



Variability in gene-based knowledge impacts variant classification: an analysis of *FBNI* missense variants in ClinVar

Linnea M. Baudhuin¹ · Michelle L. Kluge¹ · Katrina E. Kotzer¹ · Susan A. Lagerstedt¹

Received: 20 August 2018 / Revised: 13 May 2019 / Accepted: 21 May 2019 / Published online: 21 June 2019
© The Author(s), under exclusive licence to European Society of Human Genetics 2019

Abstract

Gene-specific knowledge can enhance genetic variant classification, but may not be routinely incorporated into clinical laboratory practice. For example, *FBNI* variants associated with Marfan syndrome may be variably classified depending on knowledge of *FBNI*-specific critical regions. In order to assess variability in classification of *FBNI* variants, 674 *FBNI* missense variants from 18 ClinVar submitters were compared and reanalyzed using *FBNI*-specific criteria and ACMG/AMP 2015 guidelines for variant interpretation. Conflicting variant classifications occurred in 30.7% of the missense variants that had multiple submitters. There were 451 classifications of 361 critical residue missense variants, with 80.0% (361/451) classified as likely pathogenic or pathogenic [(L)P]. Non-cysteine critical residue variants were less likely to be classified as (L)P [55.3% (78/141)] than cysteine variants [91.3% (283/310)] and were more likely to lack evidence citing the functional significance of the amino acid impacted. Application of *FBNI*-specific knowledge allowed for reclassification or discrepancy resolution in 65/361 (18.0%) critical residue variants. There were 522 classifications of 313 unique missense variants not known to impact a critical residue. Of these, 31.6% (165/522) were likely overclassified as either (L)P or uncertain significance (VUS), especially when minor allele frequency (MAF) was taken into account, and we reclassified or resolved classification discrepancies in 128/313 (40.9%) of these variants. Our results provide a refined framework and resource for *FBNI* variant classification, and further supports the more global implications of combining gene-based knowledge with ACMG/AMP criteria and appropriate MAF cutoffs for variant classification that extend beyond *FBNI*.

Introduction

In the past several years, the quality of variant classification for Mendelian disorders has been improved through data sharing which has been facilitated, in part, through databases such as ClinVar, an NIH open-access database of clinically observed variants and their classifications. However, one of the most recognized challenges with databases such as ClinVar is the number of discordantly or improperly classified variants, which has been attributed to factors such as the age of the entries, nonclinical submissions, and

differences in application of variant interpretation guidelines [1–5].

The contribution of gene-based knowledge to the process of variant classification is beginning to be more widely recognized. The American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) 2015 variant interpretation guidelines included some criteria (PM1, PVS1, and PP2) which allowed for incorporation of gene-based knowledge into variant classification [6]. In addition, ClinGen (the Clinical Genome Resource) has established working groups to develop consensus methods for variant interpretation in specific diseases and genes, producing, for example, gene-specific guidance for *MYH7*-associated inherited cardiomyopathies and RASopathies [7, 8]. Furthermore, Muñio-Mosquera et al. [9] recently published pilot disease- and gene-specific guidelines for reclassification of *FBNI* variants identified in their own laboratory. Gene-specific knowledge can improve variant classification in the following ways: (1) knowledge of critical residues for which the PM1 criteria can be applied allows for variants at these

Supplementary information The online version of this article (<https://doi.org/10.1038/s41431-019-0440-3>) contains supplementary material, which is available to authorized users.

✉ Linnea M. Baudhuin
baudhuin.linnea@mayo.edu

¹ Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

residues that would otherwise be classified as variants of uncertain significance (VUS) to be classified as likely pathogenic (LP) or pathogenic (P); (2) appropriate MAF cutoffs based on disease incidence allow for variants that would otherwise be classified as VUS (or LP) to be classified as likely benign or benign (or VUS) by using BS1 and BA1; and (3) improved inter- and intra-laboratory consistency in variant classification. Together, this leads to a reduction in the number of VUSs reported, and ultimately results in better service to practitioners and improved patient care.

Marfan syndrome (MFS), due to variants in *FBNI*, is an autosomal dominant connective tissue multisystem disorder [10]. There are over 2000 published *FBNI* variants, and many are unique to individual families. Missense variants, especially cysteine (Cys) substitutions, are the most common type of *FBNI* variant [11, 12]. The type of *FBNI* variant identified and the likelihood of that variant being pathogenic are recognized as important factors when making a diagnosis of MFS, with *de novo* (in the absence of family history), nonsense, frameshift, splicing, and missense substitutions of conserved residues considered most likely to be pathogenic [13].

In order to optimize classification of *FBNI* variants, it is important to have a thorough understanding of fibrillin-1 function and structure. Fibrillin-1 is a 350 kDa calcium-binding glycoprotein with a highly modular organization [14, 15]. Fibrillin-1 is composed of a repetition of two disulfide-rich modules: 47 epidermal growth factor (EGF)-like domains and seven transforming growth factor β -binding protein-like (TB) domains. The majority of fibrillin-1 consists of EGF-like domains, and 43/47 EGF-like domains are calcium-binding (cbEGF-like domains). There are also two hybrid domains, which contain properties of both cbEGF-like and TB domains. Cys residues play an important and well-studied role in fibrillin-1 structure, and occur in sets of six and eight in each EGF-like (calcium- and noncalcium-binding) and TB domain, respectively. *FBNI* variants that disrupt or create new Cys in these domains are known to lead to the MFS phenotype and other fibrillinopathies [13, 16–18]. Muiño-Mosquera et al. [9] highlighted the importance of these Cys residues in their suggested *FBNI* guidelines. They suggested that disruption of a Cys in the cbEGF-like domains should qualify as a strong criteria (PS3), and disruption of all other Cys residues in *FBNI*, creation of a new Cys in the cbEGF-like domains, and disruption of calcium-binding residues in the cbEGF-like domains should all qualify as a moderate criteria (PM1) [13]. However, Muiño-Mosquera et al. did not provide guidance on numerous other critical residues of fibrillin-1 that allow for further refinement of the application of the PM1 criterion (detailed in materials and methods and Fig. 1), nor did they discuss application of the PP2 criterion

(missense variant in a gene which has a low rate of benign missense variation and in which missense variants are a well-known disease mechanism).

While there are many readily accessible tools and variant databases, such as ClinVar, that provide clinical laboratories access to previously reported genetic variants, these tools may not provide sufficient evidence to perform an accurate independent variant classification. Furthermore, clinical laboratories may lack the resources to gather the necessary knowledge regarding critical residues in their gene of interest in order to develop gene-specific criteria for variant classification.

In order to better understand how clinical laboratories are classifying variants affecting critical and noncritical residues in *FBNI*, we analyzed *FBNI* variant classifications and associated evidence submitted to the ClinVar database. After this analysis, we applied gene-specific guidelines, including refined criteria for PM1 based on known critical residues, to reanalyze the variants and determine if application of these guidelines and knowledge help to reclassify or resolve discrepant classifications of variants listed in ClinVar.

Materials and methods

Data retrieval

ClinVar data were downloaded on Dec 12, 2017 for a total of 1405 unique *FBNI* variants with 1963 classifications from 18 different submitters. On February 13, 2019, downloaded data were reanalyzed and updated, and 21 additional missense variants were added from another search of ClinVar. Names of submitting groups were blinded for the analysis and variant submission evidence provided was reviewed manually.

