



OPEN ACCESS

EDITED BY

Daniel Christian Hoessli,
University of Karachi, Pakistan

REVIEWED BY

Dora Janeth Fonseca,
Rosario University, Colombia
Anna Sowińska-Seidler,
Poznan University of Medical Sciences, Poland
Hassan Dastsooz,
Italian Institute for Genomic Medicine (IIGM),
Italy
Lauren Wasson,
Harvard Medical School, United States

*CORRESPONDENCE

Sara J. Cooper,
✉ sjcooper@hudsonalpha.org

RECEIVED 20 May 2024

ACCEPTED 29 November 2024

PUBLISHED 16 December 2024

CITATION

Gordon ER, Felker SA, Coleman TF,
Sosonkina N, Pugh J, Cochran ME, Hurst ACE
and Cooper SJ (2024) Novel variant alters
splicing of *TGFB2* in family with features of
Loeys-Dietz syndrome.
Front. Genet. 15:1435734.
doi: 10.3389/fgene.2024.1435734

COPYRIGHT

© 2024 Gordon, Felker, Coleman, Sosonkina,
Pugh, Cochran, Hurst and Cooper. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Novel variant alters splicing of *TGFB2* in family with features of Loeys-Dietz syndrome

Emily R. Gordon¹, Stephanie A. Felker^{1,2}, Tanner F. Coleman^{1,3},
Nadiya Sosonkina⁴, Jada Pugh^{1,3}, Meagan E. Cochran^{1,3},
Anna C. E. Hurst² and Sara J. Cooper^{1*}

¹HudsonAlpha Institute for Biotechnology, Huntsville, AL, United States, ²Department of Genetics, University of Alabama at Birmingham, Birmingham, AL, United States, ³Smith Family Clinic for Genomic Medicine, Huntsville, AL, United States, ⁴Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, United States

Loeys-Dietz syndrome (LDS) is a connective tissue disorder representing a wide spectrum of phenotypes, ranging from isolated thoracic aortic aneurysm or dissection to a more severe syndromic presentation with multisystemic involvement. Significant clinical variability has been noted for both related and unrelated individuals with the same pathogenic variant. We report a family of five affected individuals with notable phenotypic variability who appear to have two distinct molecular causes of LDS, one attributable to a missense variant in *TGFB2* and the other an intronic variant 6 bp upstream from a splice junction in *TGFB2*. We tested the functional impacts of the variant identified in the proband alongside other variants in the region reported in ClinVar using a splice reporter system, which resulted in non-canonical splicing products for several variants including the proband. Molecular validation of the splicing products suggests that the *TGFB2* variants tested impact splicing by reducing efficiency of the canonical acceptor in favor of an alternate acceptor within the exon. These data combined with clinical phenotypes and segregation of the variant with disease support the conclusion that this intronic *TGFB2* variant may cause LDS in this patient and her mother. These analyses demonstrate that underappreciated intronic variants that alter splicing can be relevant for clinical phenotypes of connective tissue disease. This case highlights the importance of prompt familial cascade testing, clinical evaluation with detailed dysmorphology exam, comprehensive genetic testing, and collaboration between clinicians and scientists to characterize variants of uncertain significance to properly assess risk in LDS patients.

KEYWORDS

splicing, Loeys-Dietz, *TGFB2*, noncoding variation, WGS, connective tissue disorder, SpliceAI, clinical genomics

1 Introduction

Loeys-Dietz syndrome (LDS) is a rare connective tissue disorder with a prevalence of 1 in 100,000 individuals (Loughborough et al., 2018). Affected individuals present with a wide spectrum of phenotypes, ranging from isolated thoracic aortic aneurysm or dissection to a severe syndromic presentation with multisystemic involvement. Due to the incidence of aortic aneurysm and dissection, the life expectancy of an individual with LDS averages 37 years old, although the range is highly variable depending on severity (Loeys et al., 2006;

Mazzella et al., 2017). LDS diagnosis typically requires angiograms, cardiac imaging, basic family history, medical exams, and genetic testing for molecular diagnosis.

LDS results from pathogenic heterozygous loss-of-function variation in one of several TGF-beta pathway members: *SMAD2*, *SMAD3*, *TGFB2*, *TGFB3*, *TGFB1*, or *TGFB2*. The majority of cases (75%) are the result of *de novo* mutations in one of these six genes and the remainder are inherited in an autosomal dominant manner. Significant clinical variability has been noted for both related and unrelated individuals with the same pathogenic variant, and there is documented reduced penetrance, so even when a LDS-causing variant is inherited, parents may not show clinical features of the condition (Loeys and Dietz, 2008).

While the pathway affected in LDS patients is well-described, only about half of LDS patients have a molecular diagnosis (Renner et al., 2019). Inability to provide a genetic diagnosis results from a variety of challenges including a limited ability to predict the impact of non-coding, intronic, and synonymous variants. Advancements in whole genome sequencing (WGS) have significantly shortened the diagnostic odyssey and increased the frequency of molecular diagnosis in rare diseases cases, including LDS. In LDS, genetic diagnosis improves management of patients because treatment and imaging recommendations vary depending on the specific cause (Ikonomidis, 2017; MacCarrick et al., 2014; Krohg-Sørensen et al., 2017; Isselbacher et al., 2022). In our study, we explore the function of variants that could alter splicing in LDS-associated genes. We use SpliceAI predictions (Jaganathan et al., 2019) to prioritize variants that may alter splicing (Hirschi et al., 2024). Our study characterizes the functional impact of intronic variants of the LDS-associated gene *TGFB2* including one observed in a unique family with an LDS diagnosis and other individuals with features of LDS.

We describe a family with three confirmed LDS diagnoses explained by a missense variant in *TGFB2*. Additional family members were identified with features of LDS, but genetic testing ruled out the familial *TGFB2* variant. Instead, genome sequencing revealed an intronic variant in *TGFB2* [NM_003238.6 (*TGFB2*):c.755-6T>C], another LDS-associated gene. We used a splice reporter assay to characterize this variant which occurs 6bp upstream of the splice acceptor in exon 5 of *TGFB2*. Our data provides new evidence that this splice-altering allele may lead to LDS-like features in this unique family that appears to have two distinct variants in LDS-associated genes resulting in similar LDS-associated phenotypes. This case demonstrates the value of thorough genetic evaluation and emphasizes the potential impact of non-coding sequences to further our understanding of LDS.

2 Methods

2.1 Clinical evaluation of connective tissue disorder

Patients were evaluated for features of connective tissue disorders based on the standard of care process. Evaluation by a medical geneticist included a physical exam, a 92-gene connective tissue disorder panel, and targeted sequencing as described in detail in the Results section. Full details of the clinical evaluation are available in [Supplementary Material S3](#).

