Trends in Genetics

CellPress OPEN ACCESS

Forum

Everything, everywhere: FSHD as a model for complex genetic disease

Valentina Salsi [1], Francesca Losi [1], and Rossella Tupler [1,*]

Medical genetics can reveal how genetic variations shape human biology by addressing a critical question: how does a genetic lesion become a phenotype? Facioscapulohumeral muscular dystrophy (FSHD), exemplifies how a seemingly simple genetic lesion can affect multiple layers of cellular regulation, affecting 'everything, everywhere all at once'.

From genetic lesion to distributed regulation

Facioscapulohumeral muscular dystrophy (FSHD) results from alterations in the D4Z4 macrosatellite repeat at the 4q35 subtelomeric region, which lead to aberrant expression of nearby genes. A gene-centric unifying model has been proposed in which either D4Z4 repeat contraction (FSHD1) or mutations in epigenetic regulators such as structural maintenance of chromosomes flexible hinge domain containing 1, (SMCHD1), (FSHD2), converge on a shared mechanism: loss of epigenetic repression at the D4Z4 locus, resulting in the activation of the double homeobox 4, (DUX4) transcription factor and its downstream transcriptional effects [1]. While this framework has been instrumental in explaining the genetic basis of FSHD, DUX4 should not necessarily be regarded as the key player in disease pathogenesis but rather as one of several possible outputs generated by D4Z4 derepression within a broader regulatory system. However, this model fails to explain why patients with similar 4q alleles display markedly divergent clinical phenotypes [2] (Box 1). This mismatch challenges the idea of a simple linear relationship between the primary lesion and the clinical outcome and suggests additional explanations are required. We argue that FSHD is better understood as a propagation problem, in which a localized perturbation is progressively amplified across interconnected regulatory layers. This approach might break new ground in investigating and understanding molecular mechanisms underlying complex genetic diseases.

Genome organization as a propagation layer

The earliest consequence of D4Z4 contraction is not limited to transcription but also involves a reconfiguration of chromatin architecture at the 4q35 subtelomere, consistent with position effects linked to telomeric proximity (telomere position effect (TPE)-like mechanisms) [3]. This includes increased chromatin accessibility (ATAC-seq) and altered long-range interactions and nuclear positioning of 4q35 (4C/Hi-C, DNA fluorescence in situ hybridization (DNA-FISH)), consistent with a reorganization of higher-order chromatin architecture [4,5].

The presence of numerous D4Z4-like sequences in the human genome [6] raises the question of whether this reorganization affects long-range chromatin interactions between repeat-rich regions. Indeed, early DNA-FISH studies of 4q35 architecture demonstrated disease-specific interaction profiles, supporting the idea that the locus operates within a nuclear interaction network rather than as an isolated regulatory element [4].

Consistent with this view, genome-wide chromatin accessibility profiling reveals extensive differences in FSHD myoblasts compared with cells from healthy individuals [7]. The widespread changes preferentially affect heterochromatin-associated regions, including lamina-associated regions, echoing earlier data demonstrating that 4q35 displays lamin-dependent organization and preferential localization at the nuclear and nucleolar peripheries [5,8]. Notably, the altered chromatin accessibility in FSHD myoblasts is largely attenuated upon differentiation into myotubes (see Box 2). These data suggest that regulatory events in earlier developmental stages exert their effects later in the course of FSHD development. Interestingly, perturbation of chromatin regulators such as SMCHD1 can be associated with clinically distinct disorders, including FSHD2 and Bosma arhinia microphthalmia syndrome, while D4Z4 hypomethylation has also been reported in immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome, suggesting that similar chromatin alterations may produce different biological outcomes depending on context.

