HHS Public Access

Author manuscript

*Eur J Med Genet*. Author manuscript; available in PMC 2024 December 26.

Published in final edited form as:

*Eur J Med Genet*. 2023 July ; 66(7): 104775. doi:10.1016/j.ejmg.2023.104775.

# Disruption of *FBN1* by an *Alu* element insertion: A novel genetic cause of Marfan syndrome

**Benjamin M. Helm**[a,b,*], **Amanda M. Smith**[c], **Kelly Schmit**[a], **Benjamin J. Landis**[c], **Matteo Vatta**[d], **Stephanie M. Ware**[a,c]

[a]Department of Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

[b]Department of Epidemiology, Indiana University Fairbanks School of Public Health, Indianapolis, IN, USA

[c]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, USA

[d]Invitae Corporation, San Francisco, CA, USA

## Abstract

*Alu* elements are retrotransposons with ubiquitous presence in the human genome that have contributed to human genomic diversity and health. These approximately 300-bp sequences can cause or mediate disease by disrupting coding/splicing regions in the germline, by insertional mutagenesis in somatic cells, and in promoting formation of copy-number variants. *Alu* elements may also disrupt epigenetic regulation by affecting non-coding regulatory regions. There are increasing reports of apparently sporadic and inherited genetic disorders caused by *Alu*-related gene disruption, but Marfan syndrome resulting from *Alu* element insertion has not been previously described. We report a family with classic features of Marfan syndrome whose previous *FBN1* genetic testing was inconclusive. Using contemporary next-generation sequencing and bioinformatics analysis, a pathogenic/disruptive *Alu* insertion occurring in the coding region of the *FBN1* gene was identified (c.6564_6565ins*Alu*; p. Glu2189fs) and was confirmed and specified further with Sanger sequencing. This identified the molecular basis of disease in the family that was missed using previous genetic testing technologies and highlights a novel pathogenic mechanism for Marfan syndrome. This case adds to the growing literature of Mendelian

[*]Corresponding author. Department of Medical & Molecular Genetics, Indiana University School of Medicine, 975 W, Walnut St., IB-130, Indianapolis, IN, 46202, USA. bmhelm@iu.edu (B.M. Helm).

CRediT authorship contribution statement

**Benjamin M. Helm:** Conceptualization, Resources, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Amanda M. Smith:** Investigation, Formal analysis. **Kelly. Schmit:** Resources, Writing – original draft, Writing – review & editing. **Benjamin J. Landis:** Conceptualization, Investigation, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Visualization. **Matteo. Vatta:** Conceptualization, Resources, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Stephanie M. Ware:** Conceptualization, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors have no conflicts of interest to disclose. Author MV is an employee and stockholder of Invitae.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2023.104775.

diseases caused by *Alu* retrotransposition, and it also shows the growing capability of genomic technologies for detecting atypical mutation events.

**Keywords**

FBN1; Marfan; *Alu* ; Transposable element; Thoracic aortic aneurysm; Mitral valve prolapse

## 1. Introduction

*Alu* transposable elements, a class of short interspersed elements (SINEs), have contributed significantly to human and non-human primate genome diversity and evolution (Deininger, 2011). *Alu* elements persist in modern human populations and influence diversity and health via somatic insertional events and promotion of non-allelic homologous recombination (Ade et al., 2013). Given their ubiquitous nature in the genome, they may also exert other epigenetic influences and/or act on gene expression and somatic mutagenesis (Deininger, 2011).

*Alu* elements are approximately 300-bp in length and have a common structure including an RNA polymerase III promoter region promoting the retrotransposition process (Ade et al., 2013). An estimated 11% of the human genome consists of these non-autonomous retrotransposable elements (Lander et al., 2001). They insert in the genome using a process called target primed reverse transcription (TPRT) that requires an RNA intermediate, and this results in the ability for *Alu* elements to propagate widely throughout the genome. Despite this, of the >1 million *Alu* elements in the human genome, only a fraction are transcriptionally active or have high amplification capability (Deininger, 2011). Ongoing research is highlighting the roles of *Alu* retrotransposition in somatic insertional mutagenesis and cancer. However, *Alu* elements may have broader impact by influencing genome instability by creating local genomic contexts that promote non-allelic homologous recombination and copy-number variant (CNV) formation (Deininger, 2011). It has been estimated that *Alu*-mediated CNVs may account for ~0.3% of human genetic diseases, though this is likely an underestimate (Deininger and Batzer, 1999a,b; Song et al., 2018).

Estimates suggest that there is one new *Alu* insert per 20 human births, leading to about one in every 1000 incident human genetic diseases (Deininger and Batzer, 1999a,b; Xing et al., 2009). Insertions can cause disease by disrupting a coding region or splicing signal (Deininger, 2011; Hancks and Kazazian, 2012). Examples of rare genetic diseases caused by *Alu*-mediated gene disruption include adrenoleukodystrophy (*ABCD1*), Menkes disease (*ATP7A*), X-linked severe combined immunodeficiency (*IL2RG*), Mowat-Wilson syndrome (*ZFHX1B*), familial adenomatous polyposis (*APC*), lipoprotein lipase deficiency (*LPL*), CHARGE syndrome (*CHD7*), Apert syndrome (*FGFR2*), hereditary breast/ovarian cancer (*BRCA1/2*), and neurofibromatosis type 1 (*NF1*) (Deininger, 2011; Hancks and Kazazian, 2012). There are rare reports of familial retrotransposon-mediated genetic diseases inherited from a parent with germline and somatic mosaicism (van den Hurk et al., 2007). It is possible that as genomic and bioinformatics capabilities improve, additional apparently

sporadic (and familial) disorders will be shown to be caused by inherited *Alu*-mediated genetic disruptions.

