

Review Article

The non-coding genome in genetic brain disorders: new targets for therapy?

 Eva Medico-Salsench¹, Faidra Karkala², Kristina Lanko¹ and  Tahsin Stefan Barakat¹

¹Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands; ²Department of Neuroscience, Erasmus MC University Medical Center, Rotterdam, The Netherlands

Correspondence: Kristina Lanko (k.lanko@erasmusmc.nl) or Tahsin Stefan Barakat (t.barakat@erasmusmc.nl)



The non-coding genome, consisting of more than 98% of all genetic information in humans and once judged as ‘Junk DNA’, is increasingly moving into the spotlight in the field of human genetics. Non-coding regulatory elements (NCREs) are crucial to ensure correct spatio-temporal gene expression. Technological advancements have allowed to identify NCREs on a large scale, and mechanistic studies have helped to understand the biological mechanisms underlying their function. It is increasingly becoming clear that genetic alterations of NCREs can cause genetic disorders, including brain diseases. In this review, we concisely discuss mechanisms of gene regulation and how to investigate them, and give examples of non-coding alterations of NCREs that give rise to human brain disorders. The cross-talk between basic and clinical studies enhances the understanding of normal and pathological function of NCREs, allowing better interpretation of already existing and novel data. Improved functional annotation of NCREs will not only benefit diagnostics for patients, but might also lead to novel areas of investigations for targeted therapies, applicable to a wide panel of genetic disorders. The intrinsic complexity and precision of the gene regulation process can be turned to the advantage of highly specific treatments. We further discuss this exciting new field of ‘enhancer therapy’ based on recent examples.

Introduction

The human genome contains more than 20,000 protein-coding genes whose expression needs to be precisely regulated in order for normal development and physiology to occur. A key component of this well-tuned orchestra of gene regulation are non-coding regulatory elements (NCREs), which modulate the magnitude, timing and cell specificity of gene expression. The cross-talk between coding and non-coding genome relies on multiple factors, from DNA sequence to 3D chromatin structure. As one of the most complex organs of the human body, the brain is particularly dependent on the correct timing, location and level of gene expression, enabling its development and functioning during life. It is therefore not surprising, that alterations in the regulatory processes of gene regulation, including chromatin modifications, often lead to neurodevelopmental disorders [1–6]. Constantly improving genetic diagnostic techniques, based on next-generation sequencing (NGS), are currently the most effective in identifying disease-causing (e.g. pathogenic) variants in the coding sequences [7]. However, in more than 50% of individuals affected by neurodevelopmental disorders a genetic cause remains elusive despite a high clinical suspicion for a genetic disorder, and it seems plausible that the explanation for this ‘missing heritability’ can be found in the non-coding genome in a significant number of cases [7]. The limited functional annotation of non-coding sequences and limitations in our understanding of NCRE function impedes and perplexes the interpretation of non-coding variants in a diagnostic setting. This, however, is anticipated to change as we witness more studies bridging the fundamental and clinical aspects of NCREs. Also, more than

Received: 24 February 2021
Revised: 12 July 2021
Accepted: 26 July 2021

Version of Record published:
27 October 2021

90% of disease-associated single nucleotide polymorphisms (SNPs) identified by genome-wide association studies (GWASs) are located in non-coding regions of the genome [8], reinforcing the need to better understand mechanisms of gene regulation, NCREs and the non-coding genome.

In this review, we first discuss the basics of gene regulatory mechanisms and methods that can help to functionally annotate NCREs. We then highlight a number of examples where alterations of NCREs lead to disease and discuss potential options to modify NCREs as future treatment modalities for genetic disorders.

Mechanisms of gene regulation by the non-coding genome in a nutshell

Both proximal and distal NCREs interact with target genes to control gene expression, and these include promoters, enhancers, silencers and insulators. Promoters are located next to the transcription start site of a gene, and switch the gene transcription on and off. Spatio-temporal regulation of promoter activity is enabled by the interaction of promoters with other distal NCREs, such as enhancers. Enhancers are relatively short sequences (100–1000 bp) that can reside in exons, introns or intergenic regions, which can be located in close proximity of target genes, but also being distantly located up to megabases (Mb) away from their targets [9]. Enhancers contain multiple binding sites for transcription factors (TFs), which recruit cofactors that together activate transcription of target genes, independently of the enhancer orientation [9,10]. Other predictive features of enhancers include their location in open chromatin when active and epigenetic modifications like enrichment of histone modifications, such as monomethylation of lysine 4 of histone H3 or acetylation of lysine 27 of histone H3 (H3K4me1 and H3K27ac) [11]. Indeed, TF binding capacity to enhancers can be influenced by DNA methylation and chromatin organization, although some TFs also have pioneer factor activity, which means that they can actively open-up closed chromatin regions, enabling activation of genes during development [12,13]. Enhancers are ultimately brought into close proximity to their target promoters by chromatin looping, enabling gene expression [9,14,15]. These enhancer–promoter interactions are mainly confined to topologically associated domains (TADs) and occur in open-chromatin regions. TADs are genomic domains with an average size of ~1 Mb that encompass regions that preferentially physically interact with other sequences within the same TAD, being delimited and maintained by boundary elements (insulators) [16,17] (Figure 1A). In contrast with enhancers, silencers are NCREs that upon binding of regulatory proteins repress gene expression [18].

There is accumulating evidence that many genes are regulated by more than one enhancer. This enhancer redundancy has been first described in *Drosophila* and is referred to as ‘shadow enhancers’, which drive (partially) overlapping spatio-temporal activity, conferring a level of redundancy and enforcing resistance to genetic variations [19,20]. Recently, the term ‘super enhancer’ (SE) has been coined which describes clusters of putative enhancers in close genomic proximity with unusually high potential to activate transcription of key genes that define cell identity [21–24]. SEs were proposed to differ from putative enhancers in size, TF density and sensitivity to perturbation [21,22,25,26]. Despite this interesting concept, a number of studies have challenged this view and have proposed that SEs are simply a collection of multiple enhancers located in close proximity, that together do not result in more enhancer activity than the sum of its parts [24,27]. Their precise role needs to be further investigated.

How to identify putative and functional NCREs

To identify putative NCREs, one can take advantage of the several distinct features at different epigenome layers that allow to distinguish NCREs from other non-coding sequences, such as specific chromatin marks, binding of TFs, chromatin accessibility and structural organization, and a variety of techniques that can detect these features.

Chromatin immunoprecipitation (ChIP) was introduced to study protein–DNA interactions, and its combination with NGS allows to determine genome-wide binding profiles of proteins of interest [28,29]. When applied to NCRE-associated proteins, such as lineage-specific TFs or enhancer-associated histone modifications such as H3K4me1 and H3K27ac, it allows to identify the location of putative enhancers genome-wide [29,30]. It is, however, important to note that not all of these putative enhancers will turn out to be ‘*bonafide*’ active enhancers. For example, several studies have shown that not all H3K27ac marked sites function as enhancers in validation experiments, and the genome-wide depletion of some of the enhancer associated marks seem to strikingly impact very little on gene expression [31–34]. This emphasizes that even though ChIP-seq of histone modifications and enhancer-associated TFs is very useful to identify putative enhancers, functional validation remains crucial to confirm these findings [31,35–37].