Categorization of *FBNI* critical residue variants

Missense variants were categorized by domain and then investigated as to whether they impacted consensus/critical amino acid residues or otherwise were likely to impact fibrillin function (Table 1). Such missense variants that occurred in each domain (Fig. 1) included those that: (1) created or destroyed a Cys residue in a cbEGF-like domain [14, 16–19] (Fig 1a); (2) altered any of the five residues in the calcium-binding consensus sequence D/N-x-D/N-E/Q-Xm-D/N-Xn-Y/F in a cbEGF-like domain [note: an asparagine (N) to serine (S) at the second D/N position maybe tolerated] [18, 20–25] (Fig 1a); (3) altered the critical glycine (Gly)/alanine (Ala) residue between Cys2–Cys3 in a subset of cbEGF-like domains [26–30] (Fig 1a); (4) altered interdomain packing residues in a subset of cbEGF-like

A

Exon (s)	Domain	5' domain	5'AA	amino acid sequence	3'AA	3' domain
2-3	EGF1		81	IVPT CRHS ----CGDGF-- CSR P-----NM CTCP SGQ---IAP-- SCG	112	-EGF
3-4	EGF2	EGF-	115	SIQH CNIR ----CMNGG S - CS D-----DH CLCQ KGY---IGT-- HCG	146	-EGF
4-5	EGF3	EGF-	147	-QP VCESG ----CLNGGR- CV AP-----NR CACT YGF---TGP-- QCE	178	-TB
7	cbEGF1	HYB-	246	DVDE CQAIPGL-C Q GN-- CINT V- GSF CK CP AGHKLN EV SQK- CE (D)	287	-cbEGF
8	cbEGF2	cbEGF-	288	DIDE CS TIPGI - CE GG- CINT V- SSY CK CP PGFY TS PD GTR - CI (D)	329	-TB
11	EGF2	TB-	449	VT DY Q QLV RYL- CQ NR-- CIP TP- GSCR CE NC KGF QLD LRG-- EC -	489	-cbEGF
12	cbEGF3	EGF-	490	DVDE CE KNP --- CAG GE-- CIN Q- GSY TC QCR AG Y Q ST LT RTE - CR (D)	529	-cbEGF
13	cbEGF4	cbEGF-	530	DIDE CL QNGRI - CN NR-- CINT DD- SFHC VC NA GH VTR DG KN - CE (D)	571	-cbEGF
14	cbEGF5	cbEGF-	572	DMDE CS IRNM -- CL NGM-- CIN ED- SF CK CP KG QL AS DGRY - CK (D)	612	-cbEGF
15	cbEGF6	cbEGF-	613	DINE CE TPGI -- CM NR-- CV NTD- SRY CE CP GL AV LD RV -- CV (D)	653	-TB
18	cbEGF7	TB-	723	DINE CA LD DP CI - CP NGI-- CEN LR- GTY K CI NS G EV D ST GKN - CV (D)	764	-cbEGF
19	cbEGF8	cbEGF-	765	DINE CV LNSLL - CD NGQ-- CR NTP- SFV CT CP K G - IY K PD L KT - CE (D)	806	-cbEGF
20	cbEGF9	cbEGF-	807	DIDE CE SSP --- CIN GV-- CN SP- SF ICE CS SE STL DP T K TI - CI (E)	846	-HYB
23	cbEGF10	HYB-	910	DIDE CE VFP GV- CK NGL-- CV NTR- GSF CK CP SG MTL D AT G RI - CL (D)	951	-TB
25	cbEGF11	TB-	1028	DINE CK MPISL - CT HGK-- CR NTI- GSF CK CD SG ALD SE ERN - CT (D)	1069	-cbEGF
26	cbEGF12	cbEGF-	1070	DIDE CR IS PD L - CG RQ-- CV NTP- DFE CK CE DE ES CF MM K NCM (D)	1112	-cbEGF
27	cbEGF13	cbEGF-	1113	DIDE CQ RD PL L - CR GGV-- CH NTE- GSY RC CE PP GL SP NI SA- CI (D)	1154	-cbEGF
28	cbEGF14	cbEGF-	1155	DINE CE LSAHL - CP NGR-- CV NLI- KY Q CA CP GH ST PD RL F - CV (D)	1196	-cbEGF
29	cbEGF15	cbEGF-	1197	DIDE CS IM GG- CE TF-- CIN SE- GSY EC CP GP AL MP DQR S- CT (D)	1237	-cbEGF
30	cbEGF16	cbEGF-	1238	DIDE CE DNP NI- CD GQ-- CT NIP- GEY RC LY D G MA SE DM KT - CV (D)	1279	-cbEGF
31	cbEGF17	cbEGF-	1280	DVNE CD LNP NI- CL SGT-- CEN TR- SF IC CD MG SK K GKT G- CT (D)	1321	-cbEGF
32	cbEGF18	cbEGF-	1322	DINE CE IGA HN- CG KH AV - CT NTA- SF CK CS CP GI D G IK-- CT (D)	1362	-cbEGF
33	cbEGF19	cbEGF-	1363	DLDE CS NG THM- CS QH AD - CK NTM- GSY RC CK EG TG D GFT -- CT (D)	1403	-cbEGF
34	cbEGF20	cbEGF-	1404	DLDE CS EN LNL- CG NQ-- CK NAO- GT RC CE DM G V PS AD GKA - CE (D)	1445	-cbEGF
35	cbEFG21	cbEGF-	1446	DIDE CS L PN II -- CV FGT-- CH NLP- LFR CE CE IG EL DR S GG N - CT (D)	1486	-cbEGF
36	cbEGF22	cbEGF-	1487	DVNE CL DP TT-- CIS GN-- CV NTP- SYI CD CP DF EL NP TR VG- CV (D)	1527	-TB
39	cbEGF23	TB-	1606	DIDE CQ EL PL L - CQ CGK-- CINT F- GSG Q CR CP TY L N ED TRV - CD (D)	1647	-cbEGF
40	cbEGF24	cbEGF-	1648	DVNE CE TPGI -- CG PT-- CY NTV- NYT CI CP PD Y M QV NN GN NC M (D)	1688	-TB
43	cbEGF25	TB-	1766	DIDE CR EIP GV- CEN SV-- CIN MV- GSF RC CP V G F Y ND KLLV - CE (D)	1807	-cbEGF
44	cbEGF26	cbEGF-	1808	DIDE CQ NG PV-- Q R NAE - CINT A- GSY RC CK PG FR FT ST G Q -- CN (D)	1848	-cbEGF
45	cbEGF27	cbEGF-	1849	DRNE CQ EIP NI- CL SGT-- CID TV- SF Y CL CH TK NT DD Q T M- CL (D)	1890	-cbEGF
46	cbEGF28	cbEGF-	1891	DINE CE RDA -- CG NGT-- CR NTI- SF NC RC NH GL SH NND -- CI (D)	1929	-cbEGF
47	cbEGF29	cbEGF-	1930	DVDE CA SG NG NL CR NG Q-- CINT V- SF Q C Q NE GE VA PD GRT - CV (D)	1972	-cbEGF
48	cbEGF30	cbEGF-	1973	DINE CL LE PR K - CA PT-- CON LD- GSY RC IC PP GL SL QNEK -- CE (D)	2012	-cbEGF
49	cbEGF31	cbEGF-	2013	DIDE CV EE PE I - CA LGT-- CS NTE- SF RC CL PE GF SL SS SG RR - CQ (D)	2054	-TB
52	cbEGF32	TB-	2127	DMDE CK EP DV-- CK HGQ-- CINT D- GSY RC CE PF YL AG NE -- CV (D)	2165	-cbEGF
53	cbEGF33	cbEGF-	2166	DTDE CS VGN P-- CG NGT-- CK NVI- GF E CT CE EG EP GP MM T -- CE (D)	2205	-cbEGF
54	cbEGF34	cbEGF-	2206	DINE CA QN PL L - CA FR-- CV NTY- GSY EC CK PV G V LR ED RRM - CK (D)	2246	-cbEGF
55	cbEGF35	cbEGF-	2247	DEDE CE EG KHD- CT E K Q ME CK NLI - TY M CI CG PG Q RR PD GEG - CV (D)	2290	-cbEGF
56	cbEGF36	cbEGF-	2291	DENE CQ TK PG I - CEN GR-- CL NTR- SYT CE CN D GFT AS PN Q DE - CL (D)	2332	-TB
58	cbEGF37	TB-	2402	DIDE CK VI H DV - CR NGE-- CV NDR- GSY H CI CK TG TP D I TG TS- CV (D)	2443	-cbEGF
59	cbEGF38	cbEGF-	2444	DLNE CN QA PK P - CN FII-- CK NTE- SY Q CS CP KG IL Q ED GRS - CK (D)	2484	-cbEGF
60	cbEGF39	cbEGF-	2485	DLDE CA TK QH N - CQ FL-- CV NTI- GF T CK CP PG T QH HTS-- CI (D)	2523	-cbEGF
61	cbEGF40	cbEGF-	2524	DNNE CT SD IN L - CG SK GI - CON TP- SF T CE Q RG SL D Q T GS S - CE (D)	2566	-cbEGF
62	cbEGF41	cbEGF-	2567	DVDE CE GN HR-- CQ HG-- CON TI- GSY RC CP Q GL Q H Y Q W Q - CV (D)	2606	-cbEGF
63	cbEGF42	cbEGF-	2607	DENE CL SA HI-- CG AS-- CH NLT- GSY K CM CP AG Q VE Q F SG G - CQ	2647	-cbEGF
63-64	cbEGF43	cbEGF-	2648	DINE CG SA Q AP - CSY G-- CS NTE- GV L CG CP PGY FR IG Q H -- CV	2687	