Currently there are no known published reports of Marfan syndrome (MFS) caused by an *Alu*-specific insertional disruption of the *FBN1* gene. MFS (OMIM# 154700) is a multi-systemic connective tissue disorder, with wide phenotypic variability, but it is most characterized by ocular, skeletal, and cardiovascular complications. These can include myopia, ectopia lentis, a characteristic body habitus, pectus malformations, scoliosis, and aortic dilation/dissection (Dietz, 2001). We present a novel case of familial MFS resulting from an *Alu* insertion into the coding region of *FBN1* detected with contemporary genetic testing and bioinformatics. Notably, the *Alu* insertion went undetected using previous genetic testing technologies. This report highlights a novel pathogenic mechanism for a disorder familiar to genetics professionals, and it adds to the list of currently recognized Mendelian disorders caused by *Alu* retrotransposition.

## 2. Clinical report

### Methods:

Written informed consent was obtained from the patient/family for publication, including cardiac imaging and de-identified health information as approved by the institutional review board at Indiana University (IRB#1811364611). The family was identified after the proband (see pedigree in Fig. 1, indicated by an arrow) was referred for evaluation in a multidisciplinary aortopathy clinic in 2015, and follow-up evaluations occurred in 2020 following children born to the proband. Detailed clinical evaluation, dysmorphology exam, and records review were completed by board-certified clinical genetics providers. Cardiac phenotyping was completed using standard evaluations by a pediatric cardiologist. Aortic diameters were measured in systole by transthoracic echocardiography according to practice guidelines (Lopez et al., 2010), and these were normalized for body surface to generate Z-scores based on reference data from Boston Children's Hospital (Sluysmans and Colan, 2005).

### Genetic Testing Methodology:

In the original genetic testing completed in the proband in 2015, all coding regions and splice sites of the following genes were sequenced by the Indiana University Molecular Genetics Diagnostic Laboratory (IUMGDL), including transcripts in parentheses used for analysis with human genome reference hg19: *FBN1* (NM_000138.4), *TGFBR1* (NM_004612.2), and *TGFBR2* (NM_003242.5). Genomic DNA was processed by hybridization-based target enrichment (Nextera library-preparation and TruSight One Enrichment kit, Illumina) followed by next generation sequencing (NGS) using sequencing-by-synthesis (Illumina MiSeq). Variant calls were generated using Burrows-Wheeler Aligner followed by Genome Analysis Toolkit (GATK) analysis (Li and Durbin, 2009; McKenna et al., 2010). This test detects 100% of substitution variants (95% CI = [82%, 100%]) and small indels (95% CI = [98.5%, 100%]). Sanger sequencing was used to cover bases with coverage <15x depth and for bases in some known regions of genomic segmental duplications. All clinically significant variants were confirmed by independent Sanger

sequencing. This test was developed with its performance characteristics determined by the IUMGDL (CLIA: 15D0681476).

In 2020, peripheral saliva samples were processed and evaluated by NGS using a 27-gene Aortopathy Comprehensive Panel (Invitae Corp., San Francisco, CA). Saliva was processed (Volume 600ul) using Oragene DNA-OG-500 according to manufacturer instructions. Genes were targeted with oligonucleotide baits (Agilent Technologies, Santa Clara, CA; Roche, Pleasanton, CA; IDT, Coralville, IA) to capture all coding exons, plus 10–20 bases of flanking intronic sequences and noncoding regions of clinical interest. Baits were balanced to obtain a minimum of 50x and an average of 350x depth-of-sequence read coverage. A bioinformatics pipeline was utilized that incorporated both standard and custom algorithms to identify single-nucleotide variants, small indels, large indels, structural variants with breakpoints in target sequences, and exon-level copy-number variants (CNVs). In addition to standard GATK-based alignments and analysis, validated coverage-based CNV detection algorithms designed to flag possible split-read signals were applied. Once verified, the variant call format file was updated and interpreted. Variants were classified using a point-based scoring system congruent with the system for grading evidence for pathogenicity as recommended by the American College of Medical Genetics and Genomics. Orthogonal confirmation of CNVs was performed using gene-centric array-CGH. All testing for the NGS gene panel was performed at Invitae Corp., which is accredited by the Clinical Laboratory Improvement Amendments and College of American Pathologists.

## Clinical Findings:

The following summarized the relevant clinical findings for the affected family. Fig. 1 depicts the limited pedigree and references the index case proband (Individual II-3) and his two affected offspring (Individuals III-1 and III-2).

**Individual II-3 (Proband):** This individual was initially evaluated at the request of cardiothoracic surgery when he was admitted at age 18 for ascending aorta graft with aortic root replacement with valve replacement (25 mm ATS conduit and coronary artery reconstruction (Bentall procedure) in 2015). He reported being given a clinical diagnosis of MFS at age 3, including surgery to treat ectopia lentis. As a child he had annual echocardiograms at an outside facility which had shown progressive aortic root enlargement (details unavailable). He was managed on atenolol (50 mg daily) but reported decreased compliance. The family history was significant for a clinical diagnosis of MFS in the proband's mother, who died at age 18 of an unverified cardiac event when the proband was at the age of 10 months (presumed by the clinical team to have been an aortic dissection).

At the time of referral to cardiothoracic surgery, it was noted that an echocardiogram in 2010 had shown an aortic diameter of approximately 4 cm at the level of the aortic root. Echocardiograms within 4 months prior to his aortic surgery in 2015 showed aortic root diameter up to 4.6 cm and no significant aortic valve regurgitation. Imaging was also notable for mild prolapse of the anterior mitral valve leaflet with no significant mitral regurgitation. Representative pre-operative echocardiographic images are shown in Fig. 1. The aortic valve

was replaced during aortic root surgery because of an unexpected intra-operative finding of a fenestration in the right coronary leaflet.

The physical exam for the genetics evaluation was limited because of the patient's post-operative condition. The patient was noted to have facial features of MFS including dolichocephaly, downslanting palpebral fissures, malar hypoplasia, and retrognathia. His palate was intact and normally arched and the uvula was midline and normally formed. There was a pectus carinatum and a midline sternotomy. A wrist sign and thumb sign were both present, and the patient also had arachnodactyly. There was a report of mild scoliosis. Despite the limited nature of the initial post-operative exam, the patient fulfilled clinical criteria for MFS on the basis of his aortic dilation and a score of 7 for systemic features (facial features, wrist and thumb sign, pectus carinatum, mitral valve prolapse) based on the revised Ghent nosology (Loeys et al., 2010). Additionally, he was reported to have an affected first-degree family member in his mother, although her diagnosis could not be verified.

