Supplemental Data

Supplemental Data Title Page

Druggable transcriptional networks in the human neurogenic epigenome:

Transcription factors and chromatin remodeling complexes reveal novel drug-disease pathways

Gerald A. Higgins, Aaron M. Williams, Alex S. Ade, Hasan B. Alam and Brian D. Athey

Department of Computational Medicine and Bioinformatics (G.A.H., A.S.A., B.D.A.)

Department of Psychiatry (B.D.A.)

Department of Surgery (A.M.W., H.B.A.)

University of Michigan Medical School, Ann Arbor, MI 48109

Supplemental Data Abstract

This Supplemental Data section provides additional information to support the concepts and conclusions presented in “Druggable transcriptional networks in the human neurogenic epigenome: Transcription factors and chromatin remodeling complexes reveal novel drug-disease pathways”. **Supplement Data Text** and **Supplemental Data Figure 1** are presented to further illustrate the 3D spatial genome concept, then supplemented with illustrative examples further relating this concept to pharmacogenomics and pharmacokinetics. In addition, super enhancers are more thoroughly described which assists in the understanding of **Figure 2**. The bioinformatics methods used in this review and additional results are presented in **Methods and Results, Figure 5**. **Supplemental Data Table 1: Gene Names** shows the ones for **Figure 4** in the main text and **Supplemental Data Table 2: Gene Names** shows those for **Figure 5** of the main text. **Supplemental Data Table 3, 3.1-3.4** show additional results for **Figure 4,** and **Supplemental Data Table 4, 4.1 - 4.4** show additional results for **Figure 5**. **Supplemental Data Figure 1** shows the hierarchy of the spatial genome as determined by the Hi-C method and referred to in the text. **Supplemental Data Figure 2**, which supports the conclusions, presented in **Figures 4 and 5**, and the related discussion in the review. **Supplemental Data Figure 2** shows that the GO-enriched terms for the 275-gene network containing the 89 transcription factors, nuclear receptors and other molecules and the upstream xenobiotic regulators of this larger network as shown in **Figure 5**. This is derived from a consensus of results from the PsychENCODE Consortium publications.

SUPPLEMENT DATA TOC

Supplemental Data Text

A. Chromosome territories (CTs)

Chromosomes are contained within the nucleoplasm of differentiated cells as large, rope- like coils of genomic DNA and histone proteins encapsulated in chromatin. Chromosome territories (CTs) exist in 3D space where spatial proximity and chromatin state determine regulatory interactions, not distance as measured in linear DNA sequence. Although CTs do not overlap to any great degree, there are multiple spatial interactions between different CTs that are functional. These include complex transcriptional hubs, also called regulatory communities (Dai et al., 2016), that consist of multiple genes, regulatory elements including enhancers and promoters, and DNA-binding proteins such as transcription factors (Dai et al., 2016; Tjong et al., 2016). Other inter-chromosomal spatial contacts may only involve promoter-promoter pairs or enhancer-promoter interactions. Depending on the tissue microenvironment, developmental state, differentiation, lineage commitment, and pathology, inter-chromosomal spatial interactions between CTs in 3D space are altered leading to targeted changes in gene expression. For example, when epithelial cells become terminally differentiated and are integrated into epithelium, nuclear shape becomes more rigid, leading to re-orientation of CTs and activation of specific transcriptional programs driving lineage commitment (Le at al., 2016). Although researchers have developed 3D atlases that they believe represent canonical inter-chromosomal structures (Dai et al., 2016; Tjong et al., 2016), all transcriptional networks exhibit properties such as adaptation and rewiring (Hu et al., 2016). Thus, any methods that capture inter-chromosomal 3D structure at a single time point should not be considered definitive in any cell type, given the dynamic nature of chromosome interactions (Dekker et al, 2017; Reid and Rajapakse, 2017).

B. A and B compartments in chromatin

In general, compartment A contains euchromatin and more active gene transcription and compartment B corresponds to heterochromatin and is gene poor. Compartment B subsumes lamina-associated domains (LADs) located at the periphery of the nucleus. These are specific to chromosome territories, and appear to be a largely invariant feature of chromatin organization, as they are not disrupted when the organization of topologically associated domains (TADs) or lamina-associated domains (LADs) are destroyed using genome-editing methods (see below).

C. Topologically associating domains (TADs)

Within the cell's nucleus, recent studies detailing the hierarchy of transcriptional components have reinforced the adage that structure equals function. The fundamental spatial domain of transcription is the TAD located within chromosome territories (CTs), spanning approximately 1 MB in linear sequence (ranging from 0.2 MB to > 4 MB in length), and containing a variable number of genes, pseudogenes, long non-coding RNAs (lncRNAs), and DNA-dependent RNA polymerase (POLR2A). TADs are conserved as 3D structures among cell types and tissues (Lieberman-Aiden et al., 2009; Dixon et al., 2012; Dixon et al., 2016), but differ in which genes are activated in a cell type- specific and developmental manner (Dai et al., 2016; Dixon et al., 2016). For example, in humans, 60-70% of TAD structures are conserved between embryonic stem cells (ESCs) and adult, differentiated cells (Dixon et al., 2016; Ji et al, 2016). In general, TADs exhibit specific histone modifications (Ji et al., 2016; Dixon et al., 2015), are units of DNA replication timing (Pope et al., 2015) and specific TADs comprise hormone-responsive co-regulation modules (Le Dily et al., 2014). Inter-TAD functional interactions consisting of enhancer-promoter and promoter-promoter pairs are more frequent than are inter-TAD interactions (Cremer and Cremer, 2010; Dixon et al., 2015), and regulation in *cis* by enhancers constrained to promoters within a TAD seems to be a common theme (Gonzalez-Sandoval and Gasser, 2016;

Bonev and Cavalli, 2016). Transcriptionally active TADs are usually found in the interior of the nucleus and at the surface of chromosome territories (Wang et al, 2016). These TADs are highly enriched in inter-chromosomal contacts with a variety of different active promoters and enhancers on other chromosomes (Dai et al., 2016; Tjong et al., 2016; Fortin and Hansen, 2015). These trans-interacting TADs maintain strong intra-domain structure and show higher than expected frequencies of contacting individual TADs on the neighboring chromosome, suggesting that TADs interact with other chromosomes in *trans* as structurally intact units.

