Recessive Spondylocarpotarsal Synostosis Syndrome Due to Compound Heterozygosity for Variants in *MYH3*

Sophia R. Cameron-Christie,[1,17] Constance F. Wells,[1,2,3,17] Marleen Simon,[4,5] Marja Wessels,[5] Candy Z.N. Tang,[1] Wenhua Wei,[1] Riku Takei,[1] Coranne Aarts-Tesselaar,[6] Sarah Sandaradura,[7,8] David O. Sillence,[7,9] Marie-Pierre Cordier,[10] Hermine E. Veenstra-Knol,[11] Matteo Cassina,[12] Kathrin Ludwig,[13] Eva Trevisson,[12] Melanie Bahlo,[14,15] David M. Markie,[16] Zandra A. Jenkins,[1] and Stephen P. Robertson[1,*]

Spondylocarpotarsal synostosis syndrome (SCTS) is characterized by intervertebral fusions and fusion of the carpal and tarsal bones. Biallelic mutations in *FLNB* cause this condition in some families, whereas monoallelic variants in *MYH3*, encoding embryonic heavy chain myosin 3, have been implicated in dominantly inherited forms of the disorder. Here, five individuals without *FLNB* mutations from three families were hypothesized to be affected by recessive SCTS on account of sibling recurrence of the phenotype. Initial whole-exome sequencing (WES) showed that all five were heterozygous for one of two independent splice-site variants in *MYH3*. Despite evidence indicating that three of the five individuals shared two allelic haplotypes encompassing *MYH3*, no second variant could be located in the WES datasets. Subsequent genome sequencing of these three individuals demonstrated a variant altering a 5′ UTR splice donor site (rs557849165 in *MYH3*) not represented by exome-capture platforms. When the cohort was expanded to 16 SCTS-affected individuals without *FLNB* mutations, nine had truncating mutations transmitted by unaffected parents, and six inherited the rs557849165 variant in *trans,* an observation at odds with the population allele frequency for this variant. The rs557849165 variant disrupts splicing in the 5′ UTR but is still permissive of *MYH3* translational initiation, albeit with reduced efficiency. Although some *MYH3* variants cause dominant SCTS, these data indicate that others (notably truncating variants) do not, except in the context of compound heterozygosity for a second hypomorphic allele. These observations make genetic diagnosis challenging in the context of simplex presentations of the disorder.

Introduction

Spondylocarpotarsal synostosis syndrome (SCTS [MIM: 272460]) is a rare skeletal dysplasia characterized by the association of vertebral fusions with normal segmentation of the ribs, sometimes leading to severe scoliosis, as well as carpal and tarsal fusions. Other clinical features include short stature, mildly dysmorphic features, cleft palate, deafness, joint laxity, and dental enamel hypoplasia.[1–3] SCTS was first identified as an autosomal-recessive disease caused by homozygous or compound-heterozygous loss-of-function mutations in the gene encoding filamin B (*FLNB* [MIM: 603381]).[4] However, the condition exhibits genetic heterogeneity given that many affected individuals do not carry detectable pathogenic variants in *FLNB* and some families appear to inherit the disease in a dominant pattern.[5]

FLNB was the sole identified locus for SCTS until recent reports of monoallelic variants in *MYH3* (MIM: 160720) in families affected by vertical transmission of a SCTS phenotype.[6] An unrelated individual with simplex SCTS and affected members of a three-generation family have also been reported to be heterozygous for protein-altering *MYH3* mutations,[7] including an in-frame deletion of a single amino acid and a frameshift variant leading to a presumptive premature termination of protein translation. Before these observations, monoallelic missense *MYH3* mutations transmitted in a dominant fashion were also shown to underlie cases of distal arthrogryposis type 8 (DA8, otherwise known as autosomal-dominant multiple pterygium syndrome [MIM: 178110]).[8] This phenotype can include scoliosis with spinal fusions and hence partially phenocopy SCTS. Dominant mutations in *MYH3* also cause distal arthrogryposis types 1 (DA1

[1]Department of Women's and Children's Health, Dunedin School of Medicine, University of Otago, Dunedin 9054, New Zealand; [2]Paris Diderot University, Sorbonne Paris Cité, Faculty of Medicine, Paris 75007, France; [3]Département de Génétique Médicale, Maladies Rares et Médecine Personnalisée, Centre de Référence Anomalies du Développement et Syndromes Malformatifs, Centre Hospitalier Universitaire de Montpellier, Université de Montpellier 34295, Montpellier Cedex 5, France; [4]Department of Genetics, University Medical Center Utrecht, 3508 GA Utrecht, the Netherlands; [5]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam 3015 CE, the Netherlands; [6]Amphia Hospital, 4819 EV Breda, the Netherlands; [7]Department of Clinical Genetics Children's Hospital at Westmead, Sydney, NSW 2145, Australia; [8]Discipline of Child and Adolescent Health, University of Sydney, Sydney, NSW 2006, Australia; [9]Discipline of Genetic Medicine, University of Sydney, Sydney, NSW 2006, Australia; [10]Clinical Genetics, Hôpital Femme Mère Enfant, Hôpitaux de Lyon, Lyon 69677, France; [11]Department of Medical Genetics, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, the Netherlands; [12]Clinical Genetics Unit, Department of Women's and Children's Health, University of Padova, Padova 35128, Italy; [13]Cardiovascular Pathology Unit, University Hospital of Padova, Padova 35128, Italy; [14]Population Health and Immunity Division, Walter and Eliza Hall Institute of Medical Research, University of Melbourne, Melbourne, VIC 3052, Australia; [15]Department of Medical Biology, University of Melbourne, Melbourne, VIC 3052, Australia; [16]Department of Pathology, Dunedin School of Medicine, University of Otago, Dunedin 9054, New Zealand
[17]These authors contributed equally to this work
*Correspondence: stephen.robertson@otago.ac.nz
https://doi.org/10.1016/j.ajhg.2018.04.008.

[MIM: 108120]),[9] 2A (DA2A [MIM: 193700]), and 2B (DA2B [MIM: 601680]).[10]

MYH3 encodes the embryonic heavy chain myosin MYHC-Emb,[11] a protein that assembles as homo- and hetero-dimers to form the thick filament in the sarcomere. The implication of *MYH3* in the etiopathogenesis of SCTS suggests functions that extend to osteogenesis, and indeed, Zieba et al. have recently demonstrated its role in the regulation of TGFß activity in the scleromyotome; this role is reminiscent of the mechanism that underpins the *FLNB*-deficient form of the condition.[7]

Here, we investigated the molecular basis of SCTS in a cohort of individuals with no *FLNB* variants accounting for their phenotypes. These individuals were phenotypically similar to those with SCTS caused by mutations in *FLNB* but with the notable exception of the presence of contractures that were often clinically prominent at birth. We demonstrate that, in some instances, the basis for this condition is compound heterozygosity for a truncating variant in *trans* with a non-coding variant at a splice donor site that flanks a non-coding exon in the 5′ UTR of *MYH3*. The definition of a recessive basis for this form of SCTS presents clinical challenges in the context of the discovery of monoallelic variants in *MYH3* in the diagnostic setting not only because both dominant and recessive conditions can result from variation at this same locus and result in similar phenotypes but also because the modifying splice variant is non-coding and is not captured on most clinically deployed exome-capture platforms.