2.2 Genome sequencing

Whole Genome Sequencing was performed by HudsonAlpha Clinical Services Lab, LLC (“CSL”) (CLIA: #01D2086581; CAP: #8051488; Lab Director: Ghunwa Nakouzi, PhD), using the Illumina NovaSeq 6000 sequencing platform. DNA was checked for quality and concentration and sonicated to a specific fragment size. After size selection, DNA fragments were ligated to a unique pair of Illumina flow cell-specific index adapters. No amplification of the library was performed. The prepared library was quality-checked for adequate yield with fluorescence methods and quantitative polymerase chain reaction (qPCR) and for appropriate fragment size distribution using microfluidics-based platforms. The library was pooled with similar type libraries, denatured, and loaded onto Illumina NovaSeq 6000 flow cells and sequenced using standard Illumina reagents and protocols.

2.3 Data processing and quality control

Genome sequencing data were collected using Illumina sequencing and reads were aligned to the human reference GRCh38. Quality control metrics were calculated, with low quality reads being removed. Variant calling proceeded using unique, high quality reads. See [Supplementary Material S3](#) for additional details.

2.4 Genomic analysis

Sequence variants were analyzed using a custom software analysis application called Codicem (v. 5.7.4) for interpretation. See [Supplementary Material S3](#) for filtering restrictions and details. Variants were classified based on the ACMG/AMP guidelines (Richards et al., 2015). All primary, actionable secondary, and complex indel variants were confirmed by an orthogonal technology (Sanger (dideoxy) sequencing). Other reported variants were assessed by a validated algorithm and orthogonally confirmed when metrics indicated a possible false positive (Holt et al., 2021) Patient-specific orthogonal confirmation details are available upon request.

2.5 Analysis of variants submitted to ClinVar database

Variants near the proband variant were acquired from ClinVar (July 2023), and intronic and synonymous variants with no predicted amino acid or termination change were annotated using Ensemble’s Variant Effect Predictor (VEP, McLaren et al., 2016) and the SpliceAI (Jaganathan et al., 2019) web server (<https://spliceailookup.broadinstitute.org/>) to predict the effect of each variant on splicing within 500 bp of the variant. We used an additional computational predictor, Combined Annotation Dependent Depletion (CADD) score, to further filter variants for consideration (Rentzsch et al., 2021). One likely pathogenic positive control variant [NM_003238.6 (*TGFB2*):c.755-5_755-2delinsG] and two common negative control variants (NM_003238.6 (*TGFB2*): c.

TABLE 1 Variant table including SpliceAI scores.

HGVS DNA reference	Variant type	ClinVar ID	ACMG classification	gnomAD v3.1.2 allele count	CADD score	SpliceAI Δ Score acceptor gain	SpliceAI Δ Score acceptor loss	SpliceAI Δ Score donor gain	SpliceAI Δ Score donor loss
NM_003238.6 (TGFB2): c.755-6T>C	SNV	2,017,460	likely benign	0	19.78	0.03	0.1	0	0
NM_003238.6 (TGFB2): c.755-5_755-2delinsG	Deletion	528,887	likely pathogenic	0	N/A	0.14	0.91	0	0
NM_003238.6 (TGFB2): c.755-5dup	Duplication	2,165,076	benign	7	14.51	0	0	0.03	0
NM_003238.6 (TGFB2): c.747A>G (p.Arg249 =)	SNV	387,890	likely benign	45	14.98	0	0	0.02	0.18
NM_003238.6 (TGFB2): c.644-18T>A	SNV	511,039	likely benign	0	22.3	0.02	0.57	0	0

755-5dup and NM_003238.6 (TGFB2):c.747A>G (p.Arg249 =) were selected. An additional test variant classified as Likely Benign [NM_003238.6 (TGFB2):c.644-18T>A] was also selected for analysis, due to positive computational predictors. The ClinVar submitters of these variants were contacted for phenotype information, when available.

2.6 Molecular validation

We used the pSpliceExpress vector system (Addgene #32485) and methods previously established by (Kishore et al., 2008) to conduct a splice reporter assay (Supplementary Material S3). Sequences were synthesized as gBlocks HiFi gene fragments (IDT). The gBlock harbored the three exon-intron-exon-intron-exon junctions of interest along with flanking 200 base pairs of introns and attB1 sites. Six sequences were tested: reference (GRCh38), NM_003238.6:c.755-6T>C (proband), and four sequences reported in ClinVar not previously characterized by splice reporter (Table 1). Each of these variants falls in the intronic sequence near the variant NM_003238.6:c.755-6T>C (Supplementary Table S1). Cloned and purified plasmids were sequence verified with whole plasmid sequencing using Oxford Nanopore Technology (Plasmidsaurus, Eugene, OR).

HEK293T cells were seeded at a density of 5×10^5 cells per well of a 6 well plate 24 h prior to transfection in 2 mL of media. Transfection into HEK293T cells took place per well using 1 μ g of reporter plasmid, 12.5 μ L Lipofectamine 2000 (ThermoFisher #11668019), 10 μ L Plus reagent (ThermoFisher #11514015), and 400 μ L OptiMem (ThermoFisher #11058021). Twenty-four hours following transfection, the cells were harvested, and RNA was extracted with the Norgen Total RNA Extraction Kit following the manufacturer's protocol including DNase treatment followed by freezing at -80°C (Norgen #37500 and #25710).

Two step RT-PCR using 300 ng of RNA was performed. cDNA was made using random hexamer (50 ng) primed Superscript first-

strand synthesis system (Invitrogen #11904-018) and frozen at -20°C . Thirty nanograms of template cDNA was amplified by PCR using primer sequences in the rat insulin exons (RIE2 or RIE3) (5'-CCTGCTCATCCTCTGGGAGC-3' and 5'-AGGTCTGAAGGTCACGGGCC-3') using Platinum SuperFI PCR MasterMix following manufacturer's protocol for a 20 μ L reaction (Invitrogen #12368010). PCR conditions were as follows: initial denaturation 98°C for 30 s; 30 cycles of 98°C for 10 s, 66°C for 10 s, and 72°C for 1 min; and a final extension at 72°C for 5 min. PCR products were frozen at -20°C . Products were gel purified using 2% agarose gel electrophoresis with SYBR Green I gel electrophoresis dye per manufacturer's recommendations (Sigma #S9430). Gels were visualized (Bio-Rad ChemiDoc Imaging System) and bands were manually cut out. DNA was extracted from individual bands (Zymo #D4007) and sent for Sanger sequencing (Molecular Cloning Laboratory) (Supplementary Table S2). Intensity of gel bands was quantified using ImageJ 1.53 K (<http://imagej.nih.gov/ij>), background signal was normalized, and fraction of total intensity for each band was calculated. Graphs and statistics were generated using GraphPad Prism (v10). Significance was determined with an unpaired parametric *t*-test. All steps were performed sequentially in quadruplicate unless otherwise noted.