Another feature of enhancers is their need to be accessible for TFs and structural proteins, and therefore, they reside in open chromatin when active. These accessible DNA regions can be identified by multiple techniques, including the Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq). ATAC-seq takes advantage of transposons such as Tn5, which preferentially integrate into open chromatin regions and inserts

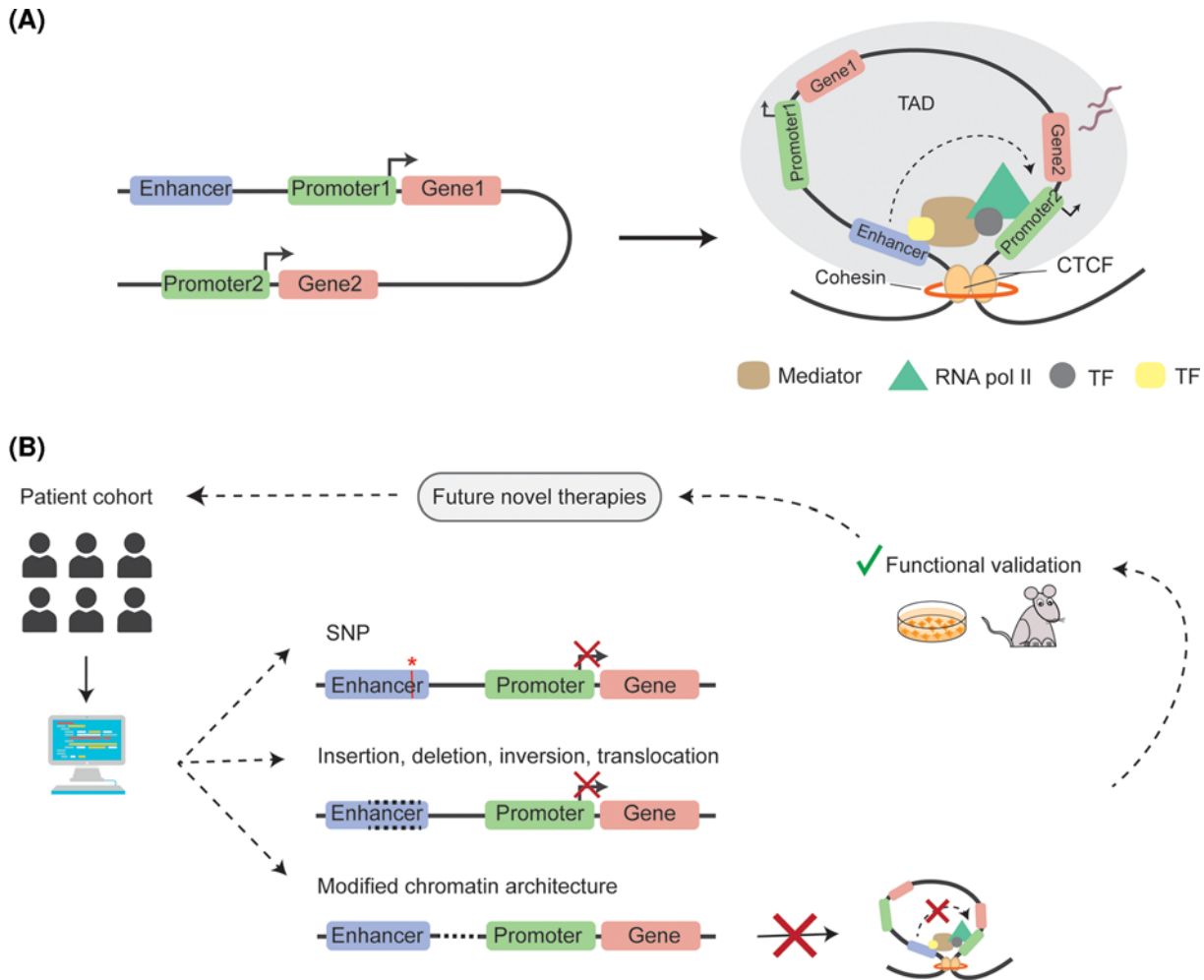


Figure 1. Enhancer–promoter interaction regulates gene expression

(A) Overview of enhancer–promoter interaction restricted within the TAD. The genome (black line) is organized by domains named TADs, which are established and defined by CTCF insulators and their interaction with the cohesin complex. To establish the enhancer–promoter interaction loop, TFs bind to both enhancer and the target promoter where the RNA Pol II complex is assembled to start transcription of the target gene. Other proteins, such as the Mediator complex, connect enhancer and promoter by interactions with TFs and the transcriptional machinery. (B) Flowchart illustrating how genetic variants in NCRE can contribute to disease, and how this knowledge could be utilized to develop future therapies. Through NGS-based studies and other future diagnostic approaches, variants in NCREs are identified. Computational analysis will help to prioritize possibly disease causing variants. These can include SNPs, insertions, deletions, inversions or translocations of enhancers; and alterations in chromatin looping preventing the proper formation of TADs, all potentially influencing or disrupting gene regulatory mechanisms and, therefore, gene expression output. The functional validation of identified variants can help to establish their pathogenicity and can provide cues for the development of future novel therapies targeting NCREs and gene regulatory mechanisms benefitting patients.

adapters by tagmentation [38]. These regions are then amplified and sequenced by NGS. One of the main advantages of ATAC-seq is that with minimal input material it can also identify other NCREs, such as promoters, which reside in accessible chromatin as well.

Chromatin architecture plays an essential role in gene regulation, facilitating enhancer–promoter interactions. Several methods exist to assess chromatin conformation. These include, among others, high-throughput chromosome conformation capture (Hi-C), which allows high-throughput contact mapping of genome-wide interactions, which allows to identify enhancer–promoter interactions and TADs [39]. Other techniques include methods that study the genome-wide interactions of selected regions of interest (such as promoter capture Hi-C [40]) or targets bound by a

protein of interest (such as HiChIP [41]). Together, these and other approaches allow to map chromatin conformation interactions, and the knowledge obtained from this can be used to predict the location of putative enhancers.

Although ChIP-seq, ATAC-seq and chromatin conformation assays are able to identify putative NCREs, functional validation remains crucial as not all putative enhancers in fact are real enhancers [42–44], even when localized in open-chromatin regions [33,45]. Classically, this validation was done by low throughput reporter assays, but new technologies enable us to significantly upscale functional validations. These include massively parallel reporter assays (MPRAs), such as STARR-seq [32], that allows the determination of enhancer activity on an episomal plasmid in a high-throughput manner, and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-based screens in which NCREs can be edited in high-throughput at endogenous loci, allowing to directly assess effects on endogenous gene expression [46]. MPRAs were first developed by Patwardhan et al. in 2009 for promoter assays and enables the analysis of thousands of regulatory elements in a single experiment [47,48]. Libraries of thousands of enhancers are *en masse* cloned in a reporter construct containing a unique barcode for each enhancer candidate [49]. The RNA-seq readout of these libraries after expression in cells or even *in vivo* in animal models is used to determine the enhancer function by linking the abundance of barcodes in mRNA to the element's activity [47]. This not only allows the testing of pre-selected genomic regions of choice, but also the screening of randomly fragmented sequences open chromatin regions, regions associated with TFs or histone marks, or even whole genomes [31,32,45,50–52]. Similar approaches can also be used to study the effects of genetic variants on function of putative enhancers via saturation mutagenesis experiments [53]. Together, with all these different currently available tools, the field now seems for the first time ready to start deciphering the regulatory grammar of the non-coding genome, and provide functional annotation to non-coding sequences. Knowledge obtained from this will enable us to better interpret genetic variation encountered in patients and might elucidate unknown causes of disease in the field of human genetics (Figure 1B).

Non-coding variants in the context of genetic brain disease

The fact that evolution has resulted in such complex, well-organized processes of gene regulation, and the obvious need for these processes for the development of an organism [54], let it not come as a surprise that disturbance of these mechanisms can lead to human disease. Aberrations of NCREs have indeed been implicated in the pathogenesis of human disorders. A classic example is pre-axial polydactyly, where variants in a limb-specific enhancer, the so-called ZRS region located in intron 5 of the *LMBR1* gene, causes disturbed expression of *SHH*, located more than 1 Mb away, resulting in pre-axial polydactyly [55,56]. Other examples include a variant in a silencer of *NOTCH1*, which was shown to contribute to Tetralogy of Fallot (a congenital heart disease) [57] and translocations of the NCREs of *PAX6* that lead to aniridia [56]. Multiple disease examples show that the human brain seems particularly sensitive to gene expression disbalance [4,43,58], and many neurodevelopmental disorders are linked to chromatin modifiers [5,59–64] and architectural proteins with roles in establishing chromatin conformation, such as CTCF, YY1 and STAG1 [1,3,4,65,66]. In this section, we will focus on a number of recent examples that illustrate the wide range of alterations of gene regulatory processes that can cause brain-related disorders, and the novel approaches to identify them. These alterations can range from deletions or duplications of NCREs identified by whole-genome sequencing (WGS) or copy number analysis, point mutations in NCREs detected by WGS or targeted sequencing, disruption of TAD boundaries encountered by chromatin conformation studies, and dysfunction of proteins involved in gene regulatory processes, just to mention a few (for further in-depth review, see [1,43,67,68]) (Figure 1B).

