 Nucleic Acids Research, 1998, Vol. 26, No. 1 229–233

Marfan Database (third edition): new mutations and new routines for the software

Gwenaëlle Collod-Béroud[1,2], Christophe Béroud[1,3], Lesley Ades[4,5], Cheryl Black[6], Maureen Boxer[7], David J. H. Brock[8], Katherine J. Holman[4], Anne de Paepe[9], Uta Francke[10,11], Ulrich Grau[12], Caroline Hayward[8], Hanns-Georg Klein[13], Wanguo Liu[11], Lieve Nuytinck[14], Leena Peltonen[15], Ana Beatriz Alvarez Perez[16], Terhi Rantamäki[15], Claudine Junien[1,17] and Catherine Boileau[1,17,*]

[1]INSERM U383, Hôpital Necker-Enfants Malades, Université René Descartes, Paris V, 149-161 rue de Sèvres, 75743 Paris Cedex 15, France, [2]INSERM U129, Institut Cochin de Génétique Moléculaire, CHU Cochin Port Royal, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France, [3]Pharmacie, Hopital Broussais, 96 rue Didot, 75014 Paris, France, [4]Department of Medical Genetics and [5]Department of Paediatrics and Child Health, Royal Alexandra Hospital for Children New, PO Box 3515, Parramatta, Sydney 2124, Australia, [6]Molecular Genetics Laboratory, Department of Molecular and Cellular Pathology, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK, [7]North Thames (East) Clinical Molecular Genetics Laboratory, Unit of Clinical Genetics and Fetal Medicine, Institute of Child Health, 30 Guilford Street, London, UK, [8]Human Genetics Unit, Department of Medicine (WGH), University of Edinburgh, Molecular Medicine Center, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK, [9]Cardiological Sciences, Saint George's Hospital Medical School, Cranmer Terrace, London SW17 ORE, UK, [10]Department of Genetics and [11]Howard Hughes Medical Institute, Stanford University Medical Center, Stanford, CA 94305-5323, USA, [12]Department of Cardiovascular Surgery and [13]Institute of Clinical Chemistry, Klinikum GroBhadern, University of Munnich, D-81366 München, Germany, [14]University Hospital Gent, OK5, Center for Medical Genetics, De Pintelaan 185, B-9000 Gent, Belgium, [15]Department of Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166, FIN-00300 Helsinki, Finland, [16]UNIFESP-EPM, Rua Vergueiro, 1921 AP. 111, 04101-000 Sao Paulo SP, Brazil and [17]Laboratoire Central de Biochimie, d'Hormonologie et de Génétique moléculaire, Hôpital Ambroise Paré, 9 avenue Charles de Gaulle, 92104 Boulogne, France

Received October 1, 1997; Accepted October 3, 1997

ABSTRACT

The Marfan database is a software that contains routines for the analysis of mutations identified in the FBN1 gene that encodes fibrillin-1. Mutations in this gene are associated not only with Marfan syndrome but also with a spectrum of overlapping disorders. The third version of the Marfan database contains 137 entries. The software has been modified to accommodate four new routines and is now accessible on the World Wide Web at http://www.umd.necker.fr

FIBRILLIN, MARFAN SYNDROME AND TYPE 1 FIBRILLINOPATHIES

Fibrillin-1 is the principal structural element of a class of connective tissue microfibrils that have a widespread distribution (1). In elastic tissues, fibrillin microfibrils play a key role in elastic fibrillogenesis and are components of elastic fibers which generate elastic recoil (2,3). In non-elastic tissues, they are proposed to play an anchoring role (4). Fibrillin-1 is encoded by a relatively large and fragmented gene (65 exons distributed over ~110 kb) located at 15q15–q21.1 (5–8). It has a complex multi-domain structure comprising 47 epidermal growth factor (EGF)-like modules (43 of which have a calcium-binding consensus sequence) interspersed with seven '8-cysteine' repeats with homology to the TGF-β1 binding protein and two 'hybrid' modules.

Marfan syndrome (MFS) is an autosomal dominant disorder affecting mainly the cardiovascular, skeletal and ocular systems (9). The reported incidence is at least 1 per 10 000 with >25% of cases being the result of new mutations. The disease is associated with mutations in the gene encoding fibrillin-1 (FBN1). More recently, defects in this gene have been shown to cause a wide spectrum of microfibrilopathies, called 'type-1 fibrillinopaties', ranging from isolated skeletal features of Marfan syndrome or familial ectopia lentis to neonatal Marfan syndrome at the most severe end.

*To whom correspondence should be addressed at: INSERM U383, Hôpital Necker-Enfants Malades, Clinique Maurice Lamy, 149–161 rue de Sèvres, 75743 Paris Cedex 15, France. Tel: +33 1 44 49 44 84; Fax: +33 1 47 83 32 06; Email: boileau@necker.fr