RNA as an amplification layer for regulatory perturbations

In our view, the initial perturbation of genome organization in FSHD is disseminated and amplified by noncoding RNAs (ncRNAs). Multiple ncRNAs are transcribed from the 4q35 loci, including D4Z4 Binding Element Transcript (DBE-T), which acts locally *in cis* to modulate chromatin structure at the D4Z4 locus [9]. The nearby FSHD region gene 2 (FRG2) locus encodes long ncRNAs that function *in trans*, interacting with nuclear compartments and distal chromatin regions [8]. Additionally, the D4Z4 array exhibits asymmetric bidirectional transcription and generates multiple RNA species, including small RNA fragments, indicating that RNA production is an intrinsic feature of the locus [10,11].

In DUX4 overexpression or inducible models, DUX4 activation is associated with a widespread reconfiguration of the transcriptome, including the activation of intergenic loci and the emergence of numerous novel isoforms, many derived from

Box 1. The same disease, really?

Pedigree analyses frequently reveal discordant phenotypes among individuals carrying the same D4Z4 repeat size, ranging from asymptomatic carriers to severely affected individuals, and a mode of inheritance divergent from the expected autosomal-dominant pattern as shown in Figure I. Notably, the FSHD molecular signature [a D4Z4-reduced allele (DRA) combined with a 4qA allele plus a poly(A) signal] is a common polymorphism with a frequency of 3–6%. This is 1000-fold greater than the disease prevalence (5 in 100 000), indicating that other factors must contribute to disease onset.

In addition, besides an identical D4Z4 allele at the 10q telomere, thousands of other D4Z4-like sequences are scattered throughout the human genome. All these repetitive elements are polymorphic in sequence and are present in variable numbers across individuals and populations. Furthermore, myopathic subjects carrying a DRA are clinically heterogeneous. They can be stratified into several clinical categories based on the presence of typical and atypical features (age and time of disease onset, distribution of muscle weakness, satellite/nonmuscular signs [2]) and on different disease trajectories. This heterogeneity can complicate clinical diagnosis and differential diagnosis, occasionally leading to overlap with other neuromuscular disorders and motivating the development of standardized clinical classification systems such as the Comprehensive Clinical Evaluation Form (CCEF).

How many genomic features influence disease appearance?

Understanding which mechanisms underlie these differences is essential for understanding which basic functions are involved in the development of disease and for creating translationally meaningful studies.

