Original articles

An exclusion map of Marfan syndrome

Susan H Blanton, Mansoor Sarfarazi, Hans Eiberg, Jacques de Groote, Peter A Farndon, Michael W Kilpatrick, Anne H Child, F Michael Pope, Leena Peltonen, Clair A Francomano, Catherine Boileau, Marion Keston, Petros Tsipouras

Abstract

The combined genetic data between the Marfan syndrome and 75 informative loci on 18 autosomes were used to construct an exclusion map for this disorder. Data are also presented for a further two unmapped markers. The most likely location of the Marfan syndrome gene is highlighted and all the unexcluded areas of the genome are displayed in a graphical form. This exclusion map shows that almost 75% of the genome has been excluded as a likely location for the Marfan syndrome gene in the majority of the families studied. Apart from chromosomes 8, 13, 21, and 22, for which no data were available, other regions not excluded yet include 5p, 6p, 9p, 10p, 12p, 15, 17p, 18, and 20p. Future linkage analysis using markers located in the highlighted regions should facilitate the identification of the site of the Marfan syndrome gene.

Department of Pediatrics, University of Connecticut Health Center, Farmington, CT 06032, USA.
S H Blanton*, M Sarfarazi, P Tsipouras

Institute of Medical Genetics, University of Copenhagen, Denmark.
H Eiberg

Clinical Genetics Unit, Birmingham Maternity Hospital, Birmingham.
J de Groote, P A Farndon, M W Kilpatrick

Department of Cardiological Sciences, St George's Hospital Medical School, London.
J de Groote, A H Child

Clinical Research Centre, Harrow, Middlesex.
J de Groote, F M Pope

Laboratory of Molecular Genetics, National Public Health Institute, Helsinki, Finland.
L Peltonen

Division of Medical Genetics, Johns Hopkins Medical School, Baltimore, MD, USA.
C A Francomano

Unité 73, INSERM, Paris, France.
C Boileau

Clinical Genetics Service, Western General Hospital, Edinburgh.
M Keston

Correspondence to Dr Sarfarazi.

*Present address: Graduate School of Biomedical Sciences, University of Texas Health Sciences Center, Houston, TX 77225, USA.

Received for publication 16 August 1989.
Revised version accepted for publication 26 October 1989.

Marfan syndrome (MS) is a serious heritable disorder of connective tissue associated with significant morbidity and mortality.[1] The mapping of the gene for MS will help significantly in our eventual understanding of its pathogenesis and, more importantly, it will be essential in the design of a rational therapy. Since 1976 a number of reports on linkage studies in MS have appeared, none of which showed linkage of MS to a marker.

More recently (June 1989) various groups involved in linkage analysis in MS were invited to participate in the first International Genetic Linkage Workshop on Marfan syndrome. At the meeting it was agreed that a significant advance towards the mapping of the gene would be made by the pooling of all linkage data for MS to date. In particular, this would avoid unnecessary duplication of negative results and highlight those areas of the genome not yet studied.

As a result of this meeting a consortium of those involved with MS was formed. This paper presents a more detailed exclusion map for MS based on previously published reports, and on the data provided by other members of the consortium; some of the individual data sets are given in full in the other papers in this issue of the Journal.

Source of linkage data

The linkage data presented here were compiled either during the first International Workshop by the collaborating consortium groups or were subsequently sent to us. In addition previously published data were incorporated. The pooled data on MS consist at present of 77 markers on 18 autosomes. Seventy-five of these markers have been localised with various degrees of confidence and the remaining two (APOH and C6) are still unmapped. These two markers have been excluded from our analysis. Three of the groups

Table 1 The overall linkage data for Marfan syndrome. The first five columns of data were used as input to the EXCLUDE program.