B

Exon (s)	Domain	5'AA	amino acid sequence	3'AA
9-10	TB1	329	ID VR PG Y C Y T A L T NG R ----- CS N QL P QS I T K MQ -- CC D AG R C W SP G V T V A PE M CP I R A T E DF N KL C SV P M V I P R G PE Y	401
16-17	TB2	654	D TH M R S T C Y G G Y K R G Q----- CI K PL F G A V T K SE-- CC CA ST - E Y A F G -- EP C Q P CP A Q N S A E Y Q A L C SS G P G M T S A G S	722
24	TB3	951	L D I R I L E T C F L R Y E D E E----- CT L I A GR H R M D A-- CC CS V -- G A A W G -- EE CE C P M R NT P E Y E L CP R G F A T K I IN G K P FP F K	1027
37-38	TB4	1528	-D T R S G N C Y L D I R P R G D NG D T A C S N E I G V G S K A S -- CC CS L -- G K A W G -- TP CE L C PS V NT S EY K IL C PG G E G FR PN P I T V I L E	1605
41-42	TB5	1689	-D M R R S L C Y R N Y A D N Q T ---- CD G EL L FN M T K K M-- CC CS Y N I G R A WN -- KP CE Q C PI P ST D E F A T L C GS Q R P G F Y I D Y T G L P V	1765
50-51	TB6	2054	Q D L R M S Y C Y A K F E G G K ----- CS SP K S R N S H Q E-- CC CA LK - G E G W G -- DP CE L C P T E P D E A F R I CP Y G S G I V G P D S A V	2126
57	TB7	2333	-D N R E G Y C F T E V L Q N M----- CQ I G S S N R N P V T K SE CC CD G-- G R G W G -- PH CE L C P F Q T V A F K KL C PH G R G R M T NG A	2401

Fig. 1 *FBN1* domain alignment and critical residues. Amino acid residues known to be critical for function are either bolded or highlighted. **a** EGF/cbEGF-like domain alignment. Cysteines (C) bolded in red; calcium-binding (D/N-x-D/N -E/Q-Xm-D/N-Xn-Y/F) highlighted in yellow [note: an asparagine (N) to serine(S) at the second D/N position maybe tolerated; see text]; Gly (G)/Ala (A) at residue 3 between obligatory Cys2 and Cys3 in a subtype of cbEGF-like domains highlighted in blue; cbEGF interdomain packing (Gly between Cys3 and Cys4 and aromatic residue Y/F/H/W between Cys5

and Cys6) highlighted in green (note: the Gly interacts with the Y/F/H/W in the upstream cbEGF domain); “EGF core” which comprises the major beta-sheet of the EGF fold is underlined (note: only highlighted/bolded amino acids in the EGF core are currently known to be critical). **b** TB domain alignment. 5' domain = 5'-adjacent domain; 3' domain = 3'-adjacent domain; AA = amino acid; cbEGF = calcium-binding EGF-like domain; EGF = noncalcium-binding EGF-like domain; TB = transforming growth factor beta domain; HYB = hybrid domain

domains [20, 29, 31–33]; (5) created or altered a Cys residue in a TB domain (Fig. 1b) [18, 34]; (6) created or altered a Cys residue in an EGF domain [35, 36] (Fig. 1a); (7) created or altered a Cys residue in the second hybrid (hybrid2) domain [37] [note: we did not consider Cys in the hybrid1 domain to meet criteria for PMI because this domain has an extra Cys of unclear function and studies

have shown that hybrid1 domain-deleted mice have a normal lifespan [38]]; (8) impacted a critical residue in the fibrillin-unique N terminus (FUN) site [39–41]; or (9) impacted a residue in the Arg-Gly-Asp (RGD) integrin binding site in the TB domain [42, 43]. However, ClinVar did not contain entries for variants in the RGD region, so knowledge of these critical residues could not be assessed.

Table 1 *FBN1* critical missense residue type, frequency, and ClinVar clinical classification

Domain	Critical residue type	Type and number of variant classifications						
		L(B) (n)	L(B) (%)	VUS (n)	VUS (%)	L(P) (n)	L(P) (%)	Total (n)
cbEGF-like	Cys	0	0	23	10.0	208	90.0	231
	Calcium-binding (D/N-x-D/N-E/Q-Xm-D/N-Xn-Y/F)	3	2.7	41	37.2	66	60.0	110
	C2–C3 Gly/Ala	0	0	12	57.1	9	42.9	21
	Interdomain packing	0	0	7	71.4	2	28.6	9
TB	Cys	0	0	3	8.1	34	91.9	37
	RGD integrin binding	0	0	0	0	0	0	0
Hybrid2	Cys	0	0	0	0	9	100	9
EGF	Cys	0	0	1	3.8	25	96.2	26
Fibrillin	FUN	0	0	0	0	8	100	8
	Total	3	0.7	87	19.3	361	80.0	451

cbEGF-like calcium binding EGF-like domain, *Cys* cysteine residue, *C2–C3 Gly/Ala* Glycine/Alanine at position 3 between Cys2 and Cys3 in subset of cbEGF-like domains, *TB* TGF-beta binding protein domain, *Hybrid2* second hybrid domain, *FUN* fibrillin-unique N terminus

FBN1 variant reanalysis and classification

FBN1 variant nomenclature was based on RefSeq NM_000138.4. Variants were classified in ClinVar based on a five-level classification system as per the ACMG/AMP 2015 variant classification guidelines: P, LP, VUS, likely benign (LB), or benign (B) (Supplemental Table 1) [6]. A single variant’s classification submitted by multiple submitters was considered discrepant if the classifications disagreed between (L)B versus VUS, (L)B versus (L)P, or VUS versus (L)P.

