Exome sequencing identifies an MYH3 mutation in a family with distal arthrogryposis type 1

David M. Alvarado
Washington University School of Medicine in St. Louis

Jillian G. Buchan
Washington University School of Medicine in St. Louis

Christina A. Gurnett
Washington University School of Medicine in St. Louis

Matthew B. Dobbs
Washington University School of Medicine in St. Louis

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Part of the Medicine and Health Sciences Commons

Recommended Citation
Exome Sequencing Identifies an MYH3 Mutation in a Family with Distal Arthrogryposis Type 1

David M. Alvarado, PhD, Jillian G. Buchan, BS, Christina A. Gurnett, MD, PhD, and Matthew B. Dobbs, MD

Investigation performed at Washington University School of Medicine, St. Louis, Missouri

Background: Few genes responsible for distal arthrogryposis type 1 are known, although genes coding for the proteins in the sarcomere have been implicated in other types of distal arthrogryposis. Cost-effective sequencing methods are now available to examine all genes in the human genome for the purpose of establishing the genetic basis of musculoskeletal disorders.

Methods: A multigenerational family with distal arthrogryposis type 1 characterized by clubfoot and mild hand contractures was identified, and exome sequencing was performed on DNA from one of the affected family members. Linkage analysis was used to confirm whether a genetic variant segregated with distal arthrogryposis.

Results: Exome sequencing identified 573 novel variants that were not present in control databases. A missense mutation in MYH3 (a gene coding for the heavy chain of myosin), causing an F437I amino acid substitution, was identified that segregated with distal arthrogryposis in this family. Linkage analysis confirmed that this MYH3 mutation was the only exome variant common to all six affected individuals.

Conclusions: Identification of an MYH3 mutation in this family with distal arthrogryposis type 1 broadens the phenotype associated with MYH3 mutations to include distal arthrogryposis types 1, 2A (Freeman-Sheldon syndrome), and 2B (Sheldon-Hall syndrome). Exome sequencing is a useful and cost-effective method to discover causative genetic mutations, although data from extended families may be needed to confirm the importance of the hundreds of identified variants.

Clinical Relevance: Distal arthrogryposis type 1 should be considered in the differential diagnosis of isolated clubfoot, particularly when hand contractures are present in any family member or when the clubfoot is severe and resistant to treatment.

Distal arthrogryposis is characterized by contractures of the distal regions of the hands and feet. The severe types of distal arthrogryposis, types 2A (also called Freeman-Sheldon syndrome) and 2B (also called Sheldon-Hall syndrome), include facial involvement and scoliosis, whereas distal arthrogryposis type 1 (DA1) does not. Distal arthrogryposis type 1 affects approximately one in 10,000 children and represents the most common type.

Multiple genes encoding proteins in the sarcomere, including myosin heavy chain 3 (MYH3), myosin heavy chain 8 (MYH8), tropomyosin 2 (TPM2), troponin I2 (TNNI2), troponin T3 (TNNT3), and myosin binding protein C1 (MYBPC1), have been implicated in distal arthrogryposis syndromes. Although MYH3 mutations account for nearly all cases of Freeman-Sheldon syndrome (type 2A) and nearly one-third of all cases of Sheldon-Hall syndrome (type 2B), to our knowledge MYH3 mutations have not been described in patients with distal arthrogryposis type 1. Previous studies have shown that mutations in known genes are rare causes of distal arthrogryposis type 1, therefore, genetic heterogeneity is expected and additional causative genes remain to be identified.

Because of the large number of genes as well as the relatively large size of the genes already implicated in distal arthrogryposis syndromes, new methods to sequence all genes are

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. In addition, one or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete Disclosures of Potential Conflicts of Interest submitted by the authors are always available with the online version of this article at jbjs.org.
necessary to effectively define the genetic basis of distal arthrogryposis in individual families. Exome sequencing is a new genetic tool that uses next-generation sequencing methods to identify mutations within the coding exons of the entire human genome (the "exome")\(^1\). Because the exome comprises just 1% of the genome, the cost of sequencing only the exons is currently only a fraction of the cost of whole-genome sequencing.

The next-generation exome capture and sequencing methods that were used in this study represent a substantial advance in research methods and, despite their limitations, these methods are likely to soon have a major impact on the diagnosis of patients with inherited musculoskeletal disorders. For instance, our identification of a mutation in MYH3 in the family described in this study broadens the phenotype associated with MYH3 mutations to include distal arthrogryposis type 1 and suggests that there may be substantial overlap between types 1, 2A, and 2B. Distal arthrogryposis type 1 should be considered in the differential diagnosis of isolated clubfoot, particularly in the presence of even minor hand contractures in the patient or family members.

