Journal of Biochemistry and Cell Biology

Review Article Open Access

Finding Genes for Neurodevelopmental Disorders by Mapping Genome-Wide Long-Range Interactions in Neural Cell Chromatin

Carl L*

Department of Molecular Biology, Institute of Health Biosciences, Sweden

Abstract

Variations in the DNA sequence are referred to as SNPs, or single nucleotide polymorphisms. In neurodevelopmental disorders (NDDs) and traits, copy number variants (CNVs) that map to putative transcriptional regulatory elements like enhancers are common. However, the genes that these enhancers control are still a mystery. In the past, it was thought that the gene promoter that was closest to an enhancer was affected by its activity and any possible changes caused by sequence variants. The discovery of genome-wide long-range interaction maps in the chromatin of neural cells challenges this idea because they demonstrate that enhancers frequently connect to promoters farther apart and skip genes in between. This perspective focuses on a number of recent studies that have used HiC, RNApolII ChIA-PET, Capture-HiC, or PLACseq to generate long-range interaction maps and overlap the identified long-range interacting DNA segments with DNA sequence variants associated with NDD (like schizophrenia, bipolar disorder, and autism) and traits (intelligence). Using this method, it was possible to connect the function of the enhancers that house the NDD-related sequence variants to a connected gene promoter that was away from the linear chromosome map. By identifying mutations in the gene's protein-coding regions (exons), some of these enhancer-connected genes had already been identified as contributing to the diseases, which validated the method. However, a significant number of the connected genes also contain genes whose exons had not previously been found to be mutated, suggesting novel NDD and trait causes. As a outcome, the DNA variants and long-range interaction maps uncovered by NDD can be used as "pointers" to locate novel candidate genes associated with the disease. Methods based on CRISPR-Cas9 are beginning to investigate the functional significance of the identified interactions as well as the enhancers and genes involved by functionally manipulating the long-range interaction network that includes promoters and enhancers. As a direct consequence of this, our comprehension of the pathology of neural development is improving.

Keywords: DNA Sequence Variants; Neurodevelopmental Disorders (NDD); Crispr-Cas9; Gene Regulation

Introduction

The human and mouse genomes were mapped with hundreds of thousands of potential enhancers thanks to high-throughput sequencing techniques like ChIPseq and ATACseq. Enhancers are fundamental DNA sequences that play a role in gene transcription and interact with gene promoters in a useful way to control gene transcription. Promoters and enhancers work with a variety of transcription factors, including RNApolII, to regulate the transcriptional output of a gene. Because they have transcription factor binding sites, histone modifications, and an "open" chromatin configuration, these sequences are epigenetic "enhancer marks." DNA sequence variants that were previously linked to inherited disorders by genome-wide association studies (GWAS) frequently mapped onto or were very close to these "epigenetic enhancers" at the genome-wide scale after the initial studies using genome-wide functional genomics [**1-3**]. As a outcome, the possibility that enhancer activity might be altered by closely related disease-associated sequence variants or mutations was raised. On the other hand, which genes are regulated by these enhancers and possibly deregulated by the accompanying mutations? The experimental determination of genome-wide long-range interaction maps in the human and mouse genomes challenged the conventional notion that an enhancer controls the gene promoter that is closest to it. These maps actually demonstrated that enhancers in a gene's intron can be connected to a different promoter than the one they map to and that epigenetic enhancers are physically and functionally connected to distant gene promoters, skipping genes in between [**4**].

Chromatin's long-term interactions are plotted

On the linear chromosome map, long-range interaction maps show the physical connections between distant DNA sequences. By crosslinking chromatin following DNA fragmentation, ligating the physically associated DNA fragments, and high-throughput sequencing, they are based on the "freezing" of the physical proximity. The studies discussed here map interactions using a variety of approaches. In theory, a single reaction using Hi-C can map all interactions between each pair of DNA fragments. Hi-C not only revealed large-scale chromosomal structures like A/B compartments and topologically associated domains, but it also provided a global view of 3D chromosomal topology. HiC libraries, on the other hand, are extremely challenging because of their global nature (all pairwise contacts between 106 to 108 fragments in a mammalian genome, depending on the restriction enzyme used). In turn, this necessitates an impractical sequencing depth for reliable and accurate chromosomal contact detection. Both RNApolII-Chromatin Interaction Analysis by Paired-End Tagging (ChIA-PET) and its variant, RNApolII-in situ ChIA-PET, make use of paired-end tagging. uses an anti-RNApolII antibody to initiate chromatin immunoprecipitation prior to proximity ligation, concentrating on interactions that are actually connected to gene transcription [**5-7**]. In zebrafish and mouse in vivo transgenesis assays, the majority of connected enhancers detected by this method in brain-derived neural stem cells (NSC)

