Received: 16 January 2023 | Revised: 24 May 2023 | Accepted: 6 June 2023

DOI: 10.1002/bies.202300010

BioEssays

REVIEW ESSAY

Prospects & Overviews

Disruption of regulatory domains and novel transcripts as disease-causing mechanisms

Lila Allou[1,2] | **Stefan Mundlos**[1,2,3]

[1]RG Development & Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany

[2]Institute for Medical and Human Genetics, Charité-Universitätsmedizin Berlin, Berlin, Germany

[3]Berlin-Brandenburg Center for Regenerative Therapies, Charité-Universitätsmedizin Berlin, Berlin, Germany

Correspondence
Lila Allou and Stefan Mundlos, RG Development & Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany.
Email: allou@molgen.mpg.de; mundlos@molgen.mpg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: MU 880/16-1; FOXG1 Research Foundation, Grant/Award Number: P.S.7905

Abstract

Deletions, duplications, insertions, inversions, and translocations, collectively called structural variations (SVs), affect more base pairs of the genome than any other sequence variant. The recent technological advancements in genome sequencing have enabled the discovery of tens of thousands of SVs per human genome. These SVs primarily affect non-coding DNA sequences, but the difficulties in interpreting their impact limit our understanding of human disease etiology. The functional annotation of non-coding DNA sequences and methodologies to characterize their three-dimensional (3D) organization in the nucleus have greatly expanded our understanding of the basic mechanisms underlying gene regulation, thereby improving the interpretation of SVs for their pathogenic impact. Here, we discuss the various mechanisms by which SVs can result in altered gene regulation and how these mechanisms can result in rare genetic disorders. Beyond changing gene expression, SVs can produce novel gene-intergenic fusion transcripts at the SV breakpoints.

KEYWORDS
3D genome organization, enhancer-promoter communication, gene-intergenic fusion transcript, structural variation, topologically associated domain

INTRODUCTION

Deletions, duplications, insertions, inversions, and translocations, collectively called structural variations (SVs), contribute significantly to human genome variability.[1–4] Pathogenic SVs have been involved in many conditions,[5] including developmental and sensory disorders, and are increasingly recognized as important causes of inherited diseases.[6] They can lead to copy number changes (gains and losses of DNA), disruption of protein-coding genes, enhancer-promoter communication rewiring, and production of fusion genes.

With the emergence of novel sequencing technologies, SV identification has significantly improved.[7] On average, over 27,000 SVs are identified per human genome.[2] The finding that SVs can alter the communication between regulatory elements and genes by disrupting topologically associated domains (TADs) has significantly improved our interpretation of the pathogenic impact of SVs.[5] Nevertheless, the medical interpretation of SVs and the prediction of their functional consequences, as well as dissecting the mechanisms by which they exert their functional impact, still need to be improved. Indeed, the vast majority of SVs occur outside the protein-coding portion of the genome.[2] Thus, the incomplete functional annotation of non-coding DNA sequences and the complexity of gene regulatory landscapes

Abbreviations: 3D, three-dimensional; Array CGH, array comparative genomic hybridization; Bp, base pairs; CBS, CTCF binding site; CGR, complex genomic rearrangement; CNV, copy number variant; CRE, *cis*-regulatory element; CTCF, CCCTC-binding factor; DHMN1, distal hereditary motor neuropathy; hiPSCs, human induced pluripotent stem cells; HPP, human proteome project; Indels, small insertions and deletions; Kb, kilobases; LncRNA, long non-coding RNA; Mb, megabases; RNA-seq, RNA-sequencing; SNV, single-nucleotide variant; SV, structural variation; TAD, topologically associated domain; WGS, whole genome sequencing.

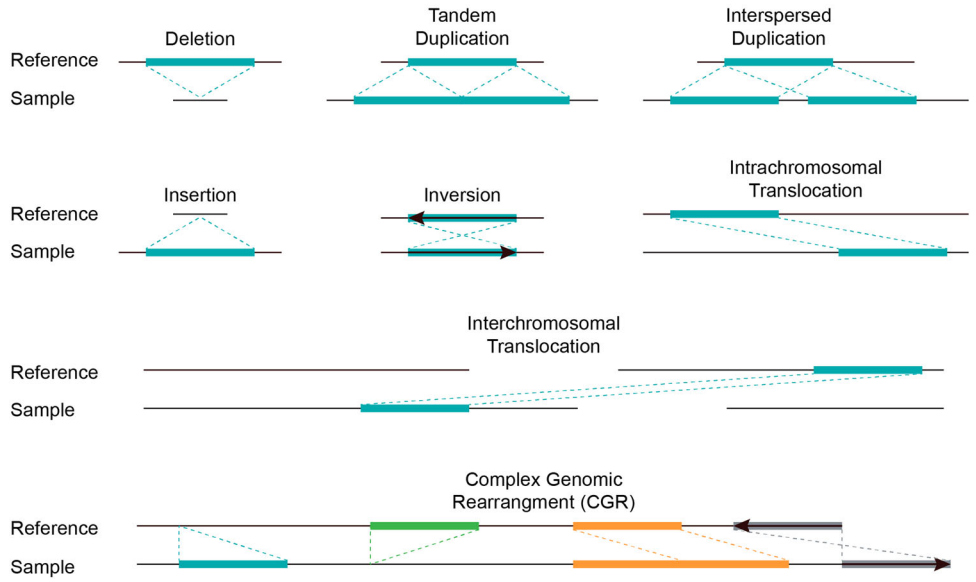

FIGURE 1 Types of SVs. Schematic representations of the different structural variants discussed in this review: deletions, duplications, insertions, inversions, translocations, and CGRs. "Reference" indicates the initial DNA sequence, and "Sample" indicates the sequence generated by each SV category. The sequence change is indicated using dashed lines. The black arrow indicates the DNA sequence orientation. The represented CGR harbors an insertion (blue), a deletion (green), a tandem duplication (orange), and an inversion (gray) simultaneously.

severely hamper interpreting their pathogenic impact. Moreover, the recent advances in gene regulation and 3D genome organization indicate that additional regulatory layers (e.g., enhancer-promoter specificity) control the communication between genes and *cis*-regulatory elements (CREs) beyond TADs.[8] These additional regulatory layers still need to be fully understood to fully predict the functional consequences of SVs with respect to gene regulation and human disease etiology. Additionally, even a single SV can have a vast array of consequences with simultaneous copy number changes, rewiring of TAD 3D architecture, and generation of novel fusion transcripts, which makes them harder to interpret.[9] Finally, the incomplete functional annotation of human genes further complicates this.[10]

In this review, we focus on rearrangements rather than aneuploidies and use the term "SV" to represent deletions, duplications, insertions, inversions, chromosomal translocations, and any combination of these events. We also exclusively focus on germline SVs. We first discuss how these SVs can alter gene expression in *cis*. We then discuss a new mechanism through which SVs impacting non-coding DNA sequences exert their pathogenic effect by producing a new gene-intergenic fusion transcript. Finally, we also discuss the interpretation of SV impact in light of the latest advances in gene regulation and 3D genome organization.

Structural variation

Human genetic variations include single-nucleotide variants (SNVs), small insertions and deletions (indels; <50 bp), and SVs.[1] SVs can broadly differ in size, ranging from ~50 bp to several megabases, affecting more of the genome per nucleotide changes than any other sequence variant.[4,11–15] They are highly diverse in type (Figure 1), including deletions, duplications, and insertions of genetic material, known as unbalanced copy number variants (CNVs). They also include inversions and interchromosomal and intrachromosomal translocations, known as balanced rearrangements. Unbalanced CNVs change the dosage of DNA, while balanced rearrangements do not change the DNA dosage. In addition to these described events, SVs include mobile element insertions (MEIs), tandem repeats, multiallelic CNVs of highly variable copy numbers, segmental duplications, complex genomic rearrangements (CGRs), as well as aneusomy and aneuploidy.[16–20]

SVs arise by homologous recombination (HR) between repeated sequences (recurrent SVs) or by non-homologous mechanisms that occur throughout the genome (non-recurrent SVs).[21] Indeed, SVs are usually caused by erroneous DNA replication and damage repair, as well as the activities of repetitive elements. Several mechanisms are known to form SVs in germline and somatic cells. Such mechanisms include non-allelic homologous recombination (NAHR), non-homologous end-joining (NHEJ), double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), break-induced replication (BIR), single-strand annealing (SSA), microhomology-mediated end joining (MMEJ), breakage-fusion-bridge (BFB) cycle, replication slippage, fork stalling and template switching (FoSTeS), microhomology-mediated break-induced replication (MMBIR), and so on. We direct the reader to some excellent reviews for more details on the mechanisms underlying SV formation.[22–24]

Various experimental techniques have been developed to identify SVs. Karyotyping was routinely performed to diagnose and screen genetic diseases.[25] Nowadays, it is rarely used for SV discovery due to its low sensitivity and precision. Instead, array comparative genomic hybridization (array CGH) is commonly used in routine clinical diagnostics for genetic disorders.[26,27] However, this technique

has limitations, including its relatively low resolution, the inability to detect balanced rearrangements, and the low efficacy in detecting mosaicisms.[28] Due to its limitations, array CGH is now mainly replaced by second-generation short-read sequencing technologies, which became available more than 10 years ago. However, due to the short-read length limitation and the human genome's repetitive nature, SVs in repetitive and segmental duplication regions remain difficult to identify.[29] The third-generation long-read sequencing and imaging technologies can overcome the limitations of second-generation short-read sequencing techniques.[30,31] We refer the reader to another review for an in-depth discussion of strategies and algorithms for SV detection.[7]

Current estimates suggest that a human genome may harbor >27,000 SVs compared to the reference genome.[2] These SVs primarily impact non-coding DNA sequences.[2] Therefore, completing the functional annotation of non-coding regions and understanding the basic principles underlying gene regulation and the genome's spatial 3D organization is crucial to comprehend their pathogenic potential fully. In the following section, we provide a brief overview of the current concepts regarding the hierarchical structures of the 3D genome. For more details on the organization and function of the 3D genome and the methods for understanding how chromatin is folded in the nucleus, we refer the reader to some excellent reviews.[32–36]

3D genome organization

Cytological and microscopy techniques revealed that individual chromosomes occupy separate chromosome territories in the nucleus during interphase.[37] Chromosome conformation capture (3C) technologies such as HiC rely on quantifying interaction frequencies between loci that lie in close spatial proximity in the nucleus independently of their linear genomic distance.[38–43] HiC technology uncovered that the genome is hierarchically organized into multiple layers in the nucleus (Figure 2). On a course scale, the genome is compartmentalized into A and B compartments corresponding to the spatial segregation of open and closed chromatin.[42,44] The A-compartment largely correlates with accessible, transcriptionally active euchromatin. It is mainly present in the interior nuclear space. Conversely, the B-compartment largely correlates with compacted, transcriptionally silent heterochromatin predominantly located in the periphery of the nucleus as lamina-associated domains (LADs) or nucleoli as nucleolar-associated domains (NADs).[45,46] The separation of A and B compartments has been proposed to reflect phase separation between active and inactive chromatin based on their inherent physical properties.[47]

Compartments are subdivided into regions called TADs.[48–50] TADs are megabase-long genomic regions that interact with themselves at a higher frequency than with the rest of the genome. The regions preventing contact between neighboring TADs are called boundaries or borders (Figure 2). Boundaries are mainly formed by the site-directed divergent binding of the CCCTC-binding factor (CTCF). TAD formation and function are mainly achieved by the combined action of the cohesin

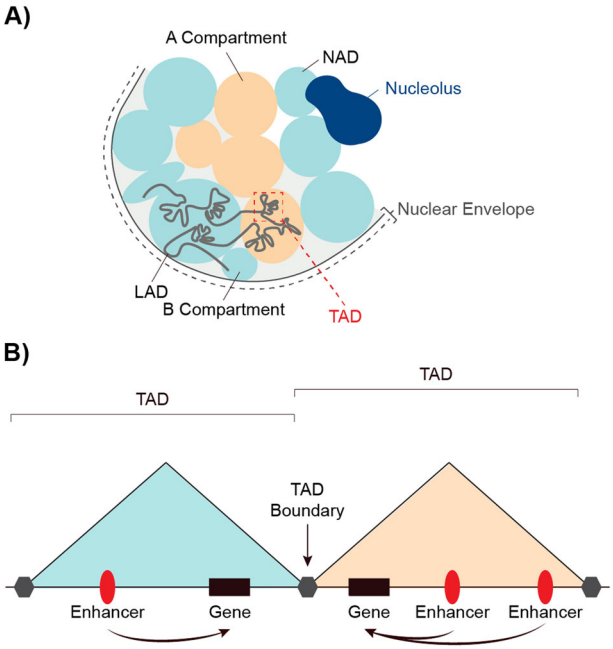

FIGURE 2 3D hierarchical organization of the genome. (A) Schematic representation of A and B compartments. At higher-order scales, chromatin with a transcriptionally active or repressive signature separates into A (orange) and B (blue) compartments, respectively. The A compartments coincide with non-lamina-associated domains (non-LADs) and are mainly present in the interior nuclear space. The B compartments frequently overlap with nucleolar-associated domains (NADs) and LADs. (B) Schematic representation of the HiC view of TADs. At smaller scales, enhancers (red ellipses) transmit regulatory information to genes (black boxes) by physical proximity within, but not between, TADs. Separated by insulating boundaries, TADs preferentially internally self-associate to create discrete functional and structural blocks.

