FULL LENGTH | Iranian Biomedical Journal 28 (5 & 6): 297-304 September and November 2024

Mutations in *COL6A* Gene Family Responsible for Muscular Dystrophies in Three Unrelated Families

Nasibeh Soltani[1,2], Zahra Shahbazi[1], Morteza Karimipoor[1], Mohammad Sadegh Fallah[2], Fatemeh Zafarghandi Motlagh[2], Masoume Amini[2], Mojdeh Jamali[2], Hamideh Bagherian[2], Razie Zeinali[2], Sirous Zeinali[2*]

[1]Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran; [2]Department of Human Genetics, Kawsar Human Genetics Research Center, Tehran, Iran

OPEN ACCESS

Received: September 30, 2023
Accepted: November 27, 2023
Published online: December 3, 2023

Citation:
Zeinali S, Soltani N, Shahbazi Z, Karimipoor M, Fallah MS, ZafarghandiMotlagh F, Amini M, Jamali M, Bagherian H, Zeinali R. Mutations in *COL6A* Gene Family Responsible for Muscular Dystrophies in Three Unrelated Families. *Iranian biomedical journal* 2024; 28(5&6): 297-304.

ABSTRACT

Background: Muscular dystrophy is an inherited disease with clinical and genetic heterogeneity. Muscle weakness is the primary symptom of these disorders that often leads to disability and death. The overall prevalence for all types of muscular dystrophies worldwide is 19.8-25.1 per 100,000 population. Autosomal recessive types of muscular dystrophies are more common in Iran, likely due to the high rate of consanguineous marriage. We aimed at deciphering molecular defects in three unrelated families with muscular dystrophies not related to Duchene MD or limb girdle muscular dystrophies. We are reporting families having affected children with MD owing to the mutations in three genes related to the *COL6A* (collagen type VI, alpha subunit) gene family.

Methods: Three unrelated families, who had at least one member affected with MD and for whom a definite molecular diagnosis was not provided by routine methods, were investigated by WES and confirmed by Sanger sequencing.

Results: In the first family, a homozygous variant was found in the *COL6A3* gene (NM_004369.4:c.4390C>T:p.Arg1464Ter), which explains the clinical symptoms observed in this family. In the second family, two homozygote missense variants with possible relevance to the patient's phenotype were identified in *COL6A1* and *COL6A2* genes (NM_001848.2:c.803A>G: p.Glu268Gly and NM_001849.3:c.2489G>A:p.Arg830Gln). Also, a heterozygous pathogenic variant in the *COL6A2* gene (NM_001849.3: c.1053+1G>T) was detected in the third family.

Conclusion: WES can serve as an effective method for detecting the causative mutations in families with unresolved cases of MD. The data provided herein broadens the spectrum of mutations causing MD in Iran.
DOI: 10.61186/ibj.4018

Keywords: Collagen Type VI, Exome sequencing, Muscular dystrophies

Corresponding Author: Sirous Zeinali
Kawsar Human Genetics Research Center, Tehran, Iran; E-mail: zeinali@gmail.com

INTRODUCTION

Muscular dystrophies are a heterogeneous group of genetically inherited degenerative disorders of the muscle, characterized by varying degrees of progressive muscle weakness and wasting, from partial mobility impairment to severe complications, affecting the quality of life and life expectancy[1-5]. The overall prevalence for all types of muscular dystrophies worldwide is 19.8-25.1 per 100,000 population[6]. According to the structure and

List of Abbreviations:
ACMG: American College of Medical Genetics and Genomics; **AMP:** Association for Molecular Pathology; **EMG:** electromyography; **MD:** muscular dystrophy; **MLPA:** multiplex ligation-dependent probe amplification; **NCV:** nerve conduction velocity; **SNV:** Single nucleotide variation; **UCMD:** Ullrich congenital muscular dystrophy ; **VUS:** variant of unknown significance

cellular/molecular basis of skeletal muscle cells, several reasons can cause MDs, including disruption of cytoskeleton-extracellular matrix connection in Duchenne/Becker MD and recessively inherited limb-girdle muscular dystrophies, aberrant glycosylation of α-dystroglycan in a heterogeneous group of congenital muscular dystrophies, disruption of the extracellular matrix in UCMD, and Bethlem myopathy results from defective collagen VI (*COL6A1*, *COL6A2*, and *COL6A3*), defective sarcolemma repair in dysferlinopathies or Limb gridle MD 2B LGMD2B)[7].