***Corresponding author:** Carl L, Department of Molecular Biology, Institute of Health Biosciences, Sweden E-mail: carl.edu@26uppsala.co.in

Received: 07-Jan-2023, Manuscript No: jbcb-23-85986, **Editor assigned:** 10-Jan-2023, PreQC No: jbcb-23-85986 (PQ), **Reviewed:** 24-Jan-2023, QC No jbcb-23-85986, **Revised:** 27-Jan-2023, Manuscript No: jbcb-23-85986 (R), **Published:** 31-Jan-2023, DOI: 10.4172/jbcb.1000171

Citation: Carl L (2023) Finding Genes for Neurodevelopmental Disorders by Mapping Genome-Wide Long-Range Interactions in Neural Cell Chromatin. J Biochem Cell Biol, 6: 171.

(14/15 tested) function as brain-active enhancers. By enriching HiC material for interactions that involve, at least on one end, restriction fragments of interest ('baits' or 'viewpoints,') such as all annotated gene promoters, the original limitation of HiC is attempted to be overcome by combining HiC with targeted capture and sequencing (Capture HiC, CHiC). Antibodies against the histone modification H3K4me3, which is found on active promoters, are used in an immunoprecipitation step to enrich for interactions involving.

Variants in the DNA sequence

Frequently, functional abnormalities in the regulation of gene expression or the encoded protein are linked to DNA variants found within genes. Multifactorial or monogenic diseases that are passed down through families may be associated with these abnormalities. The discovery of hundreds of thousands of variants that map to putative enhancers, also known as non-coding DNA, has raised the possibility that these variants are involved in both the disease and the regulation of the gene that is controlled by the enhancer. To investigate this aspect, the following steps are required: i) to determine which gene is connected; ii) to find out if the enhancer actually controls (increases or decreases the gene's activity); and (iii) to compare the enhancer activity of the variant to that of the wild type **[8]**. It can be hypothesized that the variant enhancer contributes to the disease by causing abnormal regulation of the connected gene, which can then be considered a candidate disease gene, if these points are correctly demonstrated.

Both novel genes that were specifically discovered in each study and candidate genes for NDD disease that spanned multiple studies were identified in the subsequent studies by utilizing a variety of neural cell types and developmental stages. This demonstrates that the complex brain system must investigate a wide range of cell types and developmental stages in order to accurately cover the spectrum of potential mutations.

In the majority of genome-wide informatic studies of DNA sequence variants, HiC or related variants of the method, which identified a large number of epigenetic enhancers carrying DNA variants and connected with gene promoters by interactions in neural cells (here referred to as "connected enhancers"), were utilized. HiC discovered overlapped enhancer–promoter chromatin contacts in the human cerebral cortex in the middle of gestation. These contacts were linked to non-coding variants (SNPs) found in schizophrenia (SCZ) by GWAS. As a outcome, a number of potential SCZ risk genes and pathways were identified. Independent analyses of expression quantitative trait loci (eQTL) revealed a number of these genes as potential risk factors. The study demonstrated that a significant number of genes interacted with enhancers in a brain-specific manner and that the majority of HiC-identified enhancers did not interact with adjacent genes, highlighting the significance of identifying chromatin interactions that are relevant to tissue. About 500 genes that were not adjacent to index SNPs were detected by HiC-detected contacts between gene promoters and enhancers, indicating that linear chromosome organization alone is not responsible for the majority of regulatory interactions. Postsynaptic proteins, acetylcholine receptors, neuronal differentiation proteins, and chromatin remodellers are abundant in these potential SCZ genes **[9]**. Particularly notable were the transcription factors SOX2, FOXG1, EMX1, TBR1, SATB2, CUX2, and FOXP1, mapped long-range interactions by HiC in neural progenitor cells (NPCs) derived from human induced pluripotent stem cells (hIPSC), as well as in neurons and astrocyte-like glia derived from NPC differentiation (accessed on January 1, 2023) within the framework of the work of the PsychENCODE consortium. SCZ risk loci were matched with their maps. There were a total of 224 genes with the risk loci that overlapped; Long-range interactions with distant epigenetic enhancers did not overlap with the SCZ risk loci for 580 genes, including 240 in NPCs, 227 in neurons, and 113 in astroglia. The current network of known genes that overlap risk sequences was expanded by as much as 150,000 percent thanks to these genes, which are referred to as "risk-locus connect." The genes associated with synaptic plasticity, neuronal connectivity, and chromatin-associated proteins, including TFs, were enriched by the disease-related "connectome" that was discovered.

