The Diagnostic Role of Next-Generation Sequencing on Facioscapulohumeral Muscular Dystrophy (FSHD) and Related Phenotypes

by

Beyza Yavuzcan

A Dissertation Submitted to the
Graduate School of Health Sciences
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Cellular and Molecular Medicine

KOÇ
ÜNİVERSİTESİ

February, 2025

The Diagnostic Role of Next-Generation Sequencing on Facioscapulohumeral Muscular Dystrophy (FSHD) and Related Phenotypes

Koç University

Graduate School of Health Sciences

This is to certify that I have examined this copy of a master's thesis by

Beyza Yavuzcan

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final examining committee have been made.

Committee Members:

Assist. Prof. Umut Altunoğlu (Advisor)

Prof. Hülya Kayserili

Prof. Zehra Piraye Oflazer

Prof. Zehra Oya Uyguner

Prof. Sibel Aylin Uğur İşeri

Date: 27.02.2025

To my lovely parents

ABSTRACT

The Diagnostic Role of Next-Generation Sequencing on Facioscapulohumeral Muscular Dystrophy (FSHD) and Related Phenotypes

Beyza Yavuzcan

Master of Science in Cellular and Molecular Medicine

February 27, 2025

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common genetic neuromuscular disorder exhibited by progressive muscle weakness, primarily affecting the face, scapular, and especially upper arm muscles. In approximately 95% of cases, classified as type 1 (FSHD1), the disease results from contractions of the D4Z4 macrosatellite repeats on chromosome 4q35. The remaining cases, classified as type 2 (FSHD2), are primarily caused by pathogenic variants in *SMCHD1*, with a smaller subset involving *DNMT3B and LRIF1* genes. These variants not only lead to D4Z4 hypomethylation and aberrant DUX4 activation but may also influence the disease severity in FSHD1 patients with borderline D4Z4 repeat contractions.

The diagnostic algorithm of FSHD involves the examination of D4Z4 Repeat Units (RU) and the analysis of the FSHD-associated genes to confirm or exclude diagnosis. The Molecular Combing (MC) technique has been a crucial tool for the precise and effective analysis of the D4Z4 region, while NGS provides a broader genetic landscape, specific to FSHD2 cases. This approach enables the identification of variants in *SMCHD1*, *LRIF1*, and *DNMT3B*, as well as the detection of variants in other neuromuscular disease-related genes. It enhances differential diagnosis, prevents misclassification, and improves the interpretation of genotype-phenotype correlations.

A total of 96 unrelated patients' D4Z4 RUs were determined using MC. In this group, 82 patients had contracted D4Z4 RUs (67/82 with pathogenic 1–7 RUs and 15/82 with borderline 8–10 RUs). Of the remaining 14 patients, 12 presented with ≥11 RUs while in the remaining 2 cases, only the 4qB allele was detected. Twelve patients from this uncontracted group were selected for Whole Exome Sequencing (WES) analysis based on their disease severity and complex MC results. One unaffected and two affected relatives of one patient were included in this study group for segregation and confirmation purposes.

Pathogenic variants in *SMCHD1* were detected in three patients confirming the clinical diagnosis of FSHD. Pathogenic and/or likely pathogenic *CAPN3* variants were identified in two patients, leading to a revised diagnosis of Limb-Girdle Muscular Dystrophy 2A (LGMD2A). In the remaining three patients, variants of unknown significance in *MYH2*, *TK2*, and *FHL1* genes were identified.

The results of this study largely align with the literature regarding the prevalence of contracted D4Z4 alleles in patients with a clinical diagnosis of FSHD and underscore the significance of NGS for ascertaining a definitive diagnosis in the small subset without contracted D4Z4 alleles. It is crucial to explore the impact of NGS to reveal overlapping genetic factors that may contribute to the clinical heterogeneity of FSHD and its differential diagnoses.

ÖZETÇE

Fasiyoskapulohumeral Musküler Distrofi (FSHD) ve İlgili Fenotiplerde Yeni Nesil Dizilemenin Tanısal Rolü

Beyza Yavuzcan

Hücresel ve Moleküler Tıp, Yüksek Lisans

February 27, 2025

Fasiyoskapulohumeral müsküler distrofi (FSHD), yüz, omuz ve üst kol kaslarını etkileyen ilerleyici kas güçsüzlüğü ile karakterize, en sık görülen üçüncü genetik nöromusküler hastalıktır. FSHD tip 1 (FSHD1) olarak sınıflandırılan vakaların yaklaşık %95'inde hastalık, 4q35 kromozomu üzerindeki D4Z4 tekrar bölgesinin kısalmasından kaynaklanır. Geri kalan vakalar, FSHD tip 2 (FSHD2) olarak sınıflandırılır ve bunlar temel olarak *SMCHD1* genindeki patojenik varyantlardan, daha küçük bir alt küme ise *LRIF1* ve *DNMT3B* genindeki varyantlardan kaynaklanır. Bu varyantlar, sadece D4Z4 hipometilasyonuna ve anormal DUX4 aktivasyonuna yol açmakla kalmaz, aynı zamanda sınırda D4Z4 tekrar kısalması olan FSHD1 hastalarında fenotipi de etkileyebilir. FSHD'nin tanı algoritması, tanıyı doğrulamak veya dışlamak ve ayırıcı tanı gerektiren vakaları ayırt etmek için D4Z4 tekrar ünitesi analizini ve FSHD ile ilişkili genlerin analizini içerir. Moleküler taraklama (MT) tekniği, D4Z4 tekrar kısalmalarının hassas ve etkili analizi için kritik bir araç olmuştur. Öte yandan, NGS bu yaklaşımı tamamlayarak özellikle FSHD2 vakalarına özgü daha geniş bir genetik perspektif sunar. Bu yaklaşım, hastalık belirtilerini ve şiddetini etkileyebilecek *SMCHD1, LRIF1 ve DNMT3B* değişimlerinin yanı sıra nöromüsküler hastalık ilişkili diğer genlerdeki değişimlerin belirlenmesini sağlar. Bu, ayırıcı tanıyı geliştirir, yanlış sınıflandırmayı önler ve genotip-fenotip korelasyonlarının etkin yapılmasını sağlar.

Bu çalışmada, 96 hastanın D4Z4 tekrar üniteleri MT tekniği ile belirlenmiştir. Bu grupta D4Z4 kısalması saptanmadığı veya sınırda tekrar sayısı görüldüğü için FSHD2 şüphesi olan 12 hastaya Tüm Ekzom Sekanslama uygulanmıştır. MT analizi, analiz edilen hastaların 82'sinde kısalmış D4Z4 4qA tekrar ünitesi (1–10 tekrar ünitesi), bunlardan 15'inde sınırda kısalma (8–10 tekrar ünitesi) tespit etmiştir. 12 hastada kısalma gözlenmezken, 2 hastada ise yalnızca 4qB tekrar alleli bulunduğu ortaya çıkmıştır. *SMCHD1*'deki patojenik varyantlar, sınırda gruptan iki hastada ve D4Z4 tekrar ünitesinde kısalma gözlenmeyen gruptan bir hastada olmak üzere üç hastada saptanmıştır. Ek olarak, iki hastada *CAPN3* varyantları tespit edilmiş ve Limb-Girdle Müsküler Distrofi 2A (LGMD2A) tanısını düşündürmüştür. Geri kalan vakalarda genetik bulgular sonuçsuz olsa da 3 vakada *MYH2, TK2, FHL1* genlerinde bulunan belirsiz öneme sahip varyantlar dikkat çekmiştir.

Bulgularımız, klinik FSHD tanısı almış hastalarda kısalmaya uğramış D4Z4 alellerinin yaygınlığı konusunda mevcut bilimsel çalışmalarla büyük ölçüde uyumludur. Bu durum, özellikle D4Z4 alellerinde kısalma görülmeyen küçük bir hasta grubunda kesin tanı için Yeni Nesil Dizileme (NGS) teknolojisinin ne kadar önemli olduğunu vurgulamaktadır. Dahası, NGS'nin FSHD'nin klinik olarak neden bu kadar farklı seyrettiğini (klinik heterojenite) ve FSHD'ye benzeyen diğer hastalıkları (ayırıcı tanı) anlamamıza yardımcı olabilecek örtüşen genetik faktörleri ortaya çıkarma potansiyelini araştırmanın kritik önem taşıdığını belirtmektedir.

ACKNOWLEDGEMENTS

First of all, I sincerely thank all the FSHD patients and their families for their participation and support.

I am especially grateful to my PI, Dr. Serpil Eraslan for her invaluable guidance throughout my study and to my advisor Assist. Prof. Umut Altunoğlu for his constant support and mentorship.

I appreciate Dr. Şahin Avcı for his contributions to patient care and recruitment. I also extend my gratitude to Prof. Z. Piraye Oflazer and the KUH Muscle Diseases Center team for their support in providing patient data and resources.

A special thanks to my dear friend, Manar Kaptan, with whom I worked closely throughout this project for all her efforts and support.

I am deeply grateful to Prof. Hülya Kayserili Karabey, whose vision brought together all the essential pieces—and people—that made this project possible. Her leadership not only shaped the foundation of this work but also created an inspiring and supportive environment at the Genetic Diseases Evaluation Center, where I had the privilege of working alongside wonderful colleagues throughout my time at KUH.

Lastly, I want to express my deepest gratitude to my family and all my friends for their unwavering emotional support throughout this journey.

Funding

This project was funded by TUBITAK EJP-RD-JTC2020 (Grant No: 121N274) research fund. Our sincere thanks to Koç University School of Medicine for their support during the establishment of our FSHD Diagnosis & Research Unit.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ACMG	American College of Medical Genetics and Genomics
BAMS	Bosma Arhinia Microphthalmia Syndrome
CK	Creatine Kinase
CNV	Copy number variations
CSS	Clinical Severity Scale
FSHD	Facioscapulohumeral Muscular Dystrophy
ddPCR	droplet digital PCR
DRA	D4Z4 Reduced Allele
DMD	Duchenne Muscular Dystrophy
EDMD	Emery-Dreifuss Muscular Dystrophy
EMG	Electromyography
ERV	Endogenous Retroviral Elements
FSHD-HI	FSHD Health Index
FSHD-CSS	FSHD Clinical Severity Score
HMW	High Molecular Weight
ICF	Immunodeficiency-Centromeric Instability-Facial Anomailes
LGMD	Limb-Girdle Muscular Dystrophy
LOVD	Leiden Open Variation Database
MAF	Mean Allele Frequency
MC	Molecular Combing
MD	Muscular Dystrophy
MLPA	Multiplex Ligation-dependent Probe Amplification
MW	Moleculer Weight
n.a	Non applicable
NGS	Next-Generation Sequencing
OGM	Optical Genome Mapping
PAS	Polyadenylation Site

RU	Repeat Unit
SB	Southern Blot
SNV	Single Nucleotide Variations
UBL	Ubiquitin-like
VUS	Variant Uncertain Significance
WBC	White Blood Cells
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Chapter 1:
INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) represents the third most frequently diagnosed form of muscular dystrophy, affecting an estimated 1 in 8,000 to 20,000 people (Statland & Tawil, 2014). This condition is defined by a gradual decline and weakening of muscle tissue. Muscle deterioration is a unifying feature of all muscular dystrophies but some specific types, including FSHD, are distinguished by unique genetic origins, molecular mechanisms, and clinical presentations (Tawil & van der Maarel, 2006). The nomenclature of FSHD arises from its defining feature: the predominant weakness affecting muscles in the face, shoulder, and upper arm regions, corresponding to the 'facio-', 'scapular-', and 'humeral' components of the name. The disorder is primarily attributed to genetic variants leading to the aberrant expression of the *DUX4* gene, which under normal conditions remains epigenetically silenced in somatic muscle tissue (Lemmers et al., 2012). Muscle cell apoptosis and progressive muscular atrophy are the results of a cytotoxic cascade initiated by the misregulated activation of *DUX4* (Banerji et al., 2020).

Clinical presentation of FSHD1 and FSHD2 are nearly indistinguishable, making molecular evaluation indispensable for accurate classification, as these subtypes stem from distinct genetic mechanisms that impact disease severity and inheritance patterns (Lemmers et al., 2012). FSHD1 is caused by D4Z4 repeat contractions, while FSHD2 results from epigenetic dysregulation due to pathogenic variations in chromatin-modifying genes (Hartweck et al., 2013). Clinical features such as progressive muscle weakness and scapular winging can lead to misdiagnosis, because these represent overlapping clinical features with several other muscular dystrophies, such as limb-girdle muscular dystrophies (LGMD), myotonic dystrophy, and congenital myopathies (Statland & Tawil, 2014).

Muscular dystrophies (MDs) encompass a heterogeneous group of monogenic neuromuscular disorders characterized by progressive muscle degeneration, weakness, and loss of function (Wang et al., 2020). These conditions arise from pathogenic variants in genes encoding structural or functional muscle proteins,

leading to muscle fiber degeneration and impaired musculoskeletal integrity (Statland et al., 2015). The clinical spectrum of muscular dystrophies is broad, encompassing varying degrees of muscle weakness, contractures, loss of mobility, and cardiopulmonary complications (Tawil & Statland, 2018). The severity, rate of progression, and pattern of muscle involvement are largely dictated by the underlying genetic etiology and age of onset. Despite being classified within the broader spectrum of muscular dystrophies, FSHD exhibits unique pathogenic mechanisms and clinical presentations, necessitating precise molecular differentiation from other dystrophic conditions. This distinction is particularly crucial for patients with ambiguous genetic or clinical profiles (Sacconi et al., 2019).

As precision therapies continue to advance, proper molecular classification of FSHD subtypes and differential diagnosis of other muscular dystrophies will become increasingly important for personalized treatment approaches and genetic counseling (Sacconi et al., 2019). Molecular classification plays a critical role in gene therapy research, influencing patient selection, therapeutic targets, trial design, and overall treatment efficacy (Tawil & Statland, 2018). Therefore, molecular genetic testing serves as an essential diagnostic tool, complementing clinical evaluation, assessment of family history and mode of inheritance, and muscle biopsy.

1.1 Clinical background of FSHD

FSHD is characterized by progressive and asymmetric muscle weakness, with symptoms typically manifesting before the age of 20 in approximately 90% of affected individuals (Statland & Tawil, 2014). The pattern of muscle involvement is distinct from other muscular dystrophies, requiring careful clinical evaluation and genetic confirmation. One of the earliest and most characteristic signs of FSHD is scapular winging, which results from preferential weakness of the lower trapezius and serratus anterior muscles. This weakness leads to upward displacement of the scapula during arm elevation, resulting in sloped shoulders, straight clavicles, and pectoral muscle atrophy (Tawil et al., 2014). Facial muscle weakness is another hallmark feature, typically more pronounced in the lower facial muscles than the upper. Common early indicators include difficulty whistling, incomplete eye closure during sleep, and

reduced lip pursing (Statland et al., 20159. These symptoms are clinically important for distinguishing FSHD from other muscular dystrophies, as eye movements, eyelid function, and bulbar muscles remain largely unaffected (Wang et al., 2020). In the upper limbs, muscle involvement follows a distinctive pattern, with the deltoid muscle often spared in the early stages. At the same time, biceps and triceps weakness leads to significant upper arm atrophy (Figure 1.1). In later stages of the disease, the forearm muscles often remain relatively unaffected while other muscles atrophy, leading to a characteristic physique sometimes described as having disproportionately large forearms (Sacconi et al., 2019). Weakness of the abdominal muscles is another defining feature, contributing to lumbar hyperlordosis and abdominal protrusion (Moreira et al., 2017). Beevor's sign, a notable clinical indicator of FSHD, is characterized by an upward shift of the navel when a patient in supine position attempts to flex their neck. This occurs because FSHD selectively affects the lower rectus abdominis muscle (Statland & Tawil, 2014). Due to its high specificity for FSHD, Beevor's sign is often employed as a practical diagnostic marker during clinical examinations.

Beyond muscle involvement, FSHD is linked to a range of symptoms outside of the muscular system, such as breathing difficulties, retinal blood vessel abnormalities, auditory impairment, discomfort, and exhaustion. While uncommon, respiratory impairment can occur in advanced cases, particularly in patients with severe proximal muscle weakness or spinal deformities. Restrictive lung disease has been observed in 38% of cases, primarily due to expiratory muscle weakness (Moreira et al., 2017). Reports indicate the occurrence of sleep-disordered breathing in FSHD, encompassing conditions like hypoventilation syndrome and obstructive sleep apnea. The association of these respiratory disturbances to disease stage is yet uncertain.

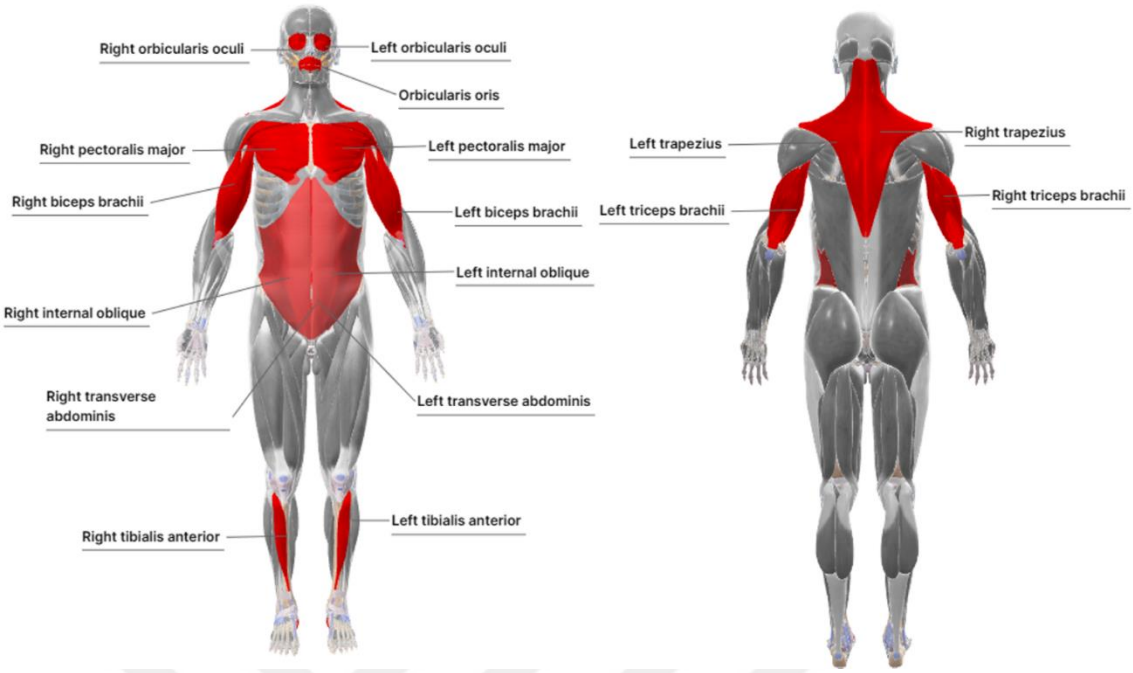

Figure 1.1: The illustration of the muscles commonly affected by FSHD.

Retinal changes and hearing impairment are more prevalent in early-onset cases and chronic musculoskeletal pain and fatigue significantly impact the quality of life, with fatigue largely resulting from postural compensation rather than primary metabolic dysfunction (Padberg et al., 1995; Lutz et al., 2013).

Gender-related differences in disease severity have also been observed, with males generally exhibiting a more severe phenotype than females, despite carrying the same genetic defect (Tawil et al., 2014). Females tend to have a later onset and a milder clinical course, possibly due to hormonal influences and epigenetic modifications, including higher baseline levels of D4Z4 methylation (Scionti et al., 2012). However, a subset of female patients can experience severe manifestations, particularly in cases where FSHD2-associated *SMCHD1* variants further exacerbate chromatin relaxation and *DUX4* misexpression (Sacconi et al., 2019). These observations suggest that although gender does not determine FSHD inheritance, it may influence clinical variability, disease progression, and therapeutic considerations.

1.2 Genetic Background of FSHD

FSHD is a genetically heterogeneous disorder associated with the cytotoxic overexpression of *DUX4* (*Double Homeobox 4*) in muscle cells. This overexpression results from structural changes in the chromosome 4q35 region or pathogenic variants in chromatin-modifying genes, primarily *SMCHD1*. Multiple factors play a role in the pathophysiology of the disease and complicate its genetic diagnosis.

1.2.1 The Role of DUX4

DUX4 is a transcription factor belonging to the homeobox gene family, playing a crucial role in regulating gene expression during early embryonic development (Geng et al., 2012; Snider et al., 2010). It contains two homeodomains that enable DNA binding and the activation of specific target genes (Yao et al., 2014). It is involved in the activation of early developmental and germline genes, including those associated with stem cell regulation, embryonic genome activation, and transposable element regulation (Bosnakovski et al., 2017; Das & Chadwick, 2016). Under normal conditions, *DUX4* expression is tightly regulated and restricted to specific embryonic stages, after which it is epigenetically silenced in somatic tissues (Geng et al., 2012). However, when ectopically expressed, *DUX4* disrupts cellular homeostasis and regulatory networks, leading to the activation of genes involved in apoptosis, oxidative stress, and immune responses, making its regulation critical for normal cellular function (Yao et al., 2014; Snider et al., 2010) (Figure 1.2).

The *MYOD1* and *PAX7* pathways, which are essential for muscle differentiation and satellite cell maintenance, are adversely affected by *DUX4* overexpression (Snider et al., 2010; Dandapat et al., 2014). This results in impaired myogenesis, increased apoptosis, and mitochondrial dysfunction, which are hallmarks of FSHD pathophysiology. Additionally, *DUX4* activates endogenous retroviral elements (ERVs) and LINE-1 transposons, contributing to genomic instability and inflammation, further exacerbating muscle degeneration (Rickard et al., 2015). One of the critical cytotoxic mechanisms of *DUX4* is its ability to induce oxidative stress and DNA damage response pathways, leading to increased susceptibility of myonuclei to apoptosis (Bosnakovski et al., 2008). *DUX4* directly upregulates genes such as *TP53* and *CASP3*, which promote apoptosis while simultaneously downregulating

DNA repair mechanisms (Choi et al., 2016). Moreover, recent studies have demonstrated that *DUX4*-expressing myocytes exhibit mitochondrial dysfunction, leading to an energy-deficient state and increased reactive oxygen species (ROS) accumulation, further contributing to muscle fiber atrophy (Campbell et al., 2018) (Figure 1.2).