Elevated levels of creatine kinase or creatine phosphokinase, an intracellular enzyme with high expression in the skeletal muscle, can precede other symptoms of MD and are usually used as a primary diagnostic biomarker for all types of MDs[8]. MD patients with identical genetic variants may have variations in the age of onset, as well as the severity, progression, and prognosis of the disorder[1,4,5,9]. Due to these overlapping clinical presentations and the wide range of phenotypes in patients with MDs, establishing an accurate diagnosis is challenging. Hence, it is essential to design a reliable prevention plan by providing genetic counseling, prenatal diagnosis, or early treatment, if available[1,9].

Most conventional molecular methods of diagnosing heterogeneous genetic diseases, including MDs, involve very expensive and time-consuming procedures. Additionally, despite progressive advances in clinical diagnosis of these disorders, molecular causes remain unknown in a large number of muscular dystrophies. WES, which focuses on sequencing protein-encoding regions of the human genome, provides a high-throughput, efficient and rapid method for deciphering the molecular basis of the disease, and in some cases, it is the only strategy to identify the gene or mutation that causes particular inherited disorder[10,11]. The present study aimed to identify the genetic causes of undiagnosed muscular dystrophies in three Iranian families using WES, followed by Sanger sequencing.

MATERIALS AND METHODS

Three unrelated families having six affected individuals with muscular weakness were recruited in the study. They had been referred to Dr. Zeinali's Medical Genetic Laboratory, Kawsar Human Genetics Research Center, Tehran, Iran. Clinical features, parental consanguinity, family history, histological findings, EMG results, creatine kinase level, mode of inheritance, signs and symptoms, and disease onset and progression were collected under the supervision of the experienced neurologists and/or a skilled genetic counselor. Family pedigrees were drawn using Progeny 9 software (https://www.progeny genetics.com).

Exome sequencing, data analysis, and interpretations

A total of three patients (proband of each family) were candidates for examination by WES. Genomic DNA was extracted from peripheral blood leukocytes using the salting out method[12] and/or proteinase K method[13]. WES was carried out using the Illumina HiSeq platform and SureSelect Human all exon V6 with a mean coverage of 100×. Data analysis was performed based on a pipeline[14] and SNVs, insertion/deletions, and duplication (<20 bp) were detected. Bioinformatics analysis of the sequencing results was performed using the Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org), locus specific mutation databases, ClinVar (https://www.ncbi.nlm.nih.gov/clinvar), HGMD (https://www.hgmd.cf.ac.uk/ac/validate.php), and standard bioinformatics software. Only rare (less than 1% of allelic frequency) coding and splicing variants were considered in this study, which included variants with predicted high (stop, frameshift, and splice essential) and moderate (missense, in frame, and splice region) impact on protein function. We performed final analysis on variations in the phenotypically relevant genes, i.e. genes with a known link to Mendelian disorders (according to Online Mendelian Inheritance in Men). The functional effect of the identified variants was assessed using in silico prediction tools, including Polymorphism Phenotyping v2 (PolyPhen-2; http://genetics.bwh.harvard.edu/pph2/), Sorting Intolerant from Tolerant (SIFT; https://sift.bii.a-star.edu.sg/), Mutation Taster (https://www.mutationtaster.org/), Functional Analysis through Hidden Markov Models v2.3 (FATHMM; http://fathmm.biocompute.org.uk/), Combined Annotation Dependent Depletion (CADD; https://cadd.gs.washington.edu) and Provean (https:// www. jcvi.org/). Additionally, the impact of splice site variants was predicted using the Human Splicing Finder (https://bio.tools/human_splicing_finder) and dbscSNV (Ada and RFscores)[15]. Selected variants were classified according to ACMG and the AMP (ACMG/AMP) standards and guidelines for the interpretation of sequence variants[16]. Based on the ACMG criteria for classifying pathogenic variants including very strong (PVS1), strong (PS1-PS4), moderate (PM1-PM6), and supporting (PP1-PP5), variations were classified as a pathogenic, likely pathogenic, VUS, likely benign, or benign[17,18]. Only the pathogenic, likely pathogenic, and VUS variants were considered for further investigation.

[DOR: 20.1001.1.1028852.2024.28.5.8.1]
[DOI: 10.61186/ibj.4018]

Variation confirmation and segregation analysis

Primers were designed and checked using bioinformatics tools and guidelines published previously[19] and were synthesized by Metabion (Steinkirchen, Germany). PCR products were purified, and Sanger sequencing was performed using Big Dye Terminator Cycle v1.1 Sequencing Kit (Thermo Fisher Scientific, California, USA) and ABI Prism 3130/xl Genetic Analyzer (ABI). Sequencing data were analyzed by Gene Runner (http://www.generunner.net/), Chromas (http://technelysium.com.au), and NCBI-Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi) software.

