DNA Methylation Profiling by Bisulfite Sequencing in FacioScapuloHumeral Muscular Dystrophy (FSHD)

by

Manar Kaptan

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

June 20, 2025

DNA Methylation Profiling by Bisulfite Sequencing in FacioScapuloHumeral Muscular Dystrophy (FSHD)

Koç University

Graduate School of Health Sciences

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

Manar Kaptan

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:

Prof. Hülya Kayserili (Advisor)

Prof. Zehra Piraye Oflazer

Prof. Zehra Oya Uyguner

Date: 20/06/2025

Dedicated to my family

For their endless love, support, and encouragement

ABSTRACT

DNA Methylation Profiling by Bisulfite Sequencing in Facioscapulohumeral Muscular Dystrophy (FSHD)

Manar Kaptan

Master of Science in Cellular and Molecular Medicine

June 20, 2025

Facioscapulohumeral Muscular Dystrophy (FSHD) is a progressive neuromuscular disorder caused by the aberrant expression of the *DUX4* gene, which is normally epigenetically silenced in somatic cells. This aberrant expression is directly linked to hypomethylation of the D4Z4 macrosatellite repeat region at chromosome 4q35, a key epigenetic alteration that plays a critical role in disease pathogenesis. FSHD is classified into two subtypes: FSHD1, which results from contraction of the D4Z4 array from 11-150 repeat units down to 1-10 repeat units, and FSHD2, which is associated with pathogenic variants in epigenetic regulators such as *SMCHD1*, *DNMT3B*, and *LRIF1*. Despite their distinct genetic origins, both forms exhibit a significant loss of DNA methylation at the D4Z4 locus, leading to *DUX4* activation and muscle degeneration.

This study aimed to implement and optimize PCR-based bisulfite sequencing, a high-resolution, single-base method for assessing cytosine methylation patterns, in order to characterize D4Z4 methylation profiles in FSHD patients, enhance diagnostic accuracy, and evaluate CpG profiling as a biomarker for disease severity. A subset of 33 patients from a larger cohort of 112 individuals was selected for detailed methylation analysis, stratified by D4Z4 repeat size and clinical subtype. Six healthy individuals were included as controls. Methylation status was evaluated in the DR1 and DR2 regions, with special focus on CpG density and correlation to clinical and genetic backgrounds. Our findings showed that 17 FSHD1 and 6 FSHD2 patients exhibited significantly reduced methylation compared to controls, while 10 patients had borderline or normal methylation levels. These results support the utility of methylation profiling in distinguishing between FSHD subtypes and clarifying cases with ambiguous genetic findings. Additionally, integration of next-generation sequencing (NGS) data from a selected group of patients enabled further investigation into potential genetic modifiers.

Methylation analysis in FSHD is crucial, as it reveals the epigenetic alterations that drive aberrant *DUX4* expression, enabling precise differentiation between FSHD subtypes and enhancing diagnostic accuracy. Our study highlights the value of bisulfite sequencing as a complementary clinical and research tool that not only improves diagnostic precision but also contributes to a deeper understanding of the epigenetic landscape in FSHD, thereby informing targeted patient management and guiding potential therapeutic interventions.

ÖZETÇE

Fasiyoskapulohumeral Musküler Distrofi (FSHD)'de Bisülfit Dizilemesi ile DNA Metilasyon Profillemesi

Manar KAPTAN

Hücresel ve Moleküler Tıp, Yüksek Lisans

Haziran 20, 2025

Fasiyoskapulohumeral Kas Distrofisi (FSHD), normalde somatik hücrelerde epigenetik olarak susturulmuş olan *DUX4* geninin anormal ekspresyonu sonucu ortaya çıkan ilerleyici bir nöromüsküler hastalıktır. Bu anormal ekspresyon, hastalık patogenezinde kritik bir rol oynayan ve önemli bir epigenetik değişiklik olan, kromozom 4q35'teki D4Z4 makrosatelit tekrar bölgesinin hipometilasyonu ile doğrudan ilişkilidir.

FSHD iki alt tipe ayrılır: FSHD1, D4Z4 dizisinin 11-150 tekrar biriminden 1-10 tekrar birimine kısalması sonucu oluşur; FSHD2 ise *SMCHD1*, *DNMT3B* ve *LRIF1* gibi epigenetik düzenleyicilerdeki patojenik varyantlarla ilişkilidir. Genetik kökenleri farklı olsa da, her iki form da D4Z4 bölgesinde belirgin bir DNA metilasyon kaybı gösterir ve bu durum *DUX4*'ün aktifleşmesine ve kas dejenerasyonuna yol açar.

Bu çalışma, sitozin metilasyon paternlerini değerlendirmek için yüksek çözünürlüklü ve tek baz duyarlılığına sahip bir yöntem olan PCR tabanlı bisülfit dizileme metodunu uygulamak ve optimize etmek, böylece FSHD hastalarında D4Z4 metilasyon profillerini karakterize etmek, tanısal doğruluğu artırmak ve CpG profil analizini hastalık şiddetinin bir biyobelirteci olarak değerlendirmeyi amaçlamıştır. 112 bireyden oluşan daha geniş bir kohorttan seçilen 33 hasta, D4Z4 tekrar uzunluğu ve klinik alt tipe göre gruplandırılarak ayrıntılı metilasyon analizine dahil edilmiştir. Altı sağlıklı birey de kontrol grubu olarak çalışmaya alınmıştır. Metilasyon durumu DR1 ve DR2 bölgelerinde değerlendirilmiş, özellikle CpG yoğunluğu ve klinik ile genetik arka planla ilişkisi üzerinde durulmuştur.

Elde ettiğimiz bulgular, 17 FSHD1 ve 6 FSHD2 hastasının kontrollerle karşılaştırıldığında anlamlı derecede azalmış metilasyon seviyelerine sahip olduğunu göstermiştir. Diğer yandan, 10 hasta ise sınırda veya normal metilasyon düzeylerine sahipti. Bu sonuçlar, metilasyon profil analizinin FSHD alt tiplerinin ayrımında ve genetik olarak belirsiz olguların netleştirilmesinde değerli bir araç olduğunu desteklemektedir. Ayrıca, seçilmiş hasta grubuna ait yeni nesil dizileme (NGS) verilerinin entegrasyonu, potansiyel genetik değiştirici (modifier) faktörlerin araştırılmasını da mümkün kılmıştır.

FSHD'de metilasyon analizi oldukça kritiktir; çünkü bu analiz, anormal *DUX4* ekspresyonuna neden olan epigenetik değişiklikleri ortaya koyarak, FSHD alt tipleri arasında hassas ayrım yapılmasına ve tanının daha doğru konulmasına olanak sağlar. Çalışmamız, bisülfit dizilemenin yalnızca tanısal hassasiyeti artıran bir yöntem olmakla kalmayıp, aynı zamanda FSHD'nin epigenetik yapısını daha iyi anlamaya katkı sağlayarak, hasta yönetimini yönlendiren ve potansiyel tedavi stratejilerine ışık tutan tamamlayıcı bir klinik ve araştırma aracı olduğunu vurgulamaktadır.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to my PI, Dr. Serpil Eraslan, for her mentorship and invaluable advice throughout this journey. Her academic expertise and leadership illuminated our path and made this thesis possible.

I am deeply thankful to my thesis advisor, Prof. Hülya Kayserili, whose remarkable academic vision and dedication to research have greatly enriched this project. Her support has not only contributed to my academic growth as a researcher but also created an inspiring academic environment at the Koç University Hospital Genetic Diagnosis Evaluation Center.

I would like to extend my sincere thanks to Dr. Şahin Avcı for his valuable insights on the clinical aspects of this study and for always sharing his expertise when needed. I am also grateful to Prof. Z. Piraye Oflazer and the Koç University Muscle Diseases Center team for their collaboration and for giving us access to such an important patient cohort. I sincerely thank all the FSHD patients and their families for their participation and support.

I feel incredibly fortunate to have shared this MSc journey with Beyza Yavuzcan, who has provided invaluable support. I would also like to thank my dear friends and colleagues, Ece Çepni (balım), Hilal Pırıl Saraçoğlu, and Şeyma Tur, for their emotional support, laughter, and constant encouragement during difficult times. You helped transform this process into a less lonely and more joyful experience.

I am grateful to my lifelong friends, Ceyda Sönmez, Aşkın Nur Özkan, and Aleyna Xhemaili, for their unwavering friendship and understanding. Special thanks to İlknur Temizkan, my childhood friend, who has walked with me through every phase of life. My last thanks to Hasan Akgöl for always being by my side and supporting me.

Most importantly, I would like to express my endless love and gratitude to my family: to my father Yılmaz Kaptan, for being the most supportive "girl dad"; to my mother Pınar Kaptan, who has always been my rock, best advisor, and source of unconditional love; to my sisters Yasenya Kaptan and my forever baby Dilara Kaptan, for being sunshine during

the darkest moments; and to my dear cat Titan, who reminded me once again what unconditional love truly means.

I am also grateful for the financial support provided by the TÜBİTAK EJP-RD-JTC2020 (Grant No: 121N274), which made the completion of this study possible. I would like to thank the Koç University School of Medicine for its institutional support during the establishment of our FSHD Diagnosis & Research Unit.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ABI	Applied Biosystems Instrument
AD	Autosomal Dominant
BAMS	Bosma Arhinia Microphthalmia Syndrome
CAPN3	Calpain-3
CCEF	Comprehensive Clinical Evaluation Form
CK	Creatine Kinase
CSS	Clinical Severity Score
CpG	Cytosine-phosphate-Guanine dinucleotide
D4Z4	D4Z4 Repeat Element
DES	Desmin
DNA	Deoxyribonucleic Acid
DNMT3B	DNA (Cytosine-5)-Methyltransferase 3 Beta
DR1	Distal Repeat 1
DR2	Distal Repeat 2
DUX4	Double Homeobox 4
EMG	Electromyography
FHL1	Four and a Half LIM Domains 1
FSHD	Facioscapulohumeral Muscular Dystrophy
HMW	High Molecular Weight
LGMD2A	Limb-Girdle Muscular Dystrophy Type 2A
LRIF1	Ligand-dependent Nuclear Receptor Interacting Factor 1
MC	Molecular Combing
MLPA	Multiplex Ligation-dependent Probe Amplification
MRI	Magnetic Resonance Imaging
NGS	Next Generation Sequencing
OMIM	Online Mendelian Inheritance in Man
PAS	Polyadenylation Site
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PolyPhen-2	Polymorphism Phenotyping v2
RUs	Repeat Units

SMCHD1	Structural Maintenance of Chromosomes Flexible Hinge Domain Containing 1
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SB	Southern Blot
TBE	Tris-Borate-EDTA
TE	Tris-EDTA
VUS	Variant of Uncertain Significance
WES	Whole Exome Sequencing

Chapter 1:
INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common inherited muscular disorder, following Duchenne muscular dystrophy and myotonic dystrophy. It is a genetic disorder characterized by progressive muscle weakness and atrophy, predominantly affecting the face, shoulder, and upper arm muscles, with symptoms worsening over time (Tawil et al., 2006). In approximately 20% of patients, the severity of muscle weakness necessitates wheelchair use in adulthood (Hobson-Webb et al., 2006). The disease exhibits variable expressivity, influenced by underlying genetic pathology, as well as epigenetic modifications, hormonal factors, and lifestyle choices, all of which impact disease onset, progression, and symptom severity (Jones et al., 2014; Sacconi et al., 2019). FSHD primarily follows an autosomal dominant inheritance pattern with incomplete penetrance, meaning that not all individuals carrying a pathogenic variant will manifest the disease (Lemmers et al., 2012).

FSHD is a genetically heterogeneous disorder with two primary subtypes: FSHD1 (OMIM #158900) and FSHD2 (OMIM #158901). Although clinically indistinguishable, these subtypes differ in their underlying genetic and epigenetic mechanisms. FSHD1, which accounts for approximately 95% of cases, results from the contraction of the D4Z4 repeat array on chromosome 4q35, leading to chromatin relaxation and aberrant *DUX4* expression. In contrast, FSHD2 arises from pathogenic variants in epigenetic regulatory genes, primarily *SMCHD1* (OMIM #614982) on chromosome 18p11, which accounts for 85% of FSHD2 cases and leads to D4Z4 hypomethylation without repeat contraction (Lemmers et al., 2010). Additional genetic subtypes have also been identified. FSHD3 (OMIM #620244) is associated with variations in *LRIF1* on chromosome 1p13, whereas FSHD4 (OMIM #606063) is caused by variations in *DNMT3B* on chromosome 20q11 (van den Boogaard et al., 2016). These subtypes further emphasize the role of epigenetic dysregulation in FSHD pathogenesis (Figure 1.1).

Despite these distinct genomic alterations, FSHD1 and FSHD2 converge on a shared molecular pathway, where epigenetic dysregulation leads to aberrant *DUX4* (Double Homeobox 4) activation in skeletal muscle (Tawil et al., 2014; Himeda et al., 2015).

Under normal conditions, *DUX4* remains epigenetically silenced in somatic tissues. However, in FSHD, its inappropriate activation leads to transcriptional deregulation, including the ectopic activation of genes typically restricted to early embryogenesis. This results in widespread cellular toxicity, oxidative stress, and progressive myofiber degeneration, culminating in muscle dysfunction (Jones et al., 2014; Banerji et al., 2017).

In more recent years FSHD is also recognized as a digenic disorder influenced by both repeat size and epigenetic modifications, reinforcing the necessity of integrating genetic and epigenetic data for improved diagnostic precision and therapeutic development (Hobson-Webb et al., 2006). A deeper understanding of these molecular mechanisms is crucial for refining diagnostic criteria and advancing targeted epigenetic therapies (Šikrová et al., 2023). Figure 1.1 illustrates the molecular interplay between genetic and epigenetic deregulation in FSHD.

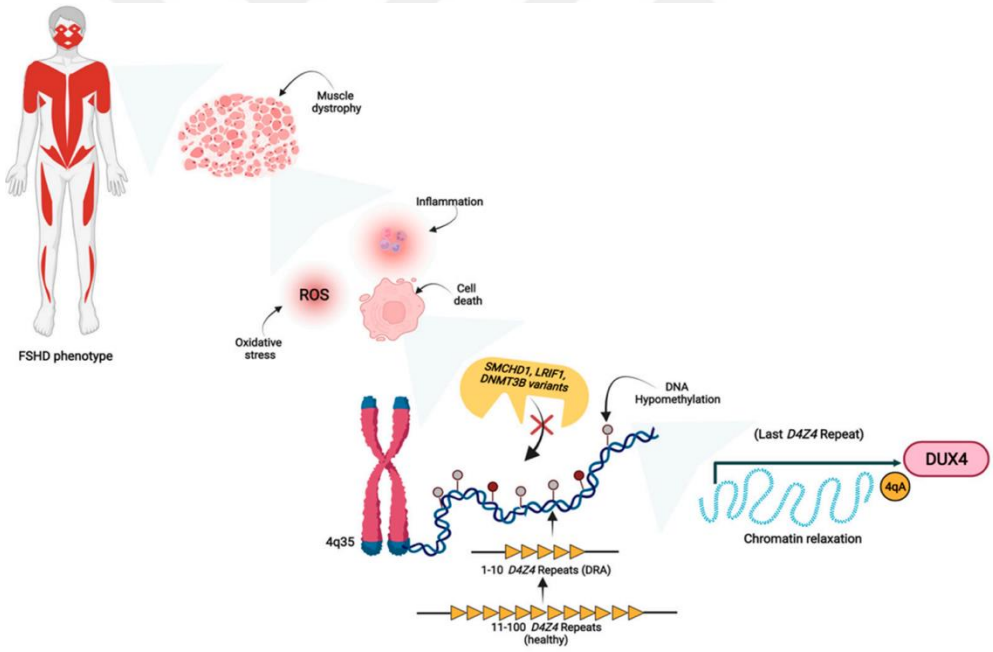

Figure 1.1: Illustration of underlying genetic and epigenetic mechanisms. (Caputo et al., 2022).

1.1 Historical Background

The understanding of FSHD has evolved significantly over time, shaped by key discoveries in clinical presentation, genetics, and molecular pathology. The condition was

first described clinically in 1884 by French physicians Louis Landouzy and Joseph Dejerine, who recognized its distinct muscle weakness pattern (Castellano et al., 2008). Their 1886 publication provided the first documentation of the familial nature of FSHD, suggesting a hereditary component to the disease. However, it was not until 1952 that definitive clinical criteria for FSHD were established, based on studies of a large, affected family in Utah, which also led to the formal introduction of the term Facioscapulohumeral Muscular Dystrophy.

Between 1952 and the 1980s, research focused on understanding FSHD inheritance patterns, confirming autosomal dominant inheritance. The first genetic insights came in the 1980s when linkage studies in FSHD-affected families identified chromosome 4 as the primary genetic locus associated with the disease. In 1992, researchers localized the FSHD causative gene region to 4q35, and in 1993, they discovered that this region contained tandem D4Z4 repeat units, with disease manifestation linked to contraction of the D4Z4 repeat array to fewer than 11 units. Initially, researchers hypothesized that D4Z4 repeat contraction resulted in the loss of function of genes in the 4q35 region, but subsequent studies demonstrated that rather than a loss-of-function mechanism, FSHD is driven by the aberrant expression of the *DUX4* gene, encoded within the last repeat unit of the D4Z4 array. In 1996, the FSHD Region Gene 1 *(FRG1*) was discovered adjacent to the D4Z4 repeat, and in 1999, the *DUX4* gene was identified, marking a major milestone in understanding FSHD pathology (Gabriëls et al., 1999).

A pivotal breakthrough came in 2007 when *DUX4* was shown to be toxic to muscle cells, with studies confirming that FSHD patients exhibit aberrant *DUX4* expression in muscle tissue. In 2009, a paradigm shift occurred with the classification of FSHD into two main subtypes: FSHD1, resulting from D4Z4 repeat contractions, and FSHD2, arising from variations in epigenetic regulators such as *SMCHD1*, *LRIF1*, and *DNMT3B* (Lemmers et al., 2009). The discovery of FSHD2 as an epigenetic disorder was further solidified in 2012 when researchers identified *SMCHD1* variations as the primary cause of D4Z4 hypomethylation, leading to chromatin relaxation and *DUX4* expression. This finding redefined FSHD as both a genetic and epigenetic disorder, influencing modern diagnostic and therapeutic strategies (Lemmers et al., 2012).

Ongoing research continues to refine the molecular mechanisms underlying *DUX4* activation and its downstream pathological effects. Advances in genetic testing and diagnostic technologies have significantly improved FSHD diagnosis, leading to increased accuracy and earlier detection. Moreover, current therapeutic efforts are focused on *DUX4* inhibition strategies, such as antisense oligonucleotides and CRISPR-based gene silencing, as well as epigenetic modification therapies aimed at restoring D4Z4 methylation to suppress *DUX4* expression. These developments mark a promising step toward targeted treatments that could potentially alter the disease course. The relentless efforts of researchers in FSHD genetics and molecular biology have paved the way for potential therapeutic strategies, offering new hope for modifying *DUX4* expression and disease progression.

1.2 Clinical and Paraclinical Findings

FSHD is characterized by a range of clinical and pathological features that primarily affect skeletal muscle structure and function. The disorder follows a progressive course, with considerable variation in severity and age of onset among affected individuals.

1.2.1 Clinical Characteristics

FSHD presents with characteristic muscle weakness and atrophy, predominantly affecting the facial, shoulder, and upper arm muscles. One of the earliest and most defining clinical features is facial muscle weakness, which manifests as difficulty closing the eyes, forming facial expressions, or smiling. This weakness can result in the "myopathic face" appearance, a hallmark of the condition (Loonen et al., 2021).

Another key feature is weakness in the scapulohumeral muscles, which typically begins in the shoulder girdle and upper arms. This leads to difficulty lifting the arms, reaching overhead, or performing activities that require upper body strength. A hallmark of FSHD is its asymmetric muscle involvement, where one side of the body is often more affected than the other, contributing to functional limitations and postural abnormalities (Statland & Tawil, 2014).

Although FSHD primarily affects the face, shoulders, and upper arms, the disease can progress to involve the lower extremities and abdominal and pelvic muscles, affecting mobility (Figure1.2). Patients often experience difficulty climbing stairs, walking long distances, or maintaining balance while standing. Routine tasks, such as lifting objects or maintaining stability, become increasingly challenging as muscle function declines (Tian et al., 2015).

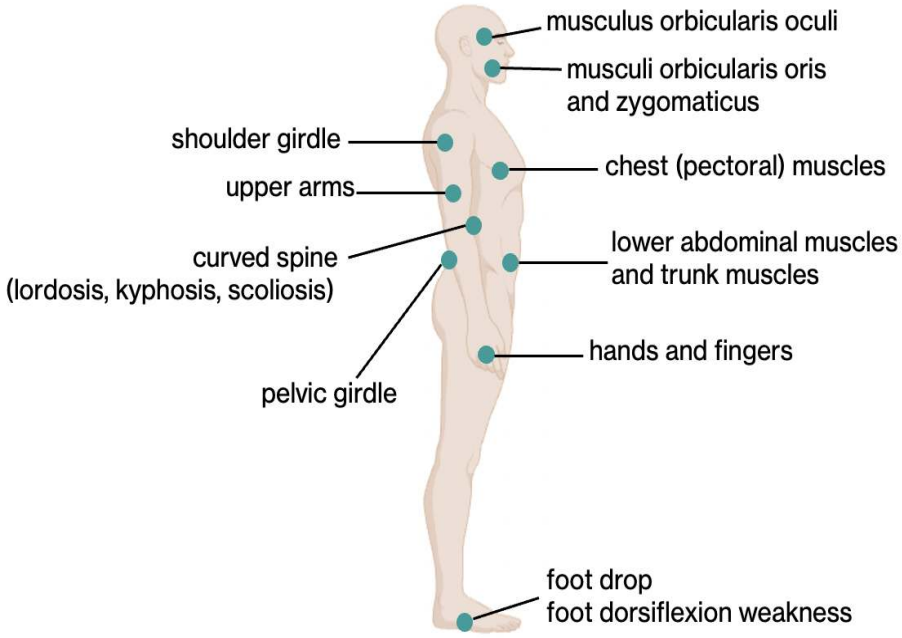

Figure 1.2: Muscle involvement in FSHD shows progressive, asymmetric weakness from the face and upper body to the lower limbs. FSHD progresses at a variable rate, with some individuals exhibiting mild symptoms, while others develop significant muscle weakness and disability over time. Onset typically occurs in the teenage years or early adulthood, but cases of childhood and late-adult onset have been reported. Diagnosis is based on clinical evaluation, genetic testing, and, in some cases, muscle biopsy or imaging studies. [Adapted from (Schätzl et al., 2021)].

1.2.2 Paraclinical Findings

The pathological hallmarks of FSHD are structural and molecular alterations in muscle tissue, influenced by genetic and epigenetic mechanisms (Schätzl et al., 2021).

The D4Z4 repeat contraction in FSHD1 and epigenetic deregulation in FSHD2 lead to chromatin relaxation and aberrant expression of the *DUX4* gene, which drives disease progression.

Muscle biopsy remains a valuable tool for assessing histopathological changes in FSHD. Inflammatory infiltrates are frequently observed in FSHD muscle biopsies, suggesting an immune response to muscle damage, though inflammation is not uniformly present. Studies show varying degrees of muscle fiber necrosis, regeneration, and fibrosis, indicating ongoing cycles of muscle injury and repair. Over time, fatty infiltration of muscle tissue becomes a characteristic feature, wherein healthy muscle fibers are replaced by fat deposits, which is detectable on MRI scans and correlates with disease progression (Gerevini et al., 2016).