complex and CTCF through a process called loop extrusion.[51,52] During this process, cohesin is loaded into chromatin and progressively extrudes a loop until a roadblock factor stalls it or the cohesin unloading factor Wapl (Wings apart-like protein homolog) releases it.[53–60] Although transcription start sites (TSSs) and transcription termination sites (TTSs), as well as other proteins such as RNA polymerase II, might act as cohesin barriers,[60,61] the most common cohesin roadblock factor is formed by convergent CTCF binding sites (CBSs) in vertebrates.[56,62]

TADs constrain the genomic regions an enhancer can act upon,[63–66] and most enhancer-gene interactions occur within TADs through chromatin loops that connect distant DNA fragments.[48,67] Thus, TADs represent fundamental regulatory units facilitating the communication between enhancers and their target genes in the same domain. At the same time, TAD boundaries act as insulators that prevent enhancers from contacting non-target genes in other domains. Indeed, several observations support the relevance of TADs for enhancer function. For example, enhancers and their target genes usually lie within the same TAD,[48,64,67] and disruption of TAD boundaries by SVs or mutations in CBSs can result in novel enhancer-promoter combinations and ectopic gains in gene expression.[63,68–71]

However, this does not seem to apply to all genome regions. First, there is a significant fraction of the genome wherein TADs cannot be clearly detected, such as near *Wnt6* (*Wingless-type MMTV integration site family, member 6*), which is located in a gene-dense region composed of weakly defined, less insulated structures.[48,72] Second, TAD fusion was observed following boundary elimination at the *Sox9-Kcnj2* locus, and TADs loss was observed following CTCF or cohesin depletion. However, this did not cause significant gene expression changes.[8,54,57] Third, using mouse embryonic stem cells (mESCs), a recent study found that enhancer-promoter interactions and transcription are largely insensitive to acute depletion of CTCF, cohesin, WAPL, or YY1 (Ying Yang 1).[73] Fourth, genes within a TAD are not always co-regulated, and the degree of promoter responsiveness to enhancers can vary based on position within a TAD.[66,74–76] These data and others (see section "*Enhancer adoption: rare or overlooked?*") indicate that additional regulatory layers control the responsiveness between genes and enhancers beyond inclusion in the same domain. Overall, this delineates the organizational and functional complexity of the gene regulatory landscapes, hampering the interpretation of SVs impacting non-coding DNA sequences and, thus, our understanding of human disease etiology.

SVs ALTERING GENE REGULATION IN *CIS*

The most studied and best-understood mechanisms through which SVs can cause disease are those directly affecting coding sequences (i.e., disrupting genes, changing gene dosage, and creating fused genes). Nevertheless, the latest technological advances allowed us to decipher new complex mechanisms through which SVs can cause disease. Indeed, detailed studies demonstrated that SVs can exert their pathogenic effect by rewiring 3D genome organization and enhancer-gene communication. Following, we will summarize the mechanisms by which SVs can alter gene regulation in *cis*, considering 3D chromatin architecture and TAD organization.

Leaving the TAD structure intact

SVs that occur within TADs and affect CREs can result in a tissue-specific loss of function or gain of function of their endogenous target genes, which can be located several hundred kilobases away[77–84] (Figure 3A–C). These SVs usually cause phenotypic features resembling selective phenotypic sub-features of coding mutations. For example, the *engrailed-1* (*En1*) gene loss of function mutations cause severe limb and brain malformations. In contrast, intra-TAD homozygous deletions encompassing its limb-specific CRE, the long non-coding RNA (lncRNA) locus *Maenli* (Master activator of *engrailed-1* in the limb), cause a partial phenotype restricted to the limb[85] (Figure 3D). Of note, duplications occurring within TADs and affecting CREs can also result in tissue-specific misexpression of their endogenous target genes and disease[77,79,86–89] (Figure 3C). This is, for example, the case for duplications involving regulatory elements within the *Ihh* (*Indian hedgehog*) regulatory domain. These duplications are associated with craniosynostosis and synpolydactyly. In this case, the craniosynostosis strictly associates with *Ihh* overexpression, whereas the observed synpolydactyly results from *Ihh* tissue-specific misexpression.[89]

One major prerequisite to interpreting and deciphering the pathogenicity of intra-TAD SVs is a complete functional annotation of non-coding DNA sequences. For example, the *Maenli* lncRNA locus was not annotated as an *En1* limb-specific CRE.[85] This hampered the clinical interpretation of the pathogenic impact of the 28-kb non-coding homozygous deletions identified in the vicinity of the *En1* gene, especially since *En1* was not involved in the disease until then.[85] Another example is the deletions/duplications identified downstream of the *SHOX* (*Short Stature Homeobox*) gene. CNVs affecting one or more of the seven *SHOX* annotated non-coding regulatory elements represent one of the most frequent causes of *SHOX*-haploinsufficiency.[90] Recently, two partially overlapping small deletions of ~12- and ~8-kb, approximately 15-kb downstream *SHOX*, were identified in four unrelated individuals with clinical features strongly suggestive of *SHOX*-haploinsufficiency. However, these deletions do not include previously annotated *SHOX* regulatory elements, suggesting they may affect yet unidentified *SHOX*-specific CREs. Thus, the incomplete functional annotation of *SHOX*-specific CREs hampers fully comprehending the pathogenic potential of the identified CNVs.[90]

Moreover, it is essential to consider that tissue-specific transcription is generally controlled by multiple CREs that work totally or partially redundant to achieve their target genes' precise tissue-specific expression patterns. This confers phenotypic robustness to loss of function mutations in individual enhancers.[91,92]

Creating a new gene-regulatory landscape

Enhancer adoption

SVs occurring between TADs can rearrange the 3D chromatin organization of a locus (Figure 4A–D). Indeed, deletions encompassing TAD boundaries will fuse the two adjacent TADs (TAD fusion) (Figure 4B).[63,93–96] Duplications that include TAD boundaries will create new TADs ("neo-TAD" formation) (Figure 4C),[68,71,94,96–101] and inversions encompassing TAD boundaries will result in the exchange of regulatory material between TADs (TAD shuffling) (Figure 4D). Finally, chromosomal translocations can result in the fusion of TADs (TAD fusion), or relocate TAD boundaries, thereby swapping regulatory material between TADs (TAD shuffling).[63,100–102] 3D chromatin rewiring may lead to gene misexpression through a process termed enhancer adoption or enhancer "hijacking," in which ectopic expression of a gene is driven by an enhancer that normally regulates another gene located in a different regulatory domain. Thus, inter-TAD SVs can create a new gene-regulatory landscape, causing gene misexpression and disease. Several examples have been reported in the literature, such as deletions at the *Epha4-Pax3* locus. These deletions rewire 3D chromatin organization at the locus (TAD fusion) and bring enhancers from the *Epha4* (*Ephrin type-A receptor 4*) TAD to the vicinity of *Pax3*

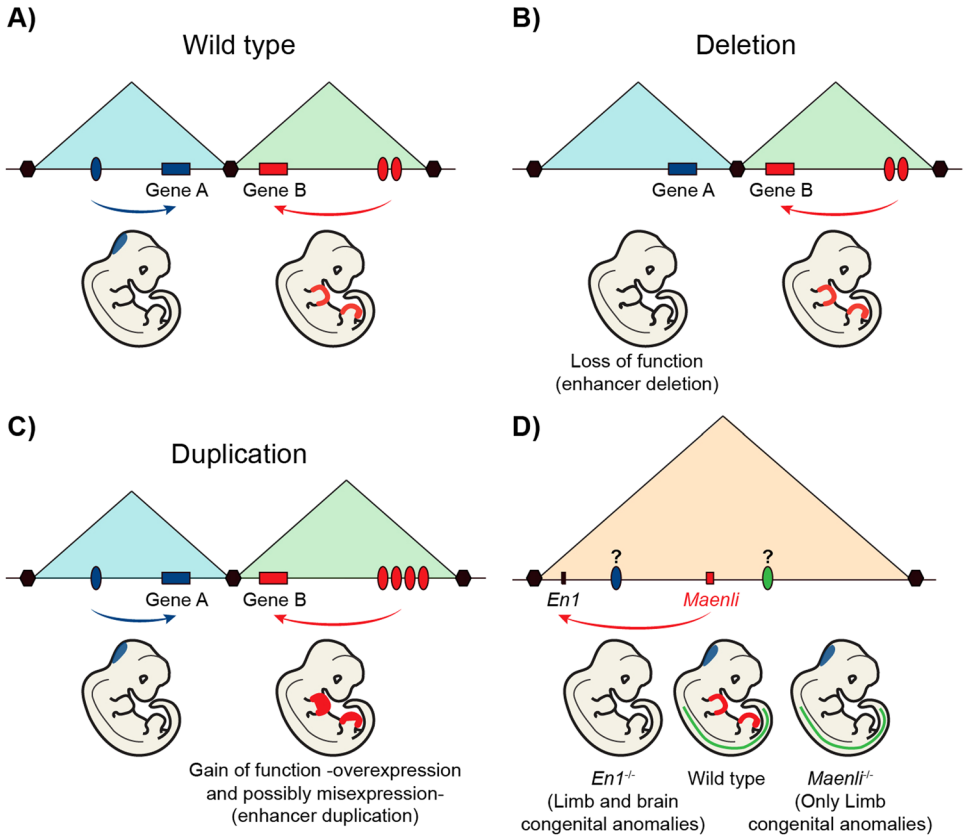

FIGURE 3 SVs leaving TADs' structure intact. Structural variants that occur within TADs (intra-TAD) do not alter the TAD 3D architecture. Nevertheless, they can change enhancer dosage, resulting in loss or gain of function of their endogenous target gene. (A) Schematic representation of the wild-type genomic locus showing gene A expressed in the developing brain (blue) under the control of its tissue-specific enhancer (blue ellipse) and gene B expressed in the developing limb bud (red) under the control of its tissue-specific enhancers (red ellipses). Genes A and B and their respective tissue-specific enhancers are located in two TADs (blue and green triangles) separated by a TAD boundary (black polygon). (B) The intra-TAD deletion of an enhancer element can result in tissue-specific loss of function of gene A in the brain (blue). (C) The intra-TAD duplication of an enhancer element can result in tissue-specific overexpression and possibly misexpression of gene B in the limb (red). (D) In the wild-type embryo, *En1* is expressed in the brain (blue), limb (red), and somites (green). The *En1* knockout ($En1^{-/-}$) results in a complete loss of *En1* expression in all three tissues and severe congenital limb and brain malformations. The *Maenli* lncRNA locus controls *En1* expression, specifically in the limb. Its loss ($Maenli^{-/-}$) results in a complete loss of *En1* expression, specifically in the limb at developmental stage E9.5, and severe congenital limb malformations. The question mark (?) indicates that the enhancers controlling *En1* expression in the brain and the somites are not yet characterized.

