Chromatin Profiling Techniques

Subjects: Neurosciences

Contributor: Anjali Chawla

The genetic architecture of complex traits is multifactorial. Genome-wide association studies (GWASs) have identified risk loci for complex traits and diseases that are disproportionately located at the non-coding regions of the genome. On the other hand, we have just begun to understand the regulatory roles of the non-coding genome, making it challenging to precisely interpret the functions of non-coding variants associated with complex diseases. Additionally, the epigenome plays an active role in mediating cellular responses to fluctuations of sensory or environmental stimuli. However, it remains unclear how exactly non-coding elements associate with epigenetic modifications to regulate gene expression changes and mediate phenotypic outcomes. Therefore, finer interrogations of the human epigenomic landscape in associating with non-coding variants are warranted.

Keywords: complex traits ; complex diseases ; brain ; non-coding ; epigenome ; DNA structure ; open chromatin ; transcription factors ; histone modifications ; chromatin loops

1. Introduction

Complex traits or diseases are considered to be influenced by interactions between environmental stimuli and regulation of multiple genes. Indeed, correlating allelic frequencies with complex trait variations through case-control genome-wide association studies made it abundantly clear that etiological dissection of complex diseases is non-trivial, and complex diseases are pleiotropic and polygenic. [1][2][3]. The etiological complexity of complex traits can be further influenced by the purging of large effect-size disease-mutations via negative selection, especially those present in the coding-regions. Effectively, this can result in small effect-size variants spread across hundreds of functionally-less deterministic regions [4]. Notably, more than 90% of genome-wide significant risk loci are located in the non-coding regions of the genome, which does not produce proteins, rendering their biological roles elusive [1][2][3][4][5][6]. Large-scale initiatives, such as ENCODE (Encyclopedia of DNA Elements) and REC (Roadmap Epigenomic Consortium), systematically catalogued non-coding elements, providing evidence that at least 80% of the genome is indeed functional [5][6]. As such, non-coding risk loci pose a significant challenge in their functional interpretation or in prioritizing causal variants. This is thought to be partly because of very small effect-sizes of putative risk variants and an insufficient statistical power in pinpointing the causal single nucleotide polymorphisms (SNPs) [1][2][3][4]. Therefore, understanding the regulatory roles of the epi/genome remains a priority.

In addition, the epi/genome can be influenced by the environment factors [7][8][9]. In general, changes in our lifestyle, diet, or social cues, can influence adaptive physiological responses and inter-individual variations in gene expression through epigenetic changes. Moreover, phenotypic heterogeneity in complex traits and diseases point towards the impact of private environment-epigenetic interactions [1][2][3][7]. These unique impacts may also lead to epigenetic variations or de novo mutations precipitated by environmental factors. Indeed, long-term epigenetic and transcriptomic changes have been reported in the brain cell-types of individuals with early-life adversity [8][9]. Therefore, we need improved approaches to identify the combined effects of genetic perturbations and environmental exposures in mediating predisposition to complex traits and diseases.

The epigenomic elements can be broadly defined by regions of open chromatin including cis-regulatory elements, such as insulators, promoters, enhancers, and trans-regulatory binding sites for transcription factors (TFs), as well as histone modification marks that orchestrate a regulatory ensemble, under a dynamic chromatin topology, capable of modulating the transcriptome without altering the nucleotide sequences per se. In turn, the chromatin states and structures are largely influenced by heritable, but also reversible, chemical modifications to DNA and histones, collectively referred to as epigenetic modifications [6][7]. The epigenetic and gene expression changes together regulate cell fate decisions during neurodevelopment. Thereby, the inherent cell-type and epigenetic heterogeneity makes it harder to tease-apart precise molecular modifications and masks subtle disease-related changes when investigating tissue homogenates. Indeed, cell-type proportions were found to be a major contributor to gene expression variations in studies employing bulk tissue

homogenates [10]. Hence, single-cell investigations are now increasingly employed over tissue homogenates, to delineate cell-type specific epigenetic programs. Although epigenetic mechanisms like DNA methylation are known to influence gene expression, in this review, we focus on chromatin structure and the respective profiling techniques, including chromatin accessibility, histone modifications, and chromatin topology (**Table 1**), outlining their applications in deciphering the intricate architecture of complex traits and diseases. To our knowledge, this is the most comprehensive review summarizing these techniques, including state-of-the-art approaches to apply them at single-cell resolution in the brain.

The Chromatin Environment

The chromatin is structurally and functionally active (**Figure 1**). The two major structural categories of chromatin, associating with distinct histone modifications, include the open regions of chromatin that associate with active gene regulatory mechanisms, collectively called euchromatin, while the nucleosome-dense heterochromatin states are important for defining transcriptionally-inactive regions. A nucleosome is the basic unit of chromatin; a histone protein octamer that wraps ~147bp of DNA, repeated periodically throughout the genome. The accessibility of chromatin (unwound open chromatin) and/or nucleosome positioning at genomic loci are indicative of their regulatory potential, and can be examined using chromatin accessibility techniques, such as DNase-seq (DNase I hypersensitive sites sequencing) and MNase-seq (micrococcal nuclease digestion of chromatin followed by sequencing) (**Figure 1**B). Typically, active regulatory regions are thought to be depleted of nucleosomes to allow RNA polymerases or TFs to bind in mediating gene expression. Additionally, the DNA around nucleosomes can transiently unwrap to allow regulatory factors to bind, known as “DNA breathing” [11].

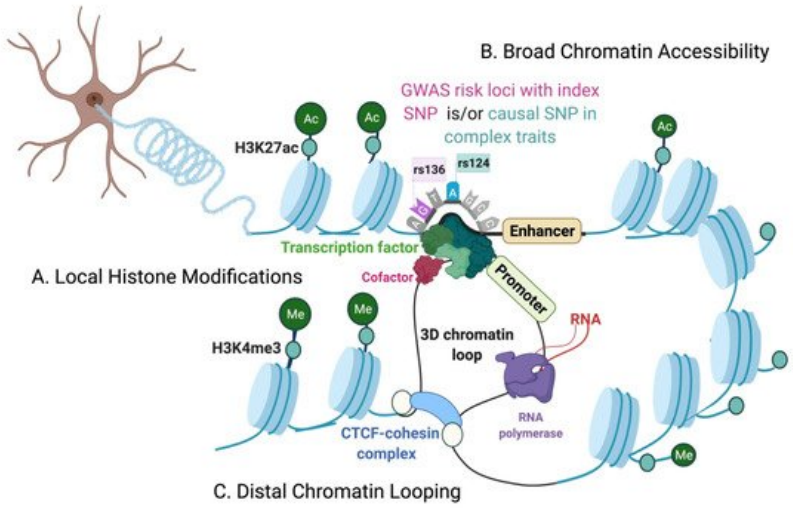

Figure 1. The Chromatin Environment. The open chromatin includes both coding and non-coding aspects of the genome. The interactions between cis- and trans- non-coding, regulatory elements and genes can occur at different genomic scales: locally (such as by histone modifications) or distally (such as by 3-dimesional interactions). The dynamics and functions of the chromatin environment can be mapped using chromatin profiling techniques. (**A**) Local histone modifications (such as acetylation or methylation): induce changes in the chromatin permissiveness, allowing binding of regulatory proteins like transcription factors, impacting expression of the nearby genes. The binding of transcription factors and histone modifications can be assayed using ChIP-seq, CUT&RUN, or CUT&TAG. (**B**) Broad chromatin accessibility: involve significant remodeling of the chromatin landscapes and redistribution of multi-nucleosomes that can directly or indirectly impact expression of multiple genes in the neighborhood. The chromatin environment, cis-regulatory elements and nucleosome distribution can be assayed using ATAC-seq, MNase-seq, DNase-seq, or FAIRE-seq. Genome-wide association studies (GWAS) risk loci for complex traits also largely map to the open non-coding genome, where the index or lead single-nucleotide polymorphism (statistically most significant SNP at a risk loci) may or may not be the disease causative SNP. Identifying regulatory roles of the epigenomic elements associating with risk variants can ascertain causal epi/genetic mechanisms of the complex traits. (**C**) Distal chromatin looping: facilitates long-range gene regulation by DNA elements located farther apart from gene promoters (more than 1–2 kbps), involving 3D changes in the chromatin topology. The spatially interacting genomic regions can be mapped using 3C, 4C, 5C, or HI-C. Additionally, genome-wide chromatin looping interactions of a regulatory protein can be assayed by ChIA-PET, 3C-ChIP, HiChIP, or PLAC-seq.