Position on the chromosome	Locus name	Data type	θ (cM)	Lod score (Z)	r	s	Chromosome localisation	Reference
1.03	*D1S2*	1	0·20	−4·72	24·4	48·7	1pter–p31	5
1.06	*D1S47*	1	0·10	−4·64	10·5	20·9	1p35–p22	5
1.12	*PND*	1	0·10	−2·85	6·4	12·8	1p36	5
1.13	*D1S7*	1	0·10	−1·83	4·1	8·2	1p35–p33	6
1.19	*Rh*	1	0·10	−4·29	9·7	19·3	1p36.2–p34	5
1.23	*D1S57*	1	0·20	−3·47	17·9	35·8	1pter–p31	5
1.26	*GLUT*	1	0·10	−0·60	1·4	2·7	1p35–p31	5
1.27	*D1S44*	1	0·10	−0·78	1·8	3·5	1p35–p31.3	5
1.34	*PGM1*	1	0·10	−2·66	6·0	12·0	1p22.1	2
1.40	*D1S38*	1	0·10	−3·20	7·2	14·4	1p22–p13	5
1.41	*NGFB*	1	0·10	−1·95	4·4	8·8	1p22.1	6
1.58	*PUM*	1	0·10	−3·37	7·6	15·2	1q21–q24	6
1.60	*FY*	1	0·10	−4·81	10·8	21·7	1q22–q23	2
1.73	*D1S53*	1	0·10	−4·58	10·3	20·6	1q23–q25	5
1.74	*F13B*	1	0·20	0·28	0·7	3·3	1q32	*
1.95	*D1S8*	1	0·10	−0·73	1·6	3·3	1q42–q43	6
2.05	*ACP1*	1	0·10	−4·37	9·8	19·7	2p25 or 2p23	2
2.21	*D2S28*	1	0·10	−1·93	4·3	8·7	2p16–2p14	5
2.24	*D2S6*	1	0·10	−2·33	5·3	10·5	2q32–q36	5
2.46	*D2S21*	1	0·10	−4·52	10·2	20·4	2q11–q13	5
2.50	*D2S44*	1	0·10	−13·23	29·8	59·6	2q12–q32	5
2.64	*D2S39*	1	0·10	−4·95	11·2	22·3	2q22–q24	5
2.80	*COL3A1*	1	0·10	−10·35	23·3	46·7	2q31–q32.3	7 8†
2.80	*COL5A2*	1	0·10	0·84	0·5	5·3	2q31–q32.3	8
2.86	*ELN*	1	0·01	0·43	0·0	1·6	2q31–qter	†
2.90	*FN1*	1	0·30	0·12	1·0	3·4	2q34–q36	†
2.94	*D2S3*	1	0·10	−2·96	6·7	13·3	2q35–q37	5
2.97	*COL6A3*	1	0·10	−1·30	2·9	5·9	2q37	5†
3.15	*D3S17*	1	0·10	−0·12	0·3	0·5	3p24–p23	5
3.31	*D3S12*	1	0·10	−7·29	16·4	32·9	3p21.3–p21.1	5
3.50	*D3S13*	1	0·23	0·44	1·5	6·6	3q11–q12	5
3.66	*D3S14*	1	0·10	−3·38	7·6	15·2	3q21–q22	5
3.74	*TF*	1	0·10	−0·57	1·3	2·6	3q21	*
3.87	*CHE2*	1	0·10	−0·59	1·3	2·7	3q26	*
3.96	*AHSG*	1	0·10	−2·26	5·1	10·2	3q27–q29	*
4.02	*D4S107*	1	0·10	−4·02	9·1	18·1	4p16	5
4.23	*D4S103*	1	0·10	−2·29	5·2	10·3	4p13–p11	5
4.33	*GC*	1	0·10	−1·82	4·1	8·2	4q12–q13	2 9*
4.62	*D4S101*	1	0·10	−1·46	3·3	6·6	4p24–p27	5
4.76	*MNSs*	1	0·10	−5·69	12·8	25·6	4q28–q31	2 9
5.38	*D5S51*	1	0·10	−1·70	3·8	7·7	5q14–q21	5
6.15	*F13A*	1	0·30	0·21	1·8	5·9	6p24–p21.3	*
6.20	*HLA*	1	0·10	−1·33	3·0	6·0	6p21.3	2
6.21	*BF*	1	0·10	−0·30	0·7	1·4	6p21.3	*
6.22	*GLO*	1	0·10	−3·15	7·1	14·2	6p21.31–p21.1	2
6.96	*PLG*	1	0·30	0·20	1·7	5·6	6q26–q27	*
7.16	*D7S62*	1	0·10	−1·12	2·5	5·0	7pter–q22	5
7.55	*D7S15*	1	0·10	−2·35	5·3	10·6	7q21–q22	5
7.56	*PON*	1	0·10	−1·08	2·4	4·9	7q21–q22	*
7.59	*COL1A2*	1	0·10	−7·31	16·5	33·0	7q21.3–q22.1	5 7 8 †
7.95	*D7S54*	1	0·10	−8·58	19·3	38·7	7q35–qter	5
9.92	*ABO*	1	0·10	−2·95	6·6	13·3	9q34.1–q34.2	2 9
9.93	*AK1*	1	0·30	0·22	1·8	6·2	9q34.1–q34.2	2
9.95	*ORM*	1	0·10	−0·51	1·1	2·3	9q31–qter	*
10.80	*D10S12*	1	0·10	−10·35	23·3	46·7	10q24.1	5
11.07	*D11S48*	1	0·10	−2·02	4·6	9·1	11p13	6
11.24	*D11S134*	1	0·10	−4·88	11·0	22·0	11p14	5
11.47	*D11S141*	1	0·10	−4·60	10·4	20·7	11q13.5–q14.2	5
11.58	*D11S137*	1	0·10	−2·80	6·3	12·6	11q14.3–q22.3	5
11.63	*D11S1184*	1	0·10	−2·27	5·1	10·2	11q22	6
11.70	*D11S132*	1	0·10	−2·75	6·2	12·4	11q22–q23	5
12.51	*COL2A1*	1	0·10	−14·10	31·8	63·6	12q14.3	5 7 8
14.85	*D14S27*	1	0·10	−0·17	0·4	0·8	14q31	5
14.91	*PI*	1	0·10	−3·925	8·8	17·7	14q32.1	2 *
14.99	*GM*	1	0·10	−3·10	7·0	14·0	14q32.3	2

