Tobet SA, Lara HE, Lucion AB, Wilson ME, Recabarren SE, Paredes AH. Hormonal programming across the lifespan. *Horm Metab Res*. 2012; 44:577–586. [PubMed: 22700441]
- O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet*. 2008; 40:1053–1055. [PubMed: 18677311]
- O'Donovan MC, Craddock NJ, Owen MJ. Genetics of psychosis; insights from views across the genome. *Hum Genet*. 2009; 126:3–12. [PubMed: 19521722]

- Orom UA, Shiekhatter R. Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell*. 2013; 154:1190–1193. [PubMed: 24034243]
- Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytynicki M, Notredame C, Huang Q, Guigo R, Shiekhatter R. Long noncoding RNAs with enhancer-like function in human cells. *Cell*. 2010; 143:46–58. [PubMed: 20887892]
- Owen MJ, O'Donovan MC, Thapar A, Craddock N. Neurodevelopmental hypothesis of schizophrenia. *Br J Psychiatry*. 2011; 198:173–175. [PubMed: 21357874]
- Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N, Program NCS, Black BL, Visel A, Pennacchio LA, Collins FS. National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors, NISC Comparative Sequencing Program Authors. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc Natl Acad Sci USA*. 2013; 110:17921–17926. [PubMed: 24127591]
- Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee SI, Cooper GM, Ahituv N, Pennacchio LA, Shendure J. Massively parallel functional dissection of mammalian enhancers in vivo. *Nat Biotechnol*. 2012; 30:265–270. [PubMed: 22371081]
- Pedrosa E, Sandler V, Shah A, Carroll R, Chang C, Rockowitz S, Guo X, Zheng D, Lachman HM. Development of patient-specific neurons in schizophrenia using induced pluripotent stem cells. *J Neurogenet*. 2011; 25:88–103. [PubMed: 21797804]
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, Plajzer-Frick I, Akiyama J, De Val S, Afzal V, Black BL, Couronne O, Eisen MB, Visel A, Rubin EM. In vivo enhancer analysis of human conserved non-coding sequences. *Nature*. 2006; 444:499–502. [PubMed: 17086198]
- Peters AH, Schubeler D. Methylation of histones: playing memory with DNA. *Curr Opin Cell Biol*. 2005; 17:230–238. [PubMed: 15780602]
- Pineda S, Gomez-Rubio P, Picornell A, Bessonov K, Marquez M, Kogevinas M, Real FX, Van Steen K, Malats N. Framework for the integration of genomics, epigenomics and transcriptomics in complex diseases. *Hum Hered*. 2015; 79:124–136. [PubMed: 26201698]
- Plongthongkum N, Diep DH, Zhang K. Advances in the profiling of DNA modifications: cytosine methylation and beyond. *Nat Rev Genet*. 2014; 15:647–661. [PubMed: 25159599]
- Prabhakar S, Noonan JP, Paabo S, Rubin EM. Accelerated evolution of conserved noncoding sequences in humans. *Science*. 2006; 314:786. [PubMed: 17082449]
- Prokunina L, Castillejo-Lopez C, Oberg F, et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet*. 2002; 32:666–669. [PubMed: 12402038]
- Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 2011; 43:977–983. [PubMed: 21926972]
- Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R, De T, Coin L, de Silva R, Cookson MR, Singleton AB, Hardy J, Ryten M, Weale ME. UK Brain Expression Consortium North American Brain Expression Consortium. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci*. 2014; 17:1418–1428. [PubMed: 25174004]
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013; 8:2281–2308. [PubMed: 24157548]
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014; 159:1665–1680. [PubMed: 25497547]
- Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA. Genome-wide location and function of DNA binding proteins. *Science*. 2000; 290:2306–2309. [PubMed: 11125145]
- Ricano-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet*. 2013; 14:325–353. [PubMed: 23834318]

