The FBN2 gene: new mutations, locus-specific database (Universal Mutation Database FBN2), and genotype-phenotype correlations.

Melissa Yana Frédéric, Christine Monino, Christoph Marschall, Dalil Hamroun, Laurence Faivre, Guillaume Jondeau, Hanns-Georg Klein, Luitgard Neumann, Elodie Gautier, Christine Binquet, et al.

To cite this version:
Melissa Yana Frédéric, Christine Monino, Christoph Marschall, Dalil Hamroun, Laurence Faivre, et al.. The FBN2 gene: new mutations, locus-specific database (Universal Mutation Database FBN2), and genotype-phenotype correlations.. Human Mutation, Wiley, 2009, 30 (2), pp.181-90. 10.1002/humu.20794. inserm-00343886

HAL Id: inserm-00343886
https://www.hal.inserm.fr/inserm-00343886
Submitted on 20 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The *FBN2* Gene: New Mutations, Locus-Specific Database (Universal Mutation Database *FBN2*), and Genotype-Phenotype Correlations

Melissa Yana Frédéric,1,2 Christine Monino,1,2 Christoph Marschall,3 Dalil Hamroun,4 Laurence Faivre,5–7 Guillaume Jondeau,5–11 Hanns-Georg Klein,3 Luitgard Neumann,12 Elodie Gautier,6,13 Christine Binquet,6,13 Cheryl Maslen,14 Maurice Godfrey,15 Prateek Gupta,16 Dianna Milewicz,18 Catherine Boileau,6,17–19 Mireille Claustres,1,2,4 Christophe Béroud,1,2,4 and Gwenaelle Collod-Béroud1,2+

1INSERM, U827, Montpellier, F-34000 France
2Université Montpellier 1, Unité Fonctionnelle de Recherche (UFR) de Médecine, Laboratoire de Génétique Moléculaire, Montpellier, F-34000 France
3Center of Human Genetics and Laboratory Medicine, Martinsried, Germany
4CHU Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Génétique Moléculaire, Montpellier, F-34000 France
5CHU Dijon, Hôpital d’Enfants, Centre de Génétique, Dijon, F-21000 France
6Centre d’Investigation Clinique—Épidémiologie Clinique/Essais Cliniques, Dijon, F-21000 France
7Univ de Bourgogne, Dijon, F-21000 France
8AP-HP, Hôpital Bichat, Consultation Marfan, Paris, F-75018 France
9AP-HP, Hôpital Bichat, Service de Cardiologie, Paris, F-75018 France
10INSERM, Paris, France
11Université Paris 7, UFR Médecine, Service de Cardiologie, Paris, F-75018 France
12Charite Campus Virchow-Klinikum, Berlin, Germany
13INSERM, CIE1, Dijon, F-21000 France
14Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon
15Center for Human Molecular Genetics, University of Nebraska Medical Center, Omaha, Nebraska
16Department of Internal Medicine and Neurosurgery, The University of Texas Medical School, Houston, Texas
17INSERM, U781, Paris, F-75015 France
18AP-HP, Hôpital Ambroise Paré, Laboratoire Central de Biochimie, d’Hormonologie et de Génétique Moléculaire, Boulogne, F-92100 France
19Université Versailles Saint Quentin-en-Yvelines, UFR Paris-Ile-de-France-Ouest (PIFO), Garches, F-92000 France

**ABSTRACT:** Congenital contractual arachnodactyly (CCA) is an extremely rare disease, due to mutations in the *FBN2* gene encoding fibrillin-2. Another member of the fibrillin family, the *FBN1* gene, is involved in a broad phenotypic continuum of connective-tissue disorders including Marfan syndrome. Identifying not only what is in common but also what differentiates these two proteins should enable us to better comprehend their respective functions and better understand the multitude of diseases in which these two genes are involved. In 1995 we created a locus-specific database (LSDB) for FBN1 mutations with the Universal Mutation Database (UMD) tool. To facilitate comparison of identified mutations in these two genes and search for specific functional areas, we created an LSDB for the *FBN2* gene: the UMD-*FBN2* database. This database lists 26 published and six newly identified mutations that mainly comprise missense and splice-site mutations. Although the number of described *FBN2* mutations was low, the frequency of joint dislocation was significantly higher with missense mutations when compared to splice site mutations. The database is freely available at http://umd.be.