**Materials and Methods**

The study was approved by our institutional review board, and all subjects gave informed consent. DNA was obtained from all affected and several unaffected members of a family that included six individuals with distal arthrogryposis type 1. Exome capture was performed for one of the affected family members (the proband) with use of the SureSelect All Exon 38Mb kit (Agilent Technologies, Santa Clara, California). The DNA was sequenced on a single lane of a HiSeq 2000 sequencer (Illumina, San Diego, California). The sequence was aligned to the hg18 assembly of the genome with use of Novoalign software (Novocraft Technologies, Selangor, Malaysia), and variants were identified with use of SAMtools\(^12\) and SeattleSeq\(^13\). Only variants with at least sixfold read coverage and a Phred-scaled single-nucleotide polymorphism quality of ≥30 were included for further analysis. Variants present in the 1000 Genomes Project (http://www.1000genomes.org), dbSNP (build 129; National Center for Biotechnology Information, Bethesda, Maryland), or our in-house exome database of twenty patients with unrelated disorders were removed. Confirmation of selected variants was performed by means of Sanger sequencing with use of an ABI 3730 Sequencer (Life Technologies, Carlsbad, California).

Linkage analysis was performed with array genotyping data from six affected and six unaffected family members, assuming 80% penetrance and a 0.1% frequency of the allele responsible for the disease. The genotyping was performed with use of a Mapping 10K XbaI array (Affymetrix, Santa Clara, California) by the National Institutes of Health Neuroscience Microarray Consortium (Los Angeles, California). Linkage analysis was performed as previously described\(^9\).

**Source of Funding**

Grants from Shriners Hospitals for Children, the Children's Discovery Institute, the March of Dimes (Basil O'Connor Starter Scholar Research Award), the St. Louis Children's Hospital Foundation, the Pediatric Orthopaedic Society of
North America, the Orthopaedic Research and Education Foundation, and the National Institutes of Health were used to fund the basic science experiments.

**Results**

A multigenerational family presented with incompletely penetrant, autosomal-dominant distal arthrogryposis type 1 characterized by clubfoot and mild hand contractures. Five family members had clubfoot, which was bilateral in four individuals and unilateral in one (Fig. 1, A). The proband (identified as III:3 in the family pedigree), who had unilateral clubfoot, had initial correction of the clubfoot with the Ponseti method\(^{14}\) (Fig. 1, B) but had two relapses and required repeat Achilles tendon tenotomies to maintain correction. The other family members with clubfoot had been treated with extensive soft-tissue releases. Two of the family members with clubfoot (individuals II:7 and II:9) also had mild hand contractures predominantly involving the distal interphalangeal joints (Fig. 1, C and D). In addition, one family member (individual II:1) developed contractures at the distal interphalangeal joints during adolescence but did not have clubfoot or other lower-extremity involvement. Other signs of hand involvement, including absent flexion creases, overriding fingers, and proximally placed thumbs, were absent in the proband and the other individuals who exhibited only clubfoot. Because facial involvement was absent, a diagnosis of distal arthrogryposis type 1 was made.

Sequencing of myosin binding protein C1 (MYBPC1), a gene that has recently been implicated in some cases of distal arthrogryposis type 1,\(^9\) did not reveal any causative mutations. However, because few genes responsible for distal arthrogryposis type 1 have been identified\(^1,9\), we performed exome sequencing of the DNA of the proband on a research basis. Analysis of a single lane of Illumina HiSeq paired-end sequence resulted in 142,366,935 reads, of which 93% could be aligned with the human genome. This provided 93% coverage of the sequence at a minimum sequencing depth of 6\(\times\). With this method, a total of 24,323 variants that would cause nonsynonymous amino acid substitutions, splice-site variants, insertions, or deletions were identified in the exome of the proband. After the exclusion of variants that were present at a frequency of >5% in control databases, 524 single-nucleotide variants and forty-nine indels (insertions or deletions) remained.