3 Results

3.1 Clinical presentation and family history: NM_003238.6 (TGFB2):c.412T>G p.Cys138Gly

The first individual (II:3, Figure 1A) evaluated at the Smith Family Clinic for Genomic Medicine in Huntsville, AL was a 53-year-old male evaluated for possible connective tissue disorders due to bicuspid aortic valve and thoracic aorta aneurysm of 4.6 cm. His medical history was significant for microretrognathia, surgically repaired as a teenager, and kyphoscoliosis. On physical

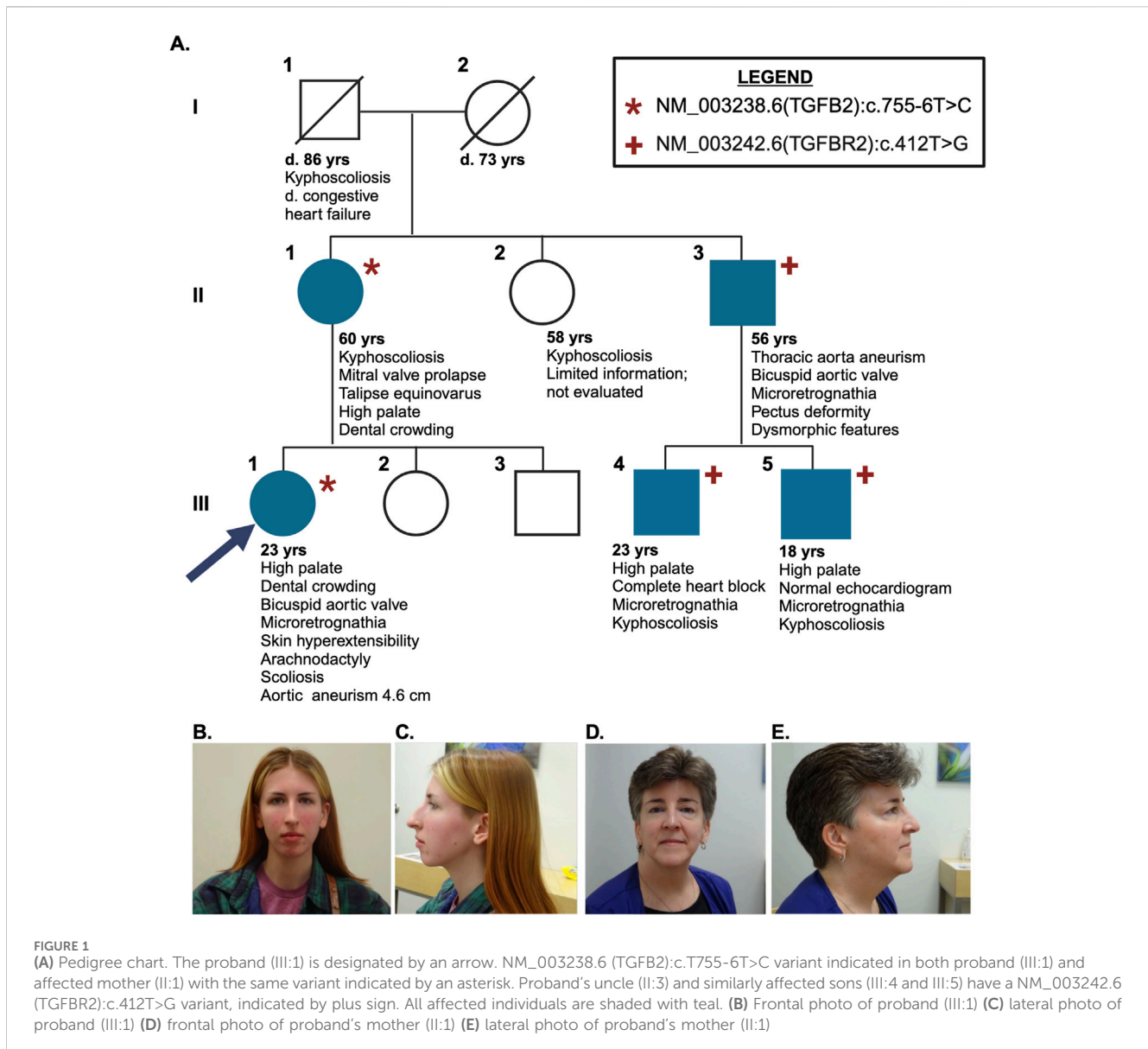


FIGURE 1

(A) Pedigree chart. The proband (III:1) is designated by an arrow. NM_003238.6 (TGFB2):c.755-6T>C variant indicated in both proband (III:1) and affected mother (II:1) with the same variant indicated by an asterisk. Proband's uncle (II:3) and similarly affected sons (III:4 and III:5) have a NM_003242.6 (TGFBR2):c.412T>G variant, indicated by plus sign. All affected individuals are shaded with teal. (B) Frontal photo of proband (III:1) (C) lateral photo of proband (III:1) (D) frontal photo of proband's mother (II:1) (E) lateral photo of proband's mother (II:1)

examination, he was noted to have malar flattening, atypical appearance of the chest (narrow upper thorax, sloping shoulders, S-shaped sternum), wide neck with webbing, slightly low-set ears with thick helix, and down slanting palpebral fissures. He did not meet the established clinical diagnostic criteria for Marfan syndrome (Loeys et al., 2010) (Supplementary Material S3).

Individual II:3 underwent testing using a commercial panel of 92 connective tissue disorder genes, revealing a variant in *TGFBR2* (NM_003242.6 (TGFBR2):c.412T>G) classified as likely pathogenic by a commercial genetic testing lab with a proprietary variant calling software based on ACMG guidelines. No other pathogenic or likely pathogenic variants were reported in genes associated with any connective tissue disorders. He was subsequently diagnosed with LDS based on the presence of the likely pathogenic *TGFBR2* variant, his aortic root enlargement, and exclusion of connective tissue disorders with overlapping clinical features (Loeys and Dietz, 2008).