**KEY WORDS:** fibrillin; FBN2; Beals-Hecht syndrome; congenital contractual arachnodactyly; CCA; database
**Introduction**

Congenital contractual arachnodactyly (CCA; MIM# 121050), or Beals-Hecht syndrome, is an autosomal dominant disease related to Marfan syndrome (MFS; MIM# 154700). These two similar syndromes are heritable connective tissue disorders caused by mutations in two genes belonging to the same family, the fibrillin family: FBN1 and FBN2, encoding fibrillin-1 and fibrillin-2, respectively. CCA was first described by Epstein et al. [1968] and subsequently by Grenier et al. [1969], several years before it was differentiated from MFS by Beals and Hecht [1971]. However, it has been suggested that the first CCA patient was reported, in fact, by Antoine-Bernard Marfan in 1896 [Marfan, 1896] with the description of the little girl Gabrielle who probably presented with CCA rather than MFS. Clinical diagnosis is difficult because there are numerous common characteristics shared between MFS and CCA, such as a so-called marfanoid appearance constituted by tall, slender, aortic aneurysms [Milewicz et al., 1996], and autosomal dominant inheritance [Ades et al., 2004; Kainulainen et al., 1994; Lonnqvist et al., 1994], 'MASS phenotype', familial or isolated forms of aortic dissection [Ades et al., 2004; Kainulainen et al., 1994; Lonnqvist et al., 1994], and autosomal dominantly inherited ectopia lentis (MIM# 129600) [Faivre et al., 2003]. These two genes display strong homology and are present at the surface of cells, permitting cells to anchor to the extracellular matrix. One of these sequences is situated in exon 24 [Ramirez and Pereira, 1999], and another one, specific to CCA, such as a so-called marfanoid appearance constituted by tall, slender, aortic malformation, and kyphoscoliosis.

CCA is a disease with a relatively good prognosis because, despite the fact that aortic dilatation has been documented in a few patients (3/47; see Supplementary Table S1; available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat), the progression of dilatation to aortic dissection has not been documented. Lethal forms of CCA are rare and are related to additional abnormalities on top of skeletal and ear features; either in the gastrointestinal system (duodenal or esophageal atresia, intestinal malrotation) or in the cardiovascular system (atrial or ventricular septal defects, interrupted aortic arch, single umbilical artery), consequences of alteration in development.

CCA is an autosomal dominant disorder. Its incidence is unknown, as well as the percentage of cases caused by de novo mutations. As with MFS, CCA does not have gender or ethnic predilection. It appears to be fully penetrant. Finally, germline mosaicism has been observed [Putnam et al., 1997; Wang et al., 1996].

To date, little is known about FBN2 mutations and their clinical spectrum. More data are available on FBN1 mutations that are implicated in MFS and numerous related pathologies such as severe neonatal Marfan syndrome [Kainulainen et al., 1994], autosomal dominantly inherited ectopia lentis (MIM# 129600) [Ades et al., 2004; Kainulainen et al., 1994; Lonnqvist et al., 1994], isolated skeletal features (or Marfanoid skeletal syndrome) [Milewicz et al., 1995], atypical MFS without cardiac involvement, involvement of the tracheal valve, aorta, skeleton, and skin (the "MASS phenotype"), familial or isolated forms of aortic aneurysms [Milewicz et al., 1996], and autosomal dominant Weill-Marchesani syndrome (MIM# 227600) [Faivre et al., 2003]. The Universal Mutation Database (UMD)-FBN1 locus-specific database (LSDB) today contains more than 1,700 mutations [Colloed-Beroud et al., 2003; Faivre et al., 2007].

The aim of our work was to create a new UMD database, to better analyze FBN2 mutations and to compare them with FBN1 mutations. Although the two genes display strong homology and similar modular organization, it is still unknown exactly what differentiates them functionally. The different roles of FBN1 and FBN2 could be explained by different promoters allowing spatiotemporally different expression and/or by different protein-protein interactions. Identifying mutations in specific regions of the two genes could allow identification of the functional regions of these two proteins. In order to accumulate and easily compare molecular and clinical data, we created a new LSDB, the UMD-FBN2 database.

**The FBN2 Gene is Homologous to FBN1**

The FBN2 gene was discovered while cloning the gene implicated in Marfan syndrome, the FBN1 gene located in 15q15–21.3. FBN2 maps in 5q23–31 [Lee et al., 1991]. The demonstration that mutations in the FBN2 gene were involved in CCA was made by genetic linkage analysis [Lee et al., 1991; Tsipouras et al., 1992] and the first implication of a FBN2 mutation in CCA was reported by Putnam et al. [1995].

FBN2 is a 279,57-kb gene (contig NT_034772) that encodes a 10,166-bp transcript and a 2,912–amino acid protein. The FBN1 and FBN2 genes are highly homologous at the nucleotide level; on comparison of 58% of the sequence, we obtained 84% homology between FBN1 and FBN2. The fibrillin family contains another member, FBN3, identified from a human fetal brain library [Nagase et al., 2001]. FBN3 is located in 19p13.3–19p13.2 and is composed of 65 exons. Its structure substantially resembles that of FBN1 and FBN2: amino acid identity is high, including conservation of all cysteine residues; and domain organization is consistent throughout each of the proteins [Corson et al., 1993].

