# Genetic Variation in Long-Range Enhancers


**John F. Fullard, Samir Rahman, and Panos Roussos**

## Contents



J. F. Fullard
Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Department of Genetics and Genomic Science and Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

S. Rahman
Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Department of Genetics and Genomic Science and Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

P. Roussos (✉)
Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Department of Genetics and Genomic Science and Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Mental Illness Research, Education, and Clinical Center (VISN 2 South), James J. Peters VA Medical Center, Bronx, NY, USA
e-mail: Panagiotis.roussos@mssm.edu



Curr Topics Behav Neurosci (2019) 42: 35–50
DOI 10.1007/7854_2019_110
Published Online: 9 August 2019

**Abstract** *Cis*-regulatory elements (CREs), including insulators, promoters, and enhancers, play critical roles in the establishment and maintenance of normal cellular function. Within each cell, the 3D structure of chromatin is arranged in specific patterns to expose the CREs required for optimal spatiotemporal regulation of gene expression. CREs can act over large distances along the linear genome, facilitated by looping of the intervening chromatin to allow direct interaction between distal regulatory elements and their target genes. A number of pathologies are associated with dysregulation of CRE function, including developmental disorders, cancers, and neuropsychiatric disease. A majority of known neuropsychiatric disease risk loci are noncoding, and increasing evidence suggests that they contribute to disease through disruption of CREs. As such, rather than directly altering the amino acid content of proteins, these variants are instead thought to affect where, when, and to what extent a given gene is expressed. The distances over which CREs can operate often render their target genes difficult to identify. Furthermore, as many risk loci contain multiple variants in high linkage disequilibrium, identification of the causative single nucleotide polymorphism(s) therein is not straightforward. Thus, deciphering the genetic etiology of complex neuropsychiatric disorders presents a significant challenge.

**Keywords** Disease · Enhancers · Epigenome · Transcription

## 1 Introduction

The human body is made up of hundreds of different cell types, each of which shares an essentially identical genome. Human nuclei are typically between 2 and 10 μm in diameter and contain approximately 2 m of DNA. In order to fit inside the nucleus, DNA is packaged into chromatin. This packaging is not random and differs between cell types, leading to the exposure of the distinct repertoire of *cis*-regulatory elements (CREs) required to generate cell-type-specific transcriptomes (reviewed in Cremer and Cremer 2001; Bulger and Groudine 2010; Heinz et al. 2015; Romanoski et al. 2015). Thus, the spatial organization of the genome plays a key role in the regulation of gene expression, the disruption of which may lead to disease (Dekker and Mirny 2016; Dixon et al. 2016; Lupianez et al. 2016).

## 2 Structure of Chromatin and 3D Genome

The fundamental unit of chromatin organization is the nucleosome, an octamer of histone proteins, which accommodates approximately 150 nucleotides of DNA that wrap around the complex. Nucleosomes coalesce into chromatin fibers, which are further compacted into chromosomes. Chromosomes are highly condensed during mitosis but become de-condensed during the interphase of the cell cycle. The

conformation of the genome during interphase, which comprises the major fraction of the cell cycle, during which a cell grows and expresses its genes, is not disorganized but highly structured. The three-dimensional (3D) structure of the genome is implicated in cell-type-specific regulation of gene transcription by facilitating contact between distal regulatory elements called enhancers and target promoters. In 1885, Carl Rabl proposed the "chromosome territory" model, which held that each chromosome occupies a defined volume of the nucleus and only shares nuclear space with a small number of adjacent chromosomes (Rabl 1885). This model was not validated until the 1980s, when Cremer and co-workers induced local DNA damage in the nucleus using a focused laser and, by providing radioactively labeled nucleotides to mark the chromosomal regions as the cells repaired the DNA damage, were able to observe DNA damage on only a few chromosomes (Cremer et al. 1982). Subsequent studies using fluorescent in situ hybridization confirmed that chromosomes occupy distinct regions within the nucleus during interphase (Cremer and Cremer 2001; Habermann et al. 2001; Parada et al. 2002, 2004; Tanabe et al. 2002; Meaburn and Misteli 2007).

The genome is broadly partitioned into megabase-sized compartments that consist of gene-rich regions of open chromatin, corresponding to transcriptionally active regions and, conversely, gene-poor regions of closed chromatin, corresponding to transcriptionally silent regions (Lieberman-Aiden et al. 2009). The genome is further segregated into locally self-interacting neighborhoods that are insulated from each other by boundary elements enriched for the architectural protein CTCF (Dixon et al. 2012; Rao et al. 2014). These neighborhoods, referred to as topologically associated domains (TADs), are cell-type invariant and conserved from mouse to human. The boundary elements insulating TADs are enriched for housekeeping genes, suggesting that other factors, in addition to CTCF, may serve as barriers in genomic organization. TAD boundaries are also known to play a role in gene expression, as deletion or disruption of boundary elements promotes aberrant enhancer-promoter interactions across adjacent TADs, leading to defective transcription (Nora et al. 2012; Lupianez et al. 2015). Indeed, disruption of TAD boundaries can produce severe physiological defects. Genomic deletions and inversions causing human limb malformation were shown to disrupt TAD boundary elements, an observation that was replicated in CRISPR-/Cas9-modified mice (Lupianez et al. 2015). Thus, TADs are evolutionarily conserved and regulate enhancer-promoter specificity.