(*cont.*)

Table 1—cont.

Position on the chromosome	Locus name	Data type	θ (cM)	Lod score (Z)	r	s	Chromosome localisation	Reference
15.55	*D15S49*	1	0·10	0·71	0·4	4·4	15q15–q23	5
16.11	*PGP*	1	0·10	−1·23	2·8	5·5	16p13	*
16.80	*HP*	1	0·10	−1·97	4·4	8·9	16q22	9 *
17.57	*D17S37*	1	0·22	0·22	0·7	3·0	17q21	5
17.67	*COL1A1*	1	0·10	−3·77	8·5	17·0	17q21.3–q22	5 8
17.91	*D17S79*	1	0·35	0·87	15·4	43·9	17q24–q25	5
18.44	*Jk*	1	0·10	−1·80	4·1	8·1	18q11.1–q11.2	2 9
19.16	*C3*	1	0·20	0·07	0·2	0·8	19p13.3–13.2	*
19.73	*SE*	1	0·10	−2·81	6·3	12·7	19q12–q13	2 9
20.78	*ADA*	1	0·10	−0·82	1·8	3·7	20q13.2–qter	2
The following two markers are unassigned								
?.??	C6	1	0·20	0·30				*
?.??	APOH	1	0·30	0·28				*

r=equivalent number of recombinants.
s=equivalent number of informative meioses.
*Data from Dr Hans Eiberg (personal communication).
†Data from Dr Leena Peltonen (Kainalainen *et al. Hum Genet*, in press).

contributed linkage data with chromosome 1 markers, in addition to those previously reported by Mace.[2] The two other well studied chromosomes were 2 and 3.

We used the program EXCLUDE[3 4] on the combined linkage data to construct the exclusion map for Marfan syndrome. The input to the program consists of chromosomal position, locus name, data type, and the linkage data in the form of recombination fraction and the corresponding lod score (first 5 columns of table 1). The chromosomal position in table 1 is defined by the chromosomal number, a full stop, and the location of the marker on that chromosome expressed as a percentage of pter to qter. When the localisation of a marker is expanded over a region of a chromosome, the site of that marker is taken to be in the middle of that region. If the lod score was uniformly negative for all values of θ, the lod score (Z) at θ=0·10 was used. When Z was positive the value of Z_{max} and the corresponding θ_{max} were used.

As the EXCLUDE program reads the input data for each marker on a separate line, it calculates the equivalent number of informative meioses (s), number of recombinants (r), and the ratio r/s (table 1). The program then finds the peak likelihood distribution of each position, either directly from

Table 2 Results from program EXCLUDE, showing the most likely location of the MS gene.