Wang et al (2016) used fluorescence *in situ* hybridization (FISH) to map the spatial location of a subset of TADs, and found their localization consistent with A and B compartmentalization of chromatin. They provided the locations of a subset of TADs on human chromosomes 20, 21, 22, and X in the IMR90 cell line. More specifically, limited measurements taken to date of TAD size in the human diploid fetal lung cell line IMR90 showed a range from 928,950 bases (chromosome 21) to 1,996,000 bases (chromosome X). In addition to the presence of 2 of each TAD in the diploid human genome, TADs can be assigned to either A and B compartments, which appear to be spatially organized into these compartments in a polarized, side-by-side fashion in individual chromosomes.

One distinguishing characteristic of TADs is that have discrete boundaries which provide a barrier to inter-TAD spatial contacts, first observed using light microcopy and Hi-C in human cell lines (Lieberman-Aiden et al., 2009; Dixon et al., 2012). Available evidence shows that TADs are bounded by architectural proteins such as CCTCCC-binding factor (CTCF), of which only approximately 15% CTCF binding site genomewide in humans are found at TAD boundaries, but CTCF is present at most TAD boundaries (Rao et al., 2014). Another common structural protein found at TAD boundaries is cohesin, including its commonly mapped submits of RAD21 and SMC proteins. Although it has been proposed that CTCF recruits cohesin to TAD boundaries, knock-out studies show that removal of CTCF at TAD boundaries increases inter-TAD spatial contact frequency, but intra-TAD contact numbers are reduced (Sexton and

Cavalli, 2015). TAD boundaries often contain active transcription marks such as H3K4me3 and H3K36me3, nascent transcripts, housekeeping genes (present in ~34% of TAD boundaries), and repeat elements such as LINE elements (Sexton and Cavalli, 2015; Zuin et al., 2014). This suggests a general property of TADs is that CTCF acts as an insulator at TAD boundaries, although this DNA-binding protein binds in many other regions in the human genome. In contrast, depletion of cohesin subunits has more generic effects, but TAD architecture is preserved (Dixon et al., 2015). The boundaries of TADs are enriched with other molecules including TFIIIC (Transcription Factor IIIC Subunit 1), a transcription factor involved in cellular differentiation, PDRM (PR/SET Domain 5) (Van Bortle et al., 2014), and during the cell cycle, a consensus subset of transcription factors involved in chromatin organization during cell development are highly enriched at TAD boundaries that bind to the promoters of genes with maternal or paternal allele-biased expression during human development (Chen et al., 2017).

D. Lamina associating domains (LADs)

LADs are concentrated at the periphery of the nucleus, contain heterochromatin characterized by the presence of the repressive histone modifications H3K9me2 and H3K9me3, and are genomic regions that typically are engaged in spatial contact with the nuclear lamina. Human LADs are massive, varying from 10 KB -10 MB in length (~0.5 MB median size), and containing up to 2,500 nucleosomes, and are distributed over all chromosome territories (van Steensel and Belmont, 2017). LADs are gene-poor domains, that are associated with the heterochromatin-enriched, repressive B compartment, contain LINE retro-elements, are A/T-rich, and may be located within peri-nucleolar heterochromatin as well as the nuclear periphery. DNA adenine methyltransferase Identification (DamID) analysis in proliferating human fibroblasts has shown that ~40% of the genome is engaged in the formation of LADs.

One of the primary effects of histone deacetylase (HDAC) inhibitors including valproic acid trichostatin A, is to cause chromatin decompaction (Felisbino et al., 2014) and acetylation of H3K9 (Herzoni et al.,

2011), leading to disruption of LAD structure and altering the spatial connections of the genes located within the LADs (van Steensel and Belmont, 2017). Although LADs are globally similar between cell types, key genes related to cell identity (e.g., developmental genes in embryonic stem cells (ESCs)) within LADs become 'loose' as nuclear volume changes, causing activation of genes that program cell fate (Higgins et al., 2017b; Stephens et al., 2017).

Knockdown of YY1 or lamin A/C, but not lamin A, leads to a loss of lamina association, suggesting that lamin C targets genes for silencing and chaperones them to LADs (Harr et al., 2015). This appears to be mediated by the LMNB receptor with lamin C acting as a chaperone for genes targeted for silencing to the nuclear periphery containing LADs. The organization and maintenance of LADs is dependent on YYI (YY1 Transcription Factor), a core component of the chromatin remodeling complex INO80 (Inositol requiring 80), a DNA-binding protein often found at gene promoters. INO80 is a poorly characterized ATP-dependent chromatin remolding complex in humans that contains many subunits, including YY1 and action, although it may serve to reposition nucleosomes and may regulate >1000 genes in the human genome (Yao et al., 2016), it is not been considered a target for epigenetic drug discovery. The *Xist* lncRNA controls X chromosome inactivation, causing remodeling of the 3D structure of the chromosome, and silences gene expression on the inactivated X chromosome by translocation of chromosome territory X to LADs at the periphery of the nucleus (Chen et al., 2016).

One set of genes located within the same TAD in humans includes members of the *CYP3A* family, which metabolize about 30% of medications and other xenobiotics. The *CYP3A* gene cluster, including *CYP3A4*, which encodes the most active pharmacokinetic enzyme, is contained within a TAD that is actively expressed in some tissues, such as human liver and brain, but not others, such as fibroblasts. Zullo at al (2012) found that the *CYP3A* loci contain discrete DNA regions that associate chromatin with the nuclear lamina and repress gene activity in a human fibroblast cell line (NIH 3T3), but not in a human liver cell line (FL83b) where the *CYP3A* genes where localized more centrally within the cell. They

also observed that LMNB1 levels were increased in proximity to *CYPA3* in the fibroblast cell line, concomitant with *CYPA3* gene silencing at the nuclear periphery, but not in the liver cell line, consistent with the tissue-specific expression of this set of related pharmacokinetic genes. Thus, adult human cells silence ectopic expression of pharmacokinetic genes through sequestration of the gene to a LAD. For a visual summary of these concepts see **SD Figure 1**.