## 3 Structural Organization of CREs

Within TADs, chromatin structure is variable depending on cell type and cell context. The functional diversity of cells is accomplished, in part, through regulatory elements residing in regions of open chromatin that determine isoform selection of genes during transcription and the timing and extent of their expression (Heinz et al. 2015). In 1981, Banerji and colleagues described a DNA element capable of enhancing transcription in HeLa cells (Banerji et al. 1981). This activity was

mediated by a (virally derived) 72 bp repeat sequence that could act in either orientation and in multiple positions relative to the transcription start site (TSS). The same group subsequently described the first mammalian enhancer (Banerji et al. 1983). This element was shown to influence expression of the beta-globin gene and could function over a broad range of genomic distances (several hundred to several thousand bp, either up- or downstream of the promoter). The element also displayed cell-type specificity as, among the cell lines tested, it was only functional in lymphocyte derived cells. A subsequent boom in enhancer identification occurred, and their mechanism of action in numerous biological processes began to emerge (Shlyueva et al. 2014).

Although the human genome contains hundreds of thousands of potential regulatory elements, only a subset is active in a given cell type at any time (ENCODE Consortium 2012). The distribution of CREs throughout the genome, therefore, provides an additional layer of transcriptional regulation in eukaryotic cells, as loop structures within chromatin facilitate direct and specific interactions between otherwise distal DNA elements, often separated by many kilobases along the linear genome (Gaszner and Felsenfeld 2006; Sanyal et al. 2011, 2012; Benabdallah et al. 2016). The 3D genome differs between cell types, is dynamic, and can change both during development and in response to external stimuli. Regions of open chromatin are enriched in enhancers, which facilitate the specificity of gene expression through interactions with sequence-specific transcription factors (TFs) (Vernimmen and Bickmore 2015). Each gene may be regulated by multiple enhancers depending on the cellular context. Similarly, a single enhancer can also contribute to the expression of a number of different genes, in some cases even if those genes reside on different chromosomes (Spilianakis et al. 2005). Thus, the manner and extent to which the genome is utilized is pivotal in defining the specifics of cellular identity and function.

## 4 Implication of CREs in Disease

Numerous studies link dysregulation of the epigenome to disease (e.g., see Robertson 2005; Williamson et al. 2011; Ward and Kellis 2012; Spielmann and Klopocki 2013; Spielmann and Mundlos 2013; Shen et al. 2014; Albert and Kruglyak 2015; Mirabella et al. 2015; Chatterjee and Ahituv 2017) and genetic variation within the noncoding, regulatory regions of the genome have been associated with a growing list of diseases (Epstein 2009; Manolio et al. 2009; Parker et al. 2013; Scacheri and Scacheri 2015), including developmental disorders (Benko et al. 2009; Gordon et al. 2009), heart disease (Goring et al. 2007; Postma et al. 2016), cancer (Herz et al. 2014; Northcott et al. 2014; He et al. 2015; Khurana et al. 2016; Lin et al. 2016), diabetes (Scott et al. 2007), Crohn's (Libioulle et al. 2007), and asthma (Moffatt et al. 2007; Manolio et al. 2008). In addition, defects in higher-order chromatin structure have also been implicated in disease. For example, mutations

within the cohesin complex (Deardorff et al. 2012; Gervasini et al. 2013), which functions to facilitate interactions between distal promoters and enhancers (Kagey et al. 2010), have been associated with Cornelia de Lange syndrome, which is characterized by physical and cognitive defects.

