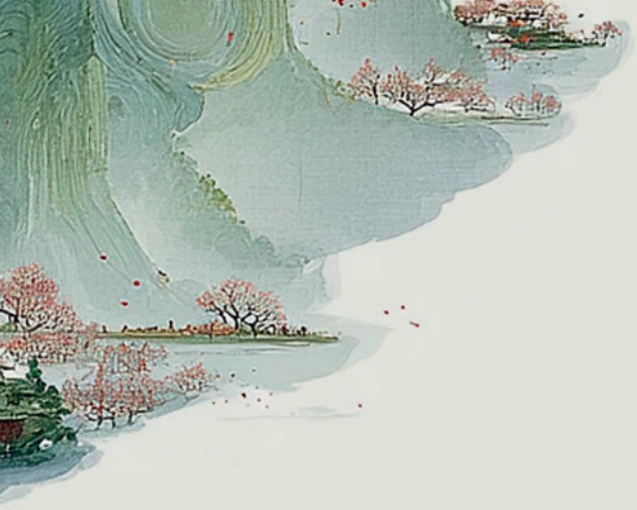
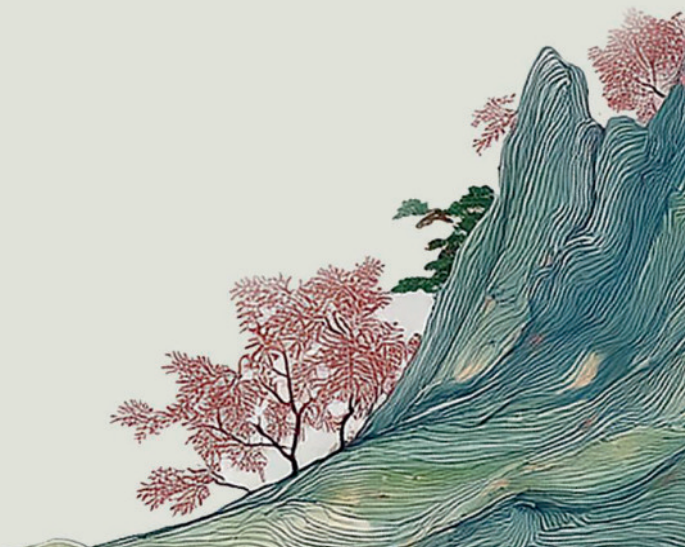
Elucidating DUX4-mediated molecular mechanisms underlying FSHD pathophysiology using multiomics approaches

Zheng, D.



Appendices



English summary

Facioscapulohumeral muscular dystrophy (FSHD) is a progressive muscular dystrophy characterized by often asymmetric muscle weakness and atrophy, primarily in the muscles of the face, shoulders, and upper arms. FSHD is caused by the aberrant expression of DUX4, a germline transcription factor, in skeletal muscle. Under normal physiological conditions, DUX4 is a key driver of zygotic genome activation (ZGA) during early embryonic development, where it plays crucial roles in inducing ZGA gene expression and modulating chromatin accessibility to facilitate the maternal-to-zygotic transition. However, in FSHD patients, the pathological reactivation of DUX4 in mature skeletal muscle triggers extensive downstream cascades that disrupt normal muscle differentiation, regeneration, and repair processes, ultimately leading to cell death and the characteristic FSHD phenotype.

While the role of DUX4 as the primary pathogenic factor in FSHD has been established, how DUX4 orchestrates the complex transcriptional landscape in the myogenic context remains poorly understood. The heterogeneous nature of skeletal muscle and the focal pattern of muscle involvement in FSHD suggest that disease mechanisms operate at multiple cellular and molecular levels that have not been fully characterized. Additionally, the lack of physiologically relevant disease models has limited our ability to translate mechanistic insights into therapeutic strategies.

This thesis addresses these challenges by systematically investigating cellular heterogeneity and DUX4-mediated transcriptional changes in FSHD through complementary approaches. We employed cutting-edge genomic technologies to provide comprehensive insights into disease mechanisms related to DUX4 at unprecedented resolution.