The patient's similarly affected sons also harbor this *TGFBR2* variant, which was detected using targeted testing. The older son (III:

4, Figure 1A) has a history of complete heart block at 10 months old, microretrognathia, kyphosis, and high palate. The younger son (III:5, Figure 1A) has microretrognathia, kyphosis, high palate, and normal echocardiogram. A diagnosis can be made in at-risk relatives of a person with a clinical diagnosis of LDS by identification of the pathogenic variant even if the other features are not yet present (Loeys and Dietz, 2008). The proband's two sons (Individuals III:4 and III:5, Figure 1A) were diagnosed with Loays-Dietz based on the presence of the diagnostic variant identified in their father.

The *TGFBR2* variant NM_003242.6 (TGFBR2):c.412T>G p. Cys138Gly is present in gnomAD v3.1.2 at extremely low frequency (1/152,174, 0.0007%), and computational meta-predictors suggest that this variant is likely to be disruptive (CADD 26.3, REVEL 0.93) (Rentzsch, et al., 2021; Ioannidis et al., 2016). In addition to this proband and his sons, this variant has been reported as likely pathogenic in one additional patient with clinical features of *TGFBR2*-related disorders in the commercial laboratory's internal database, and was reported to

ClinVar as likely pathogenic by a single submitter, Labcorp Genetics (formerly Invitae), with assertion criteria (one star review status). Another variant impacting the same position and reported (ClinVarID 213915) (NM_003242.6 (TGFB2):c.412T>C p. Cys183Arg) is further evidence supporting the likely pathogenic classification. This is a *de novo* variant in a patient with spontaneous cervical artery dissection and recurrent knee subluxations (Pezzini et al., 2011).

3.2 Identification of NM_003238.6 (TGFB2):c.755-6T>C

Proband III:1, the niece of II:3, also presented for evaluation at the Smith Family Clinic for Genomic Medicine (Figures 1A–C). Her medical history was significant for bicuspid aortic valve, microretrognathia requiring surgical repair, scoliosis requiring bracing, and dental crowding requiring extraction of 12 adult teeth. A recent echocardiogram revealed no aortic dilation or aneurysm. Upon physical exam, she was noted to have several features consistent with a connective tissue disorder. She has a long facial appearance, slightly downslanting palpebral fissures, very high-arched palate, and prominent sternum with inferior bilateral prominence at the rib attachments. Additionally, she has hyperextensible skin (4.5 cm neck, 1.5 cm volar surface of forearm), hypermobility of the shoulders and thumbs, and arachnodactyly (positive wrist and thumb sign) with long toes. She did not meet the established clinical diagnostic criteria for Marfan syndrome (Loeys et al., 2010).

The proband underwent targeted testing for the familial NM_003242.6 (TGFB2):c.412T>G p. Cys138Gly variant but it was not observed. She subsequently underwent testing using the same commercial 92 gene panel for connective tissue disorders, which was also negative but did detect a likely benign variant NM_003238.6 (TGFB2):c.755-6T>C. This variant was submitted to ClinVar (RCV002835216.2) but not reported to the proband because of its likely benign classification at that time. The family then sought whole genome sequencing through the HudsonAlpha Clinical Services Lab (CLIA Number 01D2086581, CAP Number 8051488). The same TGFB2 variant [NM_003238.6 (TGFB2):c.755-6T>C] was identified, now known to be inherited from her mother (II:1; Figures 1A, D, E), confirmed by Sanger sequencing in both the proband and her mother (Supplemental Figure S1). Her mother is more mildly affected with mitral valve prolapse, kyphoscoliosis, congenital club foot, high palate with dental crowding, deviated nasal septum, and myopia. The mother also tested negative for the NM_003242.6 (TGFB2):c.412T>G p. Cys138Gly variant identified in her brother. No other variants of interest were revealed upon analysis.

The NM_003238.6 (TGFB2):c.755-6T>C variant is located in the 3' splice acceptor region between exons 4 and 5 of the TGFB2 protein. This variant is absent from gnomAD v3.1.2, and has mild predictive scores in SpliceAI (acceptor loss: 0.10, +6bp; acceptor gain: 0.03, +121 bp, CADD score: 19.78) indicating a low probability of this variant resulting in a 115 bp deletion from exon 5 compared to the canonical transcript. A formal diagnosis of LDS is not possible without the identification of a pathogenic variant in an associated gene (Loeys and Dietz, 2008). The TGFB2 variant [NM_

003238.6 (TGFB2):c.755-6T>C] was classified as a variant of uncertain significance by the HudsonAlpha Clinical Services Lab (PM2_Supporting). The variant was thus reported to the family as a possible molecular diagnosis given the classification as a variant of uncertain significance, the overlap of the patient's symptoms (presence of characteristic craniofacial, skeletal, and cutaneous features); segregation in the family; and exclusion of a diagnosis of overlapping clinical features with TGFB2-related disorders, and segregation with related phenotypes in the family.

3.3 Characterization of previously identified TGFB2 variants with predicted splicing impacts

Given that NM_003238.6 (TGFB2):c.755-6T>C is near the exon-intron border, we hypothesized that it could impact splicing. We also identified additional variants in the region using the ClinVar database and assessed their potential to alter splicing. These included a common benign and likely benign TGFB2 variants unlikely to cause disease as negative controls, a positive control variant classified as pathogenic based on known splicing impact, and a likely benign variant with positive predictors of pathogenicity (Table 1).

The positive control variant NM_003238.6 (TGFB2):c.755-5_755-2delinsG has a ClinVar (528,887) ACMG classification of likely pathogenic, is absent from gnomAD v3.1.2, has paired acceptor gain and loss of 0.14 and 0.91, according to SpliceAI, and predicts the same alternative splicing product as the proband.

Two negative control variants were chosen based on their increased allele count (AC) in gnomAD v3.1.2. The first is an intronic variant, NM_003238.6 (TGFB2):c.755-5dup, with a gnomAD AC of 7, CADD score of 14.51, and a predicted SpliceAI donor gain score of 0.03, 207 bp downstream of the variant. The second variant is exonic, NM_003238.6 (TGFB2):c.747A>G (p.Arg249 =), with a gnomAD AC of 45, CADD score of 14.98, and a splice gain and loss scores of 0.02 and 0.18, respectively.

We also tested a likely benign variant identified in ClinVar, NM_003238.6 (TGFB2):c.644-18T>A that is absent from gnomAD v3.1.2, a CADD score of 22.3, and has a paired splice acceptor loss (SpliceAI score: 0.57, 18 bp downstream of variant) and gain (SpliceAI score 0.02, 2 bp downstream), which predicts an alternative splicing product with a 16 bp insertion in exon 4 of the transcript, suggesting a possible splicing impact.