Variant classifications were reanalyzed by our group for pathogenicity based on *FBN1*-specific modified 2015 ACMG/AMP/CAP criteria [6]. These modifications included: application of PM1 for variants that affect any of the nine types of residues discussed above and an automatic minimum classification of (L)P for any variant impacting a Cys residue in a cbEGF-like or TB domain. In addition, we applied PP2 for those variants that met additional pro-pathogenicity criteria and for which we were unable to apply BS1, BP4, or other pro-benign criteria. The PP2 criterion is applicable for *FBN1* since it has an ExAC missense constraint z-score of 5.33, which is above the ClinGen recommended z-score of 3.09 for application of PP2 [44]. In addition, the minor allele frequency (MAF) was analyzed using gnomAD population data (<https://gnomad.broadinstitute.org/>) [44]. A cutoff for applying the BS1 criteria (allele frequency greater than expected for the disorder) was set conservatively at 0.01%, based on using the 1/5000 as the population frequency of MFS (commonly stated to range from 1/5000 to 1/10,000) [45–48], and also considering that no single *FBN1* variant is responsible for >50% of all of MFS. As an added measure of

stringency, we only applied BS1 if a variant was present at a frequency of >0.01% and occurred in at least three alleles in the specific gnomAD population it was detected in. We used the HGVS nomenclature Version 15.11 for functional and clinical classification to describe variant effects (<https://vamos.momen.hgvs.org/>). Classified variants (*n* = 674) were uploaded to the *FBN1*-specific Global Variome shared Leiden Open Variation Database (LOVD) (<https://databases.lovd.nl/shared/genes/variants> 0000459080_0000459753).

Results

There were 1426 *FBN1* variants with 1999 total variant classifications from 18 submitters (mostly clinical laboratories) in ClinVar. There were 11 submitters with 45 or more *FBN1* variant submissions. There were 266 variants that had multiple (>1) submissions, and 80 of these (30.1%) had discrepant variant classifications.

Overview of missense variants

Since *FBN1* missense variants are generally more challenging to classify compared to nonsense, splicing, or frame-shift variants, we focused our analysis on missense variants. Of the 1426 *FBN1* variants listed in ClinVar, 674 (47.3%) were unique missense variants with a total of 973 variant calls. Of the 674 missense variants, 140 of them (20.8%) had more than one submitter and 43/140 (30.7%) had discrepant classifications.

Variants that occur at critical residues or otherwise impact fibrillin-1 function and are not present in population

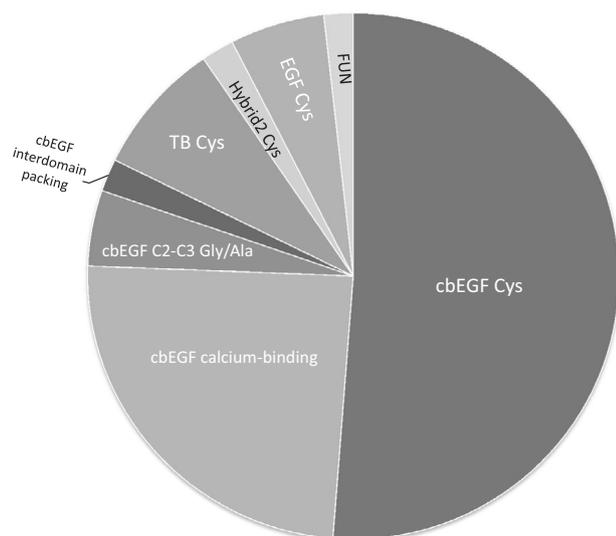


Fig. 2 *FBNI* critical region missense variant frequency by location. cbEGF = calcium binding EGF-like domain; Cys = cysteine residue; cbEGF C2–C3 Gly/Ala = Gly/Ala at residue 3 between obligatory Cys2 and Cys3 in a subtype of cbEGF-like domains; TB = TGF-beta-binding protein domain; Hybrid2 = second hybrid domain; FUN = fibrillin unique N-terminal domain

databases would be expected to be functionally classified as, at a minimum, probably affects function (PAF), or clinically classified as LP due to their functional significance [13]. Of the 674 missense variants in ClinVar, 361 (53.6%) occurred at a known critical residue or created a Cys in relevant domains of fibrillin-1 (see “categorization of *FBNI* critical residue variants” in methods section), for a total of 451 calls (Table 1; Fig. 2; supplemental table). Of these 451 calls, 361 (80.0%) gave a (L)P clinical classification; however (L)P classification frequency was dependent on which type of critical residue the variant impacted (Fig. 3a). For the well-known cbEGF-like domain Cys residues, 90.0% (208/231) of the variant classifications were correctly clinically classified as (L)P. In fact, overall, missense variants that impacted a Cys residue were highly recognized as (L)P variants (283/310 calls or 91.3%) (Fig. 3a). On the other hand, for non-Cys critical residues, only 55.3% (78/141) of clinical classifications were (L)P. Variants impacting the critical glycine/alanine and interdomain packing residues (not taking into account aromatic residue substitutions; see below) in the cbEGF-like domain were the least likely critical residue variants to be classified as (L)P (39.3% or 11/28 variant calls). Overall, we were able to resolve discrepancies in all nine variants with conflicting multi-submitter classifications and reclassify 56 of the remaining 352 (15.9%) of these functionally important variants, as described in more detail below. Overall this led to a revision of 17.2% (77/451) of variant classifications for functionally important variants.