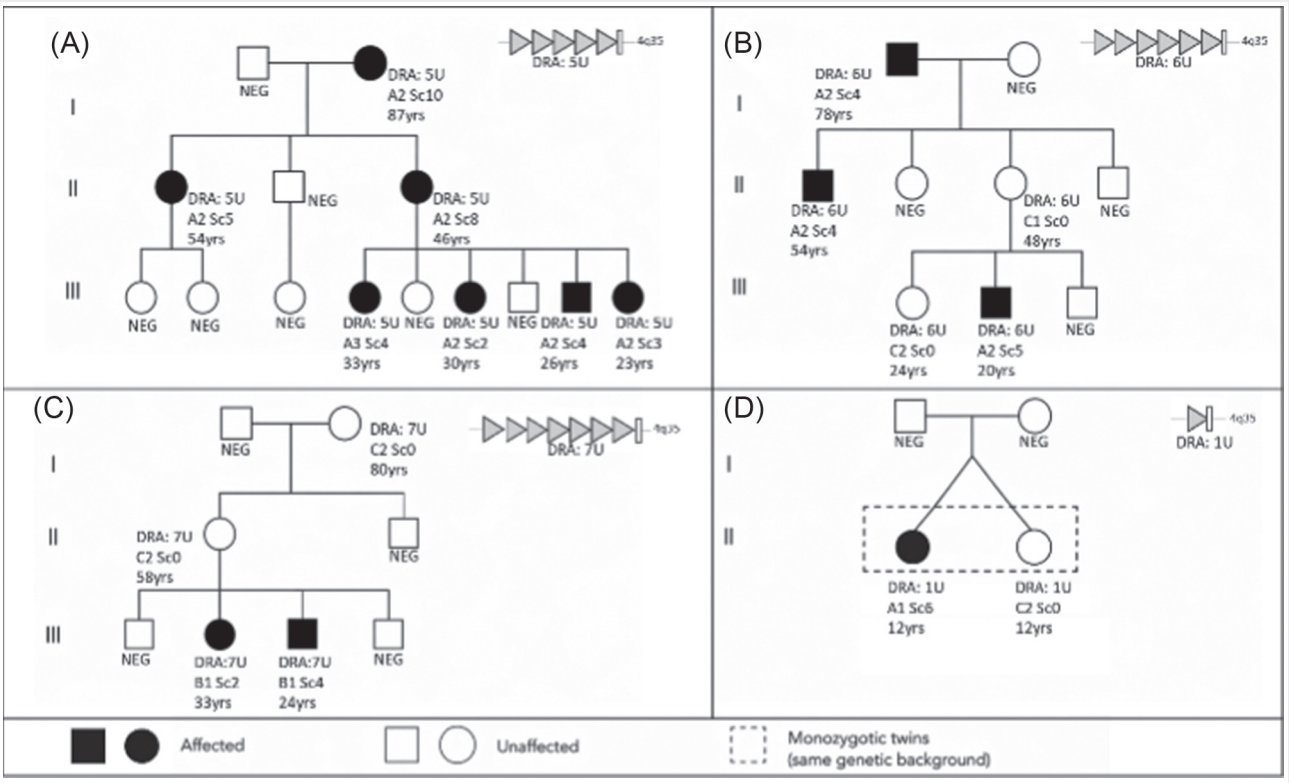

Figure I. Observed phenotypes in pedigrees in which a DRA segregates: (A) full penetrance, (B) reduced penetrance, (C) recessive-like inheritance pattern, and (D) discordant monozygotic twins. NEG: negative.

repeat-rich regions [12]. These changes reflect one layer of regulatory reconfiguration rather than a primary causal axis. DUX4-dependent transcription is therefore embedded within a broader network of interacting regulatory processes, whose combined effects shape disease manifestation.

Concurrently, repeat-derived RNAs such as human satellite II (HSATII) are bidirectionally transcribed and can form double-stranded RNA species that accumulate in intranuclear foci and recruit RNA-binding proteins, contributing to nuclear aggregation and altered RNA processing. Consistent with

Box 2. A tale of two cellular states

Propagation is strongly dependent on cellular context (see Figure I). Proliferating myoblasts provide a permissive state in which multilayer perturbations emerge, interact, and amplify, whereas differentiated myotubes largely reflect their cumulative outcome.

Single-cell and spatial transcriptomics support this view [16], showing that disease-associated programs arise in subsets of cells along divergent trajectories rather than as uniform responses. Proteomic data further indicate that protein-level alterations are established early and are not fully explained by transcription alone [8,14]. The biological landscape is not uniform across muscles. Different muscles differ in developmental history, regenerative activity, and physiological demands, creating distinct contexts in which the same primary perturbation may be amplified, buffered, or redirected to different extents.

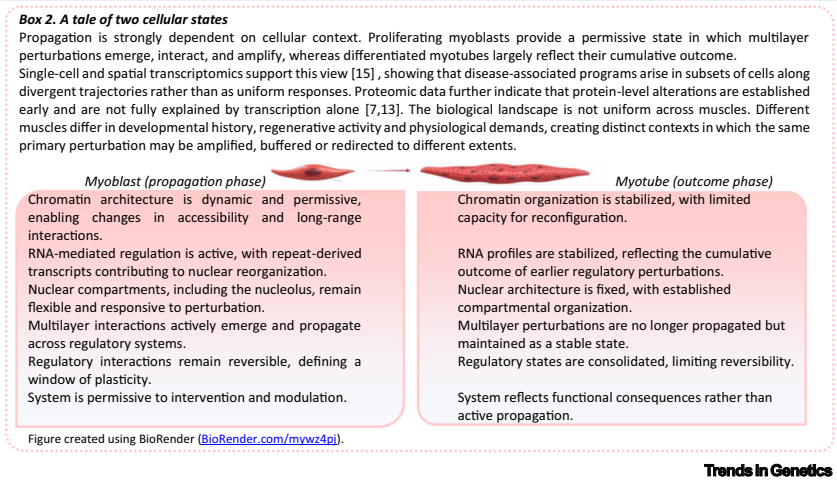

Figure I. Myoblast and myotube regulatory states. The figure was created using BioRender (BioRender.com/mywz4pj).