Primary familial brain calcification (PFBC) is a rare microvascular calcifying disorder presenting with neuropsychiatric symptoms, which is caused by haploinsufficiency of an inorganic phosphate transporter, *SLC20A2*. Recently, a deletion encompassing an enhancer region upstream of *SLC20A2* was identified in three patients with PFBC [69], and resulted in decreased *SLC20A2* expression and phosphate uptake to a similar level as observed in *SLC20A2* haploinsufficiency. Similar enhancer deletions have been found in proximity of *GABRD*, a susceptibility factor for juvenile myoclonic epilepsy [70], in a study that performed WGS analysis in an epilepsy cohort [71]. In 28.8% of 198 patients, copy number variants were found near known epilepsy genes, indicating that disruption of the gene regulatory landscape of these genes might be the cause of epilepsy [71]. Two studies have employed targeted sequencing of putative NCREs and provided evidence of single nucleotide alterations at NCREs being causative of developmental disorders. Based on targeted NCRE sequencing of 7930 individuals from the Deciphering Developmental Disorders (DDD) study, it was estimated that 1–3% of patients with neurodevelopmental disorders carry *de novo* pathogenic variant in fetal brain-active NCREs [72], although the number of NCREs investigated in this study was rather limited, possibly underestimating the real contribution of genetic variants in NCREs. Using another targeted sequencing approach, bi-allelic variants in NCREs were found in approximately 5% of consanguineous cases of autism spectrum

disorder (ASD) [73]. Interestingly, these variants are located in human accelerated regions (HARs), which are conserved genomic loci with elevated divergence in humans. Won et al. further confirmed that non-coding elements in HARs and human gained/lost enhancers interact with genes involved in human brain development, including risk genes for ASD, developmental delay and schizophrenia [74].

One of the challenges in the field remains the linking of encountered non-coding variants from WGS or GWAS to their target genes and, therefore, to neurodevelopmental disorders, as functional annotation of NCREs is still limited. Rapidly accumulating chromatin interaction data may change this. For example, several GWAS risk loci could be assigned to schizophrenia-associated genes based on 3D chromatin interaction maps of developing human cortex derived using Hi-C [75]. Hi-C is even about to enter the clinical setting for individual patients. In a recent study, Hi-C maps of cells derived from nine patients presenting with developmental disorders and harboring structural variants revealed altered TADs and regulatory elements [76]. One of these structural variants from a patient caused altered levels of *FOXG1*. Since variants in *FOXG1* and its regulatory elements were reported in congenital Rett syndrome and Rett syndrome-like phenotypes, the new variant is likely underlying the condition of this individual [77]. Interestingly, TAD-disruption appears to be a common mechanism underlying developmental disorders [78]. Deregulation of gene expression was reported in 7.3% of 219 cases of neurodevelopmental disorders, including ten individuals with structural variants disrupting the regulation of *MEF2C*, which causes intellectual disability (ID), epilepsy and cerebral malformations [79–83]. Patients carrying TAD-disrupting variants distal to *MEF2C* had lower *MEF2C* expression (as assessed in lymphoblastoid cell lines), which suggests TAD disruption as a likely explanation of patient condition.

Another challenging point in variant analysis and functional validation is the tissue-specificity of NCREs (reviewed in [43,84]), and this might be overcome by integrating data derived from relevant *in vitro* models, such as induced pluripotent stem cell (iPSC)-derived or primary neuronal cells. Combination of chromatin accessibility analysis in iPSC-derived excitatory neurons, and risk variants identified in GWAS, allowed the identification of 108 potentially functional loci in NCREs, that might contribute to schizophrenia [85]. A risk variant located at the *MIR137* locus, at a neuron-specific TF-binding site, was confirmed to alter the chromatin accessibility of the *MIR137* promoter. This resulted in altered gene expression which changed dendrite and synapse formation. Another comprehensive study in four iPSC-derived neuronal cell types integrated Hi-C, open chromatin and RNA-seq data to identify long-range enhancer–promoter interactions connecting NCREs to their target genes [86]. The approach was confirmed by CRISPR interference (CRISPRi) targeting cell type-specific regulatory regions of *CDK5RAP3*. *CDK5RAP3* was down-regulated in excitatory neurons when CRISPRi was directed to regions 1 and 2, whereas the effect of interfering with region 3 was only observed in astrocytes and motor neurons. The generated dataset allowed to link 70% of SNPs from GWASs for 11 neuropsychiatric disorders to the promoter-interacting region in one or more cell type. Nott et al. identified NCREs using ATAC-seq and ChIP-seq in major brain cell types from cortical brain tissues obtained from six individuals [87]. Risk variants for neurological and psychiatric disorders were then mapped to these regions. The functional effect of a deletion of the microglia-specific *BIN1* enhancer that contained an Alzheimer's disease-risk variant, was confirmed in microglia but, not in neurons or astrocytes, which were differentiated from the same CRISPR-Cas9 edited human iPSC. Nevertheless, the high-throughput validation of these data remains challenging as each enhancer variant needs to be tested in the respective cell type.

Interestingly, it is becoming clear that enhancers can also predict the pathogenicity of genes that are associated with them [88]. Using machine learning, Wang et al. established an enhancer-domain score (EDS), which takes several parameters into consideration, including the enhancer-domain size, enhancer redundancy, the number of functional nucleotides, TF-binding sites and tissue-specific activity [89]. This revealed that genes regulated by enhancers with a higher EDS are more likely associated with human diseases than genes with a lower score. Also, a higher EDS was linked to genes resilient against genetic disturbance. In other words, functionally important genes might overcome the effects of functionally disturbing genetic variants by using a set of redundant NCREs for their regulation. Furthermore, it was shown that enhancers with a higher EDS regulate genes which are depleted of *cis*-expression quantitative trait loci (*cis*-eQTLs) [89]. *cis*-eQTLs are loci near a gene of interest which partly explain the variation of gene-expression levels [90]. As EDS correlates variations in regulatory regions with phenotypic effects, it could be used as a tool to find the most clinically relevant enhancers and complement other methods which measure the importance of dosage-sensitive genes. This approach integrated with open chromatin data in iPSC-derived neurons revealed potential causative variants of schizophrenia which also influenced chromatin accessibility [91].

In summary, considerable progress has been made in identifying and linking variants in NCREs to their target genes in various cell types in the context of neurological disorders. The integration of these data with functional validation would not only be crucial to enhance our understanding of the functioning of NCREs and their role in neurodevelopment, but also of immense value to improve diagnostics. One could hypothesize that upon functional annotation and