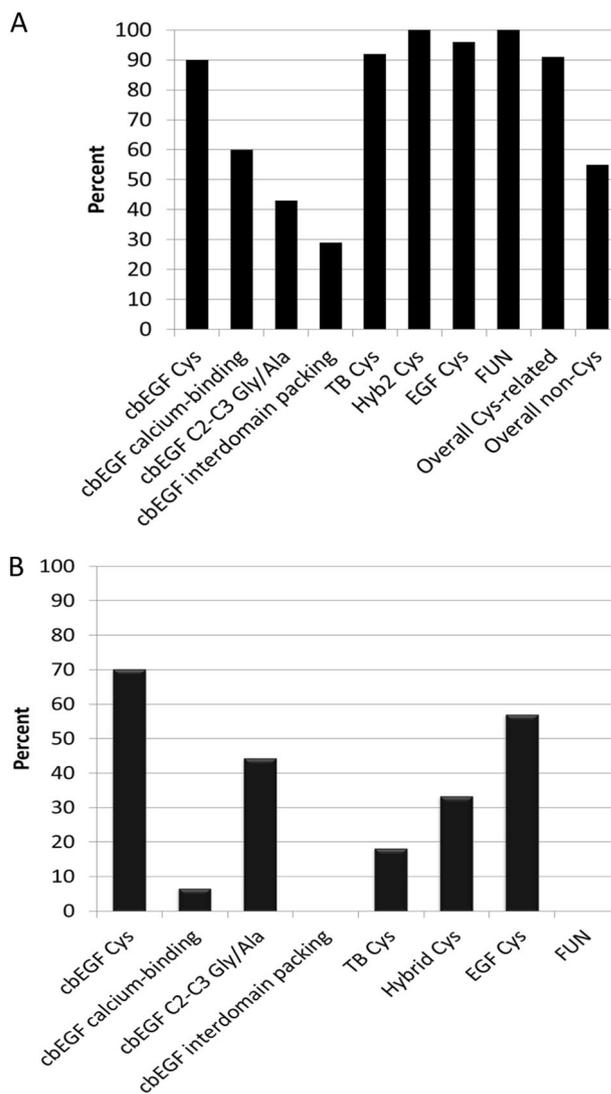


Fig. 3 **a** Percent of critical region missense variants classified as (L)P. **b** Percent of provided evidence stating critical residue as factor supporting (L)P classification. cbEGF = calcium binding EGF-like domain; Cys = cysteine residue; cbEGF C2–C3 Gly/Ala = Gly/Ala at residue 3 between obligatory Cys2 and Cys3 in a subtype of cbEGF-like domains; TB = TGF-beta binding protein domain; Hyb2 = Hybrid2 domain; FUN = fibrillin unique N-terminal domain

Missense variants impacting Cys residues in the cbEGF-like domain

As mentioned above, 90% (208/231) of cbEGF-like Cys clinical classifications were (L)P (Table 1; Fig. 3a). Of these, 130 (62.5%) provided evidence regarding variant classification. Justification for 74.6% of these calls (97/130) included verbiage related to the importance of cbEGF-like domain Cys (Fig. 3b). For the remaining 25.4% of calls, the majority (25/33) was from 2014 or earlier.

There were 20 cbEGF-like Cys missense variants that were classified 23 times as VUS by five different

submitters. One variant with a dual submission had a VUS and an LP call, and the rest had two VUS calls. Most (18/20) of the VUS variants were submitted to ClinVar between 2013 and 2017, with two variants submitted in 2009. All variant calls that were not classified as (L)P were reclassified by us to (L)P or PAF (supplemental table).

Missense variants impacting calcium-binding residues in the cbEGF-like domain

Variants impacting the D/N-x-D/N-E/Q-Xm-D/N-Xn-Y/F calcium-binding residues in the cbEGF-like domains comprised the next largest grouping of critical residue *FBN1* missense variants in ClinVar with a total of 110 variant calls for 80 variants. As described below, some of the missense changes in this region involved the substitution of a known or potentially tolerated amino acid and we did not consider these variants to be L(P). This situation impacted 9 variants with 13 clinical classifications [11 VUS and 2 L(B)]. Of the remaining 71 variants with 97 variant calls, 68.0% (66/97) were called (L)P by 9 different submitters (Fig. 3a). Of these, 39 (59.1%) classifications had meaningful submitter evidence provided. However, only 4/39 (10.2%) submissions highlighted the importance of the calcium-binding residues (Fig. 3b). Twenty-two (30.6%) of the (L)P variants were at the first D/N of the calcium-binding consensus region. However, not all submitters recognized the relevance of this first D/N, since a variant at this position was also clinically classified as VUS eight other times. In fact one submitter initially classified a “first D/N” variant as (L)P and then later classified a different “first D/N” variant as VUS.

Not including variants that we considered to be VUS (as described above and below), the majority (25/31; 80.1%) of LB or VUS clinical classifications at calcium-binding residues was submitted between 2015 and 2018, with one submission in 2008 and five submissions from 2011 to 2014. For the VUS and LB submissions, 18/25 submitters provided variant classification evidence, with a variety of supportive evidence statements, including in some cases that the variant was not a Cys-based variant in a cbEGF-like domain.

There were only four discrepant multisubmitter clinical classifications (one occurrence of LB versus VUS and three occurrences of VUS versus LP) of the calcium-binding residues. For the variant that was classified as LB and VUS, neither submission included meaningful evidence to support classification. For the VUS versus LP discrepant variant calls ($n = 13$), 8/13 calls had associated evidence in their submission, and all of these were from more recent years (2016–2018), but only one mentioned the calcium-binding critical residue.

Nine variants with 13 calls were kept at or revised to VUS (or functional effect unknown). Three of these variants

were substitutions of a D for N or vice versa at a D/N position: p.Asp1197Asn, p.Asp1240Asn, and p.Asn1907Asp. The other six variants were Asn > Ser changes at the second D/N in the D/N-x-D/N-E/Q-Xm-D/N-Xn-Y/F consensus sequence: p.Asn615Ser, p.Asn1030Ser, p.Asn1282Ser, p.Asn1489Ser, p.Asn2526Ser, and p.Asn2650Ser (Table 2). During the course of this study, we observed that some of these Asn > Ser variants had relatively high minor allele frequencies in gnomAD. For example, p.Asn1030Ser is observed in 0.22% (68/30616 alleles, including two homozygotes) of the South Asian population. We examined all asparagine residues in the second D/N position across all cbEGF-like domains to ascertain the presence and frequency of serine substitutions in the gnomAD population. Of the 43 cbEGF-like domains, 18 (41.9%) had an asparagine (versus aspartate) at that position, and 9/18 (50%) had an Asn > Ser change described in ClinVar and/or present in gnomAD (Table 2). Given that seven of the nine p.Asn > Ser variants are present in gnomAD and since Asn and Ser are both similar polar uncharged amino acids with a Grantham distance of 46 (range 0–215), it is possible that a serine at this position is tolerated (i.e., D/N-x-D/N/S-E/Q-Xm-D/N-Xn-Y/F) and should be categorized as VUS or of uncertain functional effect. However, functional studies are needed to fully assess the impact of a serine at this position. Overall, of the 44 calcium-binding variant calls that were (L)B or VUS in ClinVar, 75% (33/44) were reclassified to (L)P.

Missense variants impacting Cys residues in the TB domain

The third largest group of critical residue missense variants impacted Cys residues in the 8-Cys TB domain. There were 34/37 (91.9%) L(P) variant clinical classifications for 30 TB domain Cys variants (Fig. 3a) and three VUS classifications. Of the 34 (L)P classifications, 22 provided evidence, but only four of these acknowledged the importance of the 8-Cys motif of the TB domain in their classification evidence (Fig. 3b). For the VUS classifications, one of the submitters provided evidence supporting their VUS classification (submitted to ClinVar in 2017), but did not acknowledge the importance of the 8-Cys motif of the TB domain.

Critical glycine (Gly)/alanine (Ala) between Cys2–Cys3 in a subset of cbEGF-like domains

There were 21 clinical classifications of 19 different variants that impacted the Gly/Ala in the third position between Cys2 and Cys3 in a subset of cbEGF-like domains. Overall only 9/21 (42.9%) classifications were (L)P (Fig. 3a). The majority (17/19) of the variants had a single submission, whereas two variants had two submissions [VUS and (L)P].