In the context of neuropsychiatric disease risk, both genetic and environmental factors are known to contribute (Gandal et al. 2016). Genome-wide association studies (GWAS) of complex neuropsychiatric diseases, such as schizophrenia (SCZ; Schizophrenia Working Group of the Psychiatric Genomics 2014), bipolar disorder (BD) (Stahl et al. 2018), Alzheimer's disease (AD) (Lambert et al. 2013), and major depressive disorder (Wray et al. 2018), have identified numerous risk loci, most of which are noncoding. Examining over 35,000 cases and 110,000 controls, the Schizophrenia Working Group of the Psychiatric Genomics Consortium identified 108 genome-wide significant loci associated with increased risk of SCZ (Schizophrenia Working Group of the Psychiatric Genomics 2014). This list has subsequently been expanded, to include over 200 risk loci (ongoing analysis by PGC3), and is expected to grow further with additional studies. The majority of the identified loci are noncoding and are enriched in functional elements including enhancer sequences detected in human brain tissue (Gusev et al. 2014; Roussos et al. 2014). SCZ risk loci also co-localize with expression quantitative trait loci (eQTLs) (in particular, brain-derived eQTLs) (Roussos et al. 2014; Fromer et al. 2016; Hauberg et al. 2017), thereby implicating specific genes. However, these efforts have limited spatiotemporal resolution as most studies performed, thus far, have been restricted to homogenate brain tissue or include only broadly defined neuronal and non-neuronal populations (Fullard et al. 2017). As it is known that CREs display tissue- and cell-type specificity, with variants therein often only affecting those cells and tissues relevant to a given disease (Maurano et al. 2012; Trynka et al. 2013), an imperative of future studies will be to employ cell-type-specific or single-cell assays in order to more comprehensively understand the impact of these variants. For instance, SCZ-associated genetic loci are enriched within promoter and enhancer regions of neuronal cells (Roussos et al. 2014; Fullard et al. 2017). Furthermore, variants located in H3K4me3 (a marker of active promoters) sites specific to neuronal cells were more abundant when compared to those of non-neurons (Tansey and Hill 2018). In the case of AD, single nucleotide polymorphisms (SNPs) associated with increased risk for disease (Lambert et al. 2013) have been shown to be enriched at sites of open chromatin in immune cells and microglia (Tansey et al. 2018). These open chromatin sites contain DNA-binding motifs for specific TFs, including SPI1 and MEF2. Taken together, these observations are consistent with the fact that regulatory elements display tissue- and cell-type specificity. As such, identifying the specific cell types affected by a given variant will be an important step toward a more comprehensive understanding of the genetic etiology of neuropsychiatric disease.

## 5 Large-Scale Efforts to Study the Noncoding Genome

A number of large-scale efforts have sought to examine the role played by the noncoding genome in the regulation of cellular function. The ENCyclopedia Of DNA Elements (ENCODE) project (ENCODE Consortium 2012), the NIH Roadmap Epigenome Mapping Consortium (REMC) (Bernstein et al. 2010; Roadmap Epigenomics et al. 2015), and FANTOM5 (FANTOM Consortium 2014) have made great strides toward systematically cataloguing CREs within the human genome. Unfortunately, much of this data is not readily applicable to the study of gene expression in neuropsychiatric disease. The ENCODE project did not include brain tissue but focused, instead, on a variety of actively dividing cell lines and tissues. Although both REMC and FANTOM5 did include brain specimens, their use was limited to homogenate tissue isolated from controls. Within brain tissue, however, the different subclasses of neurons coexist alongside other cell types, including microglia, oligodendrocytes, and astrocytes. Since CRE-mediated transcriptional regulation has been shown to be cell-type specific (Heintzman et al. 2009; Cheung et al. 2010; Maurano et al. 2012; Roadmap Epigenomics et al. 2015), data derived from such studies is of limited use when applied to the study of a tissue as complex as the human brain. SCZ-associated abnormalities have been demonstrated in specific populations of brain cells, including neocortical neurons (Benes and Berretta 2001), astrocytes (Schnieder and Dwork 2011; McCullumsmith et al. 2015), oligodendrocytes (Haroutunian et al. 2014; Roussos and Haroutunian 2014; Mighdoll et al. 2015), and microglia (Bernstein et al. 2015). As such, the study of homogenate tissue samples may fail to distinguish signals unique to specific cell types, potentially missing critical changes in less abundant cell populations. Large-scale collaborative projects such as PsychENCODE (Akbarian et al. 2015) will attempt to address this issue by applying state-of-the-art omics approaches to specific cell populations (neurons and non-neurons in the case of PsychENCODE), in an effort to further our understanding of the complex genetic mechanisms that contribute both to normal brain function and to disease (see below).

## 6 Approaches to Detect Enhancer Sequences

Although the full extent to which the 3D structure of chromatin influences cell function remains unclear (Cattoni et al. 2015), the advent of methodologies that incorporate chromatin immunoprecipitation and next-generation sequencing technologies (ChIP-seq) is likely to further our understanding of this critical regulatory process (Fullwood et al. 2009; Fullwood and Ruan 2009). ChIP-seq utilizes antibodies against epigenetic markers and can be used to assess enrichment for histone 3Lys4 tri-methylation (H3K4me3), histone 3Lys27 acetylation (H3K27ac), and histone H3 at lysine 4 mono-methylation (H3K4me1) to identify putative active promoters and enhancers and active or primed enhancers, respectively.

