European Journal of Human Genetics (1999) 7, 377–385
© 1999 Stockton Press All rights reserved 1018–4813/99 $12.00
http://www.stockton-press.co.uk/ejhg

ARTICLE

Linkage analysis of multiple sclerosis with candidate region markers in Sardinian and Continental Italian families

Sandra D'Alfonso[1], Lorenza Nisticò[2], Patrizia Zavattari[1], Maria Giovanna Marrosu[3], Raffaele Murru[3], Marina Lai[3], Luca Massacesi[4], Clara Ballerini[4], Donella Gestri[4], Marco Salvetti[5], Giovanni Ristori[5], Roberto Bomprezzi[5], Maria Trojano[6], Maria Liguori[6], Domenico Gambi[7], Aldo Quattrone[8], Doriana Fruci[2], Francesco Cucca[9], Patricia Momigliano Richiardi[1] and Roberto Tosi[2]

[1]*Chair of Human Genetics, Department of Medical Sciences, University of Piemonte Orientale 'A. Avogadro', Novara*
[2]*Institute of Cell Biology, CNR, Rome*
[3]*Chair of Neurophysiopathology, Department of Neuroscience, University of Cagliari*
[4]*Department of Neurological and Psychiatric Sciences, University of Firenze*
[5]*Department of Neurological Sciences, University 'La Sapienza', Rome*
[6]*Department of Neurological and Psychiatric Sciences, University of Bari*
[7]*Department of Clinical Neurology, University 'G. D'Annunzio', Chieti*
[8]*Department of Medical Sciences, University of Catanzaro and Institute of Experimental Medicine and Biotechnology, CNR*
[9]*Institute of Clinics and Biology of Growing Age, University of Cagliari, Italy*

Previous genome screens in multiple sclerosis have shown some evidence of linkage in scattered chromosomal regions. Although in no case the evidence of each single study was compelling and although in general the linkage 'peaks' of the different studies did not coincide, some regions can be considered likely candidates for the presence of MS risk genes because of the clustering of MLS scores and homology with *eae loci*. We performed a linkage analysis of markers in these regions and of intragenic markers of some individual candidate genes (*HLA-DRB1, CTLA-4, IL9, APOE, BCL2, TNFR2*). For the first time, Southern European populations were targeted, namely Continental Italians and Sardinians. A total of 69 multiplex families were typed for 67 markers by a semi-automatic fluorescence-based assay. Results were analysed for linkage by two non-parametric tests: GENEHUNTER and SimIBD. In general, the linkage scores obtained were low, confirming the conclusion that no gene is playing a major role in the disease. However, some markers, in 2p11, 3q21.1, 7p15.2 and 22q13.1 stood out as promising since they showed higher scores with one or the other test. This stimulates further association analysis of a large number of simplex families from the same populations.

Keywords: multiple sclerosis; linkage analysis; multiplex families

Correspondence: Roberto Tosi, Istituto di Biologia Cellulare, Via Ramarini 32, 00016 Monterotondo Scalo, Rome, Italy. Tel: +39 06 90091201; Fax: +39 06 90091260; E-mail: rtosie@ibc.rm.cnr.it
Received 10 August 1998; revised 1 October 1998; accepted 6 November 1998

Introduction

Multiple sclerosis (MS) is a complex disease of the central nervous system (CNS) characterised by relapsing remitting or chronic progressive course and inflammatory demyelinating lesions of CNS white matter. In western countries MS represents, after trauma, the most frequent cause of disability among young adults.[1] A large amount of evidence indicates that an immune-mediated mechanism is responsible for damage to the nervous tissue.[2–4]

MS is caused by an interplay of environmental and genetic factors. Their relative weight can be evaluated by three approaches: population epidemiology, twin concordance and family aggregation. Epidemiological studies show that environmental factors, likely infectious agents, play a major role as demonstrated by change in MS risk consequent on migration from high to low risk areas and vice versa, and occasional 'epidemics' in small communities after contact with groups of individuals from high risk areas.[5] However, they also demonstrate the importance of genetic factors in that some ethnic groups maintain their relative resistance to MS even when they reside in areas where MS is common (eg Gypsies in Hungary, Blacks and Orientals in USA, Maoris in New Zealand, Lapps in Scandinavia).[6]

Twin studies clearly demonstrate the role of genetic factors since MZ concordance is substantially above DZ concordance (25–30% versus 3%) but also show the importance of the environment since the concordance level in MZ twins is well below 100%.

