ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

CLINICAL EXAMINATION OF A FAMILY WITH HEREDITARY SPASTIC PARAPARESIS AND ELUCIDATION OF THE *ERLIN1* GENE IN THE PATHOPHYSIOLOGY OF THE DISEASE

GÜLŞAH ŞEBNEM ÖZKÖSE
M.Sc. THESIS

DEPARTMENT OF GENOME STUDIES

SUPERVISOR
Prof. Özden Hatırnaz Ng

SECONDARY SUPERVISOR
Prof. Yasemin Alanay

ISTANBUL-2024

ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

CLINICAL EXAMINATION OF A FAMILY WITH HEREDITARY SPASTIC PARAPARESIS AND ELUCIDATION OF THE *ERLIN1* GENE IN THE PATHOPHYSIOLOGY OF THE DISEASE

GÜLŞAH ŞEBNEM ÖZKÖSE

M.Sc. THESIS

DEPARTMENT OF GENOME STUDIES

SUPERVISOR
Prof. Özden Hatırnaz Ng

SECONDARY SUPERVISOR
Prof. Yasemin Alanay

ISTANBUL-2024

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stage from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

24.06.2024
Gülşah Şebnem Özköse
Signature

PREFACE AND ACKNOWLEDGEMENT

Firstly, I would like to extend my deepest gratitude to my supervisor, Prof. Dr. Yasemin Alanay. Your endless knowledge, experience, and unwavering support have been the cornerstone of my master's journey.

To my co-supervisor, Prof. Dr. Özden Hatırnaz Ng, your expertise in molecular genetics and your willingness to guide me at every stage have been invaluable. Your detailed insights have been a beacon of light whenever I faced challenges.

My heartfelt thanks to my advisor Assoc. Prof. Dr. Özlem Akgün Doğan. Your academic and moral support in writing this thesis was instrumental in helping me complete this process.

To my advisor, Dr. Özkan Özdemir, your guidance in the lab has been a source of immense strength and positive envy.

I am deeply grateful to Prof. Dr. Uğur Özbek and Dr. Kaya Bilgüvar and Assoc. Prof. Dr. Cemaliye Akyerli Boylu whose influence sparked my interest in genome studies even before I began my education in the Department of Genome Studies.

Special thanks to my dear friend, Intern Dr. Beril Ay, Bio. Ceren Erdoğan, deeply grateful to PhD (c) İlayda Şahin and PhD (c) Gizem Önder, as well as all my classmates who have shared this journey with me.

To MSc. Psych. Ferahnaz Culum and MSc. Bengü Kaptanoğlu, my cherished friend who has lifted my spirits and brought joy during every challenging period, yours friendship means the world to me.

Finally, to my family: my mother, Hatice Şen Özköse, and my father, Hakan Özköse, for providing me with unwavering support and opportunities throughout my life; my beloved siblings, Ömer Tarık Özköse and Zeynep Ece Özköse; and my fiancé, Ali İyigel, who has stood by my side through every step of this journey.

I am profoundly lucky and endlessly grateful to have each of you in my life. Your support has made this accomplishment possible, and I carry your love and encouragement with me always.

Also thank Prof. Dr. Yasemin Topçu for her support of our research. This study was supported by the Acıbadem Mehmet Ali Aydınlar University BAP Coordination Unit. Project No: 2023/01-02.

TABLE OF CONTENTS

DECLARATION iii

PREFACE AND ACKNOWLEDGEMENT iv

TABLE OF CONTENTS v

LIST OF ABBREVIATIONS AND SYMBOLS vii

LIST OF FIGURES viii

LIST OF TABLES ix

ÖZET 1

ABSTRACT 2

1 INTRODUCTION AND AIM 3

2 BACKGROUND 5

2.1 Human Genome Project 5
2.2 Genomic (Personalized) Medicine 6
2.3 Genetic Testing 7
2.3.1 Sanger sequencing 9
2.3.2 Next generation sequencing 10
2.3.2.1 Targeted sequencing 11
2.3.2.2 Whole exome sequencing 12
2.3.2.3 Whole genome sequencing 12
2.3.3 Deep phenotyping 12
2.3.4 Genetic counseling 13
2.4 ACMG Guidelines, Variant Interpretation and Prioritization 13
2.5 Rare and Undiagnosed Diseases 14
2.5.1 Neurological and neurodegenerative disorders 15
2.5.1.1 Amyotrophic lateral sclerosis 16
2.5.1.1.1 Juvenile ALS & early-onset ALS 16
2.5.1.2 Hereditary spastic paraplegia 62 17
2.5.1.2.1 *ERLIN1* gene 18

3 MATERIALS AND METHODS 21

3.1 Materials 21
3.2.1 Ethical approval 23
3.2.2 Deep-phenotyping by a single geneticist 24
3.2.3 Peripheral 24

4 RESULTS 32

4.1 Deep Phenotyping 32
4.2 WBC, DNA and RNA Isolation 36
4.3 PCR and Agarose Gel Electrophoresis 37
4.4 Bidirectional Sanger Sequencing 38
4.5 Gene Expression Analysis 39

5 DISCUSSION 41

6 CONCLUSION 49

7 REFERENCES 50

8 APPENDIX 53

APPENDIX 1 53

APPENDIX 2 53

APPENDIX 3 53

9 CURRICULUM VITAE 59

LIST OF ABBREVIATIONS AND SYMBOLS

ACU	Acibadem University
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal dominant
ALS	Amyotrophic Lateral Scoliosis
AR	Autosomal recessive
ERLIN1	Endoplasmic Reticulum Lipid Raft Associated 1
HGMD	Human Gene Mutation Database
HGP	Human Genome Project
HSP	Hereditary Spastic Paraplegia
qPCR	Quantitative Polymerase Chain Reaction
NIH	National Institutes of Health
WHO	World Health Organization

LIST OF FIGURES

Figure 1. Molecular dysfunction mechanism of HSP (45) 20
Figure 2. *ERLIN1* protein structural diagram and reported mutations related to HSP62 (45).. 20
Figure 3. The gender and age distribution of the extended family 21
Figure 4. Study's workflow (45)...24
Figure 5. Agarose gel electrophoresis image part I...37
Figure 6. Agarose gel electrophoresis image part II..37
Figure 7. Confirmed variants of the nuclear family..38
Figure 8. Comparison of *ERLIN1* gene expression levels across zygosity.................39

LIST OF TABLES

Table 1. Primer information used in PCR 27
Table 2. Deep-phenotyping with HPO terms 34
Table 3. Concentration values of isolated DNA and RNA samples 36
Table 4. Zygosity information of 30 family members 38
Table 5. Confirmed zygosity infor according to Sanger results................................. 39

ÖZET

Herediter Spastik Paraparezili Bir Ailenin Klinik İncelenmesi ve *ERLIN1* Geninin Hastalığın Patofizyolojisindeki Rolünün Aydınlatılması

Acıbadem Üniversitesi Tıp Fakültesi Çocuk Genetik Bilim Dalına 2017 yılında parmak ucunda basarak dengesiz yürüme, ayak deformitesi ve konuşma bozukluğu şikayetleriyle başvuran yedi yaşındaki erkek çocukta, tüm ekzom dizileme analizi ile *ERLIN1* geninde prematüre stop kodona neden olan homozigot varyant saptandı. Anne ve babanın bu varyant için heterozigot taşıyıcı olduğu gösterildi. İndeks vakanın tanısı sırasında asemptomatik olan erkek kardeşinde benzer bulgular nedeniyle yapılan analizde aynı patojenik varyant homozigot olarak belirlendi. Aile öyküsünde, geniş ailedeki bazı bireylerde farklı klinik şiddette nörolojik bulguların olduğu öğrenildi. Literatürde *ERLIN1* ilişkili biallelik varyantlar herediter spastik paraparezi (HSP) ve erken başlangıçlı amyotrofik lateral skleroz (ALS) ile ilişkili olarak bildirilmiş ancak *ERLIN1* geninin bu iki fenotipin patofizyolojisindeki yeri açıklanamamıştır. Bu proje ile indeks vakanın ailesinde farklı nörolojik fenotipik özellikleri olan bireylerde *ERLIN1* geninde saptanan varyantı incelenmesi planlandı. Bu amaca yönelik çalışmaya katılmayı kabul eden HSP/ALS tanısı almış ya da etkilenmemiş toplam 30 aile bireyine "Biyobanka Onam Formu" sunulmuş, projeyle ilgili bilgi verilmiş ve onamları alınmıştır. Ardından periferik kan örneği alınarak akış şemasında yer alan moleküler teknikler sırasıyla uygulanmıştır (Şekil 4). Örnekler ACU Biyobanka biriminde birim standartlara göre numaralandırılarak arşivlenmiştir. Bu proje, literatürde bildirilmiş fakat fonksiyonel düzeyde kanıtı olmayan bir varyantın ilk destekleyici kanıtını ortaya koymayı, genotip-fenotip korelasyonunu kurarak *ERLIN1* ilişkili klinik spektrumu genişletmeyi ve hastadan elde edilen numuneyle yapılacak çalışmalarla *ERLIN1* ilişkili fonksiyonel çalışmaların öncüsü olmuştur. Çalışmamızla *ERLIN1* geninin HSP patofizyolojisinde oynadığı rolü aydınlatmakta, monoalelik (heterozigot) olarak taşınan varyantların nörolojik fenotiplerle olası ilişkisini genişletmekte ve monoalelik/bilalelik varyant varlığında *ERLIN1* geninin kandaki ekspresyon seviyelerindeki farklılıkları tespit etmektedir.

Anahtar Sözcükler: Herediter spastik paraparezi, *ERLIN1*, spastisite, endoplasmik retikulum.

ABSTRACT

Clinical Examination of a Family with Hereditary Spastic Paraparesis and Elucidation of the *ERLIN1* Gene in the Pathophysiology of the Disease

A homozygous variant causing premature stop codon in the *ERLIN1* gene was detected by whole exome sequencing analysis in a seven-year-old boy who presented to Acibadem University Faculty of Medicine, Department of Pediatric Genetics in 2017 with complaints of unstable gait, toe walking, foot deformity and speech disorder. Both parents were shown to be heterozygous carriers for this variant. His brother, who was asymptomatic at the time of diagnosis of the index case, was homozygous for the same pathogenic variant in the analysis performed due to similar findings. In the family history, it was learned that some members of the extended family had neurologic findings with different clinical severity. In the literature, *ERLIN1*-related biallelic variants have been reported to be associated with hereditary spastic paraplegia (HSP) and early-onset amyotrophic lateral sclerosis (ALS), but the role of the *ERLIN1* gene in the pathophysiology of these two phenotypes has not been elucidated. This project was planned to examine the variant detected in the *ERLIN1* gene in individuals with different neurologic phenotypic features in the family of the index case. The 'Biobank Consent Form' has been provided to a total of 30 family members, whether diagnosed with HSP/ALS or unaffected, who agreed to participate in this purpose-driven study. This was followed by peripheral blood sampling and sequential application outlined in the flow diagram (Fig 4). Samples were numbered and archived in the ACU Biobank Unit according to unit standards. This project provides initial evidence for a variant that has been reported in the literature but has not been demonstrated at the functional level, to establish genotype-phenotype correlations to broaden the clinical spectrum associated with *ERLIN1*, and to pioneer *ERLIN1*-related functional studies using patient samples. Our study elucidates the role of the *ERLIN1* gene in the pathophysiology of HSP, expands the potential association of monoallelic (heterozygous) variants with neurological phenotypes, and identifies differences in the expression levels of the ERLIN1 gene in the presence of monoallelic/biallelic variants.

Keywords: Hereditary Spastic Paraplegia, *ERLIN1,* spasticity, endoplasmic reticulum.

1 INTRODUCTION AND AIM

Genomic technology entered a new era with the sequencing of the entire human genome, followed by significant advancements in areas such as high-throughput sequencing, genome engineering, gene therapy, big data analysis, and storage (1). These developments have had a major impact on clinical diagnostics, particularly for complex, rare, and unknown disorders (2).

Rare diseases affect over 300-400 million individuals globally, resulting in chronic illnesses, disabilities, and premature mortality. The substantial number of patients who remain undiagnosed for years, coupled with many who succumb without an accurate diagnosis, has propelled researchers to enhance their methodologies (3). Genetic disorders are categorized into several sub-types, including Mendelian, chromosomal, and multifactorial disorders. Mendelian disorders also referred to as single-gene disorders, are attributed to errors in the DNA sequence of a gene and exhibit various inheritance patterns, such as autosomal dominant (AD), autosomal recessive (AR), X-linked, and mitochondrial inheritance. Chromosomal disorders, on the other hand, stem from numerical and structural aberrations. Multifactorial diseases arise from the interplay of genetic factors inherited from ancestors and environmental influences (4). Despite their diverse classifications, these disorders converge at a critical juncture: the aberrant behavior of genes during essential biological processes, which affects human developmental stages, gene regulation, and adaptation (5).

Discovering the relationship between genes and diseases and applying this knowledge in clinical settings for patients is akin to mining for gold. This process elucidates the mechanisms of diseases, which has numerous applications, including diagnosis, treatment, and prevention. High-throughput experimental techniques generate vast amounts of data, while computational prediction algorithms offer cost-effective and efficient methods for analysis (6).

Beyond personalized approaches in healthcare, certain strategies have a broader societal impact. These strategies consider the ethnic origins of specific groups, the

geographical regions where their ancestors resided, their dietary habits, exposure to chemicals, and human-made environmental disasters. The combination of this information with an individual's shared genetic background can yield novel insights for public health, such as detailed identification of infectious diseases, pathogen genomics, and the effects of these pathogens on large populations. This knowledge can accelerate the development of vaccines, improve antibiotic selection, and inform public health policies. As findings emerge, population-scale studies expand to the national level, enhancing drug usage policies, dosing guidelines, and precision medicine with minimal adverse effects (7).

The aim of this thesis is to investigate the genetic basis and phenotypic spectrum of spastic paraplegia type 62 (HSP62) and its potential overlap with juvenile or early-onset amyotrophic lateral sclerosis (ALS) in a consanguineous family. This study focused on the identification and characterization of the *ERLIN1* (NM_001100626.1), c.763C>T (p.Arg255*) variant and its role in disease manifestation. By conducting comprehensive genetic and phenotypic analyses, including in-depth clinical evaluations, the research aimed to elucidate the pathogenic mechanisms underlying these neurodegenerative conditions. The study provided insights on how to establish more robust gene-disease associations, inform the development of new diagnostic criteria, and propose guidelines for disease management and treatment based on the pathophysiology of HSP62 and early-onset ALS.