Table 1. Each line represents a single FBN1 mutation report

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
10	2	165	55	GGA	del83c	Stop at 100	248 +1 G->A	Gly	Frameshift	NH2 unique region	?	?	+	?	?	?	19
111	3	331	111	TGT	CGT	T->C	C111R	Cys	Arg	EGF-like #1	+	+	-	-	-	-	NP3
37	4	364	122	CGC	TGC	C->T	R122C (1)	Arg	Cys	EGF-like #2	+	+	-	?	+	?	32
85	4	364	122	CGC	TGC	C->T	R122C (2)	Arg	Cys	EGF-like #2	+	+	+	-	+	-	NP2
40	4	386	129	TGC	TAC	G->A	C129Y	Cys	Tyr	EGF-like #2	+	+	+	+	+	?	13
41	5	497	166	TGT	TTT	G->T	C166F	Cys	Phe	EGF-like #3	?	?	+	?	?	?	13
63	5	497	166	TGT	TCT	G->C	C166S	Cys	Ser	EGF-like #3	?	?	?	?	?	?	38
20	6	649	217	TGG	GGG	T->G	W217G	Trp	Gly	Hybrid motif #1	+	+	+	?	?	?	24
133	6	718	240	CGC	TGC	C->T	R240C	Arg	Cys	Hybrid motif #1	+	+	+	-	-	-	NP4
128	11	1421	474	TGT	del48b	del	G1468T	Cys	Frameshift	EGF-like #4	+	+	+	-	+	-	NP8
26	11	1426	476	TGC	GGC	T->G	C476G	Cys	Gly	EGF-like #4	+	+	+	?	?	?	25
51	13	1604	535	TTA	del1b	Stop at 578	1604delT	Leu	Frameshift	cb EGF-like #04	?	?	?	?	?	?	13
112	13	1633	545	CGC	TGC	C->T	R545C (1)	Arg	Cys	cb EGF-like #04	+	+	+	-	-	-	NP3
135	13	1633	545	CGC	TGC	C->T	R545C (2)	Arg	Cys	cb EGF-like #04	+	+	+	+	+	-	NP4
6	13	1643	548	AAC	ATC	A->T	N548I	Asn	Ile	cb EGF-like #04	+	+	+	?	+	?	19
74	13	1693	565	CGA	TGA	C->T	R565X	Arg	Stop	cb EGF-like #04	?	?	?	?	?	?	43
76	14	1760	587	TGT	TAT	G->A	C587Y	Cys	Tyr	cb EGF-like #05	+	+	-	?	?	?	44
134	14	1794	598	TGC	TGG	C->G	C598W	Cys	Trp	cb EGF-like #05	+	+	+	-	+	+	NP4
102	14	1836	612	AAA	del1c	Stop at 624	1836delA	Lys	Frameshift	cb EGF-like #05	+	+	+	-	-	-	NP7
12	15	1879	627	CGT	TGT	C->T	R627C (1)	Arg	Cys	cb EGF-like #06	+	+	+	-	-	-	21
75	15	1879	627	CGT	TGT	C->T	R627C (2)	Arg	Cys	cb EGF-like #06	?	?	?	?	?	?	43
56	16	1981	661	TGC	CGC	T->C	C661R	Cys	Arg	8-Cys #2	?	?	?	?	?	?	35
87	17	2113	705	GCG	ACG	G->A	A705T	Ala	Thr	8-Cys #2	+	+	+	-	+	-	48
86	17	2132	711	TGC	TAC	G->A	C711Y	Cys	Tyr	8-Cys #2	+	+	+	-	+	-	48
8	18	2168	723	GAT	GCT	A->C	D723A	Asp	Ala	cb EGF-like #07	+	+	+	?	?	?	19
42	18	2237	746	TAT	TGT	A->G	Y746C	Tyr	Cys	cb EGF-like #07	?	?	?	?	?	?	13
90	18	2243	748	TGT	del51b	del	2293 +2 T->C	Cys	Frameshift	cb EGF-like #07	+	-	+	?	?	?	41
13	18	2248	750	TGC	GGC	T->G	C750G	Cys	Gly	cb EGF-like #07	+	+	+	-	-	-	21
105	20	2447	816	TGC	TCC	G->C	C816S	Cys	Ser	cb EGF-like #09	+	+	+	-	-	-	NP1
27	21	2584	862	TGT	CGT	T->C	C862R	Cys	Arg	Hybrid motif #2	+	+	+	?	?	?	26
124	21	2668	890	TGC	CGC	T->C	C890R	Cys	Arg	Hybrid motif #2	+	+	+	-	-	-	50
43	23	2776	926	TGT	CGT	T->C	C926R	Cys	Arg	cb EGF-like #10	?	?	?	?	?	?	13
150	24	2950	984	GTC	ATC	G->A	V984I	Val	Ile	8-Cys #3	+	+	+	-	+	+	15, NP6
84	24	2986	996	TGT	CGT	T->C	C996R	Cys	Arg	8-Cys #3	+	+	+	-	+	-	NP2
44	24	3037	1013	GGA	AGA	G->A	G1013R (1)	Gly	Arg	8-Cys #3	+	+	+	?	?	?	13
70	24	3037	1013	GGA	AGA	G->A	G1013R (2)	Gly	Arg	8-Cys #3	?	?	?	?	?	?	40
104	24	3037	1013	GGA	CGA	G->C	G1013R (3)	Gly	Arg	8-Cys #3	+	+	+	-	-	-	NP1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