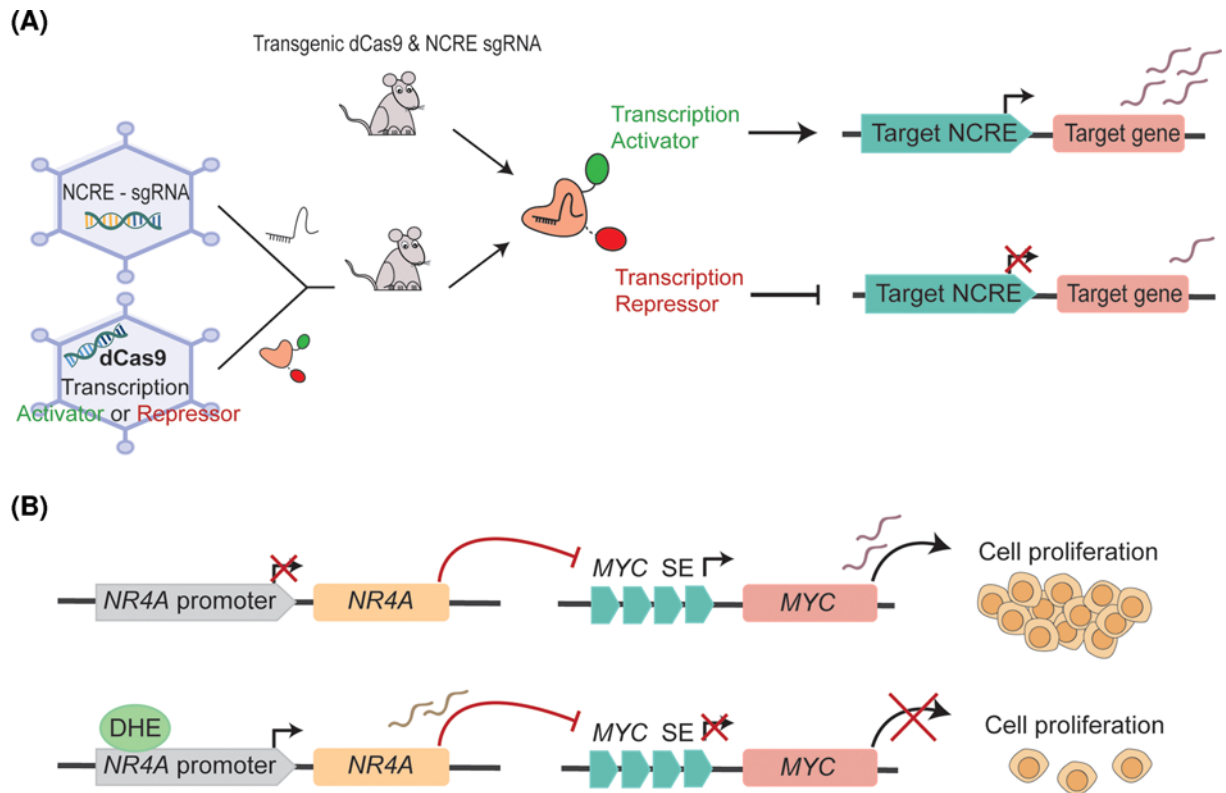


Figure 2. Different approaches to target enhancers as therapeutic agents

(A) Schematic representation of targeting enhancer activity via dCas9 fused with a transcription activator or repressor to treat disease. Transgenic mice carrying the dCas9 and the single-guide RNA (sgRNA) can either activate or repress the target NCRE and, therefore, promote or down-regulate gene expression, respectively. Both, dCas9 fused with the transcription repressor or activator and sgRNA can also be delivered by two independent recombinant adeno-associated viruses (rAAVs) in mice, which could potentially be translated to the clinic. **(B)** Enhancer activity can be controlled by small molecules like dihydroergotamine (DHE). In absence of DHE, the *NR4A* promoter remains inactive allowing activation of *MYC* SE cluster which results in enhanced *MYC* expression and, consequently, cell proliferation. DHE induces *NR4A* expression by binding to its promoter, and *NR4A* suppresses the *MYC* SE thus *MYC* expression and cell proliferation in acute myeloid leukemia are inhibited. Similar approaches could be designed for other disorders, based on knowledge obtained from gene regulatory mechanism studies. Abbreviation: dCas9, nuclease-deficient Cas9.

validation of more NCREs, it will become feasible to sequence NCREs at high throughput in genetically unexplained patients with neurogenetic disorders as a routine diagnostic approach, or to use the information derived from NCRE studies as a computational filter in WGS analysis, thereby increasing the diagnostic yield in clinical genetics (Figure 1B).

Modulating enhancer activity for future disease treatment

The increasing knowledge of NCREs relevance in neurodevelopmental diseases opens novel strategies for therapeutic interventions. So far, genome-editing strategies to rescue a disease phenotype have mainly focused on protein-coding regions [92]. Although less progress has been made on modifying or influencing NCREs, a number of recent studies show great promise for such ‘enhancer treatment’, either by influencing NCRE activity by chemical compounds or directly altering NCREs using CRISPR-Cas9-based approaches (Figure 2). The examples that we discuss in this section have clearly benefitted from an in-depth functional understanding of NCREs and gene regulatory mechanisms obtained from fundamental studies, reinforcing the need of these to fully benefit translational studies.

An interesting example of the use of small molecules to modify NCRE activity comes from the oncology field. Activation of the oncogene *MYC* is a key driver for acute myeloid leukemia (AML), and is driven by a cluster of enhancers which activity is negatively regulated by the nuclear receptor *NR4A* [93]. *NR4A* expression can be induced

by dihydroergotamine (DHE), an FDA-approved drug against migraine [94]. This NR4A-inducing activity of DHE was leveraged as an enhancer targeting therapy and resulted in reduced enhancer activity with dismissal of enhancer bound co-activators, reduced H3K27ac and diminished enhancer RNA transcription, ultimately resulting in reduced *MYC* expression and AML cell proliferation [95] (Figure 2B). This illustrates the yet-to-be unraveled potential of small molecules in enhancer treatment, especially when identified in drug repurposing screens, allowing rapid bench to bedside translation.

In the gene therapy field, recombinant adeno-associated virus (rAAV) are promising vectors to re-establish gene expression in genetic disorders to compensate for haploinsufficiency, but this might lead to ectopic transgene expression [96]. CRISPR-mediated gene activation (CRISPRa), in which a nuclease-deficient Cas9 (dCas9) is fused to a transcriptional activator, such as VP64, and is targeted using guide RNAs (gRNAs) to a promoter or enhancer, might overcome these limitations [97]. This approach was recently applied to mice that show obesity due to haploinsufficiency of the genes *SIM* and *Mc4r* [98]. By using gRNAs targeting the hypothalamic enhancer and promoter of *Sim* and the promoter of *Mc4r*, expression of these genes could be up-regulated which was sufficient to rescue the phenotype. As a pre-clinical proof of concept, also rAAV-mediated delivery of CRISPRa was performed demonstrating the possible clinical translation of this approach to humans to treat other gene-dosage abnormalities [98]. A similar approach was applied to a mouse model of Dravet syndrome (DS), a severe epileptic encephalopathy caused by loss-of-function mutations in *SCN1A*. Using rAAV-mediated delivery to the brain of DS mouse pups, dCas9 fused to the transcriptional activator VP160 was targeted by gRNAs to the *Scn1a* promoter and resulted in up-regulated *Scn1a* expression from the wildtype allele [99]. As dCas9 can also be fused to repressive domains [100], similar approaches could also be developed to silence genes, either by directly silencing the promoter or to interfere with other NCREs, and might therefore also provide options to treat disorders caused by gain-of-function mutations [100]. It will be interesting to test whether these approaches could also be further developed to influence gene expression in an allele-specific manner, or to induce expression of specific gene isoforms which are normally not expressed but could rescue phenotypes. The latter could provide a targeted therapy for a severe epileptic encephalopathy caused by a recurrent mutation in *UGP2* (*Barakat–Perenthaler syndrome*, OMIM# 618744), in which the shorter UGP2 isoform is not expressed in brain due to the mutation, and up-regulation of the longer UGP2 isoform which is normally absent from brain could potentially result in a therapy [101].

Obviously, all the approaches discussed in this section will require extensive optimizations and pre-clinical testing to ensure that the right expression levels can be achieved to benefit patients. It is however promising and inspiring for the field, that a first enhancer focused gene therapy that uses CRISPR-Cas9 technology has even recently made it into the clinic [102]. Transfusion-dependent β -thalassemia and sickle cell disease are difficult to treat genetic disorders primarily affecting red blood cells, caused by mutations in hemoglobin β [103,104]. Previous work had shown that induction of fetal hemoglobin (HbF) can improve symptoms in both disorders [105]. *BCL11A* is a repressor of γ -globin expression and HbF production in adult erythrocytes, and its expression is under control of the *BCL11A* enhancer. Using CRISPR-Cas9, the *BCL11A* enhancer was deleted in CD34⁺ hematopoietic stem and progenitor cells, enabling autologous stem cell transplantation in β -thalassemia and sickle cell disease patients. This resulted in long-lasting engraftment of the enhancer deleted cells, resulting in increased HbF expression and improved clinical outcomes [102]. This elegantly illustrates how similar approaches could also be designed for other genetic disorders, including those affecting the nervous system.

Conclusion

The increasing number of examples where variants in NCREs are linked to human diseases, including neurodevelopmental phenotypes, certainly warrants their evaluation in a clinical diagnostic setting. To further exploit this potential, not only for diagnostics but also for therapies, the field still needs to address many key questions. This includes but is not limited to deciphering the logic behind tissue-specific NCRE activity, the causes and consequences of enhancer redundancy, the establishment of 3D genome organization and the full annotation of all functional non-coding sequences. The many technology advancements in the field, including functional genomics assays and CRISPR-Cas9-based tools, will certainly help to achieve this. But of equal importance will be studying the genetics of human beings, either in large cohorts of healthy individuals to understand normal genetic variation outside of protein-coding genes, or in the context of rare genetic disorders that impact on gene regulatory mechanisms, and which might inform us on new targets and pathways for future therapies. Despite the tremendous potential that NCREs have for therapy, it is clear that it will not be straight forward to develop such treatments easily, as numerous parameters like off-target effects, cell-type specificity, and levels of gene expression need to be defined in order to apply ‘enhancer treatment’ in a wide range of patients. However, the new studies involving enhancer treatment in animal models as well as the first human clinical trials give hope for overcoming the hurdles in the foresight future.