Remarkably, familial aggregation is largely controlled by genetic factors since the disease risk in relatives closely follows the degree of genetic similarity with the proband, according to polygenic inheritance model, and is proportional to their kinship coefficient also for second and third degree relatives who are less likely to have shared the same environmental exposure.[7] More direct evidence that familial aggregation is mostly or completely due to genetic factors comes from adoption[8] and half-sib[9] studies. The current interpretation is that although MS is heavily influenced by environmental factors, they are not family specific. Hence the study of multiple-case families is most likely to provide information on genetic factors. The rather high level of λs (20–40) compares favorably with that of type 1 diabetes ($\lambda s = 15$) in which linkage studies have been largely successful. Moreover, in MS the contribution of MHC is lower than in diabetes leaving more 'room' for the contribution of other genes.

Prompted by this promising background, a systematic search of MS genes was undertaken in several populations by whole genome linkage screens. The conclusions were that no predominant susceptibility gene is involved, that no locus (not even HLA) can be consistently shown to be MS linked in all screens and that the presence of minor loci (with $\lambda s < 1.5$) cannot be excluded in any sizeable portion of the genome.

Nevertheless, for some chromosome regions multiple indications of possible linkage have accumulated, particularly on 2p, 5p, 5q, 17q and 19q.[10,11] They are weak, in no case reaching a level of 'suggestive evidence'[12] and they do not necessarily concern the same marker in the different screens, but in some cases they coincide with regions of homology to *eae* genes in the mouse[13] and in the rat.[14] Thus these regions are a logical starting point for further analysis and replication attempts.

We performed a linkage analysis in Southern European populations with the following rationale. All studies have been so far done on Northern Caucasoids. However, linkage analysis is very sensitive to the disease gene frequency. For a gene of moderate to low effect 10–100 fold fewer sib pairs may be needed to show linkage in a population where the frequency of the disease allele is intermediate compared with one where it is either high or low.[15] Therefore the choice of the population can be critical. We experienced with IDDM12 in type 1 diabetes a precedent in which a genomic region, where weak or no evidence for linkage was reported in Northern Caucasoids, showed significant linkage in Southern Caucasoids.[16,17]

The present study includes a panel of multiplex families from Continental Italy and from Sardinia. Sardinians belong to a rather homogeneous and isolated population with a genetic background substantially different from that of other populations;[18] features that are relevant in view of subsequent association studies. Moreover, the incidence of MS in Sardinia is higher than in other Italian regions and close to Northern European incidence rates.[19,20]

Markers in candidate regions were selected on the basis of previous linkage data and included a few candidate genes, namely *HLA-DRB1*, *CTLA-4*, *IL9*, *APOE*, *BCL2* and *TNFR2*. *HLA* was chosen as the only locus that showed consistently genetic association and, less consistently, linkage to MS. *CTLA-4* has been shown to be associated with IDDM[16] and other autoimmune diseases[21] and, like *BCL2* and *TNFR2*, is involved in the regulation of lymphocyte activation. *IL9* is a representative of the *IL* gene cluster that

resides within an *eae*-linked region in the rat.[14] *APOE* is located in a region showing some evidence of linkage to MS and its variation is relevant in other neurological diseases.

Materials and Methods

Diagnostic Criteria

Index patients and affected siblings with a diagnosis of definite multiple sclerosis according to Poster *et al*,[22] were enrolled in the Hospitals of Bari, Cagliari, Catanzaro, Chieti, Florence and Rome universities. Each patient was submitted to clinical evaluation by a trained neurologist and to cerebro spinal fluid and magnetic resonance imaging analysis.