More stable nucleosome post-translational modifications are facilitated by ATP-dependent chromatin remodeling complexes, such as SWI/SNF (Switch/Sucrose non-fermentable) and nucleosome remodeling and deacetylase complex (NuRD). More commonly, histone-remodeling enzymes, such as histone acetyl- or methyl- transferases, can lead to covalent modifications at the N-terminal tails or core of the histone proteins [12]. The activity of histone remodeling

enzymes in repositioning nucleosomes at chromatin regions can be regulated by the availability of metabolic cofactors. For example, histone acetyltransferases (HAT) depend on acetyl-CoA to neutralize positive charge of lysine-rich histone tails by adding an acetyl group, destabilizing electrostatic interactions with the DNA, and opening the local chromatin. In contrast, the histone deacetylases (HDAC) are dependent on the availability of Zn^{2+} or NAD^{+} cofactors to remove the acetyl groups, restabilizing the chromatin structure [12]. Thereby, histone modifications regulating changes in the chromatin environment are conducive to the binding of transcriptional repressor or activators and can be assessed by ChIP-seq (chromatin immunoprecipitation with sequencing) (**Figure 1**A) and alternative techniques (**Table 1**).

In general, histone modification patterns at regulatory sites, such as at promoters, can effect local chromatin permissiveness to TFs in regulating proximal gene activity, while large-scale histone modifications and nucleosome redistribution either directly or indirectly leading to the remodeling of chromatin accessibility landscapes can impact long-range gene regulation, and can be assessed by ATAC-seq (assay for transposase-accessible chromatin coupled to sequencing) or ChIA-PET (chromatin interaction analysis by paired-end tag sequencing). Moreover, these interactions can be reversible (e.g., to maintain cellular functions) or stable to define cell lineages (e.g., during neurodevelopment) [7].

The non-coding elements commonly effect distal gene expression through 3-dimensional (3D) chromatin interactions or loops, involving shifts in the chromatin topology. Chromatin loops spatially juxtapose functional loci and gene promoters to facilitate long-distance gene expression or insulate genomic regions with diverse chromatin states. These higher-order chromatin interactions can be mapped by chromatin conformation techniques, such as the 3C or Hi-C (**Figure 1**C).

Of note, the CCCTC-binding factor (CTCF) is a transcription factor that colocalize with ring-shaped cohesin complexes to organize the formation of 3D chromatin loops (**Figure 1**C), as well as the topologically associated domains (TADs). TADs are structural units comprising genomic regions with high interaction frequencies. Additionally, the CTCF-cohesin complexes also act as transcriptional insulators, blocking enhancer-promoter interactions, and repressing gene expression. Importantly, genetic mutations in the CTCF complexes are linked to neurodevelopmental delays [13]. Overall, the chromatin-profiling techniques for assaying distinct epigenetic features are thoroughly compared and reviewed (**Table 1**).

Table 1. Comparison of chromatin profiling techniques to assaying epigenomic features.

Epigenomic Features	Techniques	Methods Overview	Benefits	Limitations	Single-Cell and Cell-Types
1. Open chromatin regions. 2. Cis-regulatory elements.	**DNase I hypersensitive sites sequencing (DNase-seq). [14]**	**DNase I digested fragments are extracted using biotin-streptavidin complex.**	**1. High signal-to-noise ratio compared to FAIRE-seq. 2. No prior knowledge of locus-specific sequences, primers, or epitope tags is required. 3. Efficiently maps non-coding regions proximal to genes.**	**1. DNase I sequence-specific cleavage biases may determine cleavage patterns at the predicted transcription factor (TF) binding sites or footprints. This complicates correctly assessing true transcription factor binding at open chromatin. [15] 2. Requires high number of cells (ideally >= 1 M cells) [14] and a high sequencing depth. 3. Maps relatively low distal regulatory sites compared to formaldehyde-assisted isolation of regulatory elements with sequencing (FAIRE-seq). [16]**	**Single-cell (sc)-DNase-seq. [17]**

Epigenomic Features	Techniques	Methods Overview	Benefits	Limitations	Single-Cell and Cell-Types
1. Nucleosome positioning. 2. DNA-bound protein binding sites.	Micrococcal nuclease digestion of chromatin followed by sequencing (MNase-seq) [18], (alternative: nucleosome occupancy and methylome sequencing (NOME-seq). [19]	Cross-linking to covalently link proteins to the DNA, followed by micrococcal nuclease digestion to remove free DNA.	1. MNase-seq can map DNA-protein binding for both histone and non-histone proteins. 2. Indirectly maps chromatin accessibility. 3. The digested fraction of accessible chromatin can be repurposed for chromatin immunoprecipitation-based assays (Native-ChIP).	1. Requires a broad range of sequencing read-out (25 bps to 150 bps) to capture both sub-nucleosome and nucleosome fragments. [20] 2. High dependency on optimized MNase enzyme digestion for reproducibility between experiments. 3. MNase enzyme produces AT cleavage bias that needs bioinformatic corrections. 4. Requires large number of cellular input (ideally >= 1 M cells).	scMNase-seq, and scNOME-seq. [21,22,23,24]
1. Open chromatin. 2. Cis-regulatory elements. 3. Nucleosome distribution.	Assay for transposase-accessible chromatin coupled to sequencing (ATAC-seq). [25]	Tn5 transposases-based cutting and tagging of open chromatin.	1. Low input (ideally <= 50,000 cells) 2. Short and easy to use protocol. 3. Very high signal-to-noise ratio compared to other chromatin accessibility techniques.	1. Tn5 sequence insertion bias can lead to mapping and/or TF footprinting biases and needs bioinformatic corrections. 2. Mitochondrial contamination of reads (although Omni-ATAC [26] is optimized for lower mitochondrial reads).	Flow cytometry-based approaches and single cell/nucleus ATAC-seq. [27,28,29,30,31,32]
1. Protein-DNA interactions. 2. Histone post-translational modification.	Chromatin immunoprecipitation with sequencing (ChIP-seq). [33,34,35]	Formaldehyde crosslinked (X-ChIP) or micrococcal digested fragments (Native-ChIP) followed by immunoprecipitation.	1. Gold standard to map genome-wide, direct DNA-protein interactions. 2. Single-nucleotide resolution (compared to ChIP-qPCR and ChIP-chip). 3. An ultra-low-input micrococcal nuclease-based native ChIP (ULI-NChIP) can profile genome-wide binding sites of histone proteins with as few as 1000 cells. [36]	1. Cross-linking and sonication steps (X-ChIP) can lead to high background noise, requiring higher cellular input for optimal signal-to-noise ratio. [33] 2. Relies on the availability and quality of specific antibodies and can suffer from epitope masking due to cross-linking of fragments (X-ChIP). 3. Requires appropriate control experiments to minimize detection of false-positive protein-DNA binding sites.	sc-ChIP-seq [37]