RESULTS

Affected individuals, in all three families, were the result of consanguineous marriage or their parents were from the same region. All the patients had an elevated serum creatine phosphokinase in biochemical blood tests. EMG and NCV (EMG/NCV) and muscle biopsy results showed myopathic and/or dystrophic features in all cases. Affected individual in Family 3 was wheelchair bound since the age of 16. Those in Family 1 (V2) and Family 2 (III-1) presented with difficulty in ambulation, and Family 2 (III-5 and III-7) presented with minimal change in walking at the time of recruitment in the study.

WES results and variant confirmation

WES was performed only for one affected individual in each family. Mutation details and other information are presented for each family below:

Family 1: The index patient was a 4.5-year-old girl, who was born from a consanguineous marriage. She presented with hip dislocation, inability to walk, and impaired sitting balance. There was similar presentation in her uncle who had died at the age of 14. A high level of serum creatine phosphokinase was observed in this patient, and the result of EMG-NCV showed a significant neurogenic process in both limbs at one years old. The possibility of spinal muscular atrophy was ruled out by the absence of deletion using the MLPA (MRC Holland, Amsterdam, Netherland) and Sanger sequencing of SMN1 gene. Two different homozygous pathogenic variants (*COL6A3*;NM_004369.4:c.4390C>T:p.Arg1464Ter) and VUS (ITGA7;NM_002206.3:c.3268C>T:p.Gln1090Ter) were identified in WES results. A segregation study in the family showed a homozygote variant in the ITGA7 gene in her unaffected father (IV:3) and brother (V:1), shown in Figure 1, which rules out the pathogenicity. Variation in the *COL6A3* gene was heterozygous in her parents (IV:3 and IV:4) and brother (V:1). This nonsense variant has not been reported in the HGMD database and has been submitted in ClinVar (RCV000579094.3) with no clinical information. Based on the bioinformatic studies, this mutation is located in the functional VWFA8 domain of the α3 subunit (1436-1609 aa) and causes the N-terminus to be blocked because of truncated protein formation. This finding is in favor of the pathogenicity of the identified mutation and the autosomal recessive mode (PVS1, PM2, PP4, and PP1).

Family 2: The index patient was a 30-year-old woman with muscle weakness from 12-13 years old and winging scapula. Muscle biopsy showed MD compatible with limb-girdle type, and EMG/NCV result was compatible with MD. She is from related parents with a similar presentation in her cousins. Two homozygote missense variants with possible relevance to the patient's phenotype were identified in *COL6A1* (NM_001848.2 :c.803A>G: p.Glu268Gly) and *COL6A2* (NM_001849. 3:c.2489G>A: p.Arg830Gln)

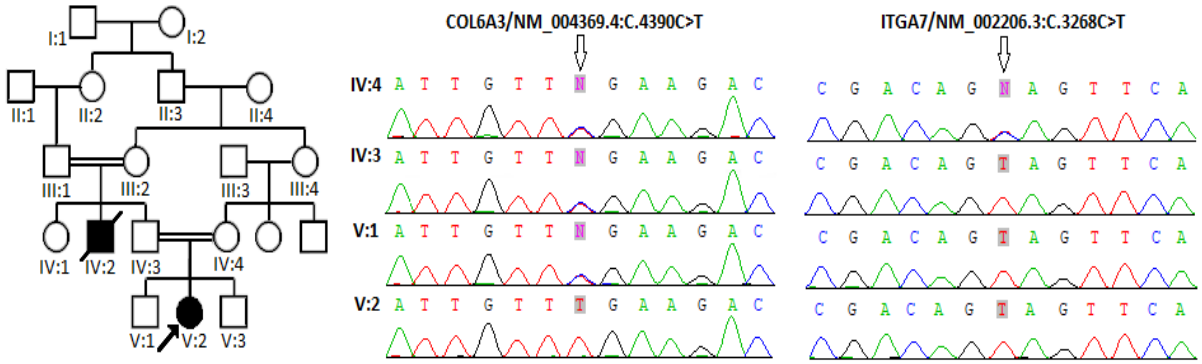

Fig. 1. Family 1: pedigree and sequencing results.