Families

Relatives of index cases and affected siblings were recruited in the same centres as part of a programme for collecting genomic material from Italian MS multiplex families sponsored by the Italian Multiple Sclerosis Foundation (FISM). Details are shown in Table 1. Enrolment of patients and relatives followed their informed consent.

Genotyping

DNA was extracted from peripheral blood by standard techniques. Samples were analysed for 67 microsatellites grouped in six pools. Preference was given to the markers studied in the previous genome screens. Most of them (49) were (CA)n repeats belonging to the Genethon map.[23] The remainder were included in the CHLC map.[24] Polymerase chain reaction (PCR) was performed in 10 μl using 40 ng of DNA, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5–2 mM MgCl2, 200 mM of each dNTP, 4 pm of each primer and 0.5 U of Taq DNA polymerase.

Forward primers were 5'-labelled with one of the four ABI (Forster City, CA, USA) fluorescence dyes (FAM, HEX, TET, Tamra). PCR reactions were performed in a 9600 Perkin Elmer (Norwalk, CT, USA) thermal cycler for 30 s at 95°C, 30 s at 52–57°C, 30 s at 72°C for 30 cycles, followed by a final extension step of 5 min.

PCR products of 11–13 microsatellites were pooled according to their non-overlapping fluorescence colour and size ranges. The pools were treated with T4 DNA polymerase (0.8 U in 3 μl at 37°C for 30 min). Two μl pool were added to 3 μl of loading buffer (72% formamide, 14% dextran blue, 14% GENESCAN-350 Tamra or GENESCAN-350 Rox (ABI) as an internal size standard), denatured at 95°C for 5 min and loaded on a 6% denaturing polyacrylamide gel. Electrophoresis was performed in a 24 cm gel for 8 h at 2500 V 40 mA and 30 W in an ABI 373A sequencer. Data were collected and sized using the Genescan 672 and Genotyper 2.1 software (ABI).

Data obtained from different gel electrophoreses were tested for uniformity by comparing the allele specific size values assigned to two reference samples loaded in each gel, namely the CEPH 134702 cell line and a local DNA sample.

Allele standardisation and control of family segregation were performed using the GAS package version 2.0 (Alan Young, Oxford University, 1993–95).

HLA-DRB1 typing was performed by the Sequence Specific Primer (SSP) technique (Dynal DR low resolution SSP; Dynal, Oslo, Norway) and by the sequence specific oligonucleotide (SSO) technique according to the XII Histocompatibility Workshop protocol.

APOE genotypes were determined by digestion with the restriction enzyme Hhal of the amplified fragment containing the polymorphic exon IV of the *APOE* gene (primers: 5'-TCG GCC GCA GGG CGC TGA TGG-3' and 5'-CTC GCG GGC CCC GGC CTG GTA-3'). Electrophoresis on 2% Metaphor (FMC, Philadelphia, PA, USA) agarose plus 1% standard agarose and staining with ethidium bromide were used.

Statistical Analysis

The data were analysed for linkage by two non-parametric methods.

The GENEHUNTER program[25] was adopted because it allows extraction of linkage information from all relatives, including families with affected members other than sibs, and lacking one or both parents. NPL_{all} scores were determined. Single-point analysis was performed for all markers. The program calculates *P* values based on the asymptotically normal statistic. NPL scores for X-linked markers were calculated by the xgh version of the program. Multipoint analysis was also performed in the chromosomal regions where several closely located markers were tested. Genetic distances between the markers were calculated by the LIK2P module of the GAS 2.0 package and were in substantial agreement with the Généthon map distances from CEPH families. When the linkage lod score was not significant (i.e. < 3) we utilised the sex-averaged genetic distances from the Généthon map[23] for Généthon markers and genetic distances from LDB database[26] for the remaining markers, except for MAOB whose distances were derived from Davies *et al*.[27]

Table 1 Composition of tested multiplex families

	Total no. of families	*Sib pairs*	*Sib trios*	*with 2 parents*	*with 1 parent*	*without parents*	*families with other affected relatives*	*no. affected individuals*	*no. analysed individuals*
Continental Italy	41	37	1	22	9	7	3[a]	87	186
Sardinia	28	24	2	13	8	5	2[b]	57	158

[a]one uncle/nephew pair and two first cousin pairs; [b]first cousins.