- Kundaje A, Meuleman W, et al. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015; 518:317–330. [PubMed: 25693563]
- Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods*. 2007; 4:651–657. [PubMed: 17558387]
- Robicsek O, Karry R, Petit I, Salman-Kesner N, Muller FJ, Klein E, Aberdam D, Ben-Shachar D. Abnormal neuronal differentiation and mitochondrial dysfunction in hair follicle-derived induced pluripotent stem cells of schizophrenia patients. *Mol Psychiatry*. 2013; 18:1067–1076. [PubMed: 23732879]
- Rockman MV, Hahn MW, Soranzo N, Goldstein DB, Wray GA. Positive selection on a human-specific transcription factor binding site regulating IL4 expression. *Curr Biol*. 2003; 13:2118–2123. [PubMed: 14654003]
- Sajan SA, Hawkins RD. Methods for identifying higher-order chromatin structure. *Annu Rev Genomics Hum Genet*. 2012; 13:59–82. [PubMed: 22703176]
- Sandborn WJ, Ghosh S, Panes J, Vranic I, Su C, Rousell S, Niezychowski W. Study A. 063 Investigators (2012) Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. *N Engl J Med*. 3921; 367:616–624. [PubMed: 22894574]
- Sandoe J, Eggan K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nat Neurosci*. 2013; 16:780–789. [PubMed: 23799470]
- Santiago M, Antunes C, Guedes M, Sousa N, Marques CJ. TET enzymes and DNA hydroxymethylation in neural development and function – how critical are they? *Genomics*. 2014; 104:334–340. [PubMed: 25200796]
- Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature*. 2012; 489:109–113. [PubMed: 22955621]
- Scharf JM, Yu D, Mathews CA, et al. Genome-wide association study of Tourette’s syndrome. *Mol Psychiatry*. 2013; 18:721–728. [PubMed: 22889924]
- Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res*. 2012; 22:1748–1759. [PubMed: 22955986]
- Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet*. 2011; 43:969–976. [PubMed: 21926974]
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014; 511:421–427. [PubMed: 25056061]
- Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O’Neill LP, Turner BM, Delrow J, Bell SP, Groudine M. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev*. 2004; 18:1263–1271. [PubMed: 15175259]
- Schug J. Using TESS to predict transcription factor binding sites in DNA sequence. *Curr Protoc Bioinformatics*. 2008; Chapter 2(Unit 2.6)
- Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Ulrich MA, Chen H, Lin S, Lin Y, Jung I, Schmitt AD, Selvaraj S, Ren B, Sejnowski TJ, Wang W, Ecker JR. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature*. 2015; 523:212–216. [PubMed: 26030523]
- Sexton T, Kurukuti S, Mitchell JA, Umlauf D, Nagano T, Fraser P. Sensitive detection of chromatin coassociations using enhanced chromosome conformation capture on chip. *Nat Protoc*. 2012; 7:1335–1350. [PubMed: 22722369]
- Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, Wood WG, Higgs DR. Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. *EMBO J*. 1992; 11:4565–4572. [PubMed: 1425590]
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, Ren B. A map of the cis-regulatory sequences in the mouse genome. *Nature*. 2012; 488:116–120. [PubMed: 22763441]

- Shin J, Ming GL, Song H. DNA modifications in the mammalian brain. *Philos Trans R Soc Lond B Biol Sci.* 2014; 369:20130512. [PubMed: 25135973]
- Shull AY, Noonepalle SK, Lee EJ, Choi JH, Shi H. Sequencing the cancer methylome. *Methods Mol Biol.* 2015; 1238:627–651. [PubMed: 25421684]
- Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet.* 2006; 38:1348–1354. [PubMed: 17033623]
- Sisk CL, Zehr JL. Pubertal hormones organize the adolescent brain and behavior. *Front Neuroendocrinol.* 2005; 26:163–174. [PubMed: 16309736]
- Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen CH, Zhang W, Jian X, Wang J, Zhang L, Looney TJ, Zhang B, Godley LA, Hicks LM, Lahn BT, Jin P, He C. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol.* 2011; 29:68–72. [PubMed: 21151123]
- Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu GL, Jin P, He C. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell.* 2013; 153:678–691. [PubMed: 23602153]
- Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CC, O'Donovan MC, Bray NJ, Mill J. Methylomic trajectories across human fetal brain development. *Genome Res.* 2015; 25:338–352. [PubMed: 25650246]
- Spilianakis CG, Flavell RA. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol.* 2004; 5:1017–1027. [PubMed: 15378057]
- Splinter E, de Wit E, van de Werken HJ, Klous P, de Laat W. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods.* 2012; 58:221–230. [PubMed: 22609568]
- Steidl U, Steidl C, Ebralidze A, et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. *J Clin Invest.* 2007; 117:2611–2620. [PubMed: 17694175]
- Stergiakouli E, Hamshere M, Holmans P, Langley K, Zaharieva I, Hawi Z, Kent L, Gill M, Williams N, Owen MJ, O'Donovan M, Thapar A. Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am J Psychiatry.* 2012; 169:186–194. [PubMed: 22420046]
- Stewart SE, Yu D, Scharf JM, et al. Genome-wide association study of obsessive-compulsive disorder. *Mol Psychiatry.* 2013; 18:788–798. [PubMed: 22889921]
- Sun YV. The influences of genetic and environmental factors on methylome-wide association studies for human diseases. *Curr Genet Med Rep.* 2014; 2:261–270. [PubMed: 25422794]
- Sun M, Ma F, Zeng X, Liu Q, Zhao XL, Wu FX, Wu GP, Zhang ZF, Gu B, Zhao YF, Tian SH, Lin B, Kong XY, Zhang XL, Yang W, Lo WH, Zhang X. Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer. *J Med Genet.* 2008; 45:589–595. [PubMed: 18417549]
- Sun W, Zang L, Shu Q, Li X. From development to diseases: the role of 5hmC in brain. *Genomics.* 2014; 104:347–351. [PubMed: 25205306]
- Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI, Vasanthakumar A, Godley LA, Chang Q, Cheng X, He C, Jin P. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci.* 2011; 14:1607–1616. [PubMed: 22037496]
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009; 324:930–935. [PubMed: 19372391]
- Tanaka Y, Maeshima K, Yamaoka K. In vitro and in vivo analysis of a JAK inhibitor in rheumatoid arthritis. *Ann Rheum Dis.* 2012; 71(Suppl 2):i70–i74. [PubMed: 22460142]
- The Psychiatric GWAS Consortium Steering Committee. A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol Psychiatry.* 2009; 14:10–17. [PubMed: 19002139]

- Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K. An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet.* 2003; 35:341–348. [PubMed: 14608356]
- Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development.* 2011; 138:811–820. [PubMed: 21247965]
- Trabzuni D, Ramasamy A, Imran S, Walker R, Smith C, Weale ME, Hardy J, Ryten M. North American Brain Expression Consortium. Widespread sex differences in gene expression and splicing in the adult human brain. *Nat Commun.* 2013; 4:2771. [PubMed: 24264146]
- Ulahannan N, Grealis JM. Genome-wide assays that identify and quantify modified cytosines in human disease studies. *Epigenetics Chromatin.* 2015; 8:5. [PubMed: 25788985]
- Vanhille L, Griffon A, Maqbool MA, Zacarias-Cabeza J, Dao LT, Fernandez N, Ballester B, Andrau JC, Spicuglia S. High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat Commun.* 2015; 6:6905. [PubMed: 25872643]
- Vawter MP, Evans S, Choudary P, Tomita H, Meador-Woodruff J, Molnar M, Li J, Lopez JF, Myers R, Cox D, Watson SJ, Akil H, Jones EG, Bunney WE. Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. *Neuropsychopharmacology.* 2004; 29:373–384. [PubMed: 14583743]
- Verlaan DJ, Berlivet S, Hunninghake GM, et al. Allele-specific chromatin remodeling in the ZBP2/GSDMB/ORMDL3 locus associated with the risk of asthma and autoimmune disease. *Am J Hum Genet.* 2009; 85:377–393. [PubMed: 19732864]
- Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res.* 2012; 22:446–455. [PubMed: 22234890]
- Wagner M, Steinbacher J, Kraus TF, Michalakis S, Hackner B, Pfaffeneder T, Perera A, Muller M, Giese A, Kretschmar HA, Carell T. Age-dependent levels of 5-methyl-, 5-hydroxymethyl-, and 5-formylcytosine in human and mouse brain tissues. *Angew Chem.* 2015; 54:12511–12514. [PubMed: 26137924]
- Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res.* 2010; 124:192–199. [PubMed: 20889312]
- Wang T, Pan Q, Lin L, Szulwach KE, Song CX, He C, Wu H, Warren ST, Jin P, Duan R, Li X. Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum Mol Genet.* 2012; 21:5500–5510. [PubMed: 23042784]
- Wang JL, Shamah SM, Sun AX, Waldman ID, Haggarty SJ, Perlis RH. Label-free, live optical imaging of reprogrammed bipolar disorder patient-derived cells reveals a functional correlate of lithium responsiveness. *Transl Psychiatry.* 2014; 4:e428. [PubMed: 25158003]
- Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012; 40:D930–D934. [PubMed: 22064851]
- Weickert CS, Elashoff M, Richards AB, Sinclair D, Bahn S, Paabo S, Khaitovich P, Webster MJ. Transcriptome analysis of male-female differences in prefrontal cortical development. *Mol Psychiatry.* 2009; 14:558–561. [PubMed: 19455171]
- Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry.* 1987; 44:660–669. [PubMed: 3606332]
- Weirauch MT, Cote A, Norel R, Annala M, Zhao Y, Riley TR, Saez-Rodriguez J, Cokelaer T, Vedenko A, Talukder S, Bussemaker HJ, Morris QD, Bulyk ML, Stolovitzky G, Hughes TR. DREAM5 Consortium. Evaluation of methods for modeling transcription factor sequence specificity. *Nat Biotechnol.* 2013; 31:126–134. [PubMed: 23354101]
- Wen L, Tang F. Genomic distribution and possible functions of DNA hydroxymethylation in the brain. *Genomics.* 2014; 104:341–346. [PubMed: 25205307]