Subjects and Methods

Recruitment, Diagnosis, and Consent of Affected Individuals

Participants in this study were enrolled through physician-initiated referral and consented under a protocol approved by the Southern and Multi-Regional Health and Disability Ethics Committee of New Zealand. A diagnosis of SCTS was made after the observation of multiple contiguous vertebral fusions on radiographs in the absence of significant deformity of the ribs or appendicular skeleton. Carpal and tarsal fusions were also obligatory features for a definitive diagnosis, although some individuals who did not have this sign were enrolled. For all individuals, the exons and exon-intron boundaries of *FLNB* were subjected to Sanger sequencing, which identified no pathogenic variants.

Exome and Genome Sequencing

DNA was extracted from leucocytes and enriched with exonic regions with Roche Nimblegen SeqCap EZ2 for individuals 1–3, with Agilent SureSelect Human All Exon V4 for the parents of individuals 1, 2, 4, and 5, and with Agilent SureSelect Human All Exon V5+UTR Capture Kit for all other samples. Sequencing was performed with 100 bp paired-end reads on the Illumina HiSeq 2000 sequencing platform. Mean coverage was 61×, and a mean >94% of the exome had a minimum depth of 10×. Genome sequencing was performed on samples from individuals 1–3 on the Illumina X-10 platform at the Kinghorn Centre for Clinical Genomics. The mean coverage was 40×, and >96% of the genome had a minimum depth of 30×.

Sequence Analysis

Raw data from the whole-exome sequencing (WES) and whole-genome sequencing (WGS) were processed according to the Genome Analysis Toolkit (GATK) Best Practice Guidelines. The Burrows-Wheeler Aligner (MEM algorithm) v.0.7.12, Picard v.1.140, and GATK were used for realignment, marking of duplicate reads, and recalibration of base quality scores. Individual variant calls were undertaken with the GATK Haplotype Caller, followed by multi-sample genotyping and recalibration of variant quality scores. Annotation of gene context was carried out with SnpEff and GATK VariantAnnotator, and allele frequencies were obtained from 1000 Genomes phase 1[12] and the Exome Aggregation Consortium (ExAC) Browser.[13] GATK SelectVariant and SnpSift v.4.1L (SnpEff) were used for sequential filtering of the multi-sample VCF file. Common and false-positive variants with an allele count > 2 in a control cohort that was processed in parallel with the samples were removed from exome sequence datasets. Rare (minor allele frequency [MAF] < 0.001) variants in coding regions, including splice sites for all affected individuals and 5′ and 3′ UTRs for individuals 6–10 and the parents of individuals 6–8, were selected according to the frequencies or absence of the variants in 1000 Genomes and/or the ExAC Browser. All variants of interest were confirmed by Sanger sequencing.

SNP Array Analysis

Array genotyping was performed with the Omni1-Quad and Omni5 chip platforms (Illumina), and genotypes were called with GenomeStudio (Illumina) and processed for Mendelian errors with Linkdatagen.[14] For estimation of inbreeding coefficients, Linkdatagen was used to select one maximally informative SNP every 0.5 cM (average MAF = 0.48 in the CEU [Utah residents with ancestry from northern and western Europe] population from 1000 Genomes) to minimize linkage disequilibrium between markers. For analysis of haplotype sharing, contiguous stretches of shared markers were called with an in-house script and then sorted and plotted with R software.

Deletion Prediction

GenomeStrip[15] and Lumpy[16] were used to identify the deletion in individuals 7 and 8, and then the deletion was verified by visual inspection with the Integrative Genomics Viewer. The breakpoints forming the deletion were confirmed by bridging PCR using primers situated on either side of the deletion site (as predicted by Lumpy).

Transcript Splicing Analysis

Both wild-type and c.−9+1G>A *MYH3* minigenes were constructed to contain exons and introns spanning exons 1–5 and cloned into the pcDNA3.1V5/His expression vector. Plasmid was transfected into HEK293FT cells with Lipofectamine 2000 (ThermoFisher) and cultured in DMEM and 10% fetal calf serum in 5% CO_2 at 37°C. RNA was extracted after 20 hr of growth with the Nucleospin RNA Plus Kit (Machery Nagel) and converted to cDNA with Superscript III (ThermoFisher). Primary amplicon PCR was carried out with primers within the vector-specific 5′ UTR and exon 3 of *MYH3*. Primers also contained Illumina adaptor linkers; amplification was run to the early log phase. Products were purified on Agencourt AMPure XP (Beckman Coulter), and