Microdissection and FACS sorting of excitatory neurons (eN), interneurons (iN), radial glia (RG), and intermediate progenitors (IPC) from mid-gestation human cortex with the appropriate antibody combinations against SOX2, PAX6, EOMES, and SATB2 were used to isolate these cells. After that, interactions from the germinal zone (GZ) and cortical plate (CP) were profiled using H3K4me3-PLAC-seq. For RG, IPC, eN, and iN, this revealed 35,552, 26,138, 29,104, and 22,598 interactions, respectively. Song et al.'s interacting promoters and enhancers were found to overlap with SNPs associated with seven brain genetic disorders: With the exception of ASD and AD, the interacting sequences significantly enhanced all traits, including AD, ADHD, ASD, BD, IQ, and schizophrenia used RNApolII-ChIA-PET to profile RNA-polII-mediated long-range interactions in the chromatin of ex vivo neural stem/progenitor cells (NSC) from the mouse neonatal forebrain in mice whose developing nervous system contained a Nestin-Cre transgene that deleted the SOX2 transcription factor-encoding gene, which is a gene whose mutation causes NDD in humans. These mice were used for this study because their It was found that SOX2 is necessary to maintain the integrity of the 3D interactome and the expression levels of some genes involved in SOX2-dependent interactions when SOX2+ and SOX2-NSC interactions were compared. The significant enrichment of SOX2-bound interactions in the enhancers and promoters of genes associated with NDD upon mutation was particularly noteworthy. 14 of the 15 tested connected enhancers were active, directing reporter gene expression toward the developing forebrain, making up the majority of the connected enhancers found in transgenic zebrafish and mice. This demonstrated that active brain in vivo functional enhancers could be found using RNApolII ChIA-PET. On the other "anchor," there were approximately 10,000 interactions between gene promoters and epigenetic enhancers in humans, but only about 7500 of those interactions were conserved **[9]**. The enhancer and promoter regions of the human genome could be remapped onto a syntenic region to accommodate these interactions; Additionally, activity-related epigenetic marks (H3K27Ac;) were found on the majority of human enhancers and promoters. H3K4me1) was linked to DNA sequence variants that are associated with NDD, both SNV and CNV, and the connected enhancers that were discovered by RNApolII ChIA-PET in NSC had a counterpart in the epigenetic enhance In interaction maps. As demonstrated by the overlap with CNV in ASD patients, microdeletions can overlap with enhancers connected to distant genes that are not involved in the deletion itself. In other instances, connective enhancers were distinguished from their target genes by longer CNVs; Even though it was not directly involved in the CNV, the gene may have been deregulated due to the absence of regulatory interactions. Additionally, numerous SNV previously associated with intelligence, bipolar disorder (BD), and schizophrenia (SCZ) overlapped with the ChIA-PET-identified connective enhancers; Some genes were linked to genes with mutations in the protein coding region that had already been identified as contributing factors to NDD, while others were linked to genes that were active in the developing nervous system and in neurons but had not been involved in the pathogenesis of NDD. Despite the fact that some of these genes had

previously been discovered in studies mapping interactions with HiC, ChIA-PET was used to discover approximately half of these genes for the first time. This suggests that a wide range of complementary mapping strategies can be used to comprehensively chart NDD-relevant chromatin interactions.

Studies on Multiple Enhancer Modulation

Functional studies of these enhancers have so far been "low-throughput," focusing on a single enhancer at a time, despite the fact that a number of enhancers with variants that may be relevant to NDD have been identified. The method of simultaneously targeting a large number of enhancers and analyzing a single cell transcriptome as a means of monitoring changes in gene expression, which was initially developed in hematopoietic cells, could be adapted to study multiplex enhancers in human neural cells as well.