E. Super-enhancers and stretch enhancers

The original definition of a super-enhancer was obtained from measurements in mouse embryonic stem cells (mESCs). Enhancers within 12.5 KB of each other were stitched together to form contiguous large enhancers which bound an order of magnitude more mediator complex (MED1) than average enhancers, and harbored hotspots of master transcription factors, which control the expression of a hierarchy of transcription factors that are involved in cell fate determination during development (e.g., MYOD1, NEUROD1, NR1D1, OCT4, SOX2), and were associated with large amounts of the histone modification H3K27ac. Thus, super-enhancers were enriched for sequence motifs corresponding to cell type–specific master transcription factors relative to normal enhancers. Using these criteria, super-enhancers bound 28-fold more MED1 than enhancers, and had a median length of 8.7 KB versus 0.703 KB for the median size of an average enhancer in a mESC cell.

Super-enhancers are cell type-specific and are involved in controlling the fate of the cells. They are also associated with frequently interacting regulatory elements (FIREs) (Schmitt et al., 2016). Super-enhancers are a subset of highly occupied transcription (HOT) domains in the human genome (Li et al., 2016). Genomewide surveys of super-enhancers in the human and mouse showed not only close correspond to master transcription factors, but also demonstrated that they exhibit significant association with genes which encode ATP binding cassette genes, calcium voltage-gated channels, potassium voltage-gated channels, glutamate receptors and solute carriers, all of which are current or potential drug

targets. Stretch enhancers may be different than super-enhancers and were first identified using ChIP-Seq results that detected several histone modifications indicative of active regulatory elements, and which spanned unexpectedly large (>3-kb) genomic regions (Parker et al., 2013). The number of stretch enhancers exceeds that of super-enhancers by an order of magnitude. However, like super-enhancers, they are cell type- specific and are associated with distinct tissue-specific master transcription factors such as NEUROD1 (human brain and pancreas).

Previous research from our laboratory has identified the central regulatory network in human brain that mediates lithium response (Higgins et al., 2015b). A detailed examination of the CNS lithium pathway (Higgins et al., 2015b) revealed that SNPs in the gene members of this pathway explained many of the adverse drug events associated with lithium therapy. Since a part of the SNP imputation and bioinformatics method (Higgins et al., 2015a) was to include 2 SNPs associated with the side effects of lithium treatment, it was not surprising that a few adverse event-linked gene variants were found in this glutamatergic pathway (Higgins et al., 2017b). However, it was remarkable that many more promoters and enhancers of the genes in this pathway have variants that have been previously associated with adverse drug events in the published literature. These pathway members contain SNPs that have been associated with lithium's side effects, but in many cases the spatial contacts are inter-chromosomal and exert their physiological consequences in other tissues such as epidermis and the cardiovascular system, not in human brain. In addition, comorbid disorders often found in conjunction with bipolar disorder are the same as, or related to, disease risk or lithium response SNPs (Higgins et al., 2017a).

Methods and supplemental results, Figure 5

The genes examined using our pharmaco-informatics pipeline (PIP) network analysis had been reported as being involved in neurogenesis in normal and pathological neurons and in developing and adult human brain tissue, and were contained in TAD loops and pathways, and gene sets were extracted from the main text, figures and Supplementary data from 12 recently published articles from the PsychENCODE Consortium (An et al., 2018; Chen et al., 2018; Gandal et al., 2018; Malysheva et al., 2019; Meng et al., 2018; Rajarajan et al., 2018; Rhie et al., 2018; de la Torre-Ubieta et al., 2018; Wang et al., 2018; Yoon et al., 2018; Zhu et al., 2018). The total number of genes reported was 425, and the union of common genes shared between the results of these publications comprised 275 genes. Of these, a subset of 92 genes was selected that encode transcription factors, nuclear receptors and chromatin remodelers based on gene network interconnectivity (p=1x10E-130, Fisher's exact test) and spatial network analysis (Higgins et al., 2019).

The criteria for gene set selection for the 89 members of the neurogenic transcription factor network and the larger interconnected network of 275 genes in which it was subsumed were: (1) 3 different software programs had to determine that these genes were significantly and directly interconnected in human brain (Higgins et al., 2019; Kanehisa et al., 2016; Krämer et al., 2013); and (2) Each gene had to be significantly associated with human neurogenesis (Bonferroni corrected) as determined by the biological process ontology label neurogenesis in *Homo sapiens* determined by Gene Ontology and correlated molecular function (Gene Ontology Consortium, 2018). The GO-enriched terms for this larger 275-gene network containing the 89 transcription factors, nuclear receptors and other molecules and the upstream xenobiotic regulators of this larger network are shown in **Figure 5**. This is derived from a consensus of results from the PsychENCODE Consortium publications and is shown in **6. Supplemental Data Table 1- 4: Results, Figure 5**

Supplemental Data Table 1: Gene Names (Figure 4)

Gene Symbol - human	Entrez Gene Name	Location	Type
ACTB	Actin beta	Cytoplasm	Other
ACTL6A	Actin like 6A	Nucleus	Other
ACTL6B	Actin like 6B	Nucleus	Other
ARID1A	AT-rich interaction domain 1A	Nucleus	Transcription factor
ARID1B	AT-rich interaction domain 1B	Nucleus	Transcription factor
ARID2	AT-rich interaction domain 2	Nucleus	Transcription factor
BCL11A	B cell CLL/lymphoma 11A	Nucleus	Transcription factor
BCL11B	B cell CLL/lymphoma 11B	Nucleus	Transcription factor
BCL7A	BCL tumor suppressor 7A	Nucleus	Transcription factor
BCL7B	BCL tumor suppressor 7B	Nucleus	Transcription factor
BCL7C	BCL tumor suppressor 7C	Nucleus	Transcription factor
DPF1	Double PHD fingers 1	Nucleus	Transcription factor
DPF2	Double PHD fingers 2	Nucleus	Transcription factor
DPF3	Double PHD fingers 3	Nucleus	Transcription factor
PBRM1	Polybromo 1	Nucleus	Transcription factor
PHF10	PHD finger protein 10	Nucleus	Transcription factor
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	Nucleus	Transcription factor
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Nucleus	Transcription factor
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	Nucleus	Transcription factor
SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin subfamily c member 1	Nucleus	Transcription factor
SMARCC2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin subfamily c member 2	Nucleus	Transcription factor
SMARCD1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1	Nucleus	Transcription factor
SMARCD2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 2	Nucleus	Transcription factor
SMARCE1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	Nucleus	Transcription factor
SOX10	SRY-box 10	Nucleus	Transcription factor
SS18	SS18, nBAF chromatin remodeling complex subunit	Nucleus	Transcription factor