(*Paired box 3*). In this case, *Pax3* is under the control of *Epha4* enhancers leading to a pathological gain of *Pax3* expression in an *Epha4*-like pattern and disease.[63] Another example is the duplications at the *Sox9-Kcnj2* locus. In this case, a region including the *Kcnj2* (*Potassium inwardly-rectifying channel, subfamily J, member 2*) gene and *Sox9* (*SRY (sex determining region Y)-box 9*) enhancers, normally located in two distinct TADs, is duplicated, resulting in the formation of a new TAD ("neo-TAD"), in which *Kcnj2* is under the control of the duplicated *Sox9* enhancers leading to gene misexpression and disease.[68] Moreover, translocations involving the *PAX3* (*Paired Box 3*) locus on chromosome 2 and the *FOXO1* (*Forkhead Box O1*) locus on chromosome 13, bringing together the *PAX3* and *FOXO1* genes, have been recently reported to rewire 3D chromatin organization (TAD shuffling). The *PAX3* promoter of the fused *PAX3-FOXO1* gene is thus controlled by *FOXO1* regulatory elements, causing *PAX3-FOXO1* oncogene activation in non-*PAX3* territories and disease.[102]

Enhancer disconnection

SVs occurring between TADs can also rearrange the 3D chromatin organization of a locus without resulting in enhancer adoption. This is, for example, the case for an 89-Mb heterozygous inversion on chromosome 6, causing branchio-oculo-facial syndrome (BOFS).[103] Using patient-specific human induced pluripotent stem cells (hiPSCs), the authors demonstrated that the inversion shuffles enhancers controlling *TFAP2A* (*Transcription Factor AP-2 Alpha*) gene expression with novel genes within the same TAD. However, this 3D chromatin rewiring does not result in enhancer adoption. In this case, the disease is caused by a loss of *TFAP2A* expression due to the disconnection of the gene from its cognate enhancers.[103] Thus, 3D chromatin rewiring by SVs may lead to gene loss of function through enhancer disconnection (Figure 4D). This mechanism is increasingly recognized as the cause of developmental disorders. For example, using whole genome

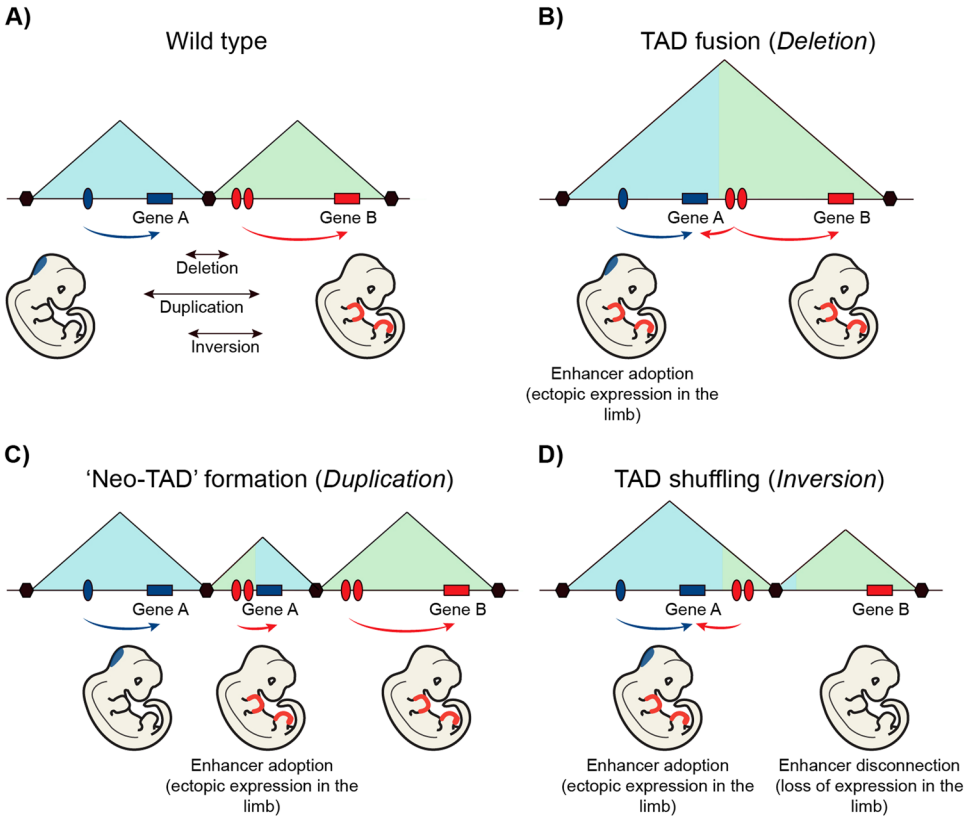

FIGURE 4 SVs creating a new gene-regulatory landscape. Structural variants that occur between TADs can rewire the TAD 3D architecture causing enhancer adoption and/or enhancer disconnection, resulting in gene misexpression and/or gene loss of function and disease. (A) Schematic representation of the wild-type genomic locus showing gene A expressed in the developing brain (blue) under the control of its tissue-specific enhancer (blue ellipse) and gene B expressed in the developing limb bud (red) under the control of its tissue-specific enhancers (red ellipses). Genes A and B and their respective tissue-specific enhancers are located in two TADs (blue and green triangles) separated by a TAD boundary (black polygon). The inter-TAD rearrangements represented in (B), (C), and (D) are indicated using double arrows. (B) Deletions involving TAD boundaries can cause TAD fusion and enhancer adoption, in which gene A is under the control of gene B enhancers, resulting in gene A misexpression in the limb (red). Chromosomal translocations can also cause TAD fusion and enhancer adoption. (C) Duplications involving TAD boundaries can create a new chromatin domain ("Neo-TAD" formation) and cause enhancer adoption, in which gene A is under the control of gene B enhancers, resulting in gene A misexpression in the limb (red). (D) Inversions involving TAD boundaries can cause TAD shuffling. This can lead to enhancer adoption, in which gene A is under the control of gene B enhancers, resulting in gene A misexpression in the limb (red). It can also lead to enhancer disconnection, in which gene B is separated from its cognate enhancers, resulting in gene B loss of function in the limb. Chromosomal translocations relocating TAD boundaries can also cause TAD shuffling, resulting in enhancer adoption and/or enhancer disconnection.

sequencing (WGS) in patients with congenital anomalies, a study systematically mapped the breakpoints of 273 balanced translocations at nucleotide resolution. A >7% of the balanced translocations disrupted TADs encompassing known haploinsufficiency disease-related genes.[104]

Enhancer adoption: Rare or overlooked?

Several studies have reported enhancer adoption as a mechanism that results in the ectopic expression of a gene as the underlying molecular mechanism of rare genetic disorders.[63,68,96,97,99,101,105,106] Since its discovery, it has been tempting to speculate that this mechanism could explain a significant fraction of unsolved cases. Although it is currently difficult to accurately estimate the prevalence of this disease mechanism due to our limited understanding of the basic mechanisms underlying enhancer-promoter communication, several studies suggest that this mechanism may not be as frequent as expected. Indeed, a recent study reported that gene expression is generally not altered when shuffling or fusing TADs, indicating a certain genome robustness against these events and that enhancer adoption is rare.[107] This study used highly rearranged chromosomes spanning around 75% of the *Drosophila* genome, including eight large nested inversions, smaller inversions, and thousands of deletions. Another study analyzed 288,457 somatic SVs using WGS data from 2658 cancers across 38 tumor types to understand their effects across TADs. This study revealed that only 14% of the boundary deletions fusing TADs, resulted in a change in expression in nearby genes more than twofold.[108] Moreover, a recent study investigated 11 individuals with CGRs, including germline chromothripsis, by combining short- and long-read genome sequencing with HiC and phased RNA-sequencing (RNA-seq). This study reported that while the majority of breakpoints reshuffle

TADs, most genes located 200-kb around a breakpoint do not show significant expression changes.[109] These data suggest that the enhancer adoption mechanism may be a rare cause of disease.

INTERPRETING THE IMPACT OF SVs: 3D AND MORE

Overall, the discovery that SVs can rewire TAD 3D architecture and the communication between regulatory elements and genes without disrupting the gene or regulatory elements sequences has dramatically improved our interpretation of the pathogenic impact of SVs. Nevertheless, when assessing the pathogenic impact of SVs, it is important to consider additional critical factors. First, TADs are highly dynamic structures,[110–112] and TAD boundary positions are not strictly conserved among cell types.[113] Thus, cell-type-specific or tissue-specific HiC maps related to the diseased tissue should be considered. It is also essential to consider that TAD boundaries are not strictly impenetrable. Indeed, inter-TAD communication can also occur.[114–118] This is exemplified by the inter-TAD communication involving the *Lbx1* (*Ladybird homeobox 1*) and *Fgf8* (*Fibroblast growth factor 8*) regulatory domains, although *Lbx1* and *Fgf8* exhibit divergent expression patterns.[101] Interestingly, the communication between the two regulatory domains contributes at least partially to the disease mechanism underlying split-hand/foot malformation type 3 (SHFM3), associated with tandem duplications at the *FGF8* locus[101] (Figure 5).

When assessing the pathogenic impact of SVs, especially those disrupting TAD 3D architecture with TAD fusion, neo-TAD formation, or TAD shuffling, it is important to consider the cell-type-specific and/or tissue-specific expression profiles of the genes at the rearranged loci. This is particularly true when the candidate genes are not known to cause disease and/or their function is still unknown. Of note, around 20% of human proteins lack physiologically informative descriptions,[10] which constitutes another major limiting factor in assessing the pathogenic impact of SVs and genetic variants as a whole. Moreover, it is essential to remember that even a single SV can have many consequences. For example, it has been shown that non-coding SVs can disrupt long-range gene regulation, resulting in a combination of both gene loss of function due to enhancer disconnection and gene misexpression due to enhancer adoption (Figure 4D). This is, for example, reported for the inversion involving the *Epha4* and *Wnt6* neighboring TADs causing congenital limb malformations.[63] This inversion causes ectopic interactions between *Wnt6* and the *Epha4* enhancers, resulting in *Wnt6* misexpression in the developing limb bud. At the same time, this inversion causes a disconnection between *Epha4* and its cognate enhancers, resulting in *Epha4* silencing in the developing limb bud.[63]

Additional regulatory layers contribute to the specific and functional communication between genes and CREs beyond being in the same TAD. Deciphering these additional regulatory layers and understanding their underlying molecular mechanisms is expected to improve our interpretation of SVs' pathogenic impact significantly. For instance, striking examples in the literature illustrate the paradox in promoter responsiveness to non-cognate enhancers. Among them are the *MEF2C* (*Myocyte Enhancer Factor 2C*) and *PAX3* promoters. Several translocations involving different chromosomes distal to *MEF2C* have been reported in the literature.[104,119–121] They all disrupt the *MEF2C* TAD 3D architecture creating a new *MEF2C*-regulatory

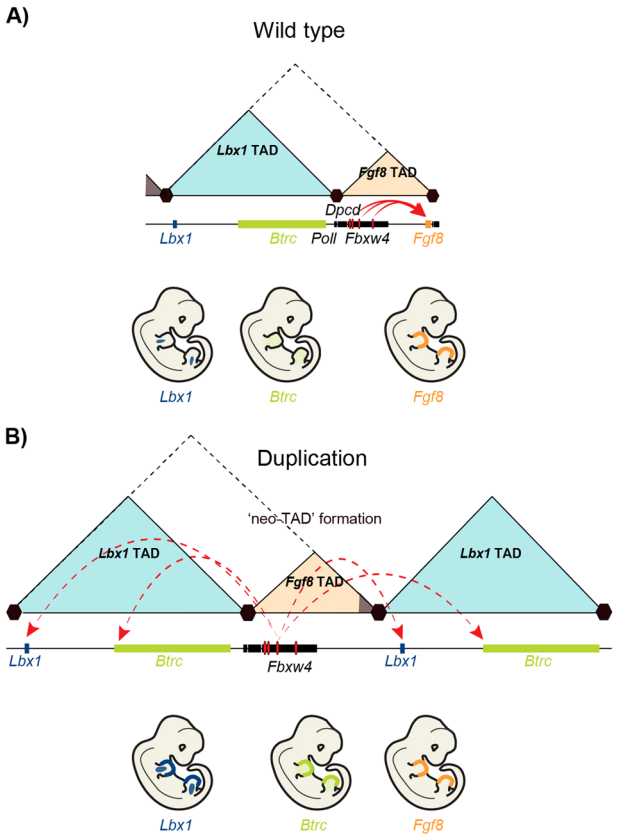

FIGURE 5 Inter-TAD communication contributes to the disease mechanism underlying SHFM3. (A) Schematic representation of the wild-type genomic locus showing the *Lbx1* TAD in blue and the *Fgf8* TAD in orange. The two TADs are separated by a TAD boundary (black polygon). A dashed triangle indicates communication between the *Lbx1* and *Fgf8* TADs. The red ellipses within the *Fbxw4* gene represent the *Fgf8* limb-specific enhancers (indicated with red arrows). The specific expression profiles of *Lbx1*, *Btrc*, and *Fgf8* during embryonic limb development are also shown. *Lbx1* is expressed in the muscle progenitor cells (blue), and *Fgf8* is expressed in the Apical Ectodermal Ridge (AER) (orange). *Btrc* is weakly expressed in the developing limb bud (light green). (B) Schematic representation of the impact of the tandem duplications causing SHFM3 on 3D genome organization and enhancer-promoter communication. The intact *Lbx1* TAD is duplicated, and a new TAD is generated between the two *Lbx1* TADs. This "neo-TAD" corresponds mainly to the *Fgf8* TAD (orange) and a duplicated genomic region centromeric of the *Lbx1* TAD boundary (brown). The dashed red arrows indicate the inter-TAD communication, resulting in ectopic interactions between the *Fgf8* limb-specific enhancers within the "neo-TAD" and the *Lbx1* and *Btrc* genes within one or both of the neighboring *Lbx1* TADs. This causes misexpression of *Lbx1* and *Btrc* in the developing limb bud in an *Fgf8*-like pattern, likely contributing to the disease. Of note, *Fgf8* expression is unaffected while *Lbx1* expression also increases in the muscle progenitor cells due to gene copy number change.

landscape specific to each rearrangement. Nevertheless, all reported SVs cause a similar haploinsufficiency phenotype due to haploinsufficient *MEF2C* expression. These data suggest that the *MEF2C* promoter is impervious to the enhancer activity of non-cognate CREs. Conversely, the *PAX3* promoter seems permeable to the enhancer activity of non-cognate enhancers. Indeed, deletions and translocations involving the *PAX3* locus lead to ectopic interactions between the *PAX3* promoter and non-cognate enhancers.[63,102] This results in gene misexpression or oncogene activation, causing disease. These examples suggest that understanding which types of promoters are permeable to the enhancer activity of non-cognate CREs and which are not, as well as whether this is absolute, would help improve our predictions of the impact of SVs on gene expression.