3.4 Genomic and molecular analysis

Given the proximity of the proband's variant to the splice site, we wanted to determine whether splicing of the TGFB2 transcript was impacted. However, TGFB2 is not highly expressed in accessible patient tissues like blood and skin (www.proteinatlas.org), so we could not assess the impact directly from patient samples using RNA sequencing or similar *in vivo* analyses. Instead, we designed a splice reporter minigene assay based on the pSpliceExpress vector system (Kishore et al., 2008). The system contains gene fragments of TGFB2 reference sequence (REF), proband, or variants of interest (Table 1) consisting of an intron 2 fragment with 5' end flanked by rat insulin exon (REI2), exon 3, intron 3, exon 4, intron 4, exon 5, and intron

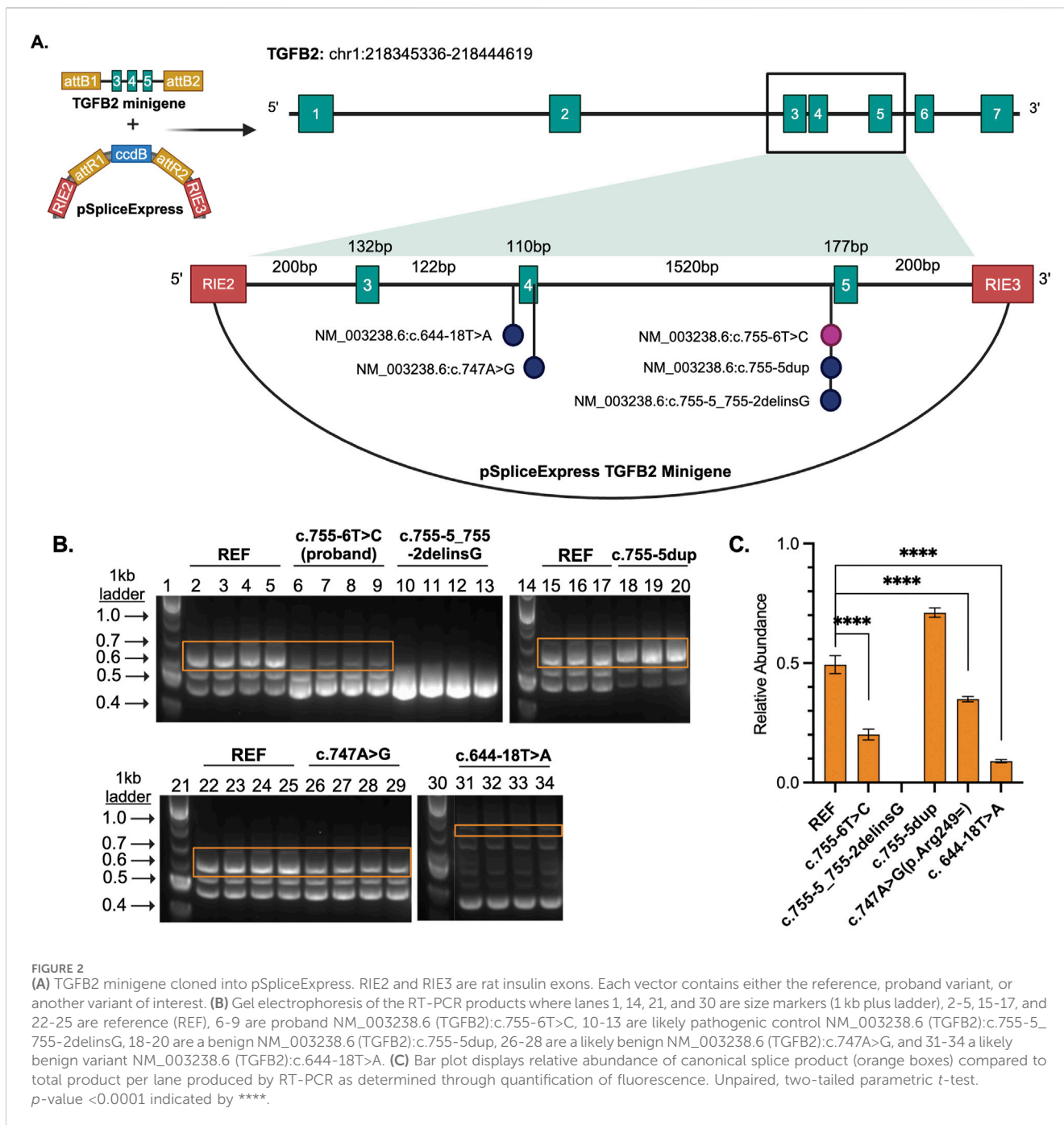


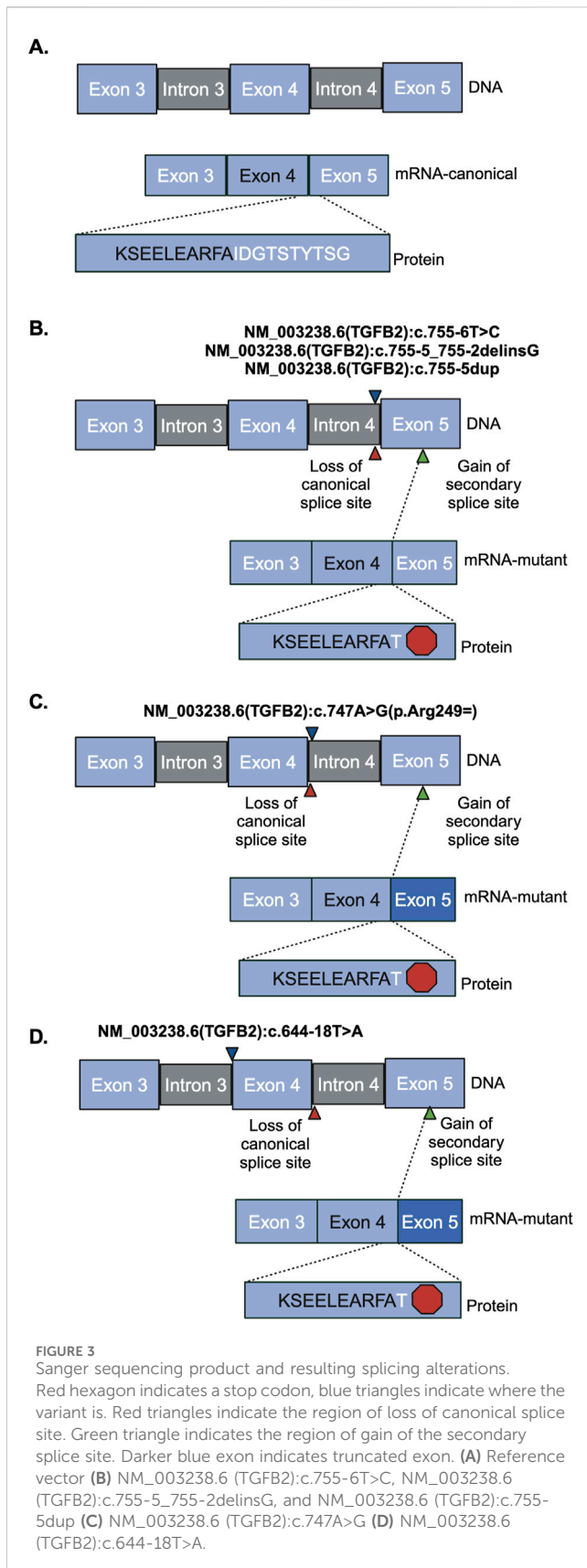
FIGURE 2

(A) TGFB2 minigene cloned into pSpliceExpress. RIE2 and RIE3 are rat insulin exons. Each vector contains either the reference, proband variant, or another variant of interest. (B) Gel electrophoresis of the RT-PCR products where lanes 1, 14, 21, and 30 are size markers (1 kb plus ladder), 2-5, 15-17, and 22-25 are reference (REF), 6-9 are proband NM_003238.6 (TGFB2):c.755-6T>C, 10-13 are likely pathogenic control NM_003238.6 (TGFB2):c.755-5_755-2delinsG, 18-20 are a benign NM_003238.6 (TGFB2):c.755-5dup, 26-28 are a likely benign NM_003238.6 (TGFB2):c.747A>G, and 31-34 a likely benign variant NM_003238.6 (TGFB2):c.644-18T>A. (C) Bar plot displays relative abundance of canonical splice product (orange boxes) compared to total product per lane produced by RT-PCR as determined through quantification of fluorescence. Unpaired, two-tailed parametric *t*-test. *p*-value <0.0001 indicated by ****.

5 fragment's 3' end flanked by rat insulin exon (REI3) (Figure 2A; Supplemental Table S1).

The region containing these variants was inserted into the splice reporter plasmid, transfected into HEK293T cells where the artificial transcript is expressed under the control of a constitutively active promoter. RNA was extracted from cells and the resulting transcripts were analyzed using gel electrophoresis and sequencing. Amplifying the transcripts with common primers resulted in a series of products representing multiple splice products. We ran the products on an agarose gel and quantified the intensity of the bands (Figures 2B, C; Supplemental Figures S2, S3) to demonstrate that there is significant reduction in full length product between the reference vector and the proband sequence,

NM_003238.6 (TGFB2):c.755-6T>C ($p < 0.0001$). We also observed reduced full length product with the variants NM_003238.6 (TGFB2):c.747A>G ($p < 0.0001$) and NM_003238.6 (TGFB2):c.644-18T>A ($p < 0.0001$). All three of these variants were classified as likely benign in ClinVar but appear to impact splicing. Sanger sequencing (Supplemental Table S2) suggests that the NM_003238.6 (TGFB2):c.755-6T>C, proband sequence, results in the use of a secondary splice acceptor 117 base pairs into exon 5. This altered transcript would result in a premature stop codon and could either lead to a truncated protein or be subject to nonsense-mediated decay (Figures 3A, B). As expected, the variant previously classified as Likely Pathogenic, NM_003238.6 (TGFB2):c.755-5_755-2delinsG, also shows no measurable level of the full