The multidomain structure of the genes belonging to the fibrillin family contains five different modules (see representation of FBN2 secondary structure in Fig. 1B). The most common is the epidermal growth factor (EGF)-like module (by homology to epidermal growth factor) with 47 repeats. This particularly conserved sequence contains six cysteines involved in three disulfide bonds [Rao et al., 1995], which are formed between cysteines 1-3, 2-4, and 5-6 [Downing, et al., 1996]. A total of 43 of these contain a conserved consensus calcium-binding sequence: D/N-X-D/N-E/Q-Xα-Xα-Xα-Xα-Y/F (n and m are variable, * indicates possible β hydroxylation) [Handford et al., 1991; Rees et al., 1988] and have been demonstrated to bind calcium [Corson et al., 1993; Dietz and Pyeritz, 1995; Glanville et al., 1994]. Another motif containing eight cysteines is repeated seven times [Yuan et al., 1997] and is named transforming growth factor binding protein (TGFβP)-like module because of its homology with TGFβ binding protein (also named 8-cys modules) [Yuan et al., 1997]. According to SwissProt, we annotated "signal peptide" from amino acid 1 to 28. By analogy with the well-described FBN1 structure, we listed other domains not listed in SwissProt: "N-terminal domain" from amino acid 29 through 82 and "C-terminal domain" from amino acid 2,732 through the end of the protein. We annotate also a 4-cysteine motif (4-cys motif), also reported in latent TGF-β binding proteins (LTBPs), from amino acid 83 to 110 [Bashir et al., 1996; Isogai et al., 2003; Olofsson et al., 1995] and a domain presenting with homology to fibrilins C-terminal domain III (FibuCTDIII-like motif) from amino acid 2,758 to 2,912 [Gilgay et al., 1999]. The protein sequence encoded by FBN2 also contains twoarginine-glycine-aspartate (RGD) sequences, only one of which is present in FBN1. These sequences allow interaction with integrin receptors, which are present at the surface of cells, permitting cells to anchor to the extracellular matrix. One of these sequences is situated in exon 24 corresponding to codons 1,061 to 1,063, the other one, specific to FBN2 sequence, is in exon 37 corresponding to codons 1,586 to 1,588 [Ritty et al., 2003]. Finally, we annotated 12 FBN2 N-glycosylation sites conserved in FBN1 that contains a total of 15 amino acid residues: 492; 1,112; 1,414; 1,529; 1,625; 1,714; 1,745; 1,756; 1,945; 2,120; 2,225; and 2,808 [Ramirez and Pereira, 1999].
**Figure 1.** Comparison of the distribution of \textit{FBN1} and \textit{FBN2} mutations. A: Distribution of \textit{FBN1} and \textit{FBN2} mutations per exon. The number for each type of mutation (missense, deletions, insertions/duplications, and nonsense) is represented for each exon. Splice mutations are not part of this graph. Distribution of mutations identified in the \textit{FBN1} gene (second graph) or associated with neonatal MFS (first graph) are compared with mutations identified in the \textit{FBN2} gene and implicated in CCA (third graph). B: Schematic representation of the deduced primary structure of fibrillin-2. Part of the fibrillin-2 protein corresponding to exons 24 to 33, where mutations implicated in CCA are clustered, is boxed.
and very conserved amino acids of unknown function include calcium binding (cb)EGF-like, TGFβP-like, EGF-like modules, as well as in 4-cys motif (Cterm) and in FibuCTDIII-like motif (Nterm). However, contrary to FBN1 and according to the FBN2 sequence, we have not identified furine/pace (consensus site R-X-K/R-R) or matrix metalloproteinase sites. To our knowledge, these sites have not been reported in the FBN2 gene.

**Expression of FBN2 and FBN1 Genes and Protein Function**

The two ~320-kD proteins encoded by FBN1 and FBN2, fibrillin-1 and fibrillin-2, respectively, are also highly homologous in terms of amino acid sequences. These large cysteine-rich glycoproteins are major components of microfibrils from extracellular matrix, which can be associated with elastin fibers. In organ culture experiments, both fibrillins are similarly expressed in most tissues, except for kidney, liver, rib anlage, and notochord [Quondamatteo et al., 2002]. However, temporal expression of the two proteins is quite different during embryogenesis. Fibrillin-2 is expressed earlier in development [Mariencheck et al., 1995], except in the cardiovascular system, and appears to have a potential role in elastogenesis, while fibrillin-1 is expressed later and has a fundamental structural role. Moreover, fibrillin-2 is expressed mostly during embryogenesis by very different types of tissues [Zhang et al., 1994, 1995]. By studying a rat model, it has been suggested that fibrillin-2 could play a role in lung development [Yang et al., 1999]. Fibrillin-1 and fibrillin-2 could perform equivalent architectural functions in some tissues while fibrillin-2 may have a distinct function in peripheral nerves [Charboneau et al., 2003].