Genetic testing of three genes (*FBN1, TGFBR1, TGFBR2*) in 2015 was performed by the IUMGDL. Testing revealed a heterozygous variant in exon 50 of *FBN1*, denoted as: c.6141C > T (p.Ser 2047 =) classified at that time as variant of uncertain significance (VUS) and later reclassified as likely benign (ClinVar ID: 1102974). This synonymous variant did not result in an amino acid change but was a rare variant not reported in 1000 Genomes or Exome Sequencing Project databases or in the Exome Aggregation Consortium (as of 2015). The patient was subsequently lost to follow-up until his sons were referred for evaluation in 2020 (see below).

**Individual III-1:** This patient was initially referred by the primary care physician in 2020 at age 3 along with his 7-month-old brother due to a family history of MFS in their father. The patient was noted to otherwise have no major health, growth, or developmental concerns. Upon the patient's arrival in the multidisciplinary aortopathy clinic, it was determined that his father (Individual II-3) had previously been evaluated by our service in 2015 although he was not present at this outpatient clinic visit. Brief review of the family history and update of the father's (Individual II-3) health history identified that in the interim since he was evaluated in 2015, he had bilateral lensectomy surgery for ectopia lentis.

Initial examination of individual III-1 revealed a height of 102.3 cm (90th centile) and weight of 16.2 kg (77th centile). The arm span-to-height ratio was <1.05. He lacked facial features of MFS though he had deep-set eyes suggestive of enophthalmos with an elongated tubular nose. His palate was intact and normally arched with a normal appearing midline uvula. Chest was symmetric with no pectus deformity. He was noted to have 5th finger camptodactyly bilaterally and a hindfoot deformity, but no other skeletal features were present. The clinical team ordered a comprehensive aortopathy genetic testing panel consisting of 27 genes (Invitae Corp.). The results demonstrated that this patient carried the paternal silent *FBN1* VUS (c.6141C > T), which as of 2020–2021 was classified as likely benign. However, this patient also was found to have a heterozygous pathogenic variant denoted as: c.6564_6565ins*Alu*; p. Glu2189fs in exon 54. His most recent evaluation at age 5 showed a height of 118.0 cm (97th centile). His chest was asymmetric with the emergence

of a pectus carinatum. There were no additional new signs or symptoms of MFS. He had not yet been seen by ophthalmology, in part due to decreased access with the COVID-19 pandemic.

The initial cardiac examination for Individual III-1 (age 3) identified moderate aortic root dilation with a diameter Z-score of +5.5, as well as mild buckling of the mitral valve. A surface 12-lead electrocardiogram was normal. Atenolol therapy was begun at the initial visit. At the time of his last follow-up (age 5) his aortic root Z-score was decreased to +4.8, and mitral valve buckling without frank prolapse or significant regurgitation was again observed (Fig. 1).

**Individual III-2:** This patient is the younger sibling of Individual III-1 and was also referred for evaluation due to the family history of MFS at age 7 months. The patient was diagnosed with hypospadias at birth and had positional plagiocephaly, but he otherwise had no medical, growth, or developmental concerns. At the initial evaluation, the patient's length was 73 cm (89th centile), and arm span was 90 cm such that the arm span-to-height ratio was <1.05. He was noted to have mild brachycephaly and no facial features of MFS. His palate was intact and normally arched with a normal appearing uvula. Chest was symmetric with no pectus deformity. He had normal palmar and plantar creases and digits with no wrist sign, thumb sign, arachnodactyly, or dolichostenomelia. There was no limitation of elbow extension, no evidence of pes planus, no skin striae, but hindfoot deformity could not be evaluated given his young age.

His initial echocardiogram was normal (at age 7 months), including an aortic root Z-score of +1.5 and no evidence of mitral valve abnormality. No electrocardiographic abnormalities were identified. In follow-up, he was started on atenolol based on the genetic diagnosis and echocardiogram suggesting mild aortic root dilation in off-axis views. At his most recent follow-up evaluation at age 2, his aortic root Z-score was +2.3, consistent with mild aortic root dilation, and his mitral valve remained normal in appearance and function (Fig. 1).

Genetic testing first completed in his brother (Individual III-1) at the initial visit was diagnostic for MFS, identifying the *Alu* insertion in *FBN1.* Subsequently, cascade testing identified the familial pathogenic *FBN1* variant in this individual as well in addition to the paternal silent *FBN1* variant, which was determined to have no effect by general splicing predictors and thus classified as likely benign. He underwent repair of his distal hypospadias and chordee at age 12 months. He was referred to orthopedics for his intoeing and metatarsus adductus and is being followed clinically with noted improvement. The most recent follow-up was at age 28 months. The patient's height was 96.0 cm (95th centile). He was developing chest asymmetry with a slight pectus carinatum. Bilateral 5th finger clinodactyly was apparent. The metatarsus adductus and intoeing with walking continued to be noted. There was no hindfoot deformity.

## Summary of Genetic Testing Results:

As noted above, the index case proband (Individual II-3) in the family initially had inconclusive genetic testing in 2015, with a synonymous VUS in *FBN1* denoted as: c.6141C > T (p.Ser 2047 =); at present, this has been classified as a likely benign variant. Individual