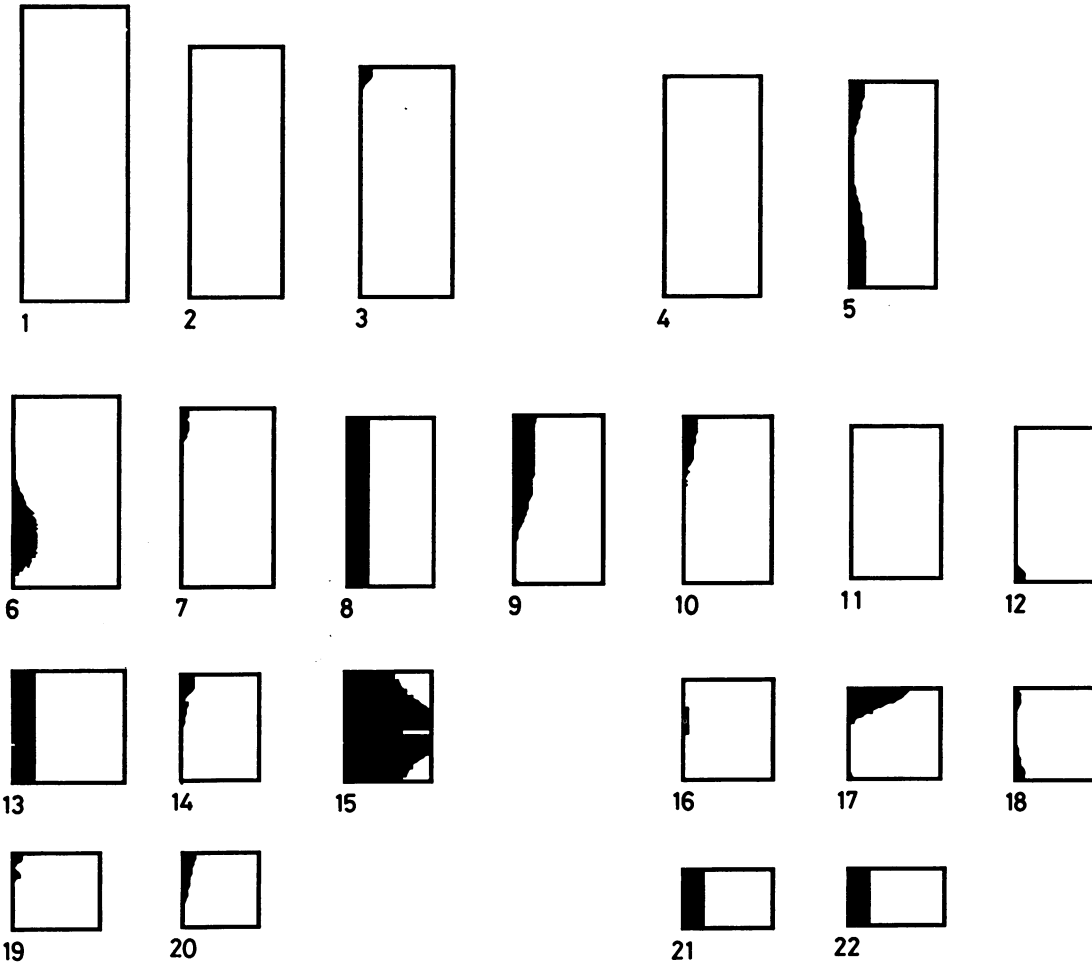
Chromosome		No of loci	Chromosomal probability		Maximum likelihood for MS location
No	Length (cM)		Relative (%)	Maximum (%)	
1	250	16	8·67	0·07	1·00
2	225	12	7·80	0·00	1·00
3	200	7	6·93	0·30	1·60
4	180	5	6·24	0·04	0·99
5	180	1	6·24	7·22	0·96
6	170	5	5·89	7·20	1·62
7	160	5	5·55	0·12	0·99
9	150	3	5·20	8·25	1·16
10	150	1	5·20	2·64	0·82
11	140	6	4·85	0·02	0·98
12	140	1	4·85	0·21	0·15
14	100	3	3·47	1·53	0·99
15	100	1	3·47	34·53	5·13
16	90	2	3·12	0·24	0·89
17	80	3	2·77	4·49	3·67
18	80	1	2·77	0·94	0·48
19	70	2	2·43	0·54	1·18
20	70	1	2·43	2·17	0·80

Maximum chromosomal probability: 34·53%.
Maximum likelihood for MS location: 5·13 (lod=0·71).

these equivalents, or by parabolic interpolation of the Z scores. As the direct count of recombinants and non-recombinants provides the optimal way of estimating the maximum likelihood and has the advantage of additivity, the program converts all the likelihoods into these equivalents. The overall positional likelihoods are then combined by multiplication, and standardised either to display a unit area or to 'fill' the diagram (figure). The EXCLUDE program not only calculates the positional likelihood of the disease locus on each chromosome, but also the percentage of the probability of a locus being on any of the 22 autosomes (table 2). The last two columns in table 2 give the maximum values in likelihood and probability of the MS locus being on that chromosome. The figure is the graphical representation of these values, which shows the exact area of each chromosome being excluded (white areas) as a likely position of the MS gene.