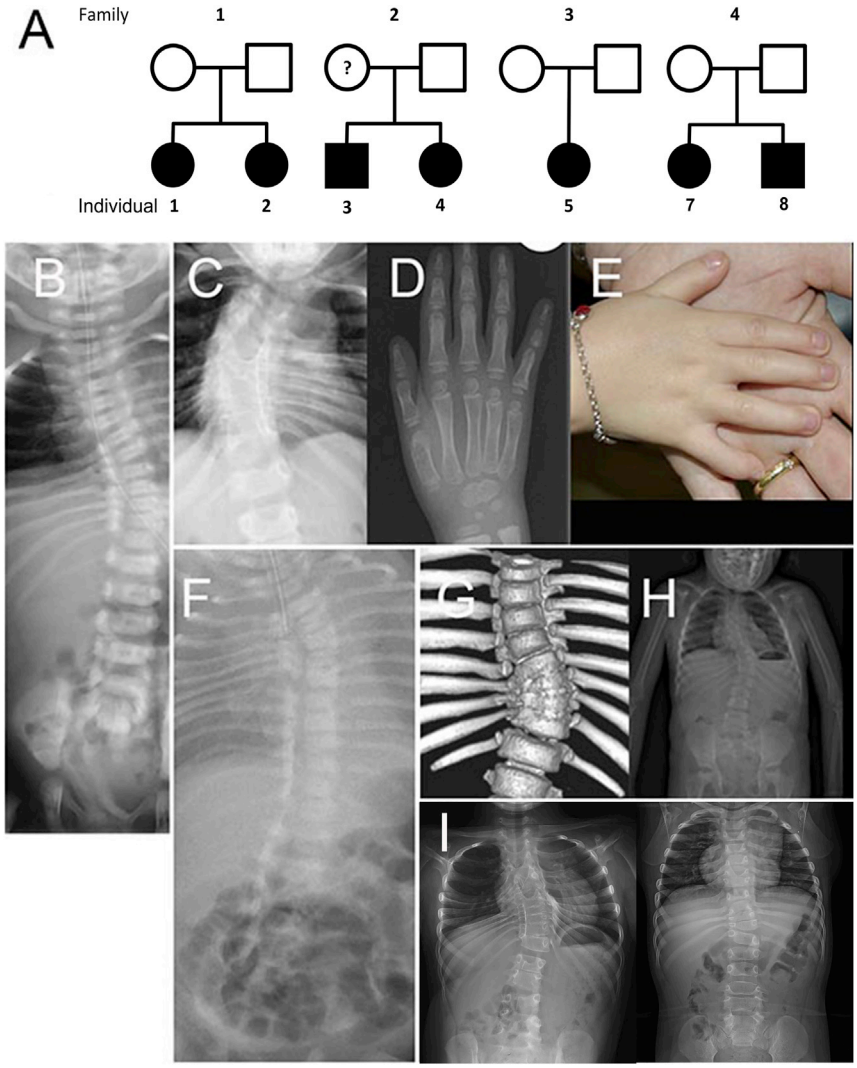

Figure 1. Pedigree Structures and Phenotypes of Individuals with SCTS
(A) Pedigrees of individuals 1–5, 7, and 8.
(B) Individual 1 in the neonatal period demonstrates scoliosis but no radiological evidence of vertebral fusion.
(C–E) Individual 2 shows scoliosis and vertebral fusions (C), capitate-hamate synostosis (D), and camptodactyly with mild interdigital webbing (E).
(F) Individual 5 with scoliosis.
(G and H) A 3D reconstructed computed tomography scan (G) and antero-posterior radiograph (H) of individual 6 with scoliosis and fusions involving the thoracic vertebrae.
(I) Individuals 7 (left) and 8 (right) with scoliosis and fusions involving the thoracic vertebrae.

concentration was determined on a Qubit with the Qubit dsDNA HS Assay Kit (ThermoFisher). Second-round index PCR was carried out for ten cycles, and products were purified as above. Samples were run with the MiSeq Reagent Nano Kit v.2 (500 cycles; Illumina). The Galaxy software hub was used for primary data analysis,[17] and paired-end-read FASTA files were imported into Sequencher for the alignment of splice variants (Sequencher v.5.3 DNA sequence analysis software, Gene Codes Corporation).

Fluorescence-Based Translational Assay

The 5′ UTRs of *MYH3* splice variants were cloned directly after the first cytomegalovirus (CMV) promoter of the dual-expression construct pSF-CMV-CMV-SbfI (Oxford Genetics) fused to mCherry. EGFP was cloned into the second CMV promoter to act as an internal control. Constructs were transfected into HEK293FT cells and cultured as above. Relative fluorescence was measured on a Victor plate reader (PerkinElmer) with excitation filters 485/14 nm (GFP) and 555/38 nm (mCherry) and emission filters 535/25 nm (GFP) and 632/45 nm (mCherry). The translational efficiency of each *MYH3* 5′ UTR splice variant was calculated as the relative mCherry fluorescence normalized to that of the dual-expressed EGFP. Experiments were carried out in triplicate on four independent plasmid clones for each construct.

Clinical Descriptions of Affected Individuals

Affected individuals (n = 16) were recruited into this study by physician-initiated referral. All affected persons had a full set of skeletal radiographs reviewed by the referring physician and S.P.R. Individual 9 has been previously described by Wiles et al.[18] (case 3). Inclusion in the study required the demonstration of vertebral fusions without evidence of a segmentation defect in addition to fusion of the carpal and/or tarsal bones. Three males and seven females with SCTS and *MYH3* variants were studied in this work. Individuals 1–4, 7, and 8 are from families with two affected individuals (Figure 1). There was no other history of the disorder in any family, except that the mother of two affected children (individuals 3 and 4) had bilateral camptodactyly of the fifth fingers. The other four individuals had simplex presentations and no previous family history of SCTS or scoliosis. Individuals 1–5 were all born in the Netherlands. Two of those families (individuals 1 and 2 and individual 5 [simplex presentation]) were born in the same region and shared a surname, although they were not knowingly related to one another. Individuals 3 and 4 also reported an unspecified history of distant consanguinity.

Results

Investigation of a Recessive Basis for This Form of SCTS: Exclusion of *FLNB*

Given the ascertainment of two families with sibling recurrence of SCTS and a third family with anecdotal evidence of a familial connection, it was hypothesized that these five affected Dutch individuals could represent a recessive form of SCTS. Genome sequencing was performed for three of the four individuals (1–3) from families with two affected children and unaffected parents. At the *FLNB* locus (chr3: 57,994,127–58,157,982; UCSC Genome Browser

hg19), there were 943 non-reference variants across these three individuals, and 24 of these were located within the exonic regions or within 50 bp of the splice sites. No individual was homozygous for any exonic variant, all variants were present in 1000 Genomes at a frequency > 0.02, and only one variant was predicted to change the protein sequence; this exceptional variant had a frequency of 0.45 in 1000 Genomes. Therefore, neither homozygosity nor compound heterozygosity for *FLNB* variants was considered a likely explanation for the SCTS phenotype in these individuals.

To further resolve the presumptive recessive basis for the condition, we sought evidence for consanguinity in these Dutch families (individuals 1–5). Parental consanguinity elevates an individual's rate of homozygosity, which is correlated with the inbreeding coefficient (f) and is dependent on the number of meioses that separate them from their shared ancestor. FEstim,[19,20] which uses a hidden Markov model to estimate homozygosity and therefore the predicted f across genome-wide marker sets, was applied to SNP array data from individuals 1 and 3–5 and their parents (no SNP array data were available for individual 2 or the parents of individual 5). f was estimated to be 0 in all individuals except for the mother of individuals 1 and 2 ($f = 0.03$, which is approximately equivalent to that of parents who are second cousins), consistent with the consanguinity reported for her family, but there was no detectable elevation in the rate of homozygosity in her affected daughter (individual 1). A non-parametric analysis of homozygosity by state for runs of SNP array markers in individuals 1, 3, and 4 (individual 5 was not analyzed because of a lack of parental DNA) likewise found that these individuals did not have significantly more homozygosity for long (>1 cM) haplotypes than healthy, unrelated Europeans ($f(1,116) = 0.31$, $p = 0.579$, ANOVA). These data indicate that if there is consanguinity between the parents of affected children in these three families, it is too remote to be detected by these methods.

Biallelic WES Analysis

We searched WES data from all affected Dutch individuals (1–5) to identify biallelic variants that comprise candidate genotypes for a recessive condition in these individuals. No rare (allele frequency < 0.01), conserved (PhastCons[21] score > 0.8 or, if PhastCons scores were not available, a PhyloP[22] score > 1) variants of predicted protein-altering effect (missense, splice-site, or truncating) were found under either a homozygous or a compound-heterozygous model, in which any combination of homozygous or heterozygous variants had to be inherited by all affected individuals (i.e., both affected siblings). A compound-heterozygous model with independent variants (not shared between families, e.g., four different alleles found between two families) in the same gene was also tested, but no two Dutch families were found to carry two rare, conserved, protein-altering variants in the same gene.

Analysis of Shared Haplotypes

Because comprehensive WES-based variant detection is limited by the capture platform and depth of each dataset, we undertook further steps to detect evidence of shared ancestry between these Dutch families. We assembled runs of contiguous SNV-array-defined alleles as haplotypes on a genome-wide scale to determine whether haplotypes were shared between the affected individuals from the Dutch families (individuals 1 and 3–5; for individual 2, SNV array data were not available). We plotted these runs by genetic size (cM) and marker number to identify extended shared genomic regions with dense SNV coverage. The higher the values of these two measures, the stronger the evidence is that an inherited segment is identical by descent (IBD). No large, shared haplotypes were identified between individual 1 and individuals 3 and 4. However, the highest-ranking haplotype that individual 1 shared with individual 5 was a 23 cM segment that overlapped 212 genes, including *MYH3* (Figure 2A). The haplotype was paternally inherited by individual 1 and was also subsequently shown to be shared by their affected sibling (individual 2).