Cap analysis of gene expression (CAGE) facilitates fine mapping of TSSs and promoter regions by sequencing the 5′ end of mature RNA (Shiraki et al. 2003; Takahashi et al. 2012). In CAGE-seq, RNA is reverse-transcribed and the 3′ ends biotinylated. Nonhybridized, single-stranded RNAs are removed by digestion with RNase, leaving 5′ complete cDNAs that are captured using streptavidin and subjected to next-generation sequencing. The FANTOM5 project has applied CAGE-seq to a wide array of cells and tissues from human and mouse, including the brain. In so doing, they have identified and quantified the activity of at least one promoter for more than 95% of annotated protein-coding genes in the human reference genome (FANTOM Consortium 2014). In addition, using H3K27ac and H3K4me1 ChIP-seq data from the ENCODE project (ENCODE Consortium 2012), FANTOM5 also applied CAGEseq data to identify a range of enhancers across human cells, through the detection of enhancer RNAs (eRNA) (Andersson et al. 2014). eRNAs are transcribed in proportion to enhancer activity, and their levels correlate with those of mRNA from nearby genes and have been shown to be differentially expressed in SCZ (Hauberg et al. 2018).

The nucleosome is known to play a central role in mediating gene expression and exists in a dynamic equilibrium between open and closed states (Mellor 2005). Nucleosome rearrangement at promoters and enhancers leads to open chromatin states and results from the binding of specific regulatory factors (Henikoff 2008). Open or accessible regions of the genome are regarded as primary positions for regulatory elements and, as such, play a critical role in regulating transcription (John et al. 2011). Approaches such as DNase-seq (DNase I hypersensitivity regions) (Song et al. 2011) or FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements) (Simon et al. 2012) have been utilized to map open chromatin (Maurano et al. 2012); however, these techniques have largely been superseded by methods that require lower amounts of input material, an important consideration when working with precious biological samples of limited availability. More recently, a tagmentation-based method called the Assay for Transposase-Accessible Chromatin followed by Sequencing (ATAC-seq) has been developed (Buenrostro et al. 2013). ATAC-seq employs a transposome complex to insert oligonucleotides into accessible regions of the genome. This facilitates the generation of sequencing libraries enriched for open chromatin with sufficient resolution to map TF occupancy and nucleosome positions in regulatory sites (Buenrostro et al. 2015a). Subsequent iterations of the method allow for chromatin structure to be profiled in as few as 500 cells (Corces et al. 2017) and those isolated from 50 μm sections and frozen archival material (including human brain (Egervari et al. 2017; Fullard et al. 2017)). In addition, the approach has been further optimized to allow chromatin profiling at the single-cell level (Buenrostro et al. 2015b; Lake et al. 2018).

## 7 Approaches to Detect the Long-Range Enhancers

Several methods based on chromosome conformation capture (3C) have been developed to assess the frequency at which any two loci in the genome are in close enough physical proximity to functionally interact (Dekker et al. 2002; Zhao et al. 2006; Dostie et al. 2006). 3C involves cross-linking of interacting DNA segments, followed by digestion with a frequently cutting restriction enzyme. Digested DNA is then religated, cross-links are reversed and the resulting 3C library subjected to PCR using primers that flank putative ligation junctions, thereby assessing the frequency at which otherwise distal genetic elements ligate to one another; a reflection of their physical proximity within chromatin. 3C can detect interactions between distal genetic elements, e.g., between a gene and an enhancer (Simonis et al. 2007; Naumova et al. 2012), and can range from target-specific (3C) to unbiased, genome-wide approaches (Hi-C) (Dekker et al. 2013), including those at the resolution of single cells (Flyamer et al. 2017; Ulianov et al. 2017). Importantly, a number of these approaches have been successfully applied to studies of the human brain (Mitchell et al. 2014; Roussos et al. 2014; Won et al. 2016).

Alterations in chromosomal loop structures have been implicated in neuropsychiatric disease (Bharadwaj et al. 2013, 2014; Roussos et al. 2014). Bharadwaj and colleagues mapped the 3D configuration of a 200 kb stretch of the human genome containing the GAD1 GABA synthesis enzyme gene locus (Bharadwaj et al. 2013), which has previously been implicated in neuropsychiatric disease (Addington et al. 2005; Straub et al. 2007). This led to the identification a 50 Kb loop structure between noncoding intergenic DNA elements and the TSS of the gene encoding GAD1. This loop was further identified in an induced pluripotent stem cell-derived neuronal cell line when compared to skin fibroblasts or undifferentiated pluripotent stem cells, indicating the cell-type specificity of the structure. The loop was decreased in the prefrontal cortex of subjects with SCZ with a concurrent decrease GAD1 expression when compared with controls.

Variations in DNA sequence can have a profound effect on the function of distal regulatory elements. Bharadwaj and co-workers have also described a conserved, methyltransferase dependent loop structure in the prefrontal cortex involved in regulating the expression of the NMDA glutamate receptor, GRIN2B (Bharadwaj et al. 2014). Disruption of the loop resulted in decreased GRIN2B expression and led to impaired cognitive function and working memory defects in a mouse model. In the above examples, the loop structure is conserved between human and mouse, indicating their functional importance.