Summary

- NCREs are important regulators of gene expression.
- Technology innovations and new model systems allow improved functional annotation of NCREs, which increases our capacity to understand genetic variation encountered outside of protein-coding genes.
- There is a rising number of genetic disorders, including neurodevelopmental phenotypes, that is linked to non-coding variants, and the first therapies targeted to NCREs are being developed.
- Increasing our knowledge on NCREs and gene regulatory mechanisms further will enable improved future diagnostics and therapy options for patients with genetic disorders.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

The Barakat lab was supported by the Netherlands Organisation for Scientific Research (ZonMW Veni) [grant number 91617021]; the NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation; the Erasmus MC Fellowship 2017; and the Erasmus MC Human Disease Model Award 2018.

Author Contribution

E.M.S., F.K., K.L. and T.S.B. performed the literature search and wrote the manuscript. E.M.S. and F.K. created the figures. K.L. and T.S.B. conceived the work. T.S.B. obtained the funding and supervised the work.

Acknowledgements

The authors would like to thank all members of the Barakat lab for helpful discussions. We apologize to those colleagues whose work could not be cited due to space limitations.

Abbreviations

AML, acute myeloid leukemia; ASD, autism spectrum disorder; ATAC-seq, Assay for Transposase Accessible Chromatin with high-throughput sequencing; ChIP, chromatin immunoprecipitation; *cis*-eQTL, *cis*-expression quantitative trait loci; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats; CRISPRa, CRISPR-mediated gene activation; CRISPRi, CRISPR interference; dCas9, nuclease-deficient Cas9; DDD, Deciphering Developmental Disorders; DHE, dihydroergotamine; DS, Dravet syndrome; EDS, enhancer-domain score; gRNA, guide RNA; GWAS, genome-wide association study; H3K27ac, acetylation of lysine 27 of histone H3; H3K4me1, monomethylation of lysine 4 of histone H3; HAR, human accelerated region; HbF, fetal hemoglobin; Hi-C, high-throughput chromosome conformation capture; iPSC, induced pluripotent stem cell; Mb, megabase; MPRA, massively parallel reporter assay; NCRE, non-coding regulatory element; NGS, next-generation sequencing; PFBC, primary familial brain calcification; rAAV, recombinant adeno-associated virus; SE, super enhancer; SNP, single nucleotide polymorphism; TAD, topologically associated domain; TF, transcription factor; WGS, whole-genome sequencing.

References

- 1 D'haene, E. and Vergult, S. (2021) Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. *Genet. Med.* **23**, 34–46, <https://doi.org/10.1038/s41436-020-00974-1>
- 2 Turan, S., Boerstler, T., Kavyanifar, A., Loskarn, S., Reis, A., Winner, B. et al. (2019) A novel human stem cell model for Coffin-Siris syndrome-like syndrome reveals the importance of SOX11 dosage for neuronal differentiation and survival. *Hum. Mol. Genet.* **28**, 2589–2599, <https://academic.oup.com/hmg/article/28/15/2589/5481587>, <https://doi.org/10.1093/hmg/ddz089>
- 3 Gregor, A., Oti, M., Kouwenhoven, E.N., Hoyer, J., Sticht, H., Ekici, A.B. et al. (2013) De novo mutations in the genome organizer CTCF cause intellectual disability. *Am. J. Hum. Genet.* **93**, 124–131, <https://pubmed.ncbi.nlm.nih.gov/23746550>, <https://doi.org/10.1016/j.ajhg.2013.05.007>
- 4 Gabriele, M., Vulto-van Silfhout, A.T., Germain, P.-L., Vitriolo, A., Kumar, R., Douglas, E. et al. (2017) YY1 haploinsufficiency causes an intellectual disability syndrome featuring transcriptional and chromatin dysfunction. *Am. J. Hum. Genet.* **100**, 907–925, <https://linkinghub.elsevier.com/retrieve/pii/S0002929717301933>, <https://doi.org/10.1016/j.ajhg.2017.05.006>

- 5 Jones, W.D., Dafou, D., McEntagart, M., Woollard, W.J., Elmslie, F.V., Holder-Espinasse, M. et al. (2012) De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am. J. Hum. Genet.* **91**, 358–364, <http://pmc/articles/PMC3415539/>, <https://doi.org/10.1016/j.ajhg.2012.06.008>
- 6 Kranz, A. and Anastasiadis, K. (2020) The role of SETD1A and SETD1B in development and disease. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194578, <https://doi.org/10.1016/j.bbagr.2020.194578>
- 7 Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A. et al. (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* **369**, 1502–1511, <https://doi.org/10.1056/NEJMoa1306555>
- 8 Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S. et al. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9362–9367, <http://www.genome.gov/gwastudies>, <https://doi.org/10.1073/pnas.0903103106>
- 9 Schoenfelder, S. and Fraser, P. (2019) Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455, <https://doi.org/10.1038/s41576-019-0128-0>
- 10 Yáñez-Cuna, J.O., Dinh, H.Q., Kvon, E.Z., Shlyueva, D. and Stark, A. (2012) Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. *Genome Res.* **22**, 2018–2030, <https://doi.org/10.1101/gr.132811.111>
- 11 Hyun, K., Jeon, J., Park, K. and Kim, J. (2017) Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, e324, <http://www.nature.com/articles/emm201711>, <https://doi.org/10.1038/emm.2017.11>
- 12 Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S. et al. (2017) Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science (80-)* **356**, eaaj2239, <http://science.sciencemag.org/content/356/6337/eaaj2239.abstract>, <https://doi.org/10.1126/science.aaj2239>
- 13 Iwafuchi-Doi, M. (2019) The mechanistic basis for chromatin regulation by pioneer transcription factors. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11**, e1427, <https://doi.org/10.1002/wsbm.1427>
- 14 Li, J., Hsu, A., Hua, Y., Wang, G., Cheng, L., Ochiai, H. et al. (2020) Single-gene imaging links genome topology, promoter-enhancer communication and transcription control. *Nat. Struct. Mol. Biol.* **27**, 1032–1040, <https://doi.org/10.1038/s41594-020-0493-6>
- 15 Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.M. et al. (2015) The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* **25**, 582–597, <http://pmc/articles/PMC4381529/>, <https://doi.org/10.1101/gr.185272.114>
- 16 Rowley, M.J. and Corces, V.G. (2018) Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800, <https://doi.org/10.1038/s41576-018-0060-8>
- 17 Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F. et al. (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64, <http://www.nature.com/articles/nature21429>, <https://doi.org/10.1038/nature21429>
- 18 Jayavelu, N.D., Jajodia, A., Mishra, A. and Hawkins, R.D. (2020) Candidate silencer elements for the human and mouse genomes. *Nat. Commun.* **11**, 1–15, <https://www.nature.com/articles/s41467-020-14853-5>, <https://doi.org/10.1038/s41467-020-14853-5>
- 19 Cannavò, E., Khoueiry, P., Garfield, D.A., Geleher, P., Zichner, T., Gustafson, E.H. et al. (2016) Shadow enhancers are pervasive features of developmental regulatory networks. *Curr. Biol.* **26**, 38–51, <https://linkinghub.elsevier.com/retrieve/pii/S0960982215014281>, <https://doi.org/10.1016/j.cub.2015.11.034>
- 20 Kvon, E.Z., Waymack, R., Elabd, M.G. and Wunderlich, Z. (2021) Enhancer redundancy in development and disease. *Nat. Rev. Genet.* **22**, 1–13, www.nature.com/nrg, <https://doi.org/10.1038/s41576-020-00311-x>
- 21 Tang, F., Yang, Z., Tan, Y. and Li, Y. (2020) Super-enhancer function and its application in cancer targeted therapy. *Npj Precis Oncol.* **4**, 2, <https://doi.org/10.1038/s41698-020-0108-z>
- 22 Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H. et al. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319, <https://www.sciencedirect.com/science/article/pii/S0092867413003929>, <https://doi.org/10.1016/j.cell.2013.03.035>
- 23 Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y. et al. (2018) Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243, <http://pmc/articles/PMC5808607?report=abstract>, <https://doi.org/10.1038/nature25461>
- 24 Hay, D., Hughes, J.R., Babbs, C., Davies, J.O.J., Graham, B.J., Hanssen, L. et al. (2016) Genetic dissection of the α -globin super-enhancer in vivo. *Nat. Genet.* **48**, 895–903, <https://doi.org/10.1038/ng.3605>
- 25 Jia, Y., Chng, W.-J. and Zhou, J. (2019) Super-enhancers: critical roles and therapeutic targets in hematologic malignancies. *J. Hematol. Oncol.* **12**, 77, <https://doi.org/10.1186/s13045-019-0757-y>
- 26 Wang, X., Cairns, M.J. and Yan, J. (2019) Super-enhancers in transcriptional regulation and genome organization. *Nucleic Acids Res.* **47**, 11481–11496, <https://doi.org/10.1093/nar/gkz1038>
- 27 Shin, H.Y., Willi, M., HyunYoo, K., Zeng, X., Wang, C., Metser, G. et al. (2016) Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat. Genet.* **48**, 904–911, 2016/07/04, <https://doi.org/10.1038/ng.3606>
- 28 Solomon, M.J., Larsen, P.L. and Varshavsky, A. (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* **53**, 937–947, [https://doi.org/10.1016/S0092-8674\(88\)90469-2](https://doi.org/10.1016/S0092-8674(88)90469-2)
- 29 Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science (80-)* **316**, 1497, <http://science.sciencemag.org/content/316/5830/1497.abstract>, <https://doi.org/10.1126/science.1141319>
- 30 Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T. et al. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657, <https://doi.org/10.1038/nmeth1068>