In addition to enhancers, several studies indicate that silencers are abundant in mammalian genomes and can repress gene expression by physically interacting with their target genes.[122–125] It is thus tempting to speculate that SVs disrupting silencer elements or silencer-gene communication can also cause disease. For example, a recent study reported a mechanism termed "silencer hijacking," whereby SVs relocate a repressive element (H3K27me3-rich region) close to the tumor suppressor gene *IKZF2* (*IKAROS Family Zinc Finger 2*), resulting in its expression downregulation in Leukemia.[126]

Taken together, the latest advances in understanding the basic principles underlying gene regulation and the genome's spatial organization highlight the challenging task of interpreting and predicting the pathogenic impact of SVs in clinical diagnostics. The newly reported mechanism through which SVs can cause disease, which we will discuss in the following section, undoubtedly corroborates this.

SVs CREATING A NOVEL GENE-INTERGENIC FUSION TRANSCRIPT

Somatic SVs are a common driver of fusion gene formation. Indeed, gene fusions have been extensively studied in cancer.[127–129] They typically involve one or more coding sequences resulting in chimeric mRNA transcripts that can form oncogenic "neo-antigens." Several SVs driving gene fusion formation have also been involved in rare genetic disorders, including congenital anomalies,[130–133] autism,[134–136] and intellectual disability.[137–142] Recently, a study described a complex SV driving a novel gene-intergenic fusion transcript formation as the underlying disease mechanism of an autosomal dominant distal hereditary motor neuropathy DHMN1[9] (Figure 6A). Using WGS, the authors identified a 1.35-Mb duplication of chromosome 7q36.3 inserted in the reverse orientation at chromosome 7q36.2. The inserted sequence fragment contained four protein-coding genes and their regulatory elements, as well as the upstream regulatory elements and a partial copy of the *UBE3C* (*Ubiquitin Protein Ligase E3C*) gene (Figure 6A). Using an in vitro human spinal motor neuron (sMN) model generated using patient-derived hiPSCs, the authors demonstrated that the genomic rearrangement alters the 3D chromatin organization at the *DHMN1* locus, causing neo-TAD formation. This neo-TAD includes the upstream regulatory elements and the partial copy of the *UBE3C* gene, as well as other genes and their respective regulatory elements (Figure 6B). Interestingly, in addition to rewiring 3D chromatin architecture, the authors demonstrated that the SV produces a novel gene-intergenic fusion transcript in which the *UBE3C* partial copy is transcribed from the reverse strand and incorporates a terminal pseudo-exon from the adjacent non-coding DNA sequences within the *DHMN1* locus (Figure 6B). Although the novel gene-intergenic fusion transcript does not undergo nonsense-mediated decay, it is unclear whether it encodes for a protein. Nevertheless, by overexpressing the novel gene-intergenic fusion transcript in *Caenorhabditis elegans*, the authors demonstrated that this transcript likely contributes to the disease mechanism through a dominant-negative reduction of the wild-type UBE3C protein. This novel disease mechanism may be overlooked and could explain the pathogenic impact of other SVs reported in the literature. For example, it is tempting to speculate that the 78-kb insertion from chromosome 8 associated with Charcot-Marie-Tooth neuropathy CMTX3 causes the disease through a similar mechanism.[143] This complex insertion corresponds to a duplication of chromosome 8q24.3, containing a partial copy of the *ARHGAP39* (*Rho GTPase Activating Protein 39*) gene, inserted into the *SOX3* (*SRY-Box Transcription Factor 3*) TAD at chromosome Xq27.1. In this case, the partial copy of the *ARHGAP39* gene may be under the control of *SOX3* enhancers generating a novel *ARHGAP39*-intergenic fusion transcript expressed in *SOX3* territories and causing disease.

Fusions involving intergenic/non-coding DNA sequences that result in cryptic (pseudo) exon formation (Figure S1) have only been reported by a few studies.[144–147] Nevertheless, they may be more prevalent than expected. Indeed, most genomic breakpoints in fusion genes are intronic or intergenic.[148–150] However, they are not typically present in mRNA or protein-coding sequences, making the detection of gene-intergenic fusions difficult. Thus, to improve the interpretation of the pathogenic impact of SVs and the identification of disease-causing gene-intergenic fusions in rare genetic disorders, it is essential to use a combination of WGS and RNA-seq patients' data, as well as develop novel algorithms for the detection of novel gene-intergenic fusion transcripts.[144] Interestingly, a recent study reported in the literature demonstrated that using rRNA-minus RNA-seq datasets is advantageous for identifying novel gene-intergenic fusion transcripts.[144] This study revealed, nevertheless, that only a fraction of all genomic rearrangements (~7%) is transcribed, and an even smaller fraction is causal for fusion genes. These data suggest that gene-intergenic fusion transcripts may be a rare cause of genetic disorders.

CONCLUSIONS

The latest technological developments have significantly improved our interpretation of the pathogenic impact of SVs. For many SVs, however, the underlying disease mechanism still needs to be fully resolved, and for many others, the prediction of their pathogenic impact still needs to be improved. Indeed, several challenges remain, hampering the medical interpretation of SVs, predicting their phenotypic consequences, and characterizing their underlying disease mechanism.

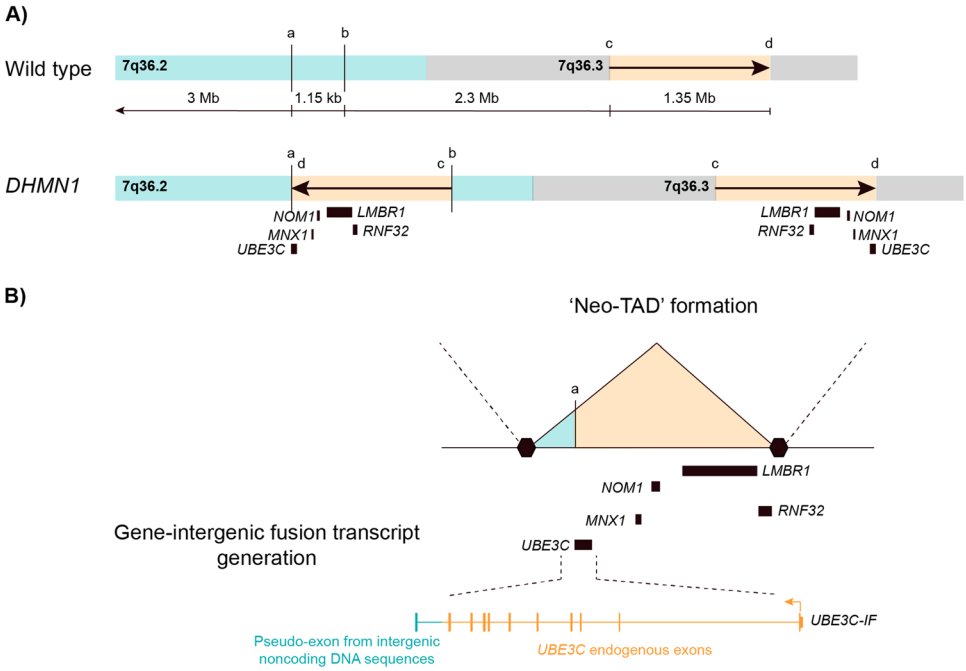

FIGURE 6 SVs producing a novel gene-intergenic fusion transcript. Structural variants that disrupt protein-coding sequences and simultaneously impact intergenic non-coding DNA sequences can produce novel gene-intergenic fusion transcripts at the SV breakpoint. (A) Schematic representation of the genomic region of chromosome 7q36.2-q36.3 (Wild type) involving the formation of the CGR associated with DHMN1 (*DHMN1*). The blue segment represents the *DHMN1* linkage region. The orange segment represents the *DHMN1* insertion sequence. The black arrow indicates the orientation of the *DHMN1* insertion sequence. The 7q36.3 duplicated sequence (orange) is inserted in an inverted orientation within the *DHMN1* locus. This duplicated sequence contains four protein-coding genes (*RNF32*, *LMBR1*, *NOM1*, and *MNX1*), their regulatory landscapes, and a partial copy of the *UBE3C* gene and its regulatory landscape. Letters indicate the proximal (a) and distal (b) ends of the insertion site, and the proximal (c) and distal (d) ends of the 7q36.3 duplicated sequence. (B) The *DHMN1* insertion sequence impacts 3D genome organization. It creates a new TAD containing the four protein-coding genes, their regulatory landscapes, and the partial copy of the *UBE3C* gene and its regulatory landscape (orange). This "Neo-TAD" also contains non-coding DNA sequences adjacent to the insertion breakpoint at the 7q36.2 locus (blue). The CGR results in the generation of a novel gene-intergenic fusion transcript (*UBE3C-IF*). The *UBE3C* partial copy is transcribed from the reverse strand and incorporates a pseudo-exon from the intergenic non-coding DNA sequences adjacent to the insertion breakpoint (a).

First, although high-throughput technologies have produced a vast amount of data in terms of protein validation, protein-protein interactions, genetic variants, gene/protein expression, and 3D structure, contributing to developing a more precise picture of the human proteome, many human proteins still have only a vague or speculative function annotated in UniProtKB/Swiss-Prot[151] and about 10% of human protein-coding genes have no function annotated at all, either predicted or experimentally confirmed.[152] The functional characterization of human proteins is a huge challenge currently undertaken using systems biology approaches combining high-throughput omics technologies and bioinformatics. However, we estimate that there were between 8 and 10 papers describing newly characterized human proteins published each month in the last five years, corresponding to a decrease in the number of uncharacterized proteins by only 25% in 5 years.[152] Thus, the scientific community should encourage collaborative projects aiming at functionalizing uncharacterized proteins to expedite the missing functional annotation of protein-coding genes. The Human Proteome Project (HPP), an international consortium dedicated to cataloguing the 20,230 proteins list, understanding the complexity of the human proteome, and making human proteomics an integrated complement to genomics and other "omics" across clinical, biomedical, and life sciences, has just committed to move in this direction.[152] In its first phase, which was focused on protein validation, the HPP project federated mostly proteomics experts. To succeed in its second phase, it will need to recruit specialists in various other human and model organism biology fields.

Second, to detect and analyze structural variation in non-coding regulatory elements, localizing these regions in the genome and validating them experimentally is necessary. Online databases are the most comprehensive and convenient method to identify regulatory regions throughout the genome because these integrate data from many assays based on biochemical properties (e.g., transcription factor (TF) binding, open chromatin, and histone modifications) to locate functional elements accurately. The main limitation of these databases is that most of the regulatory elements for which they provide information, especially enhancers, are simply putative, predicted regulatory regions, of which only a tiny portion has been experimentally validated.[153,154] Examples include the Encyclopedia of DNA Elements (ENCODE database), the Functional Annotation of the Mammalian Genome project (FANTOM5), the PsychENCODE Consortium, and the

machine learning tool RefMap.[153,154] While we expect new techniques for regulatory element discovery to emerge in the near future, the scientific community should encourage the development of high-throughput projects aiming at experimentally validating in vitro and/or in vivo the cell-type and tissue-specific predicted regulatory regions.

Third, as illustrated by the examples discussed above, disease-causing SVs do not necessarily disrupt enhancer or gene sequences but instead affect the ability of enhancers to activate their target genes specifically. Despite the tremendous progress in the field of enhancer biology during the past few decades, the biochemical processes whereby enhancers communicate specifically with their target genes still need to be clarified.[155] Thus, we anticipate that elucidating the molecular and genetic principles controlling enhancer specificity will help us to better predict the effect of genetic variation in human diseases.