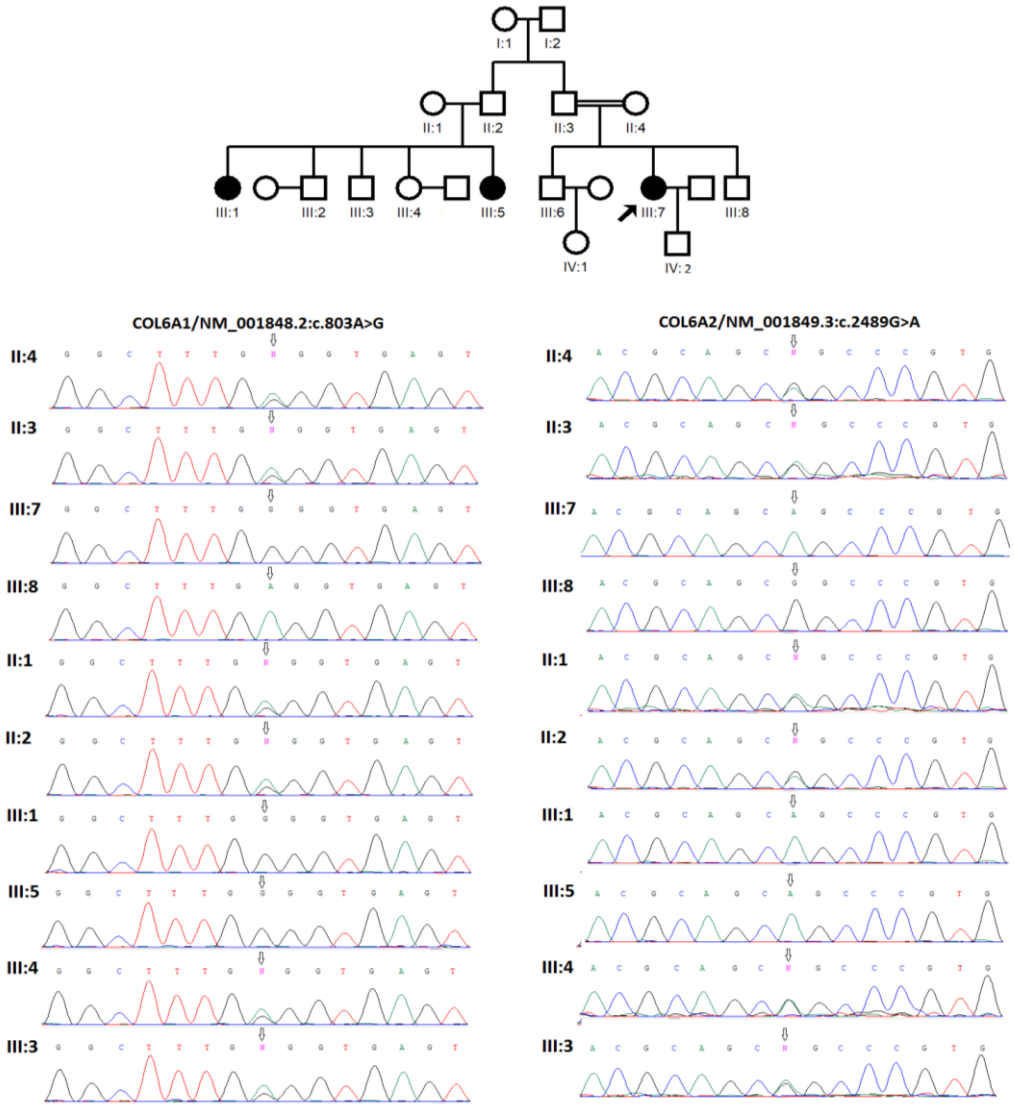

Fig. 2. Family 2: pedigree and sequencing results.

genes (Fig. 2) that was primarily classified as VUS based on the ACMG criteria. A segregating study of the variants within the family showed a homozygote change in both variations in all three affected cases. No unaffected family members were homozygotes for each variant. Following the segregation study in the family, both variants were re-classified as likely pathogenic (*COL6A1;*c.803A>G: PM2, PP3_M, PP4, PP1_M and *COL6A2*:c.2489G>A; PM5, PM2, PP3, PP4, PP1). Currently, there is no available evidence to accurately determine the disease-causing nature of these two variants, unless we use bioinformatics to predict their effects on protein structure or function. By using the Mutation Explorer tool (http://proteinformatics.org/mutation_explorer/), we attempted to check the changes in these two proteins by making mutations in the three-dimensional structure of the protein (Fig. 3). The mutation site in COL6A2 protein (Fig. 3B) is located in a more conserved region than the mutation site in COL6A1 protein (Fig. 3A). Therefore, it can be claimed that the probability of pathogenicity of mutation on *COL6A2* is higher than that on *COL6A1*.