Given the recent implication of monoallelic *MYH3* variants in some cases of SCTS[6,7] and the suspicion that recessive inheritance underpins SCTS in the families in this study, the maternally transmitted haplotypes in individuals 1 and 2 were also examined. This showed that the top maternally transmitted haplotype that individuals 1 and 2 shared with individual 5 was also inclusive of the *MYH3* locus (a 9 cM segment encompassing 28 annotated genes; Figure 2B). Many markers in all three individuals were heterozygous across this region (45% in genome-sequenced datasets aligned at more than ten reads and 60% of SNP array genotype calls), indicating that these two haplotypes are not IBD. An alternative explanation, therefore, is that both parents of individuals 1 and 2 independently share ancestry with the two parents of individual 5. Although parental DNA for individual 5 was not available, the shared surname is consistent with the notion that both fathers are ancestrally related and, by exclusion, that both mothers share a different common ancestor. The occurrence of these presumptive IBD segments around both *MYH3* alleles led to the hypothesis that compound heterozygosity for recessive alleles leads to SCTS in these individuals.

To test this hypothesis, we sought *MYH3* variants across all 16 SCTS-affected individuals without pathogenic *FLNB* variants by using WES data. Ten of these individuals were found to carry rare or private *MYH3* variants that were predicted to be protein altering (missense, nonsense, or splice-site variants or either frameshift or in-frame indels; Table 1). Nine individuals, including individuals 1–5 from the Netherlands, had a truncating variant. The paternal allele for individuals 1 and 2 had a splice-site variant that was shared by individual 5. The other Dutch family with sibling recurrence had a different splice-site variant transmitted by the minimally affected mother.

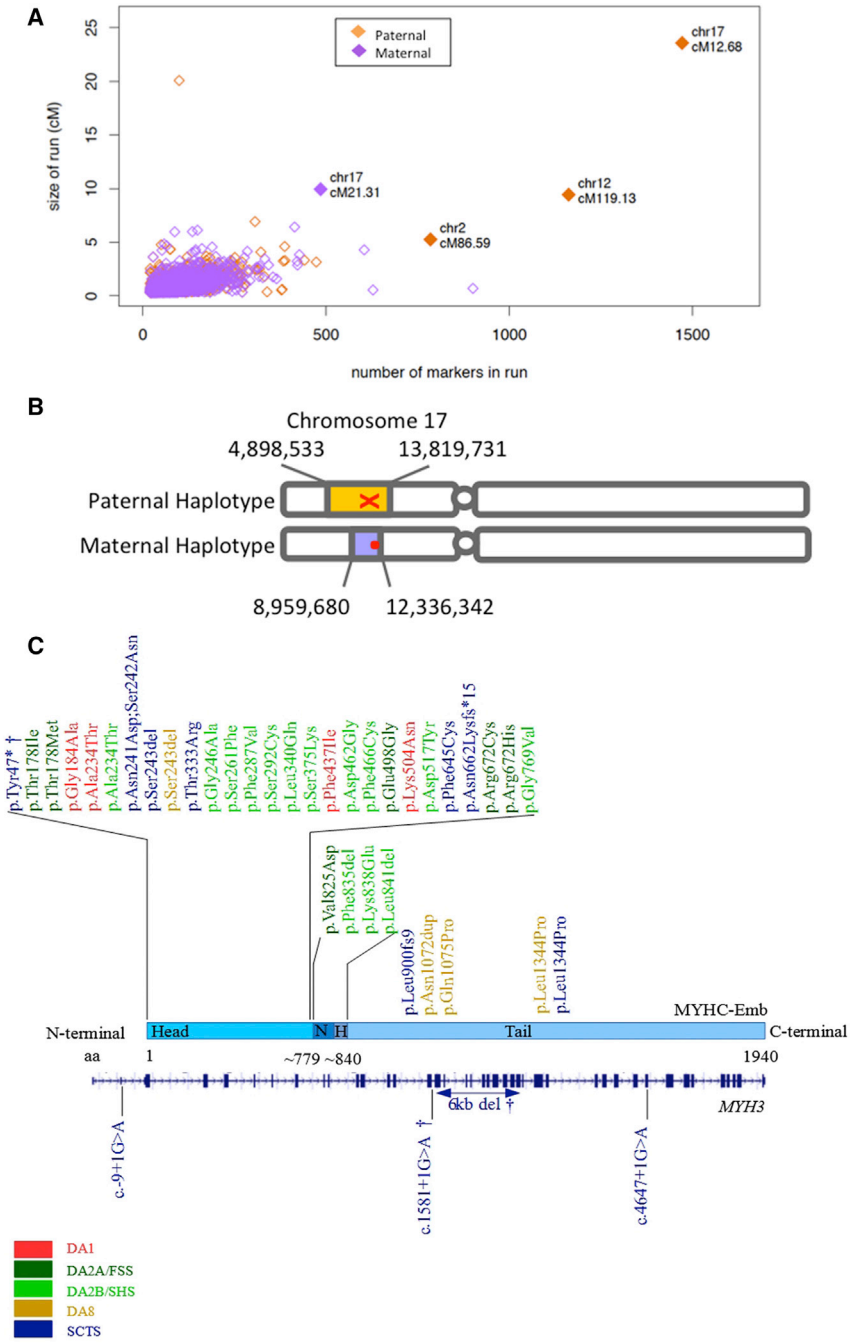

Figure 2. Mapping and Identification of Pathogenic Variants in *MYH3*

(A) Scatterplot of all genomic regions that are identical by state (as measured by continuous sharing of at least one allele) between individuals 1 and 5. Points are filled if they are in the top five most extended regions, as measured by both the number of markers and genetic distance. Labels for these points report the chromosome and centiMorgan coordinates. Paternally inherited regions in individual 1 are represented in orange, and maternally inherited regions are in purple.

(B) Shared haplotypes containing the two *MYH3* variants in individuals 1, 2, and 5 via a large, shared paternal haplotype (orange) and a shared maternal haplotype (purple). The location of *MYH3* and the truncating variant is represented by an X, and the 5′ UTR variant is represented by a circle.

(C) Schematic of *MYH3*, MYHC-EMB, and published mutations[6–10,23–26] leading to four different phenotypes. Abbreviations are as follows: DA1, distal arthrogrysosis type 1; DA2A, distal arthrogryposis type 2A; FSS, Freeman-Sheldon syndrome; DA2B, distal arthrogryposis type 2B; SHS, Sheldon-Hall syndrome; DA8, distal arthrogryposis type 8; SCTS, spondylocarpotarsal synostosis syndrome; †, mutation with c.−9+1G>A (rs557849165) on the other allele; N, neck; H, hinge; and aa, amino acid.