The two proteins can be divided into five regions called A, B, C, D, and E. The most different domains are the A domains, with 19% homology between the two fibrillins; the C domain, which is proline-rich in fibrillin-1 and glycine-rich in fibrillin-2; and the E domain, with 50% homology. The two domains B and D share more homologies, with 87% and 81% homology, respectively. This can be explained by the high representation of conserved modules presenting homology to EGF; called the EGF-like module [Corson et al., 1993; Pereira et al., 1993].

**The UMD-FBN2 Database**

Codons are numbered with respect to the FBN2 gene cDNA sequence (+1 = A of ATG). Two reference sequences have been used in previous mutation reports. These two references differ in their nucleotide sequence and by the presence of an additional amino acid (A192Q193P194 in HGMD and SwissProt; or G192P193N194R195 in GenBank NM_001999). As sequencing of more than 150 individuals only displayed "G192P193N194R195" (C. Maslen, and C. Marschall, personal communication), mutation names were renumbered with respect to the FBN2 gene cDNA sequence obtained from the GenBank database (GenBank database accession number NM_001999, complete coding sequence of human fibrillin-2). Amino acid numbers can then be different from those published (add +1 to the published amino acid [AA] number and +3 to the published nucleotide number for residues located downstream of AA 195).

Intron-exon boundaries, as well as intronic sequences, were defined by matching the cDNA sequence to the corresponding genome sequence (NT_0347772) and module organization from SwissProt (accession number P35556 with corrections for residues located downstream of AA 195). The database follows the guidelines on mutation databases of the Hugo Mutation Database Initiative including the latest nomenclature (www.hgvs.org).

We have used UMD software to create a computerized database that currently contains information about the published mutations of the FBN2 gene, mutations only reported in meeting proceedings or contributed by the coauthors of this work. The mutation records can list point mutations, large and small deletions, insertions, and mutations affecting splicing (intronic mutation) in the FBN2 gene. It cannot accommodate complex mutations. In addition, two mutations affecting the same allele will be entered as two different records linked by the same sample ID. For each mutation, information is provided at different levels: genetic (exon and codon number, wild-type and mutant codon, mutational event, mutation name), protein (wild-type and mutant amino acid, affected domain, mutation name), and clinical (skeletal, cardiovascular, gastrointestinal symptoms, when data are available).

We have annotated the FBN2 sequence with all indirect arguments to define whether potential missense mutations are really causative. These arguments list cysteine implicated in disulfide bonds in EGF-like, cbEGF-like, and TGFβP-like modules according to SwissProt [Downing et al., 1996; Rao et al., 1995], N-glycosylated amino acids [Ramirez and Pereira, 1999], amino acids implicated in RGD sequences [Ritty et al., 2003], and highly conserved amino acids of unknown function in cbEGF-like, TGFβP-like, and EGF-like modules, as well as in 4-cys motif (Cterm) and in FibuCTDIII-like motif (Nterm).

Study of the database can be done with different routines described on our internet website at www.umd.be. It is possible to analyze mutation distribution by exon, mutation distribution by mutation type, and mechanism for deletion/insertion [Beroud et al., 2000, 2005]. With the annotation highly conserved domain ("HCD"), it is possible to know if an amino acid has a known function in the protein.

The software has already been used successfully for other genes implicated in genetic diseases, such as FBN1 [Collod et al., 1996; Collod-Beroud et al., 1997, 1998, 2003, LDLR [Varet et al., 1997, 1998; Villeger et al., 2002], TGFBR2 (submitted), TGFBR1 (in construction), ALS2, ATP7B, CAPN3, CDH23, CFTR, DMD, DPYD, DYSF, E2F, EMD, EIF2B, GJB2, GPAP, LMNA, MYO7A, PCDH15, PLP1, TGFBR1, USH1C, USHA2, USH3A, SANS, SGCA, SGCG, FKRP, and ZMPSTE24 (unpublished); and for different types of genes implicated in cancer, such as TP53 [Beroud and Soussi, 1997, 1998, 2003; Beroud et al., 1996; Soussi et al., 2000], APC [Beroud and Soussi, 1996, 1997; Laurent-Puig et al., 1998], MEN1 [Wautot et al., 2002], VHL [Beroud et al., 1998; Gallou et al., 1999], WT1 [Jeanpierre et al., 1998], BRCA1, BRCA2, MEN1, MLH1, MSH2, MSH6, MYH, PMS2, and SUR1 (unpublished). More information concerning the UMD software is available at www.umd.be [Beroud et al., 2000, 2005].