- 31 Barakat, T.S., Halbritter, F., Zhang, M., Rendeiro, A.F., Perenthaler, E., Bock, C. et al. (2018) Functional dissection of the enhancer repertoire in human embryonic stem cells. *Cell Stem Cell* **23**, 276.e8–288.e8, <https://linkinghub.elsevier.com/retrieve/pii/S1934590918302960>, <https://doi.org/10.1016/j.stem.2018.06.014>
- 32 Muerdter, F., Boryn, L.M. and Arnold, C.D. (2015) STARR-seq — principles and applications. *Genomics* **106**, 145–150, <http://www.sciencedirect.com/science/article/pii/S0888754315300100>, <https://doi.org/10.1016/j.ygeno.2015.06.001>
- 33 Catarino, R.R. and Stark, A. (2018) Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes Dev.* **32**, 202–223, <https://doi.org/10.1101/gad.310367.117>
- 34 Zhang, T., Zhang, Z., Dong, Q., Xiong, J. and Zhu, B. (2020) Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol.* **21**, 45, <https://doi.org/10.1186/s13059-020-01957-w>
- 35 Kwasniewski, J.C., Fiore, C., Chaudhari, H.G. and Cohen, B.A. (2014) High-throughput functional testing of ENCODE segmentation predictions. *Genome Res.* **24**, 1595–1602, <http://pmc/articles/PMC4199366/>, <https://doi.org/10.1101/gr.173518.114>
- 36 Rickels, R., Herz, H.M., Sze, C.C., Cao, K., Morgan, M.A., Collings, C.K. et al. (2017) Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat. Genet.* **49**, 1647–1653, <http://pmc/articles/PMC5663216/>, <https://doi.org/10.1038/ng.3965>
- 37 Dorighi, K.M., Swigut, T., Henriques, T., Bhanu, N.V., Scruggs, B.S., Nady, N. et al. (2017) Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol. Cell* **66**, 568.e4–576.e4, <http://pmc/articles/PMC5662137/>, <https://doi.org/10.1016/j.molcel.2017.04.018>
- 38 Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. and Greenleaf, W.J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218, <https://pubmed.ncbi.nlm.nih.gov/24097267/>, <https://doi.org/10.1038/nmeth.2688>
- 39 Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A. et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (80-)* **326**, 289–293, <https://doi.org/10.1126/science.1181369>
- 40 Schoenfelder, S., Javierre, B.-M., Furlan-Magaril, M., Wingett, S.W. and Fraser, P. (2018) Promoter capture Hi-C: high-resolution, genome-wide profiling of promoter interactions. *J. Vis. Exp.* **2018**, 57320, <http://pmc/articles/PMC6102006/>
- 41 Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J. et al. (2016) HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922, <https://doi.org/10.1038/nmeth.3999>
- 42 Halfon, M.S. (2019) Studying transcriptional enhancers: the founder fallacy, validation creep, and other biases. *Trends Genet.* **35**, 93–103, <https://doi.org/10.1016/j.tig.2018.11.004>
- 43 Perenthaler, E., Yousefi, S., Niggel, E. and Barakat, T.S. (2019) Beyond the exome: the non-coding genome and enhancers in neurodevelopmental disorders and malformations of cortical development. *Front. Cell. Neurosci.* **13**, 352, <https://www.frontiersin.org/article/10.3389/fncel.2019.00352>, <https://doi.org/10.3389/fncel.2019.00352>
- 44 Young, R.S., Kumar, Y., Bickmore, W.A. and Taylor, M.S. (2017) Bidirectional transcription initiation marks accessible chromatin and is not specific to enhancers. *Genome Biol.* **18**, 242, <https://pubmed.ncbi.nlm.nih.gov/29284524/>, <https://doi.org/10.1186/s13059-017-1379-8>
- 45 Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M. and Stark, A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science (80-)* **339**, 1074–1077, <https://doi.org/10.1126/science.1232542>
- 46 Montalbano, A., Canver, M.C. and Sanjana, N.E. (2017) High-throughput approaches to pinpoint function within the noncoding genome. *Mol. Cell* **68**, 44–59, <https://doi.org/10.1016/j.molcel.2017.09.017>
- 47 Inoue, F. and Ahituv, N. (2015) Decoding enhancers using massively parallel reporter assays. *Genomics* **106**, 159–164, <http://www.sciencedirect.com/science/article/pii/S0888754315300082>, <https://doi.org/10.1016/j.ygeno.2015.06.005>
- 48 Patwardhan, R.P., Lee, C., Litvin, O., Young, D.L., Pe'er, D. and Shendure, J. (2009) High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* **27**, 1173–1175, <https://doi.org/10.1038/nbt.1589>
- 49 Trauernicht, M., Martinez-Ara, M. and van Steensel, B. (2020) Deciphering gene regulation using massively parallel reporter assays. *Trends Biochem. Sci.* **45**, 90–91, <https://doi.org/10.1016/j.tibs.2019.10.006>
- 50 Ashuach, T., Fischer, D.S., Kreimer, A., Ahituv, N., Theis, F.J. and Yosef, N. (2019) MPRAnalyze: statistical framework for massively parallel reporter assays. *Genome Biol.* **20**, 183, <https://doi.org/10.1186/s13059-019-1787-z>
- 51 Wang, X., He, L., Goggin, S.M., Saadat, A., Wang, L., Sinnott-Armstrong, N. et al. (2018) High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. *Nat. Commun.* **9**, 5380, <https://doi.org/10.1038/s41467-018-07746-1>
- 52 Peng, T., Zhai, Y., Atlasi, Y., ter Huurne, M., Marks, H., Stunnenberg, H.G. et al. (2020) STARR-seq identifies active, chromatin-masked, and dormant enhancers in pluripotent mouse embryonic stem cells. *Genome Biol.* **21**, 243, <https://doi.org/10.1186/s13059-020-02156-3>
- 53 Kircher, M., Xiong, C., Martin, B., Schubach, M., Inoue, F., Bell, R.J.A. et al. (2019) Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution. *Nat. Commun.* **10**, 3583, <https://doi.org/10.1038/s41467-019-11526-w>
- 54 Sonawane, A.R., Platig, J., Fagny, M., Chen, C.Y., Paulson, J.N., Lopes-Ramos, C.M. et al. (2017) Understanding tissue-specific gene regulation. *Cell Rep.* **21**, 1077–1088, <https://doi.org/10.1016/j.celrep.2017.10.001>
- 55 Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A. et al. (2003) A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735, <https://doi.org/10.1093/hmg/ddg180>
- 56 Kleinjan, D.A., Seawright, A., Schedl, A., Quinlan, R.A., Danes, S. and Van Heyningen, V. (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum. Mol. Genet.* **10**, 2049–2059, www.hgu.mrc.ac.uk/Softdata/PAX6/, <https://doi.org/10.1093/hmg/10.19.2049>
- 57 Wang, Y., Jiang, T., Tang, P., Wu, Y., Jiang, Z., Dai, J. et al. (2020) Family-based whole-genome sequencing identifies compound heterozygous protein-coding and noncoding mutations in tetralogy of Fallot. *Gene* **741**, 144555, <https://doi.org/10.1016/j.gene.2020.144555>