Table 2 *FBNI* asparagine (Asn) to serine (Ser) changes at the second D/N of the cbEGF-like calcium-binding region (D/N-x-D/N-E/Q-Xm-D/N-Xn-Y/F) observed in Clinvar and/or gnomAD

cDNA change	Amino acid change	ClinVar classification(s)	GnomAD population frequencies		
			population	Allele count	MAF (%)
c.1844A > G	p.Asn615Ser	VUS (2)	EAS	3/19946	0.015
			OTH	1/7226	0.014
c.3089A > G	p.Asn1030Ser	LB, VUS	SAS	68/30548 (two homozygotes)	0.22
			NFE	2/129150	0.00015
			OTH	2/7218	0.028
			SAS	1/34575	0.013
c.3845A > G	p.Asn1282Ser	B	NFE	7/113610	0.0062
			AMR	1/34576	0.0029
			NFE	1/113428	0.00088
c.3971A > G	p.Asn1324Ser	none	NFE	1/113428	0.00088
c.4466A > G	p.Asn1489Ser	VUS	not present		
c.6623A > G	p.Asn2208Ser	none	NFE	8/128710	0.0062
c.6878A > G	p.Asn2293Ser	none	EAS	1/18364	0.0054
c.7577A > G	p.Asn2526Ser	VUS (3)	not present		
c.7949A > G	p.Asn2650Ser	VUS	NFE	3/113634	0.0026

MAF minor allele frequency, EAS East Asian, SAS South Asian, NFE European (non-Finnish), OTH (Other), AMR (Latino)

Of the nine (L)P submissions at this position, eight (88.9%) contained evidence, but only four of these mentioned the importance of this residue (Fig. 3b). Other reasons provided for the (L)P call included that the variant was previously reported in MFS patients, absence from databases, in silico predictions, conserved residue, nonconservative change, and the presence of nearby pathogenic missense variants (including Cys variants). For VUS submitters that provided evidence, evidence supporting the call included that the variant had been reported before in association with MFS, was not observed in databases, was a nonconservative amino acid substitution, occurred at a highly conserved residue, and had deleterious in silico predictions, but did not affect a Cys residue.

EGF-like domain cysteine missense variants

There were 26 variant calls for 20 Cys missense variants in the EGF-like domain, and all but one of them were clinically classified as (L)P (Fig. 3a). The single VUS classification was a submission classified in 2009 and did not contain evidence. For the 25 (L)P calls, 14 contained evidence for classification and 57.1% (8/14) mentioned that EGF-like domain Cys variants factored into the classification (Fig. 3b).

Other critical residues

There were seven interdomain packing residues in a cbEGF-like domain with nine calls, and 2/9 clinical classifications were (L)P. Two of the interdomain packing residues

classified as VUS (p.Tyr1839Phe and p.Tyr2236Phe) involved the change from one aromatic residue (tyrosine) to another (phenylalanine), and thus we did not change their classification from VUS. One variant had multiple calls [two VUS and one (L)P]. None of the (L)P calls provided evidence stating a critical residue as a factor supporting the classification (Fig. 3b).

There were nine variants with nine (L)P variant calls that impacted Cys residues in the second hybrid (hybrid2) domain. Of the (L)P calls, four provided evidence, and two described the importance of Cys variants in general, with one mentioning the importance of Cys variants in hybrid domains.

Finally, there were eight variants with eight calls that occurred in the FUN domain. All eight of these calls were (L)P, but only one of the calls had evidence and this evidence did not state the importance of the FUN domain (Fig. 3a, b). Notably, seven of the eight calls involved disruption of a Cys, which may have influenced the (L)P call.

Missense variants at residues not known to be critical

There were 313 other missense variants with 522 clinical classifications that did not occur at a previously described critical residue (supplemental table). Of these, 224 were classified by a single submitter. For the 89 variants with multiple submitters, there were 298 variant calls. Conflicting classifications occurred with 34/89 (38.2%) variants with multiple submitters.

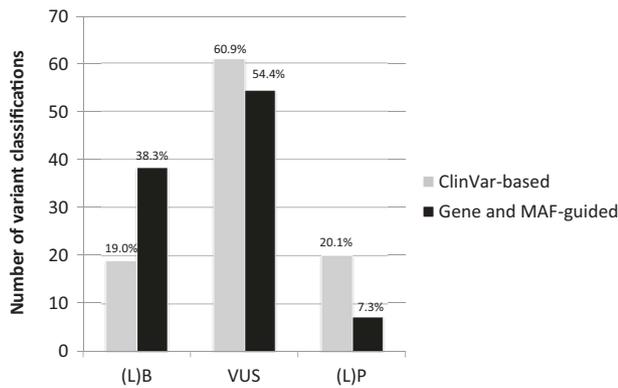


Fig. 4 Reclassification of ClinVar-classified noncritical *FBNI* missense variants. The original classification in ClinVar is represented by the “ClinVar-based” bars. The reclassifications are represented by the “Gene and MAF-guided” bars

For these noncritical residue missense variants, 105/522 (20.1%) were classified as (L)P, 318/522 (60.9%) were VUS, and 99/522 (19.0%) were (L)B. Applying review of the evidence, when available, and our refined MAF cutoff, 30.5% (97/318) of the VUS calls were revised to (L)B and 63.8% (67/105) of the L(P) calls were revised to (L)B (4/105) or VUS (63/105), respectively (Fig. 4). Overall, we were able to resolve variant discrepancies in all 34 variants with conflicting classifications and reclassify 94 (33.7%) of the remaining 279 missense variants not known to be critical or impact fibrillin-1 function.

Discussion

Through this analysis, we demonstrate the importance of applying gene-based knowledge to variant classification. Although our scope was different, we observed similarities to that of Muiño-Mosquera et al. [9], in that the application of gene-based guidelines changed the interpretation of previously classified *FBNI* variants. Unlike Muiño-Mosquera et al., which looked at *FBNI* variants of all types observed in their own laboratory, our study was focused specifically on *FBNI* missense variants submitted to ClinVar, giving a more global perspective of the impact of gene-based guidelines to *FBNI* variant classification and some of the more widespread inadequacies of *FBNI* variant assessment. We further enhanced the Muiño-Mosquera study by including additional categories of *FBNI* critical residues, incorporation of the PP2 criterion when appropriate, and a refinement of the MAF cut point to use for BS1. Furthermore, while their method resulted in an increased number of VUSs (59.7% of those outside of well-established functional sites), our approach actually led to an overall decrease in the number of VUS/effect unknown calls outside of well-established functional sites. Overall, we

were able to resolve all 43 missense variants that had ClinVar discrepant multisubmitter classifications and revise variant calls in 150 (23.8%) of the remaining 631 *FBNI* missense variants. Taken together, we revised 24.8% (241/973) of all *FBNI* missense variant classifications in ClinVar.

While the majority (80.0%) of ClinVar classifications for critical residue missense variants were (L)P, there were specific types of critical residue variants that were more often misclassified. Because of the well-known association between *FBNI* Cys variants and MFS, it is not surprising that most submitters correctly classified Cys variants as (L)P. However, when considering evidence provided supporting (L)P calls for Cys variants, there seemed to be a lowered awareness regarding the importance of Cys variants in non-cbEGF-like domains. Much of the submitter evidence supporting (L)P classification had to do with other valid reasons supporting pathogenicity, rather than the functional relevance of the residue. As noted, non-Cys critical residue variants had a lower likelihood of being classified as (L)P compared to Cys variants. Also, there was less relevant evidence highlighting the pathogenicity of these functional regions. In fact, some submitters classified these variants as VUS in part because they did not occur at a Cys. Thus, there appears to be a decreased awareness regarding the functional importance of some calcium-binding, glycine/alanine, and interdomain packing residues, as well as critical variants in the FUN domain. A caveat to this analysis is that this data is somewhat limited in its analytical power since only 59% of submitters included substantive evidence with their variant submissions, thus it is unclear what criteria were used to classify the other 41% of variants.