Electromyography (EMG) studies can aid in diagnosis, often revealing myopathic changes characterized by short-duration, low-amplitude motor unit potentials (Dorobek et al., 2013). However, EMG findings alone are not definitive and are used in conjunction with genetic testing and imaging to confirm diagnosis. Importantly, the extent of pathological changes varies significantly among individuals, reflecting the heterogeneous progression of muscle weakness and atrophy in FSHD.

The integration of genetic, epigenetic, and histopathological findings has deepened our understanding of FSHD, improving diagnostic accuracy and paving the way for targeted therapeutic approaches.

1.2.3 Phenotypic Variability in FSHD

Phenotypic variability in FSHD encompasses a broad spectrum of clinical presentations, reflecting the heterogeneous nature of the disease. The age of onset varies widely, with symptoms typically emerging in adolescence or early adulthood, though cases have been reported from early childhood to late adulthood (Hamel & Tawil, 2018; Chin et al., 2024). While FSHD primarily affects the face, shoulders, and upper arms, some individuals exhibit predominant involvement of lower extremity muscles or other specific regions, contributing to variability in functional impairment.

The severity of muscle weakness ranges from mild to profound, often exhibiting asymmetry, where one side of the body is more affected than the other. The rate of disease

progression is highly variable; some individuals experience a slow decline over decades, while others progress more rapidly. Although facial weakness is a hallmark feature, its impact on facial expressions varies, with some patients experiencing marked impairment, while others retain near-normal facial function. Beyond muscle weakness, FSHD symptoms extend to the lower extremities, pelvic, and abdominal muscles, influencing gait, balance, and core stability (Ghasemi et al., 2022). Additionally, pain and fatigue are frequent but variable non-motor symptoms that significantly impact quality of life. The degree to which daily activities are affected differs among individuals; some maintain high functional capacity, while others experience significant limitations in mobility and self-care.

1.2.4 Gender Differences and Mosaicism in Phenotypic Variability

Gender-based differences in FSHD severity and progression have been reported, with females often exhibiting later onset and milder symptoms compared to males (Tawil et al., 2014). While both sexes can experience significant functional impairment, studies suggest that women may have less severe muscle weakness overall, possibly due to hormonal influences on muscle physiology and *DUX4* expression. However, women with FSHD are more likely to report higher levels of pain and fatigue, despite having a less aggressive decline in muscle strength (Wang et al., 2022). These differences highlight the need for personalized treatment approaches that consider both genetic and gender-related factors.

FSHD exhibits significant clinical heterogeneity, a phenomenon partly attributed to somatic mosaicism. Mosaic individuals carry both a normally sized ancestral allele and a contracted D4Z4 allele, resulting in a mixture of affected and unaffected cells (Qiu et al., 2020). Depending on the D4Z4 repeat size and the percentage of impacted cells, this genetic mosaicism may develop in the proband or a clinically unaffected parent. Variable clinical expression can result from these post-zygotic variations, which most likely originate during early zygotic cell divisions and can be caused by gene conversions with or without crossover events (Lemmers et al., 2004).

The clinical implications of mosaicism in FSHD are profound. Due to the presence of unaffected muscle cells that lessen the severity of the disease, mosaic patients

frequently exhibit milder phenotypes than would be predicted based solely on D4Z4 repeat length. However, the proportion of cells carrying the pathogenic contraction influences clinical presentation, with higher mosaicism levels correlating with more severe symptoms. Gender differences have also been observed in the context of somatic mosaicism in FSHD. Studies have reported a higher prevalence of mosaicism among unaffected female carriers compared to males, potentially due to a greater tolerance for mosaic disease alleles in females, allowing them to remain asymptomatic despite carrying the genetic variation. Conversely, male mosaic patients tend to exhibit more severe clinical manifestations, suggesting a lower threshold for disease expression (van der Maarel et al., 2000; Lemmers et al., 2004).

Although FSHD is primarily associated with a genetic defect on chromosome 4, additional genetic and epigenetic modifiers influence its clinical expression (Hamel & Tawil, 2018). The distinction between FSHD1 and FSHD2, along with D4Z4 repeat size, DNA methylation levels, and regulatory gene variants, plays a crucial role in determining disease severity and progression. Recognizing and accurately detecting mosaicism is crucial for genetic counseling and disease management, as it affects both inheritance patterns and clinical outcomes. A comprehensive understanding of phenotypic variability is essential for accurate diagnosis, prognosis, and the development of personalized therapeutic strategies for individuals affected by FSHD.

1.3 Genetic Basis of FSHD

FSHD is primarily caused by the epigenetic derepression of the *DUX4* gene, located within each D4Z4 repeat unit in the subtelomeric region of chromosome 4q35. This deregulation occurs due to distinct genetic and epigenetic mechanisms in FSHD1 and FSHD2, both of which ultimately lead to *DUX4* expression in skeletal muscle and disease progression.

1.3.1 Pathogenic Interplay in FSHD1

D4Z4 Contraction, Chromatin Dynamics, and Epigenetic Regulaters: FSHD1 is the most common form of the disease, resulting primarily from the contraction of the D4Z4 repeat region within 4q35 (Gabellini et al., 2002). In healthy individuals, the D4Z4

region comprises 11–100 repeat units that maintain a repressive chromatin structure, safeguarding against *DUX4* activation in muscle cells (Lemmers et al., 2010). However, in over 95% of FSHD1 patients, the D4Z4 repeat number is reduced to 1–10 repeats on at least one 4qA allele, leading to chromatin relaxation and partial derepression of *DUX4* expression (Gabellini et al., 2002; Lemmers et al., 2010).

FSHD is further modulated by two chromosomal variants 4qA and 4qB that determine the pathogenic potential of D4Z4 repeat contractions. The 4qA haplotype, which contains a polyadenylation signal (PAS) downstream of the *DUX4* gene, facilitates the production of stable *DUX4* mRNA when the D4Z4 array contracts, thereby directly contributing to FSHD development (Tawil & van der Maarel, 2006). In contrast, the 4qB allele lacks this PAS, which limits functional *DUX4* expression even in contracted states (Jones et al., 2014). Consequently, FSHD manifests only in individuals with a contracted D4Z4 array on a 4qA allele, whereas contractions on the 4qB allele remain non-pathogenic (Smith et al., 2017) (Figure 1.3).

Recent studies underscore the emerging significance of the D4Z4 repeat region on chromosome 10q26, a region once dismissed as non-contributory due to its apparent lack of association with FSHD (Peterson et al., 2019). However, accumulating evidence indicates that rearrangements in the 10q26 region may occasionally influence the genetic architecture of FSHD, adding an extra layer of complexity to its diagnosis (Garcia & Brown, 2020). Although 10q26 does not generally possess the same pathogenic potential as the 4q35 region, its structural variations can refine our understanding of FSHD's genetic underpinnings (Miller et al., 2018). This growing recognition of 10q26's role further accentuates the need for comprehensive methylation analysis in FSHD diagnosis. In particular, variations in D4Z4 methylation-especially among asymptomatic individuals-serve as robust diagnostic biomarkers and offer valuable insights into disease severity, thereby enhancing clinical management and informing therapeutic strategies (Williams

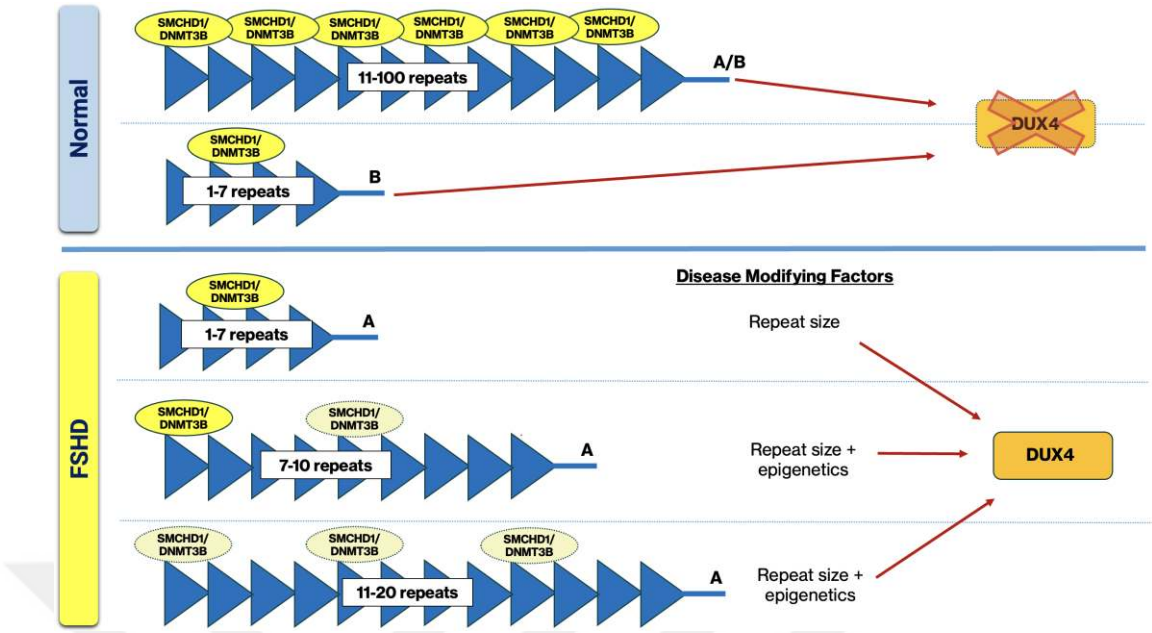

Figure 1.3: The schematic representation of the FSHD locus on chromosome 4q35. In healthy individuals, 11-100 repeats maintain a closed chromatin state, silencing *DUX4*. In FSHD, D4Z4 contraction (FSHD1) or epigenetic deregulation (FSHD2) leads to chromatin relaxation and *DUX4* aberrant expression. On permissive 4qA alleles, the final *DUX4* copy splices to an adjacent exon, stabilizing the transcript via a polyadenylation signal and driving disease progression. [Adapted from (Tawil R. and Hamel J., 2018)].

1.3.2 FSHD2 and Epigenetic Deregulation of D4Z4 Without Contraction

FSHD2 is a less common variant of the disease, resulting from variations in epigenetic regulators that normally maintain D4Z4 chromatin repression. The most frequently implicated gene in FSHD2 is *SMCHD1* (Structural Maintenance of Chromosomes Hinge Domain-Containing 1), located on chromosome 18p11, which is essential for preserving D4Z4 hypermethylation and suppressing *DUX4* expression. Additionally, variations in *DNMT3B* (DNA methyltransferase 3B) have been identified in a smaller subset of FSHD2 cases.

Unlike FSHD1, individuals with FSHD2 retain a normal D4Z4 repeat length, but due to *SMCHD1* or *DNMT3B* variations, they experience D4Z4 hypomethylation and chromatin relaxation, leading to *DUX4* aberrant expression in skeletal muscle.

Importantly, FSHD2 only manifests in individuals carrying a permissive 4qA allele, reinforcing the critical role of the 4qA haplotype in disease pathogenesis (Figure 1.3).

FSHD as a Disease Continuum: Recent research indicates that FSHD1 and FSHD2 are not entirely distinct entities but rather exist on a molecular continuum. Individuals with FSHD1 carrying a repeat size of 9-10 units often exhibit concurrent *SMCHD1* variations, demonstrating overlapping molecular mechanisms between the two subtypes. This suggests that FSHD1 and FSHD2 share a unified disease model centered on *DUX4* derepression, with varying contributions from genetic contraction and epigenetic modification. Regardless of whether the underlying cause is D4Z4 contraction (FSHD1) or chromatin modifier variations (FSHD2), *DUX4* activation in skeletal muscle remains the key driver of FSHD pathophysiology.

1.4 DUX4 Gene and Its Role in FSHD

DUX4 (Double Homeobox 4) is a transcription factor that plays a central role in the pathogenesis of FSHD. Under normal conditions, *DUX4* is epigenetically repressed in skeletal muscle, ensuring it remains inactive in adult muscle cells. However, in FSHD1, contraction of the D4Z4 repeat array on chromosome 4q35 leads to chromatin relaxation, allowing inappropriate *DUX4* expression in muscle cells (Bouwman et al., 2020). In FSHD2, variations in epigenetic regulators such as *SMCHD1* or *DNMT3B* similarly cause D4Z4 hypomethylation, leading to *DUX4* derepression (Feng et al., 2015).

Once expressed, *DUX4* disrupts muscle homeostasis by activating toxic gene networks, including germline and stem cell-associated genes typically restricted to embryonic development. This includes *ZSCAN4, TRIM43, and LEUTX*, which impair muscle differentiation, increase oxidative stress, and promote apoptosis (Bouwman et al., 2020). Additionally, *DUX4* aberrant expression alters RNA splicing, protein degradation pathways (ubiquitin ligases), and immune responses, contributing to muscle degeneration and progressive weakness. Beyond its role in FSHD pathology, *DUX4* plays a physiological role in germline development by regulating genes involved in stem cell maintenance and RNA processing. However, its aberrant expression in skeletal muscle leads to pathogenic consequences, including myogenic inhibition, immune evasion, and cellular toxicity (Feng et al., 2015).

Given its pivotal role in FSHD, therapeutic strategies are focused on blocking *DUX4* expression or activity. Current research approaches mostly involve silencing *DUX4* by antisense oligonucleotides (ASOs), siRNA, CRISPR-based epigenetic silencing, and small-molecule inhibitors targeting *DUX4*-mediated transcriptional pathways (Bouwman et al., 2020).

Understanding the molecular mechanisms of *DUX4* is crucial for developing targeted therapies and advancing precision medicine approaches for FSHD management.

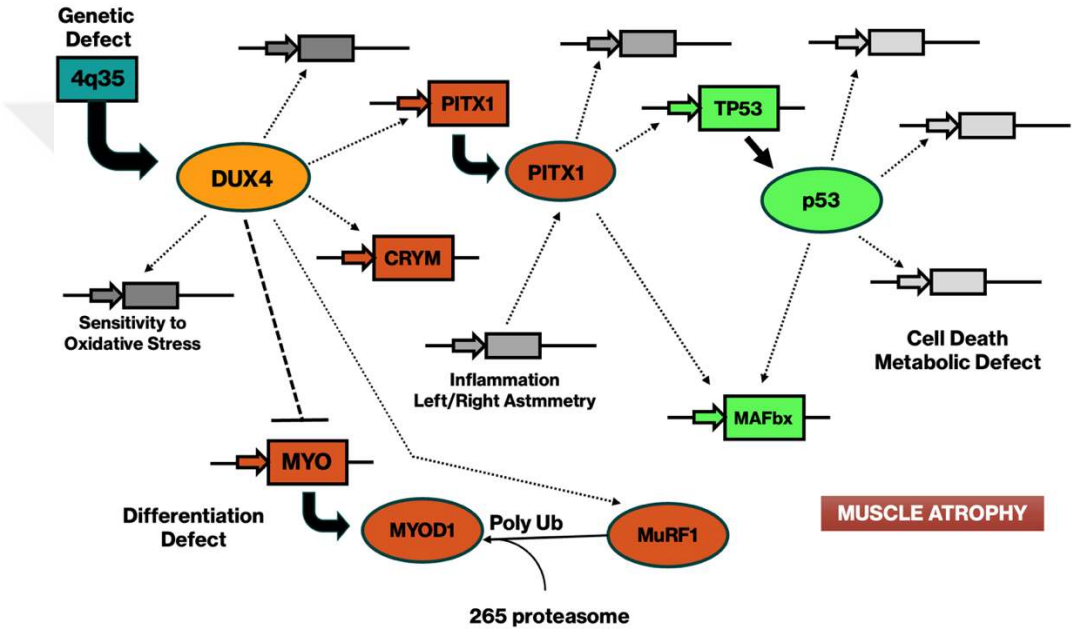

Figure 1.4: A series of events of transcriptional dysregulation in FSHD. The transcription of the *DUX4* gene is made possible by the chromatin alterations. This results in a stable, translation-ready mRNA on permissive alleles that contain the polyadenylation (poly-A) signal within the pLAM region. A transcription factor called the expressed DUX4 protein can interact with a group of target genes either directly or indirectly. The *MyoD* gene, which encodes a key transcription factor for muscle development, is suppressed by *DUX4* expression, leading to the downregulation of *MyoD* target genes in FSHD.

As a result of these events, muscle atrophy occurs. [Adapted from (Richards et al., 2012)].

DUX4 overexpression also reduces the activity of genes involved in protecting cells from oxidative stress. One key gene directly activated by *DUX4* is *PITX1*, located on chromosome 5q31. *PITX1* encodes a transcription factor known as the master regulator of hindlimb development during embryonic growth. Interestingly, unlike in 11 other neuromuscular disorders, *PITX1* is specifically expressed in the muscles of individuals with FSHD**.** Its activity leads to the production of an E3 ubiquitin ligase-an enzyme linked to inflammation and muscle wasting (atrophy) in adult skeletal muscles. One of *PITX1*'s target genes is TP53**,** a well-known regulator involved in multiple cellular functions, including metabolism, DNA repair, cell cycle control, apoptosis, and muscle atrophy.

1.5 The Epigenetic Background of FSHD

DNA methylation, the addition of methyl groups to cytosine residues within CpG dinucleotides, is a fundamental epigenetic modification that regulates gene expression by promoting chromatin compaction and transcriptional silencing. This process, catalyzed by DNA methyltransferases (*DNMT1, DNMT3A, DNMT3B*), is crucial for maintaining heterochromatin integrity in mammalian DNA (Tajima et al., 2022). The D4Z4 repeat array at chromosome 4q35, consisting of 11 to over 100 units, is normally maintained in a highly methylated and condensed heterochromatic state, particularly at CpG dinucleotides within the repeats.

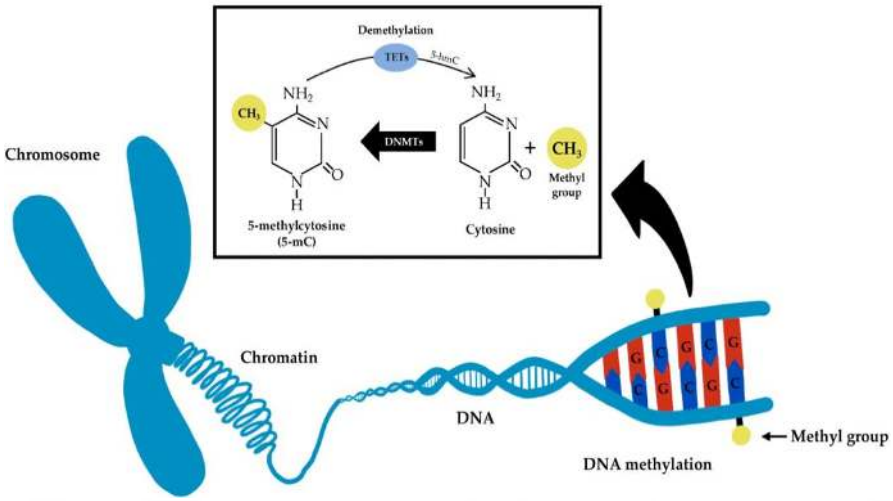

Figure 1.5: Schematic representation of DNA methylation and demethylation molecular mechanisms. (Valente et al., 2023).

Beyond DNA methyltransferases (Figure 1.5) several chromatin-modifying proteins contribute to the establishment and maintenance of this repressive state. SMCHD1 (Structural Maintenance of Chromosomes Hinge Domain-Containing 1) plays a pivotal role in long-range chromatin compaction, supporting DNA methylation stability and heterochromatin maintenance at D4Z4 (Lemmers et al., 2012; Larsen et al., 2015). *SMCHD1* cooperates with *LRIF1* (Ligand-Dependent Nuclear Receptor Interacting Factor 1) and other epigenetic regulators to ensure that D4Z4 remains transcriptionally silent by preserving methylation patterns and heterochromatin-associated histone marks, such as H3K9me3 (Zeng et al., 2009; Himeda & Jones, 2019). Loss of function in either *SMCHD1* or *LRIF1* disrupts this regulatory network (Figure 1.6), leading to D4Z4 chromatin relaxation and *DUX4* misexpression, demonstrating the redundancy and robustness of this epigenetic silencing system (Šikrová et al., 2023).

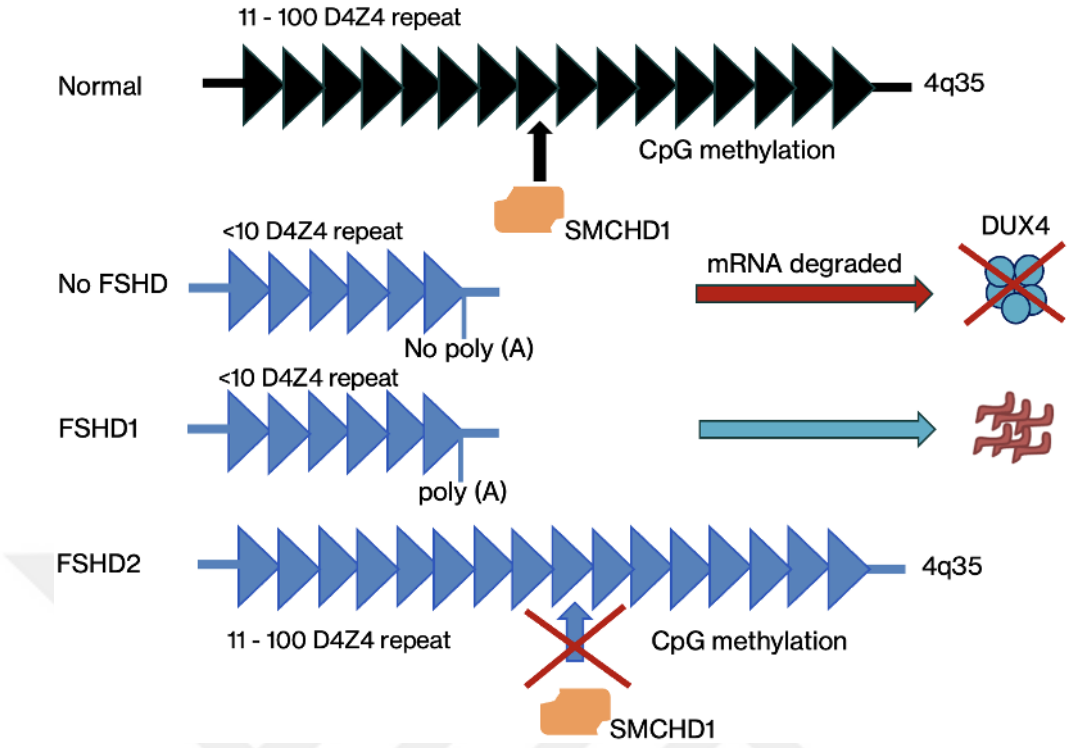

Figure 1.6: Mechanism of facioscapulohumeral muscular dystrophy (FSHD). D4Z4 repeat array on chromosome 4q35 as represented by triangles. FSHD1 patients have a D4Z4 repeat size of 1-10 units with a poly (A) signal. FSHD2 patients have 11-100 D4Z4 repeats and lack CpG methylation. [Adapted from (Bao, B., & Yokota, T., 2015)].