2 BACKGROUND

2.1 Human Genome Project

The entire genetic code of an organism is referred to as its genome, which serves as a roadmap for development and function. This structure is called DNA and is made up of adenine (A), thymine (T), cytosine (C), and guanine (G) bases. It dictates unique traits and responses to environmental factors. Organisms organize their DNA into chromosomes, with humans possessing 46 chromosomes (23 pairs). Genes, which are segments of DNA, encode proteins, influencing various functions of an organism (8).

The Human Genome Project (HGP) launched in 1990 and was completed in 2003. As a groundbreaking global scientific breakthrough, the doors of a whole new world were opened, and the most important feature of this breakthrough is the production of the first sequence of the human genome covering 92% of the human genome. Since there are a lot of complex and highly repetitive DNA parts, the whole genome has not been fully sequenced since March 2022 (9). From the perspective of a broader framework, it becomes evident that the implications of this project extend far beyond the boundaries of biology, technology, and medicine and are of social and future interest. In particular, the fields of technology and health have witnessed a marked acceleration in their respective developments, with the Human Genome Project continuing to advance at an exponential rate (10).

The initial impact of the Human Genome Project has been profound in the field of comprehensive discovery of novel genes, proteins, and non-coding regulatory RNAs. This monumental project has fueled a burgeoning interest in understanding and dissecting complex biological systems, elucidating how each component interacts within these systems, and their relationship to larger physiological networks. Consequently, the field of proteomics has emerged, focusing on the identification and organelles, tissues, and blood. Additionally, the HGP has revolutionized our comprehension of evolution by providing access to over 4,000 genome sequences. Advanced bioinformatic methods have been developed, fostering interdisciplinary

collaboration and making data readily accessible to the public through user-friendly databases and open-source software. This data sharing is crucial for advancing biology and medicine, enabling a wide array of researchers to address complex problems and develop future healthcare solutions (10).

2.2 Genomic (Personalized) Medicine

Genomic medicine is an emerging field that integrates genomic information about individuals into their clinical care. This approach involves using genomic data for diagnostic and therapeutic decision-making, while also taking into account the health outcomes and policy implications associated with its clinical application (11).

Genomic personalized medicine represents a transformative approach in healthcare, where genomic information is harnessed to tailor medical care to the individual characteristics of each patient. This field integrates comprehensive genomic data, including whole genome sequencing, exome sequencing, and targeted gene panels, to inform diagnostic, therapeutic, and preventive strategies that are specific to each patient's unique genetic makeup.

In the realm of diagnostics, genomic personalized medicine enables the identification of genetic variants that contribute to disease susceptibility, onset, and progression. By understanding these genetic underpinnings, clinicians can offer precise diagnoses, especially in cases of rare or undiagnosed conditions. This approach significantly enhances the accuracy of differential diagnoses and allows for the identification of previously unrecognized genetic disorders (11).

Genomic data also play a pivotal role in guiding therapeutic interventions. Pharmacogenomics, a key component of personalized medicine, examines how genetic variations affect an individual’s response to drugs. This knowledge allows for the optimization of drug selection and dosing, thereby maximizing therapeutic efficacy and minimizing adverse effects. For instance, in oncology, genomic profiling of

tumors can identify specific mutations that are amenable to targeted therapies, leading to more effective and personalized cancer treatments.

Preventive healthcare is another critical aspect of genomic personalized medicine. By identifying individuals at high genetic risk for certain diseases, clinicians can implement early intervention strategies. This includes lifestyle modifications, increased surveillance, and prophylactic treatments aimed at mitigating the risk of disease development. For example, individuals with BRCA1 or BRCA2 mutations can receive tailored screening and preventive measures for breast and ovarian cancer.

The integration of genomic personalized medicine into clinical practice also necessitates careful consideration of ethical and policy implications. Issues such as genetic privacy, informed consent, and equitable access to genomic technologies are paramount. It is essential to establish robust ethical frameworks and policies that protect patient confidentiality and ensure that the benefits of genomic medicine are accessible to all segments of the population.

As genomic technologies continue to advance, the scope of personalized medicine is expected to expand. Emerging fields such as epigenomics, transcriptomics, and metabolomics will further enhance our understanding of the complex interplay between genes and their expression in health and disease. Continued research and collaboration across scientific disciplines will be crucial in realizing the full potential of genomic personalized medicine, ultimately leading to improved patient outcomes and the advancement of precision healthcare (11).

2.3 Genetic Testing

Genetic testing involves analyzing DNA to identify variations that may impact an individual's health, typically using biological samples such as blood or saliva. There are four main reasons for genetic testing; to find out if there is a family history of genetic disease before symptoms appear, the likelihood of a current or future pregnancy having a genetic condition, to diagnose a genetic condition if the individual

or their child has symptoms, understand, and guide your cancer prevention or treatment plan (12). Genetic tests are categorized based on the underlying rationale for their use (13).

- Predictive testing is a procedure that is conducted for individuals who have a family member with a genetic disorder (14). The results assist in determining the probability of an individual developing the specific disorder. These tests are performed before any symptoms appear.
- Diagnostic testing is a valuable tool for confirming or ruling out a suspected genetic disorder, providing individuals with the knowledge necessary to make informed decisions about managing or treating their health (14).
- Pharmacogenomic testing provides insight into how an individual's genetic makeup influences their response to specific medications (14). It assists healthcare providers in determining the most effective treatment strategies and in avoiding adverse drug reactions.
- Reproductive testing is designed to facilitate family planning and the expansion of the family unit. This process involves administering tests to identify the genetic variants carried by the biological parents. These tests assist parents and healthcare providers in making informed decisions before, during, and after pregnancy.
- Direct-to-consumer testing can be conducted without a healthcare provider by collecting a DNA sample at home, such as by spitting saliva into a tube or taking an epithelial cell swab from the inside of the mouth and sending it to a company. The company analyzes the DNA and provides information about ancestry, consanguinity, lifestyle factors, and potential disease risk.
- Forensic testing involves the scientific analysis of various types of physical evidence, such as DNA, fingerprints, and blood spatter, collected from crime scenes, victims, or suspects. The primary goal of forensic testing is to provide information for legal purposes, including identifying biological family members and suspects, linking evidence to crime scenes, and reconstructing events related to crimes or disasters (13).

Testing options depend on factors such as patient age, family history, and patient preference, and they continually evolve with emerging technologies. Additionally, the timing of the test, whether during pregnancy or after birth, is crucial in determining which test to perform (15). The healthcare provider is primarily responsible for offering genetic testing, a responsibility that has become more challenging with the increasing number of options. Patients should make voluntary decisions about testing during pregnancy based on their values and needs. Follow-up testing should also be based on these factors (16).

2.3.1 Sanger sequencing

Sanger sequencing, known as a first-generation DNA sequencing method, remains essential in clinical genomics despite the development of next-generation sequencing (NGS) techniques that offer much higher throughput. Sanger sequencing is particularly valuable for two main purposes. First, it serves as an independent method for confirming sequence variants identified by NGS. In the clinical setting, it is used to validate NGS results by providing a reference against which NGS assays can be compared. While it is impractical to confirm every NGS-identified variant with Sanger sequencing due to the high volume of primers and reactions required, it is particularly useful for verifying variants that are biologically questionable or suspected to be spurious. Second, Sanger sequencing helps compensate for coverage gaps in regions poorly covered by NGS. Certain genomic regions, often rich in GC content, may resist sequencing due to poor capture or amplification. In such cases, Sanger sequencing can be used to ensure comprehensive coverage by targeting these difficult-to-sequence regions (17).

Bidirectional Sanger sequencing plays a critical role in variant testing by improving the accuracy, reliability, and comprehensiveness of genetic analysis, particularly in research and clinical settings where accurate variant detection is critical for diagnosis, prognosis, and treatment decisions (18).

- Error reduction: It helps minimize sequencing errors by confirming the presence of variants in both the forward and reverse directions. Consistency between the two sequences increases confidence in the detected variants and reduces the likelihood of false positives.
- Detection of heterozygous variants: Bidirectional sequencing is especially important for detecting heterozygous variants, where only one of the two alleles carries the mutation. Sequencing both strands of DNA ensures that variants present on either strand are accurately identified.
- Confirmation of variants: Variants identified in one direction can be confirmed by sequencing the complementary DNA strand. This confirmation step is essential to validate the presence of variants and distinguish them from sequencing artifacts or errors.
- Identification of insertions and deletions: Bidirectional sequencing assists in the detection of insertions and deletions (indels) by identifying mismatches between the forward and reverse sequences. This is critical for comprehensive variant analysis, especially in regions prone to indel mutations.
- Quality Control: Serves as a quality control measure, allowing researchers to assess sequencing accuracy and reliability. Consistent results from both sequencing directions indicate high-quality data, while discrepancies may indicate sequencing errors or complex genomic features.
- Comprehensive coverage: By sequencing both strands, bidirectional Sanger sequencing provides more comprehensive coverage of the target region, increasing the likelihood of detecting variants even in regions with complex sequence features or structural variations (18).

2.3.2 Next generation sequencing

In genomic research, next-generation sequencing (NGS) is a powerful tool. NGS can provide comprehensive information regarding genome structure, genetic variants, gene activity, and alterations in gene behavior by simultaneously sequencing millions of DNA fragments. Recent developments have focused on improved data analysis, reduced costs, and faster and more accurate sequencing. These advances allow for

deeper and higher-resolution information, enhancing our understanding of genomics, disease, and personalized medicine. The basic principle of NGS testing is consistent across all available platforms, combining wet lab analysis with bioinformatics expertise (19).

NGS involves several key steps to generate comprehensive genomic data. First, DNA or RNA is extracted from a biological sample and fragmented into smaller pieces. These fragments are then ligated to adapters to create a library, which is amplified using polymerase chain reaction (PCR) to increase DNA quantity. The prepared library is loaded onto a sequencing platform, where sequencing by synthesis or other methods is used to read the DNA fragments, producing short reads of DNA sequences. The raw data from the sequencer undergoes base calling to determine the nucleotide sequence in each read. These short reads are aligned to a reference genome or assembled de novo, and differences between the sequenced genome and the reference are identified, including single nucleotide variants (SNVs), insertions, deletions, and structural variants. Identified variants are annotated with information about their potential effects, filtered, and prioritized based on relevance. In clinical settings, results are interpreted by geneticists and healthcare providers to understand their implications for patient health and treatment. Finally, a comprehensive report is generated, summarizing the findings and their significance, with genetic counseling provided to patients as needed (20).

2.3.2.1 Targeted sequencing

Targeted sequencing involves selecting a set of approximately 20-200 genes related to a specific disease for laboratory analysis. This method is cost-effective and provides high-quality data by focusing on a small number of genes. However, the main disadvantage of clinical use is that it does not allow for the addition or re-analysis of genes that may become relevant in the future (21).

2.3.2.2 Whole exome sequencing

WES is an advanced, efficient, and cost-effective genetic diagnostic tool. It allows for the detection of various genetic variants, including point mutations (missense and nonsense mutations), small deletions, insertions, and canonical splicing mutations. Despite its robust capabilities, linking genetic variations to specific diseases remains a complex challenge. It also demonstrated significant diagnostic utility across a broad spectrum of genetic disorders. It has been particularly effective in diagnosing conditions with well-established molecular mechanisms, achieving high diagnostic success rates (22).

Whole exome sequencing (WES) looks at all the protein-coding parts (exons) of all human genes. Some tests focus on just the exons in about 4,000 genes known to have disease-related variants, called "clinical exomes" (CES). Both CES and WES offer a balance between depth and cost (21).

2.3.2.3 Whole genome sequencing

Whole Genome Sequencing (WGS) examines 95-98% of all human DNA without any selection steps, covering both exons and introns. This comprehensive approach can detect mutations that other methods may miss. However, WGS is the most expensive and data-intensive method and generates a substantial amount of information about intronic genetic variation, which is less well understood today (21).

2.3.3 Deep phenotyping

Deep phenotyping involves comprehensive and precise characterization of phenotypic traits, significantly enhancing the diagnostic rate of genetic testing. By correlating detailed clinical and molecular data, it helps identify subtle phenotypic variations that may be missed in standard assessments, thus improving diagnostic accuracy. Integrating deep phenotyping into genetic studies broadens the clinical spectrum of many diseases. It enables the recognition of new phenotypic subgroups

within known disorders and reveals previously unrecognized manifestations, expanding the understanding of disease variability (23).

2.3.4 Genetic counseling

Genetic counseling is a patient-centered, non-directive, and comprehensive process that empowers individuals and families to understand and cope with the medical, psychological, and familial implications of genetic factors related to disease (24).

Today, genetic counseling is provided by healthcare professionals who have been trained in medical genetics and are called "genetic counselors". Genetic counselors help identify families at risk of genetic conditions by determining the inheritance patterns of diseases based on family history and calculating the likelihood of recurrence. They provide information about genetic testing and related procedures and are trained to explain complex information about risk management, genetic testing, and diagnosis to patients and their relatives. They can provide information on the importance of genetic conditions, discuss available options, and provide information on managing the genetic condition. Additionally, they offer referrals to educational services, advocacy or support groups, other health professionals, and community or government services that may be required based on the genetic condition (14).

As precision medicine and genetic testing continue to evolve, many healthcare providers should be prepared to interpret genetic test results, establish communication, and provide decision support to their patients and their families. The ability of genetic counselors to assess the needs of their clients and the impact of genetic outcomes on their families has facilitated the development of the genetic counseling field (11).

2.4 ACMG Guidelines, Variant Interpretation and Prioritization

Rapid advancements in sequencing technology, particularly high-throughput next-generation sequencing, have made interpreting sequence variants more

important. The American College of Medical Genetics and Genomics (ACMG) has developed guidelines to enhance the understanding of variants observed in the human genome. ACMG guidelines define novel universal terminology for describing variants identified in Mendelian disorders as ‘pathogenic,’ ‘likely pathogenic,’ ‘uncertain significance,’ ‘likely benign,’ and ‘benign’. These classifications are based on population data, functional data, segregation data, in-silico and other prediction data, de novo data, allelic data, and other databases. This approach is crucial for evaluating variants detected by a laboratory within the context of the patient's and family's history, physical examinations, and previous laboratory tests, to distinguish between variants causing the patient's disorder and those that are incidental (secondary) findings or benign. Given the complexity of genetic testing, the ACMG strongly recommends that clinical molecular genetic testing be performed in laboratories approved by the Clinical Laboratory Improvement Amendments (CLIA) and interpreted by certified professionals. technology, including high-throughput next-generation sequencing (25).