- 58 Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I. et al. (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793, <http://www.nature.com/articles/ng.646>, <https://doi.org/10.1038/ng.646>
- 59 Weerts, M.J.A., Lanko, K., Guzmán-Vega, F.J., Jackson, A., Ramakrishnan, R., Cardona-Londoño, K.J. et al. (2021) Delineating the molecular and phenotypic spectrum of the SETD1B-related syndrome. *Genet. Med.*, <https://www.nature.com/articles/s41436-021-01246-2>, <https://doi.org/10.1038/s41436-021-01246-2>
- 60 Yu, X., Yang, L., Li, J., Li, W., Li, D., Wang, R. et al. (2019) De novo and inherited SETD1A variants in early-onset epilepsy. *Neurosci. Bull.* **35**, 1045–1057, <https://doi.org/10.1007/s12264-019-00400-w>
- 61 Labonne, J.D.J., Lee, K.H., Iwase, S., Kong, I.K., Diamond, M.P., Layman, L.C. et al. (2016) An atypical 12q24.31 microdeletion implicates six genes including a histone demethylase KDM2B and a histone methyltransferase SETD1B in syndromic intellectual disability. *Hum. Genet.* **135**, 757–771, <https://doi.org/10.1007/s00439-016-1668-4>
- 62 Kummeling, J., Stremmelaar, D.E., Raun, N., Reijnders, M.R.F., Willemsen, M.H., Ruiterskamp-Versteeg, M. et al. (2020) Characterization of SETD1A haploinsufficiency in humans and Drosophila defines a novel neurodevelopmental syndrome. *Mol. Psychiatry* 1–12, <https://doi.org/10.1038/s41380-020-0725-5>
- 63 Kleefstra, T., Kramer, J.M., Neveling, K., Willemsen, M.H., Koemans, T.S., Vissers, L.E.L.M. et al. (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* **91**, 73–82, <http://pmc/articles/PMC3397275/>, <https://doi.org/10.1016/j.ajhg.2012.05.003>
- 64 Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S. et al. (2015) Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin. Genet.* **87**, 252–258, <https://pubmed.ncbi.nlm.nih.gov/24527667/>, <https://doi.org/10.1111/cge.12363>
- 65 Lehalle, D., Mosca-Boidron, A.-L., Begtrup, A., Boute-Benejean, O., Charles, P., Cho, M.T. et al. (2017) STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. *J. Med. Genet.* **54**, 479, <http://jmg.bmj.com/content/54/7/479.abstract>, <https://doi.org/10.1136/jmedgenet-2016-104468>
- 66 Verheul, T.C.J., van Hijfte, L., Perenthaler, E. and Barakat, T.S. (2020) The why of YY1: mechanisms of transcriptional regulation by yin yang 1. *Front. Cell Dev. Biol.* **8**, 1034, <https://www.frontiersin.org/article/10.3389/fcell.2020.592164/full>, <https://doi.org/10.3389/fcell.2020.592164>
- 67 Smith, E. and Shilatifard, A. (2014) Enhancer biology and enhanceropathies. *Nat. Struct. Mol. Biol.* **21**, 210–219, <https://doi.org/10.1038/nsmb.2784>
- 68 Turner, T.N. and Eichler, E.E. (2019) The role of de novo noncoding regulatory mutations in neurodevelopmental disorders. *Trends Neurosci.* **42**, 115–127, <https://linkinghub.elsevier.com/retrieve/pii/S0166223618302960>, <https://doi.org/10.1016/j.tins.2018.11.002>
- 69 Cassinari, K., Rovelet-Lecrux, A., Tury, S., Quenez, O., Richard, A.C., Charbonnier, C. et al. (2020) Haploinsufficiency of the primary familial brain calcification gene SLC20A2 mediated by disruption of a regulatory element. *Mov. Disord.* **35**, 1336–1345, <https://doi.org/10.1002/mds.28090>
- 70 Macdonald, R.L., Kang, J.Q. and Gallagher, M.J. (2010) Mutations in GABAA receptor subunits associated with genetic epilepsies. *J. Physiol.* **588**, 1861–1869, <http://pmc/articles/PMC2901974/?report=abstract>
- 71 Monlong, J., Girard, S.L., Meloche, C., Cadieux-Dion, M., Andrade, D.M., Lafreniere, R.G. et al. (2018) Global characterization of copy number variants in epilepsy patients from whole genome sequencing. *PLoS Genet.* **14** (4), e1007285, <http://pmc/articles/PMC5978987/?report=abstract>, <https://doi.org/10.1371/journal.pgen.1007285>
- 72 Short, P.J., McRae, J.F., Gallone, G., Sifrim, A., Won, H., Geschwind, D.H. et al. (2018) De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature* **555**, 611–616, <http://pmc/articles/PMC5912909/?report=abstract>, <https://doi.org/10.1038/nature25983>
- 73 Doan, R.N., Bae, B.I., Cubelos, B., Chang, C., Hossain, A.A., Al-Saad, S. et al. (2016) Mutations in human accelerated regions disrupt cognition and social behavior. *Cell* **167**, 341.e12–354.e12, <http://pmc/articles/PMC5063026/?report=abstract>, <https://doi.org/10.1016/j.cell.2016.08.071>
- 74 Won, H., Huang, J., Opland, C.K., Hartl, C.L. and Geschwind, D.H. (2019) Human evolved regulatory elements modulate genes involved in cortical expansion and neurodevelopmental disease susceptibility. *Nat. Commun.* **10**, 2396, <http://pmc/articles/PMC6546784/?report=abstract>, <https://doi.org/10.1038/s41467-019-10248-3>
- 75 Won, H., De La Torre-Ubieta, L., Stein, J.L., PARIKSHAK, N.N., Huang, J., Opland, C.K. et al. (2016) Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527, <http://pmc/articles/PMC5358922/?report=abstract>, <https://doi.org/10.1038/nature19847>
- 76 Melo, U.S., Schöpflin, R., Acuna-Hidalgo, R., Mensah, M.A., Fischer-Zirnsak, B., Holtgrewe, M. et al. (2020) Hi-C identifies complex genomic rearrangements and TAD-shuffling in developmental diseases. *Am. J. Hum. Genet.* **106**, 872–884, <http://pmc/articles/PMC7273525/?report=abstract>, <https://doi.org/10.1016/j.ajhg.2020.04.016>
- 77 Allou, L., Lambert, L., Amsallem, D., Bieth, E., Ederly, P., Destrée, A. et al. (2012) 14q12 and severe Rett-like phenotypes: New clinical insights and physical mapping of FOXP1-regulatory elements. *Eur. J. Hum. Genet.* **20**, 1216–1223, <http://pmc/articles/PMC3499785/?report=abstract>, <https://doi.org/10.1038/ejhg.2012.127>
- 78 Spielmann, M., Lupiáñez, D.G. and Mundlos, S. (2018) Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467, www.nature.com/nrg, <https://doi.org/10.1038/s41576-018-0007-0>
- 79 Redin, C., Brand, H., Collins, R.L., Kammin, T., Mitchell, E., Hodge, J.C. et al. (2017) The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nat. Genet.* **49**, 36–45, <https://pubmed.ncbi.nlm.nih.gov/27841880/>, <https://doi.org/10.1038/ng.3720>
- 80 Cosgrove, D., Whitton, L., Fahey, L., Broin, P.O., Donohoe, G. and Morris, D.W. (2020) Genes influenced by MEF2C contribute to neurodevelopmental disease via gene expression changes that affect multiple types of cortical excitatory neurons. *Hum. Mol. Genet.* **30**, 961–970, <https://pubmed.ncbi.nlm.nih.gov/32975584/>, <https://doi.org/10.1093/hmg/ddaa213>
- 81 D'Haene, E., Bar-Yaacov, R., Bariah, I., Vantomme, L., Van Loo, S., Cobos, F.A. et al. (2019) A neuronal enhancer network upstream of MEF2C is compromised in patients with Rett-like characteristics. *Hum. Mol. Genet.* **28**, 818–827, <http://pmc/articles/PMC6381311/>, <https://doi.org/10.1093/hmg/ddy393>