One of these 573 variants consisted of a missense mutation in MYH3, the embryonic myosin heavy chain gene that has previously been implicated in distal arthrogryposis types 2A and 2B.\(^4\) Sanger sequencing confirmed the presence of the MYH3 F437I mutation in all of the individuals with clubfoot as well as in the individual with adolescent-onset hand contractures (Fig. 2). The A\(\rightarrow\)T missense mutation at chr17:10488494 (hg18) results in the substitution of the amino acid isoleucine for phenylalanine at residue 437. This residue is within the

---

Fig. 2
Pedigree of the family, showing members with and without distal arthrogryposis type 1. Black shading indicates clubfoot, and asterisks indicate hand contractures. The presence of the MYH3 F437I mutation is indicated by "Mut," and its absence is indicated by "WT."

Fig. 3
Alignment of MYH3 amino acids surrounding the location corresponding to the F437I mutation in sequences from multiple species. The phenylalanine at residue 437 is indicated by the rectangle and is conserved across all of the species shown, including mouse, rat, human, cow, frog, and chicken. The sequences were aligned with use of the CLUSTALW program (http://align.genome.jp).
myosin head and is highly conserved across multiple species (Fig. 3), but it lies outside the ATP binding groove where mutations appear to cause the most severe form of distal arthrogryposis, Freeman-Sheldon syndrome (type 2A) (Fig. 4). This mutation is predicted with 99.7% confidence by the PolyPhen-2 program to be “probably damaging.”

Fig. 4
Ribbon diagram of the MYH3 S2 (myosin head) region containing amino acids 4 through 1022. The location of the MYH3 F437I mutation identified in the family with distal arthrogryposis type 1 is shown in orange. The locations of five other known mutations causing distal arthrogryposis type 2A (Freeman-Sheldon syndrome) are shown in red; four of these are in the myosin head and lie in the groove that forms the ATP binding site. The image was generated by the Swiss-PDBViewer (DeepView) program (http://ca.expasy.org/spdbv) for protein accession number P11055 (MYH3).

Fig. 5
Linkage of distal arthrogryposis type 1 to chromosome 17p. A genome-wide profile of parametric multipoint LOD (logarithm of odds) scores for the family with distal arthrogryposis type 1, constructed from the Affymetrix 10K array data. The human chromosomes, running from the p terminal (left) to the q terminal (right), are concatenated on the horizontal axis. The vertical axis shows the parametric LOD score for the physical locations on the human chromosome indicated on the horizontal axis. The peak LOD (pLOD) score (defining the strongest region of linkage) of 1.86 was within chromosome 17p.
To confirm the importance of this MYH3 F437I mutation, linkage analysis was performed with the use of Affymetrix 10K single-nucleotide polymorphism data from all six affected individuals in the family. The greatest LOD (logarithm of odds) score, 1.86, was found on chromosome 17p, which contains the MYH3 gene (Fig. 5), and the MYH3 mutation was the only exome variant that was located within the approximately 3-Mb linkage interval. None of the other, less significant linkage intervals contained any of the other rare variants.

**Discussion**

Our results demonstrate that MYH3 mutations may be associated with distal arthrogryposis type 1 (as well as with types 2A and 2B as previously described). All types of distal arthrogryposis are characterized by extensive phenotypic variability between and within families, and our data suggest that there is also likely to be clinical and genetic overlap of distal arthrogryposis type 1 with types 2A and 2B. Future classification systems for these syndromes that take into account both the genotype and the phenotype are needed.

The identification of MYH3 as the location of the disease-causing mutation in this family with distal arthrogryposis type 1 is supported by the known association of this gene with distal arthrogryposis types 2A and 2B, which are disorders closely related to distal arthrogryposis type 1, and by the prediction that the F437I mutation would be damaging to the function of the protein. Furthermore, the phenylalanine at residue 437 is highly conserved across many mammalian species and the mutation was not identified in any of the controls. Most notably, segregation of the MYH3 F437I mutation in all affected family members increased our confidence that we have correctly identified the disease-causing mutation. The sequencing of variants in additional family members remains a useful, if not critical, method to increase the likelihood that a disease-causing variant has been correctly identified, and this method should continue to be offered by clinical laboratories.

The location of the MYH3 mutation may correlate with the severity of distal arthrogryposis, as suggested by Toydemir et al. The consistently mild phenotype of the distal arthrogryposis in all of the affected family members provides further support for a genotype-phenotype correlation in diseases involving MYH3. The F437I mutation is slightly outside the ATP binding groove where mutations causing the most severe form of distal arthrogryposis, Freeman-Sheldon syndrome (type 2A), are located. However, understanding the physiological impact of MYH3 missense mutations that could alter skeletal muscle contractility (by affecting nucleotide binding, ATP catalysis, or interaction with actin or regulatory proteins) has been hindered by the lack of convenient assays. Further studies are needed to better understand how these mutations cause disease.