Early research into FSHD pathogenesis initially focused on the genetic contraction of the D4Z4 repeat array, as FSHD1 patients exhibit deletions reducing the repeat number from 11-100 units to 1-10 units (Figure 1.6). However, the identification of patients with clinically identical FSHD but an unaltered D4Z4 repeat length prompted investigations into the role of epigenetic deregulation (Van Overveld et al., 2003). A landmark study by Van Overveld et al. (2003) demonstrated that D4Z4 contraction in FSHD1 patients is associated with hypomethylation, and importantly, patients with an intact D4Z4 array but clinical FSHD features also exhibited D4Z4 hypomethylation. These findings suggested that loss of D4Z4 methylation, rather than repeat contraction alone, is a key pathogenic event, reinforcing the epigenetic basis of FSHD pathogenesis.

In FSHD2, pathogenic variants in *SMCHD1* impair its ability to maintain DNA methylation at D4Z4, resulting in global hypomethylation of this region (Figure1.6).

Consequently, D4Z4 chromatin becomes decompacted, facilitating aberrant *DUX4* transcription. Unlike FSHD1, which involves D4Z4 contraction-dependent loss of methylation, FSHD2 patients exhibit a more widespread loss of methylation, not only at D4Z4 but also at additional genomic loci, underscoring the systemic impact of *SMCHD1* dysfunction (Lemmers et al., 2014; Erdmann et al., 2023).

Thus, *SMCHD1* functions as a key epigenetic regulator (Figure 1.6), ensuring that D4Z4 methylation is properly maintained, preventing chromatin relaxation and inappropriate *DUX4* expression. Its loss of function in FSHD2 disrupts this balance, further highlighting the epigenetic complexity underlying FSHD pathogenesis (Jones et al., 2015; Greco et al., 2020).

1.6 Molecular Diagnosis of FSHD and the Role of DNA Methylation Profiling

The genetic diagnosis of FSHD relies significantly on the evaluation of D4Z4 repeat units and associated methylation patterns, with methylation analysis playing a crucial role in diagnosing challenging cases. In sporadic cases, situations with unclear family history, or when a differential diagnosis is required, methylation analysis combined with genetic testing assists in confirming the disease. For familial cases, where hereditary autosomal dominant transmission is established, clinical alignment with FSHD may suffice, potentially reducing the need for extensive genetic tests. The initial approach in FSHD molecular diagnosis involves assessing the D4Z4 repeat regions, especially since FSHD1 is the leading cause of FSHD. Traditionally, Southern blot (SB) analysis has been used to detect D4Z4 repeat contractions, which was the primary technique for many years. However, SB is labor-intensive and not high-resolution. To overcome this, molecular combing (MC) provided a higher-resolution alternative, offering precise measurements of D4Z4 repeat lengths and the ability to examine the associated haplotype with improved accuracy. While MC is more accurate and faster than SB, the latest tool in FSHD analysis is optical genome mapping (OGM) (Efthymiou S et al., 2023). This high-throughput technique offers exceptional resolution for large-scale variations-including D4Z4 repeat contractions offers a more comprehensive, genome-wide structural analysis.

However, when D4Z4 repeat contractions are not detected or if borderline sizes are detected despite strong clinical suspicion, methylation analysis becomes front and center

in diagnosing FSHD. In healthy individuals, the D4Z4 region is highly methylated (60-80%), maintaining a closed chromatin conformation that silences *DUX4* (Hartweck et al., 2013). However, in FSHD patients, a marked reduction in D4Z4 methylation levels (hypomethylation) is consistently observed (Lemmers et al., 2012).

The degree of D4Z4 hypomethylation varies among FSHD cases, ranging from as low as 3% to as high as 74%, indicating that methylation variability may contribute to phenotypic heterogeneity (Jones et al., 2014). FSHD1 patients, who carry a contracted D4Z4 array (1-10 repeat units), exhibit significantly reduced methylation levels (~10-30%), whereas FSHD2 patients, who have an intact D4Z4 repeat but variations in epigenetic regulators (e.g., *SMCHD1*, *DNMT3B*, *LRIF1*), show even more severe hypomethylation (~3-15%) (van Overveld et al., 2005; Greco et al., 2020). This extreme hypomethylation in FSHD2 supports the epigenetic basis of the disease, emphasizing that DNA methylation, not just D4Z4 contraction, is central to FSHD pathogenesis (Erdmann et al., 2023).

Several studies have explored the correlation between methylation levels and disease severity. Lemmers et al. (2015) reported that lower methylation levels at the D4Z4 region correlate with more severe clinical manifestations, reinforcing the potential role of D4Z4 methylation as a disease severity biomarker. Additionally, patients with larger repeat sizes (8-10 units) but lower methylation levels often exhibit a phenotype similar to that of more contracted D4Z4 repeats (1-3 units) (Sacconi et al., 2019)(Figure 1.7). This suggests that methylation status may serve as a more accurate predictor of disease severity than D4Z4 repeat length alone (Erdmann et al., 2023).

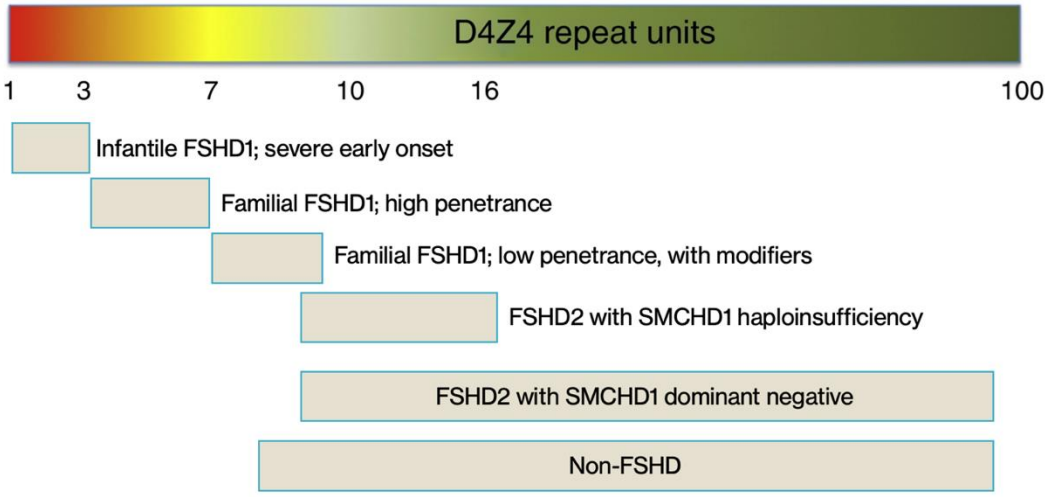

Figure 1.7: The dynamic interaction between genetic and epigenetic factors. The dynamic interaction between genetic and epigenetic factors plays a pivotal role in determining the clinical variability and severity of FSHD (red-green gradient = severe to moderate hypomethylation to normal methylation) The dynamic interaction between genetic and epigenetic factors. [Adapted from (Daxinger et al., 2015)].

Given its diagnostic and prognostic value, D4Z4 methylation analysis is being integrated into clinical workflows to differentiate FSHD1 from FSHD2 and to distinguish FSHD from other neuromuscular disorders. Advanced techniques such as next-generation bisulfite sequencing (NGS-BS) and methylation-sensitive Southern blotting are being employed to assess D4Z4 methylation status in affected individuals (Caputo et al., 2022). These findings underscore the critical role of epigenetic dysregulation in FSHD pathogenesis and highlight the potential of methylation analysis as both a diagnostic and prognostic tool for clinical applications (Hartweck et al., 2013).

Methylation analysis of the D4Z4 region, alongside sequencing of genes like *SMCHD1* and *DNMT3B*, is performed to test for FSHD2, especially when the clinical presentation matches FSHD but a clear genetic cause has not been identified (Figure 1.7). Given the growing importance of Next-Generation Sequencing (NGS) in molecular diagnostics, the expertise to perform both traditional and newer genomic techniques is rapidly becoming a necessary functional competency.

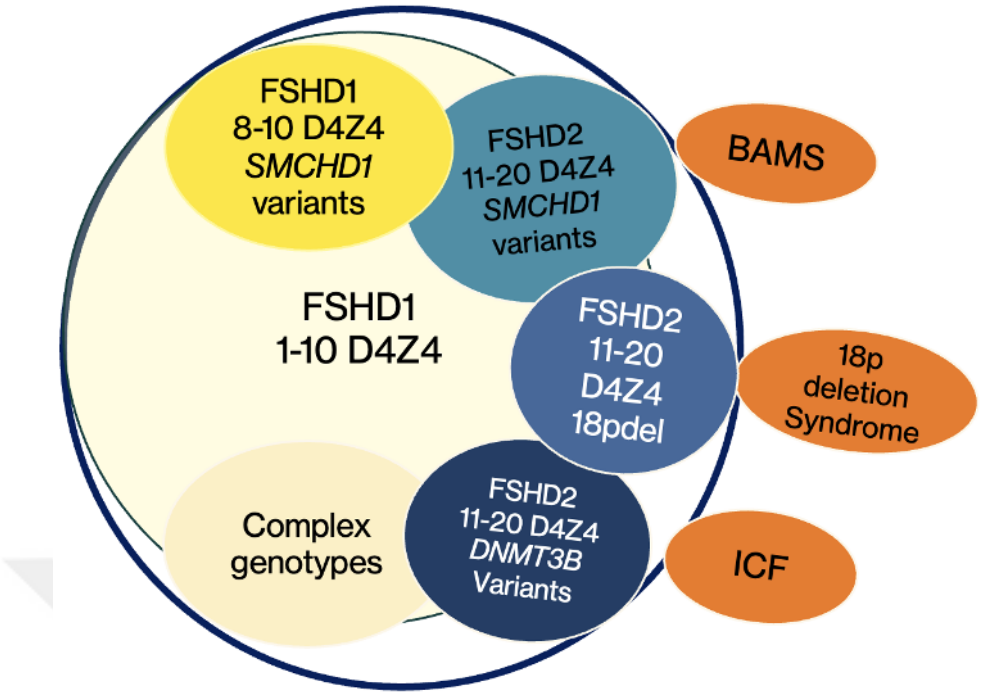

Figure 1.8: Absence of a clear molecular signature in FSHD complicates genotype-phenotype correlation. This shows the complex molecular landscape in FSHD patients with clinical manifestations. A high percentage of these patients have hypomethylation, e.g., those with *SMCHD1* or *DNMT3B* mutations. The diagram also shows that D4Z4 hypomethylation may be present in other rare disorders with *SMCHD1* mutations (BAMS), with *DNMT3B* mutations (ICF), or with deletions in the region at 18p containing the *SMCHD1* gene. [Adapted from (Salsi, V., Magdinier, F., & Tupler, R., 2020)].

Methylation levels at the D4Z4 region, which correlate tightly with disease severity, have made methylation analysis crucial for not only detecting FSHD but also for correlating clinical severity. Thus, methylation testing complements traditional genetic diagnostic tools, reinforcing its value as both a biomarker and an essential diagnostic step in complex FSHD cases.

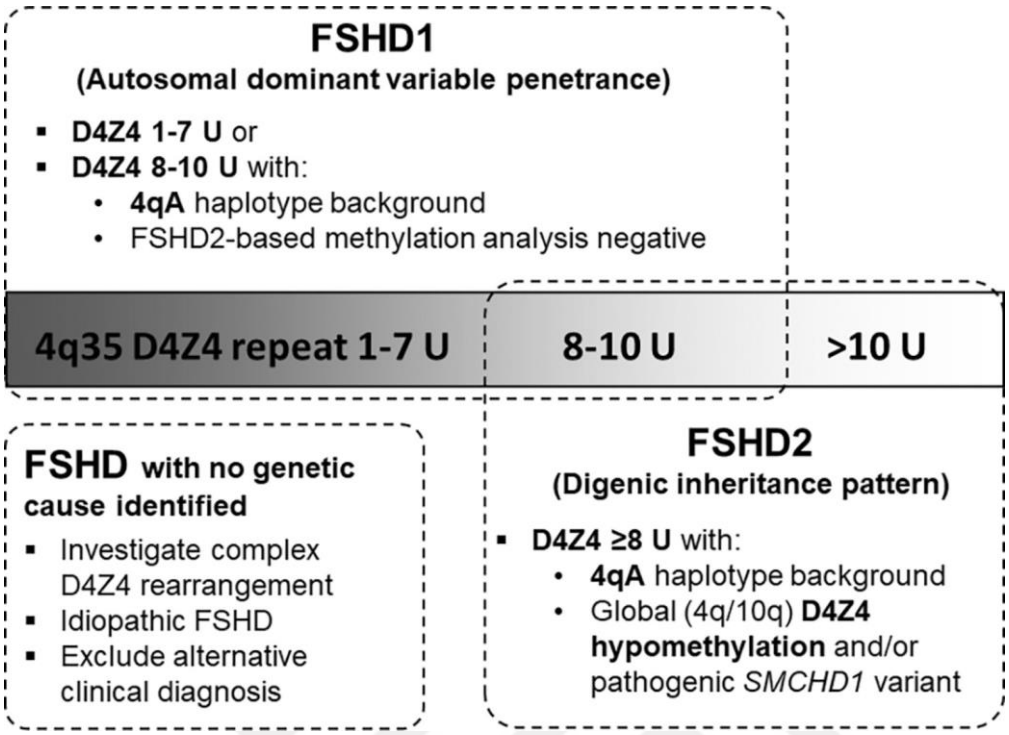

Figure 1.9: Classification of Facioscapulohumeral Muscular Dystrophy (FSHD) based on D4Z4 repeat size and genetic/epigenetic characteristics. FSHD1 is associated with 1-7 D4Z4 repeat units or 8-10 D4Z4 repeat units on a 4qA haplotype background with negative FSHD2-based methylation analysis. FSHD2 follows a digenic inheritance pattern, requiring ≥8 D4Z4 repeat units, global D4Z4 hypomethylation, and/or a pathogenic *SMCHD1* variant. Cases with no identifiable genetic cause may involve complex D4Z4 rearrangements or idiopathic FSHD. This figure illustrates the molecular classification criteria for distinguishing FSHD1, FSHD2, and uncertain FSHD cases. (Lemmers et al., 2024).

1.6.1 Methods of Methylation Analysis and Bisulfite Sanger Sequencing

Bisulfite Sanger sequencing (BS-seq) is a widely used technique for analyzing DNA methylation patterns at specific genomic loci, particularly within the D4Z4 repeat array on chromosome 4q35, a region central to FSHD pathogenesis. This method involves treating genomic DNA with sodium bisulfite, which converts unmethylated cytosines to uracil while preserving methylated cytosines, followed by PCR amplification and Sanger sequencing to determine methylation status at individual CpG sites (Jones et al., 2014).

BS-seq has proven valuable in distinguishing FSHD1 from FSHD2 by assessing D4Z4 methylation levels. FSHD1 patients exhibit partial hypomethylation, whereas FSHD2 patients show more pronounced global hypomethylation, owing to variations in epigenetic regulators such as *SMCHD1* and *DNMT3B* (Gaillard et al., 2014). A study by Gaillard et al. (2014) demonstrated that BS-seq can effectively differentiate between symptomatic FSHD patients and asymptomatic carriers, highlighting its diagnostic relevance. Additionally, Jones et al. (2014) applied BS-seq to blood and saliva samples, identifying diagnostic methylation signatures, which suggests the potential for non-invasive testing.

Despite its high accuracy and single-nucleotide resolution, BS-seq is labor-intensive and has lower throughput compared to next-generation sequencing (NGS) methods. However, it remains a gold standard for targeted methylation analysis, particularly for locus-specific FSHD profiling. While NGS-based approaches like MeDIP-seq and Nanopore sequencing provide genome-wide methylation insights, they lack the precision required for targeted D4Z4 analysis, making BS-seq a superior choice for confirming FSHD subtypes and diagnosing borderline cases (Lemmers et al., 2015; Erdmann et al., 2023) (Table 1.1).

Therefore, BS-seq remains an essential tool for understanding FSHD-related epigenetic modifications, aiding in both research and clinical diagnostics by providing accurate CpG methylation profiles within the D4Z4 region.

Table 1.1: Comparison of methylation analysis techniques in FSHD.

Technique	Resolution	Throughput	Suitability	Advantages	Limitations
BS-seq	Single nucleotide	Low	Ideal for diagnostic use	✓ High accuracy ✓ Ideal for CpG site-specific analysis ✓ Cost-effective	✓ Labor-intensive ✓ Low throughput
MSP	Single locus(qualitative)	High	Ideal for rapid diagnostic use	✓ Simple and quick ✓ Cost-effective	✓ No quantitative data ✓ Cannot distinguish between alleles
Pyrosequencing	Single nucleotide	Medium	Mostly research-based oriented	✓ Quantitative ✓ Higher throughput	✓ More expensive ✓ Requires specialized equipment
MeDIP-seq	Genome-wide	Very high	Identifies broader epigenetic changes	✓ Ideal for global methylation profiling	✓ Not suitable for targeted CpG site analysis
Nanopore Direct Methylation Sequencing	Single molecule	High	Emerging as an alternative to BS-seq	✓ Long-read capability ✓ Direct detection of 5-mC	✓ High error rates ✓ Requires bioinformatics expertise
NGS-BS	Genome-wide&targeted	High	Research-based studies	✓ High throughput ✓ Quantitative	✓ Expensive ✓ Requires extensive computational resources

(BS-seq=Bisulfite Sanger Sequencing, MSP=Methylation-Specific.PCR, MeDIP seq=Methylated DNA Immunoprecipitation, NGS-BS=Next-Generation Bisulfite Sequencing)

1.7 Treatment

Currently, there is no cure for FSHD. Therefore, symptom management and quality of life improvement remain essential to maintaining functional capacity. Physical and occupational therapy help preserve mobility and independence, while orthotic devices and respiratory support assist in managing functional impairments (Aguirre et al., 2023). Speech and swallowing therapy provide additional support for patients with facial muscle weakness. Pharmacological approaches, though still under development, include anti-inflammatory agents and myostatin inhibitors, aiming to slow disease progression (Cohen et al., 2020).

Recent advances in genetic and epigenetic therapies have fueled the development of disease-modifying treatments targeting the underlying molecular pathology of FSHD. One of the most promising approaches involves gene-silencing technologies such as antisense oligonucleotides (ASOs) and RNA interference (RNAi), which are designed to selectively inhibit *DUX4* expression-the primary pathogenic driver of FSHD (Mariot & Dumonceaux, 2022).

CRISPR-based epigenetic editing is another emerging strategy aimed at restoring normal chromatin structure and repressing *DUX4* transcription by modifying the D4Z4 regulatory region. This approach holds potential for long-term gene suppression without altering the DNA sequence, offering a targeted and durable therapeutic effect (Himeda et al., 2019).

Additionally, stem cell-based regenerative medicine is under investigation to explore its potential in replacing damaged muscle tissue and enhancing muscle regeneration. However, challenges remain in clinical translation and efficacy validation.

With continued advancements in molecular therapeutics, genetic interventions are shifting the paradigm of FSHD treatment, moving beyond symptom management toward potential curative strategies that address the underlying epigenetic dysregulation of *DUX4*. Ongoing clinical trials and translational research are critical in determining the feasibility of these next-generation therapies and their integration into FSHD clinical care.

1.8 Aim and Hypothesis

This study aims to investigate the epigenetic mechanisms contributing to the pathogenesis of FSHD and to interpret these findings concerning its genetic basis. Specifically, it seeks to optimize DNA methylation analysis of the D4Z4 locus through bisulfite sequencing, establishing it as both a diagnostic and research tool within our center.

In FSHD, disease pathogenesis is driven by hypomethylation of the D4Z4 repeat array, leading to aberrant *DUX4* expression. While healthy individuals exhibit high methylation levels that silence *DUX4*, both FSHD1 and FSHD2 are characterized by epigenetic deregulation of this region. Assessing methylation levels is essential for identifying affected individuals, especially those with borderline repeat sizes, and for diagnosing FSHD2, which occurs without repeat contraction. Methylation levels below ~35-40% at CpG sites, which are strongly associated with disease. Bisulfite sequencing enables high-resolution, quantitative methylation assessment, supporting its application in diagnostic and research settings.

The primary aim of this study is to implement and optimize DNA methylation analysis of the D4Z4 repeat array using bisulfite sequencing in a cohort of FSHD patients. The study investigates methylation profiles across patient subgroups stratified by the number of D4Z4 repeat units (RU). This comparative approach is designed to distinguish the epigenetic signatures of FSHD1 and FSHD2, improve the understanding of diagnostic methylation thresholds, and refine genotype-epigenotype correlations. Furthermore, the study evaluates the utility of bisulfite sequencing as a precise epigenetic assessment tool in both clinical and research contexts. In addition, the study incorporates the comparison of methylation findings with available next-generation sequencing (NGS) data from a subset of patients, enabling further exploration of potential genetic modifiers and epigenetic patterns.

The study cohort comprised 112 patients (92 unrelated individuals and 20 affected family members) referred to the FSHD Research and Diagnosis Unit at Koç University Hospital. D4Z4 repeat sizing was performed using the molecular combing technique.

33 patients (30 unrelated + 3 relatives) from the study cohort were selected for detailed DNA methylation profiling, and six healthy individuals were included as controls for comparative analysis using bisulfite sequencing.

By integrating epigenetic profiling with genetic data, we aim to enhance diagnostic accuracy, improve prognostic assessment, and deepen the understanding of the molecular mechanisms underlying FSHD. Ultimately, these findings will contribute to future therapeutic advancements.

Chapter 2:
MATERIAL AND METHODS

2.1 Study Cohort and Sampling

This study included 112 individuals (92 unrelated patients and 20 relatives). Informed consent was obtained from all participants before genetic testing and data collection. Clinical evaluation and preliminary diagnoses were conducted at the Neurology Department of Koc University Hospital. Clinical scoring was performed using the FSHD Comprehensive Clinical Evaluation Form (CCEF) - FSHD Evaluation Scale Section 2, as described by Ricci et al. (2016).

Following this assessment, patients were referred to the Medical Genetics Department for a detailed pedigree analysis, assessing inheritance patterns, clinical manifestations, and the potential involvement of other neuromuscular disorders. Peripheral blood samples (≥7.5 ml) were collected in K_3EDTA tubes for molecular investigations.

Our methylation study cohort consisted of 33 patients (30 unrelated, 3 family members), grouped according to four distinct genetic and clinical characteristics.

- Group 1 included 13 patients with 1-7 D4Z4 repeat units (RUs), one of whom is asymptomatic; this group represents typical FSHD1 cases, and is expected to provide insight into the methylation patterns of the FSHD1 patients
- Group 2 comprised 7 patients with 8-10 RUs, considered borderline; these patients are expected to be classified as FSHD1 or potentially FSHD1+2, based on their methylation profiles. One individual in this group is also asymptomatic.
- Group 3 included 11 patients with more than 11 RUs, representing potential FSHD2 cases or individuals requiring further testing to clarify their epigenetic status.
- Group 4 consisted of 2 patients carrying non-permissive 4qB alleles, which lack the polyadenylation signal necessary for pathogenic *DUX4* expression.

Family members with the same D4Z4 repeat unit (RU) count but differing in clinical severity were included in the methylation cohort to investigate whether variations in methylation levels could account for the observed phenotypic differences. Six healthy individuals were included as controls for comparative methylation analysis.