2.5 Rare and Undiagnosed Diseases

The European Union (EU) defines a rare disease as affecting less than 1 in 2,000 individuals. Estimates suggest that between 6,000 and 8,000 distinct rare diseases impact approximately 30 million people within the EU (26). Conversely, the United States defines a rare disease as affecting fewer than 200,000 citizens (27). Rare diseases are becoming more prevalent every day. Between 263 and 446 million people worldwide are affected by rare diseases, according to a 2019 study using the Orphanet database (28).

The definition of a rare disease lacks a universally accepted standard. Variation exists due to epidemiological factors, healthcare systems, record-keeping practices, identification and classification methodologies, and the interplay of genetic and environmental influences across geographical regions (29). To address this disparity and promote international consistency in defining and documenting rare diseases, the World Health Organization (WHO) established a dedicated numbering system within

the International Classification of Diseases (ICD). The current version, ICD-11, implemented in January 2022, encompasses over 550 rare diseases, each assigned a unique code (30).

In Türkiye, rare diseases are estimated to affect approximately 6-8% of the population, although there is no official data. This rate is compounded by the prevalence of consanguineous marriages, which increase the likelihood of inheriting genetic disorders and occur in about 25% of the population. Türkiye witnesses the identification of 3-4 new rare diseases each year, with many manifesting in early childhood or infancy. Certain diseases, such as Behçet's disease, Familial Mediterranean Fever, and Mediterranean anemia, are more frequently observed in the Turkish population despite being considered rare globally (31).

2.5.1 Neurological and neurodegenerative disorders

According to the World Health Organization, diseases affecting the central nervous system and the peripheral nervous system are classified as neurological diseases. More than 6 million people die each year from stroke. Over 50 million people suffer from epilepsy, and 47.5 million from dementia, with 7.7 million new cases of Alzheimer's disease annually (32). Neurological diseases are associated with high morbidity and mortality, representing a significant burden on healthcare systems. They are the primary cause of disability and the second leading cause of death (33).

Neurodegenerative diseases affect millions of individuals worldwide, with aging as the primary risk factor. Recent studies indicate that both genetic factors and environmental influences collaborate to increase susceptibility to neurodegenerative disorders (NDs). Specific genes can elevate the risk, while the timing and extent of neurodegeneration are influenced by the immediate environment. Furthermore, emerging research suggests that a single ND may encompass multiple concurrent diseases, which can impact the severity based on the type and stage of the disorder (34).

Neurodegeneration, characterized by the progressive loss of neurons and their functions, is central to various brain-related disorders. Despite extensive research, effective treatments for neurodegenerative diseases remain elusive, posing ongoing challenges, particularly for older individuals. The blood-brain barrier (BBB), a robust defense mechanism, restricts most foreign substances from entering the brain, complicating the development of therapies targeting the pathogenesis of these diseases. While surgical interventions show promise, concerns about their long-term benefits and potential damage to the BBB limit their widespread adoption. Nanotherapeutics, designed to cross the BBB safely, offer advantages over conventional treatments, though ongoing research seeks to optimize their effectiveness (34).

2.5.1.1 Amyotrophic lateral sclerosis

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease affecting both upper and lower motor neurons throughout the brain, brainstem, and spinal cord, typically resulting in a life expectancy of 3-5 years. Its symptoms include muscle weakness, atrophy, fasciculations, and difficulty with swallowing and speaking. Diagnosing ALS is complex, requiring the exclusion of diseases with similar symptoms. Electromyography (EMG) and nerve conduction studies (NCS) are crucial diagnostic tools. ALS encompasses various forms and variants, such as primary lateral sclerosis (PLS), progressive muscular atrophy (PMA), and progressive bulbar palsy (PBP), diagnosed using the El Escorial criteria based on clinical signs. Though its exact cause remains unclear, ALS is believed to stem from a combination of genetic and environmental factors. The mainstay of ALS treatment involves managing symptoms and providing supportive care, utilizing physical and occupational therapy, medications, and assistive devices to enhance quality of life (35).

2.5.1.1.1 Juvenile ALS & early-onset ALS

Juvenile ALS is a rare form of Amyotrophic Lateral Sclerosis (ALS) that affects individuals before the age of 25, accounting for only about 1% of all ALS cases. This

form is predominantly caused by genetic mutations. Researchers have identified mutations in several genes, including *SPG11, SOD1, SPTLC1, UBQLN2, SIGMAR1, FUS, ALS2, and SETX*, which are strongly linked to the disease (36). Symptoms and progression can vary widely, but common signs include muscle weakness, wasting, and increasing difficulty with movement. While there is no cure for juvenile ALS, treatments can help manage symptoms and improve the patient's quality of life (36).

Early-onset ALS is a broader term that encompasses ALS diagnosed before the age of 45, including juvenile ALS, but also cases identified between 25 and 45 years old. Approximately 10% of patients with ALS have a family history of the disease, with the remainder classified as sporadic cases. Early-onset ALS constitutes about 10% of all ALS diagnoses (37). Researchers have identified mutations in several genes, including *SOD1, MATR3, CHCHD10, TBK1, TUBA4A, NEK1, C21orf2, and CCNF*, although the cause may remain unknown in some cases (38). Additionally, ALS has many genetic causes and a broad spectrum with variable disease progression, thus personalized medicine is becoming essential, with treatments tailored to each patient's specific mutations. Therefore, genetic screening for known mutations will be crucial for the diagnosis, treatment, and prevention of ALS (38).

2.5.1.2 Hereditary spastic paraplegia 62

Hereditary spastic paraplegias (HSPs) are a diverse group of neurodegenerative disorders characterized by prominent lower limb spasticity due to corticospinal upper motor neuron axonopathy. These disorders manifest in different ages of onsets, both "pure" and "complex" forms with additional features. HSPs are genetically diverse, with over 70 different genetic loci and approximately 60 mutated genes identified (39). Studies on the molecular pathogenesis of HSPs underscore the importance of basic cellular functions such as membrane trafficking, mitochondrial function, organelle formation and biogenesis, axon transport, and lipid/cholesterol metabolism in axon development and maintenance (39).

Hereditary Spastic Paraplegia 62 (HSP62) is a genetic disease inherited as an autosomal recessive trait, characterized by a variety of symptoms affecting different systems of the body. It is a progressive disorder that worsens over time, with an age of onset ranging from 20 months to 13 years (40).

The clinical manifestations of HSP62 primarily affect the skeletal and neurological systems. The skeletal system, especially the spine and limbs, is significantly affected. Thoracic scoliosis, a curvature of the upper back, is common. In addition, patients may develop flexion contractures of the knees, resulting in stiffness and flexed knee joints. Muscle wasting, or amyotrophy, is a prominent muscle and soft tissue problem associated with HSP62. Neurological symptoms are most prominent. The central nervous system is severely affected, resulting in lower limb spasticity - abnormal tightness and stiffness in the legs. This condition results in a spastic gait that makes walking difficult. Because of the increased muscle tone, many patients walk on their toes. Hyperreflexia, an exaggerated reflex response, is often observed. Clonus, characterized by involuntary rhythmic muscle contractions, may also occur. Some patients have absent reflexes, such as the patellar and Achilles tendon reflexes, indicating nerve dysfunction. Dysarthria, a condition that causes slurred or slow speech due to muscle weakness, is also a symptom of HSP62 (#615681, MIM) (40).

Since it is very rare, HSP62 is mentioned in the 'Conditions/Phenotypes' section of the Genetic Testing Registry of the National Institutes of Health (NIH). This entry provides information on its clinical findings, testing options for detecting the condition, and its relationship to the *ERLIN1* gene. Additionally, it references the Orphanet registry, which classifies HSP62 under the autosomal recessive inheritance model (41, 42).

2.5.1.2.1 *ERLIN1* gene

The endoplasmic reticulum (ER) is a multifunctional eukaryotic organelle essential for numerous vital processes. It serves as the primary site for protein synthesis

and the biogenesis of membrane lipids, and it acts as the cell's internal calcium storehouse. The ER coordinates with other organelles, such as ribosomes for protein assembly and the Golgi apparatus for post-translational modifications. Its dynamic structure, consisting of interconnected tubules, sheets, and continuity with the nuclear envelope, allows it to adapt to various cellular demands and signaling cues, ensuring the efficient execution of its multifaceted roles (43).

Hereditary Spastic Paraplegia (HSP) primarily involves the degeneration of long axons in the central nervous system, particularly in the corticospinal tract. This degeneration is most noticeable in the thoracic and cervical regions of the spinal cord, resulting in demyelination. The condition can also affect the cerebellum and brainstem and is associated with a reduced number of neurons. HSP often leads to peripheral neuropathy, affecting the basal ganglia, cerebellum, and spinal cord, causing various motor dysfunctions and myelin sheath abnormalities. Other signs include cerebellar atrophy, reduced spinal cord diameter, and a thin corpus callosum (44).

HSP62 is associated with mutations in the ER lipid raft-associated protein 1 (ERLIN1/SPG62), occurring either in a homozygous or compound heterozygous state, located at the q24.31 and q24.32-33 loci (44).

In Hereditary Spastic Paraplegia (HSP), mutations in genes within the corticospinal neurons lead to several dysfunctions. These include disruption in organelle structure and transport, and mitochondrial dysfunction near the neuronal nucleus. Some mutations impair axonal transmission, while others cause myelin sheath degeneration. Additionally, mutations affecting the endoplasmic reticulum result in defective lipid metabolism. These molecular abnormalities lead to lower limb spasticity and weakness, which are characteristic symptoms of HSP phenotypes (Figure 1) (44).

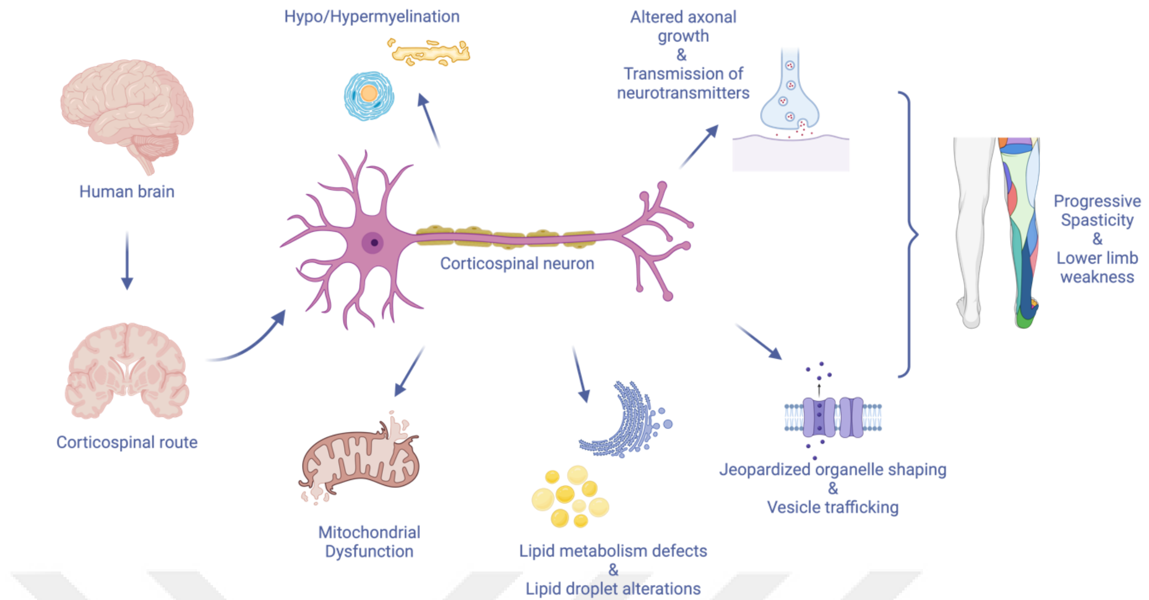


Figure 1. Molecular dysfunction mechanism of HSP (45)

In the literature, there are six variants reported in the *ERLIN1* gene so far. A termination, missense, and frameshift variants reported as follows; c.763C>T, c.149G>T, and c.862_868delACCAGG reported in 2014 (46); missense variant c.281T>C in 2017 (47); c.504+1G>A splicing site mutation reported by Zhu et al. in 2022 (48) and finally, c.195G>C missense variant reported in 2024 (49) (Figure 2).

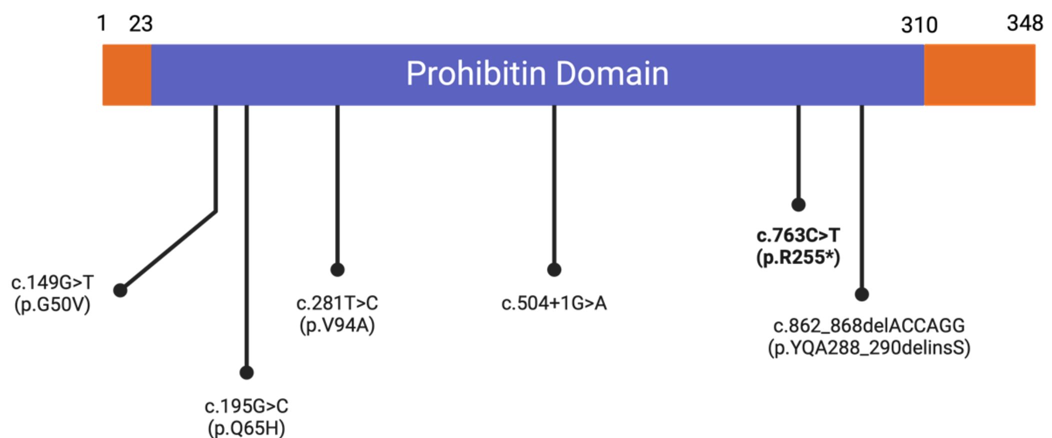


Figure 2. ERLIN1 Protein Structural Diagram and Reported Mutations Related to HSP62 (45)

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Patient and control selection

The index was 7-year-old boy, the first child of consanguineous healthy patients, with progressive walking difficulties; gait ataxia, toe walking, and pes valgus who applied to Acibadem Maslak Hospital in 2017.

Detailed phenotyping of the extended family was performed. Due to the neurologic and various systemic findings observed, we included all available members of the extended family in our experimental group. The extended family includes 30 individuals, 16 females and 14 males (Figure 3A). The age distribution of the 30 individuals in our study is as follows: 9 individuals aged 18 years or younger, 4 individuals aged 19-30 years, 11 individuals aged 31-49 years, and 6 individuals aged 50 years or older (Figure 3B).