18	24	3069	1023	AAG	AAC	G->C	K1023N	Lys	Asn	8-Cys #3	+	+	+	?	+	?	24
19	25	3083	1028	GAT	del126b	del	3208 +5 G->T	Asp	Frameshift	cb EGF-like #11	+	+	+	?	?	?	24
141	25	3095	1032	TGC	TAC	G->A	C1032Y	Cys	Tyr	cb EGF-like #11	+	-	+	-	+	-	NP2
81	25	3128	1043	AAG	AGG	A->G	K1043R	Lys	Arg	cb EGF-like #11	+	-	+	+	+	-	47
67	25	3142	1048	ATT	del3a	del	3142delATT	Ile	Frameshift	cb EGF-like #11	+	+	+	?	+	?	39
78	25	3143	1048	ATT	ACT	T->C	I1048T	Ile	Thr	cb EGF-like #11	+	-	+	-	+	-	45
66	25	3157	1053	TGC	CGC	T->C	C1053R	Cys	Arg	cb EGF-like #11	+	+	+	?	?	?	39
89	25	3163	1055	TGT	GGT	T->G	C1055G	Cys	Gly	cb EGF-like #11	+	+	+	-	+	-	48
33	25	3174	1058	GGC	ins3c	ins	3174insTGC	Gly	Frameshift	cb EGF-like #11	+	+	+	+	+	-	29
82	25	3191	1064	GAA	del1b	Stop at 1087	3192delA	Glu	Frameshift	cb EGF-like #11	+	-	+	?	?	?	47
64	26	3215	1072	GAC	GGC	A->G	D1072G	Asp	Gly	cb EGF-like #12	+	+	+	?	?	?	39
45	26	3217	1073	GAA	AAA	G->A	E1073K (1)	Glu	Lys	cb EGF-like #12	+	?	+	?	?	?	13
68	26	3217	1073	GAA	AAA	G->A	E1073K (2)	Glu	Lys	cb EGF-like #12	+	+	+	?	?	?	39
69	26	3217	1073	GAA	AAA	G->A	E1073K (3)	Glu	Lys	cb EGF-like #12	+	-	+	?	?	?	39
17	26	3220	1074	TGC	CGC	T->C	C1074R	Cys	Arg	cb EGF-like #12	+	+	+	-	+	-	24
71	26	3258	1086	TGT	TGG	T->G	C1086W	Cys	Trp	cb EGF-like #12	?	?	?	?	?	?	40
65	27	3349	1117	TGT	CGT	T->C	C1117G	Cys	Arg	cb EGF-like #13	+	+	+	?	?	?	39
28	27	3350	1117	TGT	TAT	G->A	C1117Y (1)	Cys	Tyr	cb EGF-like #13	+	+	+	?	?	?	26
106	27	3350	1117	TGT	TAT	G->A	C1117Y (2)	Cys	Tyr	cb EGF-like #13	+	+	+	-	-	-	NP1
83	27	3391	1131	AAC	TAC	A->T	N1131Y	Asn	Tyr	cb EGF-like #13	?	?	?	?	?	?	47
1	27	3410	1137	CGC	CCC	G->C	R1137P (1)	Arg	Pro	cb EGF-like #13	+	+	+	?	?	?	16
2	27	3410	1137	CGC	CCC	G->C	R1137P (2)	Arg	Pro	cb EGF-like #13	+	+	+	?	?	?	16
88	27	3458	1153	TGT	TAT	G->A	C1153Y	Cys	Tyr	cb EGF-like #13	+	+	+	-	+	+	48
79	28	3463	1155	GAC	AAC	G->A	D1155N	Asp	Asn	cb EGF-like #14	+	-	+	-	+	-	46
29	28	3464	1155	GAC	del17b	Stop at 1186	3464del17	Asp	Frameshift	cb EGF-like #14	+	-	+	?	?	?	26
97	28	3497	1166	TGC	TAC	G->A	C1166Y	Cys	Tyr	cb EGF-like #14	+	+	+	-	-	-	NP7
58	28	3509	1170	CGT	CAT	G->A	R1170H (1)	Arg	His	cb EGF-like #14	+	-	-	-	-	-	36
142	28	3509	1170	CGT	CAT	G->A	R1170H (2)	Arg	His	cb EGF-like #14	?	?	?	?	?	?	NP2
113	28	3513	1171	TGC	TGG	C->G	C1171W	Cys	Trp	cb EGF-like #14	+	+	+	-	-	-	NP3
114	28	3519	1173	AAC	AAG	C->G	N1173K	Asn	Lys	cb EGF-like #14	+	+	-	-	-	-	NP3
147	28	3524	1175	ATA	ACA	T->C	I1175T	Ile	Thr	cb EGF-like #14	+	-	+	-	+	-	NP2
96	28	3545	1182	TGC	TCC	G->C	C1182S	Cys	Ser	cb EGF-like #14	+	+	+	-	+	-	NP7
125	29	3599	1200	GAA	GGA	A->G	E1200G	Glu	Gly	cb EGF-like #15	+	+	+	-	-	-	NP8
137	29	3623	1208	TGT	del1b	Stop at 1229	3623 del G	Cys	Frameshift	cb EGF-like #15	+	+	+	-	-	-	NP4
39	29	3668	1223	TGT	TAT	G->A	C1223Y (1)	Cys	Tyr	cb EGF-like #15	+	+	+	?	?	?	34
73	29	3668	1223	TGT	TAT	G->A	C1223Y (2)	Cys	Tyr	cb EGF-like #15	+	+	+	?	?	+	42
21	30	3725	1242	TGT	TAT	G->A	C1242Y	Cys	Tyr	cb EGF-like #16	+	+	+	-	-	-	24