2.2 Kits, Consumables, and Chemicals

2.2.1 Molecular combing kits, consumables, and fine chemicals

All kits and consumables were supplied by Genomic Vision; France, and this kit comprised all buffers for agarose plug DNA isolation to obtain high molecular weight DNA. For uniform extension, DNA anchoring is provided by silanized glass coverslips (22 × 22 mm) with a hydrophobic surface. Fluorescein, digoxigenin, and biotin were fluorescently labeled probes specific to the D4Z4 repeat array, the 4qA/4qB haplotypes, and the subtelomeric regions of chromosomes 4 and 10 that were included in the FiberProbes™ mix. Detection was carried out using a set of antibodies: anti-fluorescein (FITC), anti-digoxigenin, and streptavidin-conjugated fluorophores.

Table 2.1: List of Reagents, Buffers, and Chemicals Used in the Molecular Combing.

Category	Component	Supplier / Manufacturer
DNA Extraction	FiberPrep DNA Extraction Kit	*Genomic Vision (France)*
Combing Equipment	FiberComb® Molecular Combing System	
	CombiCoverslips (silanized glass coverslips)	
Hybridization Buffers & Reagents	GV Probe	*Genomic Vision (France)*
	Buffer 2	
	Buffer 3	
Detection Buffers & Reagents	Anti-fluorescein	*Genomic Vision (France)*
	Anti-digoxin	
	Streptavidin conjugates	
	Buffer 5, Buffer 7	
Blocking	Blocking Reagent	*Genomic Vision (France)*
Denaturation	Hi-Di Formamide	*Thermo Fisher Scientific (USA)*
General Laboratory Chemicals	NH_4Cl	*Sigma-Aldrich (USA)*
	$KHCO_3$	*Sigma-Aldrich (USA)*
	Na_2EDTA (pH 7.4)	*Merck (Germany)*
	Tween 20	*Sigma-Aldrich (USA)*
	20X SSC	*Thermo Fisher Scientific*

Table 2.2: Composition of Laboratory Consumables and Equipment Used in Bisulfite Sequencing.

Consumables	Supplier / Manufacturer	Equipment	Supplier / Manufacturer
Eppendorf Tubes (1.5 / 2.0 mL)	*Eppendorf® (Germany)*	Fiber Comb	*Genomic Vision (France)*
Falcon Tubes (15 / 50 mL)	*Corning™ Falcon® (USA)*	Allegra X-14R Centrifuge	*Beckman Coulter (USA)*
Gel plug mold	*Bio-Rad (USA)*	Vortex	*Nuvemix (Nuve, Turkey)*
Pipette	*Eppendorf Research (USA)*	Slide Hybridization System	*Thermobrite StatSpin S500-12 (Abbott, USA)*
Pipette Tips	*Sarstedt (Germany), Merck (USA)*		
Microscope slides	*Thermofisher Scientific*		
Combing Reservoirs and Coverslips	*Genomic Vision (France)*		

2.2.2 Bisulfite Sequencing kits, consumables, and fine chemicals

All solid and liquid chemicals used in this study were purchased from Sigma (USA), Zymo Research, Qiagen (Germany), New England Biolabs (USA), and Merck (USA). The EZ DNA Methylation Kit from Zymo Research was used for the bisulfite treatment of DNA (Zymo Research, USA).

Table 2.3: List of Reagents, Buffers, and Chemicals Used in the Bisulfite Sequencing.

Category	Component	Supplier / Manufacturer
Enzymes	GoTaq DNA Polymerase	*Promega (USA)*
	HotStarTaq® DNA Polymerase	*Qiagen (Germany)*
Ladders	Ladder (50 bp - 1 kb)	*New England Biolabs (USA)*
Bisulfite Conversion Buffers	M-Wash Buffer	*Zymo Research (USA)*
	M-Desulphonation Buffer	
	M-Dilution Buffer	
	M-Elution Buffer	
	M-Dissolving Buffer	
	CT Conversion Reagent	
Buffer Preparation	Tris Base ($C_4H_{11}NO_3$)	*Sigma-Aldrich (USA)*
	Acetic Acid (CH_3COOH)	*Merck (USA)*
	EDTA (Na_2EDTA)	*Sigma-Aldrich (USA)*
Gel Electrophoresis Reagents	SYBR™ Safe DNA Gel Stain	*Thermo Fisher Scientific (USA)*
	TAE Buffer (1X)	*Prepared in lab*
	Agarose	*Thermo Fisher Scientific (USA)*

Table 2.4: Composition of Laboratory Consumables and Equipment Used in Bisulfite Sequencing.

Consumables	Supplier / Manufacturer	Equipment	Supplier / Manufacturer
Eppendorf Tubes (1.5 / 2.0 mL)	*Eppendorf® (Germany)*	Thermal Cyclers	*T100™ (Bio-Rad, USA), Veriti™ Thermal Cycler (Applied Biosystems, USA)*
Falcon Tubes (15 / 50 mL)	*Corning™ Falcon® (USA)*	Automated DNA Sequencing	*ABI Prism 3100 Genetic Analyzer (USA)*
PCR Tubes	*Eppendorf®*	Allegra X-14R Centrifuge	*Beckman Coulter (USA)*
Pipette	*Eppendorf Research (USA)*	Electrophoresis System	*BioRad (USA)*
Pipette Tips	*Sarstedt (Germany), Merck (USA)*	Spectrophotometer	*NanoDrop™ Thermo Scientific™ (USA)*
Zymo-Spin™ IC Columns and Collection Tubes	*Zymo Research (USA)*	Vortex	*Nuvemix (Nuve, Turkey)*

2.3 Molecular Combing Procedures

Molecular combing is a technique used to stretch and align DNA molecules uniformly on a solid surface, enabling high-resolution, in situ analysis. In this study, molecular combing was performed following the protocol provided by Genomic Vision (Molecular Combing User Guide, Version 5.3). The method relies on the surface tension generated by the receding meniscus of a liquid to exert force on DNA molecules, stretching them along glass coverslips in a parallel and linear fashion. Sample preparation for molecular combing requires approximately four days, including DNA extraction, quality control, and mounting steps to ensure suitability for downstream analysis. The protocol consists of HMW DNA isolation, molecular combing, hybridization and washing/staining and scanning. All protocols were followed as per the manufacturer's instructions (Genomic Vision, France).

2.3.1 Preparation of HMW-DNA for Molecular Combing (MC):

High Molecular Weight (HMW) DNA was required for molecular combing assays, which consist of long, intact fragments, typically ≥150–300 kb, with minimal damage or fragmentation. DNA extraction was performed using the FiberPrep® DNA Extraction Kit (Genomic Vision, Paris, France) according to the manufacturer’s guidelines.

Prior to DNA isolation, total white blood cell (WBC) counts were assessed via flow cytometry. Red blood cells were lysed, and WBCs were isolated using the following steps:

a. Incubate a minimum of 5 mL of blood (1 volume) with 3 volumes of 1× RBC lysis buffer for 10 minutes at room temperature.
b. Centrifuged at 600×g for 10 minutes, the supernatant was discarded.
c. The pellet was resuspended in 2 mL of 1× RBC lysis buffer.
d. Centrifuged again at 600×g for 2 minutes, then the supernatant was removed.
e. The WBC pellet was rehydrated in 1× PBS to achieve a concentration of 1.2×10^4-2.4×10^4 cells/µL.

To protect the genomic DNA from physical damage, the WBCs were embedded in agarose plugs at a concentration of 5×10^5-1×10^6 cells/plug. The DNA was then released from the plugs through protein digestion and membrane solubilization. Subsequently, the agarose was melted, and enzymes were used to remove it, yielding liquefied HMW DNA suitable for molecular combing analysis.

2.3.2 Combing

Samples were placed in specialized reservoirs designed to hold combing coverslips. Precautions were taken to prevent agitation, shaking, or vortexing to ensure the integrity of the DNA. Using a molecular combing device, DNA molecules were stretched uniformly onto 22 mm² barcoded and silanized FiberComb coverslips through capillary action and hydrophobic interactions. The coverslips were then baked at 60°C for 4 hours to immobilize the DNA for further analysis.

2.3.3 Hybridization, Washing, and Staining:

The immobilized DNA on the coverslips was hybridized using FiberProbe FSHD, which contains probes labeled with fluorescein, digoxigenin, and biotin. These probes specifically target the D4Z4 repeat array and the qA/qB haplotypes on chromosomes 4q35 and 10q26.

The hybridization procedure included:

1. Diluting the probe mixture in formamide for DNA denaturation at 95°C for 5 minutes.
2. Incubating the probe and coverslip at 37°C for 16-20 hours.
3. Washing the coverslips three times with 2× SSC buffer at 60°C for 5 minutes to remove excess probe.
4. Incubating the samples with a mixture of anti-fluorescein, anti-digoxigenin, and streptavidin antibodies for 20 minutes at 37°C.
5. Perform a final wash using 2× SSC + 1% Tween-20 and 1× PBS to remove unbound antibodies.

Throughout staining and washing, coverslips were protected from light to prevent photobleaching of fluorescent signals.

2.3.4 Data Acquisition and Analysis:

The stained coverslips were scanned using the FiberVision scanner, an automated, high-throughput, multi-color imaging system. The scanned data was analyzed using FiberStudio, a specialized software for molecular combing data interpretation. DNA fibers, stretched along the coverslips, displayed patterns of colored signals resembling Morse code. These signals corresponded to the D4Z4 repeat array and qA/qB haplotypes on chromosomes 4q and 10q (Figure 2.1).

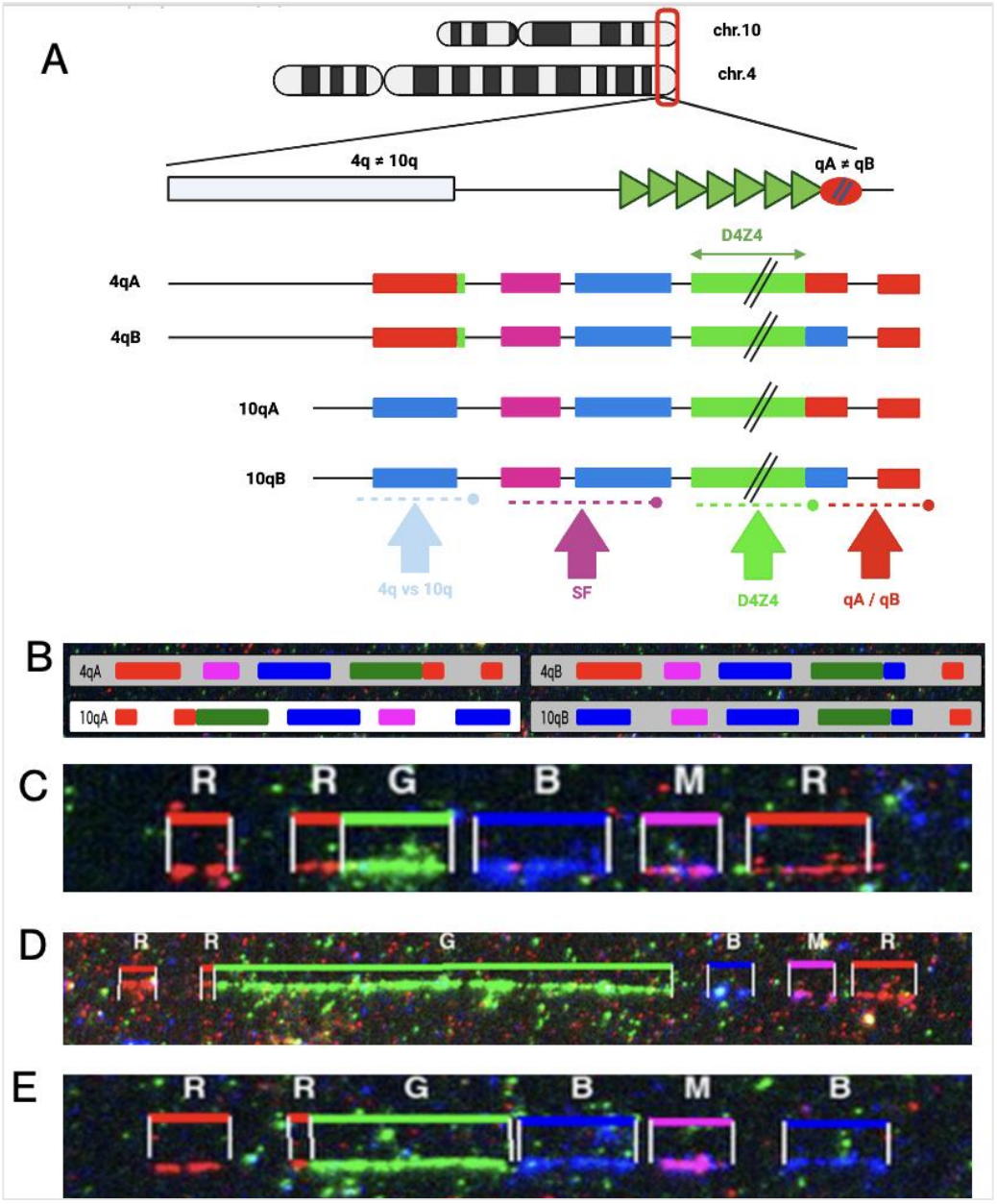

Figure 2.1: Molecular combing technique and analysis of morse code signals. (A) Schematic representation of the chromosomal regions analyzed using molecular combing. Chromosomes 4 and 10 share homologous regions (4q = 10q), and the D4Z4 repeat array is depicted with different sequence variations (qA vs. qB). The colored blocks indicate different hybridization probes targeting specific sequence elements to distinguish between 4qA, 4qB, 10qA, and 10qB alleles. (B) Expected probe hybridization patterns for distinguishing between 4q and 10q alleles. The color-coded schematic represents the unique sequence organization of each variant, aiding in the interpretation of molecular combing results. (C) Molecular combing image showing hybridization signals obtained

from fluorescently labeled probes. The Morse code pattern, formed by distinct color combinations (R: Red, G: Green, B: Blue, M: Merged), corresponds to the specific sequence organization, 4qA allele, which has contracted the D4Z4 region. (D) Detailed view of a single DNA molecule with its corresponding pattern of 4qa allele. The alignment of fluorescent signals provides an understanding of the non-contracted allele. (E) Example of molecular combing results of the 10qB allele. The consistent color-coded signals confirm the reliability of this method in distinguishing between 4q and 10q alleles.

2.4 Bisulfite Sequencing Procedures

2.4.1 Genomic DNA (gDNA) isolation

As directed by the manufacturer, DNA was extracted from 2 mL of peripheral blood using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, USA), and stored at +4°C for short-term use and at -20°C or -80°C for long-term preservation in TE buffer or molecular-grade water. Concentration and purity were assessed with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA), with acceptable purity defined by absorbance ratios of A260/A280 between 1.8 and 2.0 and A260/A230 between 2.0 and 2.2.

2.4.2 Bisulfite Modification

The EZ DNA Methylation-GoldTM Kit (USA) was used to modify genomic DNA in accordance with the manufacturer's instructions.

- **a.** According to the manufacturer’s protocol, 130 µL of CT Conversion Reagent was added to 20 µL of genomic DNA (approximately 300 ng), resulting in a final reaction volume of 150 µL. If the DNA volume was less than 20 µl, molecular-grade water was added to adjust the total volume. The recommended input DNA amount ranged between 200 and 500 ng.
- **b.** The sample was mixed, quickly spun, and placed in a thermal cycler and incubated at 98°C for 10 minutes, 64°C for 2.5 hours, and 4°C storage for up to 20 hours.

2.4.3 Clean-Up: DNA Binding, Desulphonation, Washing, and Elution

a. 600 µl of M-Binding Buffer was added to a Zymo-Spin™ IC Column placed in a Collection Tube. A total of 150 µl sample was loaded into the column and centrifuged at >10,000 x g for 30 seconds. Flow-through was discarded.

b. 100 µl of M-Wash Buffer was added to the column and centrifuged at full speed for 30 seconds. Flow-through was discarded.

c. 200 µl of M-Desulphonation Buffer was added to the column and let stand for 15-20 minutes at room temperature.

d. Step 2 was repeated, and the column was transferred onto a clean 1.5 ml microcentrifuge tube. 10 µl of M-Elution Buffer was added and centrifuged at full speed for 30 seconds to elute the DNA.

e. The elution was stored at -20 °C.

2.4.4 Quantitative Analysis of the gDNA and BSS DNA

The quality and quantity of the BSS gDNA were checked by Nanodrop 2000 (Thermo Fisher, USA).

Table 2.5: DNA Concentration and Purity Ratios of gDNA after Bisulfite Conversion.

Patient	BSS DNA Conc. (ng/ul)	260/280	260/230	Patient	BSS DNA Conc. (ng/ul)	260/280	260/230
P1	85	1.67	1.85	P21	58	1.66	2.1
P2	57	1.79	2.08	P22	80	1.72	1.99
P3	70	1.75	2.3	P23	45	1.63	1.92
P4	154	1.82	1.98	P24	45	1.8	1.98
P5	75	1.73	1.99	P25	27	1.74	1.96
P6	50	1.79	1.93	P26	43	1.72	1.85
P7	63	1.71	1.9	P27	102	1.76	1.97
P8	102	1.77	2.02	P28	140	1.75	2.18
P9	98	1.72	1.88	P29	82	1.7	2.15
P10	75	1.74	1.79	P30	69	1.72	1.65
P11	54	1.93	2.19	P31	45	1.83	1.75
P12	51	1.79	1.81	P32	65	1.8	2.1
P13	28	1.77	2.06	P33	70	1.62	1.99
P14	64	1.67	1.85	Control1	67	1.72	1.88
P15	63	1.82	2.04	Control2	81	1.82	1.79
P16	78	1.64	1.73	Control3	78	1.78	2
P17	65	1.84	1.88	Control4	82	1.8	2.05
P18	74	1.66	1.75	Control5	114	1.75	2.06
P19	100	1.82	2.3	Contol 6	72	1.7	1.89
P20	21	1.79	1.83				

2.4.5 Amplification of the Bisulfite-Modified DNA

The DR1 and DR2 regions lie immediately proximal to the D4Z4 repeat array on chromosome 4q35 and serve as critical regulatory elements. These regions are known to harbor CpG-rich sequences whose methylation status is closely linked to chromatin structure and gene silencing. Hypomethylation of DR1 and DR2 is frequently observed in FSHD1 and FSHD2 patients, reflecting a more open chromatin state. DR1 is generally regarded as more informative and discriminatory for FSHD diagnosis, especially when interpreting borderline or complex cases. DR2 can provide complementary data, but its interpretation requires caution due to variability and lower specificity. Therefore, DR1 has been established as the gold standard.

The methylation status of the patients was assessed via bisulfite sequencing. The primers used to amplify the DR1 and DR2 regions were obtained from a previously published study and target the flanking CpG-free sequences of the D4Z4 region (Hartweck et al., 2013), which contains the 4qA haplotype, allowing for equal amplification of both methylated and unmethylated alleles (Table 2.6). The composition of the amplification reaction and thermal cycler conditions are given in Table 2.7 and Table 2.8, respectively.

DR1 fragment length 251bp with 29 CpGs; located 563-814 bp from the KpnI site at the start of D4Z4, DR2 fragment length 276bp, with 22 CpGs; located 998-1271 bp from the KpnI site at the start of D4Z4.

Table 2.6: List of Primers used in PCR.

Region	Primers (5' 3')	PCR Product Length (bp)	Annealing Temp. (°C)	Number of CpGs
DR1	**DR1F:** GAAGGTAGGGAGGAAAAG **DR1R:** ACTCAACCTAAAAATATACAATCT	251 bp	51	29
DR2	**DR2F:** AAATATGTAGGGAAGGGTGTAAGTT **DR2R:** CTTAAATATACCAAACCCTCTCTCC	276 bp	54	22

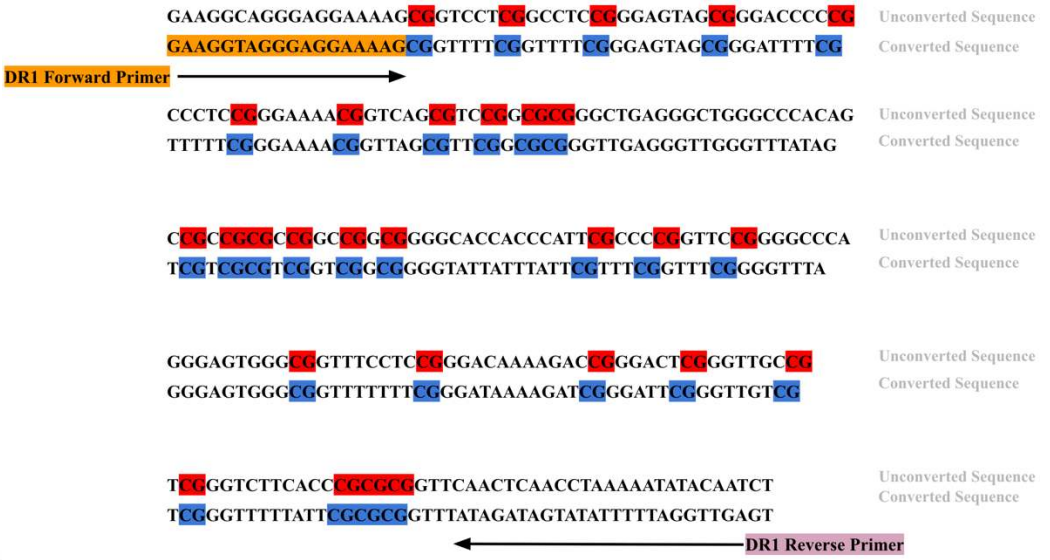

Figure 2.2: Detail of DR1 amplicon sequence. The upper line shows the unconverted genomic sequence, whereas the lower line indicates the bisulfite-converted sequence. The 29 CpG sites contained in the DR1 region are highlighted in blue and red.

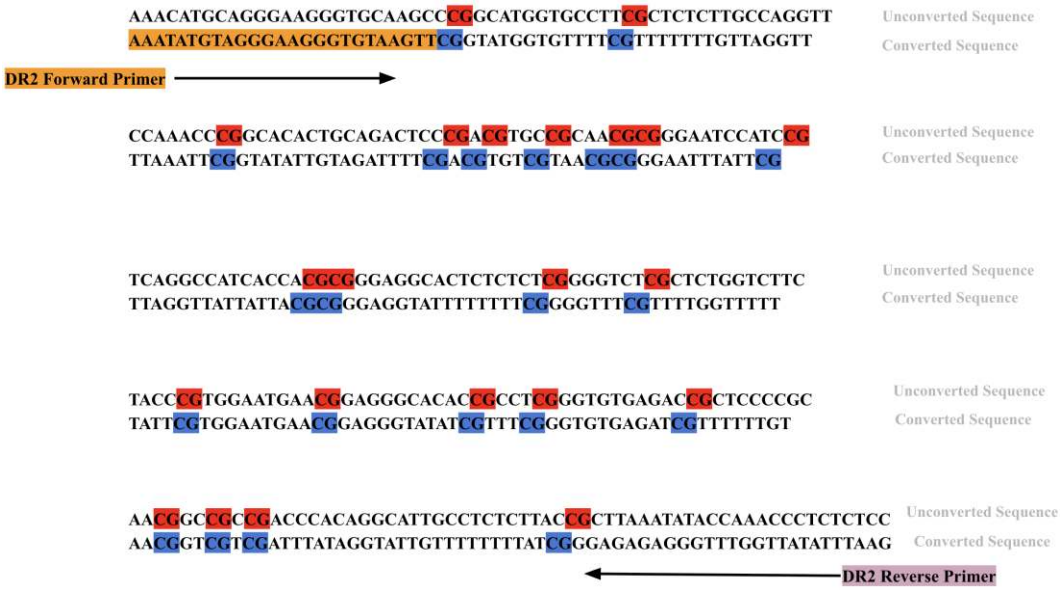

Figure 2.3: Detail of DR2 amplicon sequence. The upper line shows the unconverted genomic sequence, whereas the lower line indicates the bisulfite-converted sequence. The 22 CpG sites contained in the DR2 region are highlighted in blue and red.