The current database and subsequent updated versions are available online at www.umd.be. Notifications of omissions and errors in the current version as well as specific phenotypic data would be gratefully received by the corresponding author. The software will be expanded as the database grows and according to the requirements of its users, and new functions may be implemented.

**Mutation Analysis**

Today, the database contains 32 entries indexed in the UMD-FBN2 locus specific database (Tables 1 and 2). One mutation is only reported in a meeting proceeding (p.E390K) [Wang et al.,
Table 1. Missense, nonsense and insertion/duplication FBN2 mutations

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Exon</th>
<th>Codon</th>
<th>WT codon</th>
<th>Mutant codon</th>
<th>Event Type</th>
<th>WT AA</th>
<th>Mutant AA</th>
<th>Structure</th>
<th>HCD</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missenses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>9</td>
<td>390</td>
<td>TGT</td>
<td>A Ts Cys</td>
<td>TGFBP</td>
<td>Lys</td>
<td>TGFBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>24</td>
<td>1057</td>
<td>GGC</td>
<td>G Ac Cys</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>25</td>
<td>1093</td>
<td>ATC</td>
<td>T Ac Cys</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>3278</td>
<td>T</td>
<td>C</td>
<td>C in disulfide bonds</td>
<td>3278-3281</td>
<td>Tp</td>
<td>cb EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>26</td>
<td>1142</td>
<td>TGC</td>
<td>TAc Cys</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>3425</td>
<td>G</td>
<td>C</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1142-1156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>3535</td>
<td>G</td>
<td>C</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1142-1156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>27</td>
<td>1198</td>
<td>TGT</td>
<td>A Ts Cys</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>3593</td>
<td>G</td>
<td>C</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1185-1198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>3718</td>
<td>T</td>
<td>C</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1227-1240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>3758</td>
<td>G</td>
<td>A</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1253-1266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>3771</td>
<td>C</td>
<td></td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1246-1257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>3802</td>
<td>T</td>
<td>TGC</td>
<td>C Ac Cys</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>4274</td>
<td>T</td>
<td>TAc Cys</td>
<td>cb</td>
<td>EGF-like</td>
<td>Tp</td>
<td>1412-1425</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insertion/Duplication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>G</td>
<td>33</td>
<td>1434</td>
<td>TGT</td>
<td>InF Asn</td>
<td>cb</td>
<td>Tp</td>
<td>EGF-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonsense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>19</td>
<td>130</td>
<td>Stop</td>
<td>Stop</td>
<td>Cb</td>
<td>Stop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>31</td>
<td>142</td>
<td>Stop</td>
<td>Stop</td>
<td>Cb</td>
<td>Stop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>4</td>
<td>326</td>
<td>CCT</td>
<td>Int dup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Complex Mutation

One is a complex genomic rearrangement consisting in skipping of exon 24 and duplication of exon 23 [Gupta et al., 2002] and cannot be indexed in the database.

Nonsense and Missense Mutations

Only one reported mutation is a nonsense mutation: c.4296C>A [Gupta et al., 2002]. A total of 15 are missense mutations. A total of 11 are mutations which modify a cysteine residue implicated in a disulfide bond: c.1198G>A [Wang et al., 1995b]; c.3758G>A and c.4301G>C [Putnam et al., 1995]; c.3425G>T and c.3759T>G [Belleh et al., 2000]; c.3593G>A, c.3718T>C, c.3771C>G, and c.3802T>C [Gupta et al., 2002]; and c.3425G>A and c.4274G>T (this report). There are three mutations modifying an amino acid residue potentially implicated in Ca2+ binding, according to Dietz and Pyeritz [1995]: c.3340G>C [Dubock et al., 1998]; c.3278T>C [Park et al., 1998]; and c.3353G>T [Gupta et al., 2002]. One is a mutation that modifies an amino acid residue conserved in TGFBP-like domains: c.3710G>A [Park et al., 1998]. All 15 missense mutations (including the c.976C>T mutation not located in an HCD) are predicted to be pathogenic according to conservation and biochemical data with the UMD-Predictor tool (unpublished results).

Insertion/Duplication

Only one reported mutation is an insertion, corresponding to a small duplication of an existing three-base sequence.