Supplemental Data Table 2: Gene Names (Figure 5)

Gene Symbol, human	Entrez Gene Name	Location	Type
ACTL6A	Actin like 6A	Nucleus	Other
ARID1A	AT-rich interaction domain 1A	Nucleus	Transcription factor
ARNT	Aryl hydrocarbon receptor nuclear translocator	Nucleus	Transcription factor
ARNT2	Aryl hydrocarbon receptor nuclear translocator 2	Nucleus	Transcription factor
ARX	Aristaless related homeobox	Nucleus	Transcription factor
ASCL1	Achaete-scute family bHLH transcription factor 1	Nucleus	Transcription factor
BCL11B	B cell CLL/lymphoma 11B	Nucleus	Transcription factor
BCL6	B cell CLL/lymphoma 6	Nucleus	Transcription factor
BRDT	Bromodomain testis associated	Nucleus	Kinase
CUX1	Cut like homeobox 1	Nucleus	Transcription factor
CUX2	Cut like homeobox 2	Nucleus	Transcription factor
EMX1	Empty spiracles homeobox 1	Nucleus	Transcription factor
EOMES	Eomesodermin	Nucleus	Transcription factor
EP300	E1A binding protein p300	Nucleus	Transcription factor
ETV6	ETS variant 6	Nucleus	Transcription factor
FOXA1	Forkhead box A1	Nucleus	Transcription factor
FOXA2	Forkhead box A2	Nucleus	Transcription factor
FOXC2	Forkhead box C2	Nucleus	Transcription factor
FOXH1	Forkhead box H1	Nucleus	Transcription factor
FOXO3	Forkhead box O3	Nucleus	Transcription factor
FOXP2	Forkhead box P2	Nucleus	Transcription factor
GBX2	Gastrulation brain homeobox 2	Nucleus	Transcription factor
GLI2	GLI family zinc finger 2	Nucleus	Transcription factor
GLI3	GLI family zinc finger 3	Nucleus	Transcription factor
HDAC4	Histone deacetylase 4	Nucleus	Transcription factor
HDAC5	Histone deacetylase 5	Nucleus	Transcription factor
HDAC9	Histone deacetylase 9	Nucleus	Transcription factor
HES1	Hes family bHLH transcription factor 1	Nucleus	Transcription factor
HEY2	Hes related family bHLH transcription factor with YRPW motif 2	Nucleus	Transcription factor
HOXA2	Homeobox A2	Nucleus	Transcription factor
ID4	Inhibitor of DNA binding 4, HLH protein	Nucleus	Transcription factor
KLF16	Kruppel like factor 16	Nucleus	Transcription factor
KLF4	Kruppel like factor 4	Nucleus	Transcription factor
LHX2	LIM homeobox 2	Nucleus	Transcription factor
MAFB	MAF bZIP Transcription factor B	Nucleus	Transcription factor
MEF2C	Myocyte enhancer factor 2C	Nucleus	Transcription factor
MEF2D	Myocyte enhancer factor 2D	Nucleus	Transcription factor
MEIS1	Meis homeobox 1	Nucleus	Transcription factor
MYBL1	MYB proto-oncogene like 1	Nucleus	Transcription factor
MYCN	MYCN proto-oncogene, bHLH transcription factor	Nucleus	Transcription factor
NEUROD1	Neuronal differentiation 1	Nucleus	Transcription factor
NEUROD4	Neuronal differentiation 4	Nucleus	Transcription factor
NEUROG2	Neurogenin 2	Nucleus	Transcription factor