Table 2.7: Methylation PCR Reaction Composition.

Compositions	Final Volume (25 µl)	Final Concentration
Buffer	5 µl	1X
$MgCI_2$	2 µl	4 mM
dNTP	0.25 µl	0.2 mM (200 µM)
Forward Primer	0.5 µl	0.2 µM (200 nM)
Reverse Primer	0.5 µl	0.2 µM (200 nM)
DNA Template	X	100 ng total
Go Taq Polymerase	0.25 µl	1.25 u
Nuclease free water	to total volume	

Table 2.8: Thermal Cycling Protocol.

Stage	Temperature	Time	Cycle
Stage 1	94 °C	3 min	1 cycle
Stage 2	94 °C	40 s	40 cycles
	48 °C*	40 s	
	72 °C	1 min	
Stage 3	72 °C	5 min	1 cycle
Stage 4	4 °C	hold	infinite

* Vary according to the annealing temperature.

DR1 Primer annealing temperature: 48 °C

DR2 Primer annealing temperature: 54 °C

2.4.6 Gel Electrophoresis

A total of 5 µL of the PCR product was run on 1.5% standard agarose gels at 100 V for 25 min. The fragments were visualized using SYBR Safe DNA Gel Stain (Thermo Fisher Scientific), which is used for gel staining. Gel Doc XR+ Gel Documentation System (BioRad) was used for the visualization of gel (Figure 2.4)

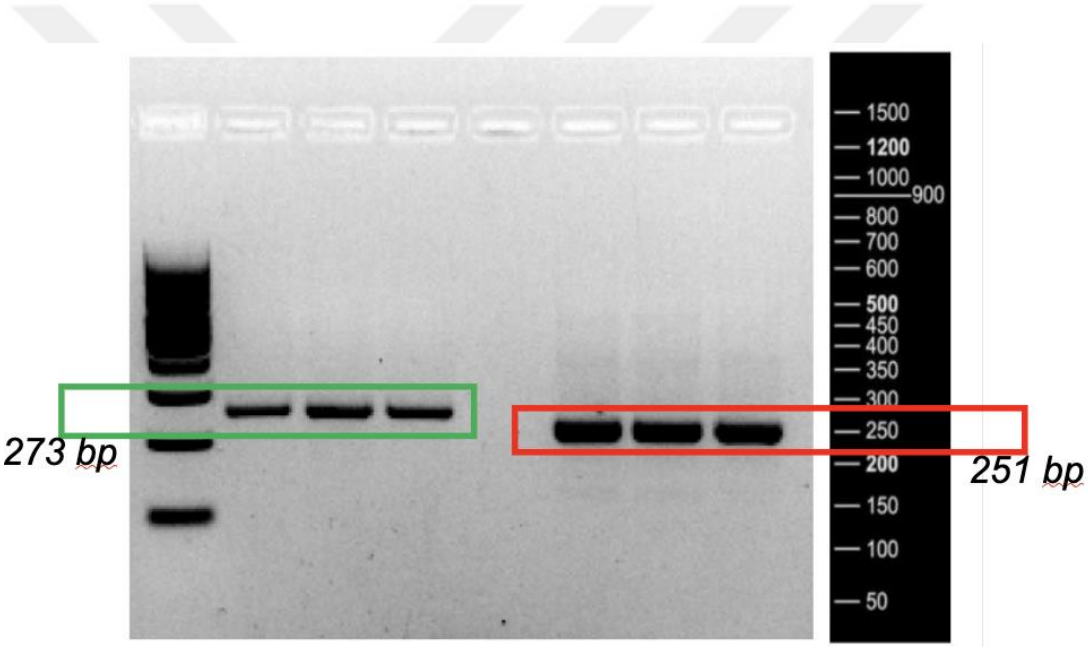

Figure 2.4: Gel electrophoresis results of PCR amplification for DR1 and DR2 regions. The 273 bp bands (highlighted in green) correspond to the DR2 region, while the 251 bp bands (highlighted in red) correspond to the DR1 region. A 50 bp DNA ladder was used as a molecular size reference.

2.4.7 BSS Sanger Sequencing

BS-PCR purifications were done using ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher, USA), and post-Sanger reactions were purified with BigDye XTerminator™ Purification Kit (Thermo Fisher, USA) and post-Sanger sequencing

BigDye™ Terminator v3.1 Cycle Sequencing Kit was used according to the manufacturer's instructions.

Table 2.9: ExoSAP cleanup and Sanger sequencing workflow, including reagent volumes, thermal cycling conditions, instrument, and software used.

Step	Reagent / Component	Volume / Conc.	Conditions / Notes
1. ExoSAP Purification	ExoSAP-IT (Applied Biosystems™, Thermo Fisher)	1 µL	Add to 5 µL PCR product.
	PCR amplicon	5 µL	
	Incubate at 37 °C for 15 min (enzymatic degradation), then 80 °C for 15 min (enzyme inactivation). Store at 4 °C (≤20 h).		
2. Sanger Sequencing Setup	BigDye Terminator v3.1 Ready Reaction Mix	2 µL	Applied Biosystems™ (Thermo Fisher) premixed dye terminators.
	5X Sequencing Buffer	3 µL	
	Primer (forward or reverse)	4 µL at 3.2 pmol/µL (~12.8 pmol total)	
	Purified PCR product	3 µL	
	Nuclease-Free Water	9 µL	Bring to final volume of 21 µL.
3. Thermal Cycling	96 °C × 1 min; 35 cycles of [96 °C × 10 s → 50 °C × 5 s → 60 °C × 4 min].		
4. Post-Run Analysis	ABI 3130/3500 Genetic Analyzer		Instrument trademarks: Applied Biosystems™ (Thermo Fisher, USA).
	FinchTV, 4Peaks, Geneious		Sequence trace visualization and methylation quantitation with online/software tools.

2.5 Analysis and CpG Calculations

Sanger sequence chromatograms were analyzed using Finch TV, 4Peaks, and Geneious Prime tools. The Cs and Ts peak areas were selected, and the cytosine methylation (Cmet) percentage was determined. Methylation percentages at specific CpG sites were defined as the cytosine peak height (Forward primer) or guanine peak height (Reverse primer), divided by the peak height sum of cytosine and thymine (Forward primer) or guanine and adenine (Reverse primer). For each sample, the methylation of 29 CpG sites in the DR1 region and 22 CpG sites in the DR2 region was calculated separately and then averaged. The obtained average value was accepted as the patient's methylation value.

2.5.1 Reference-based CpG Mapping and Methylation Profiling

The reference sequences of DR1 and DR2 were compared to the bisulfite-converted sequences. First, the chromatogram (AB1) data were imported into Geneious Prime and aligned to reference sequences specific to the DR1 and DR2 regions to facilitate the visualization of CpG sites and verify the accuracy of the amplification. By comparing cytosines and thymines, the software automatically identified methylation patterns at the described CpG patterns. Geneious also generated graphical representations of methylation density for each region. In the analysis, thymines (T) were interpreted as unmethylated, and cytosines (C) were interpreted as methylated. The ratio of cytosines to the total number of cytosines and thymines was used to determine the methylation levels for each CpG site. Additionally, statistical summaries (standard deviation, mean methylation levels) were provided. This can be expressed as: Methylation Percentage = [(Number of Methylated Reads) / (Total Reads (C + T)] × 100. When C and T peaks overlapped, chromatograms were imported into 4Peaks to confirm peak heights and ensure accurate methylation calculations. Baseline errors, noise, and ambiguous peaks were corrected or flagged for removal. Cytosine (C) (methylated) and thymine (T) (unmethylated) peak heights were measured for every CpG site (Figure 2.5). The methylation percentage was calculated using the ratio of C to (C + T), as illustrated in Figure 2.5: Methylation Percentage = [(C Peak Height) / (C Peak Height + T Peak Height)] x100.

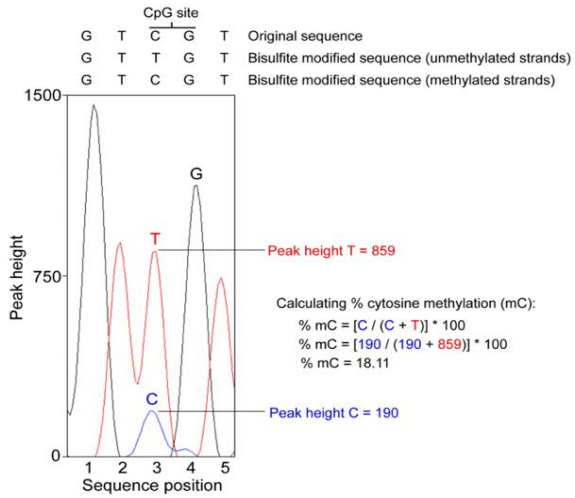

Figure 2.5: Example of methylation percentage calculation based on peak height analysis from Sanger sequencing chromatograms. The figure illustrates how cytosine (C, methylated) and thymine (T, unmethylated) peaks are identified and quantified. The methylation percentage is calculated using the formula: % mC = [C / (C + T)] × 100. In this example, the C peak height is 190, and the T peak height is 859, resulting in an 18.11% methylation value.

Methylation percentages for all CpG sites were calculated individually, and the average value was used as the final methylation percentage for each sample (Huichalaf C et al., 2014). All samples were analyzed in technical triplicate to ensure reproducibility of methylation measurements.

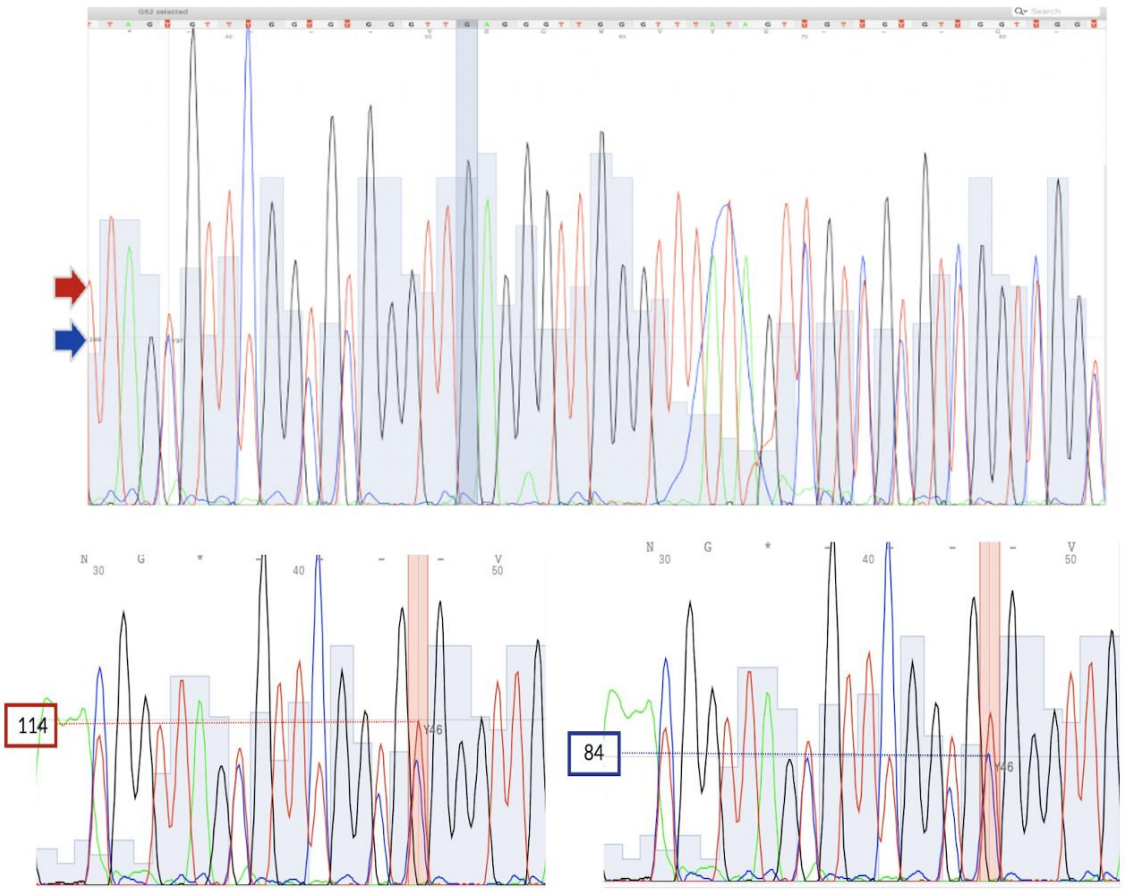

Figure 2.6: Methylation value calculation by chromatogram peak heights. An example of how to perform data analysis of a single CpG site with the use of the forward primer for sequencing. Chromatogram peaks for thymine and cytosine are compared to

determine the average level of methylation for each site within a given sample. For one CpG site, T peak height: 114, C peak height: 84. The methylation rate for this site is 42.4%.

Following each of these procedures, which were made with different tools, the ratio of cytosine to total cytosine and thymine signals was used to determine the methylation percentages. Sites having unclear sequencing data or outlier CpG sites were not included in the final analysis. The average methylation percentage for all CpG sites was used to determine the overall methylation level for the DR1 and DR2 regions. The mean methylation levels of DR1 and DR2 were examined in the patient and control samples. The possible contribution of variations in methylation patterns to the pathophysiology of FSHD was assessed.

2.6 Incorporation of NGS Data for Complementary Genetic Analysis

The study included complementary genetic data from next-generation sequencing (NGS) in addition to the methylation profiling done by bisulfite sequencing. This allowed for a more comprehensive understanding of the genomic changes associated with FSHD. NGS analysis data were obtained from Beyza Yavuzcan, MSc, as part of her master's dissertation titled 'The Diagnostic Role of Next-Generation Sequencing on Facioscapulohumeral Muscular Dystrophy (FSHD) and Related Phenotypes,' which was defended in February 2025.

Chapter 3:
RESULTS

3.1 Molecular Combing Results

All patients were evaluated by molecular combing (MC) to determine the D4Z4 array size. MC was performed for diagnostic purposes in 92 index patients and their 20 relatives to investigate FSHD1. Among the 92 index patients, 47 were male (51%) and 45 were female (49%). Based on the molecular combing results, patients were categorized as either definite FSHD1 cases or suspected FSHD cases requiring further differential diagnosis.

The disease was familial in 50 patients (54%), while 42 patients (46%) had isolated (sporadic) presentations. The age of disease onset was before 20 years in 63 patients (68%), including approximately 10% with early-onset disease (before age 10). The remaining 29 patients (32%) were classified as having late-onset FSHD (Figure 3.1).

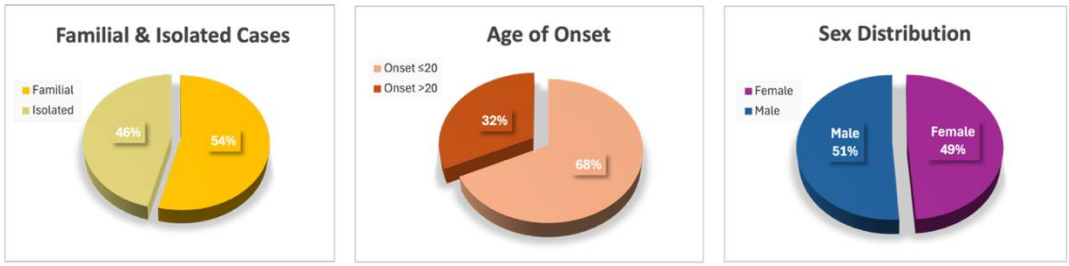

Figure 3.1: Demographic and Clinical Distribution of the FSHD Cohort.

D4Z4 profiling by MC identified FSHD1 in 78 out of 92 patients (84.8%) who had 1-10 repeat units (RUs) on the 4qA allele. Two patients (2.2%) carried only 4qB alleles and lacked the permissive 4qA haplotype. Among the 92 patients, 62 (67.4%) had 1-7 RUs, 16 (17.4%) had 8-10 RUs, and 12 (13.0%) had ≥11 RUs. (Figure 3.2). Among their 20 relatives, 7 had 1-7 RUs, 2 had 8-10 RUs, and 11 had more than 11 RUs.

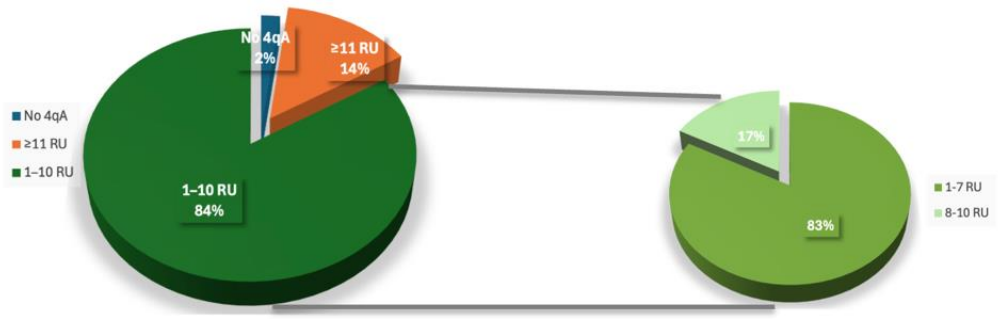

Figure 3.2: Distribution of D4Z4 Repeat Unit Sizes in the FSHD Cohort.

3.2 Clinical Evaluation Results

Comprehensive clinical scoring was performed for each patient using the FSHD Clinical Evaluation Scale (Ricci et al., 2016), which included assessments of facial, scapular, upper limb, lower limb, pelvic, and abdominal muscle involvement. Each category was scored from 0 to 2 or 0 to 5, depending on severity, and the total clinical severity score was calculated by summing all parameter scores.

Table 3.1 presents the clinical findings of 33 patients assessed using this scale. These data were later correlated with molecular findings, including D4Z4 repeat sizes, methylation levels, and NGS results, for comprehensive diagnosis and classification.

FSHD clinical scoring was used to analyze 33 patients in total. There were 13 females, ~%39 and 20 males, ~%61 in the cohort. The age of onset changed from early childhood to late adulthood; 2 patients had onset before the age of 10, and 17 patients before the age of 20. When these patients were clinically evaluated, the scapular girdle was found in all patients, whether they had a high clinical score or not. Clinical scores ranged from 1 to 15. Three patients scored the minimum value of 1; Patient 3, who had a maximum score of 15/15, is a typical illustration of severe FSHD. Six patients exhibited a severe phenotype (scores of 10 or above), whereas eight patients showed a mild phenotype (scores of 3 or less).

Table 3.1: Summary of Clinical Evaluation Scores for the FSHD Patient Cohort.

Tablo1			FSHD Comprehensive Clinical Evaluation Parameters						
Patients	Gender	Age Of Onset	Facial Weakness	Scapular Girdle Involvement	Upper Limbs Involvement	Legs Involvement	Pelvic Girdle Involvement	Abdominal Muscle Involvement	Clinical Score
Patient 1	M	17	1	2	2	0	2	0	7
Patient 2	F	12	1	3	2	2	2	0	10
Patient 3	F	5	2	3	2	2	5	1	15
Patient 4	M	n/a	1	2	0	0	0	0	3
Patient 5	M	30	2	2	1	1	2	1	9
Patient 6	F	13	2	3	1	0	2	1	9
Patient 7	M	19	1	3	2	1	2	1	10
Patient 8	M	15	2	2	2	0	1	1	8
Patient 9	F	20	2	2	0	0	0	1	5
Patient 10	M	18	0	1	0	0	0	0	1
Patient 11	F	12	2	2	0	1	1	1	7
Patient 12	M	20	1	2	1	0	1	1	6
Patient 13	M	16	1	2	1	1	2	0	7
Patient 14	F	45	1	2	0	0	2	1	6
Patient 15	F	n/a	0	1	0	0	0	0	1
Patient 16	M	17	1	2	1	1	2	1	8
Patient 17	M	37	1	2	2	0	1	1	7
Patient 18	M	16	2	2	0	0	0	1	5
Patient 19	M	34	1	3	1	0	2	1	8
Patient 20	F	12	2	3	0	1	1	1	8
Patient 21	M	12	2	3	2	0	2	1	10
Patient 22	F	40	2	3	1	0	1	1	4
Patient 23	M	8	2	3	0	0	0	1	6
Patient 24	M	17	0	1	1	0	0	0	2
Patient 25	M	55	0	2	0	2	0	0	4
Patient 26	M	10	1	3	2	2	2	1	11
Patient 27	F	13	0	1	0	0	1	0	3
Patient 28	F	15	1	2	1	0	0	0	4
Patient 29	M	44	0	1	0	0	0	0	1
Patient 30	F	19	2	3	0	0	0	1	6
Patient 31	M	44	0	2	0	0	0	0	2
Patient 32	F	24	0	3	2	2	5	1	13
Patient 33	M	39	1	1	1	0	0	0	3

3.3 NGS Results

To complement the methylation data, whole-exome sequencing data from a previous study were incorporated for selected patients to explore potential pathogenic variants. Table 3.2 summarizes the genetic alterations, including variant descriptions, ACMG classifications, and inheritance patterns.

Table 3.2: Summary of Variants Detected by Next-Generation Sequencing (NGS).

Patients	Gene	Variant position	Variant Description	Effect	ACMG Classification	Inheritance
P16	*SMCHD1*	18:2674116	NM_015295: c.2018delT	p.Ala674Leufs*9	Likely Pathogenic	Heterozygous
P18	*SMCHD1*	18p11.32p11.31	seq[GRCh37] 18p11.32p11.31(2538999-2960903)x1	421kb del (*LOC727896*, ***SMCHD1***, *EMILIN2*, *LIPN2*)	Pathogenic	Heterozygous
P25	*FHL1*	X:135289975	NM_001159699.2: c.404G>T	p.Gly135Val	VUS	Heterozygous
P27	*CAPN3*	15:42652149	NM_000070: c.146G>A	p.Arg49His	Pathogenic	Homozygous
P28	*SMCHD1*	18:2706423	NM_015295: c.610A>G	p.Lys204Glu	VUS	Heterozygous
P29	*CAPN3*	15:42680001	NM_000070: c.549delA	p.Thr184Argfs*36	Pathogenic	Heterozygous
P32	*DES*	2:219420254	NM_001927.3: c.643G>A	p.Val215Met	VUS	Heterozygous

3.4 Quantitative Analysis of Methylation Levels

The methylation patient cohort was selected from our larger group of 112 individuals, who had been previously evaluated through clinical and molecular assessments.

A total of 33 patients, 30 unrelated (12 females and 18 males) and 3 relatives, representing a spectrum of D4Z4 repeat unit (RU) sizes, were included in the study, enabling stratified analysis across subgroups. Six healthy controls (3 males and 3 females) were randomly selected.

Three patients, P4 (father of P3), P6 (brother of P7), and P18 (son of P19), were specifically selected from affected families to enable intra-familial comparison of methylation profiles and investigate potential correlations between methylation levels and clinical phenotype.

The cohort of 30 unrelated patients included 11 patients with 1-7 RUs, 6 patients with 8-10 RUs, and 11 patients with ≥11 RUs. In the remaining 2 cases, only non-permissive 4qB alleles were detected.