Siblings 7 and 8 (Figure 1) and their unaffected father carried a 6 kb intragenic deletion beginning within intron 12 and continuing to part way through exon 26 (chr17: 10,546,704–10,546,532; UCSC Genome Browser hg19; Figure S1). Individual 10 had three missense variants altering two adjacent amino acids in phase with each other (c.[721A>G;724_725delinsAA], p.[Asn241Asp];[Ser242Asn] [GenBank: NM_002470.3]). In total, where parental DNA was available, the inheritance pattern of these presumptive loss-of-function mutations indicated that they were inherited from a healthy parent in seven individuals. We identified no instances where affected individuals had biallelic deleterious variants or were homozygous for predicted or suspected loss-of-function variants.

The wider cohort of SCTS-affected individuals without pathogenic variants in *FLNB* (n = 16) had significantly more *MYH3* truncating variants than presumptively healthy control individuals from publicly available databases (Supplemental Data). In the Genome Aggregation Database (gnomAD),[13] 85 high-impact variants (stop, frameshift, and splice-site variants) are recorded in 224 individuals. In the coding sequence, there are 81 high-impact variants in 129 individuals. The most frequent of these truncating variants (c.4522C>T [p.Gln1508Ter] [GenBank: NM_002470.3]) within the exonic regions has an overall frequency of 1/13,678. No individuals homozygous for truncating variants are reported in this database. Decipher[27] also reports four individuals who have deletions encompassing *MYH3* without a documented skeletal phenotype.

Our observations that individuals 1, 2, and 5 carry the same *MYH3* splice-site variant on a shared paternal haplotype and also share a maternal haplotype in the context of

Table 1. Clinical Phenotypes and Variants in *MYH3*

	This Paper												
	Individual 1 (Family 1)	**Individual 2 (Family 1)**	**Individual 3 (Family 2)**	**Individual 4 (Family 2)**	**Individual 5 (Family 3)**	**Individual 6**	**Individual 7 (Family 4)**	**Individual 8 (Family 4)**	**Individual 9**	**Individual 10**	**Carapito et al.[6]**	**Zieba et al.[7]**	**Chong et al.[a,8]**
Origin	Dutch	Dutch	Dutch	Dutch	Dutch	North African	Bangladeshi	Bangladeshi	Australian	French	U	U	U
Sex	F	F	M	F	F	F	F	M	F	M	3 F, 1 M	2 F, 1 M	4 F, 4 M
MYH3 allele 1	c.4647+1G>A	c.4647+1G>A	c.1581+1G>A	c.1581+1G>A	c.4647+1G>A	c.141T>G	deletion of intron 12 and exon 25	deletion of intron 12 and exon 25	c.1986_1990delTTTAA	c.[721A>G;724_725delinsAA]	c.4031T>C, c.998C>G	c.1934T>G, c.727−729del, c.2699delT	c.3214_3216dup, c.3224A>C, c.727−729del
MYH3 allele 2	c.−9+1G>A	c.−9+1G>A	NAD	NAD	c.−9+1G>A	c.−9+1G>A	c.−9+1G>A	c.−9+1G>A	NAD	NAD	NAD	NAD	NAD
Vertebral fusion	T	T	C, T	C, T	T, L	T, L, S	T, L	C (T and L posterior arches only)	C, T, L	T	4/4	3/3	5/6
Carpal-tarsal fusion	yes	yes	yes	yes	U	U	no	no	yes	yes	3/3	3/3	U
Dysmorphism	yes	no	no	yes	yes	yes	no	no	no	U	4/4	0/3	U
Short neck	yes	yes	no	no	yes	yes	yes	u	yes	U	4/4	2/3	8/8
Scoliosis	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	3/4	2/3	7/8
Webbing	E, K, Fi	F	N, Fi	N, Fi	N, Fi	N	no	E	N	U	N (2/4), E (1/4), K (1/4)	E (1/3), Fi (1/3)	N (6/7), Fi (1/7), H (2/7), E (5/6)
Contractures	E, K, Sh	E, H, K, hip dysplasia	Fi	Fi	E, K, H	N	E, K	E, Fi	E	U	Fi (3/4), E (2/4)	Fi (1/3), E (1/3)	Fi (7/7), E (6/7), H (1/2)
Absent finger flexion creases	yes	yes	yes	yes	yes	no	no	yes	no	U	2/4	NA	7/7
Additional anomalies	none	cleft palate	none	none	none	none	none	club foot	none	none	none	none	cleft palate, craniosynostosis, inguinal hernia

Abbreviations are as follows: F, female; M, male; C, cervical; T, thoracic; L, lumbar, S, sacral; N, neck, E, elbows; Fi, fingers; H, hips; K, knees; Sh, shoulders; U, unknown; and NAD, no abnormality detected.[a]Individuals with clinical information and confirmed *MYH3* pathological variants.

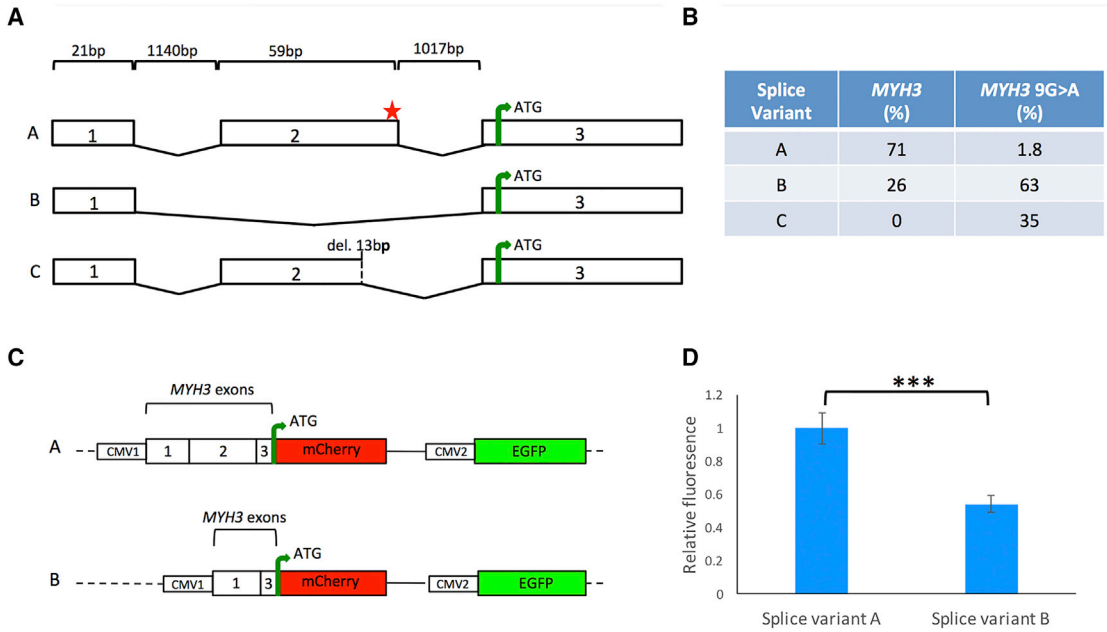

Figure 3. Analysis of Splice Variants and Translational Efficiency for Wild-Type and c.−9+1G>A *MYH3* Alleles

(A) Diagram of exons 1–3, introns, and observed spice variants designated as A, B, or C; the position of the *MYH3* c.−9+1G>A variant is indicated by a red star.