Table 1. *continued*

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
3	30	3746	1249	TGT	TCT	G->C	C1249S	Cys	Ser	cb EGF-like #16	+	+	+	?	?	?	17
91	31	3839	1280	GAT	del126b	del	3839 -1 G->T	Asp	Frameshift	cb EGF-like #17	+	?	+	?	?	?	41
35	32	3965	1322	GAC	del123b	del	3965 -2 A->T	Asp	Frameshift	cb EGF-like #18	+	-	+	?	?	?	31
36	32	3965	1322	GAC	del123b	del	4087 +1 G->A	Asp	Frameshift	cb EGF-like #18	+	+	+	?	?	?	31
151	32	4009	1337	GCT	CCT	G->C	A1337P	Ala	Pro	cb EGF-like #18	+	+	+	-	-	-	NP4
103	32	4011	1337	GCT	del1c	Stop at 1412	4011delT	Ala	Frameshift	cb EGF-like #18	+	-	+	-	-	-	NP7
52	32	4020	1340	ACC	del1c	Stop at 1412	4020delC	Thr	Frameshift	cb EGF-like #18	?	?	?	?	?	?	13
46	33	4145	1382	AAT	AGT	A->G	N1382S	Asn	Ser	cb EGF-like #19	?	?	?	?	?	?	13
115	34	4210	1404	GAC	TAC	G->T	D1404Y	Asp	Tyr	cb EGF-like #20	+	+	+	-	-	-	NP3
107	34	4270	1424	CCA	GCA	C->G	P1424A	Pro	Ala	cb EGF-like #20	+	+	-	-	-	-	NP1
130	34	4285	1429	TGT	AGT	T->A	C1429S	Cys	Ser	cb EGF-like #20	+	+	+	-	+	-	NP4
98	36	4490	1497	TGC	TCC	G->C	C1497S	Cys	Ser	cb EGF-like #22	+	+	+	-	-	-	NP7
22	36	4537	1513	TGC	CGC	T->C	C1513R	Cys	Arg	cb EGF-like #22	+	+	+	?	?	?	24
30	38	4766	1589	TGT	TTT	G->T	C1589F	Cys	Phe	8-Cys #4	+	+	+	?	?	?	26
116	39	4828	1610	TGC	GGC	T->G	C1610G	Cys	Gly	cb EGF-like #23	+	+	+	-	-	-	NP3
49	39	4857	1619	GGA	del1c	Stop at 1639	4857delA	Gly	Frameshift	cb EGF-like #23	+	+	+	?	+	?	13
4	40	4987	1663	TGT	CGT	T->C	C1663R	Cys	Arg	cb EGF-like #24	+	+	+	?	?	?	17
9	41	5137	1713	AAC	ins4a	Stop at 1735	5137ins4	Asn	Frameshift	8-Cys #5	+	+	+	?	+	?	19
148	41	5162	1721	TGT	TAT	G->A	C1721Y	Cys	Tyr	8-Cys #5	+	+	+	-	+	-	NP2
100	44	5453	1818	TGC	TAC	G->A	C1818Y	Cys	Tyr	cb EGF-like #26	+	-	+	-	+	-	NP7
99	44	5467	1823	GAA	TAA	G->T	E1823X	Glu	Stop	cb EGF-like #26	+	+	+	-	-	-	NP7
143	44	5494	1832	CGC	TGC	C->T	R1832C	Arg	Cys	cb EGF-like #26	?	?	?	?	?	?	NP2
80	44	5509	1837	CCC	TCC	C->T	P1837S	Pro	Ser	cb EGF-like #26	+	-	+	-	+	-	46
92	46	5672	1891	GAC	del117b	del	5788 +5 G->A	Asp	Frameshift	cb EGF-like #28	+	+	+	?	?	?	41
117	46	5679	1893	AAT	AAA	T->A	N1893K	Asn	Lys	cb EGF-like #28	+	+	+	-	-	-	NP3
101	46	5729	1910	GGT	GTT	G->T	G1910V	Gly	Val	cb EGF-like #28	+	+	+	-	+	-	NP7
47	46	5782	1928	TGT	CGT	T->C	C1928R	Cys	Arg	cb EGF-like #28	?	?	?	?	?	?	13
53	47	5789	1930	GAT	del129b	del	5788 +5 G->A (1)	Asp	Frameshift	cb EGF-like #29	?	?	?	?	?	?	13
54	47	5789	1930	GAT	del129b	del	5788 +5 G->A (2)	Asp	Frameshift	cb EGF-like #29	?	?	?	?	?	?	13
55	47	5789	1930	GAT	del129b	del	5788 +5 G->A (3)	Asp	Frameshift	cb EGF-like #29	?	?	?	?	?	?	13
138	47	5898	1966	CCA	del1c	Stop at 1979	5898delA	Pro	Frameshift	cb EGF-like #29	+	-	+	-	+	-	NP4