Family 3: The patient is a 32-year-old man who was born from consanguineous parents with proximal weakness and atrophy (upper and lower limbs) from four years old with repeated falling and toe walking. The muscle biopsy result was compatible with MD. EMG/NCV study showed Duchenne or Emery- Dreifuss MD. Serum creatine phosphokinase and LDH levels had been elevated and MLPA of DMD gene was normal. We detected a heterozygous variant in the *COL6A2* gene

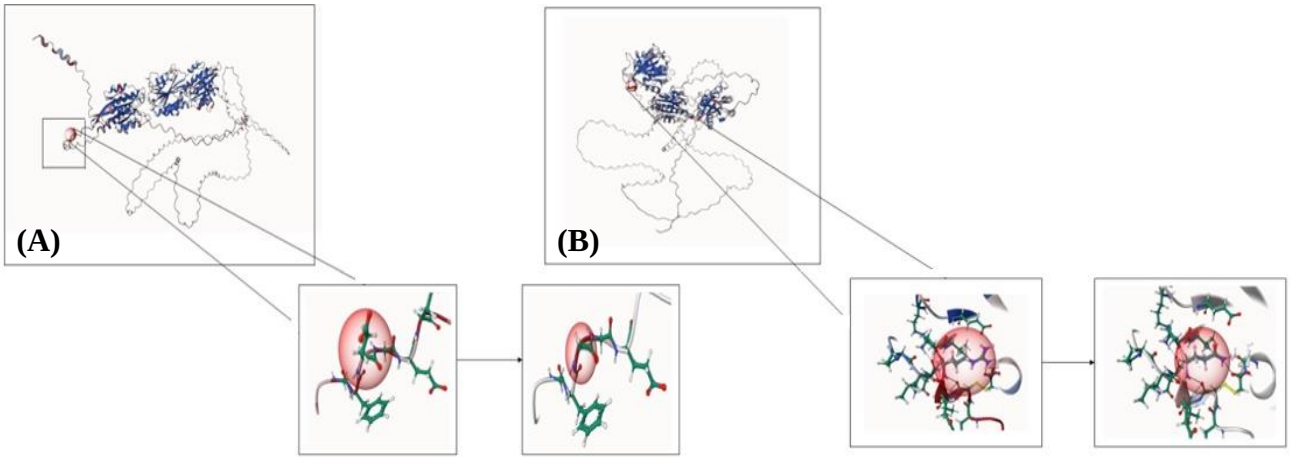

Fig. 3. Three-dimensional structure of variants effects on collagen 6 proteins: (A) COL6A1(p.Glu268Gly) and (B) COL6A2(p.Arg830Gln).

(NM_001849.3:c.1053+1G>T) that was classified as a pathogenic variant according to the recommendation of the ACMG. Sanger sequencing results confirmed the variation in the affected case. His parents were homozygote wild type. Besides, maternity and paternity were proved through haplotyping analysis. Therefore, this mutation could be considered as a de novo mutation or inherited by gonadal mosaicism (Fig. 4).

DISCUSSION

A molecular study of three unrelated Iranian families with muscular dystrophies using WES analysis showed several disease-causing variants in the *COL6A* gene family. Collagen VI is a ubiquitous extracellular matrix protein that exists in the stroma and forms a microfibrillar network in close connection with the basement membrane of most tissues. This protein includes three chains, α2, α1, and α3. The first two chains, α1 and α2, are encoded by two genes *COL6A1* and *COL6A2*, and located in chromosome 21 (21q22.3, NT_011515) 150 kb apart from each other. *COL6A3* gene encodes the α3 chain and is located at chromosomal region 2q37 (NT_005120)[20]. These three collagen family genes, i.e. *COL6A1/A2/A3*, have been associated with dominant and recessive types of muscular dystrophies. Specific variations that present in the specific regions of these genes and encode the Gly-X-Y domain, have been linked to dominant types of myopathy[21]. According to the OMIM database, pathogenic variants in *COL6A* genes cause collagen VI- related myopathy with hereditary patterns of autosomal dominant and autosomal recessive. Most affected individuals experience a weakening in their muscles and also the development of joint deformities known as contractures, which progressively limit the range of motion in the affected