Epigenomic Features	Techniques	Methods Overview	Benefits	Limitations	Single-Cell and Cell-Types
1. Protein-DNA interactions. 2. Histone post-translational modification.	ChIP with exonuclease (ChIP-exo) [38], Cleavage under targets & release using nuclease (CUT&RUN) [39], Cleavage under targets and tagmentation. (CUT&TAG) [40]	ChIP-exo: X-ChIP immunoprecipitated fragments followed by additional λ exonuclease digestion step. CUT&RUN: MNase tethered protein A, targeting specific antibody against the protein of interest. CUT&TAG: Tn5 transposase and protein A fusion protein, targeting antibody against the protein of interest.	1.ChIP-exo: with an extra exonuclease treatment, it can remove unbound and non-specific DNA, providing higher signal-to-noise ratio over ChIP-seq. [38] 2. CUT&RUN: (i) Uses enzyme-tethering to avoid cross-linking and fragmentation of DNA that greatly reduces the background noise, and epitope masking, making it lower input over ChIP. (ii) It has been validated to map H3K27me3-marked heterochromatin regions. [39] (iii)Use of enzyme-tethering also maps local environment of binding sites, making it suitable to also detect long-range interactions of the protein. 3. CUT&TAG: (i) Requires the least number of cells compared to alternatives (ideally >= 100 cells) and can be performed at single-cell level. [40] (ii) It bypasses cross-linking (compared to ChIP) and library preparation step (compared to ChIP and CUT&RUN). (iii) More sensitive, easier workflow and cost-effective compared to CUT&RUN and alternatives	1. ChIP-exo: High number of enzymatic steps in ChIP-exo makes it technically challenging and suffers from epitope masking, similar to ChIP. 2.CUT&RUN: (i) Calcium-activated MNase enzyme digestion of chromatin needs to be carefully optimized, to prevent over/under digestion of accessible chromatin. It also relies on antibody quality, like ChIP. (ii) Like X-ChIP, CUT&RUN cannot distinguish direct from indirect 3D contacts. [39] (iii) Requires higher number of cells relative to CUT&TAG (ideally >= 100,000 but can be performed with as low as 1000 cells). [39] 3. CUT&TAG: (i) A potential limitation is antibody-validation, since mapping certain protein-DNA interactions can be more efficient after cross-linking. (ii) Tn5 enzyme biases may confound detection of proteins at heterochromatin regions, since Tn5 preferentially tags accessible chromatin	CUT&TAG [40]

Epigenomic Features	Techniques	Methods Overview	Benefits	Limitations	Single-Cell and Cell-Types
3. Chromatin loops and 3D interactions.	Chromosome Conformation Capture 3C [41], 4C [42], 5C [43], and Hi-C. [44]	Formaldehyde cross-linking to covalently link physically interacting chromatin fragments.	3C/4C/5C: these progressive modifications can map increasingly more chromatin conformations, i.e., one-to-one, one-to-many, and many-to-many epigenetic features, respectively. Hi-C (all-to-all): 1. An unbiased approach that maps genome-wide 3D chromatin conformations. 2. Long-range interactions several mega-base pairs away and high-resolution inter-chromosomal contacts can also be mapped. 3. Low cellular input over 3C/4C (ideally >= 1 M cells). Easy-Hi-C: a biotin-free strategy, more sensitive and requires relatively lower cell input over Hi-C (ideally >= 50,000 cells). [45]	3C/4C/5C: 1. Maps to a limited resolution and genomic distances of interacting regions. 2. Need priori-defined regions of interests. 3. Cannot resolve long-range contacts by haplotypes (maternal/paternal) of the chromosomes. 4. Requires relatively higher number of cells (ideally >= 10M cells). Hi-C: (i) It cannot detect chromatin contacts with cell-type specificity and cannot detect functional relevance of the chromatin loops. (ii) Some proximity-ligation events can remain undetected due to low efficiency of biotin incorporation at ligation junctions. [45]	Flow cytometry-based approaches [46,47], sc-Hi-C-seq [48,49], sci-Hi-C-seq [50,51], Dip-C [52]

2. Chromatin Accessibility Techniques

Regions of open chromatin include coding and non-coding aspects of the genome. Interestingly, they harbor the majority of the genome-wide significant risk variants associated with neuropsychiatric disorders [1][2][3], and they are subject to remodeling by neuronal plasticity and therapeutic drugs [14][15]. A number of gene regulatory mechanisms can be investigated through the following techniques.

2.1. DNase I Hypersensitive Sites Sequencing (DNase-seq)

DNase-seq leverages the DNase I enzyme that digests only the open chromatin regions, and not the nucleosome-packed inactive heterochromatin, generating DNase I hypersensitive sites (DHSs). These sites encompass cis-regulatory elements, locus control regions, and transcription factor binding sites, allowing identification of functional non-coding elements. Optimal DNase I digestion is carried out to enrich for the nucleosome-free regions from the isolated nuclei. To reduce random shearing, DNase I digested DNA is embedded in low-melt gel agarose plugs, followed by synthesis of blunt ends. The extracted chromatin is ligated to biotinylated linkers for subsequent enrichment of small DNA fragments using streptavidin columns, followed by PCR amplification and hybridization to microarrays (DNase-Chip) [16] or high-throughput sequencing (DNase-seq) [17].

DNase-based high-throughput analyses of open chromatin have been widely employed to investigate regulatory functions of the non-coding regions and non-coding disease risk loci [5][18][19]. ENCODE initiatives mapped and characterized about 3 million unique DHSs using DNase-seq across hundreds of cell-types. While this represented on an average 1% genome in each cell type, it covered more than 90% ENCODE-identified binding sites of transcription factors [5]. Complex trait and disease risk variants catalogued by the National Human Genome Research Institute (NHGRI), were found to overlap strongly with ENCODE DHSs (34%), the majority of which overlapped with functional enhancers and/or the TSSs. Moreover, up to 71% of complex traits associated SNPs were found to be likely functionally causative in DHSs when

Epigenomic Features	Techniques	Methods Overview	Benefits	Limitations	Single-Cell and Cell-Types
4. Protein-bound 3D interactions	Chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) [53], HiChIP [54] and Proximity ligation-assisted ChIP-seq (PLAC-seq) [55]	Formaldehyde cross-linking, followed by antibody-based immunoprecipitation of protein-bound chromatin interactions.	ChIA-PET, HiChIP & PLAC-seq: Can illustrate regulatory roles of 3D chromatin interactions. HiChIP & PLAC-seq: Higher signal-to-noise ratio and significantly lower cell input compared to ChIA-PET.	ChIA-PET: 1. Low sensitivity in detecting 3D interactions and can have false-positives resulting by non-specific antibody binding. 2. Requires very high number of cellular input (ideally >= 100 M cells) [54,56] and high sequencing depth. 3. Ligation of DNA linkers to chromatin fragments can also lead to self-ligation of linkers and false-positive read-outs. ChIA-PET, HiChIP, and PLAC-seq: They all require a priori of target protein of interest and need bioinformatic correction for biases introduced by: ChIP procedure, different fragment lengths, and restriction enzymes cut-site biases. HiChIP and PLAC-seq also require high cell-number (ideally >= 1 M cells)	Flow cytometry approach [35,57], and multiplex chromatin interaction analysis via droplet-based and sequencing (ChIA-Drop) [58]

those in the linkage disequilibrium (LD; alleles that are non-randomly associated within a population) were included, among which 31% directly overlapped TF binding sites [5]. This demonstrated that the majority of risk SNPs associated with complex traits and diseases could potentially impact regulatory functions of the non-coding elements.