DNA methylation levels in the DR1 and DR2 regions were analyzed in technical triplicate. The mean methylation percentages along with their standard deviations (SDs) are presented in Table 3.3. For the interpretation of methylation levels, values below 40% were considered hypomethylated and supportive of FSHD; values between 40% and 60% were classified as borderline or "grey zone," requiring integration with D4Z4 repeat size,

age of onset, and clinical findings. Values above 60% were interpreted as unmethylated, not consistent with FSHD-associated epigenetic profiles.

In the DR1 region, methylation levels ranged from 12.00 ± 2.51% (Patient 18) to 57.00 ± 1.68% (Patient 31). Overall, eight patients exhibited DR1 methylation levels below 25%, while eight patients had levels exceeding 45%. Standard deviations for DR1 measurements were generally low, with 30 out of 33 patients showing standard deviations (SDs) below 5%, indicating stable methylation across CpG sites within this region.

In the DR2 region, methylation levels ranged from 39.00 ± 4.02% (Patient 16) to 73.00 ± 4.95% (Patient 31). Similar to DR1, the majority of DR2 SD values were below 5%, reflecting consistent technical reproducibility across samples. When comparing regional methylation levels, the DR2 region consistently showed higher methylation than the DR1 region across all patients. Both DR1 and DR2 methylation profiles showed variability across patients, and these measurements were used for further stratification and correlation with clinical and genetic characteristics in subsequent analyses.

Table 3.3: Mean methylation percentages and standard deviations (± SD) for DR1 and DR2 regions are presented for each patient.

Patients	DR1 Region Average Meth Level (%) ± SD	DR2 Region Average Meth Level (%) ± SD	Patients	DR1 Region Average Meth Level (%) ± SD	DR2 Region Average Meth Level (%) ± SD
Patient 1	*34.00 ± 3.68*	53.00 ± 2.75	Patient 18	*12.00 ± 2.51*	41.00 ± 4.50
Patient 2	*43.00 ± 4.90*	68.00 ± 4.85	Patient 19	*18.00 ± 4.90*	58.00 ± 4.34
Patient 3	*20.00 ± 2.50*	42.00 ± 3.97	Patient 20	32.00 ± 5.00	49.33 ± 4.96
Patient 4	*45.00 ± 4.04*	53.00 ± 5.00	Patient 21	*20.00 ± 4.51*	48.00 ± 5.00
Patient 5	*27.00 ± 2.51*	58.00 ± 3.51	Patient 22	*23.00 ± 4.50*	52.00 ± 4.30
Patient 6	*30.00 ± 4.36*	49.00 ± 4.92	Patient 23	18.00 ± 2.65	**42.00 ± 4.95**
Patient 7	*27.00 ± 3.78*	51.00 ± 2.12	Patient 24	*48.00 ± 4.04*	63.00 ± 2.83
Patient 8	*28.00 ± 4.00*	47.00 ± 4.98	Patient 25	*45.00 ± 4.50*	57.00 ± 4.91
Patient 9	*36.00 ± 3.50*	55.00 ± 3.75	Patient 26	*16.00 ± 3.79*	48.00 ± 4.03
Patient 10	*52.00 ± 3.06*	62.00 ± 5.01	Patient 27	*50.00 ± 1.63*	61.00 ± 2.07
Patient 11	*30.00 ± 3.05*	**58.00 ± 4.78**	Patient 28	*26.00 ± 4.16*	40.00 ± 4.77
Patient 12	*34.00 ± 4.04*	42.00 ± 3.50	Patient 29	*55.00 ± 1.69*	71.00 ± 5.10
Patient 13	*29.00 ± 2.94*	49.00 ± 2.85	Patient 30	*24.00 ± 4.58*	52.00 ± 4.05
Patient 14	40.00 ± 4.86	**53.00 ± 5.00**	Patient 31	*57.00 ± 1.68*	73.00 ± 4.95
Patient 15	*48.00 ± 2.86*	55.00 ± 3.93	Patient 32	*55.00 ± 1.69*	63.00 ± 3.74
Patient 16	*18.00 ± 2.08*	39.00 ± 4.02	Patient 33	54.00 ± 4.10	60.00 ± 2.90
Patient 17	*27.00 ± 3.68*	50.00 ± 2.84			

3.5 Integrated Findings for Final Diagnosis

To refine molecular classification and support accurate diagnosis, methylation profiling results were evaluated in conjunction with D4Z4 repeat unit sizing, available NGS data, clinical findings, and age of onset. (Table 3.4)

3.5.1 1-7 RU Subset Results

Among the 11 patients with 1-7 D4Z4 RUs, the average DR1 methylation level was 33.0% (SD ± 8.6%), with values ranging from 20% to 52%. When patients with DR1 methylation levels above 40% (n = 3) were excluded, the adjusted mean dropped to 29.78% (SD ± 4.8%), reflecting the core hypomethylated group. In contrast, DR2 methylation levels in this subgroup were consistently higher, with a mean of 53.0% and

a range of 42% to 68%. These results demonstrate greater variability in DR1 methylation compared to DR2, which appeared more stable among patients with contracted D4Z4 alleles. All patients in this group were ultimately diagnosed with FSHD1. Although P2 and P10 exhibited unexpectedly elevated DR1 methylation levels of 43% and 52%, respectively, they were classified based on their RU count and supportive clinical presentation. Notably, molecular combing results for P2 and P4 revealed FSHD somatic mosaicism, P2 has two 4qA alleles, with the contracted allele representing approximately 15% of the total 4qA signal, and two 4qA alleles were found in P4, one of which was contracted in three RUs and made up about 27% of the total 4qA alleles. This pattern is compatible with somatic mosaicism.

3.5.2 Subset 8-10 RU

In the subgroup of 6 patients with borderline D4Z4 repeat sizes (8-10 RUs), DR1 methylation levels showed considerable variability, with a mean of 29.5% (SD ± 13.44%) and a range of 12.0% to 48.0%. In contrast, DR2 methylation levels were more consistent, with a mean of 47.83% (SD ± 6.46%) and a range of 39.0% to 55.0%. However, the available clinical and molecular data indicated that this subgroup could not be considered clinically and genetically concordant, unlike the 1- 7 RU subset. As such, each case in this group required individual evaluation for accurate classification at final diagnosis.

FSHD1: Patients P14 and P20 had DR1 methylation levels of 40% (upper boundary) and 32%, respectively, with corresponding D4Z4 repeat counts of 8 and 9 RUs. Although NGS data were not available for P14 and pathogenic variants were not detected by WES for P20, the combination of CpG profiles and D4Z4 repeat sizes was sufficient to classify them as FSHD1. P15, a nearly asymptomatic female, had a methylation level of 48% (grey zone), while P17, a male with disease onset in his teens, showed a methylation level of 27%. Both had 9 RUs, and NGS analysis did not identify any pathogenic variants in FSHD-associated or other neuromuscular disease genes. These patients were classified based on their D4Z4 repeat sizes and/or CpG methylation profiles.

FSHD1+2: P16 and P18 had significantly low DR1 methylation levels of 18% and 12%, respectively. Both patients had disease onset in their late teens and carried 9 RUs.

Whole-exome sequencing revealed pathogenic variants in the *SMCHD1* gene, indicating genotypic features consistent with both FSHD1 and FSHD2.

3.5.3 ≥11 RU Subset Results

Among the 11 patients with ≥11 D4Z4 repeat units, four were classified as FSHD2, two as FSHD with complex genomic characteristics. Two patients were excluded from an FSHD diagnosis, and one of them was identified as having limb-girdle muscular dystrophy, the other was identified with myofibrillar myopathy, and the last three were considered likely non-FSHD cases. The clinical and molecular data indicated that this subgroup was both clinically and genetically heterogeneous. Notably, the methylation levels displayed a clustered distribution, reflecting underlying variability in pathogenic mechanisms.

-FSHD2: P21, P23, P26, and P30 were diagnosed with FSHD2, based on DR1 methylation levels well below the 40% threshold (20%, 18%, 16%, and 24%, respectively) and clinical features consistent. Although no definitive pathogenic variants were identified by NGS analysis in (P21, P23, and P26), and WES data were not available for P30, the diagnosis was supported by hypomethylation profiles.

-FSHD with complex genomic characteristics: P22 was identified as a case of FSHD based on an intermediate D4Z4 repeat size, reduced DR1 methylation level (23%), and evidence suggestive of a potential structural rearrangement between the 4q and 10q loci. No significant variants were detected in the patient's NGS analysis. P25, another atypical case, carried three 4qA alleles-two uncontracted and one contracted-with repeat units of 23, 61, and 1 RU, and allele ratios of 41%, 50%, and 9%, respectively. Whole-exome sequencing (WES) revealed a variant of uncertain significance (VUS) in the *FHL1* gene, which may contribute to the clinical phenotype (Section 3.3, pg. 45; Table 3.2).

-Differential Diagnosis: P27 and P32 were excluded from an FSHD diagnosis due to a high DR1 methylation level (50% and 55%) and the detection of variants by Whole-exome sequencing. P27 has a likely pathogenic variant in the *CAPN3*:c.146G>A ;p.(Arg49His) which supports an alternative diagnosis of limb-girdle muscular dystrophy. P32 has a heterozygous missense variant in *DES:*c.643G>A; p.(Val215Met), which is

classified as a variant of uncertain significance (VUS). Her high CpG methylation profile, clinical manifestations and NGS analysis support the rule out the FSHD-related pathogenesis and suggest DES-related myofibrillar myopathy.

-Non-FSHD: P24, P31 and P33 were classified in the uncertain diagnosis category due to inconclusive molecular findings and the absence of definitive clinical or genetic evidence. Their DR1 methylation levels were 48%, 57%, and 54%, respectively, placing three patients within the grey zone. DR2 methylation levels ranged from 60% to 73%, reflecting relatively high methylation across all cases. None of the patients exhibited a confirmed D4Z4 contraction or harbored pathogenic variants in NGS analysis. Clinical scores ranged from 1 to 13, with age of onset after the mid-20s in all cases except P24. The absence of clear hypomethylation, coupled with weak or atypical clinical findings, suggests that these patients likely fall within a broader differential diagnosis group rather than classical FSHD.

3.5.4 4qB Subset Results:

-FSHD with complex genomic characteristics: P28 was a teenage-onset female with a clinical score of 4. Molecular combing revealed that her intact D4Z4 alleles were exclusively 4qB; however, an atypical rearrangement pattern involving a 4qA allele was also observed, which prevented a definitive diagnosis based on structural criteria alone. Her DR1 methylation level was markedly reduced at 26%, significantly lower than control samples. Whole-exome sequencing identified a pathogenic variant in the *SMCHD1* gene (Table 3.2). Based on the concordant findings from NGS and methylation analysis, the patient was ultimately diagnosed with FSHD with complex rearrangement (FSHD-CR).

-**Differential Diagnosis:** P29 was excluded from an FSHD diagnosis due to a high DR1 methylation level (55%) and the detection of a likely pathogenic variant in the *CAPN3* gene, supporting an alternative diagnosis of limb-girdle muscular dystrophy.

In summary, of the remaining 19 patients, FSHD was confirmed in 13 (68.4%). The diagnosis was supported by both methylation and NGS results in 12 of these, and one patient did not have NGS data. A differential diagnosis was suggested for 3 patients

(15.8%), and in the remaining 3 cases (15.8%), additional testing is required to clarify the underlying genetic pathology.

3.5.5 Family-Based Methylation and Phenotypic Correlation

Within this cohort of unrelated patients, three individuals had family members who were also analyzed for their CpG methylation profiles to explore intra-familial variation.

-P3 and P4 (daughter-father pair): P4 is a 43-year-old affected male with a clinical score (CS) of 2. Although the exact age of onset is unclear, he did not report a severe disease course during early life. His daughter, P3, had a markedly different presentation, with a CS of 15 and disease onset at age 5, indicating early-onset FSHD. Molecular combing (MC) revealed that both carried 3 D4Z4 Rus and confirmed the presence of two 4qA alleles at P4, one of which was contracted in 3 RUs and accounted for approximately 27% of the total 4qA alleles, revealing FSHD1 with somatic mosaicism. Methylation analysis showed a DR1 level of 20% in the daughter and 45% in the father, suggesting an epigenetic difference that may partially explain the phenotypic variability.

-P6 and P7 (sibling pair): This brother and sister both had a CS between 9-10, and disease onsets are 13 and 19 years old, respectively. They were found to have 4 RUs in the D4Z4 region. Their methylation levels were similar, at 30% and 27%, indicating consistent epigenetic patterns that correlated with their comparable clinical severity.

-P18 and P19 (father-son pair): P18 is a 17-year-old affected male with a CS of 5 and disease onset at age 16. His father, P19, is 43 years old, also with a CS of 9, but a significantly later disease onset at age 34. Both were found to have 9 RUs on the D4Z4 locus. Whole-exome sequencing (WES) revealed a shared 421 kb deletion on chromosome 18 encompassing the *SMCHD1* gene. Methylation analysis showed low DR1 methylation in both individuals, 12% in the son and 18% in the father, with no substantial difference, despite their differing ages of onset.

Table 3.4: Clinical and Genetic Characteristics of the Study Cohort. Mos: Mosaic; CR: complex rearrangement.

Patients	Gender	Age of Onset	Age at Examination	Clinical Score	Family History	4qA RU Number	WES	Average CpG Methylation Level (%)	Final Diagnosis
Patient 1	Male	17	19	7	No	2	n/a	34	FSHD1
Patient 2	Female	12	32	10	No	2	n/a	43	FSHD1_mos
Patient 3	Female	5	17	15	Yes	3	n/a	20	FSHD1
Patient 4	Male	n/a	43	3	Yes	3	n/a	45	FSHD1_mos
Patient 5	Male	30	35	9	No	4	n/a	27	FSHD1
Patient 6	Female	13	28	9	Yes	4	n/a	30	FSHD1
Patient 7	Male	19	30	10	Yes	4	n/a	27	FSHD1
Patient 8	Male	15	23	8	Yes	4	n/a	28	FSHD1
Patient 9	Female	20	21	5	No	5	n/a	36	FSHD1
Patient 10	Male	18	41	1	Yes	6	n/a	52	FSHD1
Patient 11	Female	12	25	7	No	7	n/a	30	FSHD1
Patient 12	Male	20	36	6	Yes	7	n/a	34	FSHD1
Patient 13	Male	16	25	7	No	7	n/a	29	FSHD1
Patient 14	Female	45	68	6	No	8	n/a	40	FSHD1
Patient 15	Female	n/a	60	1	Yes	9	no pathogenic variant detected	48	FSHD1
Patient 16	Male	17	18	8	Yes	9	SMCHD1(NM_015295.3): Exon 15: c.2018delT; p.(Ala674Leufs*9) - Likely Pathogenic (Heterozygous)	18	FSHD1+2
Patient 17	Male	37	41	7	Yes	9	no pathogenic variant detected	27	FSHD1
Patient 18	Male	16	17	5	No	9	SMCHD1(NM_015295.3): 421 kb deletion on chr 18 - Pathogenic (Heterozygous)	12	FSHD1+2
Patient 19	Male	34	43	8	Yes	9	SMCHD1(NM_015295.3): 421 kb deletion on chr 18 - Pathogenic (Heterozygous)	18	FSHD1+2
Patient 20	Female	12	19	8	No	9	no pathogenic variant detected	32	FSHD1
Patient 21	Male	12	17	10	Yes	12	no pathogenic variant detected	20	FSHD2
Patient 22	Female	40	59	4	No	13	no pathogenic variant detected	23	FSHD_4q/10q trans.
Patient 23	Male	8	14	6	Yes	13	no pathogenic variant detected	18	FSHD2
Patient 24	Male	17	18	2	No	15	no pathogenic variant detected	48	further evaluation
Patient 25	Male	55	74	4	No	23	FHL1(NM_001159699.2): c.404G>T; p.(Gly135Val) - VUS (Heterozygous)	45	FSHD_mos
Patient 26	Male	10	22	11	No	45	no pathogenic variant detected	16	FSHD2
Patient 27	Female	13	14	3	No	48	CAPN3(NM_000070): Exon1: c.146G>A; p.(Arg49His) - Pathogenic (Homozygous)	50	LGMD2A
Patient 28	Female	15	20	4	No	4qB	SMCHD1(NM_015295): Exon5: c.610A>G; p.(Lys204Glu) - VUS (Heterozygous)	26	FSHD_C.R.
Patient 29	Male	44	44	1	No	4qB	CAPN3(NM_000070): Exon4: c.549delA; p.(Thr184Argfs*36) - Pathogenic (Heterozygous)	55	LGMD2A
Patient 30	Female	19	26	6	Yes	20	n/a	24	FSHD2
Patient 31	Male	44	45	2	No	51	no pathogenic variant detected	57	further evaluation
Patient 32	Female	24	44	13	No	48	DES(NM_001927.3): c.643G>A; p.(Val215Met) - VUS (Heterozygous)	55	MFM1
Patient 33	Male	39	43	3	No	24	no pathogenic variant detected	54	further evaluation

3.6 Correlations Between DNA Methylation and Clinical-Genetic Parameters

We investigated whether the quantity of D4Z4 RU and DNA methylation levels at the D4Z4 region are correlated to learn more about the epigenetic landscape of FSHD. Based on their 4qA RU size, patients were divided into three clinically significant subgroups:

≥ 11 RU (FSHD2 or atypical cases), 8-10 RU (borderline contraction), and 1-7 RU (typical FSHD1). To assess the overall distribution of CpG methylation levels among genetically distinct FSHD subgroups, methylation percentages were compared between patients diagnosed with FSHD1, FSHD1+2, and FSHD2 (Figure 3.3).

In FSHD1 patients, methylation values range from 20% to just above 52%. Median methylation appears around 30%. This group shows the widest variability, reflecting both typical and borderline profiles. The FSHD1+2 methylation levels are consistently low, ranging from 12% to 18%. Median is clearly below the 40% line, suggesting strong hypomethylation. This seems to be a reflection of patients with both contraction and epigenetic dysfunction. FSHD2 patients had methylation values ranging from 16% to 26%. Median is just below 25%. This data is consistent with epigenetically driven FSHD despite non-contracted D4Z4 arrays. In patients where the methylation analysis and all clinical and genotypic data suggest that these patients are unlikely to have FSHD, further investigation is warranted. The methylation levels were markedly higher, between 48% and 57%, and the median was above 50%, clearly separated from all other groups (Figure 3.3).

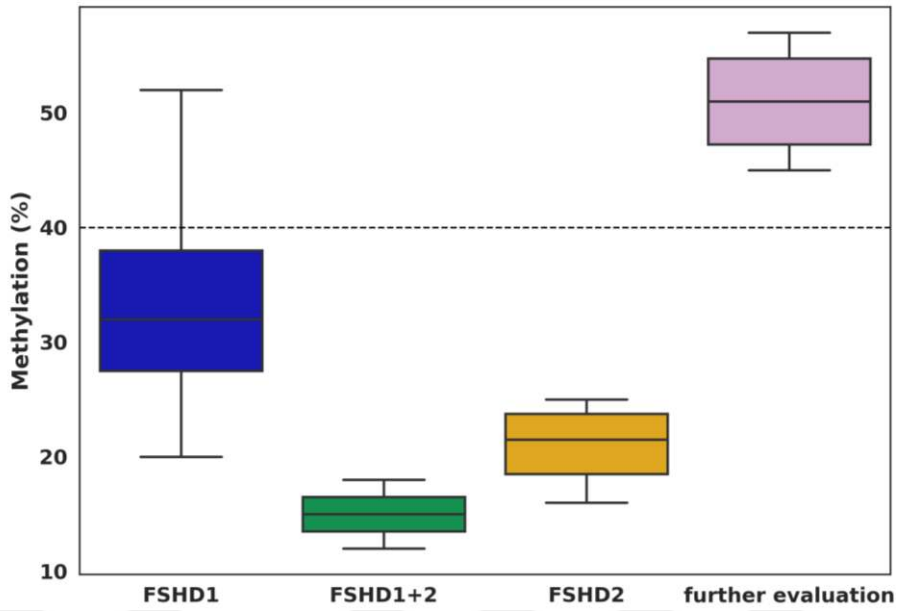

Figure 3.3: Distribution of CpG methylation levels among subgroups that have been molecularly and clinically evaluated. There is a clear inverse relationship between DR1 methylation levels and clinical severity scores (CSS) across different FSHD diagnostic groups.

Lower methylation percentages are associated with higher clinical severity, indicating that hypomethylation is generally correlated with more severe disease presentation. This trend is reflected in the downward slope of the fitted regression line, with the accompanying confidence interval suggesting a statistically meaningful association. A vertical reference line at 40% marks the commonly accepted threshold for DR1 hypomethylation. Most FSHD2 and FSHD1+2 cases fall below this line, consistent with their epigenetic profiles, while the "further testing needed" group generally exhibits higher methylation levels and lower clinical severity. FSHD1 cases, however, display greater variability, with methylation values spanning from below 20% to above 52% and a broader range of CSS, emphasizing the heterogeneity within this subgroup. Notably, a few outliers demonstrate high CSS despite borderline or elevated methylation, suggesting that methylation alone may not fully explain disease severity in all cases. Overall, this plot supports the utility of DR1 methylation profiling as a meaningful biomarker in FSHD, particularly in distinguishing subtypes and clarifying uncertain diagnoses.

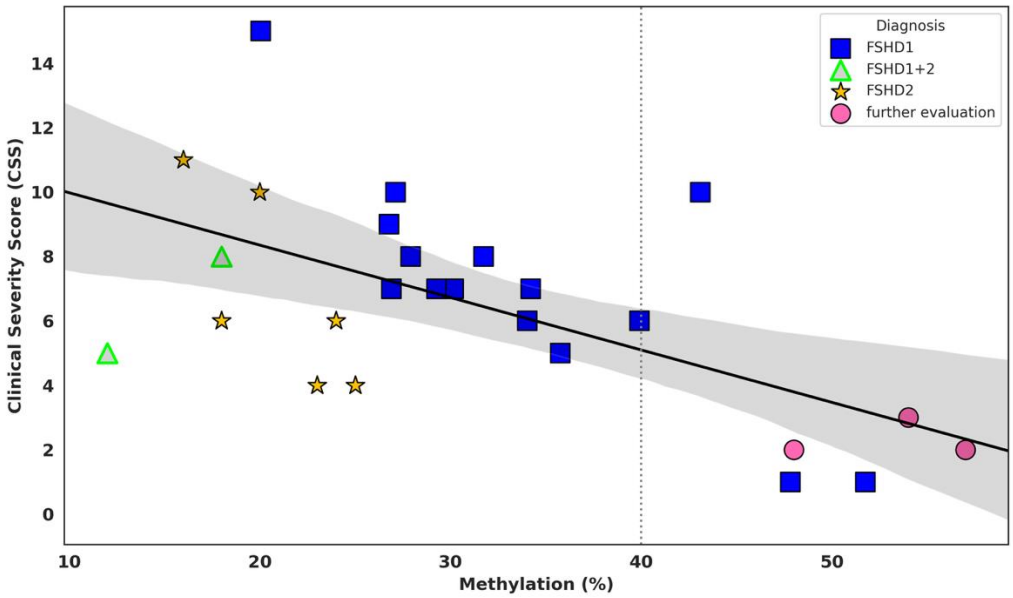

Figure 3.4: Correlation between clinical severity scores (CSS) and DR1 CpG methylation levels in patients with suspected FSHD

Each data point represents an individual, color- and shape-coded by diagnostic classification: blue squares (FSHD1), light yellow/green triangles (FSHD1+2), yellow stars (FSHD2) and pink circles (further evaluation). A statistically significant negative correlation was observed between methylation and CSS ($r = -0.621$, $p = 0.0005$), indicating that lower methylation levels tend to be related with more severe clinical presentation. The vertical dotted line at 40% represents the commonly accepted threshold for D4Z4 hypomethylation. The shaded gray area denotes the 95% confidence interval for the regression line.