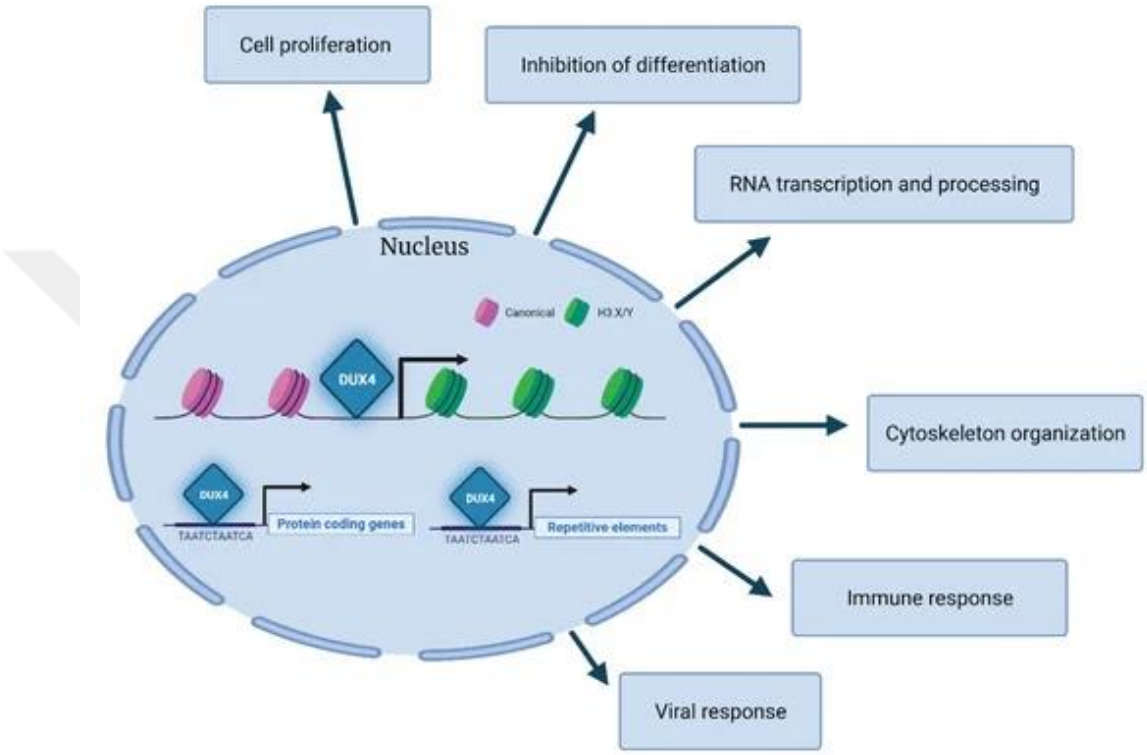

Figure 1.2: Multiple cellular processes regulated by DUX4 transcriptional pathways (Mocciaro et al., 2021).

1.2.2 Underlying Genetic Defect of FSHD1 and FSHD2

The vast majority of FSHD cases (>95%) are classified as FSHD1 and follow an autosomal dominant inheritance pattern. The underlying genetic defect is the contraction of the D4Z4 macrosatellite repeat array on chromosome 4q35. This region is enriched with heterochromatin-associated marks, including H3K9me3 and HP1γ, and is characterized by a high concentration of epigenetic modifications. Each 3.3 kb D4Z4 repeat unit is arranged in a head-to-tail configuration in a continuous sequence without interruptions, and the *DUX4* gene is embedded within this region. In healthy individuals, the number of repeat units (RUs) ranges from 11 to 100. FSHD1 arises due to deletions that reduce the number of tandem repeat units to a pathogenic

threshold of 1–10 RUs. This contraction leads to a loss of heterochromatin structure, resulting in reduced CpG methylation (hypomethylation) and the establishment of a permissive chromatin state, ultimately facilitating the aberrant expression of *DUX4* in skeletal muscle (van Overveld et al., 2003). FSHD2 is a phenotypically similar but genetically distinct form, accounts for ~5% of cases, and is characterized by D4Z4 chromatin relaxation in the absence of repeat contractions (Lemmers et al., 2012) (Figure 1.3). This form arises from pathogenic variations in chromatin-modifying genes, predominantly in *SMCHD1* (Structural Maintenance of Chromosomes Flexible Hinge Domain-Containing Protein 1). These loss-of-function variants contribute to an open chromatin state, *DUX4* transcription even in the absence of repeat contraction (Sacconi et al., 2013, Lemmers et al., 2012; Himeda & Jones, 2019). The loss of CpG methylation at the D4Z4 array facilitates the recruitment of RNA polymerase II, enabling the transcription of the distal permissive *DUX4* gene (Lemmers et al., 2010). The functions of these genes are further evaluated in 1.2.5 Function and Mechanism of the Genes Associated with FSHD2.

The distinct DNA methylation patterns at the D4Z4 repeat region on chromosome 4q35 serve as critical biomarkers for diagnosis and subtype differentiation. In healthy individuals, D4Z4 methylation levels are typically high, ranging from approximately 30% to 65% (Gaillard et al., 2014). FSHD1 patients with a contracted D4Z4 repeat array (1–10 units), display significantly reduced methylation levels, approximately 20% lower than those observed in healthy controls (Hiramuki et al., 2022). FSHD2 patients with variants in *SMCHD1* or *DNMT3B* and no substantial D4Z4 contraction, display even more pronounced hypomethylation, with levels ranging from 5% to 27% (Lemmers et al., 2012). Borderline cases, with D4Z4 repeat units between 8 and 10, present intermediate methylation levels that do not allow for definitively classifying individuals as FSHD or healthy, complicating diagnosis (Sacconi et al., 2015). These methylation thresholds are integral to the epigenetic framework of FSHD, influencing *DUX4* gene expression and contributing to the disease's pathogenesis (van der Maarel et al., 2011).

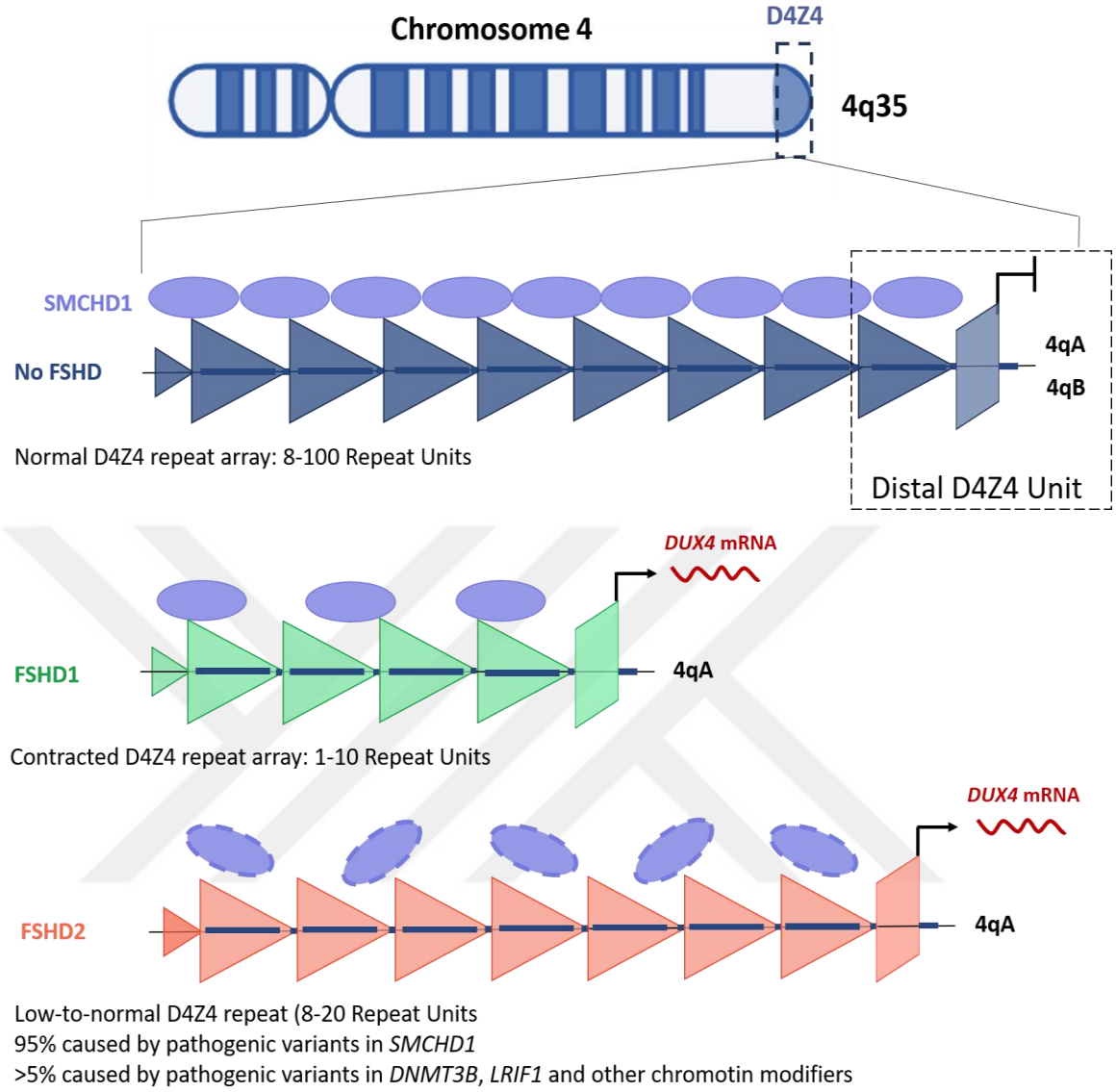

Figure 1.3:Genetic mechanisms of FSHD1 and 2 and DUX4 Activation. D4Z4 region located on 4q. The triangles represent 3,3kb repeat units. *DUX4* activation is due to repeat contractions in D4Z4 for FSHD1 and pathogenic variants in associated chromatin modifier genes in FSHD2 [adapted from (Tihaya et al., 2023)].

1.2.3 Importance of the permissive allele

The *D4Z4* region on chromosome 4q35 exists in two distinct haplotypes, 4qA and 4qB. This is another key element, which plays a crucial role in the pathogenesis of

both FSHD1 and FSHD2. The 4qA allele is essential for disease manifestation and it is defined by the presence of a polyadenylation signal (PAS) situated distal to the D4Z4 region. PAS is critical because it stabilizes DUX4 mRNA, thereby facilitating the production of a functional DUX4 protein (Lemmers et al., 2002; Tawil et al., 2014). In contrast, the non-permissive 4qB haplotype lacks the PAS, preventing DUX4 protein accumulation even in individuals with contracted D4Z4 repeats (Lemmers et al., 2007) (Figure 1.4). Therefore, specifically, the 4qA haplotype is recognized as the pathogenic haplotype

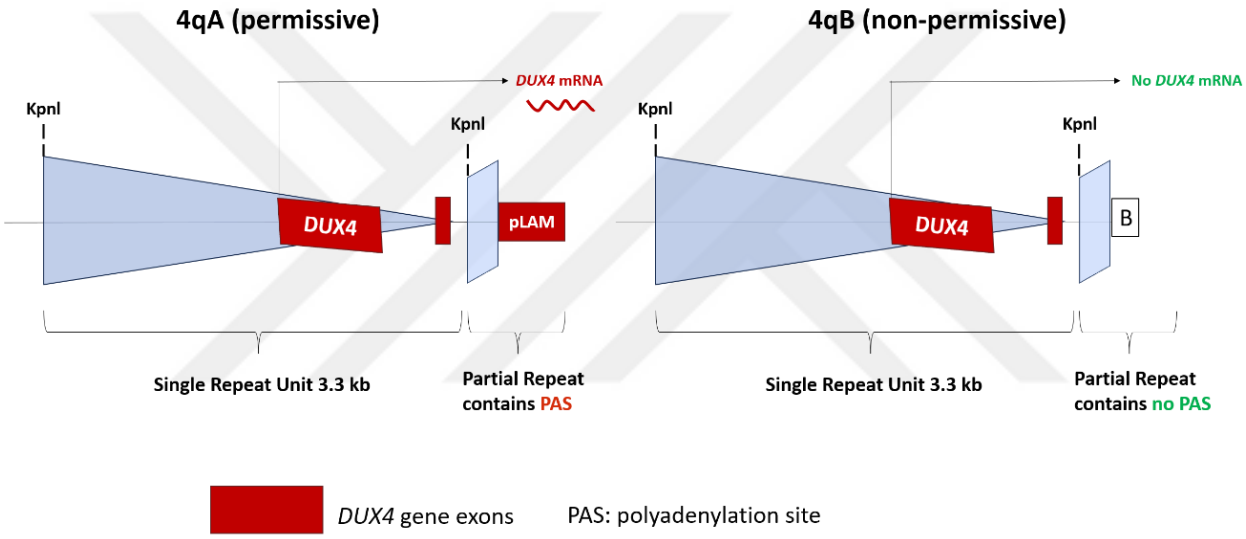

Figure 1.4: Allele composition at the Distal D4Z4 unit and DUX4 Transcription. 4qA haplotype contains the *DUX4* 3'-UTR and pLAM region, which includes the essential polyadenylation signal (PAS) required for mRNA stabilization. [adaped from (Tihaya et al., 2023)].

1.2.4 D4Z4 Homologous Region on 10qA

The D4Z4 macrosatellite repeat array is present on both chromosomes 4q35 and 10q26, sharing almost 95% sequence homology. However, despite their structural similarity, the pathogenic significance of these loci differs. While contractions of D4Z4 on 4q35 under a permissive 4qA haplotype are causative for FSHD1, contractions on 10q26 are generally considered non-pathogenic. However, contractions in 10qA or 10qB haplotypes do not cause FSHD, except in rare complex

chromosomal rearrangements (Lemmers et al., 2007). Like 4q35 locus, the 10q26 locus exists in qA and qB haplotypes. While 10qA contains a PAS in a similar fashion to 4qA, DUX4 expression from 10qA does not contribute to disease, possibly due to differences in chromatin environment or regulatory elements (Figure 1.4). The underlying mechanism for this difference has not been fully understood, but this genomic variability complicates FSHD genetic diagnosis. Additional phenomena comprising gonosomal mosaicism, deletions in probe-binding regions, and chromosomal rearrangements may further complicate molecular testing (van Overveld et al., 2003).

1.2.5 Function and Mechanism of the Genes Associated with FSHD2

SMCHD1

The *SMCHD1* gene encodes a chromatin-modifying protein with a MW of 226 kDa, essential for epigenetic gene silencing by regulating DNA methylation and chromatin conformation (Gendrel et al., 2013). SMCHD1 as a protein plays a key role in FSHD2 and acts as a modifier of disease severity in FSHD1 (Lemmers et al., 2012). Due to its distinct ATPase domain and linear domain architecture, the large and multidomain SMCHD1 protein (*Figure 1.5*) is classified as a non-canonical member of the SMC family, and functions importantly in epigenetic silencing by regulating chromatin architecture (Pedersen et al., 2019). It is involved in maintaining DNA methylation at repetitive sequences, is required for X-chromosome inactivation in females (Blewitt et al., 2008), and suppresses gene transcription through chromatin compaction.

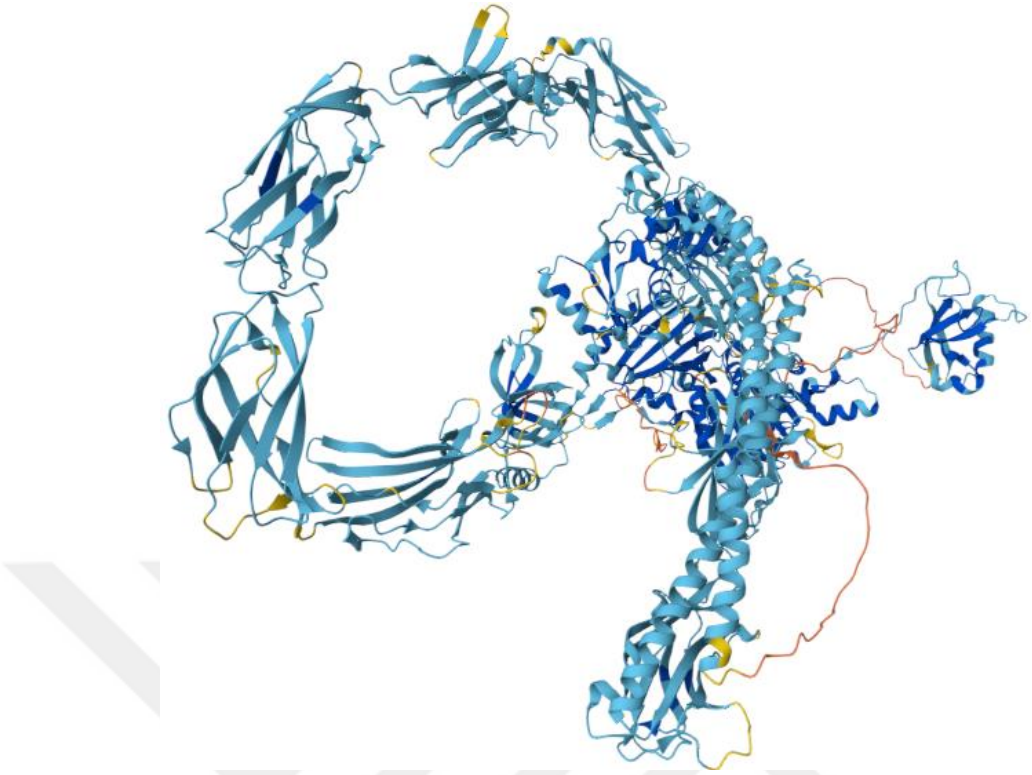

Figure 1.5: Three-dimensional structure of the SMCHD1 protein, as predicted by AlphaFold (accession number A6NHR9 from the AlphaFold Protein Structure Database). It reveals a complex architecture, rendered in shades of blue, showcasing the intricate folding patterns of the polypeptide chain with visibly intertwined regions of alpha-helices and beta-sheets, typical secondary structural elements forming a multi-domain organization.

As a chromatin remodeler, SMCHD1 contributes to repressive chromatin maintenance at various genomic loci, including the D4Z4 repeat array. SMCHD1 ensures that *DUX4* remains epigenetically silenced in healthy individuals (Sacconi et al., 2013). Pathogenic *SMCHD1* variants lead to D4Z4 hypomethylation, facilitating a permissive chromatin state that leads to misexpression of DUX4, even without D4Z4 repeat contractions. These variants cause FSHD2, a clinically indistinguishable but genetically distinct subtype of the disease.

FSHD2 is typically caused by single-base deletions, insertions, missense or nonsense variants, splice-site variants, and deep intronic variants in *SMCHD1*, as well as exonic or whole gene deletions. Despite possessing a D4Z4 array of normal size, FSHD2 patients exhibit greater D4Z4 hypomethylation than FSHD1 patients, affecting both 4q and 10q loci (Lemmers et al., 2012). Importantly, FSHD2 also

requires a permissive 4qA allele, as DUX4 expression remains dependent on the PAS for transcript stabilization. FSHD2 accounts for ~80–90% of *SMCHD1* variant carriers. Pathogenic *SMCHD1* variants are also detected in ~5% of FSHD1 patients, exacerbating disease severity in what is termed the FSHD1+2 genotype (Sacconi et al., 2019; Larsen et al., 2020). These patients often exhibit an earlier onset, more severe muscle weakness, and faster disease progression compared to those with FSHD1 alone. D4Z4 methylation levels in FSHD1+2 cases are lower than FSHD1 (14–64%) but higher than FSHD2 (5–27%), supporting the role of *SMCHD1* as both a primary causative gene in FSHD2 and a modifier in FSHD1 (Larsen et al., 2020).

Beyond its role in FSHD, *SMCHD1* is also implicated in Bosma Arhinia Microphthalmia Syndrome (BAMS)—a disorder characterized by craniofacial abnormalities due to loss-of-function variants affecting non-repetitive genes (Watson et al., 2019). While *SMCHD1* variants are implicated in both FSHD2 and BAMS, the distribution of these variants differs. In FSHD2, variants are spread across the *SMCHD1* gene, whereas BAMS-associated variants are concentrated in the N-terminal region, particularly within the extended ATPase domain. This clustering in BAMS suggests that altered ATPase activity may be central to BAMS's disease mechanism. Additionally, *SMCHD1* has been linked to tumor suppression, with studies suggesting a role in DNA damage repair and chromatin stability (Jansz et al., 2018).

A key distinction between SMCHD1 and canonical SMC proteins lies in their dimerization. Canonical SMC proteins in mammals form heterodimers, assembling from two different SMC monomers, primarily interacting at their hinge domains and to some extent at their ATPase regions. These SMC heterodimers then form closed ring structures, proposed to entrap DNA, crucial for their roles in chromosome conformation (Figure 1.6).

In contrast, SMCHD1 primarily forms homodimers, complexes made of two identical SMCHD1 protein monomers (Figure 1.7). This homodimerization of SMCHD1 occurs at the C-terminus through its SMC hinge domain, and also, uniquely, at the N-terminal ATPase region via a ubiquitin-like (UBL) domain exchange not found in canonical SMCs. These structural differences in dimerization

mechanisms suggest potential functional divergences between SMCHD1 and canonical SMC proteins.