(B) MiSeq data showing the proportions of observed splice variants A–C as a percentage of the total reads for both wild-type *MYH3* (687 paired reads) and c.−9+1G>A *MYH3* (749 paired reads; abbreviated as 9G>A). Variants were excluded if they represented ≤1% of the total reads.

(C) Diagram of dual-expression plasmid constructs designed so that the variant *MYH3* 5′ UTRs are fused to the mCherry reporter and normalized to a dual-expressed EGFP. Construct A includes *MYH3* exons 1 and 2 and 6 bp of exon 3 up to the ATG codon, and construct B includes *MYH3* exon 1 spliced directly to 6 bp of exon 3.

(D) Relative translational efficiency of *MYH3* 5′ UTR splice variants transfected into HEK293FT cells is represented as relative mCherry fluorescence normalized to that of EGFP. ***$p < 0.0001$, two-tailed t test. Error bars represent standard deviation from the mean.

sibling recurrence suggest that a maternally derived pathogenic variant remains to be defined. We therefore performed WGS on individuals 1–3 and then prioritized and filtered *MYH3* variants for high impact (SnpEff). Individuals 1 and 2 shared a variant (c.−9+1G>A) in the 5′ UTR of *MYH3* (Table 1; Figure 3). This variant (rs557849165) affects the canonical donor splice site at the untranslated boundary between exon 2 and intron 2. It is present in two individuals in the 1000 Genomes Project cohort (2/2,535) and in 85 heterozygous (but not homozygous) individuals in gnomAD (allele frequency = 0.003), making it the most common high-impact variant at the *MYH3* locus. In individuals 1 and 2, the c.−9+1G>A variant was maternally inherited on the haplotype shared with individual 5.

We screened the remainder of the SCTS cohort with no pathogenic *FLNB* variants but with *MYH3* high-impact variants (Table 1) for this c.−9+1G>A variant and found it in six of ten instances (individuals 1, 2, and 5–8). As with individuals 1 and 2, the c.−9+1G>A variant was invariably in *trans* with the truncating mutation in instances where parental DNA was available (individuals 1, 2, and 6–8). In individual 5, who also carried the c.−9+1G>A variant, parental DNA was not available, but the variant was confirmed to be in *trans* with the *MYH3* splice-site variant through haplotypic analysis. In total, of the 16 individuals in this cohort, nine have a truncating *MYH3* mutation, and of those nine individuals, six also have the c.−9+1G>A variant (rs557849165) in *trans*, a statistically significant enrichment compared with the frequency in population control individuals ($p < 10^{-34}$; see Supplemental Note). Siblings 3 and 4 did not carry the c.−9+1G>A variant even though they inherited a splice-site variant at an extremely conserved base from their mother, who was affected only by camptodactyly. From the WGS data, individual 3 also carried three moderate-impact variants in *MYH3* and 18 low-impact or modifier (intragenic, 5′ UTR, upstream, and/or downstream) variants that were not present in individual 1 or 2 (Table S1). The significance of these variants in the development of SCTS remains unknown, and further study is necessary to determine whether the splice-site variant in this instance acts in a dominant manner with variable expressivity or whether a second recessive variant was inherited paternally.

Although the c.−9+1G>A variant is statistically associated with the SCTS phenotype and it alters a canonical splice donor site, we sought further functional evidence to implicate it in recessive SCTS. *MYH3* exons 1–5, together with the intervening introns, were cloned into a mammalian expression vector, and the rs557849165 variant was introduced by site-directed mutagenesis (Figure 3A). The construct was transfected into HEK293FT cells, and cDNA was prepared and subjected to massively parallel

sequencing on a MiSeq instrument. Three predominant transcript isoforms were observed (Figures 3A and 3B). Splicing inclusive of exons 1–3 (spliceform A) was observed in 71% of reads from the wild-type construct, and 26% of reads exhibited skipping of exon 2 (spliceform B). When rs557849165 was introduced, negligible splicing of exons 1–3 was observed, and the transcript excluding exon 2 predominated. A third splice variant was also observed in 35% of transcripts. These transcripts deployed a cryptic splice donor site 13 bp 5′ of the canonical site (spliceform C).

To determine whether this alteration in 5′ UTR structure in the presence of rs557849165 resulted in a net change in translational efficiency, we cloned spliceforms A and B fused to m-cherry and co-expressed them from a plasmid also expressing eGFP to normalize for transfection efficiency (Figure 3C). Quantitation of the ratio of red to green fluorescence served as a readout of translational efficiency for each 5′ UTR. These data demonstrate that the 5′ UTR containing spliceform B, equivalent to the most prevalent spliceform in the presence of rs557849165, displayed a translational efficiency of 54% relative to that of the wild-type ($p < 0.0001$, two-tailed t test; Figure 3D), confirming that this variant confers a deleterious effect to the translation of *MYH3* transcripts.

Discussion

Previously, *FLNB* was the only gene implicated in the causation of SCTS and then only in recessive presentations of the disorder. However, recent work has shown that dominantly inherited *MYH3* variants can also lead to SCTS. To these data we now add evidence of another recessive form of the condition, caused by compound heterozygosity for variants in *MYH3*.

The existence of many control individuals who are heterozygous for truncating variants in *MYH3* must be reconciled with the implication of truncating mutations in the development of SCTS both in this paper and in two recent reports.[6,7] In those publications, it was suggested that this incongruity could be explained by truncations in different parts of the transcript either because some 3′ truncations lead to the production of a protein that retains some function or because frameshifts could confer dominant-negative effects in affected individuals. These explanations might apply in some instances (e.g., individual 9 in this study and one previously reported individual[7]) where only a single truncating mutation has been found to possibly be sufficient to cause SCTS. However, the presence in our cohort of multiple transmissions of truncating mutations from healthy parents to affected children and of sibling recurrence in three families and the presence in public databases of presumptively healthy individuals carrying *MYH3* deletions that do not confer a skeletal phenotype are further evidence that null alleles in *MYH3* are usually insufficient in isolation to produce the SCTS phenotype. Here, we have shown that in over a third of our cohort of individuals without pathogenic *FLNB* variants, a hypomorphic allele inherited in *trans* with truncating alleles confers a recessive form of SCTS. Of the nine individuals with a truncating allele, six share the same 5′ UTR variant in *trans* with truncating mutations, whereas three (siblings 3 and 4 and individual 9) have no identified equivalent candidate variant on the other allele. The recurrence of the disease in siblings 3 and 4 (a male and a female) suggests that recessive inheritance remains a plausible explanation in this instance, although in the absence of the description of further individuals carrying the same variants, it will be difficult to identify a causative variant from the list of current candidates identified by WGS given that none have a predicted high impact by SnpEff. Similarly, for individual 9, in whom only a monoallelic truncating *MYH3* variant has been identified and no WGS is yet available, there could still be another *trans* variant that has not been detected by WES. Genetic counseling should therefore proceed with a recessive etiology as a possible mode of inheritance in this instance.