SimIBD[28] uses a conditional simulation approach to produce an empirical null distribution and empirical *P* values. Like GENEHUNTER it uses all available genotypes in unaffected individuals to measure identity-by-descent (IBD) sharing. In some instances, it possesses a power of extracting linkage information higher than GENEHUNTER-All statistics.[29] The 1/sqrt(*p*) weighting function was used, where *p* is the population frequency of a given allele, with a number of replicates of 1000.

Results

Two panels of multiplex families, Sardinians (28 families) and Continental Italians (41 families), were analysed by GENEHUNTER and SimIBD. Each test extracts subsets of the complete linkage information from the available data set and they may increase the power of the overall analysis and reduce the risk of type II errors.

A summary of the data is provided in Table 2. This is so organised as to facilitate comparison with data from other studies. It reports the cytogenetic localisation of the tested markers as derived from the LDB integrated map;[26] the genetic distances between the markers; the NLP_{all} scores obtained by GENEHUNTER; the empirical *P* values calculated by SimIBD and the linkage scores in genome screens performed in USA/French,[30] UK,[31] Canadian,[32] and Finnish families.[11,33] Linkage data are reported for each of the two sets of families, Sardinian (Sard) and Continental Italians (Cont) and for both sets combined (total).

Only the single point NPL statistics is shown. NPL probability levels were in all cases > 0.05 except for D7S484 in the combined set (NPL 1.66 $P = 0.048$). Suggestions of linkage, ie NPL scores equal to or higher than 1.0 and/or SimIBD *P* values < 0.05 in at least one family set were found in 2p11, 2q33 (*CTLA-4*), 3p11, 3q21, 5p15, 5p11, 5q12, 5q14, 6p21 (*HLA*), 7p15, 17q12, 22q13, Xp11. They are highlighted in Table 2. Mostly negative scores were observed in chromosomes 1p, 3p, 4q, 7q, 12p, 12q, 14q, 17p, 19q, Xq and for the individual candidate genes *TNFR2, IL9, BCL2* and *APOE*. The D17S807 marker in 17q21 was particularly interesting because it showed suggestive linkage evidence both in the Finnish[11] and in the UK[31] screens. We obtained an NPL score of about 0.9, lower than the relatively high scores of the previous studies, but close to our arbitrary threshold of 1.0.

Multipoint NPL analysis in the regions where multiple markers were tested did not substantially increase the linkage scores, despite acceptable levels of extraction of information content. The multipoint data are shown only for chromosome 5 (Figure 1). With a level of information extraction of about 70%, the maximum NPL_{all} score observed was 1.13 ($P = 0.13$) in the Sard set in the q14 region near the D5S815 microsatellite. The same marker gave an MLS 1.14 in the US/FR study. Low or negative scores were given by other chromosome 5 regions, even though this chromosome appeared the most promising as judged from data of linkage in the other studies and from the homology of the p14–p12 region with the *eae2* mouse region.

Discussion

This is the first linkage study in MS performed on Southern European Caucasoids. Our data provide some evidence of linkage in regions that were singled out by previous studies. The linkage scores are in general low, yet positive and not dissimilar from those expected from genes of low effect studied in populations of limited sizes. In fact, even the contribution of HLA, that is supported by association studies in several populations, was not detected by linkage in all studies and, within each study, in all sets.[11,30–33] Likewise, we observed an NPL score > 1 with *HLA-DRB1* in Continental Italians and a lower, although still positive score in Sardinians, even though a significant *HLA* association was observed in both populations[34] (CB unpublished results). The relative contribution of HLA to familial aggregation can be estimated in the Continental families where 19% of sibships share no alleles IBD. Following Risch[35] and assuming a multiplicative model and a λs for MS = 20, the *HLA* contribution was 9.3% by the calculation method proposed by Todd.[36] This is comparable to the figure of 13.5% deduced from the data of Sawcer *et al*[31] under the same assumptions but much lower than the data of Risch (29%) and of Haines *et al*[37] (54%). The reasons for these wide fluctuations are not easily understood and could reflect marked population variations.