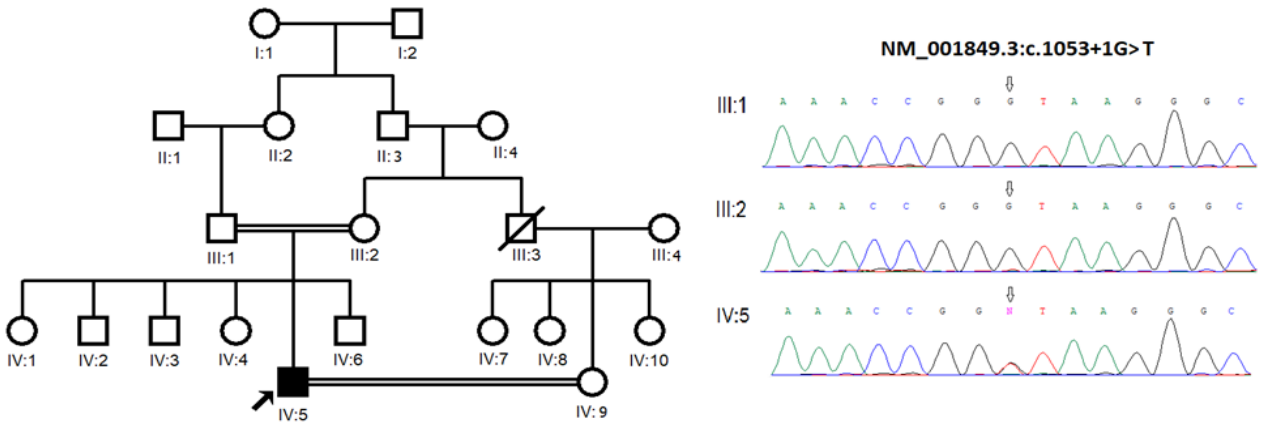

Fig. 4. Family 3: pedigree and sequencing results.

affected joints. Various types of collagen VI-related myopathies have been identified, with variable severity and age of onset. The mildest form is called Bethlem myopathy. There are also intermediate types with moderate severity, and finally, the most severe form which is known as UCMD[22]. In our study, clinical presentation and age of onset in the affected individuals were consistent with UCMD in family 1, Bethlem myopathy in family 2, and Bethlem myopathy or intermediate type in family 3, respectively. In families 1 and 3, the reported loss-of-function variants in the first family with a premature stop codon and in the third family with an effect on the splicing site led to a change in peptide size and consequently affect the quantity and quality of the final protein product.

Both missense variations, reported in *COL6A1* and *COL6A2* genes in family 2, were co-segregated within the affected family members (a sibling and two cousins). Linkage between gene families is not so uncommon and has been previously reported for different members of the collagen gene cluster[23]. *COL4A1* and *COL4A2* genes located in a gene cluster on the terminal end of chromosome 13[24], and also *COL3A1* and *COL5A2* gene[25] are the examples of this linkage. Two variations in *COL6A1* and *COL6A2* genes might be linked together to cause a change in the protein function. Variation in the *COL6A1* gene causes the replacement of glutamic acid (E) with glycine (G) at codon 268 of the protein (p.Glu268Gly). Glutamic acid is an acidic and polar negatively charged amino acid, but Glycine is a neutral and non-polar amino acid. Furthermore these two amino acids differ in size and hydrophobicity. Glycine is a smaller and more hydrophobic residue than glutamic acid. Collectively, these features might lead to changes in protein folding and conformation. Glycine is very flexible and can disturb the required rigidity of the protein at this position. This substitution in the α1 chain of collagen IV has been located within a stretch of residues and most of the non-VUS reported variations are (likely) pathogenic. In silico analysis of this missense variation using dbscSNV (Ada and RF scores) predicted that the sequence change might disrupt the consensus splice site and consequently RNA splicing. In comparison, a mutation found in the *COL6A2* gene replaces arginine (R), as a basic and polar residue with glutamine (Q), which is neutral and polar, at codon 830 (p.Arg830Gln). Since this mutation changes a positively charged polar amino acid (R) to a neutral amino acid (Q), it can affect the polarity of the protein and its dimerization. On the other hand, this variation was reported as likely pathogenic (1) and uncertain significance (3) in ClinVar database (VCV000210747. 14). Also, another variation in the same amino acid residue (p.Arg830Trp) (VCV000287934.12) has been determined to be pathogenic (PM5)[26,27]. Arginine residue at position 830 of the alpha 2 chain of collagen 6 has been situated at the beginning of the C2 domain, adjacent to a metal ion-dependent adhesion site motif and exhibits evolutionary conservation[26]. Although type VI collagen consists of three different chains, α1(VI), α2(VI), and α3(VI), the dimer formation is primarily driven by the interactions involving the α2(VI) chain. The crucial elements in this interaction include the metal ion-dependent adhesion site motif located in the α2C2 A-domain, and the GER triplet (residues 119-111, i.e. glycine, gluatamic acid, and Argenine) sequence presents in the helical domain of another α2(VI) chain[28]. These data suggests that the residue in question has clinical significance, and variants affecting this residue are likely contributed to disease. Substitution of p.Arg830Gln in *COL6A2* has been reported in four affected persons in four independent families in the form of compound heterozygous[21,26,29,30]. In at least one individual, this mutation is present on the opposite chromosome (trans) from a known pathogenic variant[30]. However, it has not been reported in homozygous form, and our study for the first time reports this mutation as homozygous state in the affected people. According to the in silico predictions available on UniProt (https://www.uniprot.org/uniprotkb/P12110/variant-viewer), this mutation is classified as probably deletrious, with a CADD score of 29.3, and the evolutionary model of variant effects predicts it as pathogenic. The MetaRNN score of this variant is equal to 0.799, and the range between 0.748 and 0.841 is classified as “supporting pathogenic” (PP3).