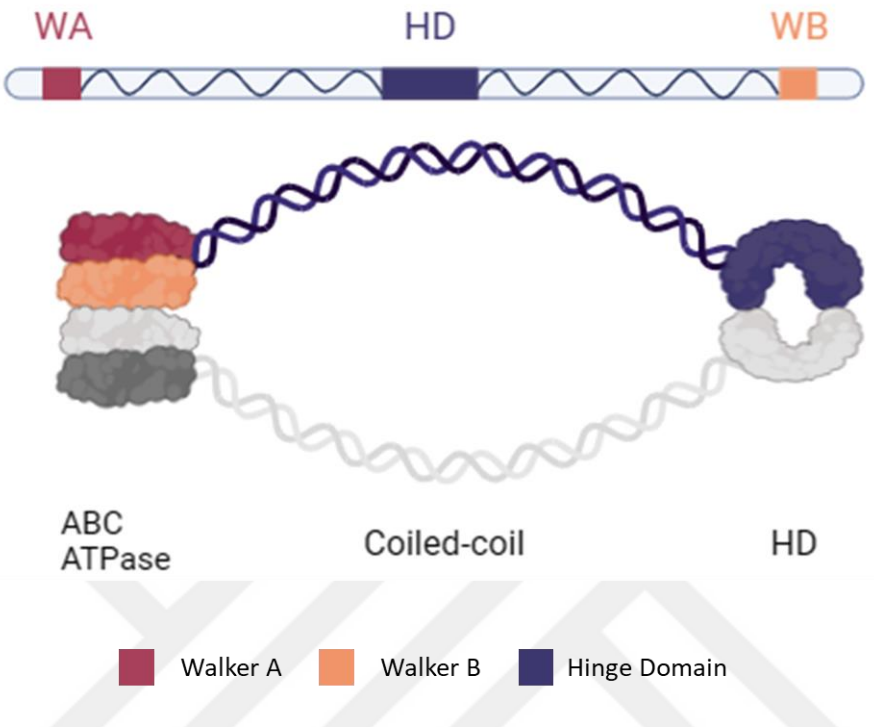

Figure 1.6: Canonical SMC Heterodimer. (adapted from (Gurzau et al., 2020].

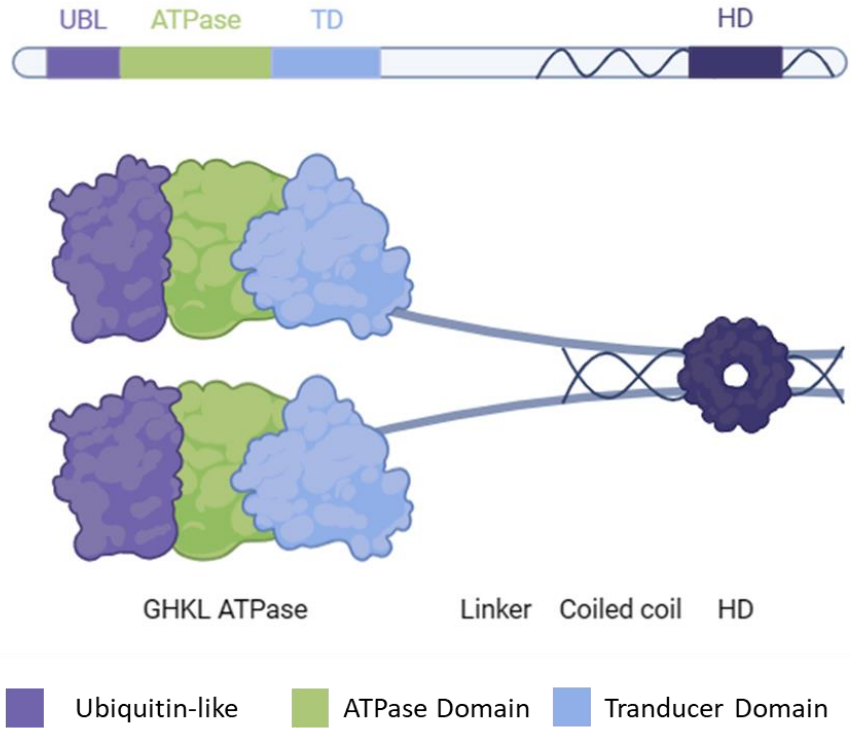

Figure 1.7: Smchd1 Homodimer (adapted from (Gurzau et al., 2020]).

Importantly, the disruption of SMCHD1 homodimerization appears to be more specific to FSHD2 variants (Gurzau et al., 2020). This suggests that impaired homodimerization could be a critical molecular mechanism underlying the phenotypic divergence between FSHD2 and BAMS. ATPase activity alone does not fully explain the distinct clinical presentations, as some disease-causing variants in both conditions lie outside the ATPase domain, and the effects of FSHD2 and BAMS variants on ATPase activity often overlap (Figure 1.8). Intriguingly, identical *SMCHD1* variants have been found in patients with both BAMS and FSHD2, pointing to a more complex picture where the resulting phenotype might be influenced by factors beyond the *SMCHD1* variant itself (Gurzau et al., 2020). Despite ongoing questions, recent discoveries about SMCHD1's dimerization and domain structure are significant for understanding FSHD2 and BAMS

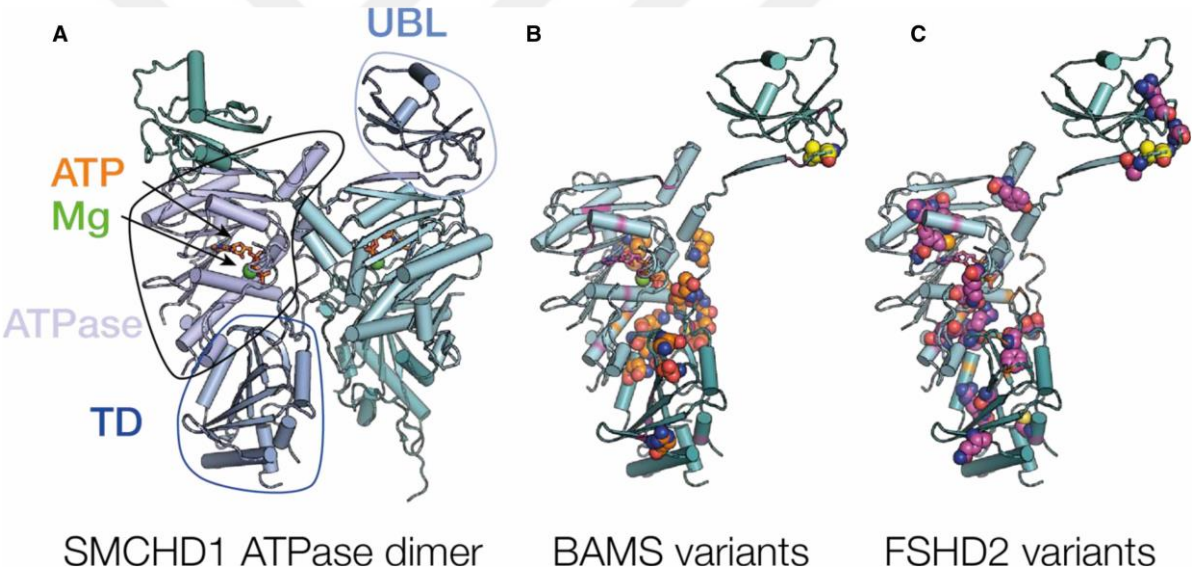

Figure 1.8: The Representation of the SMCHD1 GHKL ATPase domain in BAMS and FSHD (Gurzau et al., 2020). The figure illustrates a strand-swapped ubiquitin-like (UBL) domain, encircled for emphasis, interacting with the GHKL ATPase domain from the opposing monomer. Also highlighted with circles and labels within the same protomer are the ATPase domain itself, the ATP and magnesium (Mg) binding sites, and the transducer domains (TD). (B) and (C) pinpoint the locations of missense variants within the ATPase region. These variants are linked to BAMS (depicted in orange) and FSHD2 (depicted in magenta) and are mapped onto a single monomer of the dimer. Missense variations associated with both BAMS and FSHD2 are indicated in yellow.

DNMT3B

DNMT3B (DNA methyltransferase 3B) is a key epigenetic regulator involved in D4Z4 methylation and *DUX4* repression. While *DNMT3B* variants alone do not cause FSHD, alterations in *DNMT3B*-mediated DNA methylation at D4Z4 contribute to the hypomethylation observed in FSHD2 (Hartweck et al., 2013). *DNMT3B*, along with DNMT3A, plays a crucial role in de novo DNA methylation, ensuring proper gene silencing and chromatin stability during early embryonic development (Okano et al., 1999). In FSHD2, *DNMT3B* dysfunction contributes to global and D4Z4-specific hypomethylation, leading to chromatin relaxation and aberrant *DUX4* activation (van den Boogaard et al., 2016). *DNMT3B* cooperates with *SMCHD1* and *DNMT1*, reinforcing chromatin repression at the D4Z4 locus, though its precise role in FSHD disease severity modulation remains under investigation (Lemmers et al., 2015).

As in the case of *SMCHD1*, *DNMT3B* pathogenic variants are also associated other diseases such as Immunodeficiency-Centromeric Instability-Facial Anomalies (ICF) Syndrome and have been implicated in tumor suppression and chromatin architecture maintenance (Jansz et al., 2018).

Other Genes

Several other genes may influence FSHD progression and severity. *LRIF1*, a chromatin modifier, functions alongside *SMCHD1* to maintain heterochromatin integrity. Variants in *LRIF1* have been linked to D4Z4 hypomethylation, further supporting its role in FSHD2 pathogenesis (van den Boogaard et al., 2016). Additionally, genes such as *TP53*, *MYOD1*, and *PAX7* play critical roles in the FSHD pathomechanism. *DUX4* activation induces the p53 pathway, contributing to apoptotic cell death (Bosnakovski et al., 2008), while *MYOD1* suppression impairs myogenic differentiation, exacerbating muscle atrophy (Snider et al., 2010). *DUX4* also represses *PAX7*, a key regulator of muscle stem cell maintenance, reducing the regenerative capacity of muscle fibers in FSHD (Dandapat et al., 2014). However, while these genes modulate disease severity, none of them have been identified as direct causes of the disease.

FRG1 and *FRG2* were initially believed to be overexpressed and causative were later found to have no direct role in disease progression, with their expression changes

now considered secondary effects of *DUX4* activity (Winokur et al., 2003; Snider et al., 2010) (Figure 1.9). Similarly, *DUX4*C, a homolog of *DUX4*, was initially thought to be involved in FSHD pathogenesis but was later proven to be non-pathogenic (Ansseau et al., 2009). *Titin* (*TTN*), a key sarcomeric protein, was also investigated as a potential contributor but found to have no direct genetic link to FSHD (Tawil et al., 2014).

FAT1's potential role in FSHD lies in its involvement in muscle development and maintenance, as evidenced by mouse models exhibiting FSHD-like muscular defects upon Fat1 disruption (Caruso et al., 2013). *FAT1* variant findings in FSHD-like patients, who lack typical FSHD genetic markers, further strengthen this link (Puppa et al., 2015). Defective *FAT1*, possibly through haploinsufficiency or altered protein function due to splicing abnormalities caused by the identified variants, can contribute to an FSHD-like phenotype. This suggests *FAT1* could be a novel FSHD-related gene, potentially acting as a modifier or a primary causal gene in a subset of FSHD-like cases, highlighting a broader genetic landscape for FSHD-related neuromuscular diseases (Puppa et al., 2015).

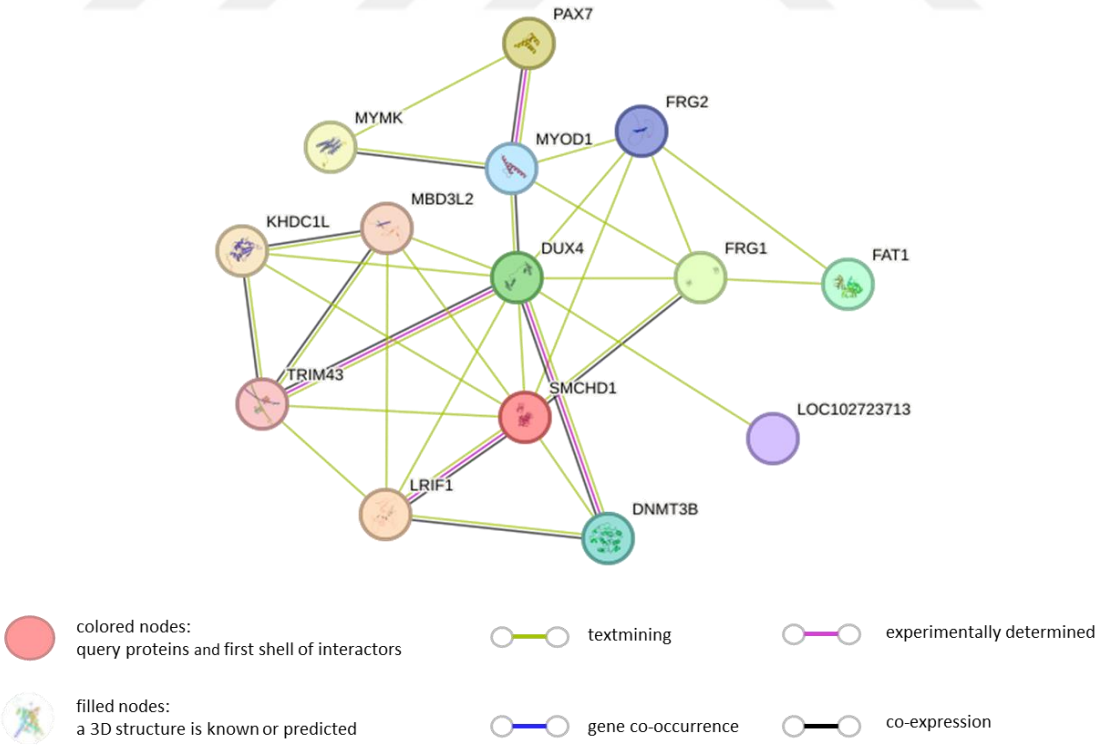

Figure 1.9: STRING Database Analysis of DUX4 Interactions.

These findings emphasize the significant role of epigenetic modifications, especially DNA methylation, in disease pathogenesis, while DUX4 misexpression remains the fundamental driver. Multiple modifiers and regulatory genes influence disease severity, progression, and cellular response, adding further complexity to the genetic landscape of FSHD. Understanding the interplay between *DUX4*, epigenetic regulators, and modifier genes is critical for developing targeted therapeutic approaches aimed at suppressing *DUX4* expression and restoring chromatin repression at the D4Z4 locus.

1.3 Diagnostic Approaches

1.3.1 Clinical Diagnosis

The clinical diagnosis of the disease is based on its hallmark features, including scapular winging, facial weakness, difficulty whistling, and an inability to fully close the eyes during sleep, as well as the extent to which the disease affects the upper and lower limb muscles, often in an asymmetric distribution.

To standardize clinical assessment and clinical scoring criteria, the Ricci Clinical Severity Scale (CSS) has been developed to evaluate muscle involvement across different anatomical regions and quantify functional impairment (Ricci et al., 1999). The CSS assigns scores ranging from 0 (asymptomatic) to 5 (severe disability), reflecting disease progression and muscle weakness severity. Other grading systems, including the FSHD Clinical Severity Score (CSS) and the FSHD Health Index (FSHD-HI), have been introduced to quantify disease burden and patient-reported functional impairment (Mul et al., 2016; Johnson et al., 2021).

Although laboratory tests such as serum creatine kinase (CK) levels and electromyography (EMG) may provide supportive evidence, they are not required for definitive diagnosis. In symptomatic individuals with FSHD, CK levels are often elevated, reaching up to five times the upper limit of normal. However, CK elevation is uncommon in asymptomatic individuals, making it an unreliable standalone biomarker (Tawil & Statland, 2014). Similarly, muscle biopsy is rarely required, as histopathological findings in FSHD are non-specific, often showing variation in fiber size, necrosis, and mild inflammatory infiltrates, which are not exclusive to FSHD (Padberg et al., 1995).

Given clinical variability and phenotypic overlap with other neuromuscular disorders, genetic testing remains essential for diagnostic confirmation, particularly in cases of sporadic disease, early-onset presentations, or atypical disease progression (Lemmers et al., 2012).

1.3.2 Genetic Diagnosis: The tools and diagnosis Algorithm

Genetic diagnosis of FSHD primarily relies on genetic testing, which is considered the gold standard for disease confirmation, particularly in sporadic cases, those with uncertain family history, or when a differential diagnosis is necessary. In familial cases, where autosomal dominant inheritance has already been genetically confirmed in a first-degree relative, additional genetic testing may not always be required, provided that clinical symptoms align with FSHD.

The first step in the molecular diagnosis of FSHD is the assessment of D4Z4 repeat units, as FSHD1 accounts for most cases (Figure 1.10). Historically, Southern blot (SB) analysis, first applied in the early 1990s, was the primary method for detecting D4Z4 repeat contractions (Upadhyaya et al., 1995). This technique involves restriction enzyme digestion of genomic DNA, followed by gel electrophoresis and hybridization with a D4Z4-specific probe, enabling repeat length determination. While highly reliable for non-complex cases, SB is labor-intensive, requires large quantities of high-quality DNA, and lacks high-resolution capabilities. To overcome these limitations, molecular combing (MC) was introduced as a higher-resolution alternative, enabling the direct visualization of individual DNA molecules stretched on a glass slide, hybridized with specific fluorescently labeled probes and allowing for precise measurement of D4Z4 repeat number and haplotype through specialized software analysis (Vasale et al., 2015). Compared to SB, MC offers greater accuracy, higher resolution, and improved efficiency in detecting repeat contractions and haplotypes.

More recently, optical genome mapping (OGM) has emerged as a high-throughput approach, capable of analyzing large structural variations, including D4Z4 repeat contractions, at an unprecedented resolution (Hamanaka et al., 2023). OGM employs labeling and imaging of ultra-long DNA molecules, providing a comprehensive genome-wide structural analysis that may facilitate a more

streamlined and scalable diagnostic workflow for FSHD1. As genomic technologies continue to advance, improvements in repeat sizing, haplotype identification, and genome-wide structural assessment are expected to increase diagnostic accuracy and accessibility, reducing reliance on older, more labor-intensive methods.

These genetic tests are typically performed on high-molecular-weight (HMW) genomic DNA, where the quality of intact long DNA molecules is crucial for reliable results.

If the D4Z4 repeat array is not contracted or has a borderline size, despite a strong clinical suspicion of FSHD, FSHD2 testing is initiated, focusing on methylation analysis of the D4Z4 region and sequencing of *SMCHD1* and *DNMT3B*. With the integration of Next-Generation Sequencing (NGS) methods into clinical practice, the diagnosis of FSHD requires expertise in both conventional Sanger sequencing and emerging genomic techniques. Given the inherent complexity of FSHD diagnosis, only a limited number of specialized laboratories provide comprehensive genetic testing, necessitating training in both classical and advanced methodologies.

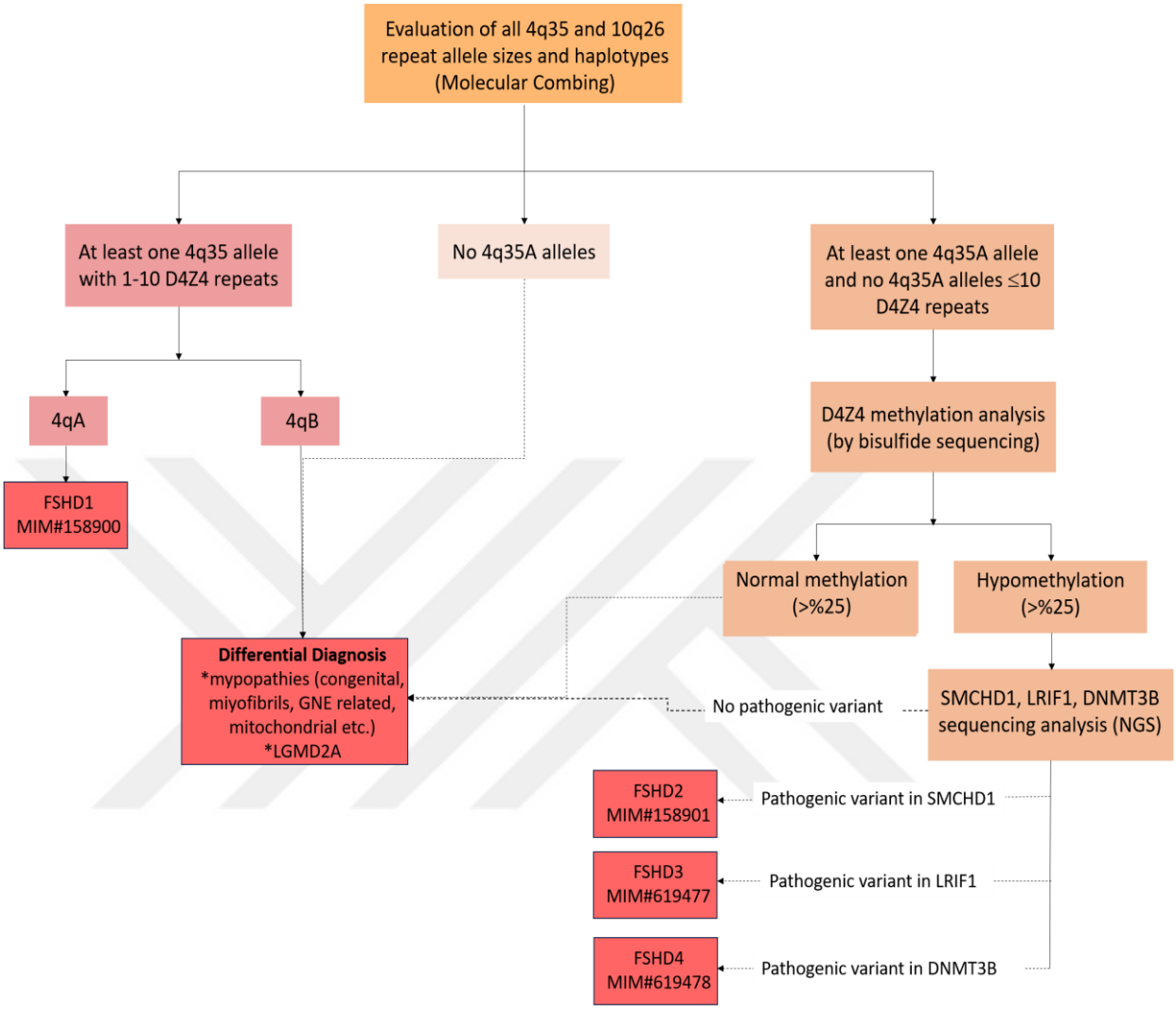

Figure 1.10: Schematic representation of the molecular diagnostic process for FSHD patients.