II-3 was lost to follow-up for approximately five years, and we did not have the opportunity to retest him in the interim. However, when a different NGS panel was completed in 2020 in the affected son (Individual III-1) by a different laboratory (Invitae Corp.), an *Alu* insertion was identified and denoted by the commercial lab as: c.6564_6565ins*Alu* (p.Glu2189fs); the predicted impact on the FBN1 protein was based on the resolved sequence data at that time. The c.6141C > T silent variant was also identified and was confirmed to be inherited from the proband father (in cis with the *Alu* insertion). Based on this finding and the consistency with the MFS phenotype (and the family history) and absence of notable variants in other genes tested, the *FBN1*:c.6564_6565insAlu (p.Glu2189fs) variant was interpreted as pathogenic and diagnostic for MFS. This variant was submitted to the ClinVar database (ClinVar ID: 1068666). Please note that the current NCBI-approved Human Genome Variation Society (HGVS) nomenclature for his variant is listed in ClinVar as: NM_000138.5(*FBN1*):c.6564_6565insTTTTTTTTTTTTTTTTTTTTNNNNNNNNNNGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGATTGGAGGTTTT (NP_000129.3: p. Glu2189 delinsPhePhePhePhePhePheXaaXaaXaaXaaValSerProPheTer); note that the poly-“N” in the nucleotide sequence and the “Xaa” in the predicted amino acid sequence represent the commercial laboratory’s lack of coverage of this segment of the *Alu* element using the next-generation sequencing platform for the initial sequencing. Please see additional detail for the Sanger confirmation studies below, including our updates for the resolved sequence and HGVS nomenclature.

Subsequently, the younger sibling (Individual III-2) also tested positive for this variant, and his most recent evaluation showed emerging features of MFS, specifically, aortic root dilation by age 2 and emerging pectus carinatum. Following the evaluations and diagnoses in the children, the affected father (Individual II-3) re-presented to the clinic, and testing confirmed that the father did also have the pathogenic *Alu* insertion in *FBN1* (c.6564_6565ins*Alu*) that was not detected in 2015. In summary, this pathogenic variant was identified in three affected relatives with phenotypes consistent with MFS. It is noteworthy that only through subsequent familial evaluations could we successfully identify and segregate the true pathogenic variant in this case, emphasizing the utility of family-based care. Notably, this variant would impact the EGF-like 33 domain motif, and disruption of such motifs/domains in the FBN1 protein are known to increase the probability of pathogenicity (Schrijver et al., 1999).

Fig. 2 depicts a screen capture from the Integrative Genomics Viewer (IGV) browser (Thorvaldsdóttir et al., 2013) showing the mapped reads in exon 54 of *FBN1* (GRCh37/hg19: NM_000138.4). The sequence change indicates a large inserted fragment of DNA is predicting a frameshift at codon 2189 (p.Glu2189Phefs*15); however, the exact size and sequence of the insertion cannot be determined by the current assay (Invitae Corp.). This insertion is expected to result in absent/disrupted FBN1 protein. While the variant is novel, we interpret this variant as pathogenic and expected to cause loss-of-function, consistent with the disease mechanism for *FBN1* (Milewicz et al., 2021). This variant is not currently present in population databases (i.e., gnomAD). Notably, the 2×150 bidirectional reads allowed partial mapping of the correct *FBN1* sequence to the reference genome, and the

single-variant miscalls indicate the mapped sequence differing from the reference and are represented by the multicolored base calls (Fig. 2). For additional sequence information and location of 5′ and 3′ insertion sites, see Supplemental file with additional images and commentary. Notably, The IGV view clearly shows the heterozygous nature of the event revealed by the split reads 5′ insertion identified by the multicolored sequence not matching the reference sequence of the *FBN1* gene at exon 54; this impacts the second base of the codon at this site. The 3' insertion site is also detailed in the Supplemental file.

### *Alu* Element Sanger Confirmation:

DNA was extracted from subject IU-TAA-882 with the Promega Maxwell® RSC Buffy Coat DNA kit (Promega Corporation, Madison, WI, USA) and amplified using primers spanning *FBN1* exon 54 and intron 54 including the *Alu* insertion originally identified by a NGS gene panel performed by Invitae (Invitae Corp.): FW (5′-GCAAGAATGTGATTGGAGGTTTTTTTTTTTT −3′) and RV (5′-CCTACAGAGAAGAGGCCACAA −3′). PCR reactions were carried out with FastStart PCR Master mix (Roche Diagnostics GmbH, Mannheim, Germany) in an Applied Biosystems Veriti thermal cycler under the following conditions: initial denaturation at 95 °C for 3 min, 35 cycles of amplification (denaturation at 95 °C for 30 s, annealing at 56 °C for 30 s, and extension at 72 °C for 1 min 30 s), and a final extension at 72 °C for 10 min. PCR amplicon was confirmed by gel electrophoresis and purified with the Nucleospin® Gel and PCR Clean-up kit (Machery-Nagel Inc., Allentown, PA, USA) following vendor's instructions. Sanger sequencing was performed with purified DNA (ACGT, Inc., Wheeling, IL, USA) via their Single Pass DNA Sequencing (SPPE) method using the same primers as above: FW (5′- GCAAGAATGTGATTGGAGGTTTTTTTTTTTT −3′) and RV (5′- CCTACAGAGAAGAGGCCACAA −3′). ACGT's system utilizes BigDye terminator version 3.1, clean-up with magnetic beads (CleanSEQ dye terminator removal kit), and analysis of extension products using ABI 3730 XL or 3730 Genetic Analyzer. ABI format chromatograms were generated and analyzed with FinchTV version 1.4.0 software (Geospiza). We also used the resolved sequence data here to then define the HGVS nomenclature and predicted impacts on the *FBN1* gene using SnapGene (SnapGene software [www.snapgene.com]) and Mutalyzer (version 3) (Lefter et al., 2021; Vis et al., 2015).