The chromosomal location of *COL6A1* and *COL6A2* genes is very close to each other (150 kb) on chromosome 21 (21q22.3, NT_011515)[20]. Therefore, the coexistence of these two variants can be the result of two independent mutations in linkage disequilibrium, though only one of the two variants could be the actual cause of the disease. RNA sequencing studies regarding the expression profile of these two genes in the affected individuals can help to determine the pathogenicity of these two variants.

A heterozygous pathogenic mutation in the *COL6A2* gene was found in individual IV-5 in the third family. According to HGMD professional 2017.1, this variant had previously been described as a cause of Bethlem myopathy by Hicks et al.[31] and Foley et al.[32]. The patient's parents did not have that mutation and maternity and paternity have been proven through short tandem repeat haplotyping. At first, IV-5 seemed to be the only case in the family, and the mutation also seemed to be de novo. However, with the appearance of some symptoms in the patient's sister, this was revoked.

[DOR: 20.1001.1.1028852.2024.28.5.8.1]
[DOI: 10.61186/ibj.4018]

DNA analysis for the patient's sister was not feasible due to parents' unwillingness to give a sample. Therefore, it was impossible to conclude whether this variant was inherited or it was a de novo mutation, since mutation in the sister should be verified to make such a conclusion.

CONCLUSION

Our findings reveal that WES can be used as one of the most effective solutions to determine the genetic cause of diseases with high genetic heterogeneity, including MDs. However, there are some limitations to the clinical application of this method; the complexity of the genetic basis of neuromuscular disorders, including collagen 6-related myopathies, can hamper the determination of the pathogenicity of each variants detected using WES alone. Therefore, determining the mode of inheritance and distinguishing the severity of the disorder, as well as genotype-phenotype correlation require comprehensive clinical, paraclinical, and genetic investigations. Herein, we attempted to provide a support on the pathogenicity of the identified variants.

DECLARATIONS

Acknowledgments

We would like to acknowledge the help of many distinguished physicians and families who participated this study. We are grateful to all of them. The paper is part of Nasibeh Soltani's PhD thesis. We would like to thank the staff of Dr. Zeinali's Medical Genetics Lab who helped us during preparation of this study. No artificial intelligence-based technologies have been used in this study.

Ethical approval

All the experimental procedures in this study were approved by the Institutional Review Board of the Kawsar Human Genetics Research Center, Tehran, Iran (ethical code: 1400-6328).

Consent to participate

All the subjects voluntarily agree to participate in this research. Written informed consents were provided by the affected cases or their parents.

Consent for publication

All authors reviewed the results and approved the final version of the manuscript.

Authors' contributions

NS: made substantial contributions to the conception, designed the work, drafted the work and/or substantively revised it; ZSh: drafted the work and/or substantively revised it; MK: design of the work; MSF: analysis and interpretation of data; FZM: acquisition of data; MA: analysis and interpretation of data; MJ: acquisition of data; HB: acquisition of data; RZ: drafted the work and/or substantively revised it. SZ: made substantial contributions to the conception, designed the work, drafted the work and/or substantively revised it.

Data availability

All relevant data can be found within the manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

Part of this study was financially supported by Pasteur Institute of Iran and Kawsar Human Genetics Research Center.

Supplementary information

The online version does not contain supplementary material.

REFERENCES

1. Emery AE. The muscular dystrophies. Lancet. 2002;359(9307):687-95.
2. Brown SC, Muntoni F, Sewry CA. Non-sarcolemmal muscular dystrophies. Brain Pathol. 2001;11(2):193-205.
3. Nigro V, Piluso G. Spectrum of muscular dystrophies associated with sarcolemmal-protein genetic defects. Biochim Biophys Acta. 2015;1852(4):585-93.
4. Carter JC, Sheehan DW, Prochoroff A, Birnkrant DJ. Muscular Dystrophies. Clin Chest Med. 2018;39(2):377-89.
5. Mercuri E, Muntoni F. Muscular dystrophies. Lancet. 2013;381(9869):845-60.
6. Theadom A, Rodrigues M, Roxburgh R, Balalla Sh, Higgins Ch, Bhattacharjee R, et al. Prevalence of muscular dystrophies: a Systematic literature review. Neuroepidemiology. 2014;43(3-4):259-68.
7. Rahimov F, Kunkel LM. Cellular and molecular mechanisms underlying muscular dystrophy. J Cell Biol. 2013;201(4):499-510.
8. Gaina G, Budisteanu M, Manole E, Lonica E. Clinical and Molecular Diagnosis in Muscular Dystrophies. Muscul Dystrophies. IntechOpen; 2019. Available from: http://dx.doi.org/10.5772/intechopen.85339
9. Dardas Z, Swedan S, Al-Sheikh Qassem A, Azab B. The impact of exome sequencing on the diagnostic yield of muscular dystrophies in consanguineous families. Eur J Med Genet. 2020;63(4):103845.
10. Rabbani B, Mahdieh N, Hosomichi K, Nakaoka H, Inoue I. Next-generation sequencing: impact of exome