Also, several techniques have been developed to measure D4Z4 methylation levels, including bisulfite sequencing, methylation-sensitive PCR, and droplet digital PCR (ddPCR), each providing varying levels of sensitivity and resolution. Given that D4Z4 methylation levels correlate with disease severity, methylation analysis serves as both a diagnostic biomarker and a complementary tool to standard genetic testing (Montagnese et al 2023).

1.4 NGS and the Next Era of FSHD Diagnosis

The genetic basis of FSHD has been extensively studied over the past three decades, evolving from early linkage analysis and SB to high-throughput NGS technologies. While the identification of the D4Z4 repeat contraction using SB was a breakthrough in the molecular diagnosis of the disease, early sequencing efforts were limited by technological constraints, as traditional sequencing methods struggled to resolve the large, repetitive D4Z4 region and its complex epigenetic modifications.

Initial studies aimed at uncovering the genetic cause of FSHD focused on linkage analysis in large, affected families, leading to the identification of chromosome 4q35 as the primary locus associated with the disease (Upadhyaya et al., 1995). At that time, sequencing approaches were largely restricted to Sanger sequencing, which, although effective for analyzing small genomic regions, was incapable of characterizing large structural variations such as repeat contractions. As the role of D4Z4 contractions in FSHD1 was well-established by the early 2000s, the underlying molecular mechanism remained unclear. Researchers discovered that not all FSHD cases could be explained by D4Z4 contractions alone, leading to the identification of FSHD2, a phenotypically identical but genetically distinct subtype (Lemmers et al., 2012). Early sequencing studies of genes within the 4q35 region, including *FRG1*, *FRG2*, and *ANT1*, initially suggested their involvement in FSHD pathogenesis but were later ruled out as non-causative (Winokur et al., 2003).

With advancements in sequencing technologies, the discovery of *SMCHD1* as the first major gene implicated in FSHD2 marked a turning point in understanding the epigenetic regulation of D4Z4 (Lemmers et al., 2012). This was followed by the identification of *DNMT3B*, another key chromatin modifier involved in DNA methylation, further expanding the molecular landscape of FSHD2 (Hartweck et al., 2013). These findings underscored that FSHD2 is primarily driven by defects in chromatin-modifying genes rather than repeat contractions, unlike FSHD1.

The introduction of NGS technologies transformed FSHD genetic research by enabling high-throughput, genome-wide analysis, allowing researchers to identify novel variants in FSHD-related genes and genetic modifiers that influence disease severity. Whole-Exome Sequencing (WES) and Whole-Genome Sequencing (WGS)

have been instrumental in detecting rare pathogenic variants in *SMCHD1*, *DNMT3B*, and other chromatin-regulating genes, contributing to disease heterogeneity (Hamanaka et al., 2023).

Over the past three decades, the genetic landscape of FSHD has evolved dramatically, transitioning from early linkage studies and Southern blotting to Next-Generation Sequencing (NGS), long-read sequencing, and genome-wide structural analysis. D4Z4 repeat contraction detection techniques remain indispensable, the integration of NGS-based approaches has revolutionized the field by enabling the identification of modifier gene variants, uncovering epigenetic mechanisms, and refining genotype-phenotype correlations. Moving forward, the convergence of long-read sequencing, methylation profiling, and high-resolution genome mapping holds immense potential for enhancing diagnostic precision, facilitating differential diagnosis, and enabling personalized therapeutic strategies. However, standardization of genetic diagnostic protocols and improving access to advanced sequencing technologies remain critical challenges in clinical implementation. As our understanding of the genetic and epigenetic underpinnings of FSHD continues to expand, cutting-edge genomic innovations will be pivotal in shaping the future of molecular diagnostics, disease classification, and targeted therapies, ultimately improving clinical outcomes for FSHD patients.

1.5 Study Aims and Objectives

This study aims to evaluate the diagnostic utility of NGS in FSHD by:

1. Identifying pathogenic variants in *SMCHD1*, *DNMT3B*, and other potential epigenetic regulators influencing D4Z4 methylation.
2. Assessing the impact of integrating NGS with conventional FSHD1 diagnostic methods to improve differential diagnosis and refine genotype-phenotype correlations.
3. Determining the clinical significance of alternative genetic diagnoses, such as Limb-Girdle Muscular Dystrophy (LGMD) and Myotonic Dystrophy (MD) in patients with unexplained FSHD-like symptoms.

4. Investigating the genetic background and phenotypic variability of FSHD, particularly in borderline D4Z4 repeat contraction cases (8–10 RU), where *SMCHD1* variants may contribute to disease severity.

To achieve the objectives of this study, a cohort of 96 affected individuals was analyzed using a comprehensive diagnostic algorithm. As a first-line test, molecular combing (MC) was employed to assess *D4Z4* repeat units (RUs), identifying contracted *D4Z4* RUs (1–10 RUs) in 82 patients, including 15 cases with borderline repeat sizes (8–10 RUs).

To further investigate selected cases with 6 non-contracted and 6 borderline contracted RUs, WES was performed on a subset of 12 patients. By integrating MC with WES, this study aims to enhance the molecular diagnosis of FSHD, refine genotype-phenotype correlations, and improve the differential diagnosis of FSHD and clinically overlapping neuromuscular disorders.

Chapter 2:
MATERIAL AND METHODS

2.1 Patient Cohort

We have recurited 96 unrelated patients. Informed consent was obtained from all participants before genetic testing and data collection in accordance with the Helsinki Protocol. Initial clinical evaluations and preliminary diagnoses were conducted at the Neurology Department of Koc University Hospital. Following this assessment, patients were referred to the Medical Genetics Department for a comprehensive family history evaluation and molecular investigation. Detailed pedigree analysis was performed to assess inheritance patterns, clinical findings, family history of neuromuscular disorders, and potential autosomal dominant transmission. Based on clinical findings and genetic evaluation criteria, patients were selected for molecular diagnostics, ensuring a standardized and systematic approach to differential diagnosis and genotype-phenotype correlation.

2.2 Sample Collection

Peripheral blood samples (7.5–10 mL) were collected from each participant in K_3EDTA vacutainers for molecular analysis. Samples were either processed immediately or stored at +4°C for a maximum of five days before DNA extraction to ensure sample integrity blood samples were collected from each individual for MC analysis.

2.3 Preparation of DNA Material

2.3.1 For PCR, Sanger Sequencing, and NGS

Genomic DNA was extracted from 2 ml peripheral blood samples using the Qiagen DNeasy Blood & Tissue Kit or QIAGEN QIAamp DNA Blood Mini Kit (Qiagen, USA), following the manufacturer's instructions. Samples were stored at +4°C for short-term use and at -20°C or -80°C for long-term in TE buffer or molecular-grade water to prevent degradation. The quality and quantity of the extracted DNA were assessed using the NanoDrop 2000 spectrophotometer (Thermo

Fisher Scientific, USA), with purity determined by A260/A280=1.8–2.0 and A260/A230=2.0–2.2 absorbance values.

2.3.2 For Molecular Combing

DNA obtained for this method should be High Molecular Weight (HMW) DNA which refers to intact, long DNA fragments, typically ≥150–300 kb, with minimal degradation or shearing. To ensure this FiberPrep® DNA Extraction kit (Genomic Vision, Paris, France) was used. All procedures were followed according to the manufacturer's instructions.

Before the use of FiberPrep®, the total WBC cell count was analyzed by flow cytometry red blood cells were removed and WBC was collected via the following protocol.

a. At least 5 ml blood sample (one volume) was incubated with three volumes of 1xRBC lysis buffer for 10 minutes at room temperature

b. Centrifuged at 600XG for 10 minutes and the supernatant was discarded,

c. 2ml 1X RBC lysis buffer was added to the supernatant and resuspended,

d. Centrifuged at 600XG for 2 minutes and the supernatant was discarded,

e. The pellet consisting of WBC was collected and resuspended in sterile 1xPBS (Biowest, France), ensuring that the cell concentration is 1,2x104-2,4x104 cells/µl.

WBCs were embedded into agarose plugs (5x105-1x106 cells/plug) to protect genomic DNA from mechanical stress. Subsequent steps involved protein digestion and membrane solubilization which releases the DNA into the plug. To liquefy the DNA the agarose was melted and then enzymatically removed, resulting in purified HMW DNA suitable for combing procedures.

2.4 Molecular Combing Procedures

The molecular combing system device, FiberPrep kit, and all MC consumables (FiberProbes, Combi Coverslips, and reservoirs) were supplied by Genomic Vision (France), and all protocols were applied according to the manufacturer's instructions.

The compositions of all in-house buffers and chemicals are detailed in Buffers and Solutions.

Combing

Each sample was placed into special reservoir sized to accommodate the combing coverslips. At this stage, extreme care was taken to avoid agitation, shaking, or vortexing to prevent any damage to the DNA. The DNA was then stretched onto 22 mm² barcoded and silanized coverslips using a Molecular Combing device. The DNA solution was carefully applied to the coverslip, where controlled capillary action and hydrophobic interactions ensured uniform stretching of long DNA molecules along the surface. They were baked at 60°C for 4 hours to crosslink the DNA to the surface, immobilizing the DNA fibers for subsequent analyses.

2.4.1 Hybridization, Wash and Stain

The stretched and immobilized DNA was subjected to FiberProbe FSHD, which contains fluorescein-, digoxigenin-, and biotin-labeled probes designed for the detection of the D4Z4 repeat array and qA and qB haplotypes on both chromosome 4q35 and 10q26 regions. The probe mix was diluted in Formamide (≥%99.0, Sigma-Aldrich®, USA) to facilitate the denaturation of genomic DNA at 95°C for 5 minutes, after which the mixture was applied onto the surface of the coverslip. Hybridization was carried out for 16–20 hours at 37°C. Detection was done with Cy3 (FITC), BV480 Streptavidin and Anti-Digoxin antibodies. Following hybridization, the coverslips were washed thrice with 2X SSC buffer at 60°C for 5 minutes. Subsequently, the samples were incubated with a mixture of anti-fluorescein, anti-digoxigenin, and streptavidin antibodies for 20 minutes at 37°C. To remove any unbound antibodies, a final wash was performed using 2X SSC + 1% Tween-20 solution, followed by 1× PBS. During and after staining coverslips were always protected from light.

2.4.2 Data Acquisition and Analysis

The coverslips were scanned on FiberVision scanner which was a high-throughput, high-resolution, automated multi-color image-capturing system. The data was sent to FiberStudio which is a specialized software for analyzing the MC data. The pattern of different colored signals along the stretched DNA fibers resembles a

Morse code, with each "dot" or "dash" representing the D4Z4 repeat array, qA and qB haplotypes on 4q and 10q (Figure 2.1).

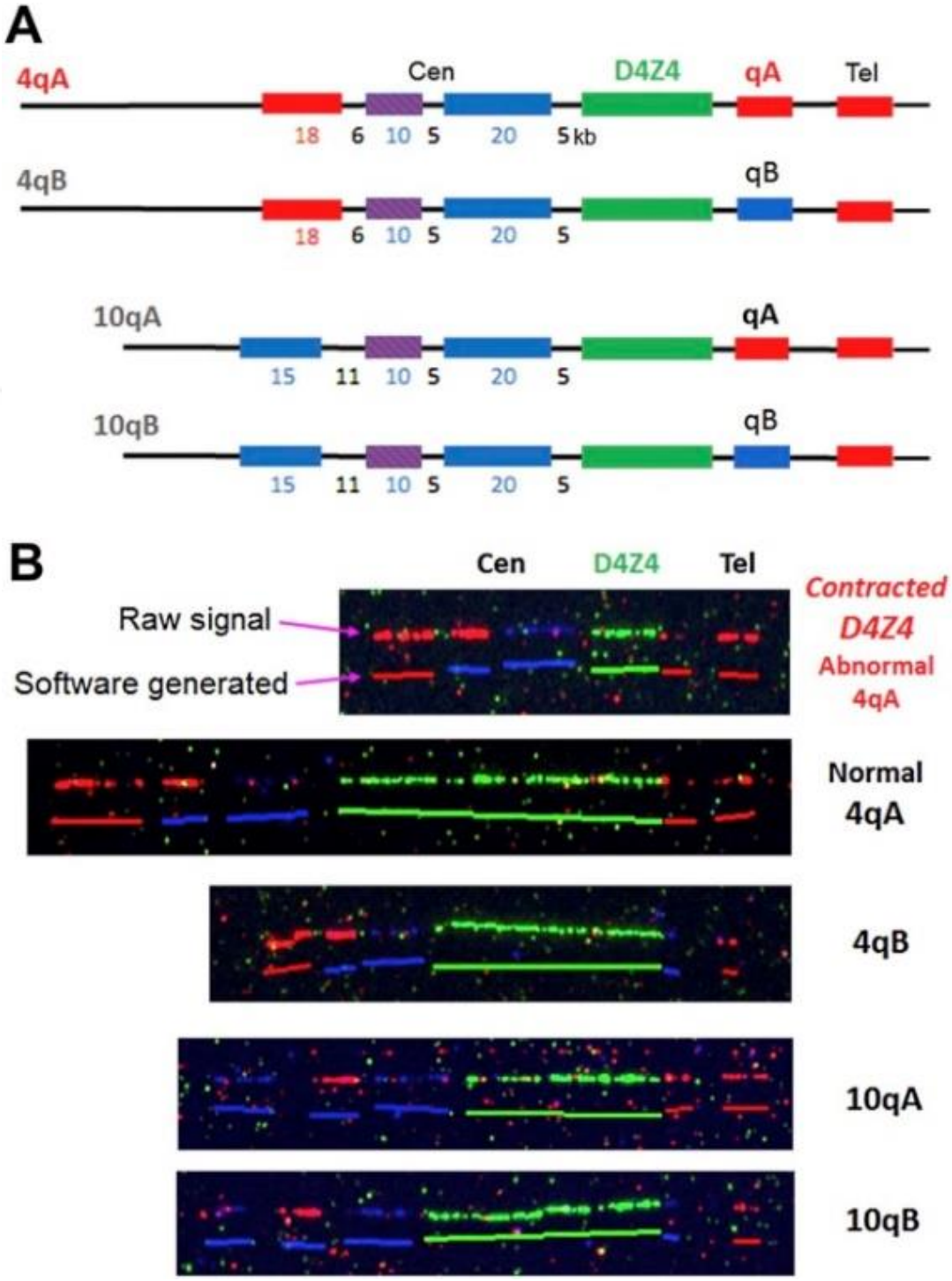

Figure 2.1: The different color patterns represent the D4Z4 repeat array, qA, and qB haplotypes on 4q and 10q.

2.5 Next Generation Sequencing

WES was performed to identify pathogenic variants in FSHD2-associated genes, including *SMCHD1* and *DNMT3B*. This approach was also used to facilitate differential diagnosis by detecting variants in genes linked to other neuromuscular disorders, including LGMDs, in patients with FSHD-like symptoms but without D4Z4 repeat contractions. An additional aim for the WES analysis was to identify potential candidate genes involved in D4Z4 epigenetic regulation, providing further insights into FSHD heterogeneity and supporting the refinement of genotype-phenotype correlations, particularly in borderline cases (8–10 RU).

WES Analysis: QIAseq Human Exome Kit (Qiagen,USA) was used, with 50 µg of genomic DNA, to target the exonic regions of the human genome, covering approximately 37 Mb. The method aimed to achieve a minimum coverage depth of 20× for ~99% of the target regions. Following the wet-laboratory process, sequencing was performed using the Illumina NovaSeq 6000 (Illumina, USA) platform.

Human refence genome alignment and variant databases and in silico prediction tools: NCBI Gene (2022-02-22), Allele Frequency Community (2019-09-25), EVS (ESP6500SI-V2), RefSeq Gene Model (2022-02-22), JASPAR (2013-11), PolyPhen-2 (v2.2.2 (HumVar)), 1000 Genome Frequency (phase3v5b), ExAC (0.3.1), iva (Dec 16 09:34), phyloP hg18 (NCBI36 (hg18) 2009-11, GRCh37 (hg19) 2014-02, GRCh38 2015-05), phyloP hg19 (NCBI36 (hg18) 2009-11, GRCh37 (hg19) 2014-02, GRCh38 2015-05), GENCODE (Release 37), CentoMD (5.3), dbVar (2021_04), OMIM (April 13, 2022), gnomAD (GRCh37 (hg19) 2.1.1, GRCh38 (hg38) 3.1.2), BSIFT (2016-02-23), TCGA (2013-09-05), ClinVar (2023-04-25), DGV (2016-05-15), COSMIC (v95), HGMD (2023.1), OncoTree (oncotree_2019_03_01), dbSNP (NCBI36 (hg18) 151, GRCh37 (hg19) 154, GRCh38 154), SIFT4G (2016-02-23).

Raw data processing: Subsequent bioinformatics analyses (including conversion from BCL to FASTQ, quality assessments, alignment to the reference genome, filtering of low-quality reads and variants, and VCF file generation) were conducted Hg19 and GATK was used for variant calling. Annovar-annotated variants were then filtered according to population frequency datasets (mean allele frequency (MAF) < 0.1%) previous reports of disease association in disease-specific databases,

pathogenicity scoring *in silico* prediction tools, Human Phenotype Ontology (HPO) ontogeny, and biological relevance.

Variant Analyses: The first step of the analysis involved examining OMIM-listed genes associated with muscular dystrophies, followed by the analysis of all remaining genes. The variant classification was performed using the American College of Medical Genetics and Genomics (ACMG) criteria, a structured approach for interpreting variants and differentiating between benign, likely benign, variants of uncertain significance (VUS), likely pathogenic, and pathogenic variants (Richards et al., 2015; Ellard et al., 2020). To support clinical variant interpretation, bioinformatics resources such as Varsome (https://varsome.com/) and Franklin (https://franklin.genoox.com/clinical-db/home) were utilized. Publicly accessible databases of genetic variation data, including ClinVar, Decipher (https://www.deciphergenomics.org/), and Leiden Open Variation Database (LOVD) served as supplementary tools. Specifically, the PP3 rule was implemented by ClinGen recommendations (Pejaver et al., 2022).

In applying the PM1 and PP2 rules, Decipher was employed to visualize gene constraint data in an accessible manner, including regional constraints for missense and loss-of-function variants (PP2), as well as to map functional domains and regulatory regions within the proteins encoded by these genes (PM1). Furthermore, a "benign cut-off frequency," obtained from Varsome, was used to compare variant frequencies across genes, facilitating the evaluation of the BS1 rule when variant frequencies surpassed gene-specific thresholds.

2.6 Segregation Analysis and Variant Confirmation

Using Sanger sequencing, we performed segregation analysis in three relatives (excluding the index patient) from one family and confirmed variants in three index patients. All amplification processes were conducted with genomic DNA isolated from the peripheral blood of patients and their family members.

Table 2.1: List of Primers used in PCR and sequencing reactions.

Gene Name and ID	Region	Sequence (5'-3')	PCR product size (bp)
CAPN3 **(NM_000070)**	Ex4-F	ATGATGCTGCTTTGGGATTC	381
CAPN3 **(NM_000070)**	Ex4-R	TTGGGTCACAGTTGGACTTG	
SMCHD1 **(NM_015295)**	Ex5-F	GGAATTTCTAGGCTAAAGGACATGA	607
SMCHD1 **(NM_015295)**	Ex5-R	AGGAAAGGGGAAAAGCTCTCAG	
SMCHD1 **(NM_015295)**	Ex15-F	TCCCTATACCATGCTTTCTTGTT	701
SMCHD1 **(NM_015295)**	Ex15-R	CCAGTCTCAGGTAGTATCTTTATAGCA	

The amplifications were performed using GoTaq® DNA Polymerase and 5X Green GoTaq® Reaction Buffer or 5X Colorless GoTaq® Reaction Buffer (Promega, USA). The reaction composition and PCR conditions are given in Table 2.2.

Table 2.2: List of components and volumes of amplification PCR mix.

Component	Final Volume	Final Concentration
5X Green or Colorless GoTaq® Reaction Buffer	10 µl	1X (1.5 mM $MgCl_2$)[2]
PCR Nucleotide Mix, 10mM each	1 µl	0.2 mM each dNTP
Upstream primer	0.5 µl	0.1–1.0 µM
Downstream primer	0.5 µl	0.1–1.0 µM
GoTaq® DNA Polymerase (5u/µl)	0.25 µl	1.25 u
Template DNA	Z µl	<0.5 µg/50µl
Nuclease-Free Water to	50 µl	

PCR reactions were performed as shown in Table 2.3.

Table 2.3: Thermal Cycling Conditions for GoTaq® DNA Polymerase-Mediated PCR Amplification.

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95°C	2 minutes	1 cycle
Denaturation	95°C	0.5–1 minute	25–35 cycles
Annealing	54–65°C*	0.5–1 minute	
Extension	72°C	1 minute/kb	
Final Extension	72°C	5 minutes	1 cycle
Soak	4°C	Indefinite	1 cycle

**Annealing temperature should be optimized for each primer set based on the primer Tm.*

PCR product efficiency and size were analyzed on 2% agarose gel prepared with ethidium bromide or SyberTMSafe Stain (1:10000, Invitrogen, catalog#S33102) in 0.5X TBE buffer. For reactions containing the 5X Green GoTaq Reaction Buffer, the amplification reaction was loaded onto the gel directly after amplification. Applied Biosystems™ ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher, USA) was used to clean up PCR products pre-Sanger sequencing. The reaction was prepared for every sample with 0.5µl ExoSAP-IT™ and 5µl PCR product. The incubation details are given in Table 2.4.

Table 2.4: Clean-up reaction for PCR products

Step	Temperature	Time	Cycle
Degradation	37°C	15 minutes	1
Inactivation	80°C	15 minutes	1

After cleaning up, BigDye reaction was performed for all clean PCR products to sequence them. Applied Biosystems™ BigDye™ Terminator v3.1 Cycle Sequencing Kit was used according to the manufacturer's protocol (Table 2.5 and

Table 2.6). After this step samples were stored on ice and protected from light.

Table 2.5: List of components for Applied Biosystems™ BigDye™ Terminator v3.1 Cycle Sequencing Kit Reaction.