In chapter 2, we used single-nucleus RNA sequencing (snRNA-seq) to explore nuclear heterogeneity within multinucleated myotubes. Through systematic analysis of three FSHD myogenic samples and one control sample, we characterized the transcriptomic heterogeneity of individual nuclei within myotubes. This approach successfully captured DUX4-affected nuclei, which were classified into two distinct populations: DUX4-affected I, characterized primarily by myogenesis inhibition, and DUX4-affected II, marked by upregulation of genes associated with oxidative stress and apoptosis signaling. While these pathways have been identified in previous studies, our data revealed that they originate from distinct nuclear populations, demonstrating the heterogeneous cellular states within nuclei. Further characterization revealed similarities between these affected clusters and early embryonic developmental cells. DUX4-affected I highly expressed markers of 4-cell stage cells, while DUX4-affected II was enriched for morula and blastocyst marker genes. Reciprocally, the marker genes of DUX4-affected clusters were also highly expressed in corresponding developmental stages, revealing that DUX4 activates early embryonic programs in myotube nuclei through processes similar to early embryonic developmental trajectories. The application of the snRNA-seq approach provided systematic insights into nuclear diversity in FSHD primary myotubes at single-nucleus resolution, enhancing our understanding of DUX4-mediated transcriptomic changes in individual nuclei.

In chapter 3, we successfully applied single-fiber RNA sequencing to muscle biopsies, increasing the resolution from bulk tissue to individual myofibers to study transcriptional heterogeneity and fiber-type-specific effects associated with DUX4 activation. This approach enabled accurate myofiber classification and revealed DUX4-signature expression in a small subset of FSHD myofibers. The development of transcriptomic profiling at the single-fiber level demonstrates the potential value of single-fiber transcriptomics for uncovering the molecular heterogeneity of FSHD and provides insights into the challenges of capturing disease-affected fibers in quality-filtered datasets. This technical advancement establishes a foundation for investigating the focal muscle involvement characteristic of FSHD and enables further study of inter-fiber heterogeneity within individual biopsies.

In chapter 4, we established 3D tissue-engineered skeletal muscles (3D-TESMs) by generating genetically matched myogenic progenitors from human induced pluripotent stem cells of three mosaic FSHD patients. These 3D-TESMs successfully recapitulated key pathological features of FSHD, including DUX4 and DUX4 target gene expression, smaller myofiber diameters, and reduced absolute forces upon electrical stimulation. Comprehensive transcriptomic characterization revealed that compared to 2D monolayer cultures, 3D-TESMs demonstrated improved cellular differentiation and elevated expression levels of the core DUX4 target genes. Notably, treatment of 3D-TESMs with three different small molecules previously identified in drug development screens using 2D muscle cultures showed no improvements, and sometimes even declines, in contractile force and sarcomere organization. These results suggest that these compounds either have detrimental effects on 3D-TESM formation or that further refinement of the 3D-TESM model is needed. Overall, we successfully developed a 3D skeletal muscle model for FSHD that provides abundant resources for preclinical research focusing on DUX4 expression and downstream pathways in relationship to contractile properties, establishing a promising platform for future drug screening applications.

In chapter 5, we performed PacBio Iso-seq on DUX4-induced myoblasts. By integrating long-read and short-read RNA sequencing data, we constructed a comprehensive isoform-resolved transcriptome specific to this cellular model with amplified DUX4 downstream cascade signals. Compared to controls, DUX4 overexpression resulted in a more complex transcriptome, characterized by an increase in novel isoforms from annotated loci and intergenic regions, a more complex alternative splicing landscape, and correspondingly more nonsense-mediated decay (NMD) target isoforms. Genes related to RNA processing and RNA metabolism contained more NMD target isoforms which also exhibited significant translational efficiency, suggesting a potential regulatory mechanism by which DUX4 reshapes the transcriptome. Additionally, we discovered that DUX4 target genes display differences in isoform usage under myogenic versus embryonic cellular contexts, indicating cellular context-specific characteristics of DUX4. Specifically expressed isoforms identified in DUX4+ transcriptomes showed excellent differential expression in *in vivo* data, suggesting potential clinical value. The unique advantages of long-read RNA sequencing enabled us to detect and analyze the intergenic isoforms. Most of these intergenic isoforms primarily originated from repetitive elements, with DUX4 binding to LTRs to induce the formation of spliced isoforms containing LTRs and downstream repetitive elements. This phenomenon was also observed in annotated loci, revealing a DUX4-mediated mechanism for increasing isoform diversity. Finally, we

constructed transcriptional and epigenetic profiles of these loci during early embryonic development, demonstrating active transcriptional characteristics during embryonic development. In conclusion, DUX4 induces extensive, profound, and complex transcriptomic changes across different cell lines. Our isoform-level characterization of these changes provides a foundation for understanding both the pathological and physiological roles of DUX4.

This comprehensive study advances our understanding of DUX4-mediated transcriptional changes and FSHD pathogenesis through multi-dimensional analysis of cellular heterogeneity and transcriptomic complexity, providing valuable insights for future therapeutic development and disease mechanism elucidation.