Likewise, collectively employing multiple databases such as ENCODE, REC, and fetal DHSs, resulted in the association of thousands of noncoding SNPs to functional DHS sites, either directly or in LD (76%), for hundreds of complex diseases, and reproducibly, 93% of DHS SNPs overlapped TF binding sites. The candidate disease risk variants were among those that mediated changes in chromatin accessibility and associated with distal gene promoters. The associations of gene promoter with DHSs were based on the significant correlations (Pearson correlation coefficient > 0.7) in their DNase I hypersensitivity signals within 500 kbps radius. This further suggested that functional DHSs that were found to be associated with complex disease risk variants could regulate distal gene promoters [19]. Taken together, these studies described an approach to identify causative SNPs at non-coding regions, whose functions otherwise are not easily understood.

Since the disruption of TF binding sites is considered to be an important mechanism by which non-coding variants mediate disease pathogenesis [5][19], many techniques have been developed for characterizing their binding to the genome. Transcription factor footprinting [20] is one such approach that can predict TF occupancy due to the relative changes in DNase cleavage events created by bound TFs along the genome, generating the resulting footprints. Employing this technique across 29 brain-tissue samples showed that TF binding contributed disproportionately to the heritability of brain-related traits and psychiatric diseases. Further, the TFs associated to those sites were found to be enriched for neurodevelopmental-related functions. However, significant footprints were relatively more variable across test samples compared to other tissue types [20], likely indicating higher cell-type heterogeneity. Therefore, future studies accounting for cellular complexity should reveal deeper insights into precise regulatory mechanisms.

Although footprinting approaches rely on the ability of TF bound sites to be more resistant to cleavage by DNase digestion, accumulating evidence suggests that TFs with shorter DNA residency time leave minimal footprints [21], illustrating a correlation between TF binding kinetics and footprinting depth. Thereby, footprinting predictions can be factor-dependent and should be carefully interpreted at dynamic timescales.

Human-specific DHSs were defined as regions with human-specific increase in DNase-seq signal compared to non-human primates. These DHSs were shown to be cell-type specific (present largely in one cell-type) and primarily enriched at distal enhancers [22]. Notably, species-specific changes to chromatin accessibility correlated with species-specific differences in gene expression and recognition sequences of TFs, such as for activator protein-1 (AP-1), a key activity-dependent TF that modulates synaptic plasticity [22]. Moreover, brain-specific DHSs that show evidence of accelerated evolution (brain-aceDHSs) were enriched for target genes with differential expression between humans and chimpanzees [23]. These brain-aceDHSs also overlapped several human-specific TF motifs, including CTCF and early growth response 1 (EGR1) motifs, important for chromatin organization and activity-dependent functions. Importantly, putative risk SNPs associated with complex traits and brain diseases also overlapped with brain-aceDHSs [23]. Taken together, these studies suggest that at least some gene-regulatory elements at open chromatin landscapes are under adaptive evolution, including those that are fundamental to neurodevelopment and cognition. Further, these regions may also confer risk to neuropsychiatric diseases through unfavorable epi/genetic variations.

A stratified LD score regression can be employed to estimate contributions of functional epigenetic elements to heritability of complex traits. Using this approach, active DHSs were shown to explain higher proportions of complex trait heritability compared to coding regions [24]. Moreover, heritability enrichments for complex traits were cell-type specific, for example, enrichment for psychiatric traits were specific to brain tissues and cell-types that overlapped histone marks associated with open chromatin and functional enhancers. These findings highlight the importance of studying tissue- and cell-specific epigenetic elements in dissecting disease etiology.

To examine cell-type specific differences in epigenomic signatures, a large number of biological replicates are required as produced by ENCODE; however, this may not be feasible for the primary tissues. Furthermore, deconvolution approaches require specific epigenetic markers for distinct cell-types, which remain approximative at best. More sensitive approaches that can allow unbiased cell-type specific investigations are inclusive of single-cell investigations.

Single-cell DNase sequencing (scDNase-seq) has been shown to generate cell-type specific DHSs. Briefly, this method involves flow cytometry based single-cell sorting, DNase I digestion, and addition of circular carrier DNA to minimize loss of digested short fragments, followed by preferential amplification of small DNA fragments and sequencing [25]. This method detected 38 thousand DHSs per cell, and was sufficient to identify cell-type specific enhancers regulating gene

expression programs. Further, this approach was successfully implemented to identify complex disease mutations at regulatory regions effecting target gene expression in specific cell-types [25]. As such, scDNase-seq can be used to identify novel cis-regulatory elements or causal risk SNPs underlying disease phenotypes with cell-type specificity and future work should consider implementing this technique.

2.2. Formaldehyde-Assisted Isolation of Regulatory Elements with Sequencing (FAIRE-seq)

FAIRE-seq, like DNase-seq, maps open regions of the chromatin. It relies on crosslinking protein bound chromatin with formaldehyde followed by nuclei isolation and lysis, sonication, and reversal of cross-links to obtain 200–1000 bp fragments. Finally, phenol-chloroform extraction can separate the organic phase containing unused covalently-linked protein complexes, from the aqueous phase with protein-free DNA. The isolated DNA can subsequently be paired with quantitative amplification (qPCR), hybridized to microarrays, or libraries can be prepared for high-throughput sequencing [26].

A combination of DNase-seq and FAIRE-seq in human cell lines encompassed 9% of human genome across cell-types and captured significantly more TF binding sites than either technique by itself. Despite the mostly overlapping nucleosome-free regions between the two techniques, there is a degree of uniqueness to each approach. FAIRE-seq captured more distal regulatory sites enriched in H3K4me1 histone marks, while DNase-seq captured open regions more proximal to TSSs enriched in H3K4me3 and H3K9ac histone marks. Together, these complementary approaches resulted in a higher-resolution mapping of cis-regulatory elements. Interestingly, open chromatin regions shared across cell lines were generally proximal to TSSs and enriched for CTCF binding sites. On the other hand, open chromatin associated with specific cell types was relatively depleted of CTCF binding sites but enriched for major cell-type defining TFs thought to coordinate cell-type specific gene expression [27]. Therefore, combining profiles of open chromatin regions from these two techniques provides deeper insight into human regulatory epigenome.

The differential properties of the FAIRE-seq and DNase-seq in mapping cis-regulatory elements are likely the result of technical differences. These include distinct regulatory proteins bound at the open chromatin regions that could impact formaldehyde cross-linking in FAIRE-seq. Likewise, relative depletion of nucleosomes proximally to genes may be more susceptible to DNase I digestion [27].

Given the accumulating evidence suggesting that risk SNPs in complex diseases are often located farther from gene bodies [20], FAIRE-seq is useful for probing distal enhancer loci. For example, FAIRE-seq-identified cis-regulatory elements in a patient-based cohort showed that the germline and somatic variants of complex diseases correlated with disruption in TF binding sites at differentially accessible enhancer regions and their accompanied altered gene expression [28]. In addition, these approaches could ascertain clinical sub-categories of the disease. FAIRE-seq combined with ATAC-seq was also used to identify key TFs that regulated distinct stages of disease progression through chromatin remodeling, whereby a loss-of-function mutation in a key disease-related TF decreased severity of the disease [29]. FAIRE-seq is not as widely implemented, possibly due to its inability in determining open chromatin regions bound to regulatory proteins (TF/RNAPII), as a result of formaldehyde cross-linking of DNA-bound proteins. Despite this, FAIRE-seq offers certain advantages, such as circumventing the requirement of an enzymatic step or nuclei suspensions, and can be paired with other chromatin techniques for investigating larger epigenomic landscapes [30].