Component	Quantity per reaction	Example Forward	Example Reverse
BigDye™ Terminator 3.1 Ready Reaction Mix	4 µL	4 µL	4 µL
BigDye™ Terminator v1.1 & v3.1 5X Sequencing Buffer	2 µL	2 µL	2 µL
Forward primer (3.2 µM)	3.2 pmol	2 µL	-
Reverse primer (3.2 µM)		-	2 µL
Deionized water (RNase/DNase-free)	Varies based on template and primer volume	10 µL	10 µL
Template	150-300 ng	2 µL	2 µL
Total Volume	20 µL	20 µL	20 µL

Table 2.6: Cycling conditions of sequencing.

Step	Temperature	Time	Cycle
Incubation	96°C	1 minutes	1
Denaturation	96°C	10 seconds	25
Annealing	50°C	5 seconds	
Extension	60°C	4 minutes	
Hold	4°C	Hold until ready to purify.	

Following the completion of the BigDye sequencing reaction, the reaction products were subjected to a purification step utilizing ZR DNA Sequencing Clean-up Kit™ (Zymo Research, USA) and Zymo-Spin™ IB Columns (Zymo Research, USA), as per the manufacturer's instructions. For each reaction, 240 µL of Sequencing Binding Buffer was added to the 5-20 µL BigDye reaction mixture. This mixture was then transferred to a Zymo-Spin™ IB Column assembled in a collection tube. To facilitate DNA binding to the column matrix, the assembly was centrifuged at 13,000

rpm (equivalent to 15,000 - 16,000xG) for 2 minutes. Subsequently, 300 µL of Sequencing Wash Buffer was applied to the column, and centrifugation was repeated at 13,000 rpm (15,000 - 16,000xG) for 5 minutes to effectively remove residual impurities. This washing step was repeated one more time. Finally, purified sequencing products were eluted by directly applying 6-20 µL of nuclease-free water onto the column matrix. The column was then transferred to a 1.5 mL microcentrifuge tube, and centrifugation was performed at 13,000 rpm (15,000 - 16,000xG) for 2 minutes to elute and recover the DNA. Before loading it onto the Applied Biosystems 3500 Genetic Analyzer sequencing instrument, the purified DNA was denatured at 95°C for 2 minutes in a thermal cycler and then held at 4°C for 3 minutes, ready for immediate loading.

All forward and reverse sequences of the amplification products were initially analyzed using Finch-TV software (version 1.4.0; Geospiza). Subsequently, these sequences were compared with reference sequences in the NCBI database using the BLASTn algorithm (blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence alignment was also performed using Benchling.

2.7 Multiplex Ligation-dependent Probe Amplification

Copy number variations (CNVs) in the *SMCHD1* and *CAPN3* genes were analyzed using multiplex ligation-dependent probe amplification (MLPA). The P493 *SMCHD1* probemix (MRC-Holland, Netherlands) was used to screen for deletions or duplications in the *SMCHD1* gene, while the P176 *CAPN3* probemix (MRC-Holland, Netherlands) was utilized for analysis of the *CAPN3* gene. Data from the MLPA assays were analyzed using the Coffalyser.Net software.

2.8 Methylation Analyses

Methylation analysis data was obtained from Manar Kaptan, BSc, as part of her master's dissertation titled 'DNA Methylation Profiling by Bisulfite Sequencing in Facioscapulohumeral Muscular Dystrophy (FSHD),' which is scheduled to be defended in March 2025.

2.9 Buffers and Solutions

RBC (Red Blood Cell) Lysis Solution: Prepared with 0,01 M Tris-HCl (pH:7,6) and 320 mM 109,54 g sucrose, 5 mM MgCl2, %1 Triton X-100 at pH 8,0 (autoclaved at 15 psi for 15 min).

TBE (Tris/Borate/EDTA) Buffer (for 5X Stock Solution): 54 g of Tris base, 27.5 g of boric acid and 20 ml of 0.5 M EDTA (pH 8.0) were added and pH was adjusted to 8.3 by HCl.

TE Solution, pH: 8,0: Prepared with 5 mL 1M Tris-HCl and 2 mL 0,5 M EDTA (pH: 8,0). It is completed to 1L and pH adjustment to 8, autoclaved at 15 psi for 15 min.

20X SSC (Saline-sodium citrate) Buffer (Thermo Fisher Scientific, USA)

Chapter 3:
RESULTS

3.1 Demographic Data of the Cohort

A total of 96 affected individuals were studied in this study. All patients had a clinical diagnosis of FSHD. However, 5 patients did not fully meet all phenotypic criteria and were referred for molecular analysis to confirm or exclude an FSHD diagnosis.

In 49 (52%) patients, the disease was familial and 47 (48%) were isolated. The gender distribution was 45 (47%) males and 51 (53%) females (Figure 3.1). The general age of onset in 72 (75%) patients was before the age of 20, within this group 8 patients were early onset (before the age of 10y/o) and the remaining 24 (25%) were late-onset patients (Figure 3.1)

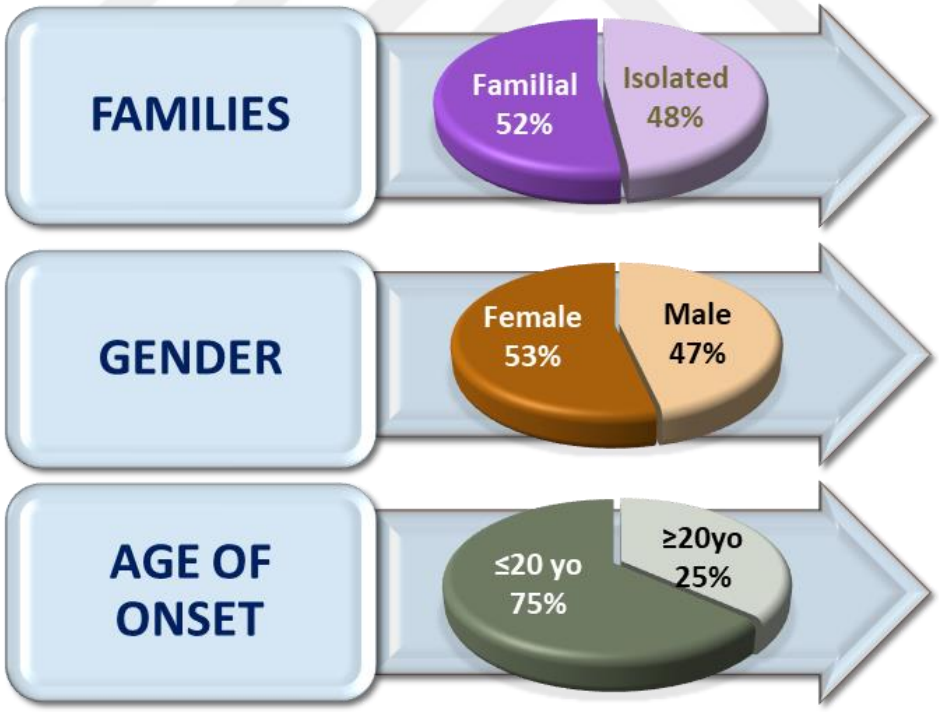

Figure 3.1: Demographic and general statistics of the patient cohort.

3.2 *D4Z4 Profiling*

MC analysis identified 82 out of 96 patients with 1–10 RUs on a 4qA allele, ascertaining a diagnosis of FSHD1. Within this group, 67 patients had 1–7 RUs, while 15 had borderline repeat sizes of 8–10 RUs. Twelve patients were found to have ≥11 RUs and the analysis also revealed two patients carrying only 4qB alleles (Figure 3.2).

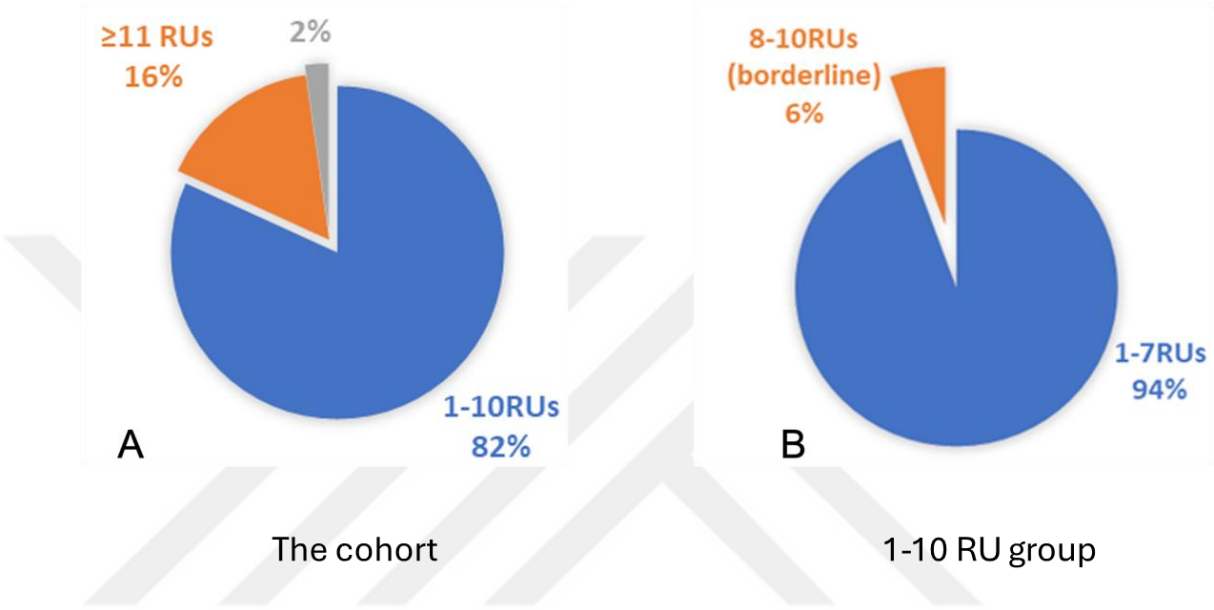

Figure 3.2: Distribution of D4Z4 RU profiles of A) the cohort and B) the 1-10 RU group.

Our initial findings confirmed the FSHD1 diagnosis of 82 patients. The D4Z4 results of affected family members were not included in this study, as we aimed to focus on selecting index patients for Whole-Exome Sequencing (WES). Only index cases were considered for diagnostic evaluation to refine WES selection criteria and optimize genetic testing. 2 affected and 1 unaffected family members were included for confirmation and segregation analyses.

3.3 *Selection of WES Patient Cohort*

Our goal was to select individuals from each group, borderline RU, non-contracted, and 4qB, who met the molecular criteria for further testing. In the borderline group, we specifically chose four patients with 9 RUs based on their age of onset and their clinical scores at their present state. All four were familial cases.

This deliberate approach allowed us to investigate whether age of onset, clinical severity, and gender influence phenotype while maintaining a consistent RU baseline for a more standardized genomic comparison.

All six patients in the non-contracted group happened to be isolated cases. Four were chosen at random and two due to their atypical MC results.

In the 4qB group, P11 was included solely for differential diagnosis, while P12 was selected due to her complex rearrangement pattern. Both patients were sporadic cases.

Table 3.1: Clinical and genetic characteristics of selected for WES. A1: Allele 1; A2: Allele 2; Mos: Mosaic; n.a: non-applicable (uncertain diagnosis); CR: complex rearrangement. Colored cells indicate patients with borderline RUs (orange), and unusual findings (green and purple). *18p deletion is a well-established CNV in FSHD cases.

Patient	Gender	Age of referral	Age of Onset	Clinical Score	Familial/ Isolated	D4Z4 size (RU)				CpG (%)	WES	Diagnosis
						4q (A1)	4q (A2)	10q (A1)	10q (A2)	Mos/ Com		
P1	F	60	51	1	+	9_qA	27_qA	20_qA	46_qA	-	43	Neg	FSHD1
P2	M	21	17	5	+	9_qA	26_qB	10_qA	38_qA	-	11	Pos	FSHD1+2
P3	M	44	42	6	+	9_qA	32_qA	9_qA	19_qB	-	20	Neg	FSHD1
P4	M	25	17	8	+	9_qA	38_qB	11_qA	19_qB	-	16	Pos	FSHD1+2
P5	M	20	16	1	-	15_qA	24_qA	15_qA	26_qA	-	40	Neg	n.a
P6	M	77	70	2	-	23_qA	61_qA	6_qA	6_qA	1_4qA (%9)	42	Pos?	FSHD1_mos EDMD-like
P7	F	17	14	2	-	48_qA	48_qA	10_qA	32_qA	-	48	Pos	LGMD2
P8	F	64	40	4	-	13_qA	15_qB	7_qA	30_qA	-	23	Neg	FSHD1_4q;10q
P9	F	65	42	8	-	43_qA	18_qB	31_qA	44_qA	-	-	Neg	n.a
P10	M	30	10	11	-	21_qA	21_qB	10_qA	10_qB	-	16	Neg	n.a
P11	M	46	xx	1	-	14_qB	37_qB	13_qA	33_qA	-	50	Pos	LGMD2
P12	F	22	xx	4	-	20_qB	27_qB	12_qA	22_qA	4qA_CR	24	Pos	FSHD1_CR

3.4 SMCHD1 Variants

In P2, P4, and P12 *SMCHD1* (NM_015295) variants were identified P2 and P4 had already been confirmed as a FSHD1 case with a borderline 9RU_4qA haplotype. P12 had a complex D4Z4 genotype (Table 3.1).

Patient 2 (P2): This is a 21-year-old male patient with symptom onset around the age of 17. He presented with bilateral mild weakness of the orbicularis oris and orbicularis oculi muscles. He had proximal upper limb involvement characterized by limited shoulder abduction (90°), and forward-sloping shoulders with bilateral scapular winging. Muscle strength in all lower limb muscle groups was 5/5. Beevor's sign was mildly positive. He was unable to bring himself to a sitting position from the supine position without getting support from the bed. His clinical severity score was 5/15. The patient's pedigree and family history indicate that his paternal aunt, father, and three siblings were similarly affected.

Although WES analysis did not identify any SNVs in this patient, CNV analysis detected a 421 kb heterozygous deletion encompassing *LOC727896, EMILIN2, LPIN2,* and *SMCHD1* genes. This finding was further validated by P493 *SMCHD1* MLPA probemix (MRC-Holland, Netherlands) analysis (Figure 3.3). Molecular analysis of the father revealed the same deletion. This deletion has been previously reported in FSHD2 cases. His father has also been subjected to MC and MLPA analysis. He also has D4Z4 9RU_4qA and whole gene deletion. The CpG level of this patient was found to be 11%.

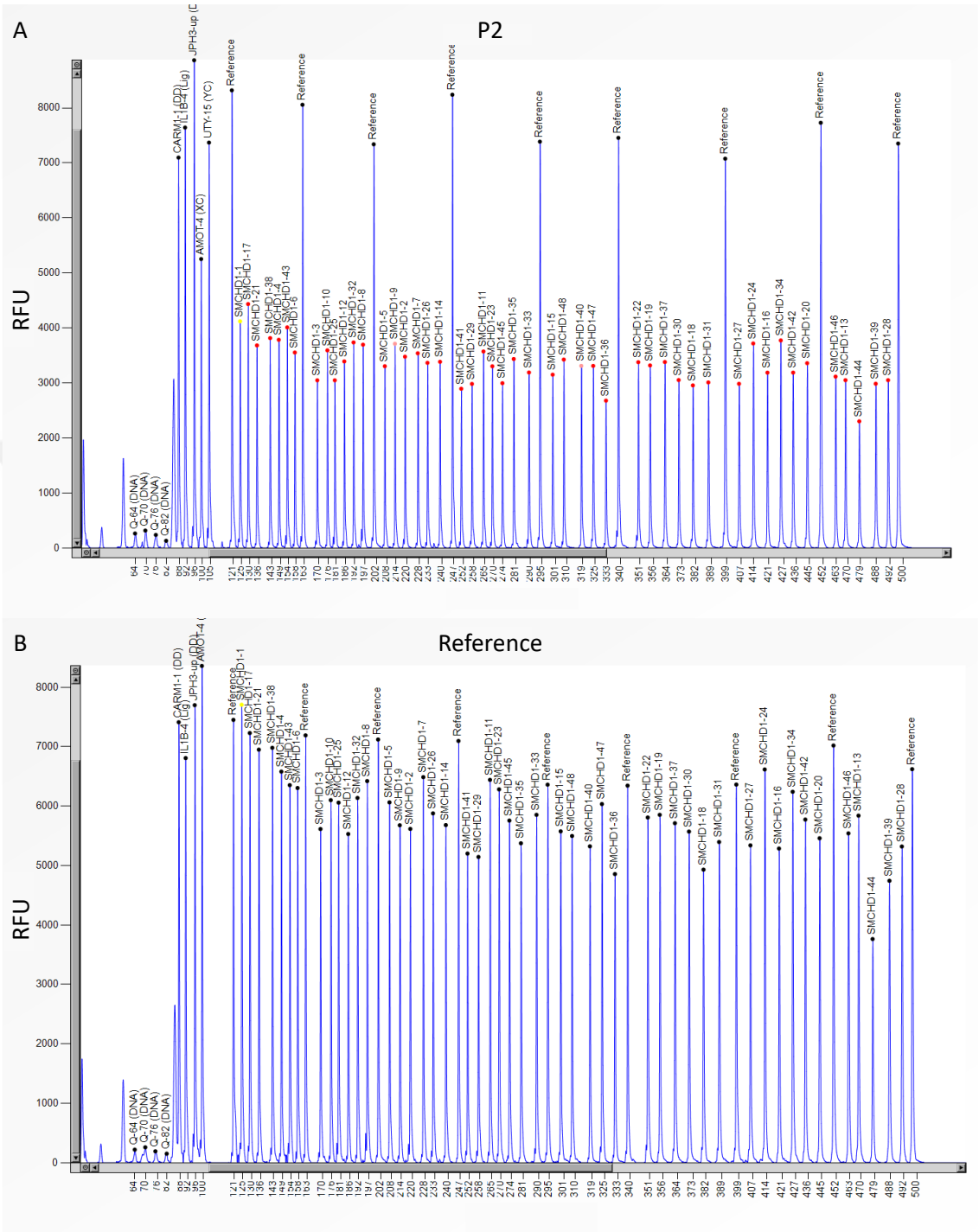

Figure 3.3: MLPA analysis of P2 confirming the whole SMCHD1 deletion.

Patient 4 (P4): This 25-year-old male patient had symptom onset around the age of 17 and displayed rapid progression of disease. He had partial weaknesses of orbicularis oculi and orbicularis oris muscles, and weakness of the right upper and lower limbs. He had limited shoulder abduction (45-90°) with bilateral winged scapula, forward sloping of shoulders, and weakness of biceps and triceps muscles. Beevor's sign was positive. He had asymmetric tibialis anterior weakness with right foot drop, and a positive Gower's sign. The clinical severity score was 8/15. Family history revealed that his brother was also affected.

Novel, heterozygous, likely pathogenic c.2018delT was identified in *SMCHD1* coding exon 15 (of 48). The confirmation and segregation analysis were performed by Sanger sequencing. This frameshift variant is predicted to result in a truncated or nonfunctional SMCHD1 protein due to premature termination, with the expected product being p.(Ala674Leufs*9). His CpG level was 16%.

The patient had a similarly affected brother (II-3), and the parents were not reported to have any clinical findings (Figure 3.4). Segregation analyses of the parents and the affected sibling revealed that the affected brother (II-3) and the mother (I-1) carried the same variant. The father was found to be negative for this variant (Figure 3.5).

SMCHD1 c.2018delT represents a novel variant since there are no reports or publications on it to the best of our knowledge. According to ACMG guidelines, this variant is classified as "likely pathogenic", supported by one very strong piece of pathogenic evidence (PVS1, null variant in a gene where loss of function is the disease mechanism) and one moderate pathogenic evidence (PM2, variant not reported in gnomAD population databases). It was also absent from TOPMed or other public variant archives including dbSNP and ClinVar.

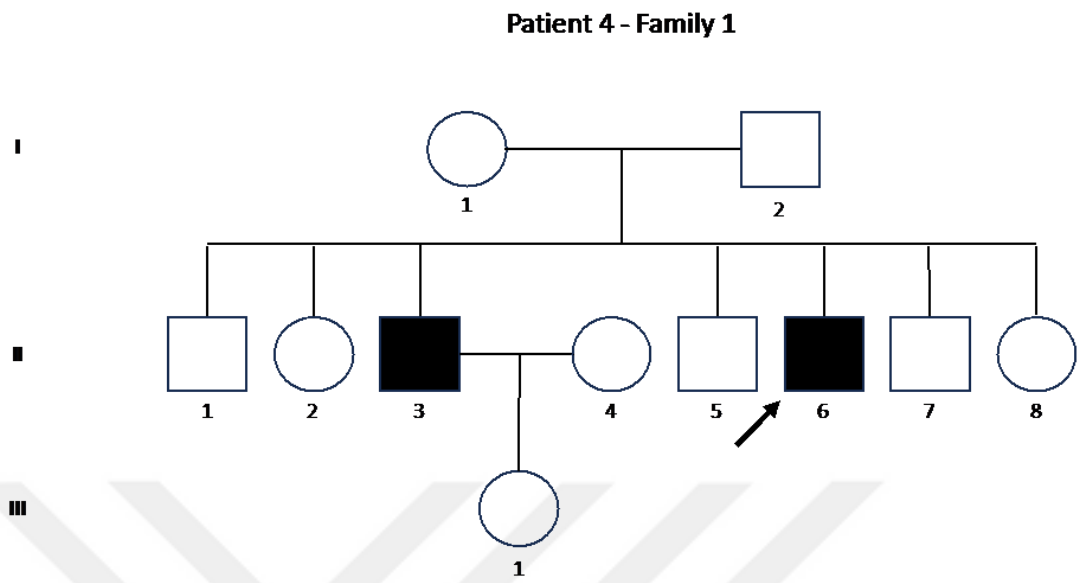

Figure 3.4: Pedigree of P4's family.