136	48	5930	1977	TGT	TAT	G->A	C1977Y	Cys	Tyr	cb EGF-like #30	+	+	+	-	-	-	NP4
144	48	5989	1997	AGA	TGA	A->T	R1997X	Arg	Stop	cb EGF-like #30	+	-	+	-	+	-	NP2
93	50	6164	2055	GAT	del126b	del	6163 +2 del16pb	Asp	Frameshift	8-Cys #6	+	+	+	?	?	?	41
118	50	6297	2099	TGC	TGG	C->G	C2099W	Cys	Trp	8-Cys #6	+	+	+	-	-	-	NP3
145	51	6314	2105	GAG	del66b	del	6314del66	Glu	Frameshift	8-Cys #6	+	+	+	-	+	-	49, NP5
119	51	6332	2111	TGT	TAT	G->A	C2111Y	Cys	Tyr	8-Cys #6	+	+	+	-	-	-	NP3
57	51	6339	2113	TAT	TAA	T->A	Y2113X	Tyr	Stop	8-Cys #6	?	?	?	?	?	?	35
23	52	6381	2127	GAT	GAA	T->A	D2127E	Asp	Glu	cb EGF-like #32	+	-	+	-	-	-	24
14	52	6431	2144	AAT	AGT	A->G	N2144S	Asn	Ser	cb EGF-like #32	+	-	+	?	?	?	22
24	52	6453	2151	TGC	TGG	C->G	C2151W	Cys	Trp	cb EGF-like #32	+	+	+	?	?	?	24
11	54	6617	2206	GAT	del123b	del	6739 +1 G->C	Asp	Frameshift	cb EGF-like #34	+	+	+	+	?	?	20
7	54	6662	2221	TGT	TCT	G->C	C2221S	Cys	Ser	cb EGF-like #34	+	+	+	?	?	?	17
120	55	6772	2258	TGT	CGT	T->C	C2258R	Cys	Arg	cb EGF-like #35	+	+	+	-	-	-	NP3
48	55	6784	2262	CAA	TAA	C->T	Q2262X	Gln	Stop	cb EGF-like #35	?	?	?	?	?	?	13
121	55	6844	2282	CGG	TGG	C->T	R2282W (1)	Arg	Trp	cb EGF-like #35	+	+	+	-	-	-	NP3
132	55	6844	2282	CGG	TGG	C->T	R2282W (2)	Arg	Trp	cb EGF-like #35	+	+	+	-	+	-	NP4
5	56	6920	2307	TGC	TCC	G->C	C2307S	Cys	Ser	cb EGF-like #36	+	+	+	?	?	?	18
94	57	6998	2333	GAC	del126b	del	6997 +1 G->A	Asp	Frameshift	8-Cys #7	+	+	+	?	?	?	41
95	58	7205	2402	GAT	del126b	del	7205 -2 A->G	Asp	Frameshift	cb EGF-like #37	+	+	+	?	?	?	41
31	59	7339	2447	GAG	AAG	G->A	E2447K	Glu	Lys	cb EGF-like #38	+	+	-	?	+	?	27
15	60	7456	2486	CTT	del366a	del	7456del366	Leu	Frameshift	cb EGF-like #39	+	+	+	-	-	-	23
122	60	7465	2489	TGT	CGT	T->C	C2489R	Cys	Arg	cb EGF-like #39	+	+	+	-	-	-	NP3
126	60	7531	2511	TGT	CGT	T->C	C2511R	Cys	Arg	cb EGF-like #39	+	+	+	?	?	?	24
108	63	7868	2623	CAC	CCC	A->C	H2623P	His	Pro	cb EGF-like #42	+	+	+	-	+	-	37
32	63	7879	2627	GGG	AGG	G->A	G2627R	Gly	Arg	cb EGF-like #42	+	+	+	?	?	?	28
59	63	8038	2680	CGC	TGC	C->T	R2680C	Arg	Cys	cb EGF-like #43	+	+	-	+	-	-	37
60	64	8052	2684	GGG	del175c	Stop at 2693	8052 -2 A->G	Gly	Frameshift	cb EGF-like #43	+	+	+	-	-	-	37
61	64	8052	2684	GGG	del175c	Stop at 2693	8051 +5 G->A	Gly	Frameshift	cb EGF-like #43	+	+	+	-	-	-	37
123	64	8176	2726	CGG	TGG	C->T	R2726W	Arg	Trp	COOH unique reg.	+	-	-	-	-	-	30
50	65	8236	2746	GAG	del2a	Stop at 2758	8236delGA	Glu	Frameshift	COOH unique reg.	+	+	+	?	+	?	13
16	65	8268	2756	TGG	TGA	G->A	W2756X	Trp	Stop	COOH unique reg.	+	+	+	?	?	?	23
38	65	8326	2776	CGA	TGA	C->T	R2776X	Arg	Stop	COOH unique reg.	+	+	+	-	-	-	33