2.3. Micrococcal Nuclease Digestion of Chromatin Followed by Sequencing (MNase-seq)

One of the most popular methods to determine nucleosome occupancy is MNase-seq. Other similar methods include nucleosome occupancy and methylome sequencing (NOME-seq) that map nucleosome position along with DNA methylation [31] or site-directed chemical cleavage of nucleosomes [32]. MNase-seq employs an endo-exonuclease called the micrococcal nuclease, isolated from *Staphylococcus aureus*, which digests linker DNA and accessible chromatin between nucleosomes, without degrading the nucleosomes. A typical MNase-seq protocol involves crosslinking chromatin with formaldehyde to prevent digestion of histone bound DNA, nuclei isolation, micrococcal digestion to remove free DNA. Subsequently, cross-linking is reversed, and proteinase K digestion is used to release histone proteins. DNA is extracted with phenol-chloroform or spin columns and used as input for microarrays [33], or high-throughput sequencing [34][35].

Employing MNase-seq in human cell lines showed that nucleosome occupancy is dependent on distinct DNA methylation and histone modification patterns [36]. For example, H3K4me3-histone marks, associated with active promoters, were generally depleted of nucleosomes, while H3K9me3-marked inactive epigenetic elements had relatively higher nucleosome occupancy [36]. On the other hand, distinct nucleosome distribution at TF binding sites can determine lineage-specific TFs. An increased nucleosome occupancy at binding sites of Stat3 and p300 TFs was found in the lineage-committed cells compared to embryonic stem cells and neural progenitor cells (NPCs) [37]. Interestingly, combining

ENCODE ChIP-seq and MNase-seq datasets led to the development of an unsupervised chromatin pattern discovery tool that predicted asymmetry and heterogeneity in distribution of nucleosomes and histone modifications flanking distinct classes of TF binding sites [38].

In general, and on an average across cell-types, most eukaryotic chromatin has a nucleosome repeat length of 185–195 bp, corresponding to ~147 bp of nucleosome DNA and ~45 bp of linker DNA. However, nucleosome spacing can also be indicative of specific cell-types and/or disease-states. For example, MNase-seq in distinct cell-types identified a shorter average nucleosome spacing in dorsal root ganglia neurons (~165 bp) compared to cortical astrocytes or oligodendrocyte precursor cells (~183 bp) [39]. Another study depicted age-dependent effects on nucleosome spacing and reported that nucleosome spacing on an average increased with age (up to 50 bp) in mammalian cortical and cerebellar neurons, but not in the glial cell-types [40]. As such, epigenetic changes (such as DNA methylation) have been shown to correlate with ageing process [9]. Given that precise nucleosome spacing at regulatory sites is an important determinant of transcriptome, it will be important to test, whether and to what extent, age-dependent changes in the neuronal epigenome relate to age-related changes in synaptic functions.

MNase-TSSs sequence capture is a modified technique to map nucleosome distribution surrounding only TSSs at a genome-wide scale. This approach identified nucleosome relocation around TSSs at early stages of the disease. This, in turn, was associated with aberrantly high TF binding and disruption of gene expression programs that mediate disease progression [41]. Moreover, alterations to nucleosome occupancy around gene TSSs has been associated with both neurological [42] and psychiatric diseases [43]. Chromatin remodelers can increase nucleosome density, displacing RNAPII and leading to gene silencing [44]. Moreover, mutations in chromatin remodelers have been reproducibly associated with neurodevelopmental and psychiatric disorders [44][45]. Taken together, nucleosome turnover by chromatin remodeling factors can impact interactions at cis-regulatory elements, dysregulating target gene expression.

Combining human *de novo* mutation datasets with MNase-seq-derived nucleosome maps revealed that non-coding regions at/around translationally stable nucleosome positioning across cell-types associate with significantly higher de novo mutation rates, INDELs, repeat elements, and a lower DNA replication fidelity of those sites [46]. This further suggests that nucleosome positioning may be an important factor in determining DNA mutation rate variations, which associate with numerous complex traits and diseases.

Recently, single-cell MNase-seq has been able to obtain nucleosome positioning and chromatin accessibility profiles from single cells [47]. Briefly, fluorescence assisted cell (FAC)-sorting of single cells can be paired with native or fixed cells and micrococcal nuclease digestion of single-cell or bulk cell suspension can be carried out depending on the amount of starting material, followed by phenol-chloroform extraction of DNA fragments. Isolated DNA is ligated with specific adapters for PCR amplifications and subsequently purified for high-throughput sequencing [47]. This approach revealed nucleosome organizing principles of cell-types, not evident in bulk MNase-seq. For example, smaller variations in the positioning of nucleosomes were detected within single cells and cell-types than those found across different cell-types. Furthermore, scMNase-seq demonstrated that the nucleosomes surrounding both the active DHSs and transcription start sites of active genes showed less positional variance across different cell-types and correlated with variations in gene expression, as compared to inactive DHSs or silenced genes [48].

Other single-cell methods include scNOMe-seq that can measure both nucleosome occupancy and DNA methylation at a genome-wide scale [49]. Multi-omics approaches, such as scNMT-seq (single-cell nucleosome, methylation and transcription sequencing), can directly identify impacts of nucleosome positioning on transcriptomic regulation at the single cell level [50]. These techniques have allowed us to integrate different but complementary levels of genomic information, providing multimodal signatures for a given cell.

2.4. Assay for Transposase-Accessible Chromatin (ATAC-seq)

ATAC-seq can capture multi-nucleosome regions of open chromatin using at least 10 times less nuclei and can obtain a higher signal-to-noise ratio compared to the previously described DNase, FAIRE, or MNase-seq. Introduced by Buenrsotro et.al, ATAC-seq requires a prokaryotic Tn5 transposase charged with point mutations to increase its enzymatic activity and adaptors to tag accessible chromatin. Tn5 transposase is applied to the isolated nuclei in bulk. Specific primer pairs can be used to amplify the cut and tagged segments of DNA, which is then followed by high-throughput sequencing. A successful ATAC-seq library shows a laddering pattern with 200 bp periodicity, corresponding to segments of DNA devoid of one (200 bp) or more nucleosomes [51]. With slight modifications, such as the use of multiple detergents and post-lysis nuclei washing with Tween-20, Omni-ATAC-seq is optimized for long-term frozen tissues and attains lower mitochondrial contamination. The use of this adapted protocol with postmortem brain tissue showed enrichments for neurological and psychiatric disease associated risk variants in regions of open chromatin [52].

ATAC-seq has become a popular technique for studying DNA structure, not only because of its ease of use, but also because of its robust findings. For example, the Common Mind Consortium (CMC)-led study in postmortem human brain identified about 9% SNP heritability in schizophrenia in the open regions of chromatin. In addition, a four-fold increase in the SNP heritability for this illness was found when including evolutionarily conserved open regions [53]. Interestingly, differences in accessibility across open regulatory regions appear to be significantly influenced by age and disease phenotypes. Cellular maturation influences the closing of regulatory loci enriched for motifs important for activity-dependent dendritic patterning and NPCs self-renewal. Schizophrenia-related phenotypic alterations were correlated with changes in open chromatin enriched in motifs important for neurogenesis and myelin regeneration [53]. Furthermore, many quantitative trait loci (QTLs) that were found to impact chromatin accessibility changes in the brains of individuals with schizophrenia, showed concordant effects with QTLs effecting gene expression changes (eQTLs), suggesting an association of specific alleles and chromatin states with gene expression alterations in diseased phenotypes. Of note, this study used a very large sample-size, but did not correct for cell-type heterogeneity in chromatin states [53].