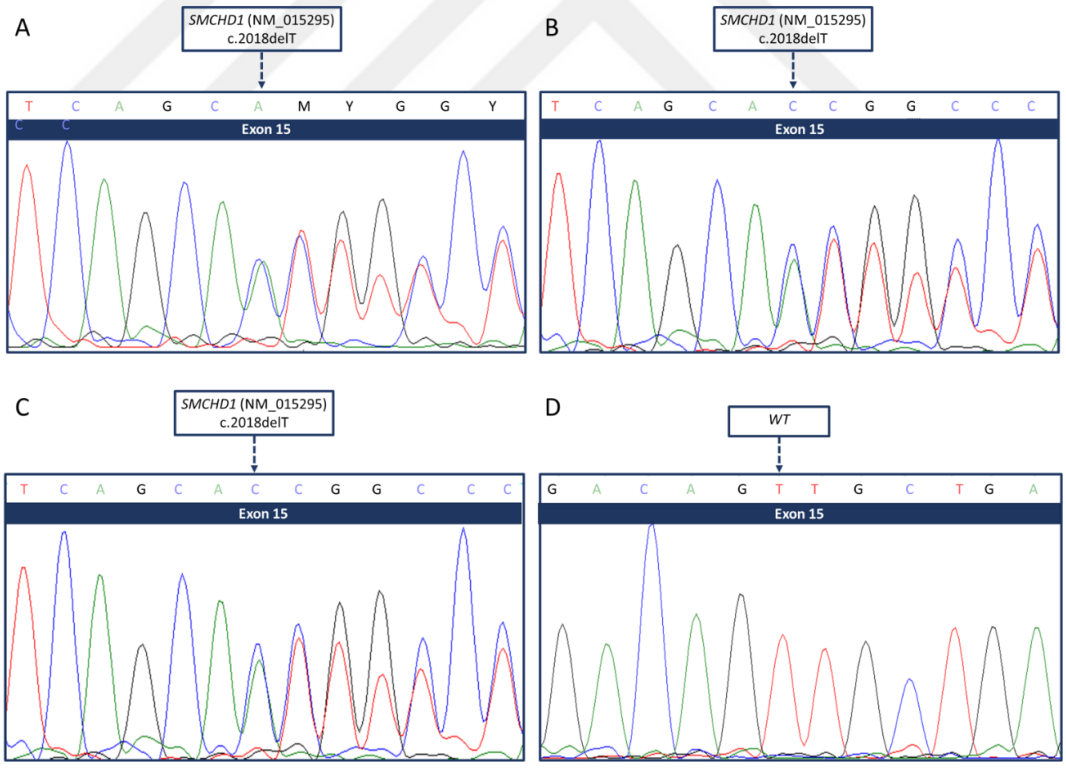

Figure 3.5: The confirmation and segregation analysis of the variant. A: Patient's variant confirmation; B: Male sibling's (II-3) segregation analysis; C: Mother's (I-1) segregation analysis; D: Father's (I-2) segregation analysis.

***Patient 12 (P12)*:** This is a 22-year-old female patient with symptom onset around the age of 16. She had weaknesses of orbicularis oris muscles more prominent on the left side, and limited shoulder abduction (90°) with bilateral winged scapula, mild weakness of bilateral triceps and pectoral muscles. She had mild, asymmetric tibialis anterior weakness with subtle left foot drop. Electromyography (EMG) findings indicated mild, generalized myopathy with chronic features. Her clinical severity score was 4/15. Pedigree analysis showed that she was an isolated case.

MC analysis revealed that this patient had a 4qB haplotype (Table 3.1) However, she also carried 4qA allele rearrangements with 6–8 RUs, followed by a 40 RU array in succession (Figure 3.6). Her CpG level was 24%.

The *SMCHD1* c.610A>G variant was detected in a heterozygous state in coding exon 5 (of 48). This is a missense variant, with the expected protein product being p.(Lys204Glu). According to ACMG guidelines, this variant is classified as "VUS", supported by one moderate (PM2), and two supporting pieces of pathogenic evidence (PP2, missense variant in a gene with low rate of benign missense mutations; PP3, computational prediction tools unanimously support a deleterious effect). It has an extremely low allele frequency of 0.000004 on TOPMed and has been reported in dbSNP (rs1184311800). Its clinical significance has not been annotated in public databases. This variant has been previously reported as disease-causing in scientific literature (Mul et al., 2018 and Lemmers et al., 2019).

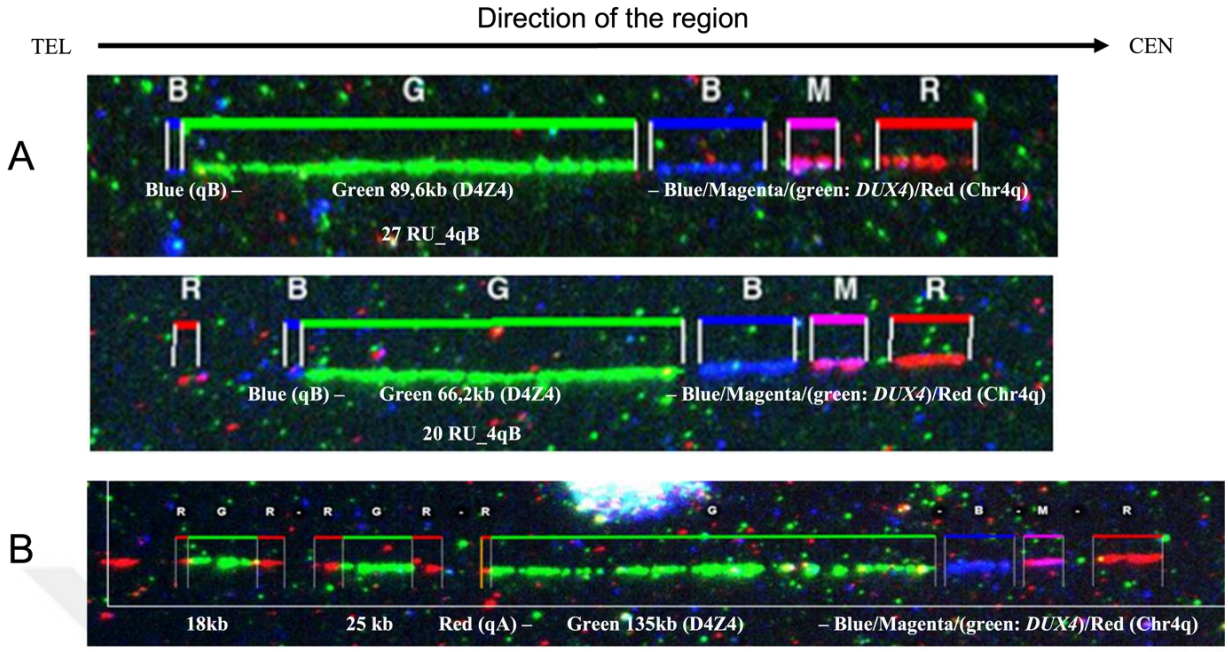

Figure 3.6: MC images of P12. A) 20 RU and 27 RU 4qB alleles. The color sequence Blue/Magenta/(green: *DUX4*)/Red indicates that the region is on chr4, blue on the telomeric region indicates the haplotype is qB, and the D4Z4 RU is calculated by the length (kb) of the green line divided by 3,3. B) At the centromeric end the color sequence indicates that this allele lies in 4q and red at the telomeric ends of each line indicates qA allele. This image shows there are consecutive lines of D4Z4 varying in length, and calculated RUs differed between 6RUs to 8RUs in individual images.

3.5 CAPN3 Variants

In P7 with non-contracted 4qA and P11 with only 4qB haplotype, *CAPN3* (NM_000070) variants were identified (Table 3.2).

Patient 7 (P7): This is a 17-year-old female patient with symptom onset around the age of 15. She exhibited mild facial involvement, characterized by asymmetric, subtle weakness of the orbicularis oris. She had bilateral winged scapula, and muscle power in upper limb muscles were normal. Lower limb findings included mild gastrocnemius weakness and atrophy of the right calf muscles. She reported difficulty and pain while climbing stairs. CK levels were persistently elevated, and EMG findings were normal. Her clinical severity score was 2/15. P7 represents an isolated case.

P7 did not have contracted 4qA. Her CpG level was 48%. WES analysis revealed the presence of homozygous *CAPN3* c.146G>A in coding exon 1 (of 24). This is a missense variant and the expected protein product is p.Arg49His. It has been previously reported as "pathogenic" both in variant databases (rs863224958) and in scientific literature. Her family members were not available for further molecular testing. We ruled out a possible trans-allelic deletion in the index by P176 *CAPN3* MLPA probemix (MRC-Holland, Netherlands), confirming homozygosity for *CAPN3* c.146G>A.

Patient 11 (P11): This is a 46-year-old male patient with symptom onset around the age of 45. He presented with an atrophy of the neck muscles. There was no facial involvement except for slightly reduced movement of the left upper lip. Bilateral trapezius and pectoral muscles appeared atrophic, and he had mild scapular winging. Muscle power examination showed normal results in all muscle groups. EMG revealed very low amplitude polyphasic potentials in the trapezius muscles. P11 was an isolated case with a clinical severity score of 1/15. FSHD molecular testing was planned to make an exclusion diagnosis.

P11 did not have contracted 4qA, the allele compositions were 14 RU_qB and 37 RU_qB. CpG level was 48%. Pathogenic *CAPN3* c.549delA variant was identified in a heterozygous state in coding exon 4 (of 24). This is a frameshift variant resulting in a truncated nonfunctional protein due to premature termination (p.Thr184Argfs*36). It was previously reported as "pathogenic" both in variant databases (rs80338800) and in the scientific literature. WES analysis did not show any other candidate SNVs or CNVs in *CAPN3*. His family members were not available for further molecular studies. We ruled out a possible trans-allelic deletion in the index by P176 *CAPN3* MLPA probemix (MRC-Holland, Netherlands), confirming homozygosity for *CAPN3* c.549delA.

3.6 Candidate Variant and re-evaluation of MC results in two patients

Two patients, P6 and P8, had uncharacteristic MC results, which were compatible with somatic mosaicism and complex rearrangement, respectively.

Patient 6 (P6): This was a 77-year-old male patient with symptom onset around the age of 55, and a history of coronary artery disease. The first clinical finding at the age of onset was asymmetric shoulder weakness, which progressed to inability to raise both arms to head level. He describes intolerance to moderate-intensity exercise (i.e., walking) in the form of fatigue and shortness of breath, which developed and slowly progressed over the past two years. He had two episodes of falling during this period. Extensive cardiological work-up showed normal results. He displayed slightly pronounced pouting of the lips, but motor functions of the orbicularis oculi and orbicularis oris muscles were normal except a slight, left-sided reduction in upper lip movement while whistling. He had scapular winging on the right with humping. Muscle power was normal in all upper limb muscle groups, including the deltoids. On lower limb muscle power examination, he showed moderate weakness of the hip abductors and adductors, mild weakness of the tibialis anterior and extensor hallucis longus muscles, and normal muscle power in the remaining groups. He exhibited a waddling, steppage gait. His clinical severity score was 2/15.

Given his mild and atypical phenotype, both FSHD and other myopathies were considered as differential diagnoses in P6. MC analysis revealed he had two non-contracted D4Z4 4qA alleles (23/61), and of all 4qA alleles counted during the study, there was also a third 4qA alleles with 1RU (Figure 3.7). The allele proportion was calculated to be 9%. CpG level was 42%. These results indicated possible somatic mosaicism in P6, which could have contributed to his late onset and atypical clinical condition. He was then included in the present study due to his unique D4Z4 composition and atypical clinical presentation. WES analysis revealed a hemizygous, missense c.404G>T, p.(Gly135Val) in *FHL1* (NM_001159699.2), which was classified as a VUS. This variant has been previously reported in ClinVar and dbSNP (rs886042453) as VUS for X-linked myopathy with postural muscle atrophy. *FHL1* is also associated with Emery-Dreifuss Muscular Dystrophy 6, X-linked (#300696, XLR), and other X-linked myopathies. The patient did not volunteer for a muscle biopsy, therefore the D4Z4 mosaicism could not be evaluated in the muscle tissue. He is an isolated case, and his family members were not available for further analysis.

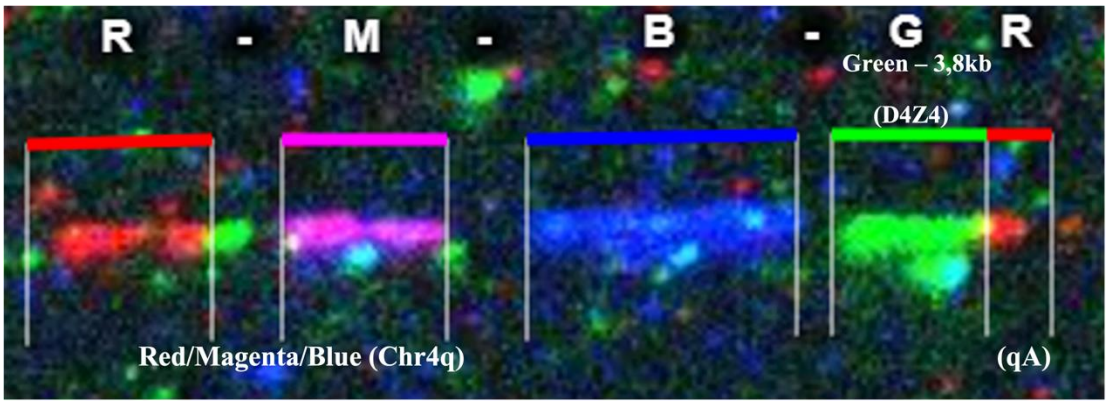

Figure 3.7: MC image of patient P6 showing 1 RU_4qA (3,8kb D4Z4).

Patient 8 (P8): This is a 64-year-old female patient with symptom onset around the age of 40. Her main complaint was the inability to raise her arms above head level, which showed very slow progression over the years. She was also diagnosed with fibromyalgia. Her EMG findings were compatible with bilateral carpal tunnel syndrome and myogenic involvement of the trapezii. CK levels were mildly elevated. She had mild ptosis on the left, subtle weakness of the left orbicularis oculi muscle, and normal orbicularis oculi function. She had bilateral, severe scapular winging most prominent on the right, and humping more prominent on the left. Muscle power examination showed mild weakness of the infraspinatus and serratus anterior muscles, the clavicular head of the pectoral muscles, and the iliopsoas muscles. Neurological examination was otherwise normal. She represents an isolated case, with a clinical severity score of 2/15.

She had one 4qA allele with 13 RUs and one 4qB allele with 15 RUs. However, her 10q results were thought to have significance with a contracted 10qA allele with 7RUs. CpG level was 23%. She was recruited for this study due to her interesting MC results. We did not find a clinically significant variant in her WES analysis.

3.7 Results Summary

We identified three *SMCHD1* variants, two *CAPN3* variants, and one potential candidate variant in the *FHL1* gene. In our 12-patient cohort, six patients received a definitive diagnosis, while two cases could be explained solely based on their MC results (Table 3.2 and

Table *3.1*). However, the underlying genetic cause remains completely unresolved for five patients with the available data (Figure 3.8).

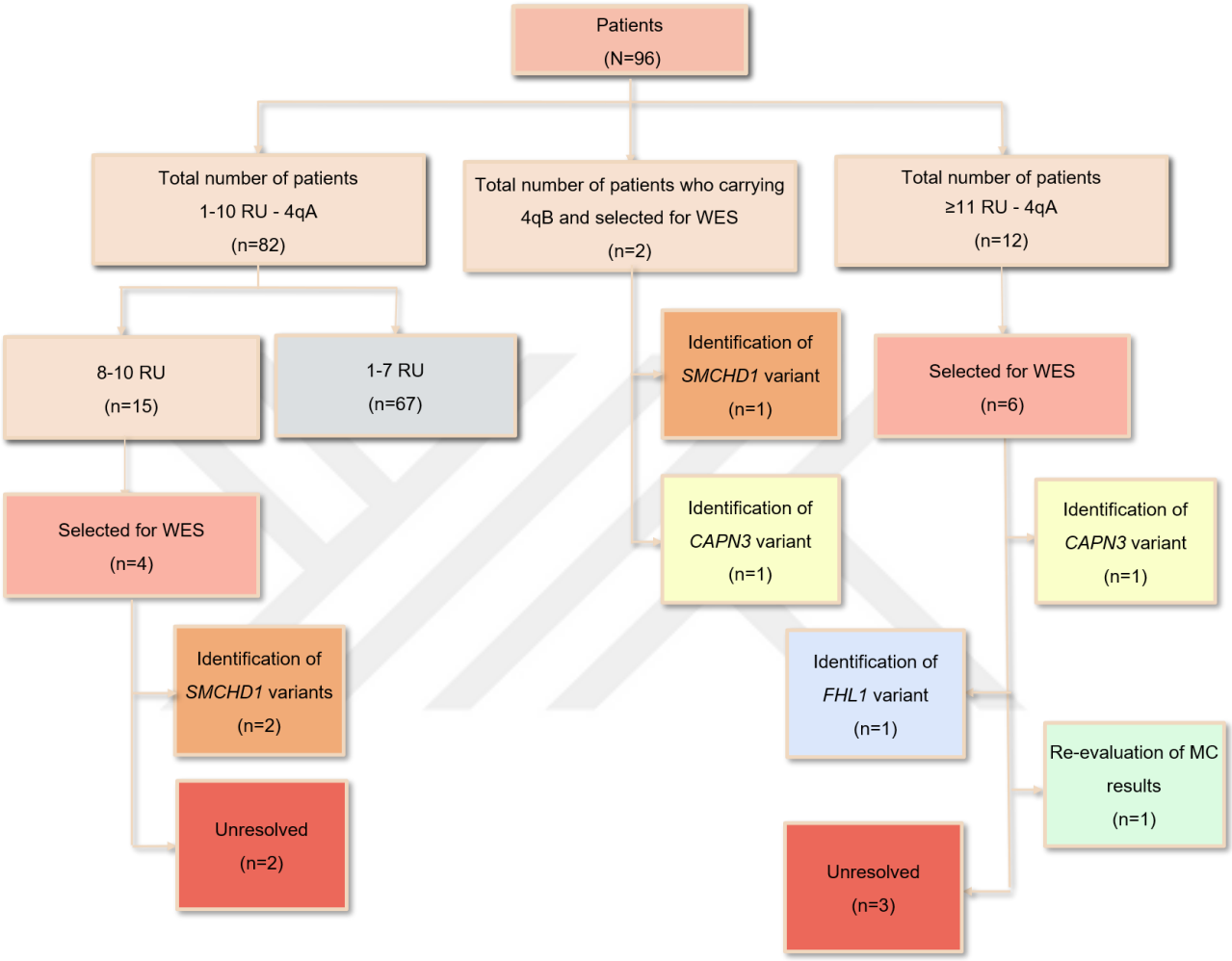

Figure 3.8: Diagram of the patient cohort.

Table 3.2: Genetic Variants Identified in 6 Patients.

Individual	Gene	Variant position	Variant Description	Effect	Population frequency	ACMG Classification	Inheritance	Variant ID
P2	*SMCHD1*	18p11.32p11.31	seq[GRCh37] 18p11.32p11.31(2538999-2960903)x1	421kb del (*LOC727896,* ***SMCHD1,*** *EMILIN2, LPIN2)*	-	Pathogenic	Heterozygous	*
P4	*SMCHD1*	18:2674116	NM_015295: c.2018delT	p.Ala674Leufs*9	-	Likely Pathogenic	Heterozygous	This study
P12	*SMCHD1*	18:2706423	NM_015295: c.610A>G	p.Lys204Glu	0.000004	VUS	Heterozygous	rs1184311800
P8	*CAPN3*	15:42652149	NM_000070: c.G146A	p.Arg49His	-	Pathogenic	Homozygous	rs863224958
P11	*CAPN3*	15:42680001	NM_000070: c.549delA	p.Thr184Argfs*36	0.000280	Pathogenic	Heterozygous	rs80338800
P6	*FHL1*	X:135289975	NN_001159699.2: c.404G>T	p.Gly135Val	-	VUS	Heterozygous	rs886042453

Chapter 4:
DISCUSSION

FSHD is known for its variable clinical expressivity, influenced by genetic, epigenetic, and potentially environmental factors. D4Z4 repeat contraction on 4qA is the primary pathogenic mechanism in FSHD1. However, disease severity and age of onset vary widely, even among individuals with the same repeat length. In addition, FSHD2, which arises from variants in chromatin-modifying genes such as *SMCHD1* and *DNMT3B*, further underscores the genetic complexity of the disease.

In this study, we analyzed a cohort of 96 unrelated patients with a preliminary or differential diagnosis of FSHD, revealing key demographic and clinical distributions (Figure 3.1). The proportion of familial cases (52%) was lower than the universally accepted range of 70%, as reported in published studies (Lemmers et al., 2012; Tawil et al., 2014). This discrepancy may be attributed to the smaller size of our cohort since published data is mostly from large cohort groups (>500). Additionally, we observed a slightly higher prevalence of female patients (53%) compared to males (47%). According to previous reports, there is a greater proportion of asymptomatic or mildly affected female carriers (Tawil et al., 2014) when compared to males which has been suggested to be due to potential hormonal differences influencing disease expressivity (Tawil et al., 2018; Mul et al., 2019). However, in general, no significant gender difference is observed between male and female patients. Early-onset cases (<10 years old) accounted for 8.3% of the cohort, while the ≤20-year-old group represented 66.7%, collectively comprising 75% of all cases. FSHD can manifest at various ages, but the typical onset occurs in the second decade of life. This finding aligns with previous reports indicating that more than 50% of individuals with FSHD exhibit symptoms by the age of 20, with nearly 10% classified as early-onset cases.

We confirmed the diagnosis of FSHD1 in 82 out of 96 patients, representing 85.4% of the cohort (Figure 3.2). Two patients in this group (P2, P4) were diagnosed with FSHD1+2 due to pathogenic *SMCHD1* variants. In the remaining 14 patients, one was diagnosed with FSHD2 (P12), two patients (P7, P11) were differentially

diagnosed as LGMD2 and one patient (P6) had an ambiguous diagnosis of FSHD1 and/or Emery-Dreifuss Muscular Dystrophy (EDMD). Our findings suggest that while FSHD1 remains the predominant form, it corresponds to a slightly lower ratio among all FSHD patients compared to the commonly reported 95% in the literature (Lemmers et al., 2012; Sacconi et al., 2019). The presence of FSHD1+2 cases reinforces the notion that *SMCHD1* variants can act as disease modifiers, exacerbating disease severity in FSHD1 individuals with borderline repeat contractions (Larsen et al., 2020). The identification of only one FSHD2 case in our study suggests that its prevalence may be lower than traditionally estimated. This aligns with recent reports indicating that FSHD2 comprises fewer than 5% of all cases when comprehensive genetic and epigenetic testing is performed (van den Boogaard et al., 2016).