In addition to the Cys versus non-Cys variant classification challenges, we observed several major themes that occurred with misclassification of critical *FBNI* residues. First, misclassification occurred more frequently in variants that were last evaluated several years ago (e.g., pre-2015). For these variants, reevaluation is likely needed by the submitter. Second, there were not only the expected inter-laboratory inconsistencies in variant classification and recognition of critical region variants, but also intralaboratory inconsistencies. The age of the classifications sometimes appeared to play a role in intralaboratory variability of classification or evidence, but was not always a factor.

There were several noncritical residue variants that had conflicting classifications and/or were deemed to be misclassified largely because of the MAF of the variant. In fact, based largely on the utilization of an appropriate MAF cutoff, we were able to greatly reduce both the number of (L)P calls (from 105 to 38; 63.8% decrease of noncritical missense variants), as well as decrease the number of reportable noncritical variants (from 423 to 322; 23.9% decrease) that might otherwise have cluttered a report. It has

been pointed out by others that there is a critical need to provide guidance surrounding MAF thresholds [7, 49]. Our results demonstrate this same phenomenon, as well as the need to reevaluate variants when updated MAF data and thresholds are made available. It is also worthwhile to point out that the evaluation of updated MAF data was helpful in determining that a serine substitution at the second D/N of the calcium-binding region in cbEGF-like domains may be tolerated.

Although the main focus of this study was to evaluate classification (and classification evidence) of *FBNI* missense variants, we noted that many ClinVar entries did not contain substantive evidence to support their classification. In fact, only 59.3% (267/450) of variant submissions from reportable variants at critical residues contained evidence. This percentage is similar to what was reported by Groth et al. [3], and was also a criticism in their report. We agree with their assessment that “variants without supporting data are of very little value”. Indeed, it was impossible for us to assess the validity of a submitter’s variant classification without any context with regards to criteria utilized for classification. When supporting data such as patient phenotype, case counts, literature citations, gene-specific details (e.g., critical region), and ACMG/AMP criteria used to classify the variant are included in a ClinVar entry, this adds value to clinicians and laboratorians alike for the process of variant curation and clinical decision-making. However, this information must also be used with discretion, since the NIH does not independently verify information submitted to ClinVar.

A recent study of a subset of *FBNI* variants in four variant databases (HGMD, UMD-FBN1, ClinVar, and UniProt) concluded that current variant databases could not be relied upon for a genetic diagnosis of MFS due to incorrectly interpreted conclusions on variants [3]. As discussed earlier, Muiño-Mosquera et al. [9] also has published on the importance of applying disease- and gene-specific guidelines for a more uniform interpretation of *FBNI* variants. Pepin et al. [2] likewise highlighted that, in the context of variant curation and interpretation, there should be collaboration between clinical laboratories, disease-, and gene-specific experts. With expanding genetic analyses becoming increasingly available and offered by multiple laboratories, our results demonstrate the importance of utilizing gene-based knowledge and applying appropriate ACMG/AMP criteria for variant classification, as well as updating classifications as more evidence becomes available, not just for *FBNI*, but other genes as well. Laboratories that perform genetic testing, including exome and genome sequencing, that includes *FBNI* analysis can use the supplemental table and Fig. 1 published here to aid in classification of previously described and novel *FBNI* variants, respectively.

While having phenotypic data available while performing variant evaluations is ideal, the reality is that this is oftentimes not the case for many clinical laboratories. Thus, our approach demonstrates that *FBNI* variant classification can be refined without always having the context of phenotype available, a situation that most clinical laboratories can relate to. Our results further emphasize, as others have previously shown, that while many variants are correctly classified in ClinVar, there is room for improvement. Collecting gene-based evidence to apply to variant classification is not a trivial task, and clinical laboratories may lack the resources to compile such evidence for each gene they encounter. Thus, articles such as this may serve as a useful resource for clinical laboratories performing *FBNI* testing as well as provide a perspective to the clinical community regarding limitations in inter- (and in some cases, intra-) laboratory variant classification.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Yang S, Lincoln SE, Kobayashi Y, Nykamp K, Nussbaum RL, Topper S. Sources of discordance among germ-line variant classifications in ClinVar. *Genet Med.* 2017;19:1118–26.
2. Pepin MG, Murray ML, Bailey S, Leistriz-Kessler D, Schwarze U, Byers PH. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. *Genet Med.* 2016;18:20–4.
3. Groth KA, Gaustadnes M, Thorsen K, Ostergaard JR, Jensen UB, Gravholt CH, et al. Difficulties in diagnosing Marfan syndrome using current FBN1 databases. *Genet Med.* 2016;18:98–102.
4. Harrison SM, Dolinsky JS, Knight Johnson AE, Pesaran T, Azzariti DR, Bale S, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med.* 2017;19:1096–104.
5. Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, et al. Performance of ACMG-AMP variant-interpretation guidelines among nine laboratories in the clinical sequencing exploratory research consortium. *Am J Hum Genet.* 2016;98:1067–76.
6. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–24.
7. Kelly MA, Caleshu C, Morales A, Buchan J, Wolf Z, Harrison SM, et al. Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: recommendations by ClinGen’s Inherited Cardiomyopathy Expert Panel. *Genet Med.* 2018;20:351–9.
8. Gelb BD, Cave H, Dillon MW, Gripp KW, Lee JA, Mason-Suares H, et al. ClinGen’s RASopathy Expert Panel consensus methods for variant interpretation. *Genet Med.* 2018;20:1334–45.