this, DUX4 expression also induces the accumulation of stable intranuclear RNAs, including nucleolar-associated transcripts, linking RNA aggregation to nucleolar perturbation [13].

In a comprehensive view, these observations show that repeat-derived RNAs introduce new interaction surfaces within the nucleus, enabling the propagation and reinforcement of regulatory perturbations across nuclear systems.

From nucleolar function to translational output

In FSHD, FRG2 long ncRNAs accumulate at the nucleolus, where they alter its organization and activity, establishing a direct mechanistic link between repeat-associated transcription and nucleolar dysfunction [8].

FRG2 transcripts are not passive byproducts of locus derepression but regulatory molecules capable of interacting with nucleolar components and chromatin-associated regions, including ribosomal DNA (rDNA). Their accumulation correlates with altered nucleolar architecture and impaired rRNA transcription, positioning FRG2 as a key mediator that connects upstream genomic perturbation to downstream functional consequences.

Alterations in rDNA activity and rRNA transcription in turn affect ribosome output and, consequently, the translational capacity of the cell. Importantly, this does not result in a uniform reduction of protein synthesis but in a redistribution of translational output, with selective effects on specific mRNA subsets.

This is consistent with proteomic studies in FSHD myoblasts, which reveal substantial changes in protein abundance that cannot be fully explained by transcription alone, including alterations in mitochondrial and ribosomal protein-related pathways [8,14]. Consistent with these observations, metabolomic studies have reported alterations in energy metabolism, mitochondrial function, oxidative stress, inflammation, and autophagy, although no consensus metabolic signature has emerged. This variability may reflect the fact that metabolic phenotypes arise from the cumulative effects of multiple intermediate regulatory layers rather than directly from the primary genetic lesion.

At this stage, the initial perturbation is converted into a functional output: altered protein production. Thus, the nucleolus operates as a critical bottleneck where distributed regulatory perturbations are functionally integrated and converted into cellular phenotypes. Muscle cells, which rely on tightly regulated translational programs to sustain differentiation and contractile function, are particularly sensitive to this imbalance.

It remains to be clarified whether the heterogeneous composition of ribosomes, in terms of proteins and RNAs, contributes to selective translation.

Individual genomic variability shapes system-level plasticity

The same perturbation does not produce the same outcome because the system in which it operates is heterogeneous. In FSHD, this variability can be illustrated by concrete genetic examples beyond D4Z4 contraction.

Sequencing-based analyses in FSHD1 cohorts indicate that additional genetic variation beyond D4Z4 contraction can contribute to phenotypic variability. Examples include variants in SMCHD1, DNA methyltransferase 3B (DNMT3B)

and ligand-dependent nuclear receptor-interacting factor 1 (LRIF1), as well as 'double trouble' cases involving calpain 3 (CAPN3), lamin A/C (LMNA), valosin-containing protein (VCP), titin (TTN), and other neuromuscular disease genes. In particular, the presence of unrelated genetic variants and comorbidities has been shown to influence disease expression, supporting a model in which distributed genetic factors modulate the clinical outcome rather than by a single secondary locus [15]. Notably, despite more than 2 decades of research, no single modifier has emerged as a universal explanation for FSHD clinical heterogeneity.

This contrasts with FSHD2, in which variants in epigenetic regulators act as primary drivers of disease, whereas in FSHD1 additional variants appear to modulate rather than define disease expression [1].

Together, these observations indicate that the same D4Z4-associated lesion can lead to different outcomes depending on the broader genomic context in which it is embedded.