The columns contain the following information and abbreviations:

A: Report number.

B: Exon number at which the mutation is located.

C: Nucleotide position at which the mutation is located, numbered with respect to the FBN1 gene cDNA sequence obtained from GenBank (GenBank accession number L13923; complete coding sequence of HUM-FIBRILLIN *Homo sapiens* fibrillin mRNA).

D: Codon number at which the mutation is located.

E: Normal base sequence of the codon in which the mutation occured.

F: Mutated base sequence of the codon in which the mutation occured.

G: Concerns base substitutions. It gives the base change, by convention, read from the coding strand. If the mutation predicts a premature protein-termination, the novel stop codon position is given, e.g. Stop at 2115.

H: Mutation name according to Beaudet *et al*. (51).
I: Wild type amino acid.
J: Mutant amino acid. Deletion and insertion mutations which result in a frameshift are designated by Frameshift. Nonsense mutations are designated by Stop.
K: Protein domain in which the mutation occured. Each module group is numbered separately and according to the position of the module with respect to the N-terminal end of the protein, e.g. cb EGF-like (for calcium-binding EGF-like modules) #1–43, EGF-like (for non calcium-binding EGF-like modules) #1–4, 8-cys (for 8-cysteine modules) #1–7, Hybrid modules #1–2 (6–8).
L–Q: Diagnostic manifestations in the systems entered with respect to the nosology proposals of Beighton *et al*. (52) recently revised by de Paepe *et al*. (53). In all these columns, '?' indicates either lack of or unspecified data until more precise information is available.
L: Presence (+) or absence (–) of skeletal manifestations.
M: Presence (+) or absence (–) of ocular manifestations.
N: Presence (+) or absence (–) of cardiovascular manifestations.
O: Presence (+) or absence (–) of pulmonary manifestations.
P: Presence (+) or absence (–) of manifestations in skin and integument.
Q: Presence (+) or absence (–) of manifestations in central nervous system.
R: Reference number indicating the publication in which the mutation is described. NP indicates unpublished mutations contributed by NP1 (Lesley Ades and Katherine J. Holman), NP2 (M. Boxer and C. Black), NP3 (Caroline Hayward and David J. H. Brock), NP4 (Anne de Paepe and Lieve Nuytinck), NP5 (Uta Francke and Wanguo Liu), NP6 (Ulrich Grau and Hanns-Georg Klein), NP7 (Ana Beatriz Alvarez Perez) (54) and NP8 (Leena Peltonen and Terhi Rantamäki).

THE MARFAN DATABASE

The mutations file of the database lists point mutations, deletions or insertions, and splice mutations in the FBN1 gene (Table 1). It contains in a standarized, easily accessible and summary form the molecular and the clinical data on the causative mutations of Marfan syndrome and type 1 fibrillinopathies. For each mutation, information is provided at several levels: at the gene level (exon and codon number, wild type and mutant codon, mutational event, mutation name), at the protein level (wild type and mutant amino acid, affected domain) and at the clinical level (absence or presence of skeletal, ocular, cardiovascular, central nervous system and other various manifestations). The present version of the database contains 137 entries corresponding to mutations either recently published or only reported in meeting proceedings or contributed by the co-authors of this paper. It is not intented to replace primary publications, although it does contain unpublished data. Forty eight new entries appear, as compared to the last update.

ANALYSIS AND LIMIT OF THE DATABASE

The global molecular analysis of the mutations file reveals that nonsense mutations (7/137), splicing errors (18/137) and small deletions (12/137) predicted to result in truncated fibrillin-1 molecules have been identified but represent a small proportion of the mutations. Insertions are surprisingly under-represented (2/137). The majority of mutations identified are missense mutations (99/137 or 72.3%) affecting primarily (73/99) the numerous calcium binding (EGF)-like modules found throughout the protein. The fibrillin gene has been identified and sequenced in two mammalian species. The identity at the amino acid level is so high (97.8% human-mouse and 96.2% human-bovine) that very often phylogenic conservation should be observed at the amino acid position affected by a given missense mutation. In effect, in all the mutations thus far reported in the FBN1 gene, the mutational event affects a conserved amino acid with respect to the mouse and bovine sequences. It is interesting to note that the only exception is the Y2113X mutation (which affects a non-conserved amino acid) leads to a truncated protein. The mutations file contains 11 recurrent mutations that have been reported either twice (R122C, R545C, R627C, C1117Y, R1137P, R1170H, C1223Y and R2282W) or thrice (G1013R, E1073K and 5788+5G→A). Until haplotype analysis is available it is unclear whether these are truly recurrent mutations or if they are carried by the same chromosome.

There is an excess of mutations in exons 25, 27 and 28 ($P < 0.001$) when comparing observed to expected mutations. This clustering is explained by the fact that almost all the mutations identified in neonatal cases of MFS1 are located within this area.

Since the present version of the software cannot accomodate two mutational events in a given individual, three mutations are not included in the current version of the mutations file: the double mutant Splice exon 51 and X2113X reported by Dietz *et al*. (12), the compound deletion del3901-4; 3908-9 reported by Nijbroek *et al*. (13), and the double mutant I1071S and E1073D reported by Wang *et al*. (14). Three other mutations are not included in the Marfan database [1588+21G→A (15), 2294–1G→C and 1837+5G→A (NP1)] until more precise information is available on mRNA splicing or stability. All things considered, 143 mutations have been described to date in the FBN1 gene.

NEW WEB VERSION OF THE SOFTWARE AND ITS NEWLY DEVELOPED ROUTINES

Four new routines now appear in the Marfan database as follows. (i) Amino acid changes: lists for each of the 20 amino acids the observed substitutions throughout the protein. (ii) Base modification: lists the observed mutations with respect to their position within the codon for each of the four bases. (iii) CpG: studies the distribution of mutations occuring at CpG sites throughout the coding sequence. The result is displayed in a graphic representation. (iv) Distribution of mutation: lists the proportion of each of the mutational events observed in a selected group of mutation records.

The software is now accessible through the World Wide Web and analyses with the various routines can be performed by users on-line. The investigation of genotype/phenotype correlations with these tools is currently difficult since clinical data is often sparse in mutation records. To facilitate the input of high quality clinical data, we are currently developing a Mutation Report Entry in the web site.

DATABASE UPDATE, SOFTWARE AVAILABILITY AND ONLINE ANALYSIS

The current database and subsequent updated versions are available on request to G.C-B. or C.Boileau on floppy disc using Apple format and Microsoft Excel®, or by Email (collod@ceylan.necker.fr). Notification of omissions and errors in the current version as well as specific phenotypic data would be gratefully received by the corresponding author. The software package is available on a collaborative basis. The software will be expanded as the database grows and according to the requirements of its users. New functions could be implemented. New web version of the Marfan Database permitting on-line analysis is accessible at http://www.umd.necker.fr . Users of the database must cite this article.