NFE2L2	Nuclear factor, erythroid 2 like 2	Nucleus	Transcription factor
NFIB	Nuclear factor I B	Nucleus	Transcription factor
NFIX	Nuclear factor I X	Nucleus	Transcription factor
NFYA	Nuclear Transcription factor Y subunit alpha	Nucleus	Transcription factor
NFYB	Nuclear Transcription factor Y subunit beta	Nucleus	Transcription factor
NKX6-1	NK6 homeobox 1	Nucleus	Transcription factor
NR0B1	Nuclear receptor subfamily 0 group B member 1	Nucleus	Nuclear receptor
NR1D1	Nuclear receptor subfamily 1 group D member 1	Nucleus	Nuclear receptor
NR2E1	Nuclear receptor subfamily 2 group E member 1	Nucleus	Nuclear receptor
NR2F1	Nuclear receptor subfamily 2 group F member 1	Nucleus	Nuclear receptor
NR2F2	Nuclear receptor subfamily 2 group F member 2	Nucleus	Nuclear receptor
NR4A3	Nuclear receptor subfamily 4 group A member 3	Nucleus	Nuclear receptor
ONECUT2	One cut homeobox 2	Nucleus	Transcription factor
PAX6	Paired box 6	Nucleus	Transcription factor
PBRM1	Polybromo 1	Nucleus	Other
PBX3	PBX homeobox 3	Nucleus	Transcription factor
POU3F2	POU class 3 homeobox 2	Nucleus	Transcription factor
POU3F3	POU class 3 homeobox 3	Nucleus	Transcription factor
POU5F1	POU class 5 homeobox 1	Nucleus	Transcription factor
PROX1	Prospero homeobox 1	Nucleus	Transcription factor
RARA	Retinoic acid receptor alpha	Nucleus	Nuclear receptor
RORA	RAR related orphan receptor A	Nucleus	Nuclear receptor
RXRA	Retinoid X receptor alpha	Nucleus	Nuclear receptor
SATB2	SATB homeobox 2	Nucleus	Transcription factor
SMAD4	SMAD family member 4	Nucleus	Transcription factor
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Nucleus	Transcription factor
SMARCE1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	Nucleus	Transcription factor
SOX10	SRY-box 10	Nucleus	Transcription factor
SOX11	SRY-box 11	Nucleus	Transcription factor
SOX2	SRY-box 2	Nucleus	Transcription factor
SOX3	SRY-box 3	Nucleus	Transcription factor
SOX4	SRY-box 4	Nucleus	Transcription factor
SOX5	SRY-box 5	Nucleus	Transcription factor
SOX6	SRY-box 6	Nucleus	Transcription factor
SOX9	SRY-box 9	Nucleus	Transcription factor
SP1	Sp1 transcription factor	Nucleus	Transcription factor
SP2	Sp2 transcription factor	Nucleus	Transcription factor
ST18	ST18, C2H2C-type zinc finger	Nucleus	Transcription factor
TBR1	T-box, brain 1	Nucleus	Transcription factor
TCF12	Transcription factor 12	Nucleus	Transcription factor
TCF3	Transcription factor 3	Nucleus	Transcription factor
TCF4	Transcription factor 4	Nucleus	Transcription factor
WNT5A	Wnt family member 5A	Nucleus	Other
ZEB1	Zinc finger E-box binding homeobox 1	Nucleus	Transcription factor
ZHX2	Zinc fingers and homeoboxes 2	Nucleus	Transcription factor
ZNF24	Zinc finger protein 24	Nucleus	Transcription factor

Supplemental Data Table 3, 3.1-3.4: Additional results, Figure 4

3.1 Biological process, Gene Ontology (Gene Ontology Consortium, 2016)

GO Biological Processes	p-value
Nucleosome disassembly	1.98E-19
Chromatin disassembly	1.98E-19
Chromatin remodeling	3.80E-19
Chromatin organization	3.92E-19
ATP-dependent chromatin remodeling	1.35E-16
Neurogenesis	7.21E-16
Positive regulation of nucleic acid-templated transcription	8.08E-14
Chromatin assembly or disassembly	2.48E-11
Nucleosome organization	2.48E-11
Nervous system development	1.73E-09

3.2 Molecular functions, Gene Ontology (Gene Ontology Consortium, 2016)

GO Molecular Function	p-value
Transcription coactivator activity	1.17E-15
Nucleosome binding	8.68E-13
RNA polymerase II distal enhancer sequence-specific DNA binding	1.43E-11
Chromatin binding	4.52E-11
Enhancer sequence-specific DNA binding	5.96E-11

3.3 Diseases / Functions annotation from IPA™ (Krämer et al, 2013)

Diseases or Functions Annotation	p-Value
Coffin-Siris syndrome	1.98E-29
Remodeling of chromatin	8.06E-23
Disassembly of nucleosomes	8.52E-23
Repair of DNA lesion	6.58E-17
Non-homologous end joining of DNA double strand break site	1.22E-16
Autosomal dominant mental retardation	8.33E-15
Congenital malformation of brain	2.64E-10
Brain oligodendroglioma	8.57E-07
Autosomal dominant mental retardation type 12	1.34E-06
Nicolaides-Baraitser syndrome	1.34E-06

3.4 Top upstream xenobiotic controller from IPA™ (Krämer et al, 2013) and KEGG (Kanehisa et al, 2016)

Upstream xenobiotic regulator	Mechanism / type	p-value of overlap
(S)-duloxetine	Antidepressant	3.59E-03
Sertindole	Antipsychotic	4.79E-03
Pomalidomide	Anti-angiogenic, anti-proliferative and Immunomodulatory effects	1.43E-02
Fluvoxamine	Antidepressant	3.65E-02
Trichostatin A	HDAC inhibitor	3.68E-02

Supplemental Data Table 4, 4.1- 4.4: Additional Results, Figure 5

4.1 Biological process of 275-gene network, Gene Ontology (Gene Ontology Consortium, 2016)

GO Biological Processes	p-value
Neurogenesis	1.66E-208
Generation of neurons	9.34E-199
Nervous system development	1.18E-186
Neuron differentiation	6.08E-146
Neuron development	6.05E-110
Regulation of nervous system development	2.67E-92
Regulation of neurogenesis	9.93E-90
Neuron projection development	4.80E-87
Regulation of neuron differentiation	4.08E-78
Neuron projection morphogenesis	1.09E-74

4.2 Molecular functions of 275-gene network, Gene Ontology (Gene Ontology Consortium, 2016)

GO Molecular Function	p-value
Sequence-specific DNA binding	8.20E-37
Transcription regulatory region DNA binding	2.90E-36
Regulatory region nucleic acid binding	3.34E-36
Transcription regulatory region sequence-specific DNA binding	1.90E-35
RNA polymerase II regulatory region sequence-specific DNA binding	1.93E-34