Disruption of long-range enhancer function has also been implicated in SCZ. Promoter and enhancer sequences are enriched in SCZ variants associated with eQTL (Roussos et al. 2014). Putative physical interactions between noncontiguous proximal and distal regulatory elements have been identified, and subsequent validation experiments confirmed the existence of a functional loop structure between a distal regulatory element and the gene encoding the L-type calcium channel (CACNA1C). Moreover, this regulatory element overlaps a disease risk locus,

confirming a functional link between SCZ-associated noncoding SNPs, the 3D genome, and the regulation of transcription in the brain. More recently, Won and colleagues generated comprehensive 3D maps of chromatin contacts during human corticogenesis, leading to the identification of hundreds of interactions between genes and distal regulatory elements (Won et al. 2016). Integrating such structural data with SCZ GWAS and eQTL data revealed a number of TF and signaling pathways associated with increased vulnerability to the disease.

Taken together, these studies demonstrate that distal regulatory elements appear to be highly dependent on 3D genomic structures to facilitate their interaction with target genes. Thus, there is a growing body of evidence to suggest that functional disruption of long-range intrachromosomal interactions might contribute to disease, thereby accounting for the observation that the majority of known neuropsychiatric disease risk loci reside within noncoding regions of the genome.

## 8 Refining the Search to Identify Active Enhancer Elements

All of the aforementioned approaches allow for the genome-wide identification of potential regulatory elements but fail to directly assess their functionality. Techniques such as STARR-seq (*s*elf-*t*ranscribing *a*ctive *r*egulatory *r*egion sequencing) (Arnold et al. 2013; Muerdter et al. 2015), and iterations thereof (Vanhille et al. 2015; Dao et al. 2017), and FIREWACh (*F*unctional *I*dentification of *R*egulatory *E*lements *W*ithin *A*ccessible *Ch*romatin) (Murtha et al. 2014) allow for the identification of active cell-type-specific regulatory elements, enabling the rapid screening of entire genomes, reviewed in Dailey (2015). STARR-seq and FIREWACh take advantage of the observation that enhancers can work independent of their relative locations and both methods use reporter assays to interrogate DNA populations for elements capable of driving transcription. By directly coupling candidate sequences to enhancer activity, these approaches enable the evaluation of millions of DNA fragments in a single experiment. A critical consideration with each approach, however, is the relevance of the cells used to carry out the assay.

## 9 Linking the Activity of Enhancers to Specific Genes

Having identified active *cis*-regulatory elements within a given cellular context, the next requirement would be to assign their activities to a specific gene or set of genes. 3C-based methodologies are a useful tool toward this purpose, as they allow for the identification of physical interactions between distal genetic elements.

The relevance of long-range chromatin interactions to human disease has been examined in a few recent studies. Jager and colleagues used an oligo capture-based approach to Hi-C (capture Hi-C) to identify chromatin interactions of 14 colorectal cancer risk (CRC) loci (Jager et al. 2015). They demonstrated that 3D contacts are

enriched for enhancers and promoters and CRC-specific TF binding sites. The role of long-range chromatin interactions in neurodevelopmental disorders like SCZ has also been investigated (Bharadwaj et al. 2014; Roussos et al. 2014; Won et al. 2016; Zhang et al. 2018).

SNPs associated with SCZ are predominantly located within distal enhancer elements. Integrating Hi-C data with noncoding SCZ GWAS variants revealed that most SNPs interact with non-proximal genes that are specifically involved in pathways related to brain development and function (Won et al. 2016). Many of these long-range chromatin interactions are detected in neuronal cells, are enriched for specific neuronal functions such as axon guidance and synaptic transmission, but are not detected in non-neuronal cells. CRISPR-/Cas9-mediated genome editing in neural progenitors validated the functional relevance of an enhancer harboring a SCZ-associated SNP that interacts with *FOXG1,* a gene encoding a TF implicated in early brain development. Deletion of the region flanking the SNP led to decreased expression of *FOXG1*, but not of the nearby *PRKD1* locus.

The effect of large deletion copy number variants (CNVs) on both the local and global chromatin interactome was investigated in a recent study (Zhang et al. 2018). Zhang and co-workers examined specific changes in chromatin interactions, histone modifications, and gene expression caused by the 3 MB chromosome 22q11.2 deletion, containing more than 60 known genes, which was shown to be associated with several psychiatric disorders, in particular SCZ and autism spectrum disorders. Hi-C analysis on cell lines derived from patients harboring the 22q11.2 deletion CNV showed that intrachromosomal contacts between regions flanking the deletion were strengthened, accompanied by respective changes in H3K27ac and H3K27me3, and associated with increased and decreased transcription, respectively. In addition, there were pronounced changes in the A/B compartment structure and TAD structure spanning the deletion region. There were also changes in inter-chromosomal *trans*-contacts, some of which could be directly attributed to the deletion CNV itself.