The age of onset and disease severity in our patients aligns with the well-established inverse correlation between FSHD severity and residual D4Z4 repeat size on chromosome 4q35, where shorter repeat arrays (1–3 RUs) are associated with earlier onset, faster progression, and more severe muscle involvement, while borderline pathogenic repeat sizes (8–10 RUs) demonstrate variable penetrance and a broader clinical spectrum. Moreover, borderline RUs are frequently influenced by additional genetic and epigenetic modifiers, including *SMCHD1* variants, methylation status, and haplotype configuration, all of which contribute to variability in disease expressivity and progression (Lemmers et al., 2012; van den Boogaard et al., 2016).

This study aimed to identify pathogenic variants in *SMCHD1*, *DNMT3B*, and other D4Z4 methylation regulators, assess the integration of NGS with conventional FSHD1 diagnostics, evaluate alternative diagnoses in FSHD-like cases, and investigate borderline D4Z4 contractions (8–10 RU), particularly the impact of SMCHD1 variants on disease severity. Therefore, we analyzed 12 FSHD patients in total chosen from main groups of borderline contracted, non-contracted and 4qB haplotypes.

Borderline Patients

Four FSHD patients, P1, P2, P3, and P4, all carrying 9 D4Z4 repeat units (RUs) on a 4qA haplotype, a borderline pathogenic configuration, confirming FSHD1 (Table 3.1). The phenotypic variability within this group suggests the influence of additional genetic or epigenetic modifying factors.

P1 was the only female patient in this cohort. She had a late onset at the age of 60, and displayed an almost asymptomatic phenotype, with an FSHD clinical severity score of 1. Despite carrying a 9RU allele, she lacked *SMCHD1* variants or other FSHD-related genetic alterations. This was a familial case and her daughter, with mild clinical features, also had 9RUs. This is also by previous studies, which report that female FSHD patients often present with milder phenotypes, later onset, and lower penetrance compared to males (Tawil et al., 2014; Scionti et al., 2012). This difference may be partially attributed to estrogen-related epigenetic regulation or the protective effects of a second X chromosome, as proposed in other neuromuscular disorders (Mul et al., 2018). Both the RU number and gender of these individuals correlate with a late-onset, mild phenotype.

The male patients in this study exhibited a wide range of disease severities, despite sharing the same 9RU contraction. P3, a male patient (onset age 43, clinical score 6) did not have an identifiable *SMCHD1* variant or CNV, indicating that D4Z4 contraction alone may be sufficient for disease manifestation in some cases. However, his later onset age and moderate severity contrast with the more severe presentations seen in the two other male patients, both of whom carried pathogenic *SMCHD1* variants. This pattern suggests that male patients may be more susceptible to disease modifiers, contributing to more severe and earlier-onset phenotypes (van den Boogaard et al., 2016).

Studies examining FSHD1 cohorts with 9RUs have demonstrated similar phenotypic variability. In a large-scale study, Scionti et al. (2012) reported that patients with 8–10 RUs often have incomplete penetrance, with a subset of individuals remaining asymptomatic, like our female patient. However, those with additional epigenetic modifications or genetic variants, particularly in *SMCHD1*, tended to develop a more severe phenotype at an earlier age.

This is exemplified by the male patient, P2, in our cohort with a 427 kb deletion on chromosome 18p, which encompassed *SMCHD1* along with other genes (Figure 3.3). His onset age (17 years) and clinical severity (score 5) suggest that haploinsufficiency of SMCHD1 likely played a role in disease acceleration. Lemmers et al. (2015) similarly described patients with FSHD1 and large deletions affecting *SMCHD1*, noting that these cases had greater D4Z4 hypomethylation, reinforcing the role of *SMCHD1* in epigenetic suppression of *DUX4*. In agreement with these, P2 also presented with marked D4Z4 hypomethylation, exhibiting a methylation level of 11%. The father of the patient also had similar clinical findings and was also found to have the same genetic composition in D4Z4 and *SMCHD1* gene.

A more severe phenotype was observed in P4 (onset at 16–17 years, clinical score 8), who carried a novel likely pathogenic base deletion variant, c.2018delT (p.Ala674Leufs*9), in *SMCHD1*. This variant is located within the coiled-coil region of the protein, which functions as an intervening domain between the N-terminal GHKL-ATPase domain and the C-terminal SMC hinge domain (Figure 4.1). The coiled-coil regions facilitate protein-protein interactions and are essential for SMCHD1 homodimerization. Truncating variants within this region can disrupt homodimerization, leading to loss of function, a mechanism that has been associated with FSHD2 (Strafella et al., 2019). The identified variant is predicted to cause a frameshift and premature protein termination, resulting in a complete loss-of-function effect. This loss-of-function variant is highly likely to exacerbate disease severity by further destabilizing repression at the 9RU D4Z4 locus, ultimately contributing to an earlier onset and more severe clinical presentation. Similar cases have been reported in FSHD2 cohorts, where frameshift or nonsense variants in *SMCHD1* lead to more aggressive disease progression, even in patients with a relatively high number of D4Z4 repeats (Zhang et al., 2022). The 16% hypomethylation level observed at the D4Z4 region in patient P4 further corroborates this finding. The severity observed in our patients is consistent with prior studies demonstrating that FSHD1 patients carrying co-occurring *SMCHD1* variants exhibit more extensive muscle involvement and earlier onset (Mul et al., 2018). To the best of our knowledge, this variant is being reported for the first time in this study.

The methylation profiles of the remaining two patients with only borderline D4Z4. P1 = 40% and P3 = 20%, align with established findings that FSHD2 patients exhibit greater hypomethylation than FSHD1 cases (Lemmers et al., 2012; Sacconi et al., 2019). The observed methylation levels also correlated with disease severity and patient clinical scores, further supporting the role of epigenetic deregulation in modifying FSHD phenotype (Gaillard et al., 2015; Hiramuki et al., 2022). Our findings emphasize the complex interplay between D4Z4 repeat contraction, *SMCHD1* function, and additional genetic modifiers in FSHD onset and progression. A borderline allele alone does not fully predict clinical severity, particularly in borderline cases. Instead, epigenetic dysregulation—often mediated by *SMCHD1* dysfunction—appears to be a key determinant of disease expressivity.

Non-contracted Patients

In the six patients from the non-contracted group (P5–P10), a pathogenic *CAPN3* variant was identified in P7. In P6, a variant of uncertain significance (VUS) in *FHL1* was detected. No significant variants were found in P5, P8, P9, and P10. However, in P8, the MC results were further re-evaluated for a potential impact on the phenotype.

P7, the 17-year-old female patient, presented with symptom onset at age 15, exhibiting mild facial weakness, isolated lower limb involvement with right calf muscle atrophy, bilateral scapular winging, and difficulty climbing stairs due to muscle weakness and pain. Her clinical severity score of 2/15 and persistently elevated CK levels raised suspicion for Facioscapulohumeral Muscular Dystrophy (FSHD1/2), though the absence of a D4Z4 contraction on 4qA led to further investigation. WES ultimately identified a homozygous pathogenic *CAPN3* c.146G>A (p.Arg49His) variant, confirming autosomal recessive Limb-Girdle Muscular Dystrophy Type 2A (LGMD2A or calpainopathy) instead.

The *CAPN3* c.146G>A (p.Arg49His) variant identified in this patient has been previously reported as pathogenic, with documented cases demonstrating early-onset, progressive lower limb weakness, calf atrophy, and elevated CK levels, aligning with her presentation (Richard et al., 1999; Fardeau et al., 2001). Unlike FSHD, where

disease progression is often slow and asymmetric, LGMD2A follows a more predictable proximal-to-distal pattern of muscle involvement (Tasca et al., 2016).

The clinical and genetic findings in P6 present a complex case with features suggestive of FSHD, yet with atypical characteristics. The patient is a 77-year-old male (onset at 55 years, clinical score 2) presented with asymmetric shoulder weakness, scapular winging, and exercise intolerance. Facial involvement was minimal, and he displayed moderate weakness of the hip adductors, abductors, and anterior distal leg muscles. This presentation is atypical for FSHD, which often manifests with more pronounced facial and scapular muscle involvement at an earlier age. MC analysis identified two non-contracted 4qA alleles and a mosaic 1RU 4qA allele, constituting approximately 9% of the alleles, and a CpG level of 40 was detected. WES revealed a hemizygous missense variant in *FHL1*, c.404G>T (p.Gly135Val).

The presence of somatic mosaicism is a phenomenon that has been well-documented in FSHD. It occurs when a post-zygotic variant results in two or more genetically distinct cell populations within an individual (van Overveld et al., 2000). In FSHD, individuals with mosaic D4Z4 contractions often exhibit reduced disease severity and later onset, likely due to a lower proportion of affected cells expressing pathogenic DUX4 (Lemmers et al., 2007). Several studies have reported that patients with somatic mosaicism for a pathogenic D4Z4 contraction frequently present with attenuated symptoms or incomplete penetrance, particularly when the proportion of contracted alleles is low (Tawil et al., 2013). P6 had a borderline CpG of 40%. Given that in healthy individuals, methylation levels could approximately be 20% lower than normal levels (30% to 65%), FSHD1 cannot definitively be ruled out in P6. Additionally, studies utilizing methylation profiling and chromatin accessibility assays suggest that mosaicism may also affect epigenetic regulation, potentially further influencing DUX4 expression levels and modifying the phenotype (van den Boogaard et al., 2016). Given these findings, the presence of a low-frequency mosaic allele in our patient (9%) may explain the atypical late-onset and milder disease course.

Variants in the *FHL1* gene have been implicated in various X-linked myopathies, notably EDMD and reducing body myopathy (Quinzii et al., 2008; Schessl et al.,

2008). The c.404G>T variant results in a p.Gly135Val substitution, affecting a highly conserved glycine residue within the third LIM domain of the FHL1 protein. LIM domains are crucial for protein-protein interactions, playing a fundamental role in the structural integrity and function of skeletal and cardiac muscle (Shalaby et al., 2009). Disruptions in these domains have been associated with abnormal sarcomere assembly and myofibrillar disorganization, leading to muscle weakness and contractures, as seen in *FHL1*-related myopathies (Schessl et al., 2010).

The p.Gly135Val variant is currently classified as a Variant of Uncertain Significance (VUS), but its location within a functionally essential domain suggests a possible pathogenic effect. Notably, other missense variants in the third LIM domain of *FHL1* have been linked to EDMD and related myopathies (Selcen et al., 2011). For instance, a study identified novel *FHL1* variants associated with EDMD6, with affected individuals presenting muscle weakness, contractures, and cardiomyopathy in late adulthood, often with a milder phenotype compared to early-onset cases (Gueneau et al., 2009). These findings support the hypothesis that even subtle alterations within the LIM domains can have clinically significant consequences, particularly in late-onset and mild phenotypes.

Given the critical role of the third LIM domain in the *FHL1* function and its established link to X-linked myopathies, the p.Gly135Val variant warrants further investigation. Functional studies are needed to determine its impact on protein stability and sarcomere assembly, and detailed clinical evaluations of carriers could clarify its contribution to disease onset and progression (Cowling et al., 2014).

The combination of somatic mosaicism for a contracted D4Z4 allele and a potentially pathogenic *FHL1* variant complicates the diagnostic picture. Given the atypical clinical features and the genetic findings, it is plausible that P6's phenotype results from a synergistic effect of both the D4Z4 contraction mosaicism and the *FHL1* variant.

P8 was a 64-year-old female with a late-onset disease course and a clinical severity score of 2/15. In this patient, the presence of 13 RUs on 4qA generally falls outside the pathogenic range for FSHD1, as FSHD1 is typically associated with 1–10 RUs on a permissive 4qA allele. However, the observed CpG methylation level of

23% is significantly lower than the expected range for healthy individuals (30–65%), suggesting the possibility of epigenetic deregulation contributing to disease manifestation. The presence of 7 RUs on 10qA further complicates the case. Studies have suggested that interchromosomal rearrangements between 4q and 10q may, in rare cases, lead to FSHD-like presentations (Gaillard et al., 2019; Lemmers et al., 2022; Ma et al., 2022). These structural alterations may transfer a permissive 4qA haplotype to chromosome 10q, enabling DUX4 expression despite the absence of a classical 4q35 contraction. The molecular findings suggest that this patient is highly likely to be an FSHD1 case due to a complex chromosomal rearrangement.

The remaining three non-contracted cases (P5, P9, and P10) could only be evaluated based on their methylation profiles and clinical scores. P5 had a clinical severity score of 1/15, with a CpG level of 40%. She did not have a definitive FSHD diagnosis, but the neuromuscular specialist aimed to make an exclusion diagnosis. Her MC, WES, and CpG results did not support FSHD, and no alternative neuromuscular disorder was identified as a differential diagnosis. Further comprehensive genetic testing, such as whole-genome sequencing (WGS) or optical genome mapping (OGM), could help uncover deep intronic variants, copy number variations, or repeat expansions that may explain her phenotype.

On the other hand, P9 and P10 had clinical severity scores of 8/15 and 11/15, respectively. P10's CpG methylation level was 16%, which is consistent with FSHD. His severe hypomethylation and typical FSHD phenotype strongly support an FSHD diagnosis. Further OGM or long-read sequencing could help identify a possible complex rearrangement not detectable by MC or WES. Unfortunately, P9's DNA material was depleted and her CpG status could not be analyzed, and since the patient was deceased, obtaining new biological material was not possible.

P11, the 46-year-old male patient presented with neck muscle atrophy, minimal facial involvement, and mild scapular winging. Electromyography revealed low-amplitude polyphasic potential in the trapezius muscles. Genetic analysis identified a heterozygous *CAPN3* c.549delA (p.Thr184Argfs*36) frameshift variant, previously reported as pathogenic in Limb-Girdle Muscular Dystrophy type 2A (LGMD R1, calpainopathy).

Traditionally, calpainopathy (LGMD R1) is autosomal recessive, requiring biallelic *CAPN3* variants for disease manifestation. However, increasing evidence suggests that certain *CAPN3* variants, particularly truncating and missense variants, may have a dominant effect in some cases. For instance, a heterozygous 21-base pair deletion (c.643_663del21) in *CAPN3* has been linked to dominantly inherited LGMD, suggesting a dominant-negative mechanism where the mutant protein interferes with the function of the wild-type allele (Chae et al., 2016).

The c.549delA variant, which results in a truncated calpain-3 protein, is among the most common pathogenic *CAPN3* variants associated with LGMD (Fanin et al., 2004; Krahn et al., 2006). While it is typically observed in homozygous or compound heterozygous states, there have been reports of heterozygous carriers exhibiting mild symptoms or even remaining asymptomatic, suggesting that additional genetic, epigenetic, or environmental factors may influence disease expression (Tasca et al., 2016). In some cases, heterozygous truncating *CAPN3* variants have been linked to a late-onset, mild LGMD phenotype, resembling autosomal dominant inheritance (Huq et al., 2019). Some truncated calpain-3 proteins can act in a dominant-negative manner, interfering with normal calpain-3 function. This phenomenon has been observed in certain LGMD-related genes, where frameshift variants disrupt protein-protein interactions, leading to a milder, later-onset phenotype in heterozygous carriers (Chae et al., 2016). It is also possible that the patient harbors an undetected second pathogenic variant, such as a deep intronic variant affecting splicing or an epigenetic regulatory defect reducing *CAPN3* expression. Such hidden biallelic contributions have been reported in LGMD (Guglieri et al., 2008). Epigenetic and environmental modifiers (Tasca et al., 2016) or mosaicism may lead to variable expressivity of heterozygous *CAPN3* variants.

Further research is needed to elucidate the pathophysiological mechanisms underlying heterozygous *CAPN3* pathogenicity and whether dominant inheritance may be more common than previously recognized.

P12 was a 22-year-old female patient with a clinical severity score of 4/15. She had typical FSHD features, including facial involvement, bilateral upper limb weakness, and scapular winging, as well as some unusual features, such as complete eyelid closure, and the presence of left foot drop with an external foot posture. The

patient was found to carry a 4qB haplotype, which is generally considered non-permissive for DUX4 expression and thus not typically associated with FSHD1 (Lemmers et al., 2012). However, we found 4qA rearrangements involving 6–8 RUs followed by a 40 RU array in succession along with a missense VUS c.610G>A in the heterozygous state (Figure 3.6). Recent studies have unveiled those complex rearrangements within the D4Z4 region, including duplications and insertions, can significantly influence the clinical manifestation of FSHD. Nguyen et al. (2017) employed molecular combing to analyze the 4q35 region in a cohort of 586 individuals presenting with FSHD-like symptoms. Their findings revealed in some patients the existence of a recurrent duplication of the D4Z4 array coupled with a 4qA haplotype and in some of these cases variants in the *SMCHD1* gene were also identified which was argued that they might act synergistically with epigenetic regulators to modify disease severity. In P12, despite lacking a classical pathogenic D4Z4 contraction, the presence of an *SMCHD1* c.610A>G variant together with the 4qA complex rearrangement baring contracted 4qA features suggests the effect between epigenetic regulation and complex genomic architecture. Notably, consistent with the literature, P12 exhibited a D4Z4 methylation level of 24%, further supporting the notion that epigenetic dysregulation, potentially influenced by the *SMCHD1* variant and complex 4qA architecture, contributes to her FSHD phenotype.

The c.610G>A variant in the *SMCHD1* gene results in the p.Lys204Glu amino acid substitution, which is located within the GHKL (Gyrase, Hsp90, Histidine Kinase, MutL) ATPase domain at the N-terminus of the SMCHD1 protein (Figure 4.1). This domain is crucial for the protein's ATP binding and hydrolysis activities, which are essential for its role in chromatin remodeling and gene silencing. Structural studies have shown that the GHKL ATPase domain facilitates dimerization of SMCHD1, a process necessary for its proper localization to chromatin and subsequent gene repression functions. Although this variant has been reported as disease-causing in the scientific literature (Mul et al., 2018 and Lemmers et al., 2019), it is still considered VUS according to ACMG criteria. Further investigation using muscle tissue with detailed functional analysis is needed to understand the genotype-phenotype correlation of this patient fully.

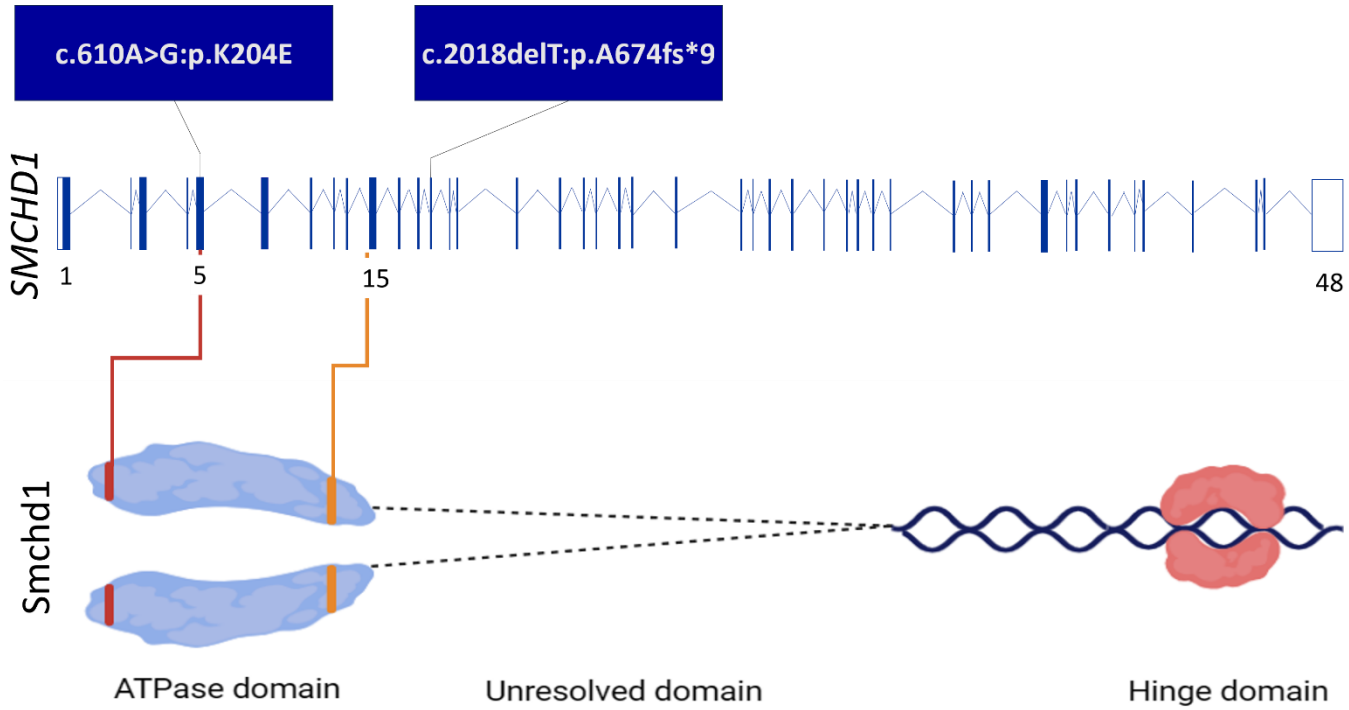

Figure 4.1: Illustration of two distinct mutations in the SMCHD1 gene (NM_015295) and their effects on the corresponding protein domains. The top panel of the figure displays the genomic structure, highlighting exons (blue boxes) and introns (blue lines) along with the precise positions of these mutations. The lower panel maps these mutations onto the SMCHD1 protein structure, illustrating their respective locations concerning the ATPase domain, the unresolved domain, and the hinge domain.