4.3 Functions annotation of 275-gene network from IPA™ (Krämer et al, 2013)

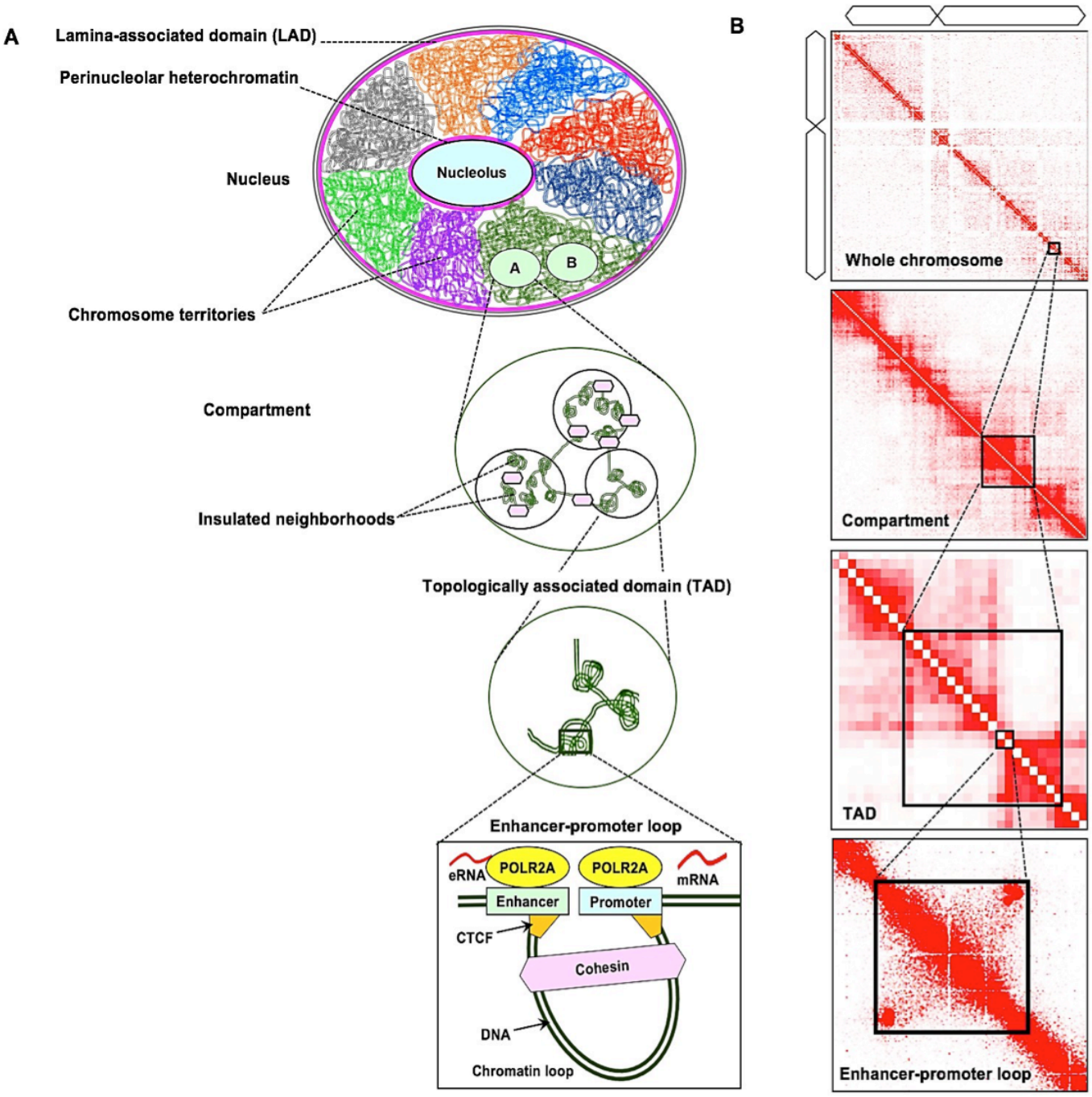
Diseases or Functions Annotation	p-Value
Development of neurons	6.03E-86
Development of central nervous system	3.98E-75
Development of head	3.58E-70
Development of body axis	8.42E-70
Morphology of nervous system	9.50E-65
Differentiation of nervous system	1.25E-62
Guidance of axons	3.64E-62
Morphogenesis of neurons	9.40E-61
Morphology of head	1.13E-60
Neuritogenesis	4.15E-60

4.4 Top upstream xenobiotic controllers of 275-gene network from IPA™ (Krämer et al, 2013) and KEGG (Kanehisa et al, 2016)

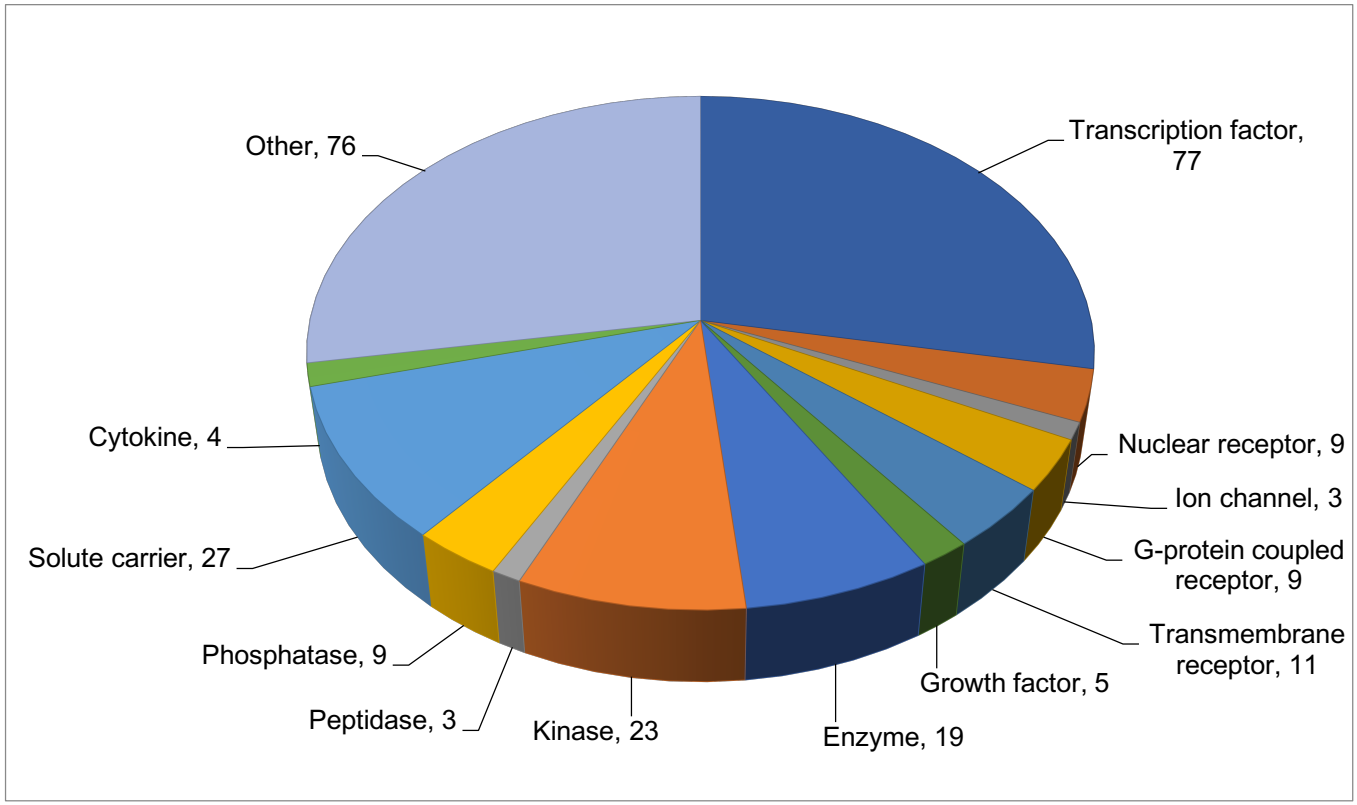
Upstream xenobiotic regulator	Mechanism / type	p-value of overlap
Bexarotene	Antineoplastic, possible antipsychotic adjuvant therapeutic	1.78E-07
Valproic acid	Anti-epileptic, mood stabilizer	1.49E-06
Ethosuximide	Anti-epileptic, absence seizures	3.12E-06
Lithium chloride	Anti-manic, mood stabilizer	1.47E-05
Napabucasin	Anti-stem cell	1.60E-05