## 10 Concluding Remarks

To better understand these relatively common neuropsychiatric disorders, therefore, requires the systematic study of the regulatory effects of noncoding mutations on gene expression. Toward this end, a number of large collaborative efforts including the CommonMind Consortium (CMC) (Fromer et al. 2016) and PsychENCODE (Akbarian et al. 2015) are underway that focus on the regulatory mechanisms driving gene expression in the human brain. The CommonMind Consortium (www.synapse.org/cmc) aims to generate and analyze large-scale transcriptome data from brain samples, including control specimens and cases with SCZ and BD. The PsychENCODE consortium was established for the purpose of studying the role played by the epigenome in neuropsychiatric diseases (Akbarian et al. 2015). The project primarily focuses on three neuropsychiatric diseases (autism spectrum

disorder, SCZ, and BD). The goal of PsychENCODE is to chart the epigenomic landscape of the brain in a large cohort of cases and controls. This large, collaborative project will employ a variety of methodologies: ChIP-seq will be used to identify putative active promoters and enhancers. In turn, ATACseq will be performed to identify open chromatin regions (Buenrostro et al. 2013), and Hi-C will be used to identify direct interactions between CREs and their downstream target genes (Dekker et al. 2013). The functional relevance of these putative enhancers will be assessed using mouse transgenesis and the STARR-seq assay (Arnold et al. 2013). Importantly, these approaches will be applied to neuronal and non-neuronal nuclei isolated at a number of different developmental time points and from multiple neocortical brain regions. Findings will be validated using neurons derived from induced pluripotent stem cells (iPSCs) and cultured neuronal cells derived from olfactory neuroepithelium (CNON cells).

In summary, there are many exciting findings from recent studies that implicate a clear role for spatial organization of the genome in gene expression regulation, aberrations of which can lead to disease. Much work remains to be done to integrate chromatin organization with gene expression and epigenome data to elucidate the systems biology of neurodevelopmental disorders. However, recent technological advances, including the development of cell-type-specific and single-cell approaches, hold great promise toward identifying the biological pathways and mechanisms that underlie disease.

## References

Addington AM et al (2005) GAD1 (2q31.1), which encodes glutamic acid decarboxylase (GAD67), is associated with childhood-onset schizophrenia and cortical gray matter volume loss. Mol Psychiatry 10:581–588

Akbarian S et al (2015) The PsychENCODE project. Nat Neurosci 18:1707–1712

Albert FW, Kruglyak L (2015) The role of regulatory variation in complex traits and disease. Nat Rev Genet 16:197–212

Andersson R et al (2014) An atlas of active enhancers across human cell types and tissues. Nature 507:455–461

Arnold CD et al (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science 339:1074–1077

Banerji J et al (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27:299–308

Banerji J et al (1983) A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell 33:729–740

Benabdallah NS et al (2016) SBE6: a novel long-range enhancer involved in driving sonic hedgehog expression in neural progenitor cells. Open Biol 6:pii: 160197

Benes FM, Berretta S (2001) GABAergic interneurons: implications for understanding schizophrenia and bipolar disorder. Neuropsychopharmacology 25:1–27

Benko S et al (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet 41:359–364

Bernstein BE et al (2010) The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 28:1045–1048

Bernstein HG et al (2015) Glial cells as key players in schizophrenia pathology: recent insights and concepts of therapy. Schizophr Res 161:4–18
Bharadwaj R et al (2013) Conserved chromosome 2q31 conformations are associated with transcriptional regulation of GAD1 GABA synthesis enzyme and altered in prefrontal cortex of subjects with schizophrenia. J Neurosci Off J Soc Neurosci 33:11839–11851
Bharadwaj R et al (2014) Conserved higher-order chromatin regulates NMDA receptor gene expression and cognition. Neuron 84:997–1008
Buenrostro JD et al (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10:1213–1218
Buenrostro JD et al (2015a) ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol 109:21.29.1–9
Buenrostro JD et al (2015b) Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523:486–490
Bulger M, Groudine M (2010) Enhancers: the abundance and function of regulatory sequences beyond promoters. Dev Biol 339:250–257
Cattoni DI et al (2015) A matter of scale: how emerging technologies are redefining our view of chromosome architecture. Trends Genet 31:454–464
Chatterjee S, Ahituv N (2017) Gene regulatory elements, major drivers of human disease. Annu Rev Genomics Hum Genet 18:45–63
Cheung I et al (2010) Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. Proc Natl Acad Sci U S A 107:8824–8829
ENCODE Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74
FANTOM Consortium (2014) A promoter-level mammalian expression atlas. Nature 507:462–470
Corces MR et al (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14:959–962
Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2:292–301
Cremer T et al (1982) Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. Hum Genet 62:201–209
Dailey L (2015) High throughput technologies for the functional discovery of mammalian enhancers: new approaches for understanding transcriptional regulatory network dynamics. Genomics 106:151–158
Dao LTM et al (2017) Genome-wide characterization of mammalian promoters with distal enhancer functions. Nat Genet 49:1073–1081
Deardorff MA et al (2012) HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature 489:313–317
Dekker J, Mirny L (2016) The 3D genome as moderator of chromosomal communication. Cell 164:1110–1121
Dekker J et al (2002) Capturing chromosome conformation. Science 295:1306–1311
Dekker J et al (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14:390–403
Dixon JR et al (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376–380
Dixon JR et al (2016) Chromatin domains: the unit of chromosome organization. Mol Cell 62:668–680
Dostie J et al (2006) Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res 16:1299–1309
Egervari G et al (2017) Striatal H3K27 acetylation linked to glutamatergic gene dysregulation in human heroin abusers holds promise as therapeutic target. Biol Psychiatry 81:585–594
Epstein DJ (2009) Cis-regulatory mutations in human disease. Brief Funct Genomic Proteomic 8:310–316