5920 particular up-and-comer enhancers were adjusted in this study utilizing stable dCAS9-KRAB lentiviruses; On average, 15+/-11 guideRNA-encoding viruses were used to transduce each cell. 164 genes were downregulated (downregulated) as a outcome of the connected enhancer's perturbation, as revealed by RNA sequencing of single cells sorted by FACS. FISH, or fluorescence in situ hybridization, is a method for quantitatively labeling individual cells that makes use of the expression of the mRNA(s) under investigation. Similar to CRISPRi-FlowFISH, which uses a pooled screen to target several putative enhancers with a variety of guideRNAs before analyzing the expression of the gene or genes of interest, this is carried out in a manner. High-throughput sequencing is used to determine the frequency of guideRNAs from each bin after cells are FACS-sorted into six bins based on the level of mRNA expression **[10]**. Lastly, the various guideRNA effects on gene expression can be estimated using the relative abundance of guideRNAs in each bin. A comprehensive investigation of disease-related GWAS variants in numerous cells and tissues, including brain cells, benefited from this approach. An Activity-By-Contact (ABC) model that takes Activity into account (measurements of chromatin accessibility (ATAC-seq or DNAse-seq), H3K24Ac histone modification (by ChIPseq), and chromatin connectivity (HiC)) was used to predict functional enhancer–gene connections in the most recent study. CRISPRCas testing demonstrated that the ABC model accurately predicted (in silico) connections between functional enhancers and promoters. These approaches provide fresh insight into the possibility that non-coding variants can cause diseases by combining their strengths.

Conclusion

Through CRISPR-Cas9-mediated genome editing, the "knock-in" of the alleged pathogenic variants in the enhancers will ultimately enable the clarification of their function within the context of a uniform genetic background. The study of gene expression changes and their effect on neurogenesis will greatly benefit from the development of protocols for the differentiation of human pluripotent stem cells into brain organoids in vitro, which will replicate important human-specific aspects of brain development.

The interactions of the newly discovered genes with NDD-variant-containing enhancers will benefit both brain organoid studies and mouse models for the study of gene function. By performing targeted mutagenesis in human cells, human cellular models can be created using CRISPR-Cas9 to investigate the function of these new NDD-relevant genes.

References

1. Schroeter H, Spencer JP, Rice-Evans C, Williams RJ (2001) Flavonoids protect neurons from oxidized low-density-lipoprotein- induced apoptosis involving c-Jun N-terminal Kinase (JNK), c-Jun and caspase-3. Biochem J 358: 547-557.
2. Won DH, Kim LH, Jang B, Yang IH, Kwon HJ, et al.(2018) In vitro and in vivo anti-cancer activity of silymarin on oral cancer. Tumor Biology 40:
3. Ramasamy, Agarwal R (2008) Multitargeted therapy of cancer by silymarin. In Cancer Lett 269: 352-362.
4. Thorgeirsson TE, Geller F, Sulem P (2008) A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nat 452(7187): 638-642.
5. Ramalho MAP, Abreu AFB, Santos JB, Nunes JAR (2012) Aplicações da Genética Quantitativa no Melhoramento de Plantas Autógamas. Lavras Ufla 90: 365-456.
6. Beaver JS, Kelly JD (1994) Comparison of selection methods for dry bean populations derived from crosses between gene pools. Crop Sci 34: 34-37.
7. Schwarz C, Mehnert W, Lucks JS, Muller RH (1994) Solid lipid nanoparticles (SLN) for controlled drug delivery. I. Production, characterization and sterilization. J Controlled Release 30(1): 83-96.
8. Puglia C, Bonina F (2012) Lipid nanoparticles as novel delivery systems for cosmetics and dermal pharmaceuticals. Expert Opinion on Drug Delivery 9(4): 429-441.
9. Walters KA, Roberts MS (2002) Dermatological and Transdermal Formulations. Int J Derm 90: 14-85.
10. Armstrong DW, Wang X, Ercal N (1998) Enantiomeric composition of nicotine in smokeless tobacco, medicinal products and commercial reagents. Chir 10: 587-591.