3.7 FSHD Cases with Complex Structural or Genetic Features

The following cases illustrate the importance of comprehensive molecular work-up and full-locus analysis in patients who do not conform to classical FSHD1 or FSHD2 definitions. These individuals exhibited structurally or genetically atypical findings**,** including 4q/10q translocations**,** allelic mosaicism, or co-existing variants in non-D4Z4-related myopathy genes**.** Their diagnostic classification required integration of methylation profiling, D4Z4 repeat sizing, haplotype determination, and next-generation sequencing to reach a conclusive interpretation. Such cases underscore the necessity of

evaluating beyond canonical contraction or methylation thresholds when complex rearrangements or overlapping phenotypes are suspected.

Complex Rearrangement of 4q/10q in Patient 22:

P22 is a 59-year-old female who began showing symptoms around the age of 40. With a low clinical severity score of 4/15 and a relatively mild disease course, she represents a borderline phenotype with features partially compatible with FSHD. This patient is distinguishable by her epigenetic profile. Despite having 13 and 15 D4Z4 RUs on 4qA and 4qB alleles, respectively, values that fall outside the classical FSHD1 contraction range, her DR1 methylation level was significantly decreased at 23%**,** consistent with a hypomethylated epigenotype**.** Her DR2 methylation was 52%, within the expected range for non-contracted alleles, reinforcing the regional methylation specificity. This distinct DR1 hypomethylation in the absence of D4Z4 contraction raises the possibility of *DUX4* expression deregulation despite apparently non-pathogenic allele lengths.

We did not detect any significant variants by whole-exome sequencing (WES), molecular combing analysis revealed a potentially relevant finding: a contracted 10qA allele with 7 RUs. While contractions on 10q are traditionally considered non-pathogenic due to the absence of a *DUX4*-permissive polyadenylation signal, the presence of significant hypomethylation raises the possibility of complex or mislocalized epigenetic regulation affecting the 4q locus or involving a rare rearrangement with the 10q region.

Somatic Mosaisism in Patient 25:

P25 is a 74-year-old male presenting with late-onset, mild muscular symptoms first noted at age 55. His clinical severity score was 4/15. MC analysis revealed two non-contracted 4qA alleles (23 and 61 RUs), and an additional 1 RU allele representing approximately 9% of the total 4qA signal-suggestive of somatic mosaicism. Crucially, DR1 methylation was measured at 45%, and DR2 at 57%, placing this patient within the epigenetic grey zone. These methylation levels fall above the hypomethylated threshold yet below the unmethylated boundary, aligning with the observed mild clinical expression and suggesting partial epigenetic deregulation. Further complexity was introduced by whole-exome sequencing, which revealed a heterozygous VUS in the *FHL1:*c.404G>T;

p.(Gly135Val), previously linked to X-linked myopathies including Emery-Dreifuss Muscular Dystrophy 6. While this variant may contribute to the phenotype, its role remains uncertain without segregation or functional studies. The combination of intermediate methylation, mosaic D4Z4 contraction, and the absence of classic FSHD features positions P25 as a case of uncertain FSHD spectrum, requiring continued monitoring and possibly re-evaluation should additional data, such as muscle biopsy or familial genotyping-become available. The methylation findings provide a compelling molecular correlate for his late onset and attenuated clinical course.,

Structurally Complex Alleles in Patient 28:

The patient is a 20-year-old female who presented with clinical features consistent with FSHD symptoms, and onset was reported at approximately age 15, and her clinical severity score was assessed as 4/15. Methylation analysis revealed a distinctly hypomethylated profile, with DR1 and DR2 CpG methylation levels of 26% and 40%, respectively. These values fall below or near the threshold typically associated with FSHD2 epigenetic profiles, supporting a diagnosis consistent with *DUX4* de-repression. Although molecular combing demonstrated a 4qB haplotype, additional complex rearrangements were observed, including a segment with 6-8 RUs adjacent to a larger 40 RU array on 4qA, suggestive of structural mosaicism. This configuration, in combination with her epigenetic findings, prompted further genetic evaluation. Whole-exome sequencing identified a heterozygous missense variant in the *SMCHD1* gene (c.610A>G, p. Lys204Glu). This variant is currently classified as a variant of uncertain significance (VUS) according to ACMG guidelines, supported by PM2, PP2, and PP3 criteria. *In silico* predictions suggest pathogenic potential, but definitive evidence remains limited. Nonetheless, this specific variant has been previously reported in the literature in association with FSHD2 cases (Mul et al., 2018; Lemmers et al., 2019). Taken together, the markedly reduced methylation at DR1, intermediate DR2 levels, suggestive structural variation at the D4Z4 locus, and presence of a potentially pathogenic *SMCHD1* variant support a diagnosis of FSHD with atypical epigenetic and structural features, potentially falling within the FSHD2 or FSHD-complex rearrangement (FSHD_C.R.) category.

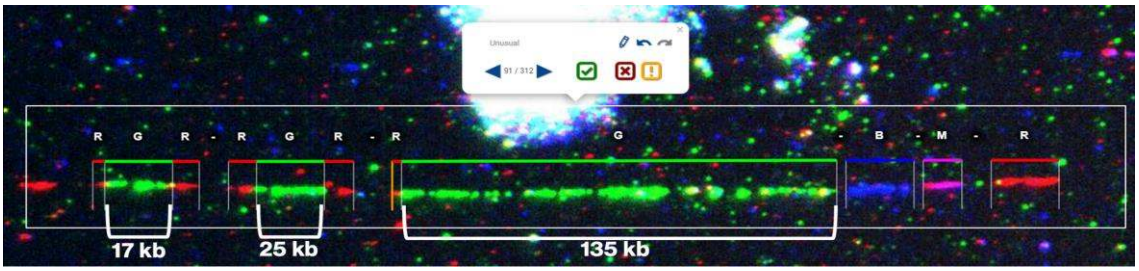

Figure 3.5: Molecular combing result image of Patient 28, reveals a complex 4q configuration with a short D4Z4 repeat (6-8 RUs) adjacent to a longer segment (~40 RUs), indicative of structural rearrangement.

Chapter 4:
DISCUSSION

FSHD is driven by both genetic and epigenetic mechanisms, making its diagnosis inherently complex and multifaceted. While FSHD1 results from contraction of the D4Z4 repeat array and FSHD2 arises from defects in epigenetic regulators such as *SMCHD1*, both subtypes share a common molecular feature: hypomethylation at the D4Z4 locus, which leads to derepression of the pathogenic *DUX4* gene. This epigenetic signature has been consistently observed in both FSHD1 and FSHD2 patients, who exhibit significantly lower D4Z4 methylation levels compared to healthy individuals (Gaillard et al., 2014; Sacconi et al., 2019). The hypomethylation is often more pronounced in FSHD2, particularly when pathogenic variants in modifier genes are present (Lemmers et al., 2012b; Lemmers et al., 2015). Over the past decade, D4Z4 methylation has emerged as a valuable diagnostic biomarker, helping to distinguish between classical FSHD1 and epigenetically driven forms like FSHD2. As such, it is increasingly incorporated into diagnostic workflows alongside genetic testing (Lemmers et al., 2012; Hartweck et al., 2013; Jones et al., 2015; Sacconi et al., 2019). In this study, we build on this foundation by investigating the correlation between repeat unit (RU) size, CpG methylation levels, and clinical severity to examine the molecular and epigenetic characteristics of our FSHD patients.

4.1 Overall Statistics and Genotype-Phenotype Correlation of Molecular Combing Results

The FSHD patient group included 112 individuals evaluated for suspected differential diagnoses of FSHD, consisting of 92 unrelated patients and 20 family members (Section 3.1, pg. 42; Figure 3.1). The general statistics show that 54% of cases were familial, while 46% appeared to be isolated. Although isolated presentations are typically reported to account for 10-30% of FSHD cases, this range may not reflect the true rate of de novo mutations, as it does not account for instances in which one or both parents are asymptomatic carriers or somatic mosaics for the underlying genetic defect. Therefore, detailed family investigations, including molecular analysis of parental

genotypes, are essential to accurately determine inheritance patterns and to distinguish truly de novo cases from those with unrecognized familial transmission. The cohort had nearly equal sex distribution, with 51% male and 49% female participants, aligning with the understanding that both FSHD1 and FSHD2 are inherited in an autosomal dominant manner and therefore affect both sexes equally. However, previous studies showed that penetrance is generally higher in males (estimated at ~95%) compared to females (70-80%) (Lemmers et al., 2012). Females are more likely to be asymptomatic carriers or exhibit milder phenotypes, particularly in cases involving borderline D4Z4 repeat units. Interestingly, our cohort did not reflect this typical pattern. The mean age of onset in females was approximately 19 years, compared to 25 years in males, and females showed slightly higher clinical severity scores. This deviation may be influenced by referral bias, with symptomatic females being more likely to seek medical attention. Alternatively, epigenetic variability may play a role. A larger and more diverse cohort would be needed to explore these differences further. Age of onset data revealed that 68% of patients experienced their first symptoms at or before the age of 20, with approximately 10% of these cases presenting in early childhood (before age 10). The remaining 32% had symptoms emerging after the age of 20, including late-onset cases (after age 40). These findings are consistent with previous publications (Tawil et al., 2014). These findings are consistent with previous studies, which report that over half of individuals with FSHD exhibit clinical symptoms by the age of 20, and early-onset is estimated to be 7% to 15%, with a more severe early onset type (Goselink et al., 2019).

Among the 92 genetically evaluated unrelated patients, FSHD1 was molecularly confirmed in 78 individuals (84.8%), all of whom carried 1 to 10 D4Z4 repeat units (RUs) on a permissive 4qA allele (Section 3.1, pg. 43; Figure 3.2). Specifically, 62 patients (67.4%) had 1-7 RUs, 16 patients (17.4%) had 8-10 RUs, and 12 patients (13.0%) had ≥11 RUs. These findings reflect the typical RU distribution seen in FSHD1 cohorts, where the majority of affected individuals have 1-7 RUs. Genotype-phenotype correlations show that the patients with 1-3 RUs have earlier onset and more severe phenotypes. Patients with 4-7 RUs often show variable clinical severity. Those with borderline contractions of 8-10 RUs are generally more mildly affected, unless there is additional involvement of epigenetic modifiers. All data correlates with other research outcomes (Ricci et al., 2013; Mul et al., 2018). Among the 20 family members analyzed, nine were

affected and 11 were unaffected. All affected relatives had the same D4Z4 repeat unit (RU) size as the index patient in their family. Thus, their diagnoses were confirmed. The unaffected relatives were referred to assess the possibility of asymptomatic carrier status; however, none were found to be carriers, and all had more than 11RUs, and FSHD was excluded.

4.2 Clinical Characteristics and Characterization of CpG Methylation Patterns in our selected FSHD Cohort

Out of a larger cohort of 112 individuals, a subset of 33 patients (30 unrelated, 3 relatives) was selected for detailed methylation profiling, alongside six healthy controls for comparison. Patients were divided into four groups based on D4Z4 repeat size and permissive allele status: typical FSHD1 cases (1-7 RUs), borderline repeat size cases (8-10 RUs), potential FSHD2 or undetermined cases (≥11 RUs), and individuals with non-permissive 4qB alleles. This classification aimed to explore correlations between CpG methylation patterns, genotype, and clinical presentation.

4.2.1 Clinical Characteristics

Clinical evaluation of 30 individuals revealed a wide range of clinical severity scores, with slightly higher mean scores in females compared to males. Based on our dataset, all patients, irrespective of their age of onset, D4Z4 unit size, and methylation patterns, had different degrees of scapular girdle involvement. The next most involved muscle groups were facial muscles (75%), pelvic girdle and abdominal muscles (65%) among the patients. These findings are consistent with the classical phenotype of FSHD (Statland, J. M., & Tawil, R., 2014). The correlation of muscle involvement according to sex revealed that, for each muscle group, females had the highest severity scores when compared to males. This is somewhat an unexpected result given the established literature, which typically reports higher penetrance and severity in males with FSHD. However, a recent analysis of the U.S. National Registry for FSHD highlighted significant gender disparities in pediatric-onset cases. Among the 166 genetically confirmed FSHD1 and FSHD2 patients studied, girls were diagnosed at a younger age than boys and exhibited more severe disease progression. In this study, pediatric-onset FSHD was defined as a diagnosis before the age of 18, with early-onset categorized as a diagnosis

before age 10 (Katz et al., 2024). When we analyzed our cohort using these criteria, the mean clinical severity score (CSS) among female patients (n=7) was 8, compared to 6.5 in male patients (n=10), which appears to be consistent with the findings of the U.S. registry study.

4.2.2 Characterization of CpG Methylation Patterns

In this study, bisulfite sequencing targeted both the DR1 and DR2 regions near the D4Z4 macrosatellite array. DR1, positioned closer to the *DUX4* gene, has been established as the more diagnostically informative site for detecting FSHD-associated hypomethylation (Hartweck et al., 2013; Sacconi et al., 2019). It is widely recognized as the gold standard for methylation-based diagnostics in FSHD and is included in several molecular testing guidelines (Gaillard et al., 2014; Sacconi et al., 2019).

Our results support this consensus. DR1 methylation levels in our FSHD-confirmed patients showed broader variability and greater correlation with clinical status than DR2. DR1 values ranged from 20% to 52%, with the majority falling below the 40% hypomethylation threshold. By contrast, DR2 methylation levels were more stable and generally higher, ranging from 47% to 78% across the cohort, including in cases with mosaicism or borderline repeat sizes. Even patients with clearly hypomethylated DR1 values (e.g., P3: 20%, P16: 18%) had DR2 levels above 40% (e.g., 42% and 39%, respectively), showing the limited sensitivity of DR2. This is consistent with prior studies showing that DR2 is less responsive to disease-associated epigenetic changes, likely due to its upstream location and lower CpG density (van Overveld et al., 2003; Hartweck et al., 2013).

Collectively, our findings reaffirm that DR1 is a more reliable and discriminatory methylation marker for both FSHD1 and FSHD2. As such, it serves as the primary focus for the detailed analyses presented in the following sections, including family-based methylation profiles and their correlation with clinical phenotypes:

1-7 RU FSHD1 Subgroup

There were 11 unrelated and 2 relatives, all diagnosed as FSHD1 (P1-P13) with 1-7 RUs, who presented with a variable clinical and epigenetic pattern. Most of these patients had either moderate or high CSS with early onset of symptoms frequently encountered during adolescence or early adulthood (Section 3.2, pg. 44; Table 3.1). The average methylation percentage in this group ranged between 20 and 52%, with most falling below the 40% threshold considered indicative of pathogenic hypomethylation. Apart from two borderline borderline grey-zone (P2=43%, P4=45%) and one outlier (P10=52%), these data conform to the previously described hypomethylation values of pathogenic alleles (Sacconi S. et al., 2019).

P2 was a 32-year-old female with disease onset at age 12, and she has CSS of 10/15. She presented with bilateral scapular winging, marked proximal upper limb weakness, and spinal curvature. She was dependent on a wheelchair. Molecular combing revealed approximately 16% somatic mosaicism for the contracted 4qA allele with 2RUs. Her DR1 methylation level was 43%, falling within the borderline range for FSHD1. This epigenetic profile was consistent with low-level mosaic FSHD1, typically associated with a milder clinical course; however, her disease severity was unexpectedly high. In light of this, the patient underwent further clinical re-evaluation. MRI and EEG findings raised suspicion of neurodevelopmental involvement, as she exhibited epileptic seizures and cognitive impairment. Unlike other high-CSS FSHD cases in our cohort, she showed no abdominal muscle involvement and only mild facial weakness, which suggested the presence of an additional neurological contributor to her symptoms. WES identified a heterozygous variant of uncertain significance (VUS) in the *NEUROD2* gene NM_006160.3:c.34C>G(p.Arg12Gly), previously implicated in developmental and epileptic encephalopathy-72 (OMIM: 618374), a disorder characterized by refractory seizures, psychomotor delay, and impaired ambulation (Runge et al., 2021). Although *NEUROD2* is primarily associated with central nervous system conditions, its potential contribution to neuromuscular involvement remains unclear. The coexistence of FSHD1 mosaicism and the *NEUROD2* VUS may have jointly contributed to her severe clinical presentation. Whether the variant functioned as a genetic modifier or indicated an

overlapping neurodevelopmental disorder requires further clarification through segregation analysis, functional studies, and extended neurological follow-up.

P4 was a 43-year-old male with a very mild disease presentation. Age of onset was uncertain, and his clinical examination revealed subtle facial involvement, mild right-sided scapular winging, and humeral weakness, without objective strength loss in any muscle group. Clinical severity score (CSS) was 3. MC analysis indicated FSHD1 with somatic mosaicism, with a 3 RU D4Z4 contraction on one 4qA allele, comprising approximately 27% of the total 4qA alleles. His DR1 methylation level was 45%, placing him in the borderline range. He was included in this study due to the presentation of a much more severe phenotype with his daughter P3. Her disease onset was at age 5. She exhibited a severe FSHD phenotype with a CSS of 15/15, the same 3RUs contraction, and a DR1 methylation level of 20%. She had been wheelchair-dependent for the past four years and showed advanced clinical features including severe scoliosis, pronounced lumbar lordosis, complete immobility of lower facial muscles, bilateral scapular winging, and a positive Beevor's sign. Additionally, she exhibited mild high-frequency hearing loss. This striking difference can be explained by the father's somatic mosaicism, a phenomenon frequently observed in de novo cases where the parent is only mildly affected or asymptomatic (van der Maarel et al., 2000). In this context, his relatively high methylation level aligns well with both his mild clinical presentation and the mosaic nature of his genetic findings. Ultimately, integrating methylation status, RU size, and sex-specific effects when interpreting FSHD1 phenotypes offers more accurate prognostic insight than genetic data alone.

P10 was a 41-year-old male with symptom onset at age 18. He reported mild upper limb weakness for over two decades. Despite this long disease duration, he retained full arm elevation and normal lower limb strength. EMG was unremarkable. His facial features were not typical of FSHD; notably, he was able to whistle and inflate his cheeks. Genetic analysis revealed 6 RUs on a permissive 4qA allele, inherited maternally. However, his D4Z4 methylation level was 52%, exceeding the established hypomethylation threshold typically associated with pathogenic FSHD alleles. His clinical findings are in line with his methylation scores; however, in contrast with his RU number. Despite the repeat contraction, the clinical profile and high methylation do not

fit the traditional diagnosis of FSHD. Even within the same family, Gaillard et al. (2014) showed that D4Z4 methylation levels can distinguish between asymptomatic carriers and FSHD1 patients with symptoms.

8-10 RU Borderline Subgroup

This subgroup consists of 6 unrelated and 1 relative (P14-P20) with 8-10 repeat units (RUs). DR1 methylation levels exhibited substantial variation, averaging 29.5% and spanning from 12.0% to 48.0%. This configuration falls within the borderline pathogenic range and, according to diagnostic algorithms, should be considered as potentially involving co-existing pathogenic variants in epigenetic modifier genes (Lemmers et al., 2024). Accordingly, WES results were incorporated for all patients in this group except P14, enhancing both diagnostic accuracy and genotype-phenotype correlation. Based on the integrated molecular and clinical findings, four patients were diagnosed with FSHD1, while two index cases and one relative were classified as FSHD1+2.

P14 was a 68-year-old female with symptom onset at age 45, classifying her as a late-onset case. She had a Clinical Severity Score (CSS) of 6 and exhibited typical FSHD-related muscle involvement. Her DR1 methylation level was 40%, placing her at the upper threshold of the hypomethylation range, and she carried 8 D4Z4 repeat units. Although NGS data were unavailable for this patient, the combination of her CpG methylation profile and D4Z4 repeat size was sufficient to support a diagnosis of FSHD1.

P15 was a 60-year-old female patient who carried 9 RUs on a permissive 4qA allele, a configuration consistent with FSHD1. With a CSS of 1, she was considered asymptomatic. WES revealed no FSHD-related variants. Her CpG methylation level was 48%, the highest within this group, placing her in the borderline methylation category. This epigenetic profile aligned with her clinical and genetic findings. Her mildly affected daughter (not included in this study due to limited material) also carried 9 RUs, suggesting a familial borderline case with reduced penetrance. Tonini et al. (2004) reported that families with 7-10 RUs often include asymptomatic gene carriers. Similarly, Lemmers et al. (2015) and Jones et al. (2015) observed that hypomethylation tends to be less

pronounced in asymptomatic individuals with borderline RU sizes, supporting the role of DNA methylation as a disease modifier.

P16 was a male patient who became symptomatic at the age of 17, initially presenting with weakness in the right arm and leg. He exhibited significant muscle wasting in both upper and lower limbs, with a positive Gowers' sign. Clinical examination revealed marked weakness in the pelvic and shoulder girdle muscles, and he was unable to lift either arm fully. His CSS was 8. MC identified a borderline D4Z4 repeat contraction of 9RUs on the 4qA allele. In addition, WES revealed a novel heterozygous frameshift variant in the *SMCHD1* gene, c.2018delT (p.Ala674Leufs*9), classified as likely pathogenic, according to ACMG criteria (PVS1, PM2). These findings supported a diagnosis of FSHD1+2. The combination of borderline D4Z4 repeat size and a significantly reduced DR1 methylation level (18%) is consistent with previous studies indicating that pathogenic *SMCHD1* variants act as disease modifiers, exacerbating clinical severity in genetically borderline cases (Sacconi et al., 2012; Lemmers et al., 2015). As reported by Lemmers et al. (2015), FSHD1 patients harboring additional pathogenic variants in the *SMCHD1* gene tend to exhibit earlier disease onset, more pronounced muscle weakness, and significantly lower D4Z4 methylation levels, supporting the role of *SMCHD1* as a disease modifier and reinforcing the concept of FSHD1+2.

P17 (male, onset at age 37, CSS = 7) and **P20** (female, onset at age 12, CSS = 8) both carried 9 D4Z4 repeat units. Their DR1 methylation levels were 27% and 32%, respectively, both falling within the hypomethylated range. Whole-exome sequencing did not reveal any significant variants. The alignment of their clinical features, genetic findings, and methylation profiles supported a diagnosis of FSHD1 in both cases.

P18 and P19, a father-son pair. The son (P18) developed symptoms at age 16, presenting with mild facial weakness, scapular winging, and limited shoulder abduction, with a CSS of 5. His D4Z4 methylation was markedly low (12%), consistent with FSHD2. His father (P19), who began showing symptoms later, at age 34, had a CSS of 8 by age 43, with more advanced signs including dysphagia, upper arm atrophy, and calf pseudohypertrophy. His methylation level was slightly higher (18%). They both received

a molecular diagnosis of FSHD1 based on the presence of a contracted D4Z4 repeat (9 RUs) on a permissive 4qA allele. WES analysis revealed a heterozygous ~421 kb deletion on chromosome 18 encompassing the *SMCHD1* gene. As previously reported, *SMCHD1* haploinsufficiency leads to *DUX4* de-repression and D4Z4 hypomethylation, thereby contributing to the FSHD2 disease mechanism even in the presence of a borderline repeat contraction (Lemmers et al., 2012; Sacconi et al., 2012). These findings supported a familial FSHD1+2 diagnosis. The observed difference in severity is likely to be related to the difference in CpG levels and age-related, which aligns with the progressive nature of FSHD, rather than suggesting variable penetrance. These findings are in concordance with prior studies demonstrating that *SMCHD1* mutations, particularly when paired with borderline D4Z4 repeat contractions, drive FSHD2 pathology through pronounced hypomethylation (Lemmers et al., 2015; Sacconi et al., 2012; Mul et al., 2018).