These observations aside, this and previous publications have shown clear evidence that an autosomal-dominant form of SCTS with, in some instances, considerable overlap with DA8 does exist and is caused by monoallelic mutations in *MYH3*.[6,7] Individual 10 had three variants in *cis*, which are predicted to change two highly conserved amino acids in the head of the myosin (Asn241 and Ser242), and did not have the c.−9+1G>A variant. This individual might be considered to have a mild form of DA8 (or alternatively, autosomal-dominant SCTS), given that in a previously published family with ten affected members over five generations,[8] three individuals were reported to carry a similar variant, c.727_729del (p.Ser243del). This is the only reported family with a mutation that affects exons encoding the head of the myosin and leads to a DA8 phenotype. Notably, these individuals did not have the same dysmorphic features as the other two families presented in this same publication; those families carried missense mutations affecting the tail of MYHC-Emb. Zieba et al.[7] also reported a three-generation, SCTS-affected family where one individual (R07-183B) had a p.Ser243del variant. Isidor et al.[5] reported a family with an autosomal-dominant SCTS transmission that was later found to have a variant resulting in a substitution in the head domain of myosin.[6] The variants found in all of these families hold the potential to confer a dominant-negative effect and further suggest that in the context of *MYH3* missense mutations, autosomal-dominant SCTS or DA8 could denote phenotypes at different positions along a phenotypic spectrum. A recent report that adds to this evolving understanding of genotype-phenotype correlations in this disorder described an individual with DA2B and partial fusions of cervical, thoracic, and lumbar vertebrae.[23] All mutations described in DA1, DA2A, and DA2B are missense or in-frame deletions, which is consistent with a dominant-negative effect or gain-of-function mechanism underlying these disorders.

Some genotype-phenotype correlation in DA1, DA2A, and DA2B has already been established. All mutations causing DA1 affect residues located in the head of MYHC-Emb (Figure 2C), and those leading to DA2A typically lie close to a groove between two parts of the ATP binding site.[10] Mutations causing DA2B affect the head, neck, or hinge of MYHC-Emb (Figure 2C) and mostly lie near the surface of the protein.[10] However, there is a high degree of phenotypic variability such that individuals with the same mutation can be mildly or severely affected, sometimes with a different type of DA (e.g., DA1 and DA2B caused by c.700G>A [p.Ala234Thr][24]).

The likelihood that the conditions formerly described as DA8 and SCTS (and the recessive entity we describe here to exhibit phenotypic overlap between the two) are phenotypically discrete conditions is further diminished by our observation that several members of our cohort (individuals 1, 2, 5, and 8, who all carry a truncating mutation and the 5′ UTR variant) were provisionally diagnosed with arthrogryposis at birth. In our cohort, most individuals with a phenotype ascribable to an *MYH3* genotype presented with contractures affecting multiple joints (Table 1). Consequently, the presence of contractures could be a discriminating feature differentiating the SCTS phenotype due to variants in *FLNB* from those due to biallelic variants in *MYH3*.

In this cohort, no individual was shown to be homozygous or compound heterozygous for a truncating variant. Moreover, no individuals in the ExAC Browser are homozygous for a loss-of-function variant. Given that MYHC-Emb is key to skeletal muscle development,[28] its absence due to truncating variants in *MYH3* could potentially be lethal during embryogenesis. This observation, together with our functional validation of the c.−9+1G>A variant, demonstrates that this 5′ UTR variant is a hypomorphic allele, and the presence of a truncating variant in *trans* could reduce the amount of MYHC-Emb below a threshold sufficient to lead to SCTS. The compound inheritance of a null allele and a hypomorphic allele as a pathogenic mechanism has precedent in thrombocytopenia absent radius syndrome[29] (MIM: 274000) and phenotypes related to variants at the *CFTR* locus.[30]

These findings bring two cautionary issues to light in the context of molecular diagnosis and genetic counseling for SCTS or DA8. Evaluation of the coding regions of *MYH3* is complicated because the first two exons, including the site of the c.−9+1G>A variant, are not covered by commonly deployed exome sequencing capture platforms, explaining why it was not observed in our initial round of exome sequencing. Second, the isolated observation of a potentially pathological variant in *MYH3* in the absence of pedigree information to confirm or preclude vertical transmission of either the phenotype or the variant is problematic. It is important to note that our data indicate that the absence of either the c.−9+1G>A variant or another clearly pathological hypomorphic variant in *trans* (e.g., individuals 9 and 10 in this study) is not exclusionary evidence of a recessive etiology for the condition.

Explaining how an embryonic muscle protein can produce a predominantly postnatal bone phenotype remains a challenge. MYHC-Emb is present in bone, but not cartilage, in fetal life.[8] Zieba et al.[7] further determined that in embryonic and postnatal day 1 mice, the protein encoded by *Myh3* is found in osseous tissues and in small muscles of the cervical and thoracic spine, which stabilize intervertebral joints. However, some features of SCTS (carpal and tarsal synostosis) appear only later in life, suggesting ongoing expression of *MYH3* long beyond the immediate postnatal period. Other muscle proteins (e.g., TNNI2) have been shown to localize to bone tissue,[31] and disruption of cross-talk between bone and skeletal muscle could also plausibly contribute to the evolution of the SCTS phenotype.[28] Further investigation on the diversity of functions of MYHC-Emb in skeletogenesis is clearly warranted.

Accession Numbers

The ClinVar accession numbers for the variants reported in this paper were not available from ClinVar as of the date this article was finalized for publication; please contact the corresponding author for the numbers.

Supplemental Data

Supplemental Data include one figure, one table, and one Supplemental Note and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.04.008.

Acknowledgments

The authors thank the families for their involvement in this study. Katherine Smith is thanked for assisting with the Linkdatagen analysis. S.R.C.-C. was supported by the Kirsty McDermott Memorial Scholarship from Curekids, and S.P.R. was supported by Curekids New Zealand.

Declaration of Interests

The authors declare no competing interests.

Received: February 23, 2018
Accepted: April 13, 2018
Published: May 24, 2018; corrected online July 27, 2019

Web Resources

1000 Genomes Project, http://www.internationalgenome.org/
Decipher, http://decipher.sanger.ac.ukWorkings
ExAC Browser, http://exac.broadinstitute.org/
Galaxy, https://usegalaxy.org/
GenBank, https://www.ncbi.nlm.nih.gov/genbank/
Gene Codes Corporation, http://www.genecodes.com/
gnomAD, http://gnomad.broadinstitute.org/
OMIM, http://www.omim.org/
UCSC Genome Browser, https://genome.ucsc.edu/