Fourth, the interpretation of the pathogenic impact of SVs may, in some cases, only be achieved by using a multi-omics approach, such as the case for the SVs producing novel gene-intergenic fusion transcripts. This combines genomic, epigenomic, transcriptomic, and functional genomic information. Such an approach has been proposed to unify variant detection and interpretation.[156] It should be considered by clinicians, clinical geneticists, and scientists to improve patients' diagnosis and understanding the basic mechanisms of disease. Indeed, while whole genome and exome sequencing have increased the identification of causal variants compared to single gene testing alone, with diagnostic rates of approximately 50% for inherited diseases,[157] integrated multi-omics analysis may further increase diagnostic yield. Moreover, advances in omics technologies have begun to enable personalized medicine at an extraordinarily detailed molecular level. These technologies have contributed medical advances that have begun to enter clinical practice. However, each technology individually cannot capture the entire biological complexity of most human diseases. Integration of multiple technologies has thus emerged as an approach to provide a more comprehensive view of biology and disease. We direct the reader to an excellent review discussing the importance of integrative omics for health and disease.[158]

Overall, substantial obstacles must be overcome to improve our prediction of the pathogenic impact of SVs and their medical interpretation. Overcoming these challenges becomes crucial. Indeed, recent studies have shown that SVs alter the expression of nearby genes with larger effect sizes than SNVs or small indels, indicating that SVs are more likely to be pathogenic.[159,160] Since these obstacles are not overcome, it is essential to experimentally validate the functional consequences of disease-associated SVs to assign pathogenicity. This is not only true for *de novo* SVs but also inherited variants. Experimentally validating the functional consequences of disease-causing SVs paves the way for characterizing novel disease mechanisms by which SVs exert their pathogenic impact, which is crucial for developing accurate computational prediction tools for SV interpretation.

ACKNOWLEDGMENTS

The authors thank Giulia Cova and Juliane Glaser for information, discussions, and corrections during the preparation of this manuscript. This study was supported by a grant from the Deutsche Forschungsgemeinschaft (MU 880/16-1) to S.M. L.A. was supported by the FOXG1 Foundation (grant P.S.7905).

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors have declared no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were generated or analyzed in this study.

ORCID

Stefan Mundlos https://orcid.org/0000-0001-7824-1945

REFERENCES

1. Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., … Lander, E. S. (2015). A global reference for human genetic variation. *Nature*, *526*, 68–74.
2. Chaisson, M. J. P., Sanders, A. D., Zhao, X., Malhotra, A., Porubsky, D., Rausch, T., Gardner, E. J., Rodriguez, O. L., Guo, L., Collins, R. L., Fan, X., Wen, J., Handsaker, R. E., Fairley, S., Kronenberg, Z. N., Kong, X., Hormozdiari, F., Lee, D., Wenger, A. M., … Lee, C. (2019). Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nature Communication*, *10*, 1784.
3. Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., Khera, A. V., Lowther, C., Gauthier, L D., Wang, H., Watts, N. A., Solomonson, M., O'donnell-Luria, A., Baumann, A., Munshi, R., Walker, M., Whelan, C. W., Huang, Y., Brookings, T., … Talkowski, M. E. (2020). A structural variation reference for medical and population genetics. *Nature*, *581*, 444–451.
4. Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Hsi-Yang Fritz, M., Konkel, M K., Malhotra, A., Stütz, A. M., Shi, X., Paolo Casale, F., Chen, J., Hormozdiari, F., Dayama, G., Chen, K., … Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, *526*, 75–81.
5. Spielmann, M., Lupiáñez, D. G., & Mundlos, S. (2018). Structural variation in the 3D genome. *Nature Reviews Genetics*, *19*, 453–467.
6. Cutrupi, A. N., Brewer, M. H., Nicholson, G. A., & Kennerson, M. L. (2018). Structural variations causing inherited peripheral neuropathies: A paradigm for understanding genomic organization, chromatin interactions, and gene dysregulation. *Molecular Genetics and Genomics Medicine*, *6*, 422–433.
7. Ho, S. S., Urban, A. E., & Mills, R. E. (2020). Structural variation in the sequencing era. *Nature Reviews Genetics*, *21*, 171–189.
8. Despang, A., Schöpflin, R., Franke, M., Ali, S., Jerković, I., Paliou, C., Chan, W. L., Timmermann, B., Wittler, L., Vingron, M., Mundlos, S., & Ibrahim, D M. (2019). Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nature Genetics*, *51*, 1263–1271.
9. Cutrupi, A. N., Narayanan, R. K., Perez-Siles, G., Grosz, B. R., Lai, K., Boyling, A., Ellis, M., Lin, R. C. Y., Neumann, B., Mao, D., Uesugi, M., Nicholson, G. A., Vucic, S., Saporta, M. A., & Kennerson, M. L. (2022). Novel gene-intergenic fusion involving ubiquitin E3 ligase UBE3C causes distal hereditary motor neuropathy. *Brain*, *146*(3), 880–897.
10. Wood, V., Lock, A., Harris, M. A., Rutherford, K., Bähler, J., & Oliver, S. G. (2019). Hidden in plain sight: What remains to be discovered in the eukaryotic proteome? *Open Biology*, *9*, 180241.

11. Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T. D, Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C. H., Kristiansson, K., Macarthur, D. G., Macdonald, J. R., Onyiah, I., Pang, A. W. C., Robson, S., … Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature*, *464*, 704–712.
12. Sudmant, P. H., Kitzman, J. O., Antonacci, F., Alkan, C., Malig, M., Tsalenko, A., Sampas, N., Bruhn, L., Shendure, J., & Eichler, E. E. (2010). Diversity of human copy number variation and multicopy genes. *Science*, *330*, 641–646.
13. Mills, R. E., Walter, K., Stewart, C., Handsaker, R. E., Chen, K., Alkan, C., Abyzov, A., Yoon, S. C., Ye, K., Cheetham, R. K, Chinwalla, A., Conrad, D. F., Fu, Y., Grubert, F., Hajirasouliha, I., Hormozdiari, F., Iakoucheva, L. M., Iqbal, Z., Kang, S., … Korbel, J. O. (2011). Mapping copy number variation by population-scale genome sequencing. *Nature*, *470*, 59–65.
14. Sudmant, P. H., Mallick, S., Nelson, B. J., Hormozdiari, F., Krumm, N., Huddleston, J., Coe, B. P., Baker, C., Nordenfelt, S., Bamshad, M., Jorde, L. B., Posukh, O. L., Sahakyan, H., Watkins, W. S, Yepiskoposyan, L., Abdullah, M. S, Bravi, C. M., Capelli, C., Hervig, T., … Eichler, E. E. (2015). Global diversity, population stratification, and selection of human copy-number variation. *Science*, *349*, aab3761.
15. Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, *12*, 363–376.
16. Carvalho, C. M. B., Ramocki, M. B., Pehlivan, D., Franco, L. M., Gonzaga-Jauregui, C., Fang, P., Mccall, A., Pivnick, E. K., Hines-Dowell, S., Seaver, L. H., Friehling, L., Lee, S., Smith, R., Del Gaudio, D., Withers, M., Liu, P., Cheung, S. W., Belmont, J. W., Zoghbi, H. Y., … Lupski, J. R. (2011). Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nature Genetics*, *43*, 1074–1081.
17. Kidd, J. M., Graves, T., Newman, T. L., Fulton, R., Hayden, H. S., Malig, M., Kallicki, J., Kaul, R., Wilson, R. K., & Eichler, E. E. (2010). A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell*, *143*, 837–847.
18. Yang, L., Luquette, L. J., Gehlenborg, N., Xi, R., Haseley, P. S., Hsieh, C.-H., Zhang, C., Ren, X., Protopopov, A., Chin, L., Kucherlapati, R., Lee, C., & Park, P. J. (2013). Diverse mechanisms of somatic structural variations in human cancer genomes. *Cell*, *153*, 919–929.
19. Li, Y., Roberts, N. D., Wala, J. A., Shapira, O., Schumacher, S. E., Kumar, K., Khurana, E., Waszak, S., Korbel, J. O., Haber, J. E., Imielinski, M., Akdemir, K. C., Alvarez, E. G., Baez-Ortega, A., Beroukhim, R., Boutros, P. C., Bowtell, D. D. L., Brors, B., Burns, K. H., … Von Mering, C. (2020). Patterns of somatic structural variation in human cancer genomes. *Nature*, *578*, 112–121.
20. Pös, O., Radvanszky, J., Buglyó, G., Pös, Z., Rusnakova, D., Nagy, B., & Szemes, T. (2021). DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects. *Biomed Journal*, *44*, 548–559.
21. Lupski, J. R. (2015). Structural variation mutagenesis of the human genome: Impact on disease and evolution. *Environmental and Molecular Mutagenesis*, *56*, 419–436.
22. Hastings, P. J., Lupski, J. R., Rosenberg, S. M., & Ira, G. (2009). Mechanisms of change in gene copy number. *Nature Reviews Genetics*, *10*, 551–564.
23. Carvalho, C. M. B., & Lupski, J. R. (2016). Mechanisms underlying structural variant formation in genomic disorders. *Nature Reviews Genetics*, *17*, 224–238.
24. Weckselblatt, B., & Rudd, M. K. (2015). Human structural variation: Mechanisms of chromosome rearrangements. *Trends in Genetics*, *31*, 587–599.
25. Wan, T. S. K. (2014). Cancer cytogenetics: Methodology revisited. *Annals of Laboratory Medicine*, *34*, 413–425.
26. Zarrei, M., Macdonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nature Reviews Genetics*, *16*, 172–183.
27. Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., Abdel-Hamid, H., Bader, P., Mccracken, E., Niyazov, D., Leppig, K., Thiese, H., Hummel, M., Alexander, N., Gorski, J., … Eichler, E. E. (2011). A copy number variation morbidity map of developmental delay. *Nature Genetics*, *43*, 838–846.
28. Wright, C. F., Fitzpatrick, D. R., & Firth, H. V. (2018). Paediatric genomics: Diagnosing rare disease in children. *Nature Reviews Genetics*, *19*, 253–268.
29. Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., Van De Vorst, M., Van Bon, B. W. M., Willemsen, M. H., Kwint, M., Janssen, I. M., Hoischen, A., Schenck, A., Leach, R., Klein, R., Tearle, R., Bo, T., Pfundt, R., Yntema, H. G., De Vries, B. B. A., Kleefstra, T., Brunner, H. G., … Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, *511*, 344–347.
30. Cretu Stancu, M., Van Roosmalen, M. J., Renkens, I., Nieboer, M. M., Middelkamp, S., De Ligt, J., Pregno, G., Giachino, D., Mandrile, G., Espejo Valle-Inclan, J., Korzelius, J., De Bruijn, E., Cuppen, E., Talkowski, M. E., Marschall, T., De Ridder, J., & Kloosterman, W. P. (2017). Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nature Communication*, *8*, 1326.
31. Huddleston, J., Chaisson, M. J. P., Steinberg, K. M., Warren, W., Hoekzema, K., Gordon, D., Graves-Lindsay, T. A., Munson, K. M., Kronenberg, Z. N., Vives, L., Peluso, P., Boitano, M., Chin, C.-S., Korlach, J., Wilson, R. K., & Eichler, E. E. (2017). Discovery and genotyping of structural variation from long-read haploid genome sequence data. *Genome Research*, *27*, 677–685.
32. Bonev, B., & Cavalli, G. (2016). Organization and function of the 3D genome. *Nature Reviews Genetics*, *17*, 772.
33. Jerkovicť, I., & Cavalli, G. (2021). Understanding 3D genome organization by multidisciplinary methods. *Nature Reviews Molecular Cell Biology*, *22*, 511–528.
34. Mohanta, T. K., Mishra, A. K., & Al-Harrasi, A. (2021). The 3D genome: From structure to function. *International Journal of Molecular Sciences*, *22*, 11585.
35. Robson, M. I., Ringel, A. R., & Mundlos, S. (2019). Regulatory landscaping: How enhancer-promoter communication is sculpted in 3D. *Molecular Cell*, *74*, 1110–1122.
36. Glaser, J., & Mundlos, S. (2022). 3D or not 3D: Shaping the genome during development. *Cold Spring Harbor Perspectives in Biology*, *14*, a040188.
37. Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysak, M. A., Fuchs, J., & Schubert, I. (2004). Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. *Chromosoma*, *113*, 258–269.
38. Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science*, *295*, 1306–1311.
39. Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., & Ohlsson, R. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature Genetics*, *38*, 1341–1347.
40. Rodley, C. D. M., Bertels, F., Jones, B., & O'sullivan, J. M. (2009). Global identification of yeast chromosome interactions using Genome conformation capture. *Fungal Genetics and Biology*, *46*, 879–886.
41. Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C., Green, R. D., & Dekker, J. (2006). Chromosome Conformation Capture Car-