Flyamer IM et al (2017) Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature 544:110–114
Fromer M et al (2016) Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci. https://doi.org/10.1038/nn.4399
Fullard JF et al (2017) Open chromatin profiling of human postmortem brain infers functional roles for non-coding schizophrenia loci. Hum Mol Genet 26:1942–1951
Fullwood MJ, Ruan Y (2009) ChIP-based methods for the identification of long-range chromatin interactions. J Cell Biochem 107:30–39
Fullwood MJ et al (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462:58–64
Gandal MJ et al (2016) The road to precision psychiatry: translating genetics into disease mechanisms. Nat Neurosci 19:1397–1407
Gaszner M, Felsenfeld G (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 7:703–713
Gervasini C et al (2013) Molecular characterization of a mosaic NIPBL deletion in a Cornelia de Lange patient with severe phenotype. Eur J Med Genet 56:138–143
Gordon CT et al (2009) Long-range regulation at the SOX9 locus in development and disease. J Med Genet 46:649–656
Goring HH et al (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 39:1208–1216
Gusev A et al (2014) Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. Am J Hum Genet 95:535–552
Habermann FA et al (2001) Arrangements of macro- and microchromosomes in chicken cells. Chromosome Res 9:569–584
Haroutunian V et al (2014) Myelination, oligodendrocytes, and serious mental illness. Glia 62:1856–1877
Hauberg ME et al (2017) Large-scale identification of common trait and disease variants affecting gene expression. Am J Hum Genet 100:885–894
Hauberg ME et al (2018) Differential activity of transcribed enhancers in the prefrontal cortex of 537 cases with schizophrenia and controls. Mol Psychiatry. https://doi.org/10.1038/s41380-018-0059-8
He H et al (2015) Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. Proc Natl Acad Sci U S A 112:6128–6133
Heintzman ND et al (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459:108–112
Heinz S et al (2015) The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol 16:144–154
Henikoff S (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet 9:15–26
Herz HM et al (2014) Enhancer malfunction in cancer. Mol Cell 53:859–866
Jager R et al (2015) Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. Nat Commun 6:6178
John S et al (2011) Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat Genet 43:264–268
Kagey MH et al (2010) Mediator and cohesin connect gene expression and chromatin architecture. Nature 467:430–435
Khurana E et al (2016) Role of non-coding sequence variants in cancer. Nat Rev Genet 17:93–108
Lake BB et al (2018) Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat Biotechnol 36:70–80
Lambert JC et al (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45:1452–1458
Libioulle C et al (2007) Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. PLoS Genet 3:e58

Lieberman-Aiden E et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293
Lin CY et al (2016) Active medulloblastoma enhancers reveal subgroup-specific cellular origins. Nature 530:57–62
Lupianez DG et al (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161:1012–1025
Lupianez DG et al (2016) Breaking TADs: how alterations of chromatin domains result in disease. Trends Genet 32:225–237
Manolio TA et al (2008) A HapMap harvest of insights into the genetics of common disease. J Clin Invest 118:1590–1605
Manolio TA et al (2009) Finding the missing heritability of complex diseases. Nature 461:747–753
Maurano MT et al (2012) Systematic localization of common disease-associated variation in regulatory DNA. Science 337:1190–1195
McCullumsmith RE et al (2015) Cell-specific abnormalities of glutamate transporters in schizophrenia: sick astrocytes and compensating relay neurons? Mol Psychiatry. https://doi.org/10.1038/mp.2015.148
Meaburn KJ, Misteli T (2007) Cell biology: chromosome territories. Nature 445:379–781
Mellor J (2005) The dynamics of chromatin remodeling at promoters. Mol Cell 19:147–157
Mighdoll MI et al (2015) Myelin, myelin-related disorders, and psychosis. Schizophr Res 161:85–93
Mirabella AC et al (2015) Chromatin deregulation in disease. Chromosoma. https://doi.org/10.1007/s00412-015-0530-0
Mitchell AC et al (2014) The genome in three dimensions: a new frontier in human brain research. Biol Psychiatry 75:961–969
Moffatt MF et al (2007) Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 448:470–473
Muerdter F et al (2015) STARR-seq - principles and applications. Genomics. https://doi.org/10.1016/j.ygeno.2015.06.001
Murtha M et al (2014) FIREWACh: high-throughput functional detection of transcriptional regulatory modules in mammalian cells. Nat Methods 11:559–565
Naumova N et al (2012) Analysis of long-range chromatin interactions using chromosome conformation capture. Methods 58:192–203
Nora EP et al (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485:381–385
Northcott PA et al (2014) Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature 511:428–434
Parada LA et al (2002) Conservation of relative chromosome positioning in normal and cancer cells. Curr Biol 12:1692–1697
Parada LA et al (2004) Tissue-specific spatial organization of genomes. Genome Biol 5:R44
Parker SC et al (2013) Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A 110:17921–17926
Postma AV et al (2016) Genetics of congenital heart disease: the contribution of the noncoding regulatory genome. J Hum Genet 61:13–19
Rabl C (1885) Uber Zelltheilung. In: Gegenbaur C (ed) Morphologisches Jahrbuch, vol 10, pp 214–330
Rao SS et al (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159:1665–1680
Roadmap Epigenomics C et al (2015) Integrative analysis of 111 reference human epigenomes. Nature 518:317–330
Robertson KD (2005) DNA methylation and human disease. Nat Rev Genet 6:597–610
Romanoski CE et al (2015) Exploiting genomics and natural genetic variation to decode macrophage enhancers. Trends Immunol 36:507–518