ACKNOWLEDGEMENTS

This work was supported by grants from Recherche Clinique-Assistance Publique Hôpitaux de Paris (Grant CRC940116), AFM (Association Française contre les Myopathies), Université René Descartes-Paris V, Ministère de l'Education Nationale, de l'Enseignement Supérieur, de la Recherche et de l'Insertion Professionnelle (ACC-SV2) and Faculté de Médecine Necker. G.C.-B. is supported by a grant from AFM (Association Française contre les Myopathies). L.N. and A.d.P. are supported by grants from the fund for Scientific Research Flanders. A.B.A.P. is supported by a grant from FAPESP (Brazil).

REFERENCES

1 Sakai, L., Keene, D.R. and Engvall, E. (1986) *J. Cell Biol.*, **103**, 2499–2509.
2 Cleary, E.G. and Gibson, M.A. (1983) *Int. Rev. Connect. Tissue Res.*, **10**, 97–209.
3 Mecham, R.P. and Heusar, J.E. (1991) The elastic fiber. In Hay,E.D. (ed.) *Cell Biology of the Extracellular Matrix*, second edn. Plenum Publishing Co., New York, pp 79–109.
4 Ramirez, F., Pereira, L., Zhang, H. and Lee, B. (1993) *BioEssays*, **15**, 589–594.
5 Magenis, R.E., Maslen, C.L., Smith, L., Allen, L. and Sakai, L. (1991) *Genomics*, **11**, 346–351.
6 Malsen, C.L., Corson, G.M., Maddox, B.K., Glanville, R.W. and Sakai, L. (1991) *Nature*, **352**, 334–337.
7 Corson, G.M., Chalberg, S.C., Dietz, H.C., Charbonneau, N.L. and Sakai, L.S. (1993) *Genomics*, **17**, 476–484.
8 Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J.R., Sykes, B., Pangilinan, T. and Bonadio, J. (1993) *Hum. Mol. Genet.*, **2**, 961–968.
9 Pyeritz, R.E. (1993) In Royce,P.M. and Steinmann,B. (eds) *Molecular, Genetic and Medical Aspects.* Wiley-Liss, New York, pp 437–468.
10 Collod, G., Broud, C., Soussi, T., Junien, C., Boileau, C. (1996) *Nucleic Acids Res.*, **24**, 137–140.
11 Collod-Broud, G., Broud, C., Ads, L., Black, C., Boxer, M., Brock, D.J., Godfrey, M., Hayward, C., Karttunen, L., Milewicz, D., Peltonen, L., Richards, R.I., Wang, M., Junien, C., Boileau, C. (1997) *Nucleic Acids Res.*, **25**, 147–150.
12 Dietz, H., Valle, D., Francomano, C., Kendzior, R.J., Pyeritz, R. and Cutting, R. (1993) *Science*, **259**, 680–683.
13 Nijbroek, G., Sood, S., McIntosh, I., Francomano, C.A., Bull, E., Pereira, L., Ramirez, F., Pyeritz, R.E. and Dietz, H.C. (1995) *Am. J. Hum. Genet.*, **57**, 8–21.
14 Wang, M., Kishnani, P., Decker-Phillips, M., Kahler, S.G., Chen, Y.T. and Godfrey, M. (1996) *J. Med. Genet.*, **33**, 1–4.
15 Grau, U., Mair, H., Detter, C., Seidel, D., Reichart, B. and Klein, H.G. (1996) *Eur. J. Pediat.* **155**, 739 (Abstract 43 P).
16 Dietz, H.C., Cutting, G.R., Pyeritz, R.E., Malsen, C.L., Sakai, L.Y., Corson, G.M., Puffenberg, E.G., Hamosh, A., Nanthakumar, E.J., Curristin, S.M., *et al*. (1991) *Nature*, **352**, 337–339.
17 Dietz, H.C., Saraiva, J.M., Pyeritz, R.E., Cutting, G.R. and Francomano, C.A. (1992) *Hum. Mut.*, **1**, 366–374.
18 Dietz, H.C., Pyeritz, R.E., Puffenberger, E.G., Kendzior, R.J.J., Corson, G.M., Malsen, C.L., Sakai, L.Y., Francomano, C.A. and Cutting, G.R. (1992) *J. Clin.l Invest.*, **89**, 1674–1680.
19 Dietz, H.C., McIntosh, I., Sakai, L.Y., Corson, G.M., Chalberg, S.C., Pyeritz, R.E. and Francomano, C.A. (1993) *Genomics*, **17**, 468–475.
20 Godfrey, M., Vandemark, N., Wang, M., Velinov, M., Wargowski, D., Tsipouras, P., Han, J., Becker, J., Robertson, W., Droste, S. and Rao, V.H. (1993) *Am. J. Hum. Genet.*, **53**, 472–480.
21 Hayward, C., Rae, A., Porteous, M., Logie, L. and Brock, D. (1994) *Hum. Mol. Genet.*, **3**, 373–375.
22 Hewett, D., Lynch, J., and Sykes, B. (1993) *Hum. Mol. Genet.*, **2**, 475–477.