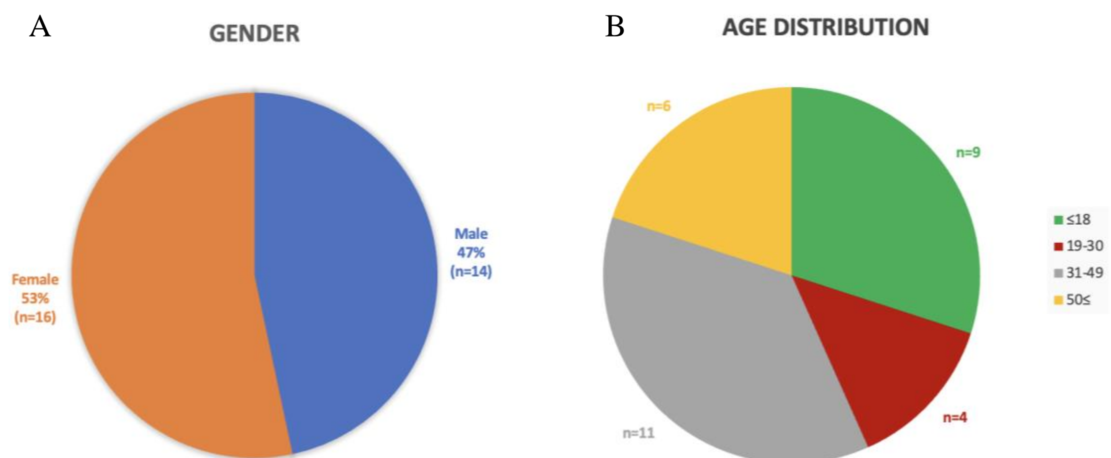


Figure 3. The gender and age distribution of the extended family

3.1.2 Chemicals

- Absolute Ethanol (100%)
- Agarose (Merck, Germany)

- Coomassie Blue dye (Merck, Germany)
- Ethanol for lab cleaning (Isolab, Germany)
- Ethanol for solution mixtures (Merck, Germany)
- Ethidium Bromide (Sigma, Germany)
- Nuclease-free water (Thermo Fisher, USA)
- Phosphate Buffered Saline (PBS) (Thermo Fisher, USA)
- Proteinase K
- 50X TAE Buffer (Sigma, Germany)

3.1.3 Consumables

- Microcentrifuge tube (Isolab, Germany)
- Micropipette (Gilson, USA)
- Parafilm M Bemis Sealing Film
- Pipettes (Thermo Fisher, USA)
- 15 mL and 50 mL falcon tubes (Isolab, Germany)

3.1.4 Laboratory Devices

- Cell counter (RVD Automatic Cell Conter
- Centrifuge
- CFX96 Real-Time System, C1000 Touch Thermal Cycler (Biorad, USA)
- ChemiDoc MP Imaging System with universal Hood III (BioRad, USA)
- Cold room
- Electrophoresis instrument (Biorad, USA)
- Freezer
- Fume Hood (Thermo Fisher, USA)
- Heat Block
- Qubit 4 Fluorometer (Thermo Fisher, USA)
- Rack types (mini and normal)
- Refrigerator (+4°C, -20°C, -80°C) (Nuve, Turkiye)

- Spectrophotometer (Nanodrop) (Thermo Fisher, USA)
- Tabletop mini centrifuge
- T100 Thermal Cycler (Biorad, USA)
- Waterbath (Weightlab, Turkey)

3.1.5 Kits

- BlasTaq 2X qPCR MasterMix (abmgood, Canada)
- Hot Start PCR Master mix (NucleoGene, Turkey)
- Reverse Transcriptase First Strand cDNA Kit (Thermo Fisher, USA)
- RNeasy Mini Kit (Qiagen, USA)
- QIAamp Blood Maxi Kit (Qiagen, USA)
- 100 bp DNA Ladder Plus Marker (Cat#: GMM100P, GeneMark)

3.1.6 Services

- Data storage service from ACU Biobank,
- Designed primer sequences were sent to Sentebiolab Biotech for primer sythesis,
- Sanger sequencing service from MSM Products and Services (via Macrogen).

3.2 Methods

3.2.1 Ethical approval

The Ethics Committee of Acibadem Mehmet Ali Aydinlar University, Istanbul, Türkiye, approved this study under the reference number ATADEK-2022/17. The study included four nuclear and 26 extended family members. Written informed consent was obtained from all participants. The clinician informed the nuclear family about the future aspects of the study, while the researcher provided detailed information to the extended family members about the project's research objectives,

potential outcomes, and importance. This ensured that all participants were fully aware of the study's goals and their role in it.

3.2.2 Deep-phenotyping by a single geneticist

The deep phenotyping process was performed by a single clinician with extensive knowledge of the index case and his brother, providing more than 10 Human Phenotype Ontology terms per patient.

Since extended family members live outside Istanbul, information about their medical history or condition was provided by family members, not the clinician. The workflow of the study, initiated after reviewing the literature on the variant observed in the family presenting to the clinician with a specific complaint, is shown in Figure 4.

Figure 4. Study's workflow (45)

3.2.3 Peripheral blood collection

Peripheral blood samples were collected from a total of 30 individuals, including diagnosed patients and their family members. Fourteen of these blood samples were

obtained from Arhavi, Artvin, while the remaining samples were collected at Acıbadem Maslak and Acıbadem Bakırköy hospitals in Istanbul.

3.2.4 WBC isolation from blood

WBC isolation began by bringing the incoming samples to room temperature (15-25°C). The centrifuge was pre-cooled. Sample registration was checked, and identification of the samples was noted for labeling. After determining the sample volume, it was transferred to a 15 mL or 30 mL falcon tube. Three times the volume of the blood was added as 1X RBC solution (e.g., 7.5 mL for a 2.5 mL sample). The tubes were gently inverted and incubated at +4°C for 25 minutes, with inversion every 5 minutes. The tubes were then centrifuged at 1500 RPM for 10 minutes, and the supernatant was removed.

1 mL of RBC solution was added to the remaining pellet and vortexed until homogenized. The homogenized sample was transferred to a microcentrifuge tube labeled appropriately and vortexed. Cold-centrifugation was performed at 5000 RPM for 5 minutes, and the supernatant was removed.

1500 µL of PBS solution were added to the microcentrifuge tubes and vortexed. Cell counter slides were labeled for cell counting. 10 µL of sample and 10 µL of trypan blue were vortexed and transferred to the slide. The sample was then divided into two microcentrifuge tubes, each containing 750 µL. The tubes were centrifuged at 5000 RPM for 5 minutes, and the supernatant was removed using a micropipette. The cell pellet was stored at -80°C.

3.2.5 DNA isolation and quantification from blood

Blood samples collected from patients in EDTA tubes were isolated using the pre-optimized QIAamp Blood Maxi kit protocol (Qiagen, USA). The concentration and purity values of the DNA obtained were measured using a spectrophotometer device (Nanodrop).

The ratio of absorbance at 260 nm (peak DNA absorption) to 280 nm (protein absorption) should be between 1.8 and 2.0. Values outside this range were not used in downstream applications.

3.2.6 RNA isolation from white blood cells

Isolating RNA from White Blood Cells (WBC) involves lysing the cells to release RNA, followed by stabilization to prevent degradation. RNA extraction method silica column purification is then used to isolate RNA from the lysate, with subsequent purification and quantification steps to ensure high-quality RNA for downstream applications. Required amounts were determined based on the cell counts of the previously isolated WBCs. The isolation procedures were then performed using the Qiagen RNeasy Mini Kit according to the Purification of Total RNA from Animal Cells Using the Spin Technology protocol.

After RNA isolation, fluorescence-based quality control assays can be performed to assess the quantity and quality of the isolated RNA.

RNA quality measurements have been performed with both Qubit 4 Fluorometer (Thermo Fisher, USA) and Spectrophotometer (Nanodrop) (Thermo Fisher, USA). According to the results, it was decided that some of the RNA samples were not suitable for the gene expression study, and only real-time qPCR was performed on the selected RNA samples.

RNA purity is assessed using two absorbance ratios. The 260/280 ratio, comparing absorbance at 260 nm (RNA peak) to 280 nm (protein absorption), ideally falls between 1.8 and 2.1, indicating minimal protein contamination. Additionally, the 260/230 ratio, comparing absorbance at 260 nm to 230 nm (contaminant absorption), typically ranges from 2.0 to 2.2, signifying low levels of other impurities.

3.2.7 Primer design and validation

Specific primers for PCR amplification were designed using Primer3 software. The target sequence for primer design was retrieved from the Ensembl transcript NM_001100626.1. Primers were designed to flank the region of interest with an optimal amplicon size of 137 bp.

Following in silico primer design, several validation steps were performed to ensure primer suitability for PCR. The primers were analyzed using SNPCheck to assess their specificity and minimize the possibility of amplifying unintended targets due to single nucleotide polymorphisms (SNPs). Additionally, the primers were evaluated in silico using the USCS In Silico PCR tool to confirm their ability to specifically amplify the desired region from the chosen reference sequence.

Finally, the primers were analyzed using the IDT OligoAnalyzer tool to assess their melting temperature (Tm), GC content, and potential for self-dimerization or hairpin formation. Primers with optimal Tm values (around 60°C), balanced GC content (between 40-60%), and minimal potential for secondary structures were selected for further experiments seen in Table 1.

Table 1. Primer information used in PCR

	Primers	Sequence	Tm	GC%
ERLIN1	Forward	GGGAGCCAGTCACCTTGTTT	60.18	55.00
	Reverse	AGCTGAGGGAATGGATTGACA	59.08	47.62
GAPDH	Forward	GCTCTCCAGAACATCATCCCT	59.20	52.40
	Reverse	GGCAGGTTTTTCTAGACGGC	59.20	55.00

3.2.8 Polymerase chain reaction (PCR)

Polymerase Chain Reaction (PCR), a well-established molecular biology technique, was employed to generate millions of copies of the target DNA sequence. NucleoGene Hot Start PCR Master Mix was utilized following the manufacturer's protocol. Prior to analyzing patient samples, the PCR experiment was optimized. This optimization process involved adjustments to reaction conditions, and verification was performed using DNA isolated from the most distantly related member of the family pedigree. Subsequently, optimized annealing temperatures were selected as 59^0 C, and 2 μL of purified DNA from each family member were used in separate PCR reactions with the optimized thermal cycler conditions.

3.2.9 Agarose gel electrophoresis

Gel electrophoresis is the separation of DNA fragments of different sizes by causing the DNA fragments to move in an electric field. A gel is a material through which DNA fragments can pass. When PCR products are applied to the gel, they move through the gel with the help of an electric current. The DNA fragments move through the gel at different speeds depending on their size. As a result, gel electrophoresis can be used to visually determine whether the DNA fragments from the PCR reaction are the correct size and whether the desired DNA fragments are present.

Gel electrophoresis was done with 100 bp DNA Ladder Plus Marker and was used to verify whether the DNA samples resulting from PCR contained the desired DNA fragments and whether the PCR products were the correct size.

3.2.10 Bidirectional sanger sequencing

Bidirectional Sanger sequencing service was purchased from Macrogen Online Sequencing Order System through MSM company.

3.2.11 cDNA synthesis from isolated RNA

To generate complementary DNA (cDNA) from RNA, reverse transcription (RT) was performed using the Invitrogen SuperScript III First-Strand Synthesis System for RT-PCR Kit (Invitrogen) following the manufacturer's instructions. This process utilizes the enzyme reverse transcriptase to create cDNA copies complementary to the RNA template, requiring primers that anneal to the RNA for initiation. Importantly, optimization of the RT reaction conditions was conducted using experimental samples, not patient samples, to ensure optimal cDNA yield and minimize potential biases before proceeding with patient RNA.

3.2.12 Real-time quantitative PCR (RT-qPCR)

Following cDNA synthesis, real-time quantitative PCR (RT-qPCR) was employed to precisely quantify gene expression levels. This sensitive technique allows for the simultaneous amplification and quantification of a specific target DNA sequence during the PCR reaction. Fluorescent probes or dyes are incorporated into the reaction, and the fluorescence emitted is directly proportional to the amount of PCR product generated. This enables researchers to precisely measure the initial amount of target DNA present in the sample. By comparing gene expression levels between different samples, qPCR allows for the analysis of how gene expression varies under different experimental conditions or within distinct biological samples. This powerful approach provides valuable insights into gene regulation and cellular processes.

For gene expression analysis, individuals representing each genotype (wild-type, heterozygous, and homozygous) were selected based on RNA concentration values. High-quality RNA is essential for accurate gene expression analysis because contaminants can inhibit reverse transcription and PCR amplification, leading to inaccurate quantification of gene expression levels. Only samples with high-purity RNA, as assessed by spectrophotometry, were chosen for further analysis.

3.2.13 Statistical analysis

Following data acquisition, various open-source resources were utilized for analysis. ClinVar database, a public archive of human variation-phenotype associations with supporting evidence, was used to address several critical questions. These included whether a variant had been seen before, the levels of evidence supporting the association, and whether the variant was present in healthy individuals. By leveraging ClinVar, researchers can determine the clinical significance of genetic variants, assess the quality and quantity of evidence, and understand the broader context of variant pathogenicity.

Sequences obtained from the bidirectional Sanger sequencing were analyzed using multiple applications, including CLC Genomics Workbench, FinchTV, and CodonCode Aligner. Additionally, further analysis was performed on the PDF file provided by the sequencing company to cross-verify the results and ensure the reliability of the data.

For real-time qPCR analysis, data from the amplification table (cycle thresholds), melting curve and peak tables, number of available data points, and their standard deviations were used. The target gene, *ERLIN1*, was compared against the control gene, *GAPDH*. *GAPDH*, a housekeeping gene with relatively stable expression across various cell types, was chosen as a control to normalize potential variations in RNA input and sample processing.

GraphPad Prism software was employed for statistical analysis of the qPCR data. A one-way analysis of variance (ANOVA) was performed to determine if there were statistically significant differences in *ERLIN1* gene expression levels between the three genotype groups: wild-type, heterozygous, and homozygous.

Following the one-way ANOVA, multiple comparison tests were conducted to identify specific groups with statistically significant differences in expression. Tukey's Multiple Comparison Test was utilized. This test is generally preferred when

comparing all possible pairs of means within the data set. It assumes equal variances across all groups. Prism uses the Tukey-Kramer variation of this test, which accounts for unequal sample sizes if present.

Prism reports the results of both ANOVA and multiple comparison tests. The ANOVA provides an F-statistic and p-value. A significant F-statistic (typically with a p-value lower than 0.05) indicates that there is an overall difference in *ERLIN1* expression between at least two of the genotype groups.