length (REF) transcript. We do note that all other sequences, including the REF, show low levels of the alternate splice product. (Figures 2B, C; Supplemental Figure S2, S3).

4 Discussion

While there have been advances in characterizing intronic and synonymous splicing variants causative of Mendelian disease, they are still underexplored as a molecular diagnosis. In this study, we explore several variants that occur near splice junctions of *TGFβ2*, a gene linked to the connective tissue disorder Loays-Dietz Syndrome. We characterize the function of variants observed in patients and prioritize using computational predictions. Our functional data provide insight into the biological consequences of those variants. This work was motivated by the identification of an individual with features of LDS who did not have a molecular diagnosis. In fact, using current diagnostic conventions, fewer than half of individuals with LDS and other connective tissue disorders receive a molecular diagnosis and in those cases diagnosis relies on clinical features and phenotype alone (Renner et al., 2019). This limitation is not specific to LDS. Genetic diagnosis of rare disease patients is possible in approximately 50% of rare disease cases with current sequencing technologies. Continuing efforts to annotate and interpret noncoding and splicing variants combined with improved technologies for detecting repeat expansions and structural variation are slowly increasing that rate (Abul-Husn, et al., 2023; Felker, et al., 2023; Hiatt, et al., 2021).

The lack of molecular diagnosis limits a clinician's ability to accurately diagnose patients and thus direct proper care. Here, we present a study of a unique family with both confirmed LDS and features of LDS. Several family members' symptoms can be explained by a missense variant in *TGFβR2*, another gene associated with LDS. The proband in our study is a member of the same family, but lacks the established *TGFβR2* variant. Instead, the proband has a rare intronic variant in *TGFβ2* (GnomAD v4.1.0 exome allele frequency 6.949e-7). This variant occurs 5 bp upstream of the splice acceptor and SpliceAI gives it a non-zero score indicating a possibility that it affects splicing. Our hypothesis is that the variants that favor the alternate acceptor reduce the level of functional protein leading to the observed phenotypes. Given the rarity of the variant, the proximity to the splice site, the SpliceAI prediction, and no other relevant variants observed; we sought to characterize the variant further. We acknowledge the statistical rarity (1 in 10, 000, 000,000) of having two independent variants in a pedigree that are causative for LDS. However, there is supporting evidence for this improbable explanation: 1) presence of only one of the two variants in affected members, 2) absence of any other candidate variants to explain the LDS phenotypes, and 3) functional data detailed in this paper.