Splice-Site Mutations

A total of 14 are splice mutations. Three mutations modify an acceptor splice-site: c.IVS33–2A>T [Wang et al., 1996], c.IVS28–3C>G (this report), and c.IVS30–5T>G [Nishimura et al., 2004], and eight modify a donor splice-site: c.IVS31–1G>C [Park et al., 1998], c.IVS26+2T>C and c.IVS26+2T>G [Gupta et al., 2002], c.IVS32+3G>A [Gupta et al., 2004], c.IVS32+2T>C (this report), c.IVS31+1G>A and c.IVS35+1G>T [Nishimura et al., 2007], and c.3343G>C (last base of exon 25, in donor splice-site consensus) [Babcock et al., 1998]. The software can compute the variation of the consensus value (CV) of the mutan "splice-site" vs. the wild-type. Wild type, mutated CVs as well as the variation between these values are indicated in Table 2. If variation is >10%, it is highly probable that the mutated site is preferred to the wild-type. As observed in Table 2, significant variations are found for 10 out of the 11 mutations. More arguments are then needed for the c.IVS30–5T>G mutation (variation of ~9,51) [Nishimura et al., 2007]. Alternative splicing has been demonstrated in only five cases at the RNA level (Table 2). Three mutations are not localized in the canonical donor or acceptor splice-site consensus sequences. Two mutations localized in –24 and –26 of intron 30 are in the vicinity of the branchpoint, c.IVS30–26G>T [Maslen et al., 1997] and c.IV-
Table 2. Splice FBN2 mutations

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Splice site type</th>
<th>Wild type sequence</th>
<th>WT CV</th>
<th>Mutant sequences</th>
<th>Mutant CV</th>
<th>Variation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acceptor splice sites</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.IVS33-2A&gt;T (c.4346-2A&gt;T)</td>
<td>Acceptor</td>
<td>gcttttgcacagA</td>
<td>94.61</td>
<td>gcttttgcacagA</td>
<td>78.69</td>
<td>−16.83</td>
<td>Skipping of exon 34, in frame</td>
</tr>
<tr>
<td>c.IVS28-3C&gt;G (3725-3C&gt;G)</td>
<td>Acceptor</td>
<td>tgttgtttcacagA</td>
<td>93.13</td>
<td>tgttgtttcacagA</td>
<td>80.71</td>
<td>−13.34</td>
<td>Not demonstrated, probable in frame abnormal splicing</td>
</tr>
<tr>
<td>c.IVS30-5T&gt;G (c.3974-5T&gt;G)</td>
<td>Acceptor</td>
<td>aataaatttctagA</td>
<td>68.16</td>
<td>aataaatttctagA</td>
<td>61.68</td>
<td>−9.51</td>
<td>Not demonstrated, more argument needed</td>
</tr>
<tr>
<td><strong>Donor splice sites</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.IVS31+1G&gt;C (c.4099+1G&gt;C)</td>
<td>Donor</td>
<td>AGgtaggt</td>
<td>68.54</td>
<td>AGgtaggt</td>
<td>70.36</td>
<td>−20.53</td>
<td>Skipping of exon 31, in frame</td>
</tr>
<tr>
<td>c.IVS26+2T&gt;C (c.3472+2T&gt;C)</td>
<td>Donor</td>
<td>GGGtaagc</td>
<td>86.73</td>
<td>GGGtaagc</td>
<td>68.54</td>
<td>−20.96</td>
<td>Not demonstrated, probable in frame abnormal splicing</td>
</tr>
<tr>
<td>c.IVS31+1G&gt;A (c.4099+1G&gt;A)</td>
<td>Donor</td>
<td>AGgtaggt</td>
<td>86.55</td>
<td>AGgtaggt</td>
<td>70.36</td>
<td>−20.53</td>
<td>Not demonstrated, probable in 86.36frame abnormal splicing</td>
</tr>
<tr>
<td>c.IVS35+1G&gt;T (c.4594_1G&gt;T)</td>
<td>Donor</td>
<td>AGgtaggt</td>
<td>83.64</td>
<td>AGgtaggt</td>
<td>65.45</td>
<td>−21.74</td>
<td>Not demonstrated, probable in frame abnormal splicing</td>
</tr>
<tr>
<td>c.3343G&gt;C (last base of exon 25)</td>
<td>Donor</td>
<td>GGGtaagc</td>
<td>86.36</td>
<td>GGGtaagc</td>
<td>72.54</td>
<td>−16.00</td>
<td>Skipping of exon 25, in frame</td>
</tr>
<tr>
<td><strong>Mutations not localised in canonical sequences</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.IVS30-26G&gt;T (c.3974-26G&gt;T)</td>
<td>No branch point modification, no creation of splice site consensus</td>
<td>caatgtgtgcaat</td>
<td>74.50</td>
<td>caatgtgtgcaat</td>
<td>74.50</td>
<td>0</td>
<td>Skipping of exon 31, in frame</td>
</tr>
<tr>
<td>c.IVS30-24A&gt;G (c.3974-24A&gt;G)</td>
<td>No branch point modification, no creation of splice site consensus</td>
<td>caatgtgtgcaat</td>
<td>74.50</td>
<td>caatgtgtgcaat</td>
<td>74.50</td>
<td>0</td>
<td>Skipping of exon 31, in frame</td>
</tr>
</tbody>
</table>

Consensus values for each potential donor or acceptor splice site are calculated for Wild Type (WT) and mutant sequence according to Senapathy et al. [1990]; and Shapiro and Senapathy [1987] (100 = strong splice site; 0 = not a splice site). Observed consequences describe the abnormalities found at the mRNA level.