Roussos P, Haroutunian V (2014) Schizophrenia: susceptibility genes and oligodendroglial and myelin related abnormalities. Front Cell Neurosci 8:5
Roussos P et al (2014) A role for noncoding variation in schizophrenia. Cell Rep 9:1417–1429
Sanyal A et al (2011) Chromatin globules: a common motif of higher order chromosome structure? Curr Opin Cell Biol 23:325–331
Sanyal A et al (2012) The long-range interaction landscape of gene promoters. Nature 489:109–113
Scacheri CA, Scacheri PC (2015) Mutations in the noncoding genome. Curr Opin Pediatr 27:659–664
Schizophrenia Working Group of the Psychiatric Genomics C (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature 511:421–427
Schnieder TP, Dwork AJ (2011) Searching for neuropathology: gliosis in schizophrenia. Biol Psychiatry 69:134–139
Scott LJ et al (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316:1341–1345
Shen E et al (2014) Regulation of histone H3K4 methylation in brain development and disease. Philos Trans R Soc Lond Ser B Biol Sci 369:20130514
Shiraki T et al (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A 100:15776–15781
Shlyueva D et al (2014) Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet 15:272–286
Simon JM et al (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc 7:256–267
Simonis M et al (2007) An evaluation of 3C-based methods to capture DNA interactions. Nat Methods 4:895–901
Song L et al (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res 21:1757–1767
Spielmann M, Klopocki E (2013) CNVs of noncoding cis-regulatory elements in human disease. Curr Opin Genet Dev 23:249–256
Spielmann M, Mundlos S (2013) Structural variations, the regulatory landscape of the genome and their alteration in human disease. Bioessays 35:533–543
Spilianakis CG et al (2005) Interchromosomal associations between alternatively expressed loci. Nature 435:637–645
Stahl E et al (2018) Genomewide association study identifies 30 loci associated with bipolar disorder. *bioRxiv*. https://doi.org/10.1101/173062
Straub RE et al (2007) Allelic variation in GAD1 (GAD67) is associated with schizophrenia and influences cortical function and gene expression. Mol Psychiatry 12:854–869
Takahashi H et al (2012) 5′ end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nat Protoc 7:542–561
Tanabe H et al (2002) Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. Proc Natl Acad Sci U S A 99:4424–4429
Tansey KE, Hill MJ (2018) Enrichment of schizophrenia heritability in both neuronal and glia cell regulatory elements. Transl Psychiatry 8:7
Tansey KE et al (2018) Genetic risk for Alzheimer's disease is concentrated in specific macrophage and microglial transcriptional networks. Genome Med 10:14
Trynka G et al (2013) Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat Genet 45:124–130
Ulianov SV et al (2017) Single-cell Hi-C bridges microscopy and genome-wide sequencing approaches to study 3D chromatin organization. Bioessays 39:1700104
Vanhille L et al (2015) High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. Nat Commun 6:6905
Vernimmen D, Bickmore WA (2015) The hierarchy of transcriptional activation: from enhancer to promoter. Trends Genet 31:696–708

Ward LD, Kellis M (2012) Interpreting noncoding genetic variation in complex traits and human disease. Nat Biotechnol 30:1095–1106

Williamson I et al (2011) Enhancers: from developmental genetics to the genetics of common human disease. Dev Cell 21:17–19

Won H et al (2016) Chromosome conformation elucidates regulatory relationships in developing human brain. Nature 538:523–527

Wray NR et al (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 50:668–681

Zhang X et al (2018) Local and global chromatin interactions are altered by large genomic deletions associated with human brain development. bioRxiv. https://doi.org/10.1101/182451

Zhao Z et al (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38:1341–1347