We measured the impact on splicing of several intronic *TGFβ2* variants. The SpliceAI algorithm predicted that the variant observed in our proband [NM_003238.6 (TGFβ2):c.755-6T>C] and a rare variant reported in ClinVar [NM_003238.6 (TGFβ2):c.644-18T>A] may interfere with canonical splicing of *TGFβ2*. Both variants reduce the level of protein-generating transcript in favor of an isoform that results in a premature stop codon by formation of an alternate acceptor 117 bp downstream in exon 5. This splice isoform was also observed when testing the likely pathogenic variant [NM_003238.6 (TGFβ2):c.755-5_755-2delinsG]. Isoforms containing this splice junction have been previously reported (NR_138148.2, NR_138149.2), and result in nonsense-mediated decay (O'Leary et al., 2016; Strausberg et al., 2002), so it is possible that these isoforms are less common but perhaps normal splicing products of *TGFβ2* expression. Other common variants in

this region did not significantly alter splicing efficiency [NM_003238.6 (TGFB2):c.755-5dup and NM_003238.6 (TGFB2):c.747A>G (p.Arg249 =)]. In addition, at least one other variant impacting splicing of *TGFB2* that introduces a similar premature stop has been associated with loss of protein function (Arnaud et al., 2019). Our data support the hypothesis that the variant observed in the proband favors an isoform that creates a premature stop ultimately reducing the level of functional protein.

Our findings provide insight into the variability in phenotype and genotype in LDS, highlighting the potential importance of splicing efficiency. Given the variable expressivity and even rare incomplete penetrance observed among patients with LDS and specifically within this family it seems plausible that phenotypic differences observed between families with different disease-causing variants could partly be explained by the level of canonical transcript. This finding allows for new ACMG codes applicable to the proband in consideration of molecular studies: PM2_Supporting and PP1_Supporting (Walker et al., 2023).

This study serves the scientific and clinical communities of LDS and connective tissue disorders by contributing to the growing body of evidence supporting the relevance of intronic splicing variants in transforming growth factors and their receptors as a molecular cause for LDS (Fujiwara et al., 2018). Our work provides new evidence supporting the potential relevance of these variants to disease and demonstrates the importance of functional assays together with computational predictions to reveal the impacts that putative disease-causing variants have on phenotype. Together with clinical information, functional data and computational predictions together will contribute to improved diagnostic rates and ultimately better patient outcomes.

Data availability statement

Publicly available data was analyzed in this study. This data can be found here: ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) under the accession number SCV003221034.2. Additional datasets presented in this article are not readily available because data sharing was not consented with the patients under the IRP protocol used.

Ethics statement

The studies involving humans were approved by the Western Institutional Review Board protocol no. 0072 under IRB Protocol #20161118. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

EG: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software,

Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing. SF: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Writing—original draft. TC: Conceptualization, Data curation, Writing—review and editing. NS: Data curation, Formal Analysis, Writing—review and editing. JP: Data curation, Writing—review and editing. MC: Data curation, Supervision, Writing—review and editing. AH: Conceptualization, Investigation, Supervision, Writing—review and editing. SC: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Support for this work comes from the HudsonAlpha Foundation Hero Fund and the State of Alabama's Research Fund to HudsonAlpha Institute. SC is supported by NIH grant UL1TR003096. SF is supported by 5T32HG008961-08.

Acknowledgments

We would like to acknowledge the institutional support at the HudsonAlpha Institute for Biotechnology that makes translational collaborations like this possible. We would also like to thank our patients at the Smith Family Clinic for Genomic Medicine for sharing their medical histories, family history, and photographs for publication. We would like to thank Carter Wright for helping connect the authors and motivate this study. We would like to thank Susan Hiatt, Owen Hirschi, and James Lawlor for their contributions to computational biology and analysis. Finally, we thank members of several labs for their feedback on this case and manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1435734/full#supplementary-material>