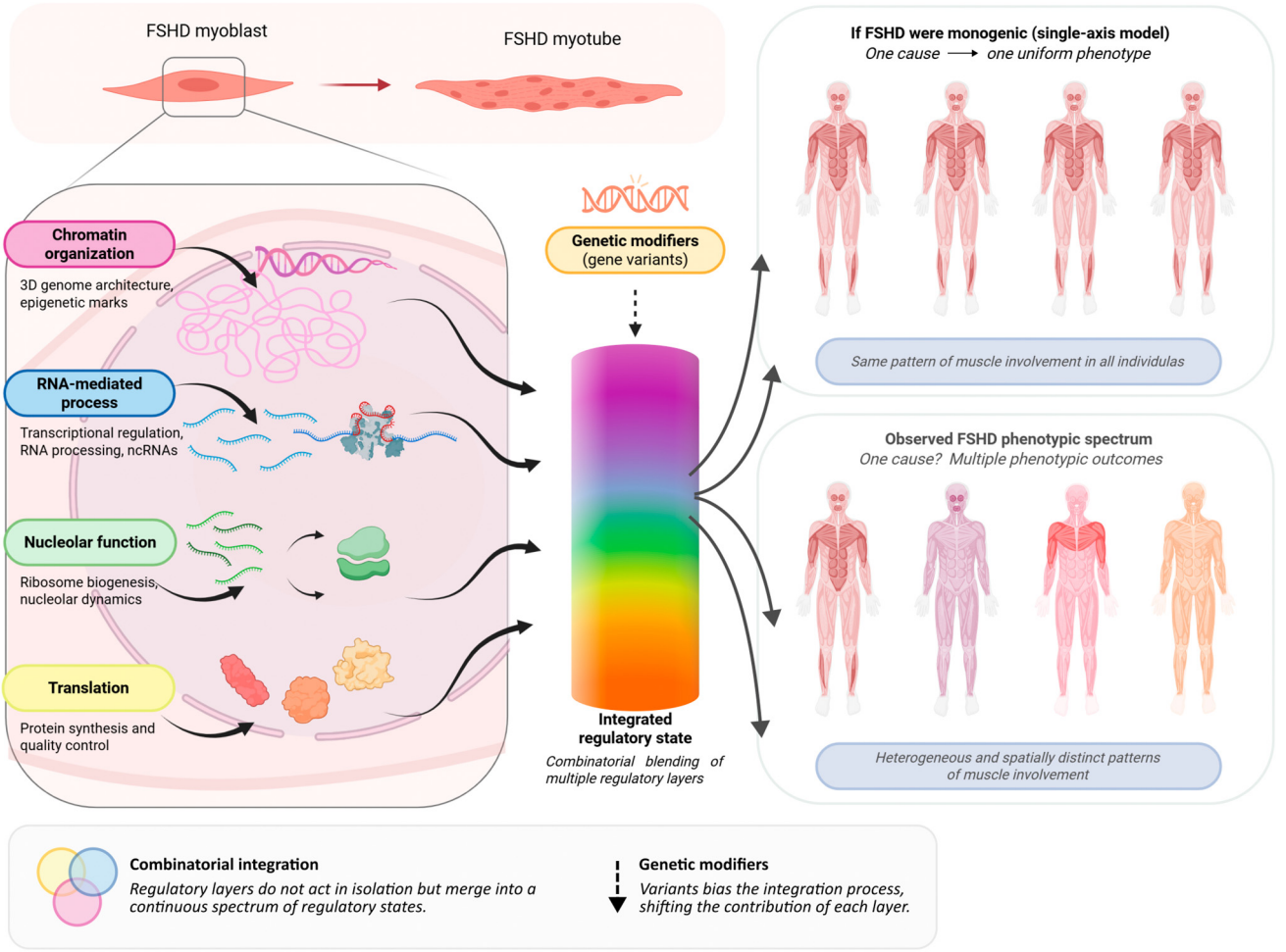

Figure 1. Multilayer propagation of regulatory perturbations in FSHD. In FSHD, each regulatory layer is represented as a distinct color contributing to the system, from chromatin organization and RNA-mediated processes to nucleolar function and translation. These components do not generate independent outputs but progressively merge, producing continuous gradients. Each observed phenotype corresponds to a unique combination of these contributions, where differences in shade reflect differences in how regulatory layers interact and are weighted. Genetic modifiers further bias this mixing process, shifting the balance among layers. Phenotypic variability thus arises not from discrete states but from the combinatorial blending of multilayer regulatory inputs into a spectrum of outcomes. Colors represent the relative contribution of individual regulatory layers (chromatin organization, RNA-mediated processes, nucleolar function, and translation) and their integration into a common regulatory state. The different shades do not correspond to specific disease subtypes but illustrate distinct combinations and relative weights of biological processes arising from a common initiating lesion. The figure was created using BioRender (BioRender.com/mywz4pj). FSHD: facioscapulohumeral muscular dystrophy; ncRNAs: noncoding RNAs.

Consistent with this view, patients with FSHD2 are generally not more severely affected compared to patients with FSHD1, despite often exhibiting greater D4Z4 hypomethylation, suggesting that disease severity cannot be explained by the magnitude of a single molecular alteration.

In this view, the genome is both the source of perturbation and the landscape that determines its propagation, defining a personalized threshold at which multilayered dysregulation translates into disease.

Concluding remarks and future directions

FSHD exposes a key limitation of gene-centered models: they fail to explain how the effects of a single 'causative mutation' can result in phenotypically diverse disease states. The extensive use of DNA sequencing to obtain unequivocal diagnoses has revealed important limitations. Despite major advances, approximately 50–75% of rare diseases remain genetically unresolved. Moreover, even when a causative gene is identified, it often does not fully explain the clinical variability observed among patients, challenging the one-gene, one-disease paradigm.