*Identified twice.
S30–24A>C [Nishimura et al., 2007], but do not involve one of the seven nucleotides of the most probable branchpoint consensus sequence (–35 to –29, agcacAt) (see Supplementary Fig. S1). One of these two mutations was confirmed to cause missplicing at the mRNA level [Maslen et al., 1997] thus suggesting the presence of a regulatory sequence in this intronic region. The natural acceptor splice site in intron 30 is weak (68.16) and the disturbance of the surrounding sequence could play an important role in splicing. Further experimental analyses are needed to demonstrate this hypothesis and identify this key sequence regulator. The last mutation not localized in a canonical sequence is c.IVS28–15A>G [Putnam et al., 1997]. This mutation in intron 28 does not involve one of the seven nucleotides of the most probable branchpoint consensus sequence (–34 and –28, acctgAc) (see Supplementary Fig. S4). Furthermore, the natural acceptor splice-site in this intron is strong (89.40). This mutation creates a potential alternative acceptor splice-site (caatgtggttgcag, 74.50). Nevertheless, abnormal splicing corresponding to the skipping of exon 29 has been reported. The mechanism of this skipping is thus unclear.

In conclusion, FBN1 mutations leading to a premature termination codon are significantly underrepresented (only one nonsense mutation is reported) contrasting with an overrepresentation of in-frame mutations (splice-site mutations are likely all in-frame and the unique duplication results in in-frame insertion). This particular distribution of mutation types is similar in the homologous FBN1 region (exons 24–32; c.f. paragraph 6). Nevertheless, contrary to FBN1, where mutations in this region are mainly missense (56.9% for 12.5% of splice-site mutations), the mutations currently reported in the FBN2 gene are evenly distributed between missense (50%) and splice-site mutations (46.9%).

Structure/Function Correlations

All known mutations in the FBN2 gene except two, c.976C>T (this report) and c.1168G>A [Wang et al., 1995a], are clustered in a limited region between exons 24 through 35 (Fig. 1A; graph 3). Curiously, the homologous region in the FBN1 gene (exons 24–32) clusters almost all mutations implicated in neonatal forms of MFS that correspond to the most severe end of the MFS clinical spectrum [Park et al., 1998] (Fig. 1A; graph 2). Clinical features found in neonatal MFS and CCA can be overlapping as congenital contractures, dolichostenomelia, arachnodactyly, and abnormal ears are found in both cases. Other FBN1 mutations, implicated in a large spectrum of clinical phenotypes, are spread throughout the gene (Fig. 1A; graph 1). Furthermore, FBN1 mutations in exons 24 to 32 not implicated in neonatal cases were associated with a more severe phenotype, including younger age at diagnosis, higher probability of ectopia lentis, ascending aortic dilatation, aortic surgery, mitral valve abnormalities, sclerosis, and shorter survival. Therefore, more than a “neonatal region,” it should be considered a “severe region.” The presence of a mutation in this limited FBN1 region appears to be the best indicator of early onset aortic risk [Fairev et al., 2007]. Only four patients with a known FBN2 mutation among 47 had aortic root dilatation at an early age and none of these individuals are known to have progressed to aortic dissection (Supplementary Table S1). This region seems then to have a different function in fibrillin-2. Nevertheless, even if the presence of a cardiac implication is rare and the prognosis for CCA patients with these abnormalities is not known, cardiac evaluation including echocardiogram must be set up for CCA patients. The most severe end of the CCA clinical spectrum corresponds to a severe/lethal form. Molecular studies of only one individual (IVS33–2A>T [Wang et al., 1996]) among the six cases reported with severe/lethal CCA [Currarino and Friedman, 1986; Godfrey et al., 1995; Lipson et al., 1974; Macnab et al., 1991; Wang et al., 1996] have been performed and the identified mutation is localized in this specific region. It would be very interesting to search for the molecular abnormalities in the other reported cases and look for their localization in or out of this limited region of the FBN2 gene.