The multiple comparison tests then identify which specific pairwise comparisons (comparisons between two groups at a time) are statistically significant. Prism typically reports adjusted p-values for each comparison, which account for the fact that multiple comparisons were performed. These adjusted p-values should be used to determine which differences in gene expression are statistically significant after considering all possible comparisons within the dataset.

4 RESULTS

This section presents the findings from the genetic analysis and gene expression studies conducted on the familial cohort. As part of a comprehensive familial study workflow shown in Figure 4.

4.1 Deep Phenotyping

The result of deep phenotyping is a categorization of clinical symptoms or diagnoses into 11 groups: immunological, blood or lymph, cardiovascular, hepatobiliary, endocrine, gastrointestinal, digestive, nephrological, respiratory, neurological, and musculoskeletal. Detailed information about the medical history or clinical findings observed in individuals is provided in Figure 6 and Table 2.

Figure 4. Extended family pedigree with deep-phenotyping

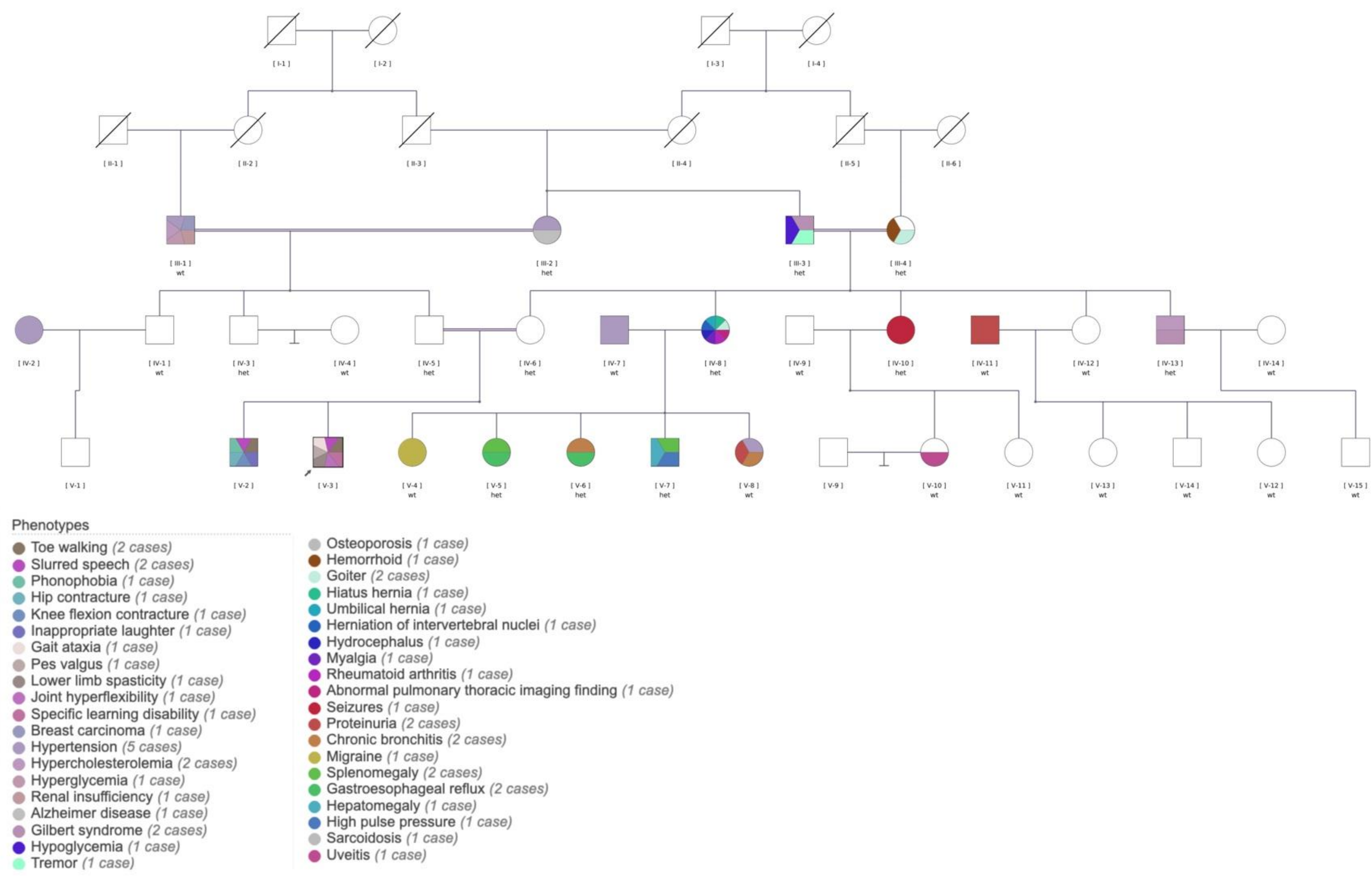


Table 2. Deep-Phenotyping with HPO terms

Patient's ID	Age	Affected Systems	Clinical Findings
III-1	78	Blood or lymph, Cardiovascular, Nephrological	Breast cancer, Hypertension, Hypercholesterolemia, Hyperglycemia, Renal insufficiency
III-2	77	Cardiovascular, Neurological	Hypertension, Alzheimer disease
III-3	66	Hepatobiliary, Endocrine, Neurological	Gilbert syndrome, Hypoglycemia, Tremor
III-4	67	Musculoskeletal, Digestive, Endocrine	Osteoporosis, Hemorrhoid, Goiter
IV-1	48	N/A	N/A
IV-2	42	Cardiovascular	Hypertension
IV-3	54	N/A	N/A
IV-5	39	N/A	N/A
IV-6	36	N/A	N/A
IV-7	51	Cardiovascular	Hypertension
IV-8	43	Endocrine, Musculoskeletal Respiratory, Gastrointestinal, Neurological	Goiter, Hiatus hernia, Umbilical hernia, Herniation of intervertebral nuclei, Hydrocephalus, Myalgia, Rheumatoid arthritis, Abnormal pulmonary thoracic imaging finding

Table 2. Deep-Phenotyping with HPO terms (Continued)

IV-9	46	N/A	N/A
IV-10	47	Neurological	Seizures
IV-11	44	Nephrological	Proteinuria
IV-12	39	N/A	N/A
IV-13	34	Hepatobiliary, Cardiovascular	Gilbert syndrome, Hypercholesterolemia
IV-14	34	N/A	N/A
V-2	8	Neurological, Musculoskeletal	Toe walking, Slurred speech, Phonophobia, Mild flexion contractures of hip and knee, Inappropriate laughter
V-3 (index)	15	Neurological, Musculoskeletal	Gait ataxia, Toe walking, Pes valgus, Lower limb spasticity, Slurred speech, Joint hyperflexibility, Mild specific learning disability
V-4	22	Immunological, Neurological	Migraine
V-5	21	Digestive, Gastrointestinal	Splenomegaly, Gastroesophageal reflux
V-6	6	Gastrointestinal, Respiratory	Gastroesophageal reflux, Chronic bronchitis

Table 2. Continue Deep-Phenotyping with HPO terms (Continued)

V-7	11	Blood or lymph, Digestive, Cardiovascular	Splenomegaly, Hepatomegaly, High pulse pressure
V-8	13	Cardiovascular, Nephrological, Respiratory	Hypertension, Proteinuria, Chronic bronchitis
V-10	27	Immunological, Respiratory	Sarcoidosis, uveitis
V-11	22	N/A	N/A
V-12	15	N/A	N/A
V-13	17	N/A	N/A
V-14	6	N/A	N/A
V-15	4	N/A	N/A

4.2 WBC, DNA and RNA Isolation

White blood cells (WBC), DNA, and RNA were isolated from each participant. The concentration and purity of the isolated DNA were measured with a Nanodrop spectrophotometer, and RNA was measured using a Nanodrop spectrophotometer and Qubit 4 Fluorometer (Table 3). Considering its compliance with the inclusion conditions sample IDs written in bold were chosen for downstream applications.

Table 3. Concentration values of isolated DNA and RNA samples

Patient's ID	DNA (ng/µL)	RNA (ng/µL)	Patient's ID	DNA (ng/µL)	RNA (ng/µL)	Patient's ID	DNA (ng/µL)	RNA (ng/µL)
III-1	41,1	17,1	IV-8	80,1	NA	V-5	136,5	NA
III-2	69,2	9,1	IV-9	133,1	82,4	V-6	79,3	NA
III-3	91,3	NA	IV-10	68,5	37,4	V-7	80,4	NA
III-4	68,2	NA	IV-11	110,6	NA	V-8	72,8	NA
IV-1	65,5	NA	IV-12	76,3	NA	V-10	105,4	17,0
IV-2	55,9	NA	IV-13	125,2	52,2	V-11	157,1	41,1
IV-3	87,6	24,6	IV-14	101,0	3,7	V-12	67,6	NA
IV-5	102,8	46,2	V-2	50,1	35,0	V-13	131,0	NA
IV-6	32,6	41,9	V-3	50,4	38,8	V-14	87,3	NA
IV-7	92,8	NA	V-4	107,7	NA	V-15	188,3	11,3

4.3 PCR and Agarose Gel Electrophoresis

Polymerase chain reaction (PCR) amplification was performed for each participant. PCR conditions were set according to the kit recommendations and optimization. The PCR products were then analyzed using agarose gel electrophoresis to confirm successful amplification. As a result of electrophoresis, each sample produces bands at the same position on an agarose gel, indicating that all the samples have amplified DNA fragments of the same size.

For Sanger sequencing following gel electrophoresis, samples with a single band between 137 bp were selected. Since all our samples met this criterion, samples from 30 individuals were sent for bidirectional Sanger sequencing.

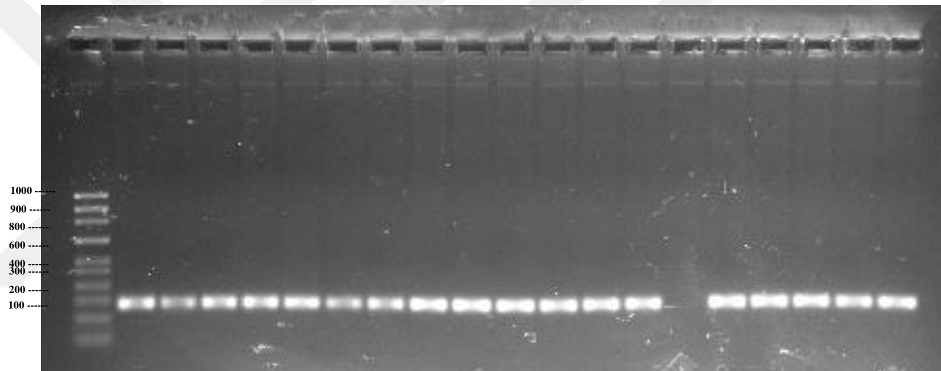


Figure 5. Agarose Gel Electrophoresis Image Part I

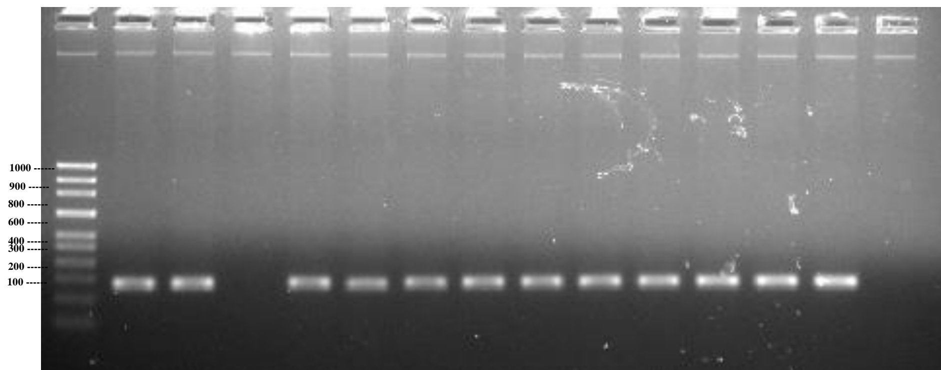


Figure 6. Agarose Gel Electrophoresis Image Part II

4.4 Bidirectional Sanger Sequencing

To identify variations in the *ERLIN1* gene, bidirectional Sanger sequencing was performed on 30 individual samples, targeting termination mutation in *ERLIN1* gene that was reported in the index case. All individuals are Sanger sequenced for the same genetic variation in the *ERLIN1* gene (NM_001100626.1), c.763C>T, p.(Arg255*), variation type is stop gain (rs876657413). ClinVar records numbered 1 with pathogenic interpretation (ClinVar Variation ID 226426).

The four individuals in the pedigree form a nuclear family. There are two brothers with the homozygous variant and parents with heterozygous status for the *ERLIN1* variant as illustrated in Figure 2. The results of Sanger sequencing conducted on 30 individuals within the extended family are as follows;

- 16 individuals were found to be wild-type,
- 12 individuals carrying c.763C>T termination variant in heterozygous state,
- 2 individuals had a homozygous variant in the *ERLIN1* gene (Table 4).

Table 4. Zygosity information of 30 family members

Zygosity					
Wild Type		*Heterozygote*		*Homozygote*	
Female	*Male*	*Female*	*Male*	*Female*	*Male*
9	*7*	*7*	*5*	*0 (zero)*	*2*

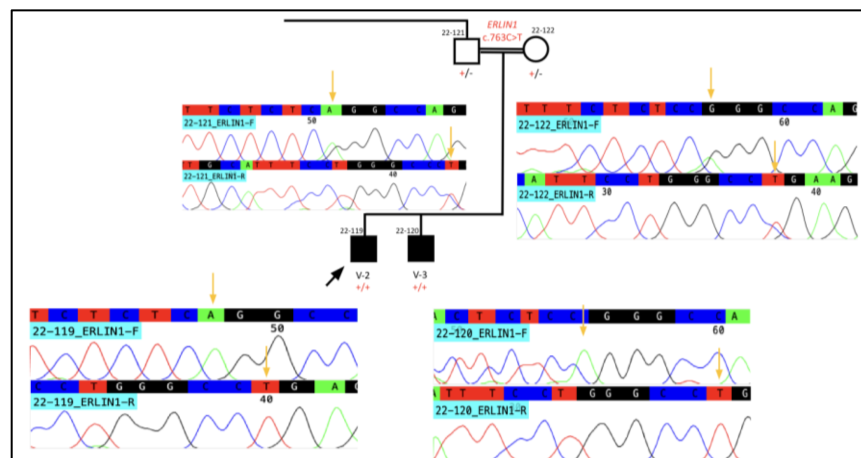


Figure 7. Confirmed variants of the nuclear family

Table 5. Confirmed zygosity info according to sanger results

Generation ID	Zygosity	Generation ID	Zygosity	Generation ID	Zygosity
III-1	wild type	IV-8	het	V-5	wild type
III-2	het	IV-9	wild type	V-6	het
III-3	het	IV-10	het	V-7	het
III-4	het	IV-11	wild type	V-8	het
IV-1	wild type	IV-12	wild type	V-10	wild type
IV-2	het	IV-13	het	V-11	wild type
IV-3	wild type	IV-14	wild type	V-12	wild type
IV-5	het	V-2	**hom**	V-13	wild type
IV-6	het	V-3	**hom**	V-14	wild type
IV-7	wild type	V-4	wild type	V-15	wild type

4.5 Gene Expression Analysis

For gene expression analysis, eight individuals representing each genotype (2 wild-type, 4 heterozygous, and 2 homozygous) were selected based on RNA purity values. Samples only with high-purity RNA, as assessed by spectrophotometry, were chosen for further analysis.