References

- Abul-Husn, N. S., Marathe, P. N., Kelly, N. R., Bonini, K. E., Sebastin, M., Ogdiss, J. A., et al. (2023). Molecular diagnostic yield of genome sequencing versus targeted gene panel testing in racially and ethnically diverse pediatric patients. *Genet. Med.* 25 (9), 100880. Epub 2023 May 6. PMID: 37158195; PMCID: PMC10789486. doi:10.1016/j.gim.2023.100880
- Arnaud, P., Hanna, N., Benarroch, L., Aubart, M., Bal, L., Bouvagnet, P., et al. (2019). Genetic diversity and pathogenic variants as possible predictors of severity in a French sample of nonsyndromic heritable thoracic aortic aneurysms and dissections (nshTAAD). *Genet. Med.* 21 (9), 2015–2024. Epub 2019 Feb 11. PMID: 30739908. doi:10.1038/s41436-019-0444-y
- Behera, S., LeFaive, J., Orchard, P., Mahmoud, M., Paulin, L. F., Farek, J., et al. (2023). FixItFelix: improving genomic analysis by fixing reference errors. *Genome Biol.* 24 (1), 31. PMID: 36810122; PMCID: PMC9942314. doi:10.1186/s13059-023-02863-7
- Felker, S. A., Lawlor, J. M. J., Hiatt, S. M., Thompson, M. L., Latner, D. R., Finnila, C. R., et al. (2023). Poison exon annotations improve the yield of clinically relevant variants in genomic diagnostic testing. *Genet. Med.* 25 (8), 100884. Epub 2023 May 6. PMID: 37161864; PMCID: PMC10524927. doi:10.1016/j.gim.2023.100884
- Freed, D., Pan, R., Chen, H., Li, Z., Hu, J., and Aldana, R. (2022). DNAscope: high accuracy small variant calling using machine learning. *bioRxiv*. [Preprint]. doi:10.1101/2022.05.20.492556
- Fujiwara, T., Takeda, N., Hara, H., Morita, H., Kishihara, J., Inuzuka, R., et al. (2018). Distinct variants affecting differential splicing of TGFBR1 exon 5 cause either Loeys-Dietz syndrome or multiple self-healing squamous epithelioma. *Eur. J. Hum. Genet.* 26 (8), 1151–1158. Epub 2018 Apr 30. PMID: 29706644; PMCID: PMC6057981. doi:10.1038/s41431-018-0127-1
- Hiatt, S. M., Lawlor, J. M. J., Handley, L. H., Ramaker, R. C., Rogers, B. B., Partridge, E. C., et al. (2021). Long-read genome sequencing for the molecular diagnosis of neurodevelopmental disorders. *HGG Adv.* 2 (2), 100023. Epub 2021 Jan 16. PMID: 33937879; PMCID: PMC8087252. doi:10.1016/j.xhgg.2021.100023
- Hirschi, O. R., Felker, S. A., Rednam, S. P., Vallance, K. L., Parsons, D. W., Roy, A., et al. (2024). Combined bioinformatic and splicing analysis of likely benign intronic and synonymous variants reveals evidence for pathogenicity. *Genet. Med. Open* 2, 101850. PMID: 37961416; PMCID: PMC10635218. doi:10.1016/j.gimo.2024.101850
- Holt, J. M., Kelly, M., Sundlof, B., Nakouzi, G., Bick, D., and Lyon, E. (2021). Reducing Sanger confirmation testing through false positive prediction algorithms. *Genet. Med.* 23 (7), 1255–1262. Epub 2021 Mar 25. PMID: 33767343; PMCID: PMC8257489. doi:10.1038/s41436-021-01148-3
- Ikonomidis, J. S. (2017). Aortic surgery in Loeys-Dietz syndrome: excellent results, but close surveillance is necessary. *J. Thorac. Cardiovasc Surg.* 153 (2), 413–414. Epub 2016 Nov 16. PMID: 27919459; PMCID: PMC5916778. doi:10.1016/j.jtcvs.2016.11.010
- Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., et al. (2016). REVEL: an Ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* 99 (4), 877–885. Epub 2016 Sep 22. PMID: 27666373; PMCID: PMC5065685. doi:10.1016/j.ajhg.2016.08.016
- Isselbacher, E. M., Preventza, O., Hamilton Black, J., Augoustides, J. G., Beck, A. W., Bolen, M. A., et al. (2022). 2022 ACC/AHA guideline for the diagnosis and management of aortic disease: a report of the American heart association/American college of cardiology joint committee on clinical practice guidelines. *Circulation* 146 (24), e334–e482. Epub 2022 Nov 2. PMID: 36322642; PMCID: PMC9876736. doi:10.1161/CIR.0000000000001106
- Jaganathan, K., Kyriazopoulou, P. S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., et al. (2019). Predicting splicing from primary sequence with deep learning. *Cell* 176 (3), 535–548.e24. Epub 2019 Jan 17. PMID: 30661751. doi:10.1016/j.cell.2018.12.015
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581 (7809), 434–443. doi:10.1038/s41586-020-2308-7
- Kishore, S., Khanna, A., and Stamm, S. (2008). Rapid generation of splicing reporters with pSpliceExpress. *Gene* 427 (1–2), 104–110. Epub 2008 Oct 1. PMID: 18930792; PMCID: PMC2821805. doi:10.1016/j.gene.2008.09.021
- Krogh-Sørensen, K., Lingaas, P. S., Lundblad, R., Seem, E., Paus, B., and Geiran, O. R. (2017). Cardiovascular surgery in Loeys-Dietz syndrome types 1–4. *Eur. J. Cardiothorac. Surg.* 52 (6), 1125–1131. PMID: 28541520. doi:10.1093/ejcts/ezx147
- Loeys, B. L., and Dietz, H. C. (2008). “Loeys-dietz syndrome,” in *GeneReviews*® M. P. Adam, J. Feldman, G. M. Mirzaa, et al. (Seattle (WA): University of Washington, Seattle), 1993–2023. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK1133/>.
- Loeys, B. L., Dietz, H. C., Braverman, A. C., Callewaert, B. L., De Backer, J., Devereux, R. B., et al. (2010). The revised Ghent nosology for the Marfan syndrome. *J. Med. Genet.* 47 (7), 476–485. PMID: 20591885. doi:10.1136/jmg.2009.072785
- Loeys, B. L., Schwarze, U., Holm, T., Callewaert, B. L., Thomas, G. H., Pannu, H., et al. (2006). Aneurysm syndromes caused by mutations in the TGF-beta receptor. *N. Engl. J. Med.* 355 (8), 788–798. PMID: 16928994. doi:10.1056/NEJMoa055695
- Loughborough, W. W., Minhas, K. S., Rodrigues, J. C. L., Lyen, S. M., Burt, H. E., Manghat, N. E., et al. (2018). Cardiovascular manifestations and complications of loeys-dietz syndrome: CT and MR imaging findings. *Radiographics* 38 (1), 275–286. PMID: 29320330. doi:10.1148/rg.2018170120
- MacCarrick, G., Black, J. H., Bowdin, S., El-Hamamsy, I., Frischmeyer-Guerrero, P. A., Guerrero, A. L., et al. (2014). Loeys-Dietz syndrome: a primer for diagnosis and management. *Genet. Med.* 16 (8), 576–587. Epub 2014 Feb 27. PMID: 24577266; PMCID: PMC4131122. doi:10.1038/gim.2014.11
- Mazzella, J. M., Frank, M., Collignon, P., Langeois, M., Legrand, A., Jeunemaitre, X., et al. (2017). Phenotypic variability and diffuse arterial lesions in a family with Loeys-Dietz syndrome type 4. *Clin. Genet.* 91 (3), 458–462. Epub 2016 Sep 13. PMID: 27440102. doi:10.1111/cge.12838
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., et al. (2016). The ensembl variant effect predictor. *Genome Biol.* 17 (1), 122. PMID: 27268795; PMCID: PMC4893825. doi:10.1186/s13059-016-0974-4
- O’Leary, N. A., Wright, M. W., Brister, J. R., Ciuffo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44 (D1), D733–D745. Epub 2015 Nov 8. PMID: 26553804; PMCID: PMC4702849. doi:10.1093/nar/gkv1189
- Pezzini, A., Drera, B., Del Zotto, E., Ritelli, M., Carletti, M., Tomelleri, G., et al. (2011). Mutations in TGFBR2 gene cause spontaneous cervical artery dissection. *J. Neurol. Neurosurg. Psychiatry* 82 (12), 1372–1374. Epub 2011 Jan 26. PMID: 21270064. doi:10.1136/jnnp.2010.231902
- Renner, S., Schüler, H., Alawi, M., Kolbe, V., Ryczynski, M., Woitschach, R., et al. (2019). Next-generation sequencing of 32 genes associated with hereditary aortopathies and related disorders of connective tissue in a cohort of 199 patients. *Genet. Med.* 21 (8), 1832–1841. Epub 2019 Jan 24. PMID: 30675029. doi:10.1038/s41436-019-0435-z
- Rentsch, P., Schubach, M., Shendure, J., and Kircher, M. (2021). CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med.* 13 (1), 31. PMID: 33618777. doi:10.1186/s13073-021-00835-9
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17 (5), 405–424. doi:10.1038/gim.2015.30
- Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., et al. (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U. S. A.* 99 (26), 16899–16903. Epub 2002 Dec 11. PMID: 12477932; PMCID: PMC139241. doi:10.1073/pnas.242603899
- Walker, L. C., Hoya, M., Wiggins, G. A. R., Lindy, A., Vincent, L. M., Parsons, M. T., et al. (2023). Using the ACMG/AMP framework to capture evidence related to predicted and observed impact on splicing: recommendations from the ClinGen SVI Splicing Subgroup. *Am. J. Hum. Genet.* 110 (7), 1046–1067. Epub 2023 Jun 22. PMID: 37352859; PMCID: PMC10357475. doi:10.1016/j.ajhg.2023.06.002