bon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Research*, *16*, 1299–1309.
42. Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*, 289–293.
43. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., De Wit, E., Van Steensel, B., & De Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Genetics*, *38*, 1348–1354.
44. Dekker, J., Marti-Renom, M. A., & Mirny, L. A. (2013). Exploring the three-dimensional organization of genomes: Interpreting chromatin interaction data. *Nature Reviews Genetics*, *14*, 390–403.
45. Pombo, A., & Dillon, N. (2015). Three-dimensional genome architecture: Players and mechanisms. *Nature Reviews Molecular Cell Biology*, *16*, 245–257.
46. Van Steensel, B., & Belmont, A. S. (2017). Lamina-associated domains: Links with chromosome architecture, heterochromatin, and gene repression. *Cell*, *169*, 780–791.
47. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., & Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. *Nature*, *547*, 241–245.
48. Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*, 376–380.
49. Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., & Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, *485*, 381–385.
50. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., & Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*, *148*, 458–472.
51. Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of chromosomal domains by loop extrusion. *Cell Reports*, *15*, 2038–2049.
52. Mirny, L. A., & Solovei, I. (2021). Keeping chromatin in the loop(s). *Nature Reviews Molecular Cell Biology*, *22*, 439–440.
53. Sanborn, A. L., Rao, S. S. P., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K. P., Gnirke, A., Melnikov, A., Mckenna, D., Stamenova, E. K., Lander, E. S., & Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *PNAS*, *112*, E6456–65.
54. Nora, E. P., Goloborodko, A., Valton, A.-L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A., & Bruneau, B. G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, *169*, 930–944.e22.
55. Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., Huber, W., Haering, C. H., Mirny, L., & Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, *551*, 51–56.
56. Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., & Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, *159*, 1665–1680.
57. Rao, S. S. P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D., Huang, X., Shamim, M. S., Shin, J., Turner, D., Ye, Z., Omer, A. D., Robinson, J. T., Schlick, T., Bernstein, B. E., … Casellas, R. (2017). Cohesin loss eliminates all loop domains. *Cell*, *171*, 305–320.e24.
58. Vian, L., Pękowska, A., Rao, S. S. P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., El Khattabi, L., Dose, M., Pruett, N., Sanborn, A. L., Canela, A., Maman, Y., Oksanen, A., Resch, W., Li, X., Lee, B., Kovalchuk, A. L., Tang, Z., … Casellas, R. (2018). The energetics and physiological impact of cohesin extrusion. *Cell*, *173*, 1165–1178.e20.
59. Haarhuis, J. H. I., Van Der Weide, R. H., Blomen, V. A., Yáñez-Cuna, J. O., Amendola, M., Van Ruiten, M. S., Krijger, P. H. L., Teunissen, H., Medema, R. H., Van Steensel, B., Brummelkamp, T. R., De Wit, E., & Rowland, B. D. (2017). The cohesin release factor WAPL restricts chromatin loop extension. *Cell*, *169*, 693–707.e14.
60. Valton, A.-L., Venev, S. V., Mair, B., Khokhar, E. S., Tong, A. H. Y., Usaj, M., Chan, K., Pai, A. A., Moffat, J., & Dekker, J. (2022). A cohesin traffic pattern genetically linked to gene regulation. *Nature Structural & Molecular Biology*, *29*, 1239–1251.
61. Banigan, E. J., Tang, W., van den Berg, A. A., Stocsits, R. R., Wutz, G.,Brandao, H. B.,Busslinger, G. A.,Peters, J.-M., & Mirny, L. A. (2022). Transcription shapes 3D chromatin organization by interacting with loop extrusion. *Biophysics and Computational Biology*, *120*, 1–12.
62. De Wit, E., Vos, E. S. M., Holwerda, S. J. B., Valdes-Quezada, C., Verstegen, M. J. A. M., Teunissen, H., Splinter, E., Wijchers, P. J., Krijger, P. H. L., & De Laat, W. (2015). CTCF binding polarity determines chromatin looping. *Molecular Cell*, *60*, 676–684.
63. Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S. A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., … Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, *161*, 1012–1025.
64. Ruf, S., Symmons, O., Uslu, V. V., Dolle, D., Hot, C., Ettwiller, L., & Spitz, F. (2011). Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nature Genetics*, *43*, 379–386.
65. Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., & Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research*, *24*, 390–400.
66. Marinić, M., Aktas, T., Ruf, S., & Spitz, F. (2013). An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Developmental Cell*, *24*, 530–542.
67. Palstra, R.-J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., & De Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nature Genetics*, *35*, 190–194.
68. Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L., Spielmann, M., Timmermann, B., Wittler, L., Kurth, I., Cambiaso, P., Zuffardi, O., Houge, G., Lambie, L., Brancati, F., … Mundlos, S. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, *538*, 265–269.
69. Hanssen, L. L. P., Kassouf, M. T., Oudelaar, A. M., Biggs, D., Preece, C., Downes, D. J., Gosden, M., Sharpe, J. A., Sloane-Stanley, J. A., Hughes, J. R., Davies, B., & Higgs, D. R. (2017). Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature Cell Biology*, *19*, 952–961.
70. Vos, E. S. M., Valdes-Quezada, C., Huang, Y., Allahyar, A., Verstegen, M. J. A. M., Felder, A.-K., Van Der Vegt, F., Uijttewaal, E. C. H., Krijger, P. H. L., & De Laat, W. (2021). Interplay between CTCF boundaries and a super enhancer controls cohesin extrusion trajectories and gene expression. *Molecular Cell*, *81*, 3082–3095.e6.
71. Weischenfeldt, J., Dubash, T., Drainas, A. P., Mardin, B. R., Chen, Y., Stütz, A. M., Waszak, S. M., Bosco, G., Halvorsen, A. R., Raeder, B.,

Efthymiopoulos, T., Erkek, S., Siegl, C., Brenner, H., Brustugun, O. T., Dieter, S. M., Northcott, P. A., Petersen, I., Pfister, S. M., … Korbel, J. O. (2017). Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nature Genetics*, *49*, 65–74.

72. Kraft, K., Magg, A., Heinrich, V., Riemenschneider, C., Schöpflin, R., Markowski, J., Ibrahim, D. M., Acuna-Hidalgo, R., Despang, A., Andrey, G., Wittler, L., Timmermann, B., Vingron, M., & Mundlos, S. (2019). Serial genomic inversions induce tissue-specific architectural stripes, gene misexpression and congenital malformations. *Nature Cell Biology*, *21*, 305–310.
73. Hsieh, T. S., Cattoglio, C., Slobodyanyuk, E., Hansen, A. S., Darzacq, X., & Tjian, R. (2022). Enhancer-promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nature Genetics*, *54*, 1919–1932.
74. Akalin, A., Fredman, D., Arner, E., Dong, X., Bryne, J., Suzuki, H., Daub, C. O., Hayashizaki, Y., & Lenhard, B. (2009). Transcriptional features of genomic regulatory blocks. *Genome Biology*, *10*, R38.
75. Kikuta, H., Laplante, M., Navratilova, P., Komisarczuk, A. Z., Engström, P. G., Fredman, D., Akalin, A., Caccamo, M., Sealy, I., Howe, K., Ghislain, J., Pezeron, G., Mourrain, P., Ellingsen, S., Oates, A. C., Thisse, C., Thisse, B., Foucher, I., Adolf, B., … Becker, T. S. (2007). Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Research*, *17*, 545–555.
76. Ringel, A. R., Szabo, Q., Chiariello, A. M., Chudzik, K., Schöpflin, R., Rothe, P., Mattei, A. L., Zehnder, T., Harnett, D., Laupert, V., Bianco, S., Hetzel, S., Glaser, J., Phan, M. H. Q., Schindler, M., Ibrahim, D. M., Paliou, C., Esposito, A., Prada-Medina, C. A., … Robson, M. I. (2022). Repression and 3D-restructuring resolves regulatory conflicts in evolutionarily rearranged genomes. *Cell*, *185*, 3689–3704.e21.
77. Lohan, S., Spielmann, M., Doelken, S. C., Flöttmann, R., Muhammad, F., Baig, S. M., Wajid, M., Hülsemann, W., Habenicht, R., Kjaer, K. W., Patil, S. J., Girisha, K. M., Abarca-Barriga, H. H., Mundlos, S., & Klopocki, E. (2014). Microduplications encompassing the Sonic hedgehog limb enhancer ZRS are associated with Haas-type polysyndactyly and Laurin-Sandrow syndrome. *Clinical Genetics*, *86*, 318–325.
78. Klopocki, E., Ott, C.-E., Benatar, N., Ullmann, R., Mundlos, S., & Lehmann, K. (2008). A microduplication of the long range SHH limb regulator (ZRS) is associated with triphalangeal thumb-polysyndactyly syndrome. *Journal of Medical Genetics*, *45*, 370–375.
79. Benko, S., Fantes, J. A., Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C. T., Mcbride, D., Golzio, C., Fisher, M., Perry, P., Abadie, V., Ayuso, C., Holder-Espinasse, M., Kilpatrick, N., Lees, M. M., … Lyonnet, S. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nature Genetics*, *41*, 359–364.
80. Birnbaum, R Y., Clowney, E. J, Agamy, O., Kim, M. J., Zhao, J., Yamanaka, T., Pappalardo, Z., Clarke, S. L., Wenger, A. M., Nguyen, L., Gurrieri, F., Everman, D. B., Schwartz, C. E., Birk, O. S., Bejerano, G., Lomvardas, S., & Ahituv, N. (2012). Coding exons function as tissue-specific enhancers of nearby genes. *Genome Research*, *22*, 1059–1068.
81. Tayebi, N., Jamsheer, A., Flöttmann, R., Sowinska-Seidler, A., Doelken, S. C., Oehl-Jaschkowitz, B., Hülsemann, W., Habenicht, R., Klopocki, E., Mundlos, S., & Spielmann, M. (2014). Deletions of exons with regulatory activity at the DYNC1I1 locus are associated with split-hand/split-foot malformation: Array CGH screening of 134 unrelated families. *Orphanet Journal of Rare Diseases*, *9*, 108.
82. Lango Allen, H., Caswell, R., Xie, W., Xu, X., Wragg, C., Turnpenny, P. D., Turner, C. L. S., Weedon, M. N., & Ellard, S. (2014). Next generation sequencing of chromosomal rearrangements in patients with split-hand/split-foot malformation provides evidence for DYNC1I1 exonic enhancers of DLX5/6 expression in humans. *Journal of Medical Genetics*, *51*, 264–267.
83. Ghiasvand, N. M., Rudolph, D. D., Mashayekhi, M., Brzezinski, J. A., Goldman, D., & Glaser, T. (2011). Deletion of a remote enhancer near ATOH7 disrupts retinal neurogenesis, causing NCRNA disease. *Nature Neuroscience*, *14*, 578–586.
84. Bhatia, S., Bengani, H., Fish, M., Brown, A., Divizia, M. T., De Marco, R., Damante, G., Grainger, R., Van Heyningen, V., & Kleinjan, D. A. (2013). Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *American Journal of Human Genetics*, *93*, 1126–1134.
85. Allou, L., Balzano, S., Magg, A., Quinodoz, M., Royer-Bertrand, B., Schöpflin, R., Chan, W. L., Speck-Martins, C. E., Carvalho, D. R., Farage, L., Lourenço, C. M., Albuquerque, R., Rajagopal, S., Nampoothiri, S., Campos-Xavier, B., Chiesa, C., Niel-Bütschi, F., Wittler, L., Timmermann, B., … Superti-Furga, A. (2021). Non-coding deletions identify Maenli lncRNA as a limb-specific En1 regulator. *Nature*, *592*, 93–8.
86. Dathe, K., Kjaer, K. W., Brehm, A., Meinecke, P., Nürnberg, P., Neto, J. C., Brunoni, D., Tommerup, N., Ott, C. E., Klopocki, E., Seemann, P., & Mundlos, S. (2009). Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *American Journal of Human Genetics*, *84*, 483–492.
87. Sun, M., Ma, F., Zeng, X., Liu, Q., Zhao, X.-L., Wu, F.-X., Wu, G.-P., Zhang, Z.-F., Gu, B., Zhao, Y.-F., Tian, S.-H., Lin, B., Kong, X.-Y., Zhang, X.-L., Yang, W., Lo, W. H.-Y., & Zhang, X. (2008). Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer. *Journal of Medical Genetics*, *45*, 589–595.
88. Ngcungcu, T., Oti, M., Sitek, J. C., Haukanes, B. I., Linghu, B., Bruccoleri, R., Stokowy, T., Oakeley, E. J., Yang, F., Zhu, J., Sultan, M., Schalkwijk, J., Van Vlijmen-Willems, I. M. J. J., Von Der Lippe, C., Brunner, H. G., Ersland, K. M., Grayson, W., Buechmann-Moller, S., Sundnes, O., … Ramsay, M. (2017). Duplicated enhancer region increases expression of CTSB and segregates with keratolytic winter erythema in South African and Norwegian families. *American Journal of Human Genetics*, *100*, 737–750.
89. Will, A. J., Cova, G., Osterwalder, M., Chan, W.-L., Wittler, L., Brieske, N., Heinrich, V., De Villartay, J.-P., Vingron, M., Klopocki, E., Visel, A., Lupiáñez, D. G., & Mundlos, S. (2017). Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nature Genetics*, *49*, 1539–1545.
90. Fanelli, A., Vannelli, S., Babu, D., Mellone, S., Cucci, A., Monzani, A., Al Essa, W., Secco, A., Follenzi, A., Bellone, S., Prodam, F., & Giordano, M. (2022). Copy number variations residing outside the SHOX enhancer region are involved in Short Stature and Leri-Weill dyschondrosteosis. *Molecular Genetics & Genomic Medicine*, *10*, e1793.
91. Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., Lee, E. A., Zhu, Y., Plajzer-Frick, I., Pickle, C. S., Kato, M., Garvin, T. H., Pham, Q. T., Harrington, A. N., Akiyama, J. A., Afzal, V., Lopez-Rios, J., Dickel, D. E., Visel, A., & Pennacchio, L. A. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*, *554*, 239–243.
92. Hörnblad, A., Bastide, S., Langenfeld, K., Langa, F., & Spitz, F. (2021). Dissection of the Fgf8 regulatory landscape by in vivo CRISPR-editing reveals extensive intra- and inter-enhancer redundancy. *Nature Communication*, *12*, 439.
93. Northcott, P. A., Lee, C., Zichner, T., Stütz, A. M., Erkek, S., Kawauchi, D., Shih, D. J. H., Hovestadt, V., Zapatka, M., Sturm, D., Jones, D. T. W., Kool, M., Remke, M., Cavalli, F. M. G., Zuyderduyn, S., Bader, G. D., Vandenberg, S., Esparza, L. A., Ryzhova, M., … Pfister, S. M. (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature*, *511*, 428–434.
94. Northcott, P. A., Buchhalter, I., Morrissy, A. S., Hovestadt, V., Weischenfeldt, J., Ehrenberger, T., Gröbner, S., Segura-Wang, M., Zichner, T., Rudneva, V. A., Warnatz, H.-J., Sidiropoulos, N., Phillips, A.