Supplemental Data: Figures

Supplemental Data Figure 1



Supplemental Data Figure 2



Supplemental Data Figure Captions

Supplemental Data Figure 1. The spatial hierarchy of transcription in the human genome. (A) Visualization of nuclear and sub- nuclear transcriptional topology, including regions of heterochromatin located at the periphery of the nucleus (lamina-associated domains (LADs)) and surrounding the nucleolus, chromosomes arranged as space-filling chromosome territories (CTs) in the cytoplasm, A (euchromatin) and B (heterochromatin) compartments within each CT, insulated neighborhoods bounded by cohesin within a compartment, topologically-associated domains (TADs) and enhancer-promoter loops. **(B)** Hi-C mapping shows multiscale hierarchy of transcriptional regulation of biological structures shown in panel **(A)**. Modified from Rao, et al. 2014.

Supplemental Data Figure 2. Distribution of protein classes encoded by the 275 genes in the human neurogenic transcriptional network expanded from **Figure 5** in the text.

Supplemental Data References

An JY, Lin K, Zhu L, Werling DM, Dong S, Brand H, Wang HZ, Zhao X, Schwartz GB, Collins RL, et al (2018) Genomewide de novo risk score implicates promoter variation in autism spectrum disorder. *Science* DOI: 10.1126/scitranslmed.aat8178.

Bonev B and Cavalli G (2016) Organization and function of the 3D genome. *Nat Rev Genet* **17**: 772–772.

Chen C, Meng Q, Xia Y, Ding C, Wang L, Dai R, Cheng L, Gunaratne P, Gibbs RA, Min S, et al (2018) The transcription factor POUF32 regulates a gene co-expression network in brain tissue from patients with psychiatric disorders. *Sci Transl Med* **10**: *pii:e*aat8178.

Chen H, Liu S, Seaman L, Najarian C, Wu W, Ljungman M, Higgins GA, Hero A, Wicha M, and Rajapakse I (2017) Parental allele-specific genome architecture and transcription during the cell cycle. *bioRxiv* DOI: http://dx.doi.org/10.1101/201715.

Chen CK, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, Cerase A, McDonel P, and Guttman M (2016) Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**: 468–472.

Cremer T and Cremer M (2010) Chromosome territories. *Cold Spring Harb Persepect Biol* DOI: 10.1101/cshperspect.a003889.

Dai C, Li W, Tjong H, Hao S, Zhou Y, Li Q, Chen L, Zhu B, Alber F, and Jasmine Zhou X (2016) Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities. *Nat Commun* **7**: 11549.

Dekker J, Belmont AS, Guttman, Leshyk VO, Lis JT, Lomvardas S, Mirny LA, O'Shea CC, Park PJ, Ren B, et al (2017) The 4D nucleome project. *Nature* **549**: 219-226.

de la Torre-Ubieta L, Stein JL, Won H, Oplan CK, Liang D, Lu D, and Geschwind DH (2018) The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* **172**: 289-304.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, and Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380.

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al (2015) Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331-336.

Dixon JR, Gorkin DU, and Ren B (2016) Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell* **62**: 668-680.

Felisbino MB, Gatti MSV, and Mello MLS (2014) Changes in chromatin structure in NIH 3T3 cells induced by valproic acid and trichostatin A. *J Cell Biochem **115***: 1937-1947.

Fortin JP and Hansen KD (2015) Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol* **16**: 180.

Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, Won H, van Bakel H, Varghese M, Wang Y, Shieh AW, et al (2018) Transcriptome-wide isoform-level dysregulation in ADS, schizophrenia and bipolar disorder. *Science* DOI: 10.1126/science.aat8127.

Gene Ontology Consortium (2018) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acis Res* **47**: D330-D338.

Girdhar K, Hoffman GE, Jian Y, Brown L, Kundakovic M, Hauberg ME, Francoeur NJ, Wang YC, Shah H, Kvanagh DH, et al (2018) Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. *Nat Neurosci* **21**:1126-1136.

Gonzalez-Sandoval A and Gasser SM (2016) On TADs and LADs: Spatial Control Over Gene Expression. *Trends Genet* **32**: 485-495.

Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ, and Reddy KL (2015) Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A- type lamins. *J Cell Biol* **208**:33– 52.

Hezroni H, Sailaja BS, and Meshorer E (2011) Pluripotency-related, valproic acid (VPA)-induced genome-wide histone H3 lysine 9 (H3K9) acetylation patterns in embryonic stem cells. *J Biol Chem* **286**: 35977-35988.