- 82 Yauy, K., Schneider, A., Ng, B.L., Gaillard, J.B., Sati, S., Coubes, C. et al. (2019) Disruption of chromatin organisation causes MEF2C gene overexpression in intellectual disability: a case report. *BMC Med. Genomics* **12**, 116, <https://pubmed.ncbi.nlm.nih.gov/31375103/>, <https://doi.org/10.1186/s12920-019-0558-8>
- 83 Bai, M., Ye, D., Guo, X., Xi, J., Liu, N., Wu, Y. et al. (2020) Critical regulation of a NDIME / MEF 2C axis in embryonic stem cell neural differentiation and autism. *EMBO Rep.* **21**, e50283, <https://pubmed.ncbi.nlm.nih.gov/33016573/>, <https://doi.org/10.15252/embr.202050283>
- 84 Gasperini, M., Tome, J.M. and Shendure, J. (2020) Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat. Rev. Genet.* **21**, 292–310, www.nature.com/nrg, <https://doi.org/10.1038/s41576-019-0209-0>
- 85 Forrest, M.P., Zhang, H., Moy, W., McGowan, H., Leites, C., Dionisio, L.E. et al. (2017) Open chromatin profiling in hiPSC-derived neurons prioritizes functional noncoding psychiatric risk variants and highlights neurodevelopmental loci. *Cell Stem Cell* **21**, 305.e8–318.e8, <http://pmc/articles/PMC5591074/?report=abstract>, <https://doi.org/10.1016/j.stem.2017.07.008>
- 86 Song, M., Yang, X., Ren, X., Maliskova, L., Li, B., Jones, I.R. et al. (2019) Mapping cis-regulatory chromatin contacts in neural cells links neuropsychiatric disorder risk variants to target genes. *Nat. Genet.* **51**, 1252–1262, <http://pmc/articles/PMC6677164/?report=abstract>, <https://doi.org/10.1038/s41588-019-0472-1>
- 87 Nott, A., Holtman, I.R., Coufal, N.G., Schlachetzki, J.C.M., Yu, M., Hu, R. et al. (2019) Brain cell type-specific enhancer-promoter interactome maps and disease risk association. *Science* **0793**, 1–12, <https://doi.org/10.1126/science.aay0793>
- 88 Chandra, V., Bhattacharyya, S., Schmiedel, B.J., Madrigal, A., Gonzalez-Colin, C., Fotsing, S. et al. (2021) Promoter-interacting expression quantitative trait loci are enriched for functional genetic variants. *Nat. Genet.* **53**, 110–119, <https://pubmed.ncbi.nlm.nih.gov/33349701/>, <https://doi.org/10.1038/s41588-020-00745-3>
- 89 Wang, X. and Goldstein, D.B. (2020) Enhancer domains predict gene pathogenicity and inform gene discovery in complex disease. *Am. J. Hum. Genet.* **106**, 215–233, <https://linkinghub.elsevier.com/retrieve/pii/S0002929720300124>, <https://doi.org/10.1016/j.ajhg.2020.01.012>
- 90 Nica, A.C. and Dermitzakis, E.T. (2013) Expression quantitative trait loci: present and future. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20120362, <https://doi.org/10.1098/rstb.2012.0362>
- 91 Zhang, S., Zhang, H., Zhou, Y., Qiao, M., Zhao, S., Kozlova, A. et al. (2020) Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. *Science (80-)* **369**, 561–565, <http://pmc/articles/PMC7773145/>, <https://doi.org/10.1126/science.aay3983>
- 92 Li, H., Yang, Y., Hong, W., Huang, M., Wu, M. and Zhao, X. (2020) Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal. Transduct. Target Ther.* **5**, 1, <https://doi.org/10.1038/s41392-019-0089-y>
- 93 Boudreaux, S.P., Ramirez-Herrick, A.M., Duren, R.P. and Conneely, O.M. (2012) Genome-wide profiling reveals transcriptional repression of MYC as a core component of NR4A tumor suppression in acute myeloid leukemia. *Oncogenesis* **1**, 19, <https://doi.org/10.1038/oncsis.2012.19>
- 94 Silberstein, S.D., Shrewsbury, S.B. and Hoekman, J. (2020) Dihydroergotamine (DHE) - then and now: a narrative review. *Headache* **60**, 40–57, <https://doi.org/10.1111/head.13700>
- 95 Call, S.G., Duren, R.P., Panigrahi, A.K., Nguyen, L., Freire, P.R., Grimm, S.L. et al. (2020) Targeting oncogenic super enhancers in MYC-dependent AML using a small molecule activator of NR4A nuclear receptors. *Sci. Rep.* **10**, 1–15, <https://doi.org/10.1038/s41598-020-59469-3>
- 96 Lykken, E.A., Shyng, C., Edwards, R.J., Rozenberg, A. and Gray, S.J. (2018) Recent progress and considerations for AAV gene therapies targeting the central nervous system. *J. Neurodev. Disord.* **10**, 16, <https://neurodevdisorders.biomedcentral.com/articles/10.1186/s11689-018-9234-0>, <https://doi.org/10.1186/s11689-018-9234-0>
- 97 Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N. et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (80-)* **339**, 819–823, <https://doi.org/10.1126/science.1231143>
- 98 Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A. et al. (2019) CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science (80-)* **363**, eaau0629, <https://doi.org/10.1126/science.aau0629>
- 99 Colasante, G., Lignani, G., Brusco, S., Di Bernardino, C., Carpenter, J., Giannelli, S. et al. (2020) dCas9-based Scn1a gene activation restores inhibitory interneuron excitability and attenuates seizures in Dravet syndrome mice. *Mol. Ther.* **28**, 235–253, <https://linkinghub.elsevier.com/retrieve/pii/S1525001619304010>, <https://doi.org/10.1016/j.ymthe.2019.08.018>
- 100 Parsi, K.M., Hennessy, E., Kearns, N. and Maehr, R. (2017) Using an inducible CRISPR-dCas9-KRAB effector system to dissect transcriptional regulation in human embryonic stem cells. *Methods Mol. Biol.* **1507**, 221–233, https://link.springer.com/protocol/10.1007/978-1-4939-6518-2_16, https://doi.org/10.1007/978-1-4939-6518-2_16
- 101 Perenthaler, E., Nikoncuk, A., Yousefi, S., Berdowski, W.M., Alsagob, M., Capó, I. et al. (2019) Loss of UGP2 in brain leads to a severe epileptic encephalopathy, emphasizing that bi-allelic isoform-specific start-loss mutations of essential genes can cause genetic diseases. *Acta Neuropathol* **139**, 415–442, <http://link.springer.com/10.1007/s00401-019-02109-6>, <https://doi.org/10.1007/s00401-019-02109-6>
- 102 Frangoul, H., Altshuler, D., Cappellini, M.D., Chen, Y.-S., Domm, J., Eustace, B.K. et al. (2021) CRISPR-Cas9 gene editing for sickle cell disease and β -thalassemia. *N. Engl. J. Med.* **384**, 252–260, <https://doi.org/10.1056/NEJMoa2031054>
- 103 Motta, I., Bou-Fakhredin, R., Taher, A.T. and Cappellini, M.D. (2020) Beta thalassemia: new therapeutic options beyond transfusion and iron chelation. *Drugs* **80**, 1053–1063, <https://doi.org/10.1007/s40265-020-01341-9>
- 104 Gardner, R.V. (2018) Sickle cell disease: advances in treatment. *Ochsner J.* **18**, 377–389, <https://doi.org/10.31486/toj.18.0076>
- 105 Ikawa, Y., Miccio, A., Magrin, E., Kwiatkowski, J.L., Rivella, S. and Cavazzana, M. (2019) Gene therapy of hemoglobinopathies: progress and future challenges. *Hum. Mol. Genet.* **28**, R24–R30, <https://doi.org/10.1093/hmg/ddz172>