In our opinion, FSHD offers the possibility of a paradigm shift, as this disease is better understood as a multilayered process in which a localized perturbation propagates across genome organization, RNA-mediated regulation, nuclear architecture, and translational control. In this framework, outlined in Figure 1, disease is not defined by a single molecular event but by how its effects spread across interconnected regulatory systems.

This perspective reframes a central question in medical genetics. Rather than asking which gene is altered, the relevant question becomes how perturbations are propagated, buffered, or amplified across regulatory layers. Disease thus emerges not as a binary consequence of a mutation but as a quantitative property of the system, determined by the balance between perturbation and resilience.

The novelty of this framework is therefore not the identification of an additional pathogenic factor but a shift in the level of interpretation: from asking which molecular event initiates disease to understanding the dynamics of the regulatory system as a whole.

This model generates testable predictions. Disease severity should reflect the extent of propagation across layers rather than the primary lesion alone. Genetic modifiers may shape interactions between layers rather than act on single pathways. Likewise, early and dynamic cellular states represent critical windows in which propagation can still be modulated, with important implications for therapeutic strategies targeting regulatory plasticity rather than fixed endpoints.

Most importantly, the framework predicts that individuals carrying similar D4Z4 alleles but displaying divergent clinical phenotypes cannot remain biologically identical across all regulatory layers. Their trajectories must diverge at one or more downstream levels, even if the specific point of divergence differs among individuals. Identifying where such divergence occurs may provide a mechanistic explanation for incomplete penetrance, intrafamilial variability, and disease progression.

Looking forward, this framework calls for a redefinition of multiomics. Rather than parallel measurements of molecular layers, multiomics should capture how information flows between them. This requires integrating genome-resolved and repeat-aware analyses with state-specific and single-cell approaches, with an emphasis on transitions rather than static profiles.