H., Schumacher, S., Kleinheinz, K., Waszak, S. M., Erkek, S., Jones, D. T. W., Worst, B. C., … Lichter, P. (2017). The whole-genome landscape of medulloblastoma subtypes. *Nature*, *547*, 311–317.

95. Hnisz, D., Weintraub, A. S., Day, D. S., Valton, A.-L., Bak, R. O., Li, C. H., Goldmann, J., Lajoie, B. R., Fan, Z. P., Sigova, A. A., Reddy, J., Borges-Rivera, D., Lee, T. I., Jaenisch, R., Porteus, M. H., Dekker, J., & Young, R. A. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, *351*, 1454–1458.
96. Meinel, J. A., Yumiceba, V., Kunstner, A., Schultz, K., Kruse, N., Kaiser, F. J., Holterhus, P.-M., Claviez, A., Hiort, O., Busch, H., Spielmann, M., & Werner, R. (2022). Disruption of the topologically associated domain at Xp21.2 is related to 46,XY gonadal dysgenesis. *Journal of Medical Genetics*, *60*, 469–476.
97. Franke, M., Daly, A. F., Palmeira, L., Tirosh, A., Stigliano, A., Trifan, E., Faucz, F. R., Abboud, D., Petrossians, P., Tena, J. J., Vitali, E., Lania, A. G., Gómez-Skarmeta, J. L., Beckers, A., Stratakis, C. A., & Trivellin, G. (2022). Duplications disrupt chromatin architecture and rewire GPR101-enhancer communication in X-linked acrogigantism. *American Journal of Human Genetics*, *109*, 553–570.
98. Melo, U. S., Piard, J., Fischer-Zirnsak, B., Klever, M.-sK., Schöpflin, R., Mensah, M. A., Holtgrewe, M., Arbez-Gindre, F., Martin, A., Guigue, V., Gaillard, D., Landais, E., Roze, V., Kremer, V., Ramanah, R., Cabrol, C., Harms, F L., Kornak, U., Spielmann, M., … Van Maldergem, L. (2021). Complete lung agenesis caused by complex genomic rearrangements with neo-TAD formation at the SHH locus. *Human Genetics*, *140*, 1459–1469.
99. De Bruijn, S. E., Fiorentino, A., Ottaviani, D., Fanucchi, S., Melo, U. S., Corral-Serrano, J. C., Mulders, T., Georgiou, M., Rivolta, C., Pontikos, N., Arno, G., Roberts, L., Greenberg, J., Albert, S., Gilissen, C., Aben, M., Rebello, G., Mead, S., Raymond, F. L., … Hardcastle, A. J. (2020). Structural variants create new topological-associated domains and ectopic retinal enhancer-gene contact in dominant retinitis pigmentosa. *American Journal of Human Genetics*, *107*, 802–814.
100. Melo, U. S., Schöpflin, R., Acuna-Hidalgo, R., Mensah, M. A., Fischer-Zirnsak, B., Holtgrewe, M., Klever, M.-K., Türkmen, S., Heinrich, V., Pluym, I. D., Matoso, E., Bernardo De Sousa, S., Louro, P., Hülsemann, W., Cohen, M., Dufke, A., Latos-Bieleńska, A., Vingron, M., Kalscheuer, V., … Mundlos, S. (2020). Hi-C identifies complex genomic rearrangements and TAD-shuffling in developmental diseases. *American Journal of Human Genetics*, *106*, 872–884.
101. Cova, G., Glaser, J., Schöpflin, R., Prada-Medina, C. A., Ali, S., Franke, M., Falcone, R., Federer, M., Ponzi, E., Ficarella, R., Novara, F., Wittler, L., Timmermann, B., Gentile, M., Zuffardi, O., Spielmann, M., & Mundlos, S. (2023). Combinatorial effects on gene expression at the Lbx1/Fgf8 locus resolve split-hand/foot malformation type 3. *Nature Communication*, *14*, 1475.
102. Vicente-García, C., Villarejo-Balcells, B., Irastorza-Azcárate, I., Naranjo, S., Acemel, R. D., Tena, J. J., Rigby, P. W. J., Devos, D. P., Gómez-Skarmeta, J. L., & Carvajal, J. J. (2017). Regulatory landscape fusion in rhabdomyosarcoma through interactions between the PAX3 promoter and FOXO1 regulatory elements. *Genome Biology*, *18*, 106.
103. Laugsch, M., Bartusel, M., Rehimi, R., Alirzayeva, H., Karaolidou, A., Crispatzu, G., Zentis, P., Nikolic, M., Bleckwehl, T., Kolovos, P., Van Ijcken, W. F. J., Šarić, T., Koehler, K., Frommolt, P., Lachlan, K., Baptista, J., & Rada-Iglesias, A. (2019). Modeling the pathological long-range regulatory effects of human structural variation with patient-specific hiPSCs. *Cell Stem Cell*, *24*, 736–752.e12.
104. Redin, C., Brand, H., Collins, R. L., Kammin, T., Mitchell, E., Hodge, J. C., Hanscom, C., Pillalamarri, V., Seabra, C. M., Abbott, M.-A., Abdul-Rahman, O. A., Aberg, E., Adley, R., Alcaraz-Estrada, S. L., Alkuraya, F. S., An, Y., Anderson, M.-A., Antolik, C., Anyane-Yeboa, K., … Talkowski, M. E. (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nature Genetics*, *49*, 36–45.
105. Giorgio, E., Robyr, D., Spielmann, M., Ferrero, E., Di Gregorio, E., Imperiale, D., Vaula, G., Stamoulis, G., Santoni, F., Atzori, C., Gasparini, L., Ferrera, D., Canale, C., Guipponi, M., Pennacchio, L. A., Antonarakis, S. E., Brussino, A., & Brusco, A. (2015). A large genomic deletion leads to enhancer adoption by the lamin B1 gene: A second path to autosomal dominant adult-onset demyelinating leukodystrophy (ADLD). *Human Molecular Genetics*, *24*, 3143–3154.
106. Mouri, K., Sagai, T., Maeno, A., Amano, T., Toyoda, A., & Shiroishi, T. (2018). Enhancer adoption caused by genomic insertion elicits interdigital Shh expression and syndactyly in mouse. *PNAS*, *115*, 1021–1026.
107. Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R R., Korbel, J. O., & Furlong, E. E. M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nature Genetics*, *51*, 1272–1282.
108. Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., Campbell, P. J., Chin, L., Dixon, J. R., Futreal, P. A, Akdemir, K. C., Alvarez, E. G., Baez-Ortega, A., Boutros, P. C., Bowtell, D. D. L., Brors, B., Burns, K. H., Campbell, P. J., Chan, K., … Von Mering, C. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nature Genetics*, *52*, 294–305.
109. Schöpflin, R., Melo, U. S., Moeinzadeh, H., Heller, D., Laupert, V., Hertzberg, J., Holtgrewe, M., Alavi, N., Klever, M.-K., Jungnitsch, J., Comak, E., Türkmen, S., Horn, D., Duffourd, Y., Faivre, L., Callier, P., Sanlaville, D., Zuffardi, O., Tenconi, R., … Mundlos, S. (2022). Integration of Hi-C with short and long-read genome sequencing reveals the structure of germline rearranged genomes. *Nature Communication*, *13*, 6470.
110. Gabriele, M., Brandão, H. B., Grosse-Holz, S., Jha, A., Dailey, G. M., Cattoglio, C., Hsieh, T.-H. S., Mirny, L., Zechner, C., & Hansen, A. S. (2022). Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, *376*, 496–501.
111. Mach, P., Kos, P. I., Zhan, Y., Cramard, J., Gaudin, S., Tünnermann, J., Marchi, E., Eglinger, J., Zuin, J., Kryzhanovska, M., Smallwood, S., Gelman, L., Roth, G., Nora, E. P., Tiana, G., & Giorgetti, L. (2022). Live-cell imaging and physical modeling reveal control of chromosome folding dynamics by cohesin and CTCF. *bioRxiv*, 2022.03.03.482826.
112. Finn, E H., Pegoraro, G., Brandão, H B., Valton, A.-L., Oomen, M E., Dekker, J., Mirny, L., & Misteli, T. (2019). Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell*, *176*, 1502–1515.e10.
113. Mcarthur, E., & Capra, J A. (2021). Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. *American Journal of Human Genetics*, *108*, 269–283.
114. Bonev, B., Mendelson Cohen, N., & Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., & Cavalli, G. (2017). Multiscale 3D genome rewiring during mouse neural development. *Cell*, *171*, 557–572.e24.
115. Paulsen, J., Liyakat Ali, T. M., Nekrasov, M., Delbarre, E., Baudement, M.-O., Kurscheid, S., Tremethick, D., & Collas, P. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nature Genetics*, *51*, 835–843.
116. Szabo, Q., Jost, D., Chang, J.-M., Cattoni, D. I., Papadopoulos, G. L., Bonev, B., Sexton, T., Gurgo, J., Jacquier, C., Nollmann, M., Bantignies, F., & Cavalli, G. (2018). TADs are 3D structural units of higher-order chromosome organization in Drosophila. *Science Advances*, *4*, eaar8082.
117. Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., & Spitz, F. (2016). The Shh topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Developmental Cell*, *39*, 529–543.
118. Galupa, R., Nora, E. P., Worsley-Hunt, R., Picard, C., Gard, C., Van Bemmel, J. G., Servant, N., Zhan, Y., El Marjou, F., Johanneau, C.,

Diabangouaya, P., Le Saux, A., Lameiras, S., Pipoli Da Fonseca, J., Loos, F., Gribnau, J., Baulande, S., Ohler, U., Giorgetti, L., & Heard, E. (2020). A conserved noncoding locus regulates random monoallelic xist expression across a topological boundary. *Molecular Cell*, *77*, 352–367.e8.