The family described in this study also demonstrates the occasional difficulty of identifying distal arthrogryposis type 1 as the cause of isolated clubfoot. Although three individuals from this family had isolated clubfoot, the presence of hand contractures in even one family member should indicate the possibility of distal arthrogryposis. Nevertheless, hand contractures may not appear until adolescence, as demonstrated in individual II:1, who had the MYH3 F437I mutation. Although worsening of contractures may occur if appropriate therapy such as bracing is not used, contractures are generally thought to be nonprogressive in cases of distal arthrogryposis. The nonprogressive and often congenital nature of contractures in distal arthrogryposis had been attributed to the location of the defects in genes such as MYH3 that were believed to be expressed exclusively during fetal development. Recent studies, however, have shown that MYH3 is also expressed in adult muscle, perhaps explaining the late-onset hand contractures in one member of this family. Distal arthrogryposis should also be considered in the differential diagnosis for families with highly penetrant, autosomal-dominant clubfoot, as this inheritance pattern is relatively rare in isolated clubfoot, with the exception of clubfoot associated with recurrent chromosome 17q23 microdeletions and microduplications that are responsible for approximately 5% of all cases of familial clubfoot. Clubfoot that is resistant to treatment should also signal the possibility of distal arthrogryposis. Although the proband’s clubfoot was treated nonsurgically with use of the Ponseti method, she developed multiple recurrences. Despite these complications, excellent outcomes are possible in patients in whom distal arthrogryposis is treated with use of the Ponseti method.

The results of this study demonstrate the power of new sequencing methods for the diagnosis of human musculoskeletal disease. Although exome sequencing of the patient in this study was done on a research basis, clinically available diagnostic testing will soon utilize similar methods to screen large gene panels for mutations related to diseases. Exome analysis is a reasonable approach for the identification of disease-causing mutations for several reasons. First, the vast majority of currently identified disease-causing mutations are located within the coding regions of genes (the exome), although this is likely to change in the future as researchers are now discovering the extent to which regions that do not code for proteins can affect human health and disease. Second, because the exome comprises only 1% of the genome, the cost of sequencing the exome is currently a fraction of the cost of whole-genome sequencing. Finally, analysis of the exome is less complicated than analysis of the entire genome because of the availability of well-established methods to predict or experimentally test for the effect of a mutation on the function of the protein.

The interpretation of variants in coding exons, while simpler than the interpretation of the results of whole-genome sequence analysis, remains complex. Analysis of exome data relies heavily on the use of filtering methods to exclude variants that are also present in some healthy (control) individuals. These normal variants are being compiled through the efforts of the 1000 Genomes Project, the goal of which is to identify rare variants in more than 1000 healthy individuals from various populations. However, such “control” databases are likely to include disease-causing mutations with low penetrance as well as mutations causing late-onset disorders that have not yet manifested in the reportedly normal individuals who provided
the control samples. Collections of disease-causing variants in databases such as the Human Gene Mutation Database will also assist in the analysis of exome sequencing data, but this database currently lists only thirteen MYH3 mutations. Be-cause mutations responsible for rare disorders such as distal arthrogryposis may be found only in single families, such data-bases may currently be of limited use for exome analysis. Eventually, however, they may become populated with variants that are known with high confidence to be associated with specific diseases. Finally, algorithms are available to predict the functional effects of gene mutations that modify proteins. However, many disease-associated mutations, including several previously described mutations in human myosin heavy chain genes, are incorrectly predicted by programs such as PolyPhen-2 to be “benign” mutations, thus limiting the overall usefulness of this type of analysis.

Despite the limitations of exome sequencing, such as the difficulty of distinguishing the disease-causing variant from the large number of other novel variants identified, we expect that these techniques will rapidly advance medical research and revolutionize patient care. As demonstrated here with MYH3, disease-associated gene variations are gradually being shown to be responsible for an expanded range of phenotypes as exome sequencing technology identifies mutations that are also found in patients with related disorders. Although MYH3 mutations are uncommon causes of isolated clubfoot, the diagnosis of distal arthrogryposis should be considered when hand contractures (even those that develop during adolescence) are present in any family members or when the clubfoot is severe and resistant to treatment. Although MYH3 sequencing is currently available as a single-gene genetic test, we anticipate that extended gene panels or exome-based sequencing will soon be available for musculoskeletal disorders such as distal arthrogryposis.

References