This middle-FBN1 region is the beginning of the longest stretch of EGF-like motifs. Immunohistochemical studies of neonatal patients’ cells’ culture showed important decrease of immunoreactive material when stained with anti-fibrillin-1 antibodies and an abnormal microfibrils morphology: fibrils were short, fragmented, and frayed [Raghunath et al., 1993; Superti-Furga et al., 1992]. Rotary shadowing electron microscopy analyses of these cell cultures had no identifiable 10-nm-diameter microfibrils at all [Lonqvist et al., 1996]. Alteration of this region of the protein then, undoubtedly has a significant and specific effect on fibrillin-1-containing microfibril formation and their stability in the extracellular matrix. Clarifying the specific functional significance of this region in FBN1 during early development may help in understanding the specific function of the corresponding FBN2 sequence.

Finally, the clustering of CCA mutations in this FBN2 region raises the question of the phenotypes associated with mutations localized outside of this region. It is not known today if these FBN2 mutations are either silent, result in different phenotypes (part of the spectrum of phenotypes associated with a FBN1 mutation or with no relationship), or are lethal. If the clinical spectrum associated with FBN2 mutations is comparable with those linked to FBN1 mutations, mutations localized in this particular region should correspond to the most severe end of the spectrum and mutations that occur outside this region would be associated with milder phenotypes.

Genotype-Phenotype Correlations

Methodology

To study the effect of mutation type, we compared patients with a missense mutation to patients with a splice-site mutation for each clinical feature: abnormal ears, micrognathia, high arched palate, arachnodactyly, joint contractures, joint dislocation, scoliosis, pectus deformity, pes deformity, dolichostenomelia, orthopedic treatment, and aortic root dilatation and refraction. As the age at diagnosis of each clinical feature was not collected, age at last follow-up was the only information available. To indirectly take into account length of patient of follow-up in this situation, we adjusted all comparisons of CCA manifestation frequencies for the ages at last follow-up, and categorized them into 10-year age groups. These adjusted comparisons were performed using the Mantel-Haenszel (MH) test [Mantel and Haenszel, 1959]. As this test is adopted only if the relationship between the mutation type and the clinical manifestation is similar in the different strata of age at last follow-up, we checked the homogeneity between strata using the Breslow-Day chi-squared test of homogeneity [Breslow and Day, 1987]. If an interaction was observed, results were presented for each category of age at last follow-up. Stata software version 8 (Stata Corp., College Station, TX) was used for all statistical analyses. Only P values <0.005 were considered as significant since multiple tests were performed.
Only mutations with complete clinical descriptions have been taken into account in these analyses (mutations c.4296C>A and c.3269_3271dup excluded). For some mutations, clinical descriptions available for other affected relatives carrying the same FBN2 mutation were included (IVS26+2T>G, IVS28–15A>G, IVS30–26G>T, IVS32+5G>A, IVS33–2A>T, IVS30–24A>C, c.3170G>A, c.3278T>C, and c.3343G>C) (see Supplementary Fig. S1).

**Results**

A total of 29 individuals with a splice-site mutation and 19 with a missense mutation were included in these analyses (a total of 23 distinct mutations are taken into account). The only significant result was the frequency of joint dislocation that was significantly higher with missense mutations when compared to splice-site mutations (MH test, \( P = 0.0028 \)).

**Polymorphisms**

To assist in diagnosis, we also list published or contributed FBN2 polymorphisms (Supplementary Tables S2 and S3). The aim is to make available a complete set of FBN2 gene variations (mutations + polymorphisms) to the community, which may be used to quickly identify a causative mutation, thus saving time and money on testing for a polymorphism in a control population. These polymorphisms are numbered according to NM_001999 with G192P, N194R, K195. Coding FBN2 gene variations are represented in Supplementary Table S2 and noncoding in Supplementary Table S3.

**Future Prospects**

The clustering of CCA mutations in the limited region between FBN2 exons 24 through 35 raises the question of the existence or not of phenotypes associated with FBN2 mutations localized inside and outside of this region. The mutational events in previously reported severe/lethal CCA cases [Currarino and Friedman, 1986; Godfrey et al., 1995; Lipson et al., 1974; Macnab et al., 1991; Wang et al., 1996], as well as in isolated associated CCA signs, such as arthrogryposis, kyphoscoliosis, and osteopetrosia, must now be investigated to define just how broad the clinical spectrum associated with FBN2 mutations is. For example, familial forms of severe scoliosis could be investigated for the presence of FBN2 gene mutations.

**Acknowledgments**

This work was supported by grants from AFM (Association Française contre les Myopathies) (AFM-R07017FF) (to C.Boi. and D.H.), Université Montpellier 1, INSERM, GIS-Institut des Maladies Rates 2004 (to C.Boi. and G.J.), and Agence Nationale pour la Recherche (ANR) (ANR-05PCOD-14-02 to C.Boi. and G.J.). M.Y.E. is supported by a grant from AFM.