Sanger sequencing results verified the *Alu* insertion in the *FBN1* exon 54 of subject IU-TAA-882. The reverse sequence is shown in the electropherogram of Supplemental Fig. 5, in which a string of thymidines (dT) are observed at the 5′ end of the sequence indicating the poly(A) region at the 3' end of the *Alu* sequence. Interestingly, some *Alu* elements have tails of 40 or more adenosine residues (dA) in length, and the length of the poly(A) provides *Alu* elements retropositional competence (Roy-Engel et al., 2002). However, we were not able to determine the precise number of dTs (conversely dAs) present in the sequence. Sanger sequencing also uncovered a 175-nucleotide region previously unidentified in the *Alu* sequence (labeled "NNNNNNNNN" by Invitae). In addition, we were able to identify only one CCACCGCGC sequence instead of the "CCACCGCGCCCACCGCGC" repeat previously reported in the *Alu* sequence by Invitae. Finally, we estimate the entire *Alu* sequence is approximately 337 base pairs. The HGVS nomenclature for the *Alu* insertion confirmed here is denoted as: NM_000138.5(FBN1):

c.6564_6565insTTTNTTTTTTTTTTTTTTTTTTTTTTTTTNTTNTTTTTTT-TTNTNGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGGGATCTCGGCTCACTGCAAGCTCCGCCTCCTGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCGCTACGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGATTGGAGGTTTT. The “N” bases noted in the poly-T region at the start indicate some ambiguity in the Sanger chromatogram traces at those sites (see Supplementary Material). Using Mutalyzer, the predicted impact on the FBN1 protein was denoted as: NM_000138.5 (NP_000129.3):p.Glu2189Phefs*78. Notably, this is the predicted impact using the resolved sequence of the *FBN1* gene with the *Alu* insertion, though additional protein/RNA sequencing studies beyond the scope of this report would be needed to verify this. Specific details allowing estimation of these impacts are provided in the Supplementary Material.

### *Alu* Subfamily Classification:

The variant nomenclature used here is based on available/ascertained sequence information for the *Alu* insertion. Use of the *BLAST-Like Alignment Tool* (BLAT) and the *RepeatMasker* track in the UCSC Browser allowed classification of this *Alu* element. The *Alu* element was specified as belonging to the *AluYa5* subfamily (Bennett et al., 2008).

## 3. Discussion

With improved genomic testing technologies and bioinformatics analyses over the last decade, there has been increasing recognition of the role of transposable elements causing rare Mendelian disorders (Deininger, 2011; Hancks and Kazazian, 2012). However, there have been no previously published reports of *Alu*-specific disruption of the *FBN1* gene causing MFS. We address this and report a family of three affected first-degree relatives with clinical diagnoses of MFS found to have a novel *Alu* insertion which disrupts the coding sequence of the *FBN1* gene. Importantly, the proband in the family had inconclusive results from NGS panel testing in 2015. The synonymous variant initially led the clinical team in 2015 to hypothesize that it may have been pathogenic by creating a cryptic splice site and altering splicing; however, RNA sequencing was not available and the proband was lost to follow-up preventing further investigation at the time. There were also no other living relatives to allow family segregation studies to aid interpretation. Classifying this VUS pre-surgically likely would have not impacted any surgical decision-making given the proband meeting revised Ghent criteria for MFS. Instead, more contemporary NGS analysis in 2020 determined that the proband’s initial VUS was likely not the cause of the phenotype. Notably, the variant-calling pipeline used by IUMGDL in 2015 could not account for variants like *Alu* insertions. If Sanger sequencing had been used initially, employing the primers and amplification protocol designed specifically to detect this genomic event, it is likely that the *Alu* insertion would have been identified in 2015; however, it is uncommon for commercial clinical genetic testing laboratories in the United States to prioritize Sanger methods in the era of next-generation sequencing (or without *a priori* knowledge of an atypical variant like this *Alu* insertion). Therefore, we completed additional studies and

confirmed the *Alu* insertion via Sanger sequencing, and this also allowed us to resolve nearly the entire insertion sequence (Supplemental File, Fig. 5). While the poly-T sequence could not be fully resolved allowing us to specify the predicted frameshift mechanism, the *Alu* insertion would be expected to significantly disrupt the *FBN1* coding sequence consistent with pathogenicity.

This case illustrates the importance of considering retesting and/or NGS data reanalysis for patients meeting clinical diagnostic criteria for a well-defined genetic disorder like the proband here. However, it should be noted that a simple data reanalysis of the original 2015 NGS panel would not have identified the *Alu* insertion, simply because of the limitations in that test's sequencing depth and variant-calling pipelines not designed to capture or confirm larger insertions. It is unlikely that interrogation of the original FASTQ or alignment files would have identified the *Alu* insertion due to limited sequencing depth and mapping quality in 2015. This may be similar to a 2018 case report of Bardet-Biedl syndrome where a pathogenic variant arising from a transposon could only be detected by manual visual inspection of NGS alignment data with a pipeline specific for detecting transposable elements (Tavares et al., 2019). Therefore, newer retesting (i.e., resequencing) methods with improved bioinformatics methods would be necessary to detect variants like this *Alu* insertion. The difference in diagnostic capability from 2015 to 2020 in this case clearly shows how rapidly the sequencing capabilities change and the need for clinicians to stay updated on diagnostic testing limitations. It also highlights the need to consider atypical mechanisms of rare diseases, like retrotransposon-mediated events or large CNVs. This case also highlights the importance of the clinical phenotype in addition to the utility of contemporary clinical genetic testing, as opposed to reflexing to research (which may not be accessible to clinicians). Dordoni et al. (2017) note that chromosomal alterations of *FBN1* are rare but had been increasingly identified with testing technologies allowing identification of them, e.g., multiplex ligation-dependent probe amplification and microarray-based diagnostic testing.

Interestingly, there is one previous report of MFS caused by a multi-exonic deletion arising from a concomitant retrotransposon insertion (Brett et al., 2017). The authors reported a four-year-old patient with aortic dilation but who otherwise did not meet clinical diagnostic criteria for MFS. Using chromosome microarray analysis, the patient had a 36.8 kb deletion at 15q21.1 affecting exons 7–9. They performed sequencing analysis of the breakpoint regions and determined that there was an associated insertion of a retrotransposon within the intron 6/intron 9 region. However, our current case differs in that a predicted sequence frameshift/alteration and predicted disruption of *FBN1* was mediated by an *Alu* element retrotransposon without causing copy-number abnormalities, whereas the case from Brett et al. (2017) was thought to be caused by a *MAST2* SINE-VNTR-*Alu* (*MAST2*-SVA) element. Hancks and Kazazian (2016) provide an in-depth review of this and other retrotransposon elements and their trans-mobilization mechanisms, including their potential to generate deletions via non-allelic homologous recombination. We direct interested readers to this work for more details beyond the scope of this case report.