Since ATAC-seq can be performed on small amounts of material, researchers have successfully used fluorescence-activated nuclei sorting (FANS) to isolate broad cell types based on antibodies against specific cell markers. Generating neuronal (NeuN+) and non-neuronal (NeuN-) populations from postmortem brain regions of healthy individuals showed that individual cell-types capture more than 50% of the variance in open chromatin brain regions, in contrast to biological sex that accounted for less than 2% variance [54]. Additionally, the neuronal open chromatin showed less overlap with the bulk DHSs than non-neuronal cells, potentially indicating higher variability among neuronal subtypes. Moreover, open chromatin regions of neurons were mostly distal and intergenic with more variable profiles across brain regions than non-neuronal open chromatin [54], suggesting region-specific distal gene regulation in neurons.

Overlapping risk loci with open chromatin regions revealed that neurons from the striatum and hippocampus were enriched for schizophrenia risk variants, while non-neuronal hippocampal regions were enriched for risk variants associated with major depressive disorder (MDD) [54]. Likewise, an organoid model of forebrain development (cell sorted by FACS) depicted both time- and lineage-specific accessibility patterns that correlated with distal enhancer accessibility (+/− 500 kbps of TSSs) of glial and neuronal marker gene expression. In terms of disease association, schizophrenia-associated risk variants were enriched across mature neuronal or non-neuronal cell-types, while those for autism spectrum disorders were enriched primarily in progenitor glial cells [55], further highlighting the importance of employing cell-type specific modalities.

Combining ATAC-seq with a more refined FANS approach by sorting for glutamatergic neurons, GABAergic neurons, oligodendrocytes, and microglia/astrocytes resulted in cell-type specific differentially open coding- and noncoding-regions [56]. For example, differentially open chromatin overlapping *Bdnf* gene was found in the glutamatergic neurons, while open chromatin of *Lhx6* gene was detected in the GABAergic neurons. In addition, cell-type specific open chromatin overlapped with regulatory regions of cell-type specific marker genes. Further, TF footprinting using ATAC-seq, such as DNase-seq, can predict binding of TFs at open chromatin. The footprinted TFs were associated with target genes by the distance of TF binding sites to TSSs. Moreover, the target genes of cell-type specific TFs were among those with cell-type specific open chromatin [56]. These results elucidate the role of accessible chromatin in influencing cellular transcriptome.

The open chromatin regions in glutamatergic neurons showed strong enrichments for risk variants associated with psychiatric phenotypes including schizophrenia and brain-related traits like neuroticism and intelligence [56]. Moreover, cell-type deconvolution of bulk ATAC-seq from the brains of individuals with schizophrenia [53] using cell-type open chromatin signatures identified in this study, further implicated glutamatergic cell-type in pathology of schizophrenia [56]. On the other hand, microglia/astrocytes cell types were enriched for Alzheimer's disease (AD) risk related SNPs. Together, these findings support the need to acquire cell-specific epigenome when investigating complex phenotypes [56].

Single-cell or nucleus ATAC-sequencing (sc/sn-ATAC-seq) can capture cells that cannot be isolated through gene markers (i.e., FANS based isolation), as well as identify landscapes of rare cell-types and/or cell-states. Using the principles of bulk ATAC-seq, scATAC-seq requires a fluidics-based chip, where single cells are captured into individual wells, followed by Tn5 transposition and amplification. Single-cells are then barcoded for cell-identification, and subsequently pooled for library generation and next-generation sequencing (NGS) [57]. Alternatively, a high-throughput droplet-based sequencing can be done using 10x chromium microfluidics, where cells are transposed in bulk, and then isolated with a gel bead matrix so every region of open chromatin from a given cell is tagged with a unique 16 bp cell specific barcode sequence. This approach was used to profile distinct regions of the developing human forebrain, revealing regulatory mechanisms essential for neurogenesis with cell-type and cell-state specific chromatin landscapes and those associating with germline and de novo disease risk variants of complex psychiatric traits [58].

A plate-based combinatorial barcoding approach called sci-ATAC-seq was established to allow multiplexing of high numbers of cells/nuclei. First, one-to-few nuclei are tagged with barcoded Tn5 in a single well of a 96-well plate, and then it is followed by a fixed number of successive barcoding events with different barcode and pools of nuclei, enabling multiplexing of cells, making it scalable and cost-efficient [59]. This approach was used to develop an atlas of 45 distinct brain regions from the adult mice, identifying almost 492,000 cis-regulatory elements, which could define 160 cell-type clusters [60]. The majority of the cis-regulatory elements (96%) were located at least 1kbp away from promoter regions. Among 1% of invariant cis-regulatory elements across the cell-types, 80% were at promoters and others mainly at CTCF binding sites. The open chromatin from mice leveraged with coordinates converted to human genome, revealed significant overlaps of complex brain disease risk variants with open chromatin regions with both regional and cell-type specificity [60].

The use of bulk-ATAC-seq captured minimal enrichments for Alzheimer's or Parkinson's disease associated risk variants, however, combining it with snATAC-seq revealed five-fold enrichment of SNPs overlaying regions of open chromatin at cell-type specific regulatory loci [61]. Further, SNP heritability for Alzheimer's and Parkinson's were mainly predicted to occur in microglial cells. Both microglia-specific TF binding sites and gene targets were found to be enriched for risk SNPs, while heritability for other neurological or psychiatric traits were mostly predicted in distinct neuronal cell-types [61]. These findings strongly point towards the importance of using single-cell techniques when studying complex disorders of the brain.

Taken together, the general patterns of chromatin accessibility and disease enrichments consistently show distal regulation of cell-type specific genes. Risk variants for psychosis-associated diseases are mainly enriched in the open regions of neurons, while neurodegenerative disease variants occur more consistently in open chromatin regions of non-neuronal cell-types. These findings hold true across distinct chromatin accessibility measuring approaches [62][63][64].