Higgins GA, Allyn-Feuer A, and Athey BD (2015a) Epigenomic mapping and effect sizes of noncoding variants associated with psychotropic drug response. *Pharmacogenomics* **16**: 1565-1583.

Higgins GA, Allyn-Feuer A, Barbour E, and Athey BD (2015b) A glutamatergic network mediates lithium response in bipolar disorder as defined by epigenome pathway analysis. *Pharmacogenomics* **16**: 1547–1563.

Higgins GA, Georgoff P, Nikolian VC, Allyn-Feuer A, Pauls B, Higgins R, Athey BD, and Alam HB (2017b). Network reconstruction reveals that valproic acid activates neurogenic transcriptional programs in adult brain following traumatic injury. *Pharm Res* **34**: 1658-1672.

Higgins GA, Ade A, Reamaroon N, Kalinin A, and Athey BD. (2019) Methods and system to reconstruct drug spatial networks from pharmacogenomic regulatory interactions and uses. USPTO, *Submitted.*

Hu JX, Thomas CE, and Brunak S (2016) Network biology concepts in complex disease comorbidities. *Nature Rev Genet* **17**: 615–629.

Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, Shachar S, Weintraub AS, Hnisz D, Pegoraro G, Lee TI, et al (2016) 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell* **18**: 262–275.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, and Morishima K (2016) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.

Kramer A, Green J, Pollard Jr J, and Tugendreich S (2013) Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30**: 523-530.

Le HQ, Ghatak S, Yeung CYC, Tellkamp F, Günschmann C, Dieterich C, Yeroslaviz A, Habermann B, Pombo A, Niessen CM, et al (2016) Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat Cell Biol* **18**: 864–875.

Le Dily F, Baù D, Pohl A, Vicent GP, Serra F, Soronellas D, Castellano G, Wright RHG, Ballare C, Filion G, et al (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev* **28**:2151–2162.

Li H, Liu F, Ren C, Bo X, and Shu W (2016) Genome-wide identification and characterization of HOT regions in the human genome. *BMC Genomics* **17**: 733.

Li G, Fullwood MJ, Xu H, Mulawadi FH, Velkov S, Vega V, Ariyaratne PN, Mohamed YB, Oo HS, Tennakoon C, et al (2010) ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. *Genome Biol **11***: R22.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al (2009) Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**: 289–294.

Malysheva V, Mendoza-Parra MA, Blum M, and Gronemeyer H (2019) Highly dynamic chromatin interactions drive neurogenesis through gene regulatory networks. bioRxiv **DOI:** https://doi.org/10.1101/303842.

Meng Q, Wang K, Brunetti T, Xia Y, Jiao C, Dai R, Fitzgerald D, Thomas A, Jay L, Eckhart H, et al (2018) The DGCR5 long noncoding RNA may regulated the expression of several schizophrenia-related genes. *Sci Transl Med* DOI: 10.1126/scitranslmed.aat6912.

Parker SCJ, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N, NISC Comparative Sequencing Program, 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.

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, et al (2015) Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**:402–405.

Rajarajan P, Borrman T, Liao W, Schrode N, Flaherty E, Casino C, Powell S, Yashawini C, LaMarca EA, Kassim B, et al (2018) Neuron-specific signatures in the chromosomal connectome associated with schizophrenia risk. *Science* DOI:10.1126/science.aat4311

Rao SPR, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**:1665–1680.

Rhie SK, Schreiner S, Witt H, Armoskus S, Lay FD, Camarena A, Spitsyna, Guo Y, Berman BP, Evgrafov OV, et al (2018) Using 3D epigenomic maps of primary olfactory neuronal cells from living individuals to understand gene regulation. *Sci Adv* DOI: 10.1126/sciadv.aav8550.

Ried T and Rajapakse I (2017) The 4D Nucleome. *Methods* **123**:1-2.

Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Ying L, Barr CL, and Ren B (2016) A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Reports* **17**:2042-2059.

Sexton T and Cavalli G (2014) The Role of Chromosome Domains in Shaping the Functional Genome. *Cell* **160**:1049-1059.

LeStephens AD, Banigan EJ, Adam SA, Goldman, RD, and Marko JF (2017) Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. *Mol Bio Cell **28***:1984-1996.

Tjong H, Li W, Kalhor R, Dai C, Hao S, Gong K, Zhou Y, Li H, Zhou XJ, et al (2016) Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. *Proc Natl Acad Sci U S A* **113**:1663–1672.

Van Bortle K, Nichols MH, Li L, Ong CT, Takenaka N, Qin ZS, and Corces VG (2014) Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* **15**:R82.

van Steensel B and Belmont AS (2017) Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**:780-791.

Wang S, Su JH, Beliveau BJ, Bintu B, Moffitt JR, Wu C, and Zhuang X (2016) Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**:598–602.

Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, Clarke D, Gu M, Emani P, Yang YT, et al (2018) Comprehensive functional genomic resource and integrative model for the human brain. *Science* DOI: 10.1126/science.aat8464.

Yao W, King DA, Beckwith SL, Gowans GJ, Yen K, Zhou C, and Morrison AJ (2016) The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Mol Cell Biol* **36**:979–991.

Yoon KJ, Vissers C, Ming GL, and Song H (2018) Epigenetics and epitranscriptomics in temporal patterning of cortical neural progenitor competence. *J Cell Biol* **217**:1901-1914.

Zhu Y, Sousa AMM, Gao T, Skarica M, Li M, Santpere G, Esteller-Cucala P, Juan D, Ferrandez-Peral L, Gulden FO, et al (2018) Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* DOI: 10.1126/science.aat8077.

Zuin J, Dixon JR, van der Reijden MIJA, Ye Z, Kolovos P, Brouwer RWW, van de Corput MPC, van de Werken HJG, Knoch TA, van Ijcken WFJ, et al (2014) Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A* **11**:996–1001.

Zullo JM, Demarco IA, Piqué-Regi R, Gaffney DJ, Epstein CB, Spooner CJ, Luperchio TR, Bernstein BE, Pritchard JK, Reddy KL, et al (2012) DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell* **149**:1474-1487.