sequencing in characterizing Mendelian disorders. J Hum Genet. 2012;57(10):621-32.
11. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet. 2011;48(9):580-9.
12. Gautam A. Isolation of DNA from Blood Samples by Salting Method. In: DNA and RNA Isolation Techniques for Non-Experts. Techniques in Life Science and Biomedicine for the Non-Expert. Springer, Cham. https://doi.org/10.1007/978-3-030-94230-4_12.
13. Qamar W, Khan MR, Arafah A. Optimization of conditions to extract high quality DNA for PCR analysis from whole blood using SDS-proteinase K method. Saudi J Biol Sci. 2017;24(7):1465-9.
14. Yap ZhY, Efthymiou S, Seiffert S, Vargas Parra K, Lee S, Nasca A, et al. Bi-allelic variants in OGDHL cause a neurodevelopmental spectrum disease featuring epilepsy, hearing loss, visual impairment, and ataxia. Am J Hum Genet. 2021;108(12):2368-84.
15. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. Nucleic Acids Res. 2014;42(22):13534-44.
16. Richards S, Aziz N, Bale Sh, 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.
17. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073-81.
18. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-9.
19. Abd-Elsalam KA. Bioinformatic tools and guideline for PCR primer design. African J Biotechnol. 2003;2(5):91-5.
20. Lampe AK, Bushby KMD. Collagen VI related muscle disorders. J Med Genet. 2005;42(9):673-85.
21. Nallamilli BRR, Chakravorty S, Kesari A, Tanner A, Ankala A, Schneider Th, et al. Genetic landscape and novel disease mechanisms from a large LGMD cohort of 4656 patients. Ann Clin Transl Neurol. 2018;5(12):1574-87.
22. Becker AKA, Mikolajek H, Paulsson M, Wagener R, Werner JM. A structure of a collagen VI VWA domain displays N and C termini at opposite sides of the protein. Structure. 2014;22(2):199-208.
23. Francomano CA, Cutting GR, McCormick MK, Chu ML, Timpl R, Hong HK, et al. The COL6A1 and COL6A2 genes exist as a gene cluster and detect highly informative DNA polymorphisms in the telomeric region of human chromosome 21q. Hum genet. 1991;87(2):162-6.
24. Cutting GR, Kazazian Jr H, Antonarakis SE, Killen PD, Yamada Y, Francomano CA. Macrorestriction mapping of COL4A1 and COL4A2 collagen genes on human chromosome 13q34. Genomics. 1988;3(3):256-63.
25. Tsipouras P, Schwartz RC, Liddell AC, Salkeld CS, Weil D, Ramirez F. Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on the long arm of human chromosome 2. Genomics. 1988;3(3):275-7.
26. Foley AR, Hu Y, Zou Y, Columbus A, Shoffner J, Dunn DM, et al. Autosomal recessive inheritance of classic Bethlem myopathy. Neuromuscul disord. 2009;19(12):813-7.
27. Westra D, Schouten MI, Stunnenberg BC, Kusters B, Saris ChGJ, Erasmus CE, et al. Panel-based exome sequencing for neuromuscular disorders as a diagnostic service. J Neuromuscul Dis. 2019;6(2):241-58.
28. Ball S, Bella J, Kielty C, Shuttleworth A. Structural basis of type VI collagen dimer formation. J Biol Chem. 2003;278(17):15326-32.
29. Gualandi F, Urciuolo A, Martoni E, Sabatelli P, Squarzoni S, Bovolenta M, et al. Autosomal recessive Bethlem myopathy. Neurology. 2009;73(22):1883-91.
30. Fan Y, Liu A, Wei C, Yang H, Chang X, Wang S, et al. Genetic and clinical findings in a Chinese cohort of patients with collagen VI-related myopathies. Clin Genet. 2018;93(6):1159-71.
31. Hicks D, Lampe AK, Barresi R, Charlton R, Fiorillo C, Bonnemann CG, et al. A refined diagnostic algorithm for Bethlem myopathy. Neurology. 2008;70(14):1192-9.
32. Foley AR, Quijano-Roy S, Collins J, Straub V, McCallum M, Deconinck N, et al. Natural history of pulmonary function in collagen VI-related myopathies. Brain. 2013;136(Pt 12):3625-33.