≥11 RU Subgroup

This subgroup comprised 11 patients whose D4Z4 repeat sizes fell within the normal range. In these cases, methylation profiling played a key role in guiding diagnostic decisions. Based on integrated analysis and additional testing, four patients (P21, P23, P26, P30) were diagnosed with FSHD2. Two patients with complex genotypes (P22, P25) were classified as having FSHD despite atypical findings. Two patients (P27, P32) received a differential diagnosis, while four others (P24, P31, P33) were excluded from the FSHD cohort due to inconclusive or non-supportive molecular and clinical data.

P21, P23, P26, and **P30** showed DR1 methylation levels well below the 30% threshold (20%, 18%, 16%, and 24%, respectively). All patients in this subgroup experienced symptom onset before the age of 20. Among them, P23 and P30 shared similar disease severity (CSS = 6) and exhibited classical FSHD features, including bilateral scapular winging, axial muscle involvement, and predominant upper limb weakness. Both patients also demonstrated hallmark facial features such as incomplete eye closure, impaired cheek puffing, and asymmetrical upper lip movement. P21 and P26 had more severe clinical presentations (CSS=10 and 11, respectively), again with a distribution of muscle weakness typical of FSHD. In all four cases, these characteristic phenotypes-together with markedly low DR1 methylation levels, supported a diagnosis of FSHD2. WES data were not available for P30, and bioinformatic analyses of the

remaining patients did not reveal any pathogenic variants in known FSHD-related genes. Recent large-cohort studies have shown that individuals with 8-20 D4Z4 RUs may present with significant hypomethylation and occasionally harbor variants in epigenetic regulator genes (Lemmers et al., 2018; Jia et al., 2021). Based on these criteria, the diagnoses of FSHD2 in P21, P23, and P30, each with ≤20 RUs, were considered valid. Although P26 had 45 RUs, which exceeded the conventional FSHD diagnostic threshold, his clinical presentation and methylation profile were highly suggestive of FSHD2. He exhibited typical muscle involvement with an early-onset course and a DR1 methylation level of just 16%. Such findings justify the diagnosis, especially in light of updated guidelines recommending further genomic analyses, such as optical genome mapping (OGM) or long-read sequencing, to detect complex rearrangements or cis-acting variants that may be missed by standard molecular combing or WES (Lemmers et al., 2024; Lemmers et al., 2022).

P22 is a female patient, 59 years old, who presented with symptoms, such as mild scapular winging and slowly progressive weakness in the arms (clinical severity score 4). Molecular combing demonstrated a contracted 10qA allele with 7 D4Z4 repeat units (RUs), 13 and 15 RUs on her 4qA and 4qB alleles, her CpG methylation level was reduced to 23%, in the usual FSHD2 diagnosis threshold of <25-30%, and reflecting epigenetic dysregulation (Sacconi et al., 2019; van den Boogaard et al., 2016). Even though this contraction is classically not pathogenic, recent literature has demonstrated that the disease is caused by repeat contraction and *DUX4* expression from chromosome 10 due to a de novo D4Z4 repeat exchange between chromosomes 4 and 10. (Lemmers et al., 2022; Delourme et al., 2023). It has been proposed in recent studies that structural alterations in the locus 10q, like translocations and cis-regulatory changes, may influence penetrance and phenotype in genetically indefinite cases of FSHD (Lemmers et al., 2022; Nguyen et al., 2019). The presence of the short allele 10qA with 7 RUs in this patient may cause an FSHD-like phenotype due to mislocalization of the regulatory elements or altered chromatin dynamics (Lemmers et al., 2022).

P25, the 74-year-old man, started having symptoms around the age of 55, initially with weakness of the right side greater than the left in the shoulders. He has a clinical score of 4/15, representing very mild disease severity. EMG findings were normal.

Molecular combing analysis demonstrated two uncontracted D4Z4 4qA alleles (23 and 61 RUs) and one third of a 4qA allele with 1 RU, representing about 9% of counted alleles and suggesting somatic mosaicism. His CpG methylation percentage was 45%, in a borderline hypomethylated range. Whole-exome sequencing revealed a heterozygous VUS in the *FHL1:*c.404G>T; p.(Gly135Val). Mutations in *FHL1* have been described to cause a scapuloperoneal phenotype mimicking FSHD, inherited in an X-dominant fashion (Quinzii et al., 2008; Sacconi et al., 2013). Lack of muscle biopsy and segregation in the family prevents definitive diagnosis, but the association of low-level D4Z4 mosaicism and an *FHL1* variant is consistent with blended or atypical myopathy. Additional research, such as muscle biopsies and segregation studies in families, is required to determine the role of these genetic elements in the clinical findings.

P27 was a 14-year-old female with onset at 13 years and 3/15 CSS. Her obvious clinical findings were mild facial weakness, scapular winging, and calf muscle atrophy. WES identified the pathogenic variant in the *CAPN3* gene as c.146G>A (p.Arg49His), and so the diagnosis of calpainopathy, or LGMD2A, as recorded in the ClinGen Evidence Repository (ClinGen, 2024) is confirmed.

P32 was a female patient who presented with aggressively progressive muscle weakness beginning at the age of 24 and wheelchair dependence by age 30. Persistent neck muscle weakness over two decades and highly elevated CK levels (3356 U/L). Molecular combing revealed she has no contracted RUs. Whole-exome sequencing revealed a heterozygous missense variant in *DES*:c.643G>A, p.(Val215Met), which is classified as a variant of uncertain significance (VUS). Mutations in *DES* are known to cause myofibrillar myopathy type I (OMIM: 601419), typically presenting with progressive limb-girdle weakness as seen in P32. In silico analysis, using tools (SIFT, PolyPhen-2, and MutationTaster) supported the pathogenic potential of this variant. All findings, such as high CpG methylation percentage (55%), and except for upper limb weakness, no additional clinical manifestations suggestive of FSHD, support the rule out the FSHD and suggest with possibility of a genetic myopathy, such as DES-related myofibrillar myopathy. This case illustrates how integrating clinical features, variant interpretation, and methylation profiling can help us in refining diagnosis. Additional

studies, such as muscle biopsy, functional assays, and familial segregation analysis, are warranted to confirm pathogenicity.

P24, P31, and P33 presented with clinical features partially suggestive of facioscapulohumeral muscular dystrophy (FSHD), yet lacked D4Z4 repeat contractions and exhibited CpG methylation levels in the normal range (Lemmers et al., 2015). P24 carried 15 repeat units on the 4qA allele and had a methylation level of 48%, a clinical severity score (CSS) of 2, supported by myopathic findings on EMG. He is clinically mild, with no involvement of the face and abdominal muscles. Only the right scapular winging is locally present and shows mild involvement. There is no pathogenic variant in the WES data. P31 (CSS = 2, 51 RUs methylation = 57%) and P33 (CSS = 3, 24 RUs, methylation = 54%) have presented with varying degrees of upper limb weakness. All three patients exhibited some degree of scapular winging or humping, and two (P31, P33) had facial involvement in the form of ptosis (P33) and asymmetrical upper lip movement (P31), which was minimal. Two patients (31 and 33) had normal or mildly elevated CK.

These cases highlight the diagnostic complexity of FSHD-like phenotypes in the absence of classical molecular findings. Despite exhibiting mild or partial clinical features of FSHD, the absence of D4Z4 contractions, normal CpG methylation levels, and unremarkable WES findings suggests the involvement of another genetic or epigenetic mechanism.

4qB Subgroup

This subgroup consists of two patients (P28 and P29). P28, which has complex genomic characteristics, has a 26% CpG methylation value and the *SMCHD1* variant. She was diagnosed within the FSHD2 spectrum, more specifically as an FSHD-complex rearrangement (FSHD_CR) based on integrated analysis of NGS data, CpG methylation value, and molecular combing analysis. P29 has a 55% CpG methylation value, and he received a differential diagnosis of LGMD2A, supported by the identification of a pathogenic variant in the *CAPN3* gene by NGS analysis.

P28 is a 20-year-old female with symptom onset at age 15 and a clinical severity score of 4/15. She had bilateral scapular winging, progressive weakness of the arms, and

mild facial weakness. EMG showed chronic, moderate myopathy. Her methylation status in the DR1 region was 26%, in the hypomethylated range characteristic of FSHD2 (Sacconi et al., 2019).

Even though her haplotype was found to be 4qB, usually regarded as non-permissive in FSHD, this pattern implies an intricate structural rearrangement that includes involvement of a qA-type segment, and possibly it occurred by inter-allelic recombination.

Her WES analysis revealed a heterozygous VUS variant in *SMCHD1*:c.610A>G, p.(Lys204Glu). The existence of an *SMCHD1* variant and the 4qA complex rearrangement characteristics in P28, in the absence of a typical pathogenic D4Z4 constriction, suggests the interaction between complex genomic architecture and epigenetic control. *SMCHD1* variations and complex rearrangements like these can play a crucial role in the development of FSHD manifestations. (Nguyen et al., 2019). Taken together, her clinical presentation, lower level of methylation, and established *SMCHD1* mutation fulfill diagnostic criteria for FSHD2. This case shows the value of the application of molecular diagnostics, methylation analysis, and *SMCHD1* sequencing in confirming and diagnosing FSHD.

P29, a 44-year-old male, has clinical findings of late-onset muscle weakness and atrophy of paraspinal muscles. His methylation status in the DR1 region was 55%. He has a heterozygous CAPN3 c.549delA (p.Thr184Argfs*36) frameshift mutation that has been previously documented as pathogenic in numerous LGMD2A patients (ClinVar, 2024). Trans-allelic deletions were excluded by MLPA analysis and thereby confirmed the variant configurations. These show the significance of extensive genetic testing, such as CAPN3 screening and MLPA, in patients with FSHD-like features, especially when the molecular diagnostic criteria for FSHD are not fulfilled.

Chapter 5:

CONCLUSION

Methylation analysis has become a valuable complement to genetic testing in the diagnosis of FSHD, enabling the distinction between FSHD1 and FSHD2. In this study, we analyzed CpG methylation levels in the DR1 and DR2 regions. DR1 demonstrated the most pronounced hypomethylation in affected individuals (Hartweck et al., 2013), whereas DR2 showed methylation levels as close to or almost the same as those of healthy controls. Despite DR1 and DR2 being only 131 bp apart, these results imply that even tiny genomic distances can produce remarkably different methylation patterns.

Advancing our understanding of the molecular mechanisms that drive epigenetic dysregulation in FSHD may inform more targeted therapeutic strategies. Critical variables such as percent CpG methylation, presence of pathogenic variants in *SMCHD1* and other modifier genes, degree of somatic mosaicism, structural haplotype complexity, and patient-specific variables (e.g., sex hormones, medical history, environmental exposures) should be integrated into clinical assessments.

Our findings suggest that future diagnostic approaches to FSHD diagnosis can shift from strict subtype classifications toward more integrative, probabilistic models that will be crucial in deciphering the complexity of FSHD and accurately differentiating it from phenotypically similar myopathies.

BIBLIOGRAPHY

1. Aguirre, J., et al. (2023). Role of physical therapy in muscular dystrophies. *Neuromuscular Disorders,* 33(4), 245-260.

2. Banerji, C. R., et al. (2017). Mechanisms of *DUX4*-induced toxicity in FSHD. *Journal of Neuromuscular Diseases,* 4(1), 43-54.

3. Bouwman, L. F., et al. (2020). *DUX4*-driven gene networks in Facioscapulohumeral Muscular Dystrophy and therapeutic implications. *Neuromuscular Disorders*, 30(8), 678-689.

4. Caputo, V., et al. (2022). Advances in DNA methylation analysis for Facioscapulohumeral Muscular Dystrophy. *Epigenomics,* 14(7), 521-537.

5. Castellano, V., et al. (2008). Historical perspectives on Facioscapulohumeral Muscular Dystrophy. *Neuromuscular Disorders,* 18(9), 675-682.

6. Chin, A. X. Y., et al. (2024). Clinical heterogeneity in Facioscapulohumeral Muscular Dystrophy: A systematic review of phenotypic variability. *Journal of Neuromuscular Diseases,* 11(2), 245-258.

7. Dorobek, M., et al. (2013). Electromyographic patterns in Facioscapulohumeral Muscular Dystrophy: A diagnostic evaluation. *Muscle & Nerve,* 48(5), 785-792.

8. Erdmann, et al. (2023). Methylation of the 4q35 D4Z4 repeat defines disease status in facioscapulohumeral muscular dystrophy, *Brain*, Volume 146, Issue 4, Pages 1388-1402

9. Feng, Q., et al. (2015). Epigenetic regulation of *DUX4* expression in FSHD. *Nature Genetics,* 47(9), 1027-1034.

10. Gabriëls, J., et al. (1999). *DUX4*, a candidate gene in Facioscapulohumeral Muscular Dystrophy, encodes a protein with homeobox domains. *Nature Genetics,* 21(2), 178-180.

11. Ghasemi, M., et al. (2022). Expanding the spectrum of Facioscapulohumeral Muscular Dystrophy: Insights into extra-muscular manifestations and systemic involvement. *Neuromuscular Disorders*, 32(7), 589-599.

12. Greco, S., et al. (2020). Epigenetic alterations in Facioscapulohumeral Muscular Dystrophy. *Frontiers in Genetics*, 11, 215.

13. Hamel J, Tawil R. (2018). Facioscapulohumeral Muscular Dystrophy: Update on Pathogenesis and Future Treatments. Neurotherapeutics. 5(4):863-871.

14. Himeda, C. L., Jones, T. I., & Jones, P. L. (2015). Facioscapulohumeral Muscular Dystrophy as a model for epigenetic regulation and disease. *Current Opinion in Genetics & Development*, 36, 50-56.

15. Hobson-Webb, L. D., et al. (2006). FSHD inheritance patterns and clinical severity. *Muscle & Nerve*, 34(3), 245-252.

16. Jones, T. I., et al. (2014). Methylation variability in FSHD1 and FSHD2. *Human Molecular Genetics*, 23(19), 5241-5249.

17. Larsen, M., et al. (2015). The role of *SMCHD1* in regulating D4Z4 methylation in FSHD2. *Neuromuscular Disorders*, 25(10), 791-799.

18. Lemmers, R. J., et al. (2009). A new paradigm for Facioscapulohumeral Muscular Dystrophy: *DUX4* expression and chromatin relaxation. *Science*, 326(5952), 595-598.

19. Lemmers, R. J., et al. (2010). Genetic and epigenetic basis of FSHD2. *Nature Genetics,* 42(6), 573-580.

20. Lemmers, R. J., et al. (2012). *SMCHD1* regulates a switch between epigenetic states at the FSHD-linked D4Z4 repeat. *Nature Genetics*, 44(12), 1385-1389.

21. Loonen, T. G. J., et al. (2021). The myopathic face in Facioscapulohumeral Muscular Dystrophy: Clinical manifestations and diagnostic significance. Journal of *Neuromuscular Disorders,* 8(2), 230-245.

22. Mariot, V., & Dumonceaux, J. (2022). Gene-editing strategies in FSHD: Current progress and future perspectives. *Trends in Molecular Medicine*, 28(4), 312-328.

23. Schätzl, T., et al. (2021). Pathological mechanisms in FSHD: A genetic and epigenetic perspective. *Orphanet Journal of Rare Diseases*, 16(1), 112.

24. Šikrová, D., et al. (2023). Advances in FSHD molecular research and targeted therapies. *Nature Reviews Neurology*, 19(5), 291-308.

25. Statland, J. M., & Tawil, R. (2014). Facioscapulohumeral Muscular Dystrophy: Molecular insights and therapeutic approaches. *Neurology Clinics*, 32(3), 721-736.

26. Tawil, R., et al. (2014). Pathogenic mechanisms of *DUX4* activation in FSHD. *Annals of Neurology,* 75(6), 849-861.

27. Tian, F., Oskoui, M., & Gagnon, C. (2015). Diagnosis, management, and clinical outcome measures in Facioscapulohumeral Muscular Dystrophy. *Canadian Journal of Neurological Sciences,* 42(Suppl 1), S121-S129.

28. Van Overveld, P. G., et al. (2003). Hypomethylation of D4Z4 in Facioscapulohumeral Muscular Dystrophy. *Annals of Neurology*, 54(4), 539-543.

29. Wang, L., et al. (2022). Gender differences in FSHD symptom progression. *Neuromuscular Disorders,* 32(3), 145-152.

30. Wilde, J. M., et al. (2016). Targeting p38 MAPK in FSHD therapy. *Molecular Therapy,* 24(4), 748-757.

31. Zeng, W., et al. (2009). Histone modifications and *DUX4* expression in FSHD. *PLoS Genetics,* 5(5), e1000559.

32. Huichalaf C et al. (2014). Micheloni S, Ferri G, Caccia R, Gabellini D. DNA methylation analysis of the macrosatellite repeat associated with FSHD muscular dystrophy at the single nucleotide level. *PLoS One*. *9*(12), e115278.

33. Tiffon C. (2018). The Impact of Nutrition and Environmental Epigenetics on Human Health and Disease. *Int J Mol Sci.*, 19(11), 3425.

34. Lemmers, R. J. L. F., Dandekar, R., & van der Maarel, S. M. (2024). Facioscapulohumeral muscular dystrophy diagnostic guidelines: 2024 update. *Neurology: Genetics*, 10(3), e000107.

35. Jia, Fangzhi Frank et al. (2021). Facioscapulohumeral muscular dystrophy type 2: an update on the clinical, genetic, and molecular findings, *Neuromuscular Disorders*, Volume 31, Issue 11, 1101 - 1112.

36. Katz, N., McHan, L., Tawil, R., McDermott, M., & Statland, J. (2024, March 3-6). Pediatric FSHD in the US: Results from the US National Registry for FSHD [Poster presentation]. 2024 MDA Clinical and Scientific Conference, Orlando, FL.

37. de Greef JC, et al. (2009). Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum Mutat.,* (10),1449-59.

38. Daxinger L, Tapscott SJ & van der Maarel SM. (2015). Genetic and epigenetic contributors to FSHD. *Curr Opin Genet Dev.,* 33:56-61

39. Jones TI, et al. (2014). Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clin Epigenetics* 6, 23.

40. Zeng W, et al. (2014). Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Hum Mutat, 35*(8), 998-1010.

41. van den Boogaard ML et al. (2016). Mutations in *DNMT3B* Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy. *Am J Hum Genet.,* 98(5), 1020-1029.

42. Hartweck LM et al. (2013). A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology*, 80(4),392-9.

43. Greco, A., Goossens, R., van Engelen, B., & van der Maarel, S. M. (2020). Consequences of epigenetic derepression in facioscapulohumeral muscular dystrophy. *Clinical Genetics*, 97(2), 199-214.

44. Calandra, P et al. (2015). Allele-specific DNA hypomethylation characterises FSHD1 and FSHD2. *Journal of Medical Genetics*, *53*(5), 348-355.

45. van Overveld et al. (2005). Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Annals of Neurology*, *58*(4), 569-576.

46. Preston MK, Tawil R & Wang LH. (2020). Facioscapulohumeral Muscular Dystrophy. In *GeneReviews.*

47. 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).

48. van der Maarel SM et al. (2000). De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am J Hum Genet,.*26-35.

49. Ricci G et al. (2016). A novel clinical tool to classify facioscapulohumeral muscular dystrophy phenotypes. *J Neurol.,*263(6):1204-14.

50. Mul, K., et al. (2018). Phenotype-genotype relations in facioscapulohumeral muscular dystrophy type 1. *Clinical Genetics*, 95(4), 529-537.

51. Mul, K., et al., (2018). Lifetime endogenous estrogen exposure and disease severity in female patients with facioscapulohumeral muscular dystrophy. *Neuromuscular Disorders,* 28(6), 508-511.

52. Tawil R, Van Der Maarel SM. (2006). Facioscapulohumeral muscular dystrophy. *Muscle Nerve*, 34(1), 1-15.

53. Tonini MM et al. (2004). Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul. Disord.,* 14, 33-38

54. Ricci, G., et al. (2013). Large-scale genotype-phenotype analysis indicates that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain, 136*(11), 3408-3417.

55. van der Maarel, S. M., Tawil, R., & Tapscott, S. J. (2011). Facioscapulohumeral muscular dystrophy and *DUX4*: Breaking the silence. *Trends in Molecular Medicine,* 17(5), 252-258.

56. Lutz, K. L., et al. (2013). Clinical and genetic features of hearing loss in facioscapulohumeral muscular dystrophy. *Neurology*, 81(18), 1571-1577.

57. Goselink, R. J. M. et al. (2019). Early onset as a marker for disease severity in facioscapulohumeral muscular dystrophy. *Neurology* 92, e378-e385

58. Lemmers, R. J. L. F., et al. (2022). Chromosome 10q-linked FSHD identifies DUX4 as principal disease gene. *Neurogenetics*, 23(1), 45-57.

59. Lemmers RJLF et al. (2018). Cis D4Z4 repeat duplications associated with facioscapulohumeral muscular dystrophy type 2. *Hum Mol Genet*.;27(20):3488-3497.

60. Sacconi, S. et al. (2012). Patients with a Phenotype Consistent with Facioscapulohumeral Muscular Dystrophy Display Genetic and Epigenetic Heterogeneity, *J. Med. Genet*., *49*(1), 41-46.

61. Sacconi S. et al. (2019). FSHD1 and FSHD2 form a disease continuum, *Neurology*, 92(19),e2273-e2285.

62. Lemmers RJ et al. (2015). Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum Mol Genet.,* 24(3),659-69.

63. Gaillard MC et al. (2014). Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. *Neurology,* 83(8),733-42.

64. Robin JD et al. (2015). *SORBS2* transcription is activated by telomere position effect over long distances upon telomere shortening in muscle cells from patients with facioscapulohumeral dystrophy. Genome Res., 25,1781-90.

65. Sacconi, S. et al. (2013). The FSHD2 gene *SMCHD1* is a modifier of disease severity in families affected by FSHD1. The American Journal of Human Genetics, 93(4), 744-751.

66. Nguyen K, et al. (2019). Deciphering the complexity of the 4q and 10q subtelomeres by molecular combing in healthy individuals and patients with facioscapulohumeral dystrophy. J Med Genet., 590-601.

67. Cohen J et al. (2020). Therapeutic Approaches in Facioscapulohumeral Muscular Dystrophy. *Trends Mol Med.* 2021 Feb;27(2):123-137.

68. Bao, B., & Yokota, T. (2015). Potential of antisense therapy for facioscapulohumeral muscular dystrophy. *Expert Opinion on Orphan Drugs*, 3(12), 1365-1374.

69. Richards, M., Coppée, F., Thomas, N., Belayew, A., & Upadhyaya, M. (2012). Facioscapulohumeral muscular dystrophy (FSHD): an enigma unravelled?. *Human genetics*, 131(3), 325-340.

70. Valente, A., Vieira, L., Silva, M. J., & Ventura, C. (2023). The Effect of Nanomaterials on DNA Methylation: A Review. *Nanomaterials*, 13(12), 1880.

72. Salsi, V., Magdinier, F., & Tupler, R. (2020). Does DNA Methylation Matter in FSHD? *Genes,* 11(3), 258.