119. D'haene, E., Bar-Yaacov, R., Bariah, I., Vantomme, L., Van Loo, S., Cobos, F. A., Verboom, K., Eshel, R., Alatawna, R., Menten, B., Birnbaum, R. Y., & Vergult, S. (2020). A neuronal enhancer network upstream of MEF2C is compromised in patients with Rett like characteristics. *Human Molecular Genetics*, *29*, 879–880.
120. Mohajeri, K., Yadav, R., D'haene, E., Boone, P. M., Erdin, S., Gao, D., Moyses-Oliveira, M., Bhavsar, R., Currall, B. B., O'keefe, K., Burt, N. D., Lowther, C., Lucente, D., Salani, M., Larson, M., Redin, C., Dudchenko, O., Aiden, E. L., Menten, B., … Talkowski, M. E. (2022). Transcriptional and functional consequences of alterations to MEF2C and its topological organization in neuronal models. *American Journal of Human Genetics*, *109*, 2049–2067.
121. Zepeda-Mendoza, C. J., Ibn-Salem, J., Kammin, T., Harris, D. J., Rita, D., Gripp, K. W., Mackenzie, J. J., Gropman, A., Graham, B., Shaheen, R., Alkuraya, F. S., Brasington, C. K., Spence, E. J., Masser-Frye, D., Bird, L. M., Spiegel, E., Sparkes, R. L., Ordulu, Z., Talkowski, M. E., … Morton, C. C. (2017). Computational prediction of position effects of apparently balanced human chromosomal rearrangements. *American Journal of Human Genetics*, *101*, 206–217.
122. Gaszner, M., & Felsenfeld, G. (2006). Insulators: Exploiting transcriptional and epigenetic mechanisms. *Nature Reviews Genetics*, *7*, 703–713.
123. Doni Jayavelu, N., Jajodia, A., Mishra, A., & Hawkins, R. D. (2020). Candidate silencer elements for the human and mouse genomes. *Nature Communication*, *11*, 1061.
124. Ngan, C. Y., Wong, C. H., Tjong, H., Wang, W., Goldfeder, R L., Choi, C., He, H., Gong, L., Lin, J., Urban, B., Chow, J., Li, M., Lim, J., Philip, V., Murray, S. A., Wang, H., & Wei, C.-L. (2020). Chromatin interaction analyses elucidate the roles of PRC2-bound silencers in mouse development. *Nature Genetics*, *52*, 264–272.
125. Pang, B., & Snyder, M. P. (2020). Systematic identification of silencers in human cells. *Nature Genetics*, *52*, 254–263.
126. Wang, X., Xu, J., Zhang, B., Hou, Y., Song, F., Lyu, H., & Yue, F. (2021). Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. *Nature Methods*, *18*, 661–668.
127. Mitelman, F., Johansson, B., & Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. *Nature Reviews Cancer*, *7*, 233–245.
128. Mertens, F., Johansson, B., Fioretos, T., & Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. *Nature Reviews Cancer*, *15*, 371–381.
129. Dubois, F., Sidiropoulos, N., Weischenfeldt, J., & Beroukhim, R. (2022). Structural variations in cancer and the 3D genome. *Nature Reviews Cancer*, *22*, 533–546.
130. Yamada, M., Suzuki, H., Watanabe, A., Uehara, T., Takenouchi, T., Mizuno, S., & Kosaki, K. (2021). Role of chimeric transcript formation in the pathogenesis of birth defects. *Congenital Anomalies*, *61*, 76–81.
131. Cousin, M. A., Smith, M. J., Sigafoos, A. N., Jin, J. J., Murphree, M. I., Boczek, N. J., Blackburn, P. R., Oliver, G. R., Aleff, R. A., Clark, K. J., Wieben, E. D., Joshi, A. Y., Pichurin, P. N., Abraham, R. S., & Klee, E. W. (2018). Utility of DNA, RNA, protein, and functional approaches to solve cryptic immunodeficiencies. *Journal of Clinical Immunology*, *38*, 307–319.
132. Oliver, G. R., Blackburn, P. R., Ellingson, M. S., Conboy, E., Pinto E Vairo, F., Webley, M., Thorland, E., Ferber, M., Van Hul, E., Werf, I. M., Wuyts, W., Babovic-Vuksanovic, D., & Klee, E. W. (2019). RNA-Seq detects a SAMD12-EXT1 fusion transcript and leads to the discovery of an EXT1 deletion in a child with multiple osteochondromas. *Molecular Genetics & Genomic Medicine*, *7*, e00560.
133. Oliver, G. R., Tang, X., Schultz-Rogers, L. E., Vidal-Folch, N., Jenkinson, W. G, Schwab, T. L., Gaonkar, K., Cousin, M. A., Nair, A., Basu, S., Chanana, P., Oglesbee, D., & Klee, E. W. (2019). A tailored approach to fusion transcript identification increases diagnosis of rare inherited disease. *Plos One*, *14*, e0223337.
134. Ceroni, F., Sagar, A., Simpson, N. H., Gawthrope, A. J. T., Newbury, D. F., Pinto, D., Francis, S. M., Tessman, D. C., Cook, E. H., Monaco, A. P., Maestrini, E., Pagnamenta, A. T., & Jacob, S. (2014). A deletion involving CD38 and BST1 results in a fusion transcript in a patient with autism and asthma. *Autism Research*, *7*, 254–263.
135. Pagnamenta, A. T., Bacchelli, E., De Jonge, M. V., Mirza, G., Scerri, T. S., Minopoli, F., Chiocchetti, A., Ludwig, K. U., Hoffmann, P., Paracchini, S., Lowy, E., Harold, D. H., Chapman, J. A., Klauck, S. M., Poustka, F., Houben, R. H., Staal, W. G., Ophoff, R. A., O'donovan, M. C., … Monaco, A. P. (2010). Characterization of a family with rare deletions in CNTNAP5 and DOCK4 suggests novel risk loci for autism and dyslexia. *Biological Psychiatry*, *68*, 320–328.
136. Loi, E., Moi, L., Blois, S., Bacchelli, E., Vega Benedetti, A. F., Cameli, C., Fadda, R., Maestrini, E., Carta, M., Doneddu, G., & Zavattari, P. (2020). ELMOD3-SH2D6 gene fusion as a possible co-star actor in autism spectrum disorder scenario. *Journal of Cellular and Molecular Medicine*, *24*, 2064–2069.
137. Yue, Y., Grossmann, B., Holder, S. E., & Haaf, T. (2005). De novo t(7;10)(q33;q23) translocation and closely juxtaposed microdeletion in a patient with macrocephaly and developmental delay. *Human Genetics*, *117*, 1–8.
138. Backx, L., Seuntjens, E., Devriendt, K., Vermeesch, J., & Van Esch, H. (2011). A balanced translocation t(6;14)(q25.3;q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum. *Cytogenetic and Genome Research*, *132*, 135–143.
139. Córdova-Fletes, C., Domínguez, M. G., Delint-Ramirez, I., Martínez-Rodríguez, H. G., Rivas-Estilla, A. M., Barros-Núñez, P., Ortiz-López, R., & Neira, V. A. (2015). A de novo t(10;19)(q22.3;q13.33) leads to ZMIZ1/PRR12 reciprocal fusion transcripts in a girl with intellectual disability and neuropsychiatric alterations. *Neurogenetics*, *16*, 287–298.
140. Moysés-Oliveira, M., Guilherme, R. S., Meloni, V. A., Di Battista, A., De Mello, C. B., Bragagnolo, S., Moretti-Ferreira, D., Kosyakova, N., Liehr, T., Carvalheira, G. M., & Melaragno, M. I. (2015). X-linked intellectual disability related genes disrupted by balanced X-autosome translocations. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *168*, 669–677.
141. Hackmann, K., Matko, S., Gerlach, E.-M., Von Der Hagen, M., Klink, B., Schrock, E., Rump, A., & Di Donato, N. (2013). Partial deletion of GLRB and GRIA2 in a patient with intellectual disability. *European Journal of Human Genetics*, *21*, 112–114.
142. Mayo, S., Monfort, S., Roselló, M., Orellana, C., Oltra, S., Caro-Llopis, A., & Martínez, F. (2017). Chimeric genes in deletions and duplications associated with intellectual disability. *International Journal of Genomics*, *2017*, 1.
143. Brewer, M. H., Chaudhry, R., Qi, J., Kidambi, A., Drew, A. P., Menezes, M. P., Ryan, M. M., Farrar, M. A., Mowat, D., Subramanian, G. M., Young, H. K., Zuchner, S., Reddel, S. W., Nicholson, G. A., & Kennerson, M. L. (2016). Whole genome sequencing identifies a 78 kb insertion from chromosome 8 as the cause of charcot-marie-tooth neuropathy CMTX3. *Plos Genetics*, *12*, e1006177.
144. Hoogstrate, Y., Komor, M. A., Böttcher, R., Van Riet, J., Van De Werken, H. J. G., Van Lieshout, S., Hoffmann, R., Van Den Broek, E., Bolijn, A. S., Dits, N., Sie, D., Van Der Meer, D., Pepers, F., Bangma, C. H., Van Leenders, G. J. L. H., Smid, M., French, P. J., Martens, J. W. M., Van Workum, W., … Jenster, G. W. (2021). Fusion transcripts and their genomic breakpoints in polyadenylated and ribosomal RNA-minus RNA sequencing data. *Gigascience*, *10*(12), 1–16.

145. Waddell, L. B., Bryen, S. J., Cummings, B. B., Bournazos, A., Evesson, F. J., Joshi, H., Marshall, J. L., Tukiainen, T., Valkanas, E., Weisburd, B., Sadedin, S., Davis, M. R., Faiz, F., Gooding, R., Sandaradura, S. A., O'grady, G. L., Tchan, M. C., Mowat, D. R., Oates, E. C., … Cooper, S. T. (2021). WGS and RNA studies diagnose noncoding DMD variants in males with high creatine kinase. *Neurology Genetics*, *7*, e554.
146. Cmero, M., Schmidt, B., Majewski, I. J., Ekert, P. G., Oshlack, A., & Davidson, N. M. (2021). MINTIE: Identifying novel structural and splice variants in transcriptomes using RNA-seq data. *Genome Biology*, *22*, 296.
147. Chen, X., Zhao, G., Zhong, P., Zhang, M., Chen, R., & Zhang, D. (2020). Chr2 30297612-ALK, a novel intergenic fusion with exon18 of ALK, responds to crizotinib. *Clinical Lung Cancer*, *21*, e564-e6.
148. Annala, M. J., Parker, B. C., Zhang, W., & Nykter, M. (2013). Fusion genes and their discovery using high throughput sequencing. *Cancer Letters*, *340*, 192–200.
149. Yun, J. W., Yang, L., Park, H.-Y., Lee, C.-W., Cha, H., Shin, H.-T., Noh, K.-W., Choi, Y.-L., Park, W.-Y., & Park, P. J. (2020). Dysregulation of cancer genes by recurrent intergenic fusions. *Genome Biology*, *21*, 166.
150. Mcpherson, A., Wu, C., Wyatt, A. W., Shah, S., Collins, C., & Sahinalp, S. C. (2012). nFuse: Discovery of complex genomic rearrangements in cancer using high-throughput sequencing. *Genome Research*, *22*, 2250–2261.
151. Breuza, L., Poux, S., Estreicher, A., Famiglietti, M. L., Magrane, M., Tognolli, M., Bridge, A., Baratin, D., Redaschi, N., & Consortium, U. (2016). *The UniProtKB guide to the human proteome*. Database.
152. Duek, P., Gateau, A., Bairoch, A., & Lane, L. (2018). Exploring the uncharacterized human proteome using neXtProt. *Journal of Proteome Research*, *17*, 4211–4226.
153. Abascal, F., Acosta, R., Addleman, N. J., Adrian, J., Afzal, V., Ai, R., Aken, B., Akiyama, J. A., Jammal, O. A., Amrhein, H., Anderson, S. M., Andrews, G. R., Antoshechkin, I., Ardlie, K. G., Armstrong, J., Astley, M., Banerjee, B., Barkal, A. A., … Weng, Z. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, *583*, 699–710.
154. Consortium EP. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*, 57–74.
155. Panigrahi, A., & O'malley, B W. (2021). Mechanisms of enhancer action: The known and the unknown. *Genome Biology*, *22*, 108.
156. Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, *18*, 83.
157. Shashi, V., Mcconkie-Rosell, A., Rosell, B., Schoch, K., Vellore, K., Mcdonald, M., Jiang, Y.-H., Xie, P., Need, A., & Goldstein, D. B. (2014). The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genetics in Medicine*, *16*, 176–182.
158. Karczewski, K. J., & Snyder, M. P. (2018). Integrative omics for health and disease. *Nature Reviews Genetics*, *19*, 299–310.
159. Scott, A. J., Chiang, C., & Hall, I. M. (2021). Structural variants are a major source of gene expression differences in humans and often affect multiple nearby genes. *Genome Research*, *31*, 2249–2257.
160. Chiang, C., Scott, A. J., Davis, J. R., Tsang, E. K., Li, X., Kim, Y., Hadzic, T., Damani, F. N., Ganel, L., Montgomery, S. B., Battle, A., Conrad, D. F., & Hall, I. M. (2017). The impact of structural variation on human gene expression. *Nature Genetics*, *49*, 692–699.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Allou, L., & Mundlos, S. (2023). Disruption of regulatory domains and novel transcripts as disease-causing mechanisms. *BioEssays*, *45*, e2300010. https://doi.org/10.1002/bies.202300010