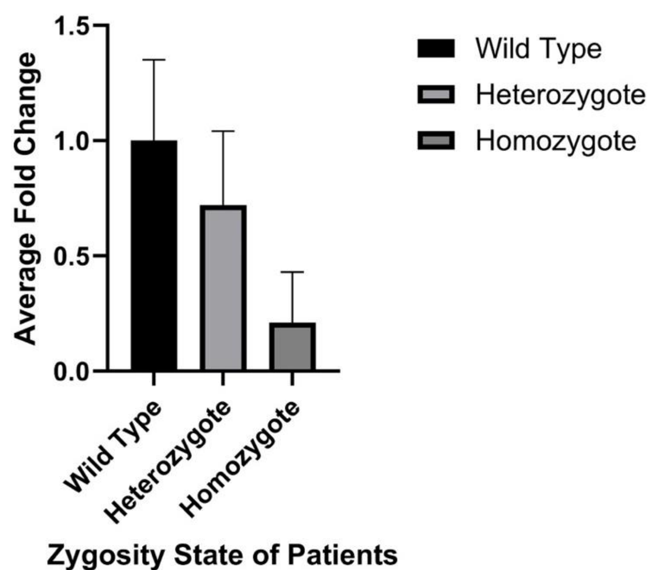


Figure 8. Comparison of ERLIN1 Gene Expression Levels Across Zygosity

In the qPCR results, the gene expression levels were normalized using the *GAPDH* housekeeping gene as a control. The average fold change in gene expression for the wild-type samples was set to 1.00. In comparison, the heterozygous state showed a reduced expression level with an average fold change of 0.72, while the homozygous state exhibited a further decreased expression level, with an average fold change of 0.21. The p-value obtained from the statistical analysis was 0.0785. Although this p-value did not reach statistical significance ($p > 0.05$), there was a clear trend of decreased expression levels in the heterozygous and homozygous groups.

5 DISCUSSION

Rare diseases are described as conditions that are seen less than 1/2000. On the other hand, the term “rare” is not fair for a group of disorders that are affecting more than 300 million people across the world. The visibility and awareness are increasing, paving the way for the next steps in rare and undiagnosed diseases: diagnosis and treatment options. Therefore, the priority is to reach as many individuals with rare diseases as possible, obtain detailed phenotypic and genotypic information about the cases and their families, and report variants associated with the diseases. These reported variants should then be supported by functional studies at the cellular level and animal models. Finally, the molecular mechanisms should be elucidated, and therapeutic developments should be achieved to have the chance to reach individuals.

In this thesis study, we aimed to elucidate the role of the *ERLIN1* gene in the pathophysiology of HSP62, to show similarities between HSP62 and juvenile or early-onset ALS, to expand the clinical spectrum associated with *ERLIN1* through deep phenotyping, and to provide evidence of genotype-phenotype effects by comparing gene expression levels in patient samples. We also explored the potential association with neurological phenotypes and aimed to identify differences in *ERLIN1* expression in the presence of monoallelic and biallelic variants, as well as in wild-type healthy individuals.

One of the most intriguing outputs of this study is the deep phenotyping approach in rare and undiagnosed cases. Here the index case has been thoroughly assessed by a single pediatric geneticist. The beginning of our research story dates back to 2017 when a seven-year-old boy presented to the Department of Pediatric Genetics at Acıbadem University School of Medicine with complaints of toe walking, unsteady gait, foot deformity, and speech disorder. Our index patient, referred by a neurologist, underwent a one-hour clinical examination during the initial consultation. During this time, deep phenotyping was performed, and seven different HPO terms were identified. Genetic counseling was provided to the family, an extended family pedigree was obtained, and testing options were discussed. After the family decided to proceed with genetic testing, they were directed to diagnostic laboratory for WES analysis. The

results revealed a homozygous *ERLIN1* (NM_001100626.1), c.763C>T (p.Arg255*) termination variant. Through deep phenotyping and HPO terms, this variant was accepted as consistent with the genetic diagnosis of spastic paraplegia 62, and the interpretation of this variant was classified as 'pathogenic' (class 1) according to ACMG recommendations. Following the receipt of the results, a literature review was conducted, and the variant and diagnostic information considered relevant were communicated to the family.

In December 2021, a reclassification report from the genetic diagnosis center was provided to the clinician. The statement in the reclassified report was as follows: 'Based on new guidelines of ACMG for variant interpretation and new evidence becoming available, the homozygous *ERLIN1* (NM_001100626.1), c.763C>T (p.Arg255*) variant has been reclassified from 'pathogenic' (class 1) to 'variant of uncertain significance' (VUS) (class 3). Following this, the clinician and researcher, curious about the condition of the index case and his sibling, invited the family for an appointment. In February 2022, a deep phenotyping and follow-up examination were performed on the brothers who came for the check-up. In our 12-year-old male index case (V-3), in addition to the previous mild specific learning disabilities were observed. In the 6-year-old male sibling (V-2), subsequently, with questions about whether this variant was truly pathogenic.

In the literature, there are six variants reported in the *ERLIN1* gene so far: c.763C>T, c.149G>T, and c.862_868delACCAGG reported by Novarino et al. in 2014; c.281T>C reported by Tunca et al. in 2018; c.504+1G>A reported by Zhu et al. in 2022; and finally, the c.195G>C variant reported by Kilic et al in 2024. We determined the *ERLIN1* (NM_001100626.1), c.763C>T (p.Arg255*) variant which causes premature stop codon. This mutation causes translation to stop before the protein is fully synthesized, producing a truncated protein that is typically non-functional. Such truncated proteins may lack essential domains necessary for their normal function, potentially resulting in loss of function or gain of deleterious properties that contribute to various diseases. In our situation, this variant is likely to produce a non-functional protein associated with the clinical findings of HSP62.

Patients with *ERLIN1* mutations can exhibit a range of symptoms from mild to severe, including early-onset spasticity and muscle weakness. Novarino et al reported different variant types such as missense, frameshift, and stop codon mutations, each correlating with varying degrees of disease severity. The cases with c.763C>T variant, exhibited HSP62 linked to the homozygous *ERLIN1* gene. In contrast, the missense variant c.149G>T shows a moderate phenotype. Lastly, the frameshift variant c.862_868delACCAGG was observed in family members with the mildest findings. These included tiptoe walking, walking alone with an unsteady gait (post-operation), walking alone with a mild scissors gait, plantar reflex, vibration sense at the ankle, and normal cognition. Comparing the first complaints seen in reported *ERLIN1*-related HSP62 patients; Novarino et al. reported leg pain, gait disturbance, bilateral pes equinovarus, tiptoe walking, and lower limb spasticity. Our patient's phenotype is similar to the findings reported in Novarino et al.'s paper, with symptoms such as spasticity, increased deep tendon reflexes (DTR), and vibration sense at the ankles. However, our patient does not exhibit thoracic kyphosis, amyotrophy, or the ability to walk independently. On the other hand, the index's younger brother who is also homozygous for the same variant presented the symptoms more severe like toe walking, slurred speech, phonophobia, mild flexion contractures of the hip and knee, and inappropriate laughter.

In another study Tunca et al., reported the homozygous c.281T>C variant that was transmitted across multiple generations, with different severities and symptoms like; walking difficulties and swallowing problems, or mild limp that needed surgical intervention, or abnormal gait, increased deep tendon reflexes (DTR), muscle atrophy, and fibrillation and fasciculation. These findings are notably similar to the *ERLIN1*-related HSP62 disease phenotype, suggesting a potential linkage between ALS and HSP62 as mentioned by Tunca et al. However, our index case's brother differentiates from this clinical spectrum with no swallowing problems and additional findings such as phonophobia, slurred speech, contractures of the hip and knee, and inappropriate laughter.

In 2022, Zhu et al. reported a c.504+1G>A splicing site mutation in a homozygous state in a 23-year-old male, with gait disturbance, reduced muscle strength, elevated muscle tone in the lower limbs, no Babinski sign, and an inability to walk without support. Other family members, apart from the index case, showed no clinical symptoms and were unaffected. He had no difficulty drinking water or swallowing. The clinical symptoms of this index case are similar to our cases V-2 and V-3.

In a very recent study, Kilic et al. reported an 11-year-old male with leg spasticity and weakness due to the homozygous c.195G>C variant. The patient exhibited tiptoe walking and spasticity in the lower limbs, with no cerebellar signs and normal cognition. It was confirmed that his parents are carriers of the variant in a heterozygous state, but no clinical findings or further information were provided in the manuscript. Due to the lack of detailed clinical findings described in this article, it is not possible to make a definite conclusion or comparison. If we only compare our nuclear family, the symptoms observed in our V-2 and V-3 cases appear to be much more severe.

In our study, the male individual V-3 exhibits similar symptoms to other cases, including unsteady gait or gait disturbance, spasticity in the lower limbs, and tiptoe walking, with no cerebellar signs. Since we began monitoring the index case in 2017, we have noticed a gradual increase in the severity of the disease. It is slowly, but definitely progressive. Interestingly, the younger brother, V-2, is much more severe than the index case. While V-3 showed mild symptoms at the age of 7, his 6-year-old brother V-2 already has tiptoe walking, slurred speech, phonophobia, mild flexion contractures of the hip and knee, inappropriate laughter, and inability to walk without support. Slurred speech, phonophobia, mild flexion contractures of the hip and knee, and inappropriate laughter symptoms are reported for the first time. This situation raises “Why does the disease present more severely in V-2, who was born later?” questions in our minds. Since the clinician observed neurological symptoms in other family members, we expected to identify another homozygous individual within the family. Looking back 10 years ago, and considering other neurological and neurodevelopmental-related disorders, it is evident that whether the variant is in the heterozygous or homozygous state, both can cause symptoms and affect individuals

with varying severity. Using the bidirectional Sanger sequencing data we obtained, we compared each family within the extended family based on zygosity. Among the three generations for which we had biological samples, only our index case and his brother were homozygous for the variant.

In 2016, Xie et al. conducted a study investigating the likelihood of having a child with autism. They observed that families with one autism spectrum disorder (ASD) child have nearly 14 times the likelihood of having a second child with ASD. Factors mentioned include the advancing age of the parents and the increased age at childbirth. It was also noted that if the older autistic child is male, the likelihood of the next child being born with autism is higher. It is significant that both affected individuals observed in our family are male children. This could similarly be considered an increased likelihood factor, as this condition may not manifest in older generations or may not have been randomly encountered among carriers. Additionally, similar to the study, there is a 7-year age gap between siblings in our nuclear family. Thus, the advancing age of the parents and increased age at childbirth could also be considered contributing factors.

ERLIN1 gene expression in humans is primarily found in various tissues, with significant levels in the brain and spinal cord. On the other hand, *ERLIN1* was shown to be expressed in all types of blood cells, the highest being eosinophils (https://v19.proteinatlas.org/ENSG00000107566-*ERLIN1*/blood). There are also studies showing the expression level changes of *ERLIN1* under different conditions, like sepsis (50). Since this is a termination variant, we also wanted to assess differences of gene expression levels in cases with different genotypes and if there is any correlation between the gene expression levels and clinical presentations. The results indicated that the gene expression level observed in wild-type individuals, who do not have the targeted mutation in the *ERLIN1* gene, is the highest. As expected, this is followed by heterozygous carrier individuals and the lowest expression is observed in homozygous individuals with this variant. According to our study, gene expression levels tend to decrease across different genotypes, but we did not have enough samples to provide statistical evidence for this. Similar situations are often encountered in

studies conducted in the field of rare and undiagnosed diseases due to the difficulty of finding affected individuals. Additionally, it is a known fact that the mRNA expression levels can not always represent the protein expression levels but we barely had materials to isolate DNA and RNA from the family members so we did not have chance to perform protein analysis.

Throughout this research process, we encountered some limitations. Firstly, the blood samples were collected outside of Istanbul and brought to our laboratory where the isolations were conducted. During this time, we used our available resources, and for safety in the village environment, the father of our index patient, voluntarily traveled to Artvin to collect the blood samples. Unfortunately, despite our efforts, the samples were not transported in a stable environment during transfer. Therefore, some of the samples in the RNA isolations had low purity and concentration levels that were not suitable for downstream applications. This extended handling time and potential exposure to varying conditions during transport may have further compromised the integrity of the samples. This condition limited the number of cases for expression analysis. Hence, although we detected a decrease between wild type, heterozygous and homozygous genotypes, the *p value* was limited to 0.07. Secondly, we did not have enough material to perform protein analysis, which would have been insightful for the study. Despite these limitations, the study provides valuable insights, but future research would benefit from more controlled and expedited sample handling procedures to ensure higher quality genetic material for analysis.

The zygosity of diseases, which can significantly affect the efficacy and safety of treatments, is increasingly being taken into account in recent advances in drug discovery. Technologies such as CRISPR-Cas9 enable precise gene editing tailored to whether a patient is homozygous or heterozygous for a particular mutation, improving therapeutic outcomes. In the field of rare and undiagnosed diseases, understanding zygosity is essential for developing effective gene therapies, which may require different dosages or delivery methods depending on the genetic context. These healthcare solutions should be followed up with ongoing clinical monitoring and genetic counseling to assess the impact of treatments and address any unexpected

outcomes. Such a comprehensive approach ensures that both the therapeutic efficacy and patient safety are continuously evaluated. At this point, it is crucial to increase the number of genetic counselors and enhance the frequency of genetic testing for healthy individuals. This will support and facilitate advancements in personalized medicine.