Given the ubiquity of these elements throughout the human genome and their ability to trans-mobilize and potentially disrupt gene coding regions, promote CNV formation, or alter

epigenetic activity (Hancks and Kazazian, 2016), it is possible that previously "diagnosis-elusive" clinical cases may have had mutational events arising from transposable elements. Until genomic technologies improved, it was simply not possible or too challenging to identify such events (e.g., Tavares et al., 2019). Therefore, we wish to raise awareness among clinicians to consider such rare mutational events, especially for cases where there is strong clinical suspicion for a disorder, but diagnostic genetic testing was inconclusive. It may also be reasonable to consider retrotransposon events for diseases associated with relatively large genes with ≥20–30 exons where these events have been previously reported (e.g., *NF1*, *CHD7*, *APC*, and *BRCA1/2*) (Deininger, 2011; Hancks and Kazazian, 2012). It is possible that the probability of *Alu* insertional disruption of a coding region would be a function of how large a gene is, providing a larger target area for a retrotransposon to insert into. However, this is conjectural and requires further investigation. It is unclear if there are genomic contexts that might increase the probability of *Alu* insertions into genic regions, and such events may be stochastic. Others have explored gene-specific susceptibility to *Alu*-mediated CNVs, and we direct readers to that work (Song et al., 2018); notably, it is unknown how frequently *Alu*-mediated CNVs occur in *FBN1*.

Using the ascertainable sequence information, BLAT, and the *RepeatMasker* track in the UCSC Genome Browser, we were able to classify this *Alu* element as belonging to the *AluYa5* subfamily (Bennett et al., 2008). Subfamily classification does not significantly inform clinical evaluation, management, or genetic counseling for this family. However, it is interesting to note that the *AluY* family and its subfamilies all have intact *Alu* core elements and have relatively high levels of genomic mobilization (Bennett et al., 2008). These higher mobilization efficiencies have led to the *AluYa5* and *AluYb8* subfamilies comprising over half of all polymorphic *Alu* elements in humans (Bennett et al., 2008).

Last, there are a few notable limitations of this report. Due to limitations of short-read sequencing, we could not ascertain the approximately 300-base pair size of the *Alu* insertion here. However, we confirmed the *Alu* insertion via orthogonal Sanger sequencing and resolved the sequence that the commercial laboratory was unable to (except for slightly decreased resolution of the poly-T tail composition, which can be variable in *Alu* elements) (Roy-Engel et al., 2002). Otherwise, there was enough sequence information to estimate a predicted protein impact and use BLAT for *Alu* classification. However, the subfamily classification does not improve current understanding of functional effect or alter diagnosis and clinical management. Also, it was unclear whether chromosome microarray could have a role in determining whether the *Alu* insertion in this family mediates a multi-exon deletion, as in Brett et al. (2017); however, orthogonal gene-specific CNV analysis was completed and normal. In conclusion, the Mendelian cause of MFS in this family was a result of the retrotransposon-mediated mutational event that was ascertained because of rapid improvements in NGS technology and analysis. This highlights the importance of *Alu* retrotransposition as a rare disease mechanism in MFS in addition to other Mendelian genetic diseases. It also highlights the utility of clinical genetic testing and multidisciplinary collaboration between clinicians and clinical genetic testing laboratories in the context of a well-defined (but previously molecular diagnosis-elusive) phenotype.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors wish to thank the patient and family for their willingness to share deidentified information. This project was made possible in part by the Indiana University Strategic Research Initiative (BMH, BJL, and SMW). We also thank members of our team for helping facilitate and troubleshoot the *post hoc* Sanger sequencing studies confirming the *Alu* insertion: Helen Bellchambers, John Wells, Maria Padua, and Courtney Vujakovich.

## Data availability

No data was used for the research described in the article.

## References

Ade C, Roy-Engel AM, Deininger PL, 2013. *Alu* elements: an intrinsic source of human genome instability. Current opinion in virology 3 (6), 639–645. 10.1016/j.coviro.2013.09.002. [PubMed: 24080407]

Bennett EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, Devine SE, 2008. Active Alu retrotransposons in the human genome. Genome Res. 18 (12), 1875–1883. 10.1101/gr.081737.108. [PubMed: 18836035]

Brett M, Korovesis G, Lai A, Lim E, Tan EC, 2017. Intragenic multi-exon deletion in the FBN1 gene in a child with mildly dilated aortic sinus: a retrotransposal event. J. Hum. Genet 62 (7), 711–715. 10.1038/jhg.2017.32. [PubMed: 28331219]

Deininger P, 2011. *Alu* elements: know the SINEs. Genome Biol. 12 (12), 236. 10.1186/gb-2011-12-12-236. [PubMed: 22204421]

Deininger PL, Batzer MA, 1999a. *Alu* repeats and human disease. Mol. Genet. Metabol 67 (3), 183–193. 10.1006/mgme.1999.2864.

Deininger PL, Batzer MA, 1999b. Alu repeats and human disease. Mol. Genet. Metabol 67 (3), 183–193. 10.1006/mgme.1999.2864.

Dietz H, 2001. Marfan syndrome. In: Adam MP, Ardinger HH, Pagon RA, et al. (Eds.), GeneReviews® [Internet]. University of Washington, Seattle, Seattle (WA), pp. 1993–2020. Updated 2017 Oct 12]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1335/.