- Wen L, Li X, Yan L, et al. Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol.* 2014a; 15:R49. [PubMed: 24594098]
- Wen Z, Nguyen HN, Guo Z, et al. Synaptic dysregulation in a human iPSC cell model of mental disorders. *Nature.* 2014b; 515:414–418. [PubMed: 25132547]
- van de Werken HJ, de Vree PJ, Splinter E, Holwerda SJ, Klous P, de Wit E, de Laat W. 4C technology: protocols and data analysis. *Methods Enzymol.* 2012a; 513:89–112. [PubMed: 22929766]
- van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Oz Y, Bouwman BA, Verstegen MJ, de Wit E, Tanay A, de Laat W. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods.* 2012b; 9:969–972. [PubMed: 22961246]
- Whitaker JW, Chen Z, Wang W. Predicting the human epigenome from DNA motifs. *Nat Methods.* 2015; 12:265–272. 267 p following 272. [PubMed: 25240437]
- Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 2013; 153:307–319. [PubMed: 23582322]
- Wockner LF, Noble EP, Lawford BR, Young RM, Morris CP, Whitehall VL, Voisey J. Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry.* 2014; 4:e339. [PubMed: 24399042]
- Wright R, Rethelyi JM, Gage FH. Enhancing induced pluripotent stem cell models of schizophrenia. *JAMA Psychiatry.* 2014; 71:334–335. [PubMed: 24382673]
- Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell.* 2012; 148:816–831. [PubMed: 22341451]
- Yoon KJ, Nguyen HN, Ursini G, et al. Modeling a genetic risk for schizophrenia in iPSCs and mice reveals neural stem cell deficits associated with adherens junctions and polarity. *Cell Stem Cell.* 2014; 15:79–91. [PubMed: 24996170]
- Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min JH, Jin P, Ren B, He C. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell.* 2012; 149:1368–1380. [PubMed: 22608086]
- Zahn-Waxler C, Shirliff EA, Marceau K. Disorders of childhood and adolescence: gender and psychopathology. *Annu Rev Clin Psychol.* 2008; 4:275–303. [PubMed: 18370618]
- Zandi PP, Avramopoulos D, Willour VL, Huo Y, Miao K, Mackinnon DF, McInnis MG, Potash JB, Depaulo JR. SNP fine mapping of chromosome 8q24 in bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet.* 2007; 144B:625–630. [PubMed: 17357146]
- Zhao Z, Tavosidana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, Pant V, Tiwari V, Kurukuti S, Ohlsson R. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet.* 2006; 38:1341–1347. [PubMed: 17033624]
- Zhao Z, Li Y, Chen H, Lu J, Thompson PM, Chen J, Wang Z, Xu J, Xu C, Li X. PD_NGSAtlas: a reference database combining next-generation sequencing epigenomic and transcriptomic data for psychiatric disorders. *BMC Med Genomics.* 2014; 7:71. [PubMed: 25551368]
- Zhi D, Aslibekyan S, Irvin MR, Claas SA, Borecki IB, Ordovas JM, Absher DM, Arnett DK. SNPs located at CpG sites modulate genome-epigenome interaction. *Epigenetics.* 2013; 8:802–806. [PubMed: 23811543]
- Zhong H, Beaulaurier J, Lum PY, et al. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet.* 2010; 6:e1000932. [PubMed: 20463879]
- Zhou T, Shen N, Yang L, Abe N, Horton J, Mann RS, Bussemaker HJ, Gordan R, Rohs R. Quantitative modeling of transcription factor binding specificities using DNA shape. *Proc Natl Acad Sci USA.* 2015; 112:4654–4659. [PubMed: 25775564]
- Zou F, Chai HS, Younkin CS, et al. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet.* 2012; 8:e1002707. [PubMed: 22685416]
- Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van Ijcken WF, Grosveld FG, Ren B, Wendt KS. Cohesin and CTCF

differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci USA*. 2014; 111:996–1001. [PubMed: 24335803]

Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosci*. 2005; 23:143–152. [PubMed: 15749241]