References

1. Sullivan, P.F.; Daly, M.J.; O'Donovan, M. Genetic architectures of psychiatric disorders: The emerging picture and its implications. Nat. Rev. Genet. 2012, 13, 537–551.
2. Tak, Y.G.; Farnham, P.J. Making sense of GWAS: Using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenet. Chromatin 2015, 8, 57.
3. Ripke, S.; Neale, B.M.; Corvin, A.; Walters, J.T.R.; Farh, K.-H.; Holmans, P.A.; Lee, P.; St Clair, D.; Weinberger, D.R.; Wendland, J.R.; et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014, 511, 421–427.
4. O'Connor, L.J.; Schoech, A.P.; Hormozdiari, F.; Gazal, S.; Patterson, N.; Price, A.L. Extreme polygenicity of complex traits is explained by negative selection. Am. J. Hum. Genet. 2019, 105, 456–476.
5. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012, 489, 57–74.
6. Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Milosavljevic, A.; Ren, B.; Stamatoyannopoulos, J.A.; et al. Integrative analysis of 111 reference human epigenomes. Nature 2015, 518, 317–330.
7. Cavalli, G.; Heard, E. Advances in epigenetics link genetics to the environment and disease. Nature 2019, 571, 489–499.
8. Lutz, P.-E.; Tanti, A.; Gasecka, A.; Barnett-Burns, S.; Kim, J.J.; Zhou, Y.; Chen, G.G.; Wakid, M.; Shaw, M.; Almeida, D.; et al. Association of a history of child abuse with impaired myelination in the anterior cingulate cortex: Convergent epigenetic, transcriptional, and morphological evidence. AJP 2017, 174, 1185–1194.
9. Labonté, B.; Suderman, M.; Maussion, G.; Navaro, L.; Yerko, V.; Mahar, I.; Bureau, A.; Mechawar, N.; Szyf, M.; Meaney, M.J.; et al. Genome-wide epigenetic regulation by early-life trauma. Arch. Gen. Psychiatry 2012, 69, 722–731.
10. Patrick, E.; Taga, M.; Ergun, A.; Ng, B.; Casazza, W.; Cimpean, M.; Yung, C.; Schneider, J.A.; Bennett, D.A.; Gaiteri, C.; et al. Deconvolving the contributions of cell-type heterogeneity on cortical gene expression. PLoS Comput. Biol. 2020, 16, e1008120.
11. Li, G.; Levitus, M.; Bustamante, C.; Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. Nat. Struct. Mol. Biol. 2005, 12, 46–53.
12. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications. Cell Res. 2011, 21, 381–395.
13. DDD Study; Konrad, E.D.H.; Nardini, N.; Caliebe, A.; Nagel, I.; Young, D.; Horvath, G.; Santoro, S.L.; Shuss, C.; Ziegler, A.; et al. CTCF variants in 39 individuals with a variable neurodevelopmental disorder broaden the mutational

and clinical spectrum. Genet. Med. 2019, 21, 2723–2733.

14. Gallegos, D.A.; Chan, U.; Chen, L.-F.; West, A.E. chromatin regulation of neuronal maturation and plasticity. Trends Neurosci. 2018, 41, 311–324.
15. Pang, B.; Qiao, X.; Janssen, L.; Velds, A.; Groothuis, T.; Kerkhoven, R.; Nieuwland, M.; Ovaa, H.; Rottenberg, S.; van Tellingen, O.; et al. Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. Nat. Commun. 2013, 4, 1908.
16. Crawford, G.E.; Davis, S.; Scacheri, P.C.; Renaud, G.; Halawi, M.J.; Erdos, M.R.; Green, R.; Meltzer, P.S.; Wolfsberg, T.G.; Collins, F.S. DNase-Chip: A high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. Nat. Methods 2006, 3, 503–509.
17. Song, L.; Crawford, G.E. DNase-Seq: A high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. Cold Spring Harb. Protoc. 2010, 2010.
18. Dorschner, M.O.; Hawrylycz, M.; Humbert, R.; Wallace, J.C.; Shafer, A.; Kawamoto, J.; Mack, J.; Hall, R.; Goldy, J.; Sabo, P.J.; et al. High-throughput localization of functional elements by quantitative chromatin profiling. Nat. Methods 2004, 1, 219–225.
19. Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; et al. Systematic localization of common disease-associated variation in regulatory DNA. Science 2012, 337, 1190–1195.
20. Funk, C.C.; Casella, A.M.; Jung, S.; Richards, M.A.; Rodriguez, A.; Shannon, P.; Donovan-Maiye, R.; Heavner, B.; Chard, K.; Xiao, Y.; et al. Atlas of transcription factor binding sites from ENCODE DNase hypersensitivity data across 27 tissue types. Cell Rep. 2020, 32, 108029.
21. Sung, M.-H.; Guertin, M.J.; Baek, S.; Hager, G.L. DNase footprint signatures are dictated by factor dynamics and DNA sequence. Mol. Cell 2014, 56, 275–285.
22. Shibata, Y.; Sheffield, N.C.; Fedrigo, O.; Babbitt, C.C.; Wortham, M.; Tewari, A.K.; London, D.; Song, L.; Lee, B.-K.; Iyer, V.R.; et al. Extensive evolutionary changes in regulatory element activity during human origins are associated with altered gene expression and positive selection. PLoS Genet. 2012, 8, e1002789.
23. Lu, Y.; Wang, X.; Yu, H.; Li, J.; Jiang, Z.; Chen, B.; Lu, Y.; Wang, W.; Han, C.; Ouyang, Y.; et al. Evolution and comprehensive analysis of DNaseI hypersensitive sites in regulatory regions of primate brain-related genes. Front. Genet. 2019, 10, 152.
24. ReproGen Consortium; Schizophrenia Working Group of the Psychiatric Genomics Consortium; The RACI Consortium; Finucane, H.K.; Bulik-Sullivan, B.; Gusev, A.; Trynka, G.; Reshef, Y.; Loh, P.-R.; Anttila, V.; et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 2015, 47, 1228–1235.
25. Jin, W.; Tang, Q.; Wan, M.; Cui, K.; Zhang, Y.; Ren, G.; Ni, B.; Sklar, J.; Przytycka, T.M.; Childs, R.; et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature 2015, 528, 142–146.
26. Giresi, P.G.; Kim, J.; McDaniell, R.M.; Iyer, V.R.; Lieb, J.D. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res. 2007, 17, 877–885.
27. Song, L.; Zhang, Z.; Grasfeder, L.L.; Boyle, A.P.; Giresi, P.G.; Lee, B.-K.; Sheffield, N.C.; Graf, S.; Huss, M.; Keefe, D.; et al. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res. 2011, 21, 1757–1767.
28. Koues, O.I.; Kowalewski, R.A.; Chang, L.-W.; Pyfrom, S.C.; Schmidt, J.A.; Luo, H.; Sandoval, L.E.; Hughes, T.B.; Bednarski, J.J.; Cashen, A.F.; et al. Enhancer sequence variants and transcription-factor deregulation synergize to construct pathogenic regulatory circuits in B-cell lymphoma. Immunity 2015, 42, 186–198.
29. Davie, K.; Jacobs, J.; Atkins, M.; Potier, D.; Christiaens, V.; Halder, G.; Aerts, S. Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-Seq and FAIRE-Seq open chromatin profiling. PLoS Genet. 2015, 11, e1004994.
30. Simon, J.M.; Giresi, P.G.; Davis, I.J.; Lieb, J.D. Using FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) to isolate active regulatory DNA. Nat. Protoc. 2012, 7, 256–267.
31. Kelly, T.K.; Liu, Y.; Lay, F.D.; Liang, G.; Berman, B.P.; Jones, P.A. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res. 2012, 22, 2497–2506.
32. Voong, L.N.; Xi, L.; Wang, J.-P.; Wang, X. Genome-wide mapping of the nucleosome landscape by micrococcal nuclease and chemical mapping. Trends Genet. 2017, 33, 495–507.
33. Ozsolak, F.; Song, J.S.; Liu, X.S.; Fisher, D.E. High-throughput mapping of the chromatin structure of human promoters. Nat. Biotechnol. 2007, 25, 244–248.

34. Ramachandran, S.; Ahmad, K.; Henikoff, S. Transcription and remodeling produce asymmetrically unwrapped nucleosomal intermediates. Mol. Cell 2017, 68, 1038–1053.e4.

35. Henikoff, J.G.; Belsky, J.A.; Krassovsky, K.; MacAlpine, D.M.; Henikoff, S. Epigenome characterization at single base-pair resolution. Proc. Natl. Acad. Sci. USA 2011, 108, 18318–18323.

36. Yazdi, P.G.; Pedersen, B.A.; Taylor, J.F.; Khattab, O.S.; Chen, Y.-H.; Chen, Y.; Jacobsen, S.E.; Wang, P.H. Nucleosome organization in human embryonic stem cells. PLoS ONE 2015, 10, e0136314.