9. Muino-Mosquera L, Steijns F, Audenaert T, Meerschaut I, De Paepe A, Steyaert W, et al. Tailoring the American College of Medical Genetics and Genomics and the Association of Molecular Pathology Guidelines for the Interpretation of Sequenced Variants in the *FBN1* Gene for Marfan Syndrome: Proposal for a Disease- and Gene-Specific Guideline. *Circ Genom Precis Med*. 2018;11:e002039.
10. Judge DP, Dietz HC. Marfan's syndrome. *Lancet*. 2005;366:1965–76.
11. Faivre L, Collod-Beroud G, Loeys BL, Child A, Binquet C, Gautier E, et al. Effect of mutation type and location on clinical outcome in 1,013 probands with Marfan syndrome or related phenotypes and *FBN1* mutations: an international study. *Am J Hum Genet*. 2007;81:454–66.
12. Baudhuin LM, Kotzer KE, Lagerstedt SA. Decreased frequency of *FBN1* missense variants in Ghent criteria-positive Marfan syndrome and characterization of novel *FBN1* variants. *J Hum Genet*. 2015;60:241–52.
13. Loeys BL, Dietz HC, Braverman AC, Callewaert BL, De Backer J, Devereux RB, et al. The revised Ghent nosology for the Marfan syndrome. *J Med Genet*. 2010;47:476–85.
14. Pereira L, D'Alessio M, Ramirez F, Lynch JR, Sykes B, Pangilinan T, et al. Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum Mol Genet*. 1993;2:961–8.
15. Handford PA. Fibrillin-1, a calcium binding protein of extracellular matrix. *Biochim Biophys Acta*. 2000;1498:84–90.
16. Schrijver I, Liu W, Brenn T, Furthmayr H, Francke U. Cysteine substitutions in epidermal growth factor-like domains of fibrillin-1: distinct effects on biochemical and clinical phenotypes. *Am J Hum Genet*. 1999;65:1007–20.
17. Vollbrandt T, Tiedemann K, El-Hallous E, Lin G, Brinckmann J, John H, et al. Consequences of cysteine mutations in calcium-binding epidermal growth factor modules of fibrillin-1. *J Biol Chem*. 2004;279:32924–31.
18. Zeyer KA, Reinhardt DP. Engineered mutations in fibrillin-1 leading to Marfan syndrome act at the protein, cellular and organismal levels. *Mutat Res Rev Mutat Res*. 2015;765:7–18.
19. Kiely CM, Shuttleworth CA. The role of calcium in the organization of fibrillin microfibrils. *FEBS Lett*. 1993;336:323–6.
20. Downing AK, Knott V, Werner JM, Cardy CM, Campbell ID, Handford PA. Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. *Cell*. 1996;85:597–605.
21. Reinhardt DP, Ono RN, Notbohm H, Muller PK, Bachinger HP, Sakai LY. Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis. A potential disease-causing mechanism in Marfan syndrome. *J Biol Chem*. 2000;275:12339–45.
22. Wu YS, Bevilacqua VL, Berg JM. Fibrillin domain folding and calcium binding: significance to Marfan syndrome. *Chem Biol*. 1995;2:91–7.
23. Handford PA, Mayhew M, Baron M, Winship PR, Campbell ID, Brownlee GG. Key residues involved in calcium-binding motifs in EGF-like domains. *Nature*. 1991;351:164–7.
24. Yuan X, Werner JM, Lack J, Knott V, Handford PA, Campbell ID, et al. Effects of the N2144S mutation on backbone dynamics of a TB-cbEGF domain pair from human fibrillin-1. *J Mol Biol*. 2002;316:113–25.
25. Glanville RW, Qian RQ, McClure DW, Maslen CL. Calcium binding, hydroxylation, and glycosylation of the precursor epidermal growth factor-like domains of fibrillin-1, the Marfan gene protein. *J Biol Chem*. 1994;269:26630–4.
26. Khau Van Kien P, Baux D, Pallares-Ruiz N, Baudoin C, Plancke A, Chassaing N, et al. Missense mutations of conserved glycine residues in fibrillin-1 highlight a potential subtype of cb-EGF-like domains. *Hum Mutat*. 2010;31:E1021–42.
27. Whiteman P, Smallridge RS, Knott V, Cordle JJ, Downing AK, Handford PA. A G1127S change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. *J Biol Chem*. 2001;276:17156–62.
28. Francke U, Berg MA, Tynan K, Brenn T, Liu W, Aoyama T, et al. A Gly1127Ser mutation in an EGF-like domain of the fibrillin-1 gene is a risk factor for ascending aortic aneurysm and dissection. *Am J Hum Genet*. 1995;56:1287–96.
29. Smallridge RS, Whiteman P, Werner JM, Campbell ID, Handford PA, Downing AK. Solution structure and dynamics of a calcium binding epidermal growth factor-like domain pair from the neonatal region of human fibrillin-1. *J Biol Chem*. 2003;278:12199–206.
30. Whiteman P, Downing AK, Smallridge R, Winship PR, Handford PA. A Gly -->Ser change causes defective folding in vitro of calcium-binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J Biol Chem*. 1998;273:7807–13.
31. Jensen SA, Corbett AR, Knott V, Redfield C, Handford PA. Ca²⁺-dependent interface formation in fibrillin-1. *J Biol Chem*. 2005;280:14076–84.
32. Jensen SA, Robertson IB, Handford PA. Dissecting the fibrillin microfibril: structural insights into organization and function. *Structure*. 2012;20:215–25.
33. Karttunen L, Raghunath M, Lonnqvist L, Peltonen L. A compound-heterozygous Marfan patient: two defective fibrillin alleles result in a lethal phenotype. *Am J Hum Genet*. 1994;55:1083–91.
34. Yuan X, Downing AK, Knott V, Handford PA. Solution structure of the transforming growth factor beta-binding protein-like module, a domain associated with matrix fibrils. *EMBO J*. 1997;16:6659–66.
35. Stahl-Hallengren C, Ukkonen T, Kainulainen K, Kristofersson U, Saxne T, Tornqvist K, et al. An extra cysteine in one of the non-calcium-binding epidermal growth factor-like motifs of the *FBN1* polypeptide is connected to a novel variant of Marfan syndrome. *J Clin Investig*. 1994;94:709–13.
36. Handford P, Downing AK, Rao Z, Hewett DR, Sykes BC, Kiely CM. The calcium binding properties and molecular organization of epidermal growth factor-like domains in human fibrillin-1. *J Biol Chem*. 1995;270:6751–6.
37. Jensen SA, Iqbal S, Lowe ED, Redfield C, Handford PA. Structure and interdomain interactions of a hybrid domain: a disulphide-rich module of the fibrillin/LTBP superfamily of matrix proteins. *Structure*. 2009;17:759–68.
38. Charbonneau NL, Carlson EJ, Tufa S, Sengle G, Manalo EC, Carlberg VM, et al. In vivo studies of mutant fibrillin-1 microfibrils. *J Biol Chem*. 2010;285:24943–55.
39. El-Hallous E, Sasaki T, Hubmacher D, Getie M, Tiedemann K, Brinckmann J, et al. Fibrillin-1 interactions with fibulins depend on the first hybrid domain and provide an adaptor function to tropoelastin. *J Biol Chem*. 2007;282:8935–46.
40. Jensen SA, Handford PA. New insights into the structure, assembly and biological roles of 10–12 nm connective tissue microfibrils from fibrillin-1 studies. *Biochem J*. 2016;473:827–38.
41. Yadin DA, Robertson IB, McNaught-Davis J, Evans P, Stoddart D, Handford PA, et al. Structure of the fibrillin-1 N-terminal domains suggests that heparan sulfate regulates the early stages of microfibril assembly. *Structure*. 2013;21:1743–56.
42. McGowan SE, Holmes AJ, Mecham RP, Ritty TM. Arg-Gly-Asp-containing domains of fibrillins-1 and -2 distinctly regulate lung fibroblast migration. *Am J Respir Cell Mol Biol*. 2008;38:435–45.
43. Zeyer KA, Zhang RM, Kumra H, Hassan A, Reinhardt DP. The Fibrillin-1 RGD Integrin binding site regulates gene expression

- and cell function through microRNAs. *J Mol Biol.* 2019;431:401–21.
44. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91.
 45. Sun QB, Zhang KZ, Cheng TO, Li SL, Lu BX, Zhang ZB, et al. Marfan syndrome in China: a collective review of 564 cases among 98 families. *Am Heart J.* 1990;120:934–48.
 46. Fuchs J. Marfan syndrome and other systemic disorders with congenital ectopia lentis. A Danish national survey. *Acta Paediatr.* 1997;86:947–52.
 47. Gray JR, Bridges AB, Faed MJ, Pringle T, Baines P, Dean J, et al. Ascertainment and severity of Marfan syndrome in a Scottish population. *J Med Genet.* 1994;31:51–4.
 48. Groth KA, Hove H, Kyhl K, Folkestad L, Gaustadnes M, Vejlstrup N, et al. Prevalence, incidence, and age at diagnosis in Marfan Syndrome. *Orphanet J Rare Dis.* 2015; 10:153.
 49. Kobayashi Y, Yang S, Nykamp K, Garcia J, Lincoln SE, Topper SE. Pathogenic variant burden in the ExAC database: an empirical approach to evaluating population data for clinical variant interpretation. *Genome Med.* 2017;9:13.