To further underscore the importance of genetic counseling and testing, it is important to recognize the options they generate for every stage of life according to requirement of the individual or family. Genetic testing generates options for early diagnosis and informed decision-making. It may guide personalized treatment plans, identify family members who are at risk, and provide peace of mind by alleviating suspicions about actions taken during pregnancy or later on. In the end, it primarily focuses on three questions:

1. Why did this happen?
2. Is it going to occur again?
3. What can the patient and their family do?

This is also important in understanding the genetic underpinnings of neurological disorders and highlight the significance of genetic testing and counseling. For instance, HSP62 and its potential connections to other neurodegenerative diseases, such as ALS, exemplify how genetic insights can lead to better disease management and therapeutic strategies. Research has shown that there may be overlapping genetic pathways and mechanisms between HSP62 and ALS, providing a foundation for future investigations and potential treatments. From a broader perspective, while there are many advantages to the use of genetic testing, the potential disadvantages can be limited or minimized through the use of genetic counseling and skilled genetic counselors. Additionally, advancements in new technologies, the adaptation of families, and the deep phenotyping by clinicians can decrease the time spent without a diagnosis and shorten the diagnostic odyssey. Here, we observed the importance of careful literature review and patient follow-up, especially since the reclassification of this variant came up as VUS (Variant of Uncertain Significance). Therefore, appropriate genetic counseling must be given to the extended family to ensure they are well-informed about the implications of these findings. This finding emphasizes the

critical role of genetic testing and counseling in diagnosing and managing neurodegenerative diseases

With this thesis project, we were able to highlight;

- the importance of deep phenotyping
- the importance of extended family screening
- the importance of continuous follow up, reanalysis and reevaluation of the cases and the variants
- and the detailed and informative genetic counseling in the support of the patients, their family members.

6 CONCLUSION

In this thesis, we explored the complex relationship between gene expression levels and various type of zygosities, exemplifying the importance of understanding the genotype-phenotype relationship with a focus on the neurodegenerative disease Hereditary Spastic Paraplegia 62 (HSP62). Our study expanded the spectrum of phenotypic information related to *ERLIN1*-associated HSP62 through deep phenotyping of a large family of 30 individuals. These insights highlight the importance of genetic testing and counseling in modern healthcare. Bidirectional Sanger sequencing revealed that these phenotypic variations could occur even in individuals who are carriers or non-carriers within the family, potentially due to lack of penetrance. To confirm these observations, further functional studies and Western blot experiments should be conducted after protein isolation to examine the effects on protein size. Our study underscores the need for further investigation into the mechanisms of HSP62, particularly in comparison with ALS mechanisms, given the variability in disease courses among different variants. Understanding these mechanisms may provide insights into potential therapeutic approaches for both HSP62 and ALS.

7 REFERENCES

1. Galas DJ, McCormack SJ. An historical perspective on genomic technologies. Curr Issues Mol Biol. 2003;5(4):123-7.
2. Cunha A. Genomic technologies-from tools to therapies. Genome Med. 2017;9(1):71.
3. Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. Genome Med. 2022;14(1):23.
4. Mahdieh N, Rabbani B. An overview of mutation detection methods in genetic disorders. Iran J Pediatr. 2013;23(4):375-88.
5. Kim J, Kim JJ, Lee H. An analysis of disease-gene relationship from Medline abstracts by DigSee. Sci Rep. 2017;7:40154.
6. Luo P, Xiao Q, Wei PJ, Liao B, Wu FX. Identifying Disease-Gene Associations With Graph-Regularized Manifold Learning. Front Genet. 2019;10:270.
7. Reisel D, Baran C, Manchanda R. Preventive population genomics: The model of BRCA related cancers. Adv Genet. 2021;108:1-33.
8. Human Genomic Variation - Fact Sheets about Genomics: National Human Genome Research Institute; [Available from: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genomic-variation.
9. Human Genome Project Fact Sheet National Human Genome Research Institute [updated August 24, 2022. Available from: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project.
10. Hood L, Rowen L. The Human Genome Project: big science transforms biology and medicine. Genome Med. 2013;5(9):79.
11. Wise AL, Manolio TA, Mensah GA, Peterson JF, Roden DM, Tamburro C, et al. Genomic medicine for undiagnosed diseases. Lancet. 2019;394(10197):533-40.
12. Genomics & Precision Health; Genetic Testing: Centers for Disease Control and Prevention; [updated June 24, 2022. Available from: https://www.cdc.gov/genomics/gtesting/genetic_testing.htm.
13. Introduction to Genomics National Human Genome Research İnstitute [updated October 11, 2019. Available from: https://www.genome.gov/About-Genomics/Introduction-to-Genomics.
14. Understanding Genetics: A New York, Mid-Atlantic Guide for Patients and Health Professional. Genetic Alliance Monographs and Guides. Washington (DC)2009.
15. Jelin AC, Sagaser KG, Wilkins-Haug L. Prenatal Genetic Testing Options. Pediatr Clin North Am. 2019;66(2):281-93.
16. Fonda Allen J, Stoll K, Bernhardt BA. Pre- and post-test genetic counseling for chromosomal and Mendelian disorders. Semin Perinatol. 2016;40(1):44-55.
17. Hagemann IS. Clinical Genomics. In: Pfeifer SKaJ, editor.2015.
18. Koboldt DC. Best practices for variant calling in clinical sequencing. Genome Med. 2020;12(1):91.
19. Satam H, Joshi K, Mangrolia U, Waghoo S, Zaidi G, Rawool S, et al. Next-Generation Sequencing Technology: Current Trends and Advancements. Biology (Basel). 2023;12(7).
20. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: overviews and challenges. Biotechniques. 2014;56(2):61-4, 6, 8, passim.
21. Gorcenco S, Ilinca A, Almasoudi W, Kafantari E, Lindgren AG, Puschmann A. New generation genetic testing entering the clinic. Parkinsonism Relat Disord. 2020;73:72-84.
22. Jalkh N, Corbani S, Haidar Z, Hamdan N, Farah E, Abou Ghoch J, et al. The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population. BMC Med Genomics. 2019;12(1):11.
23. Akgun-Dogan O, Tuc Bengur E, Ay B, Ozkose GS, Kar E, Bengur FB, et al. Impact of deep phenotyping: high diagnostic yield in a diverse pediatric population of 172 patients through clinical whole-genome sequencing at a single center. Front Genet. 2024;15:1347474.
24. National Society of Genetic Counselors' Definition Task F, Resta R, Biesecker BB, Bennett RL, Blum S, Hahn SE, et al. A new definition of Genetic Counseling: National Society of Genetic Counselors' Task Force report. J Genet Couns. 2006;15(2):77-83.
25. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of

Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.
26. Innovation ECD-GfRa. European Research on Rare Diseases: European Union; [Available from: https://research-and-innovation.ec.europa.eu/research-area/health/rare-diseases_en.
27. Disorders NOfR. Rare Disease Day: Frequently Asked Questions 2019 [Available from: https://rarediseases.org/wp-content/uploads/2019/01/RDD-FAQ-2019.pdf.
28. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet. 2020;28(2):165-73.
29. In: Field MJ, Boat TF, editors. Rare Diseases and Orphan Products: Accelerating Research and Development. The National Academies Collection: Reports funded by National Institutes of Health. Washington (DC)2010.
30. WHO. Rare diseases: World Health Organisation; [Available from: https://www.who.int/standards/classifications/frequently-asked-questions/rare-diseases.
31. TÜSEB. Rare Disease Report: Türkiye Public Health and Chronic Diseases Institute; 2019 [Available from: https://files.tuseb.gov.tr/tuhke/files/haberler/nadir-hastaliklar-raporu.pdf.
32. Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: translating evidence into policy. Lancet Neurol. 2020;19(3):255-65.
33. Schuermans N, Verdin H, Ghijsels J, Hellemans M, Debackere E, Bogaert E, et al. Exome Sequencing and Multigene Panel Testing in 1,411 Patients With Adult-Onset Neurologic Disorders. Neurol Genet. 2023;9(3):e200071.
34. Lamptey RNL, Chaulagain B, Trivedi R, Gothwal A, Layek B, Singh J. A Review of the Common Neurodegenerative Disorders: Current Therapeutic Approaches and the Potential Role of Nanotherapeutics. Int J Mol Sci. 2022;23(3).
35. Chen S, Sayana P, Zhang X, Le W. Genetics of amyotrophic lateral sclerosis: an update. Mol Neurodegener. 2013;8:28.
36. Lehky T, Grunseich C. Juvenile Amyotrophic Lateral Sclerosis: A Review. Genes (Basel). 2021;12(12).
37. Zarei S, Carr K, Reiley L, Diaz K, Guerra O, Altamirano PF, et al. A comprehensive review of amyotrophic lateral sclerosis. Surg Neurol Int. 2015;6:171.
38. Chia R, Chio A, Traynor BJ. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. Lancet Neurol. 2018;17(1):94-102.
39. Blackstone C. Hereditary spastic paraplegia. Handb Clin Neurol. 2018;148:633-52.
40. Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. MIM Number: {615681}: {10/12/2016} [Available from: https://omim.org/.
41. Rubinstein WS, Maglott DR, Lee JM, Kattman BL, Malheiro AJ, Ovetsky M, et al. The NIH genetic testing registry: a new, centralized database of genetic tests to enable access to comprehensive information and improve transparency. Nucleic Acids Res. 2013;41(Database issue):D925-35.
42. Orphanet: an online rare disease and orphan drug data base: INSERM; 1999 [Available from: http://www.orpha.net.
43. Schwarz DS, Blower MD. The endoplasmic reticulum: structure, function and response to cellular signaling. Cell Mol Life Sci. 2016;73(1):79-94.
44. Meyyazhagan A, Orlacchio A. Hereditary Spastic Paraplegia: An Update. Int J Mol Sci. 2022;23(3).
45. Created with [cited 2024. Available from: BioRender.com.
46. Novarino G, Fenstermaker AG, Zaki MS, Hofree M, Silhavy JL, Heiberg AD, et al. Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. Science. 2014;343(6170):506-11.
47. Tunca C, Akçimen F, Coşkun C, Gündoğdu-Eken A, Kocoglu C, Çevik B, et al. ERLIN1 mutations cause teenage-onset slowly progressive ALS in a large Turkish pedigree. Eur J Hum Genet. 2018;26(5):745-8.
48. Zhu ZY, Li ZY, Zhang C, Liu XL, Tian WT, Cao L. A novel homozygous mutation in ERLIN1 gene causing spastic paraplegia 62 and literature review. Eur J Med Genet. 2022;65(11):104608.
49. Kilic MA, Yildiz EP, Deniz A, Coskun O, Kurekci F, Avci R, et al. A Retrospective Review of 18 Patients With Childhood-Onset Hereditary Spastic Paraplegia, Nine With Novel Variants. Pediatr Neurol. 2024;152:189-95.

50. Huang SSY, Toufiq M, Saraiva LR, Van Panhuys N, Chaussabel D, Garand M. Transcriptome and Literature Mining Highlight the Differential Expression of ERLIN1 in Immune Cells during Sepsis. Biology (Basel). 2021;10(8).

8 APPENDIX

APPENDIX 1

The publication of the study is currently being written. It took place as a poster presentation at Course in Genetic Counselling, Bertinoro, Italy on November 12-17, 2023.

APPENDIX 2

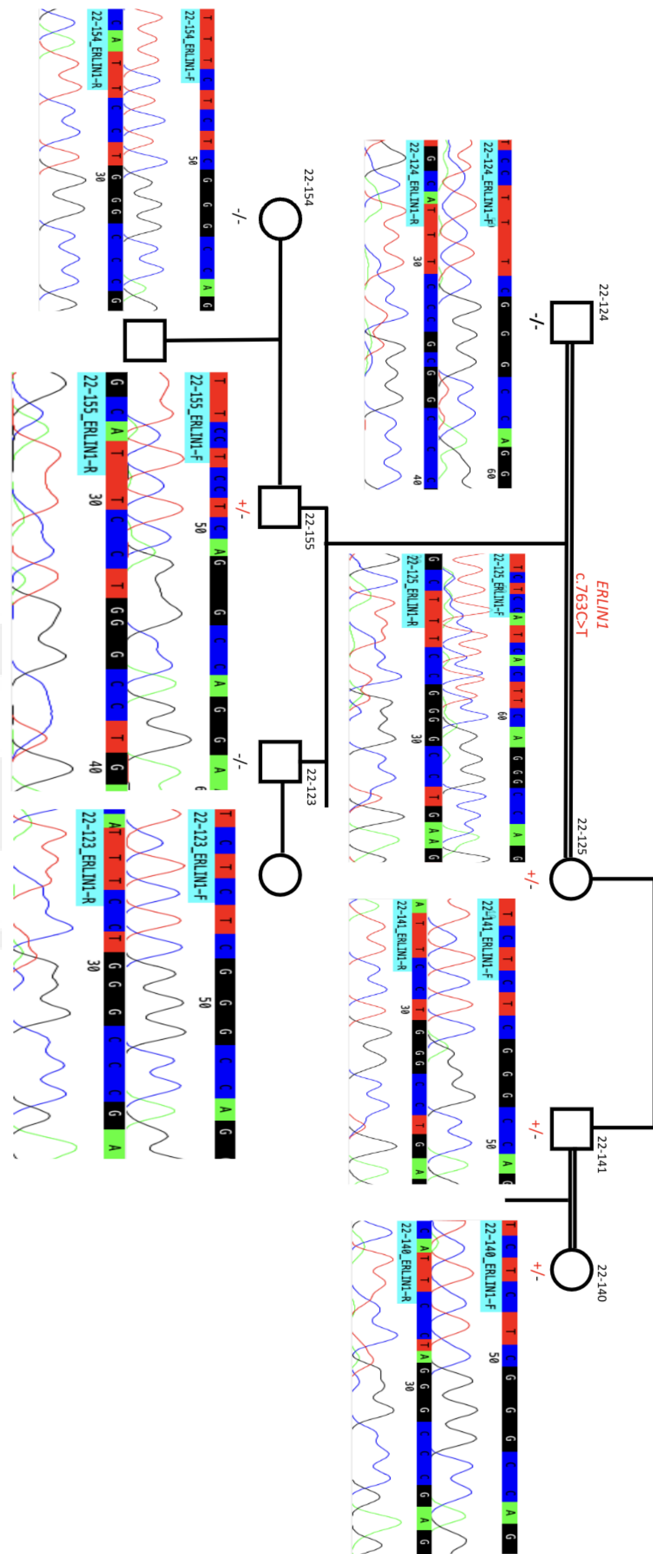
Ethic committee decision

APPENDIX 2

Ethic committee decision (Continued)