In parallel, the concept of personalization must evolve. Beyond identifying variants, the goal should be to define how an individual genome constrains or enables the propagation of perturbations. Repeat architecture, ncRNA landscapes, and their impact on nuclear and translational organization, become central to this view. FSHD thus provides a tractable model for a broader principle: phenotype emerges from the interaction between perturbation and system architecture. Understanding disease will therefore require shifting from identifying causes to quantifying how biological systems respond to them.

Acknowledgments

We thank Paul D. Kaufmann for helpful discussions. This work was supported by the Cariplo-Telethon Foundation (GJC23060 to R.T.). The research leading to these results received funding from the European Union—NextGenerationEU through the Italian Ministry of University and Research under PNRR-M4C2-I1.3, Project PE_00000019 'HEAL ITALIA' (CUP E93C22001860006) awarded to R.T. at the University of Modena and Reggio Emilia. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Declaration of interests

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

[1]Department of Biomedical, Metabolic and Neural Sciences University of Modena and Reggio Emilia, 41125 Modena, Italy

*Correspondence:
rossella.tupler@unimore.it (R. Tupler).

https://doi.org/10.1016/j.tig.2026.07.006

References

1. Lemmers, R.J.L.F. *et al.* (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329, 1650–1653
2. Ricci, G. *et al.* (2016) A novel clinical tool to classify facioscapulohumeral muscular dystrophy phenotypes. *J. Neurol.* 263, 1204–1214
3. Robin, J.D. *et al.* (2014) Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. *Genes Dev.* 28, 2464–2476
4. Gaillard, M.-C. *et al.* (2019) Analysis of the 4q35 chromatin organization reveals distinct long-range interactions in patients affected with Facio-Scapulo-Humeral Dystrophy. *Sci. Rep.* 9, 10327
5. Ottaviani, A. *et al.* (2009) The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet.* 5, e1000394
6. Salsi, V. *et al.* (2026) Rethinking genomics of facioscapulohumeral muscular dystrophy in the telomere-to-telomere era: pitfalls in the hidden landscape of D4Z4 repeats. *Eur. J. Hum. Genet.* 34, 357–367
7. Losi, F. *et al.* (2026) Stage-dependent chromatin accessibility remodeling defines an architectural state in facioscapulohumeral muscular dystrophy myoblasts. *bioRxiv* https://doi.org/10.64898/2026.03.06.710110
8. Salsi, V. *et al.* (2025) Nucleolar FRG2 lncRNAs inhibit rRNA transcription and cytoplasmic translation, linking FSHD to dysregulation of muscle-specific protein synthesis. *Nucleic Acids Res.* 53, gkaf643
9. Cabianca, D.S. *et al.* (2012) A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* 149, 819–831
10. Block, G.J. *et al.* (2012) Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One* 7, e35532
11. Salsi, V. *et al.* (2025) Posttranscriptional RNA stabilization of telomeric RNAs FRG2, DBE-T, D4Z4 at human 4q35 in response to genotoxic stress and D4Z4 macrosatellite repeat length. *Clin. Epigenetics* 17, 73
12. Zheng, D. *et al.* (2025) DUX4 activates common and context-specific intergenic transcripts and isoforms. *Sci. Adv.* 11, eadt5356
13. Arends, T. *et al.* (2026) DUX4-induced HSATII RNA accumulation drives protein aggregation, impacting RNA processing pathways. *J. Cell Biol.* 225, e202501129
14. Nishimura, Y. *et al.* (2023) Facioscapulohumeral muscular dystrophy is associated with altered myoblast proteome dynamics. *Mol. Cell. Proteom.* 22, 100605
15. Puma, A. *et al.* (2025) Double trouble: a comprehensive study into unrelated genetic comorbidities in adult patients with facioscapulohumeral muscular dystrophy type I. *Eur. J. Hum. Genet.* 33, 1006–1014
16. Chen, L. *et al.* (2024) Single-cell spatial transcriptomics reveals a dystrophic trajectory following a developmental bifurcation of myoblast cell fates in facioscapulohumeral muscular dystrophy. *Genome Res.* 34, 665–679