23 Kainulainen, K., Sakai, L.Y., Child, A., Pope, F.M., Puhakka, L., Ryhanen, L., Palotie, A., Kaitila, I. and Peltonen, L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5917–5921.
24 Kainulainen, K., Karttunen, L., Puhakka, L., Sakai, L. and Peltonen, L. (1994) *Nature Genet.*, **6**, 64–69.
25 Piersall, L., Dietz, H., Hall, B., Caddle, R., Pyeritz, R., Francomano, C. and McIntosh, I. (1994) *Hum. Mol. Genet.*, **3**, 1013–1014.
26 Tynan, K., Comeau, K., Pearson, M., Wilgenbus, P., Levitt, D., Gasner, C., Berg, M., Miller, D. and Francke, U. (1993) *Hum. Mol. Genet.*, **2**, 1813–1821.
27 Lönnqvist, L., Child, A., Kainulainen, K., Davidson, R., Puhakka, L. and Peltonen, L. (1994) *Genomics*, **19**, 573–576.
28 Karttunen, L., Raghunath, M., Lönnqvist, L. and Peltonen, L. (1994) *Am. J. Hum. Genet.*, **55**, 1083–1091.
29 Milewicz, D.M. and Duvic, M. (1994) *Am. J. Hum. Genet.*, **54**, 447–453.
30 Milewicz, D.M., Grossfield, J., Cao, S-N., Kielty, C., Covitz, W. and Jewett, T. (1995) *J. Clin. Invest.*, **95**, 2373–2378.
31 Wang, M., Price, C.E., Han, J., Cisler, J., Imaizumi, K., Van Thienen, M-N., DePaepe, A. and Godfrey, M. (1995) *Hum. Mol. Genet.*, **4**, 607–613.
32 Stahl-Hallengren, C., Ukkonen, T., Kainulainen, K., Kristofersson, U., Saxne, T., Tornqvist, K. and Peltonen, L. (1994) *J. Clin. Invest.*, **94**, 709–713.
33 Hayward, C., Porteous, M., Brock, D. (1994) *Hum. Mut.*, **3**, 159–162.
34 Hewett, D.R., Lynch, J.R., Child, A., and Sykes, B.C. (1994) *J. Med. Genet.*, **31**, 338–339.
35 McInnes, R.R. and Byers, P.H. (1993) *Curr. Opin. Genet. Dev.*, **3**, 475–483.
36 Hayward, C., Porteous, M.E.M. and Brock, D.J.H. (1994) *Mol. Cell. Probes*, **8**, 325–327.
37 Grossfield *et al*. (1993) *Am. J. Hum. Genet.* **53**, abstract 1167.
38 Tilsra, D.J. and Byers, P.H. (1993) Scientific workshop of the Marfan syndrome, Oregon, abstract.
39 Putnam, E.A., Cho, M., Zinn, A.B., Towbin, J.A., Byers, P.H. and Milewicz, D.M. (1996) *Am. J. Med. Genet.*, **62**, 233–242.
40 Mathews, K., Wang, M., Corbit, C.K. and Godfrey, M. (1995) *Am. J. Hum. Genet.*, **57**, abstract 1966.
41 Liu, W., Qian, C., Comeau, K., Brenn, T., Furthmayr, H. and Francke U. (1996) *Hum. Mol. Genet.*, **5**, 1581–1587.
42 Sood, S., Eldadah, Z.A., Krauss, W.L., McIntosh, I. and Dietz, H.C. (1996) *Nature Genet.*, **12**, 209–211.
43 Quan, F., Sakai, L. and Popovich, B.W. (1995). *Am. J. Hum. Genet.*, **57**, abstract 1936.
44 Booms, P., Vetter, U. andRobinson, P.N. (1996) *Eur. J. Pediat.*, **155**, 739 (Abstract 42 P).
45 Lönqvist, L., Karttunen, L., Rantamäki, T., Kielty, C., Raghunath, M. and Peltonen, L. (1996) *Genomics*, **36**, 468–475.
46 Milewicz, D.M., Michael, K., Fisher, N., Coselli, J.S., Markello, T. and Biddinger, A. (1996) *Circulation*, **94**, 2708–2711.
47 Wang, M., Wang, J.Y., Cisler, J., Imaizumi, K., Burton, B.K., Jones, M.C., Lamberti, J.J. and Godfrey, M. (1997) *Hum. Mut.*, **9**, 359–362.
48 Ades, L.C., Haan, E.A., Colley, A.F. and Richard, R.I. (1996) *J. Med. Genet.*, **33**, 665–671.
49 Liu, W., Qian C. and Francke, U. (1997) *Nature Genet.*, **16**, 328–329.
50 Kielty, C., Rantamäki, T., Child, A., Shuttleworth, A. and Peltonen, L. (1995) *J. Cell Sci.*, **108**, 1317–1323.
51 Beaudet, A.L. and Tsui, L.-C. (1993) *Hum. Mut.*, **2**, 245–248.
52 Beighton, P., de Paepe, A., Danks, D., Finidori, G., Gedde-Dahl, T., Goodman, R., Hall, J.G., Hollister, D.W., Horton, W., McKusick, V.A., *et al*. (1988) *Am. J. Med. Genet.*, **29**, 581–594.
53 De Paepe, A., Devereux, R.B., Dietz, H.C., Hennekam, R.C.M. and Pyeritz, R.E. (1996) *Am. J. Hum. Genet.*, **62**, 417–426.
54 Perez, A.B., Pereira,L., Brunoni, D., Passos-Buenos, M.R., manuscript submitted.