Dordoni C, Ciaccio C, Santoro G, Venturini M, Cavallari U, Ritelli M, Colombi M, 2017. Marfan syndrome: report of a complex phenotype due to a 15q21.1 contiguos gene deletion encompassing FBN1, and literature review. Am. J. Med. Genet 173 (1), 200–206. 10.1002/ajmg.a.37975. [PubMed: 27615407]

Hancks DC, Kazazian HH Jr., 2012. Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev 22 (3), 191–203. 10.1016/j.gde.2012.02.006. [PubMed: 22406018]

Hancks DC, Kazazian HH Jr., 2016. Roles for retrotransposon insertions in human disease. Mobile DNA 7, 9. 10.1186/s13100-016-0065-9. [PubMed: 27158268]

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, 2001. International human genome sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409 (6822), 860–921. 10.1038/35057062. [PubMed: 11237011]

Lefter M, Vis JK, Vermaat M, den Dunnen JT, Taschner PEM, Laros JFJ, 2021. Mutalyzer 2: next generation HGVS nomenclature checker. Bioinformatics 37 (18), 2811–2817. 10.1093/bioinformatics/btab051. [PubMed: 33538839]

Li H, Durbin R, 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14), 1754–1760. 10.1093/bioinformatics/btp324. [PubMed: 19451168]

Loeys BL, Dietz HC, Braverman AC, Callewaert BL, De Backer J, Devereux RB, Hilhorst-Hofstee Y, Jondeau G, Faivre L, Milewicz DM, Pyeritz RE, Sponseller PD, Wordsworth P, De Paepe AM, 2010. The revised Ghent nosology for the Marfan syndrome. J. Med. Genet 47 (7), 476–485. 10.1136/jmg.2009.072785. [PubMed: 20591885]

Lopez L, Colan SD, Frommelt PC, Ensing GJ, Kendall K, Younoszai AK, Lai WW, Geva T, 2010. Recommendations for quantification methods during the performance of a pediatric echocardiogram: a report from the pediatric measurements writing group of the American society of echocardiography pediatric and congenital heart disease Council.Journal of the American society of echocardiography. Off. Publ.Am. Soc. Echocardiogr 23 (5), 465–577. 10.1016/j.echo.2010.03.019.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (9), 1297–1303. 10.1101/gr.107524.110. [PubMed: 20644199]

Milewicz DM, Braverman AC, De Backer J, Morris SA, Boileau C, Maumenee IH, Jondeau G, Evangelista A, Pyeritz RE, 2021. Marfan syndrome. Nat. Rev. Dis. Prim 7 (1), 64. 10.1038/s41572-021-00298-7. [PubMed: 34475413]

Roy-Engel AM, Salem AH, Oyeniran OO, Deininger L, Hedges DJ, Kilroy GE, Batzer MA, Deininger PL, 2002. Active Alu element “A-tails”: size does matter. Genome Res. 12 (9), 1333–1344. 10.1101/gr.384802. [PubMed: 12213770]

Schrijver I, Liu W, Brenn T, Furthmayr H, Francke U, 1999. Cysteine substitutions in epidermal growth factor-like domains of fibrillin-1: distinct effects on biochemical and clinical phenotypes. Am. J. Hum. Genet 65 (4), 1007–1020. 10.1086/302582. [PubMed: 10486319]

Sluysmans T, Colan SD, 2005. Theoretical and empirical derivation of cardiovascular allometric relationships in children. J. Appl. Physiol 99 (2), 445–457. 10.1152/japplphysiol.01144.2004. [PubMed: 15557009]

Song X, Beck CR, Du R, Campbell IM, Coban-Akdemir Z, Gu S, Breman AM, Stankiewicz P, Ira G, Shaw CA, Lupski JR, 2018. Predicting human genes susceptible to genomic instability associated with*Alu/Alu*-mediated rearrangements. Genome Res. 28 (8), 1228–1242. 10.1101/gr.229401.117. [PubMed: 29907612]

Tavares E, Tang CY, Vig A, Li S, Billingsley G, Sung W, Vincent A, Thiruvahindrapuram B, Héon E, 2019. Retrotransposon insertion as a novel mutational event in Bardet-Biedl syndrome. Mol. Gen.Genom.Med 7 (2), e00521. 10.1002/mgg3.521.

Thorvaldsdóttir H, Robinson JT, Mesirov JP, 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings Bioinf. 14 (2), 178–192. 10.1093/bib/bbs017.

van den Hurk JA, Meij IC, Seleme MC, Kano H, Nikopoulos K, Hoefsloot LH, Sistermans EA, de Wijs IJ, Mukhopadhyay A, Plomp AS, de Jong PT, Kazazian HH, Cremers FP, 2007. L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet 16 (13), 1587–1592. 10.1093/hmg/ddm108. [PubMed: 17483097]

Vis JK, Vermaat M, Taschner PE, Kok JN, Laros JF, 2015. An efficient algorithm for the extraction of HGVS variant descriptions from sequences. Bioinformatics 31 (23), 3751–3757. 10.1093/bioinformatics/btv443. [PubMed: 26231427]

Xing J, Zhang Y, Han K, Salem AH, Sen SK, Huff CD, Zhou Q, Kirkness EF, Levy S, Batzer MA, Jorde LB, 2009. Mobile elements create structural variation: analysis of a complete human genome. Genome Res. 19 (9), 1516–1526. 10.1101/gr.091827.109. [PubMed: 19439515]