References

1. Coêlho, K.E., Ramos, E.S., Felix, T.M., Martelli, L., de Pina-Neto, J.M., and Niikawa, N. (1998). Three new cases of spondylocarpotarsal synostosis syndrome: clinical and radiographic studies. Am. J. Med. Genet. *77*, 12–15.
2. Langer, L.O., Jr., Gorlin, R.J., Donnai, D., Hamel, B.C., and Clericuzio, C. (1994). Spondylocarpotarsal synostosis syndrome (with or without unilateral unsegmented bar). Am. J. Med. Genet. *51*, 1–8.
3. Steiner, C.E., Torriani, M., Norato, D.Y., and Marques-de-Faria, A.P. (2000). Spondylocarpotarsal synostosis with ocular findings. Am. J. Med. Genet. *91*, 131–134.
4. Krakow, D., Robertson, S.P., King, L.M., Morgan, T., Sebald, E.T., Bertolotto, C., Wachsmann-Hogiu, S., Acuna, D., Shapiro, S.S., Takafuta, T., et al. (2004). Mutations in the gene encoding filamin B disrupt vertebral segmentation, joint formation and skeletogenesis. Nat. Genet. *36*, 405–410.
5. Isidor, B., Cormier-Daire, V., Le Merrer, M., Lefrancois, T., Hamel, A., Le Caignec, C., David, A., and Jacquemont, S. (2008). Autosomal dominant spondylocarpotarsal synostosis syndrome: phenotypic homogeneity and genetic heterogeneity. Am. J. Med. Genet. A. *146A*, 1593–1597.
6. Carapito, R., Goldenberg, A., Paul, N., Pichot, A., David, A., Hamel, A., Dumant-Forest, C., Leroux, J., Ory, B., Isidor, B., and Bahram, S. (2016). Protein-altering *MYH3* variants are associated with a spectrum of phenotypes extending to spondylocarpotarsal synostosis syndrome. Eur. J. Hum. Genet. *24*, 1746–1751.
7. Zieba, J., Zhang, W., Chong, J.X., Forlenza, K.N., Martin, J.H., Heard, K., Grange, D.K., Butler, M.G., Kleefstra, T., Lachman, R.S., et al. (2017). A postnatal role for embryonic myosin revealed by *MYH3* mutations that alter TGFβ signaling and cause autosomal dominant spondylocarpotarsal synostosis. Sci. Rep. *7*, 41803.
8. Chong, J.X., Burrage, L.C., Beck, A.E., Marvin, C.T., McMillin, M.J., Shively, K.M., Harrell, T.M., Buckingham, K.J., Bacino, C.A., Jain, M., et al.; University of Washington Center for Mendelian Genomics (2015). Autosomal-dominant Multiple Pterygium Syndrome is caused by mutations in *MYH3*. Am. J. Hum. Genet. *96*, 841–849.
9. Alvarado, D.M., Buchan, J.G., Gurnett, C.A., and Dobbs, M.B. (2011). Exome sequencing identifies an *MYH3* mutation in a family with distal arthrogryposis type 1. J. Bone Joint Surg. Am. *93*, 1045–1050.
10. Toydemir, R.M., Rutherford, A., Whitby, F.G., Jorde, L.B., Carey, J.C., and Bamshad, M.J. (2006). Mutations in embryonic myosin heavy chain (*MYH3*) cause Freeman-Sheldon syndrome and Sheldon-Hall syndrome. Nat. Genet. *38*, 561–565.
11. Schiaffino, S., Rossi, A.C., Smerdu, V., Leinwand, L.A., and Reggiani, C. (2015). Developmental myosins: expression patterns and functional significance. Skelet. Muscle *5*, 22.
12. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., McVean, G.A.; and 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.
13. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
14. Bahlo, M., and Bromhead, C.J. (2009). Generating linkage mapping files from Affymetrix SNP chip data. Bioinformatics *25*, 1961–1962.
15. Handsaker, R.E., Van Doren, V., Berman, J.R., Genovese, G., Kashin, S., Boettger, L.M., and McCarroll, S.A. (2015). Large multiallelic copy number variations in humans. Nat. Genet. *47*, 296–303.
16. Layer, R.M., Chiang, C., Quinlan, A.R., and Hall, I.M. (2014). LUMPY: a probabilistic framework for structural variant discovery. Genome Biol. *15*, R84.
17. Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. *44* (W1), W3–W10.
18. Wiles, C.R., Taylor, T.F., and Sillence, D.O. (1992). Congenital synspondylism. Am. J. Med. Genet. *42*, 288–295.
19. Leutenegger, A.-L., Prum, B., Génin, E., Verny, C., Lemainque, A., Clerget-Darpoux, F., and Thompson, E.A. (2003). Estimation of the inbreeding coefficient through use of genomic data. Am. J. Hum. Genet. *73*, 516–523.
20. Leutenegger, A.-L., Labalme, A., Génin, E., Toutain, A., Steichen, E., Clerget-Darpoux, F., and Edery, P. (2006). Using genomic inbreeding coefficient estimates for homozygosity mapping of rare recessive traits: application to Taybi-Linder syndrome. Am. J. Hum. Genet. *79*, 62–66.
21. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.
22. Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. *20*, 110–121.
23. Scala, M., Accogli, A., De Grandis, E., Allegri, A., Bagowski, C.P., Shoukier, M., Maghnie, M., and Capra, V. (2018). A novel pathogenic *MYH3* mutation in a child with Sheldon-Hall syndrome and vertebral fusions. Am. J. Med. Genet. A. *176*, 663–667.
24. Beck, A.E., McMillin, M.J., Gildersleeve, H.I., Kezele, P.R., Shively, K.M., Carey, J.C., Regnier, M., and Bamshad, M.J. (2013). Spectrum of mutations that cause distal arthrogryposis types 1 and 2B. Am. J. Med. Genet. A. *161A*, 550–555.
25. Tajsharghi, H., Kimber, E., Kroksmark, A.K., Jerre, R., Tulinius, M., and Oldfors, A. (2008). Embryonic myosin heavy-chain mutations cause distal arthrogryposis and developmental myosin myopathy that persists postnatally. Arch. Neurol. *65*, 1083–1090.
26. Al-Haggar, M., Yahia, S., Damjanovich, K., Ahmad, N., Hamada, I., and Bayrak-Toydemir, P. (2011). p.R672C mutation of *MYH3* gene in an Egyptian infant presented with Freeman-Sheldon syndrome. Indian J. Pediatr. *78*, 103–105.
27. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. Am. J. Hum. Genet. *84*, 524–533.
28. Gorski, J.P., Huffman, N.T., Vallejo, J., Brotto, L., Chittur, S.V., Breggia, A., Stern, A., Huang, J., Mo, C., Seidah, N.G., et al. (2016). Deletion of *Mbtps1* (*Pcsk8, S1p, Ski-1*) gene in osteocytes stimulates soleus muscle regeneration and increaseds size and contractile force with age. J. Biol. Chem. *291*, 4308–4322.

29. Albers, C.A., Paul, D.S., Schulze, H., Freson, K., Stephens, J.C., Smethurst, P.A., Jolley, J.D., Cvejic, A., Kostadima, M., Bertone, P., et al. (2012). Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. Nat. Genet. *44*, 435–439, S1–S2.
30. Groman, J.D., Hefferon, T.W., Casals, T., Bassas, L., Estivill, X., Des Georges, M., Guittard, C., Koudova, M., Fallin, M.D., Nemeth, K., et al. (2004). Variation in a repeat sequence determines whether a common variant of the cystic fibrosis transmembrane conductance regulator gene is pathogenic or benign. Am. J. Hum. Genet. *74*, 176–179.
31. Zhu, X., Wang, F., Zhao, Y., Yang, P., Chen, J., Sun, H., Liu, L., Li, W., Pan, L., Guo, Y., et al. (2014). A gain-of-function mutation in *Tnni2* impeded bone development through increasing Hif3a expression in DA2B mice. PLoS Genet. *10*, e1004589.