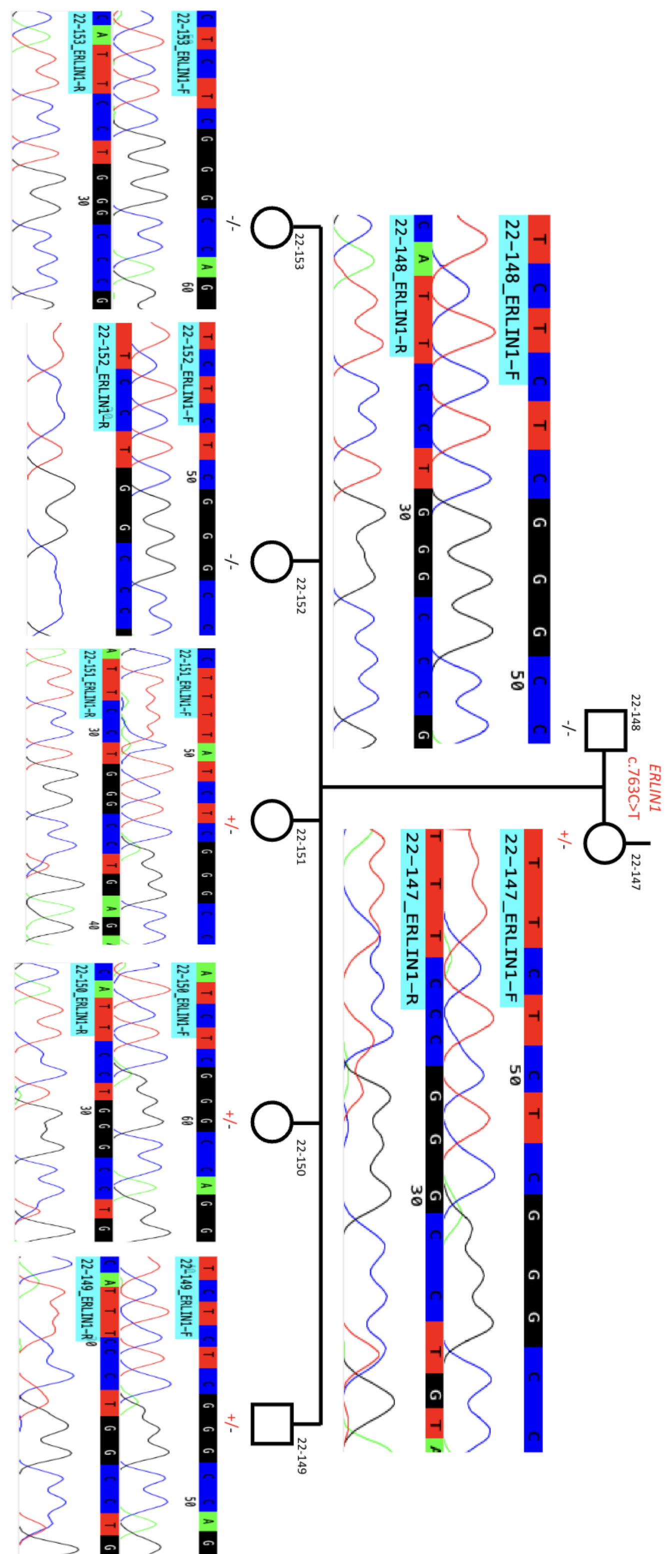
APPENDIX 3

Bidirectional Sanger Sequencing Results of Extended Family



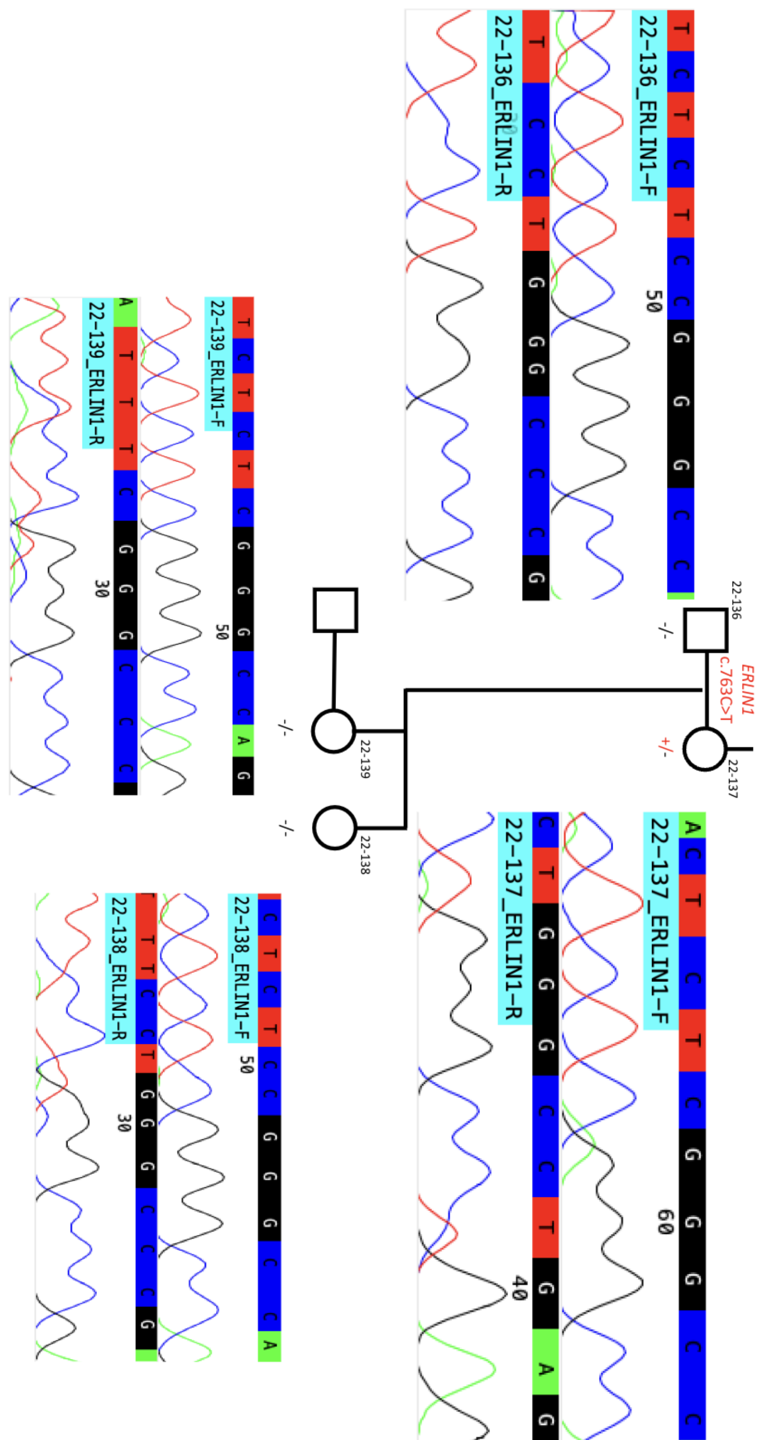
APPENDIX 3

Bidirectional Sanger Sequencing Results of Extended Family (Continued)



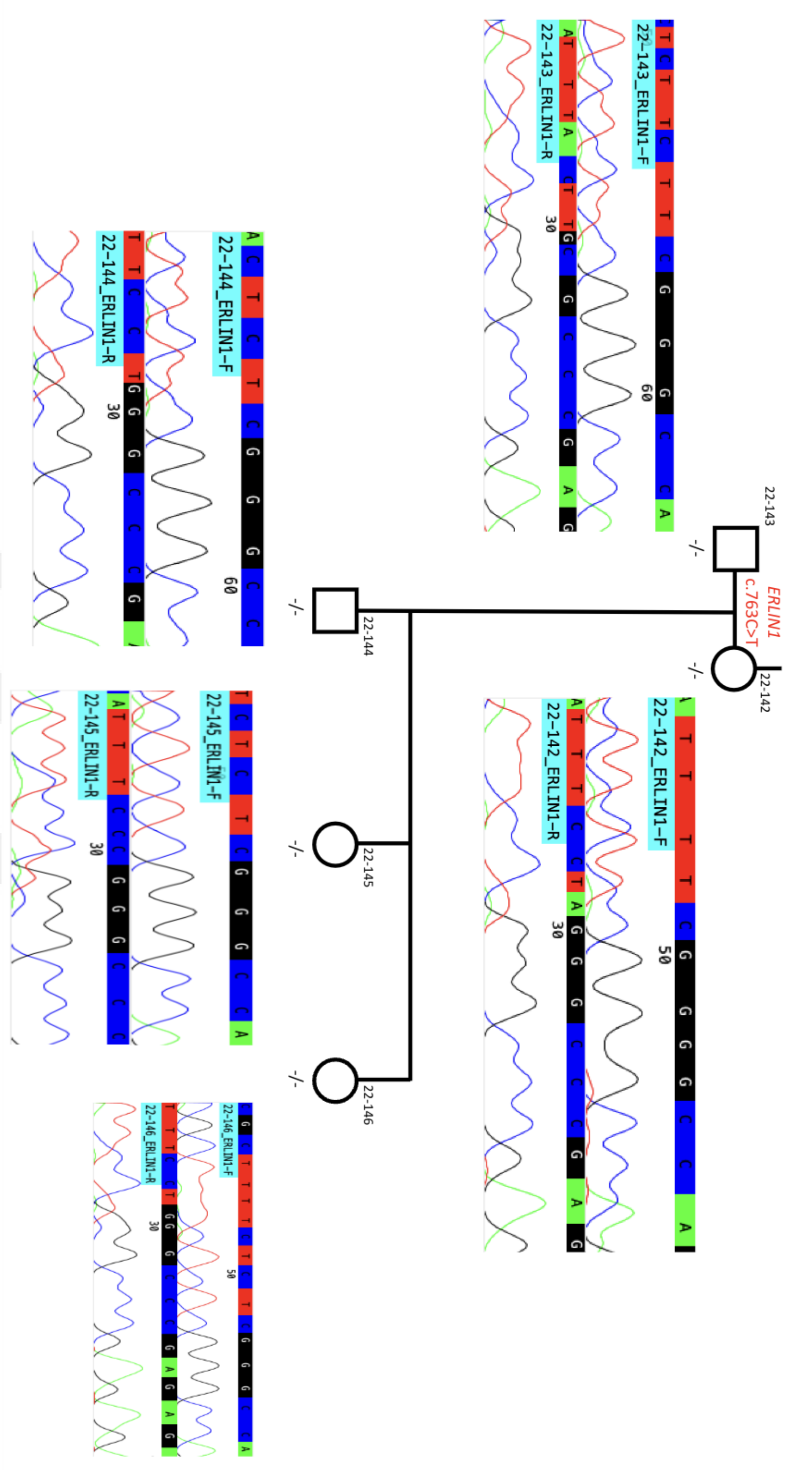
APPENDIX 3

Bidirectional Sanger Sequencing Results of Extended Family (Continued)



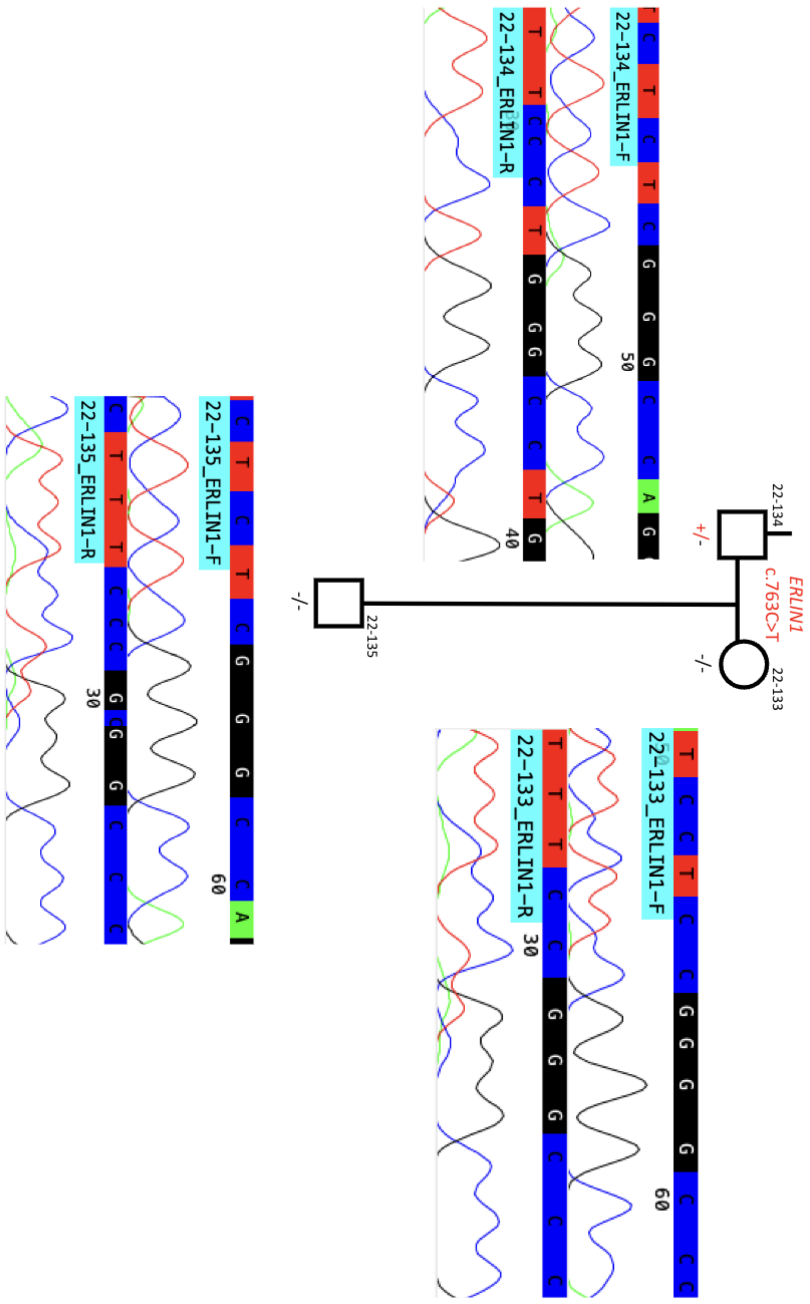
APPENDIX 3

Bidirectional Sanger Sequencing Results of Extended Family (Continued)



APPENDIX 3

Bidirectional Sanger Sequencing Results of Extended Family (Continued)


ERLIN1
c.763C>T
22-134
+/-
22-133
-/-
22-135
-/-
22-134_ERLIN1-F
22-134_ERLIN1-R
22-133_ERLIN1-F
22-133_ERLIN1-R
22-135_ERLIN1-F
22-135_ERLIN1-R

9 CURRICULUM VITAE