Results and discussion

The first genetic linkage studies on Marfan syndrome were done using conventional protein and cell surface antigens as markers.[2 9] Mace[2] reported a maximum lod score of $Z=1{\cdot}25$ obtained at $\theta=0{\cdot}24$ between *Rh* and *MS*. The combined linkage data presented here, between *MS*, *Rh*, and markers in the proximity, clearly exclude genetic linkage to that part of chromosome 1. Since no linkage with this disorder could be established, it was not until recently that investigators undertook the candidate gene



The exclusion map for MS. The shaded area of the chromosomes shows the possible location of the disease gene and the white area the excluded zones.

approach.[7 10–13] However, these studies were all negative, effectively ruling out these loci as possible sites of mutations resulting in MS.

The additional data presented by the participating members of the MS consortium, as shown in this issue, provided the basis of constructing an exclusion map for the MS gene. Before this workshop, data were available on 15 markers only. It has now been expanded to 77 markers over a total of 18 chromosomes. None of the markers studied by the participating groups shows significant linkage with MS.

This information provided the means of excluding more areas of the genome as a likely location of the MS gene. There were no linkage data available for chromosomes 8, 13, 21, and 22. These chromosomes taken together represent approximately 12% of the genome. The exclusion map indicated that the MS gene might be located on chromosome 15, 9p, 5p, 6q, 17p, 10p, or 20p. However, the possibility of the MS gene being on chromosomes 8, 13, 21, or 22, on which we had no data, remains very strong. Although the exclusion map shows that the chance of the disease gene being on any other chromosome is very remote, it should be emphasised that this is only true assuming there is a single locus responsible for the MS gene.

This indication from the exclusion map that the MS gene is located in one of the above mentioned regions of the genome will give impetus for the use of markers specific for these areas. A search for new informative polymorphisms in these areas of the genome will be very important in the process of gene localisation. However, consideration should be given to the fact that MS might be heterogeneous, which might not only lead to spurious positive lod scores but spurious exclusions as well. Although the study of a number of large families makes this unlikely, a careful heterogeneity analysis must be carried out if negative results are obtained for the part of the genome which has not been studied yet.

The authors acknowledge Ms Robin C Schwartz for her contribution to this project.

1 Pyeritz RE, McKusick VA. The Marfan syndrome: diagnosis and management. *N Engl J Med* 1979;**300**:772–7.
2 Mace M. A suggestion of linkage between the Marfan syndrome and the rhesus blood group. *Clin Genet* 1979;**16**:96–102.
3 Sarfarazi M, Huson SM, Edwards JH. An exclusion map for von Recklinghausen neurofibromatosis. *J Med Genet* 1987;**24**: 515–20.
4 Edwards JH. Exclusion mapping. *J Med Genet* 1987;**24**:539–43.
5 Schwartz RC, Blanton SH, Hyde CA, *et al.* Linkage analysis in Marfan's syndrome. *J Med Genet* 1990;**27**:86–90.
6 de Groote J, Farndon PA, Kilpatrick MW, *et al.* Linkage data for Marfan's syndrome and markers on chromosomes 1 and 11. *J Med Genet* 1990;**27**:82–5.
7 Francomano CA, Streeten EA, Meyers DA, Pyeritz RE. Marfan syndrome: exclusion of genetic linkage to three major collagen genes. *Am J Med Genet* 1988;**29**:457–62.
8 Boileau C, Jondeau G, Bonaiti C, *et al.* Linkage analysis of five fibrillar collagen loci in a large French Marfan's syndrome family. *J Med Genet* 1990;**27**:78–81.
9 Schleutermann DA, Murdoch JL, Walker BA, *et al.* A linkage study of the Marfan syndrome. *Clin Genet* 1976;**10**:51–3.
10 Tsipouras P, Borresen AL, Bamforth S, Harper PS, Berg K. Marfan syndrome: exclusion of genetic linkage to the COL1A2 gene. *Clin Genet* 1986;**30**:428–32.
11 Dalgleish R, Hawkins JR, Keston M. Exclusion of the α2(I) and the α1(III) collagen genes as the mutant loci in a Marfan syndrome family. *J Med Genet* 1987;**24**:148–51.
12 Ogilvie DJ, Wordsworth BP, Priestley LM, *et al.* Segregation of all four major fibrillar collagen genes in the Marfan syndrome. *Am J Hum Genet* 1987;**41**:1071–82.
13 Huttunen K, Kaitila I, Savolainen A, Palotie A, Peltonen L. The linkage analysis with RFLP markers of elastin and type III collagen in Finnish Marfan families. *Am J Med Genet* 1989;**30**: 244A.