37. Teif, V.B.; Vainshtein, Y.; Caudron-Herger, M.; Mallm, J.-P.; Marth, C.; Höfer, T.; Rippe, K. Genome-wide nucleosome positioning during embryonic stem cell development. Nat. Struct. Mol. Biol. 2012, 19, 1185–1192.

38. Kundaje, A.; Kyriazopoulou-Panagiotopoulou, S.; Libbrecht, M.; Smith, C.L.; Raha, D.; Winters, E.E.; Johnson, S.M.; Snyder, M.; Batzoglou, S.; Sidow, A. Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. Genome Res. 2012, 22, 1735–1747.

39. Clark, S.C.; Chereji, R.V.; Lee, P.R.; Fields, R.D.; Clark, D.J. Differential nucleosome spacing in neurons and glia. Neurosci. Lett. 2020, 714, 134559.

40. Berkowitz, E.M.; Sanborn, A.C.; Vaughan, D.W. Chromatin structure in neuronal and neuroglial cell nuclei as a function of age. J. Neurochem. 1983, 41, 516–523.

41. Druliner, B.R.; Vera, D.; Johnson, R.; Ruan, X.; Apone, L.M.; Dimalanta, E.T.; Stewart, F.J.; Boardman, L.; Dennis, J.H. Comprehensive nucleosome mapping of the human genome in cancer progression. Oncotarget 2015, 7, 13429–13445.

42. Chutake, Y.K.; Costello, W.N.; Lam, C.; Bidichandani, S.I. Altered nucleosome positioning at the transcription start site and deficient transcriptional initiation in friedreich ataxia. J. Biol. Chem. 2014, 289, 15194–15202.

43. Sun, H.; Damez-Werno, D.M.; Scobie, K.N.; Shao, N.-Y.; Dias, C.; Rabkin, J.; Koo, J.W.; Korb, E.; Bagot, R.C.; Ahn, F.H.; et al. ACF chromatin-remodeling complex mediates stress-induced depressive-like behavior. Nat. Med. 2015, 21, 1146–1153.

44. Goodman, J.V.; Bonni, A. Regulation of neuronal connectivity in the mammalian brain by chromatin remodeling. Curr. Opin. Neurobiol. 2019, 59, 59–68.

45. Ronan, J.L.; Wu, W.; Crabtree, G.R. From neural development to cognition: Unexpected roles for chromatin. Nat. Rev. Genet. 2013, 14, 347–359.

46. Li, C.; Luscombe, N.M. Nucleosome positioning stability is a significant modulator of germline mutation rate variation across the human genome. Genomics 2018. preprint.

47. Gao, W.; Lai, B.; Ni, B.; Zhao, K. Genome-wide profiling of nucleosome position and chromatin accessibility in single cells using scmnase-seq. Nat. Protoc. 2020, 15, 68–85.

48. Lai, B.; Gao, W.; Cui, K.; Xie, W.; Tang, Q.; Jin, W.; Hu, G.; Ni, B.; Zhao, K. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 2018, 562, 281–285.

49. Pott, S. Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. eLife 2017, 6, e23203.

50. Clark, S.J.; Argelaguet, R.; Kapourani, C.-A.; Stubbs, T.M.; Lee, H.J.; Alda-Catalinas, C.; Krueger, F.; Sanguinetti, G.; Kelsey, G.; Marioni, J.C.; et al. ScNMT-Seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nat. Commun. 2018, 9, 781.

51. Buenrostro, J.D.; Giresi, P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 2013, 10, 1213–1218.

52. Corces, M.R.; Trevino, A.E.; Hamilton, E.G.; Greenside, P.G.; Sinnott-Armstrong, N.A.; Vesuna, S.; Satpathy, A.T.; Rubin, A.J.; Montine, K.S.; Wu, B.; et al. An improved ATAC-Seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods 2017, 14, 959–962.

53. Bryois, J.; Garrett, M.E.; Song, L.; Safi, A.; Giusti-Rodriguez, P.; Johnson, G.D.; Shieh, A.W.; Buil, A.; Fullard, J.F.; Roussos, P.; et al. Evaluation of chromatin accessibility in prefrontal cortex of individuals with schizophrenia. Nat. Commun. 2018, 9, 3121.

54. Fullard, J.F.; Hauberg, M.E.; Bendl, J.; Egervari, G.; Cirnaru, M.-D.; Reach, S.M.; Motl, J.; Ehrlich, M.E.; Hurd, Y.L.; Roussos, P. An atlas of chromatin accessibility in the adult human brain. Genome Res. 2018, 28, 1243–1252.

55. Trevino, A.E.; Sinnott-Armstrong, N.; Andersen, J.; Yoon, S.-J.; Huber, N.; Pritchard, J.K.; Chang, H.Y.; Greenleaf, W.J.; Pașca, S.P. Chromatin accessibility dynamics in a model of human forebrain development. Science 2020, 367, eaay1645.

56. Hauberg, M.E.; Creus-Muncunill, J.; Bendl, J.; Kozlenkov, A.; Zeng, B.; Corwin, C.; Chowdhury, S.; Kranz, H.; Hurd, Y.L.; Wegner, M.; et al. Common schizophrenia risk variants are enriched in open chromatin regions of human glutamatergic neurons. Nat. Commun. 2020, 11, 5581.
57. Buenrostro, J.D.; Wu, B.; Litzenburger, U.M.; Ruff, D.; Gonzales, M.L.; Snyder, M.P.; Chang, H.Y.; Greenleaf, W.J. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 2015, 523, 486–490.
58. Ziffra, R.S.; Kim, C.N.; Wilfert, A.; Turner, T.N.; Haeussler, M.; Casella, A.M.; Przytycki, P.F.; Kreimer, A.; Pollard, K.S.; Ament, S.A.; et al. Single cell epigenomic atlas of the developing human brain and organoids. Dev. Biol. 2019. preprint.
59. Cusanovich, D.A.; Daza, R.; Adey, A.; Pliner, H.A.; Christiansen, L.; Gunderson, K.L.; Steemers, F.J.; Trapnell, C.; Shendure, J. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 2015, 348, 910–914.
60. Li, Y.E.; Preissl, S.; Hou, X.; Zhang, Z.; Zhang, K.; Fang, R.; Qiu, Y.; Poirion, O.; Li, B.; Liu, H.; et al. An atlas of gene regulatory elements in adult mouse cerebrum. Neuroscience 2020. preprint.
61. Corces, M.R.; Shcherbina, A.; Kundu, S.; Gloudemans, M.J.; Frésard, L.; Granja, J.M.; Louie, B.H.; Eulalio, T.; Shams, S.; Bagdatli, S.T.; et al. Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer’s and Parkinson’s diseases. Nat. Genet. 2020, 52, 1158–1168.
62. Mulqueen, R.M.; DeRosa, B.A.; Thornton, C.A.; Sayar, Z.; Torkenczy, K.A.; Fields, A.J.; Wright, K.M.; Nan, X.; Ramji, R.; Steemers, F.J.; et al. Improved single-cell ATAC-Seq reveals chromatin dynamics of in vitro corticogenesis. Genomics 2019. preprint.
63. Lake, B.B.; Chen, S.; Sos, B.C.; Fan, J.; Kaeser, G.E.; Yung, Y.C.; Duong, T.E.; Gao, D.; Chun, J.; Kharchenko, P.V.; et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat. Biotechnol. 2018, 36, 70–80.
64. Chen, S.; Lake, B.B.; Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. Nat. Biotechnol. 2019, 37, 1452–1457.