The integration of NGS into FSHD diagnostics has significantly enhanced the identification of pathogenic variants, particularly in FSHD2 and FSHD-like phenotypes, allowing for a more refined genotype-phenotype correlation. However, despite the advances in molecular combing (MC) and NGS, challenges remain in cases with ambiguous genetic findings, borderline D4Z4 repeat sizes, or complex structural rearrangements. Our findings emphasize the need for alternative genomic approaches, including long-read sequencing, optical genome mapping (OGM), and epigenetic profiling, which could provide a deeper understanding of chromatin

accessibility and regulatory mechanisms influencing *DUX4* expression. Additionally, VUS variants identified in *SMCHD1*, *DNMT3B*, and other chromatin regulators require functional validation through cellular and animal models to determine their true pathogenic potential. In unresolved cases, further research focusing on non-coding regions, deep intronic mutations, and additional epigenetic modifications will be crucial. Expanding muscle tissue-based analyses and integrating multi-omics approaches will provide critical insights into the pathogenic mechanisms underlying FSHD and overlapping myopathies, ultimately improving diagnostic accuracy, patient stratification, and future therapeutic strategies.

BIBLIOGRAPHY

Banerji, C. R. S., Panamarova, M., Hebaishi, H., et al. (2020). PAX7 and DUX4 regulate alternative splicing during myogenesis. Journal of Cell Science, 133(5), jcs238840.

Blewitt, M. E., Gendrel, A. V., Pang, Z., et al. (2008). SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation and transcriptional silencing. Proceedings of the National Academy of Sciences, 105(35), 14001–14006.

Blewitt, M. E., Gendrel, A. V., Pang, Z., et al. (2008). Smchd1, encoded by the Gt(ROSA)26Sor locus, is required for X inactivation and epigenetic silencing of repetitive elements. Nature Genetics, 40(5), 663-669.

Bosnakovski, D., Xu, Z., Gang, E. J., et al. (2008). An epigenetic switch regulates muscle differentiation and disease phenotype in FSHD. Developmental Cell, 15(5), 841-857.

Caruso, N., Herberth, B., Bartoli, M., Puppo, F., Dumonceaux, J., Zimmermann, A., ... & Helmbacher, F. (2013). Deregulation of the protocadherin gene FAT1 alters muscle shapes: implications for the pathogenesis of facioscapulohumeral dystrophy. *PLoS genetics*, *9*(6), e1003550.

Chae, J., Minami, N., Jin, Y., et al. (2016). Dominant-negative calpainopathy due to heterozygous CAPN3 variants. Brain, 139(8), 2154–2170.

Cohn, R. D., Greally, J. M., et al. (2022). Advances in long-read sequencing for FSHD research. Genome Research, 32(4), 580-594.

Cowling, B. S., Cottle, D. L., Wilding, B. R., D'Arcy, C. E., & McGrath, M. J. (2014). FHL1: A multifunctional protein implicated in myopathies and intracellular signaling. Frontiers in Physiology, 5, 508.

De Greef, J. C., Lemmers, R. J., van Engelen, B. G., et al. (2009). Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. Human Mutation, 30(10), 1449-1459.

Dixit, M., Ansseau, E., Tassin, A., et al. (2020). Somatic mosaicism and variable expressivity in FSHD: Insights from methylation profiling. Brain, 143(3), 693-708.

Ellard, S., Baple, E., Brewer, C., et al. (2020). ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. Association for Clinical Genomic Science (ACGS).

Fanin, M., Nascimbeni, A. C., Fulizio, L., et al. (2004). A rare variant of limb girdle muscular dystrophy presenting with scapuloperoneal distribution and caused by calpain-3 gene variants. Muscle & Nerve, 29(4), 575–579.

Fardeau, M., Hugon, G., Marty, I., et al. (2001). Limb-girdle muscular dystrophy type 2A. Clinical aspects and diagnostic approach. Acta Myologica, 20(2), 136–145.

Gabriëls, J., Beckers, M. C., Ding, H., et al. (1999). DUX4 as a candidate gene for FSHD. Nature Genetics, 21(1), 37-41.

Gaillard, MC., Broucqsault, N., Morere, J. et al. Analysis of the 4q35 chromatin organization reveals distinct long-range interactions in patients affected with Facio-Scapulo-Humeral Dystrophy. Sci Rep 9, 10327 (2019). https://doi.org/10.1038/s41598-019-46861-x

Gendrel, A. V., Apedaile, A., Coker, H., et al. (2013). Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. Developmental Cell, 23(2), 265–279.

Gendrel, A. V., Apedaile, A., Coker, H., et al. (2013). Smchd1-dependent and -independent pathways determine developmentally regulated X-chromosome inactivation in mice. Nature Communications, 4, 2282.

Geng, L. N., Yao, Z., Snider, L., et al. (2012). DUX4 activates germline genes, retroelements, and immune mediators: Implications for facioscapulohumeral dystrophy. Developmental Cell, 22(1), 38-51.

Gueneau, L., Bertrand, A. T., Jais, J. P., Salih, M. A., Stojkovic, T., Wehnert, M., ... & Bonne, G. (2009). Mutations of FHL1 gene cause Emery-Dreifuss muscular dystrophy and more severe phenotypes. Human Molecular Genetics, 18(4), 586-596.

Guglieri, M., Magri, F., D'Angelo, M. G., et al. (2008). Calpainopathy: A survey on the clinical and genetic heterogeneity of patients in a single neuromuscular unit. Neuromuscular Disorders, 18(7), 567–574.

Gurzau, A. D., Blewitt, M. E., Czabotar, P. E., Murphy, J. M., & Birkinshaw, R. W. (2020). Relating SMCHD1 structure to its function in epigenetic silencing. *Biochemical Society Transactions*, *48*(4), 1751-1763.

Hamanaka, K., Goto, K., Matsumoto, M., et al. (2023). Optical genome mapping as a novel approach for facioscapulohumeral muscular dystrophy diagnosis. Human Genetics, 142(3), 367-379.

Hartweck, L. M., Anderson, L. J., Lemmers, R. J., et al. (2013). A role for SMCHD1 in FSHD2 and Bosma arhinia microphthalmia syndrome. Human Molecular Genetics, 22(13), 2510–2523.

Hartweck, L. M., Anderson, L. J., Lemmers, R. J., et al. (2013). DNA methylation analysis for diagnosis and characterization of FSHD1 and FSHD2. Neuromuscular Disorders, 23(12), 956-963.

Hartweck, L., Anderson, L. J., Lemmers, R. J., & Miller, D. G. (2013). The role of SMCHD1 in FSHD and epigenetic regulation of D4Z4. *Human Molecular Genetics*, *22*(24), 4925–4934.

Huq, S., Regensburger, A. P., Stránecký, V., et al. (2019). Challenges in diagnosing FSHD: Mimicking disorders and the role of genetic testing. European Journal of Neurology, 26(6), 891–903.

Jansz, N., Keniry, A., Trussart, M., et al. (2018). Smchd1 regulates long-range chromatin interactions and epigenetic silencing. Nature Communications, 9(1), 4949.

Jansz, N., Nesterova, T., Keniry, A., et al. (2018). Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. Nature Structural & Molecular Biology, 25(9), 766-777.

Johnson, N. E., Quinn, C., Eastwood, E., et al. (2021). The FSHD Health Index (FSHD-HI): A disease-specific patient-reported outcome measure of functional impairment in FSHD. Muscle & Nerve, 64(5), 585-592.

Krahn, M., Béroud, C., Labelle, V., et al. (2006). Analysis of the CAPN3 gene in limb-girdle muscular dystrophy: Variantal spectrum and genotype-phenotype correlations. Human Variant, 27(4), 409–419.

Larsen, M., Rost, S., El Hajj, N., et al. (2020). Diagnostic value of D4Z4 methylation analysis in Facioscapulohumeral muscular dystrophy. Annals of Clinical and Translational Neurology, 7(2), 267–279.

Larsen, M., Rost, S., El Hajj, N., et al. (2020). FSHD2 and D4Z4 methylation: An epigenetic disorder affecting muscle gene regulation. Human Molecular Genetics, 29(10), 1603-1614.

Lemmers, R. J., Goeman, J. J., van der Vliet, P. J., et al. (2012). The role of SMCHD1 in FSHD2 and its interaction with D4Z4 chromatin structure. American Journal of Human Genetics, 91(6), 1014-1023.

Lemmers, R. J., de Kievit, P., Sandkuijl, L. A., et al. (2007). Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. Nature Genetics, 39(2), 137-138.

Lemmers, R. J., Tawil, R., Petek, L. M., et al. (2012). Digenic inheritance in facioscapulohumeral muscular dystrophy: SMCHD1 modifies the epigenetic state of the D4Z4 repeat region on chromosome 4q35. Nature Genetics, 44(12), 1370–1374.

Lemmers, R. J., Tawil, R., Petek, L. M., et al. (2012). DUX4 is a candidate gene for facioscapulohumeral muscular dystrophy. Science, 329(5999), 1650-1653.

Lemmers, R. J., Tawil, R., Petek, L. M., et al. (2010). D4Z4 repeat contraction and its role in FSHD diagnosis and chromatin regulation. Nature Genetics, 42(11), 985-990.

Lemmers, R. J., Tawil, R., Petek, L. M., et al. (2007). Somatic mosaicism in FSHD: Implications for disease severity and inheritance. Neurology, 68(6), 431-436.

Lemmers, R. J., van der Maarel, S. M., et al. (2010). D4Z4 repeat contraction and its role in FSHD diagnosis and chromatin regulation. Nature Genetics, 42(11), 985-990.

Lemmers, R. J., van der Maarel, S. M., van Deutekom, J. C., et al. (2002). Speculation on a role for DUX4 in FSHD pathogenesis. Annals of Neurology, 51(2), 163-166.

Lemmers, R. J., van der Vliet, P. J., Balog, J., et al. (2012). Digenic inheritance of an SMCHD1 variant and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. Nature Genetics, 44(12), 1370–1374.

Lemmers, R. J., van der Vliet, P. J., Balog, J., et al. (2015). *SMCHD1 variants in FSHD2: Epigenetic alterations and their clinical significance. Nature Genetics, 47*(12), 1527–1530.

Strafella C, Caputo V, Galota RM, Campoli G, Bax C, Colantoni L, Minozzi G, Orsini C, Politano L, Tasca G, Novelli G, Ricci E, Giardina E, Cascella R. The variability of SMCHD1 gene in FSHD patients: evidence of new mutations. Hum Mol Genet. 2019 Dec 1;28(23):3912-3920. doi: 10.1093/hmg/ddz239. PMID: 31600781; PMCID: PMC6969370.

Lutz, K. L., Holte, L., Kwan, J., et al. (2013). Hearing loss in facioscapulohumeral muscular dystrophy. Muscle & Nerve, 47(6), 835-840.

Mocciaro, E., Runfola, V., Ghezzi, P., Pannese, M., & Gabellini, D. (2021). DUX4 role in normal physiology and in FSHD muscular dystrophy. *Cells*, *10*(12), 3322.

Montagnese, FedericaTasca, Giorgio et al., (2023) 268th ENMC workshop - Genetic diagnosis, clinical classification, outcome measures, and biomarkers in Facioscapulohumeral Muscular Dystrophy (FSHD): Relevance for clinical trials Neuromuscular Disorders, Volume 33, Issue 5, 447 - 462

Moreira, S., Wood, L., Smith, D., et al. (2017). Respiratory dysfunction in facioscapulohumeral muscular dystrophy: A systematic review. Neuromuscular Disorders, 27(3), 199-208.

Mul, K., Heutinck, L. L., van den Boogaard, M. L., et al. (2016). The FSHD Clinical Severity Score (CSS): A validated disease severity measure for FSHD. Neurology, 87(15), 1609-1615.

Mul, K., Voermans, N. C., Lemmers, R. J., et al. (2018). Phenotypic variability and clinical progression in FSHD: The modifying role of SMCHD1 variants. *Brain*, *141*(1), 228–239.

Okano, M., Bell, D. W., Haber, D. A., Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell, 99(3), 247-257.

Padberg, G. W., Brouwer, O. F., de Keizer, R. J., et al. (1995). Facioscapulohumeral muscular dystrophy in the Netherlands: The clinical picture and its progression in late childhood and adolescence. Brain, 118(6), 1551-1562.

Pedersen, L. C., Inoue, K., Kim, S., Perera, L., & Shaw, N. D. (2019). A ubiquitin-like domain is required for stabilizing the N-terminal ATPase module of human SMCHD1. *Communications Biology*, *2*(1), 255.

Puppo, F., Dionnet, E., Gaillard, M. C., Gaildrat, P., Castro, C., Vovan, C., ... & Lévy, N. (2015). Identification of Variants in the 4q35 Gene FAT 1 in Patients with a Facioscapulohumeral Dystrophy-Like Phenotype. *Human mutation*, *36*(4), 443-453.

Quinzii, C. M., Vu, T. H. T., Min, K., & Hirano, M. (2008). X-linked reducing body myopathy: Clinicopathologic correlations in three families with novel FHL1 mutations. Neurology, 71(19), 1571-1576.

Ricci, E., Galluzzi, G., Deidda, G., et al. (1999). Progressive asymmetric limb muscle involvement in facioscapulohumeral muscular dystrophy and a new clinical scale. Neuromuscular Disorders, 9(6-7), 362-367.

Richard, I., Broux, O., Allamand, V., et al. (1999). Variants in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell, 81(1), 27–40.

Richards, S., Aziz, N., Bale, et al.; S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L.; ACMG Laboratory Quality Assurance Committee. (2015). Standards and guidelines for the

interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine, 17(5), 405–424.

Rousseau, J., Aubourg, P., Arnaud, P., et al. (2017). Advances in molecular combing for neuromuscular genetic diagnosis. Neurology, 89(10), 1010-1018.

Sacconi, S., Lemmers, R. J., Balog, J., et al. (2019). FSHD1+2: A severe clinical phenotype caused by coexisting D4Z4 contraction and SMCHD1 mutations. Brain, 142(3), 191-203.

Sacconi, S., Lemmers, R. J., Balog, J., et al. (2013). The role of SMCHD1 in FSHD2 and its implications for disease severity in FSHD1. Brain, 136(6), 1838-1850.

Sacconi, S., Lemmers, R. J. L. F., Balog, J., et al. (2013). The FSHD2 gene SMCHD1 is a modifier of disease severity in FSHD1. Nature Communications, 4, 2284.

Sacconi, S., Salort-Campana, E., Lemmers, R. J., et al. (2019). Genotype-phenotype correlations in facioscapulohumeral muscular dystrophy based on a large international cohort. Brain, 142(5), 1039-1057.

Sacconi, S., Salviati, L., Desnuelle, C., et al. (2019). Facioscapulohumeral dystrophy: Clinical and molecular diagnosis. Acta Myologica, 38(1), 45–54.

Schessl, J., Zou, Y., McGrath, M. J., Cowling, B. S., Maiti, B., Chin, S. S., ... & North, K. N. (2008). Mutations in FHL1 cause scapuloperoneal myopathy and distal myopathy with reduced body myopathy. Nature Genetics, 40(6), 635-639.

Scionti, I., Greco, F., Ricci, G., et al. (2012). Large-scale population analysis challenges the current criteria for the molecular diagnosis of facioscapulohumeral muscular dystrophy. *American Journal of Human Genetics*, 90(4), 628–635.

Scionti, I., Greco, F., Ricci, G., et al. (2012). *Large-scale population screening for FSHD: Insights into disease prevalence and genotype-phenotype correlation. PLoS One, 7*(4), e35689.

Selcen, D., Bromberg, M. B., Chin, S. S., & Engel, A. G. (2011). Reducing body myopathy with cytoplasmic and nuclear accumulation of FHL1 protein. Neurology, 76(18), 1664-1670.

Shalaby, S., Goldenberg, A., Al-Shalabi, S., Stojkovic, T., Richard, P., & Bonne, G. (2009). The LIM-only protein FHL1: Emerging functions and genetic implications in muscle disorders. Journal of Medical Genetics, 46(12), 791-802.

Snider, L., Geng, L. N., Bessling, S. L., et al. (2010). Molecular pathways involved in facioscapulohumeral muscular dystrophy. Journal of Molecular Biology, 395(2), 292-300.

Statland, J. M., Tawil, R. (2014). Facioscapulohumeral muscular dystrophy. Neurologic Clinics, 32(3), 721-728.

Statland, J. M., Wang, L. H., Graham, C. D., et al. (2015). Clinical and genetic characterization of late-onset facioscapulohumeral muscular dystrophy. JAMA Neurology, 72(11), 1401-1409.

Strafella C, Caputo V, Galota RM, Campoli G, Bax C, Colantoni L, Minozzi G, Orsini C, Politano L, Tasca G, Novelli G, Ricci E, Giardina E, Cascella R. The variability of SMCHD1 gene in FSHD patients: evidence of new mutations. Hum Mol Genet. 2019 Dec 1;28(23):3912-3920. doi: 10.1093/hmg/ddz239. PMID: 31600781; PMCID: PMC6969370.

Straub, V., Murphy, A., & Udd, B. (2018). 501st ENMC international workshop: Limb-girdle muscular dystrophies—nomenclature and reformed classification. Neuromuscular Disorders, 28(8), 702–710.

Tasca, G., Monforte, M., Ottaviani, P., et al. (2016). Magnetic resonance imaging in a large cohort of patients with calpainopathy: Pattern recognition and implications for clinical trials. Neuromuscular Disorders, 26(9), 725–733.

Tawil, R., Kissel, J. T., Heatwole, C., et al. (2014). Evidence-based guideline summary: Diagnosis and care of facioscapulohumeral muscular dystrophy: Report of the Guideline Development Subcommittee of the American Academy of Neurology and the American Association of Neuromuscular & Electrodiagnostic Medicine. Neurology, 83(24), 2218-2225.

Tawil, R., Kissel, J. T., Heatwole, C., et al. (2014). Facioscapulohumeral muscular dystrophy diagnostic criteria and genetic assessment. Journal of Neuromuscular Diseases, 1(1), 29-42.

Tawil, R., & Statland, J. M. (2014). FSHD: Mechanisms of pathogenesis and therapeutic development. Trends in Molecular Medicine, 20(2), 110-118.

Tawil, R., Statland, J. M. (2018). Genetic diagnostic criteria for facioscapulohumeral muscular dystrophy: An evolving framework. Muscle & Nerve, 57(6), 850-859.

Tawil, R., Statland, J. M., & van der Maarel, S. M. (2014). Facioscapulohumeral muscular dystrophy: The path to consensus on pathophysiology and therapeutics. *Neurology*, *83(6)*, 476–485.

Tawil, R., Padberg, G. W., Shaw, D. W., et al. (2014). Clinical and molecular aspects of facioscapulohumeral muscular dystrophy. Muscle & Nerve, 49(5), 635–648.

Tawil, R., van der Maarel, S. M. (2006). Facioscapulohumeral muscular dystrophy. Muscle & Nerve, 34(1), 1-15.

Tihaya, M. S., Mul, K., Balog, J., de Greef, J. C., Tapscott, S. J., Tawil, R., ... & van der Maarel, S. M. (2023). Facioscapulohumeral muscular dystrophy: the road to targeted therapies. *Nature Reviews Neurology*, *19*(2), 91-108.

Trevisan, C. P., Pastorello, E., Tomelleri, G., et al. (2020). Facioscapulohumeral muscular dystrophy: A clinical, genetic, and molecular approach. Acta Myologica, 39(2), 53–66.

Upadhyaya, M., Maynard, J., Rogers, M., et al. (1995). Molecular analysis of the 4q35 deletion in facioscapulohumeral muscular dystrophy (FSHD). Journal of Medical Genetics, 32(7), 552-560.

Urtizberea, J. A., Jimenez-Mallebrera, C., Cirak, S., et al. (2008). Diagnostic approach to the limb-girdle muscular dystrophies. Acta Myologica, 27(2), 100–106.

van den Boogaard, M. L., Lemmers, R. J., Balog, J., et al. (2016). Mutations in SMCHD1 compromise chromatin architecture and mimic polygenic FSHD2 in patients with FSHD1. Nature Communications, 7, 10613.

van den Boogaard, M. L., Lemmers, R. J., Balog, J., et al. (2016). Polygenic modulation of DUX4 expression and disease severity in FSHD1. *Brain, 139*(3), 655–668.

van Deutekom, J. C., Wijmenga, C., van Tienhoven, E. A., et al. (1993). Evidence for subtelomeric rearrangements on chromosome 4q in facioscapulohumeral muscular dystrophy (FSHD). Nature Genetics, 4(3), 223-227.

van Overveld, P. G., Lemmers, R. J., Sandkuijl, L. A., et al. (2000). Hypomethylation of D4Z4 in 4q-linked FSHD: A clinical epigenetic conundrum. Human Molecular Genetics, 9(15), 2313-2323.

van Overveld, P. G., Lemmers, R. J., Sandkuijl, L. A., et al. (2003). Hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. Nature Genetics, 35(4), 315-317.

Vasale, J., Boyar, F., Jocson, M., et al. (2015). Molecular combing compared to Southern blot for measuring D4Z4 contractions in FSHD. Neuromuscular Disorders, 25(12), 929–935.

Yinxing Ma et al. ,Exchange of subtelomeric regions between chromosomes 4q and 10q reverts the FSHD genotype and phenotype.Sci. Adv.10,eadl1922(2024).DOI:10.1126/sciadv.adl1922

Zhang, Y., Statland, J. M., & Tawil, R. (2022). *Emerging mechanisms and modifiers in FSHD: Implications for clinical management. Trends in Genetics, 38*(5), 409–421.

Wang, Y., Xu, W., Liang, Y., et al. (2020). Differential diagnosis of muscular dystrophies: Advances in genetic testing and emerging molecular therapies. Frontiers in Neurology, 11, 787.

Watson, C. M., Crinnion, L. A., Antanaviciute, A., et al. (2019). SMCHD1 variants associated with a phenotypic spectrum of craniofacial anomalies and developmental delay. European Journal of Human Genetics, 27(5), 779–788.

Watson, L. A., Wang, X., Elbert, A., et al. (2019). SMCHD1 regulates a subset of paternally imprinted genes and gates X chromosome inactivation. Nature Genetics, 51(2), 247-257.

Winokur, S. T., Barrett, K., Martin, J. H., et al. (2003). The role of FRG1, FRG2, and ANT1 in FSHD pathogenesis. Human Molecular Genetics, 12(3), 253-264.

Appendix A: …..

Appendix goes here.