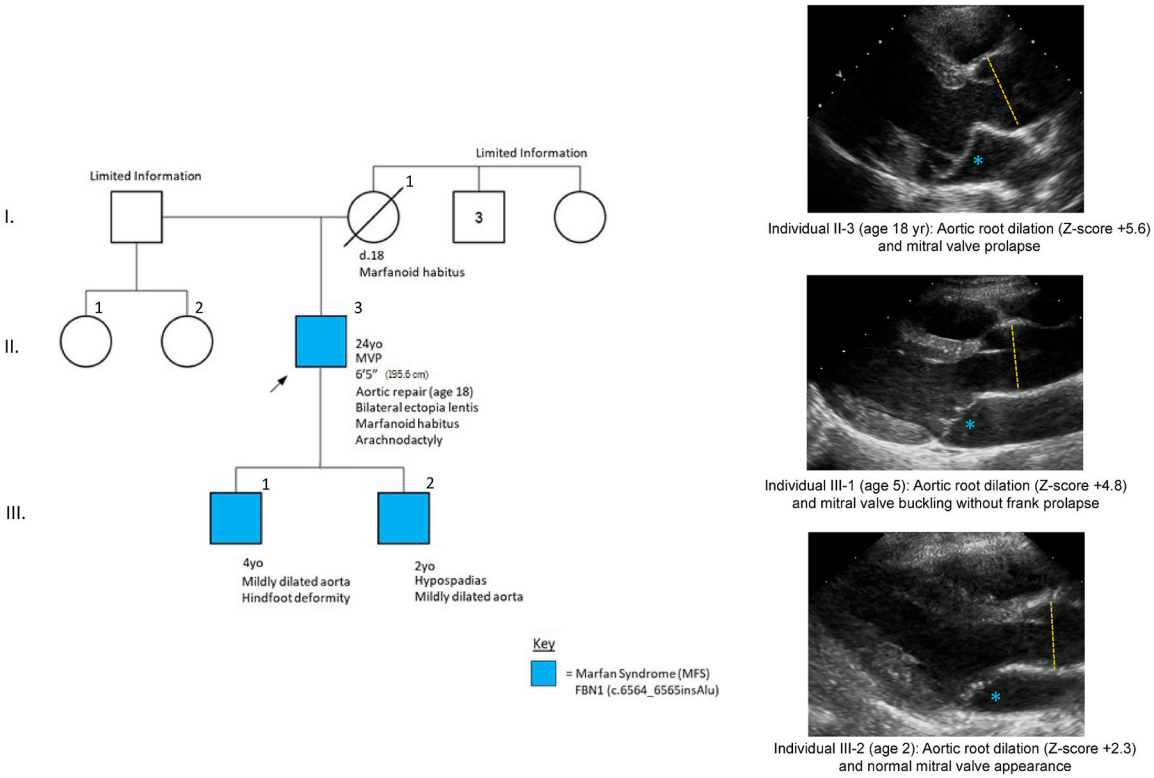


**Fig. 1.**

Selected pertinent family history information summarized in the pedigree (left). The right panels are echocardiographic images depicting affected individuals' involvement of the aortic root and mitral valves. Orange dashed lines indicate aortic root segments. Blue asterisks are located near the mitral valve leaflets during systole. Abbreviations: MVP = mitral valve prolapse. Note: height for Individual II-3 was listed in the pedigree as 6′−5′′ which is equivalent to 77 inches which is equivalent to 195.6 cm. Note: the ages in the pedigree above may vary with the clinical summaries in the text, since the latter included >1 evaluation in clinic at different times compared to when the pedigree was recorded. Please note that the ages listed in this pedigree reflect the family history collected at individuals' III-1 and III-2 clinical ascertainment (and may differ from the age at follow-up evaluations in clinic (e.g., the echocardiogram for individual III-1 (age 5) shown at right, approximately one year after presenting to clinic).

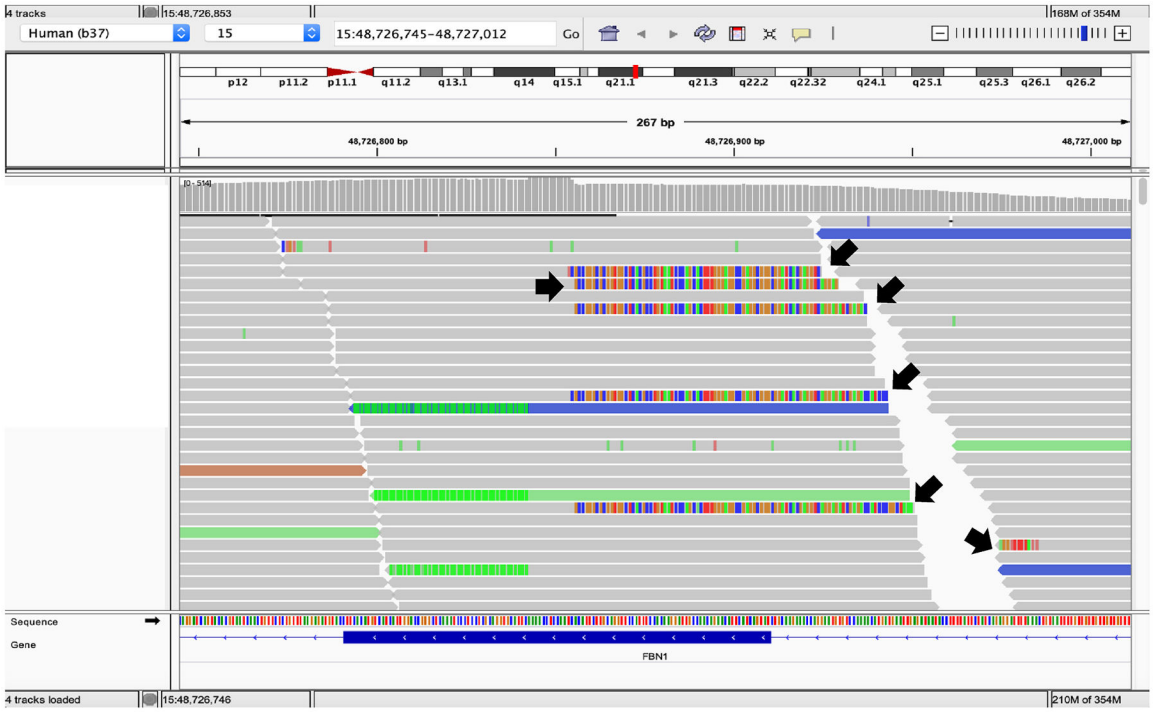


**Fig. 2.**

Image from the Integrated Genomics Viewer depicting the *Alu* insertion in *FBN1* (c.6564_6565insAlu). The image is a limited capture of the reads mapped to this region of *FBN1*, and the multicolored bases indicate several reference sequence mismatches. While the insertion's impact on the nucleotide sequence of the *FBN1* gene is known (the nucleotide start and end, as noted in the variant nomenclature), the specific sequence of the *Alu* insertion is unknown due to limitations of short-read sequencing (or without use of single-molecule sequencing). Focus is directed to the multicolored base calls in the image above (arrows). NOTE: this region was covered at 514x depth, so there are numerous reads that cannot be viewed in a single image here. See Supplemental file for additional images showing the reads and the evidence supporting the heterozygous call.