The problem of replication is strictly related to the magnitude of the genetic effect to be detected. The chance of replication decreases not linearly, but according to the square of the effect of the gene. All linkage studies so far indicate the absence of any gene with a predominant effect in MS. However, the evidence for the presence of MS genes and for them accounting for most of the family aggregation is overwhelming.[8–9] Assuming that 10 (or more) epistatic genes of equal effect contribute a λs of about 30, each of them would have a $\lambda s = 1.4$ (or less). Under these conditions,

Table 2 Single point linkage data in Sardinians, Continental Italians and combined sets

Chr.	*Cytogenetic band*	*Marker*	*Distance* cM*	*GENE HUNTER (NPL)*			*SIMIBD (P)*			*UK*	*US/FR*	*C*	*F*
				Total	*Sard.*	*Cont.*	*Total*	*Sard.*	*Cont.*	*MLS*	*lod score*	*MLS*	*lod score*
1	p36.22	TNFR2	24.4	0.10	−0.94	0.91	0.10	0.25	0.55				
	p36.12	D1S199	22.7	−0.62	−0.96	−0.01	0.43	0.21	0.59	1.2			
	p34.3	D1S201	>50[b]	−0.67	0.18	−1.03	0.80	0.30	0.88	0.93		0.95	
	p22.2	D1S216		−0.01	−0.21	0.16	0.12	0.34	0.16	1.0			
2	p16.3	D2S119	41[a]	−0.98	−0.26	−1.07	0.7	0.17	0.88			1.24	
	p11.2	D2S169	2.6	0.57	0.22	0.56	**0.019**	0.18	**0.044**	1.3			
	p11.1	D2S139	>50[b]	**0.98**	0.92	0.53	0.8	0.52	0.85	1.4			
	q33.1	CTLA-4		0.67	**1.04**	0.01	0.39	0.52	0.34				
3	p21.2	D3S1289	9	0.18	−0.41	0.58	0.16	0.34	0.15	1.2			
	p14.3	D3S1300	10	0.13	−0.30	0.41	0.24	0.23	0.44	1.3			
	p14.1	D3S1285	6	−0.03	0.14	−0.16	0.74	0.58	0.68	1.1			
	p13	D3S1261	18	−0.98	−1.00	−0.48	nt	nt	nt	0.8		0.99	
	p11.1	D3S1595	16	0.43	−0.49	**0.96**	0.31	0.89	0.15				
	q21.1	D3S1278	14	**1.10**	0.85	0.73	0.42	0.45	0.46				
	q21.3	D3S3607	14	−0.25	−0.32	−0.06	0.20	0.80	0.16				
	q22.2	D3S1309		−0.21	0.01	−0.29	0.74	0.65	0.71			1.01	
4	q32.1	D4S415	9	−1.07	−1.27	−0.32	0.95	0.9	0.79				
	q35.2	D4S426	10	−0.40	−0.19	−0.37	0.46	0.72	0.28	1.4			
5	p15.33	D5S417	6	0.93	−0.00	**1.24**	0.10	0.80	**0.002**			1.8	
	p15.31	D5S406	17	−0.27	−0.21	0.18	0.85	0.93	0.50			4.2	
	p15.2	D5S416	7	−1.06	−0.83	−0.68	0.95	0.70	0.93				3.4
	p15.1	D5S655	12	−0.85	−0.23	−0.93	0.96	0.81	0.94				
	p13.1	D5S477	7	−0.41	0.06	−0.60	0.88	0.65	0.86				1.5
	p12	D5S455	16	−0.29	−0.67	0.20	0.59	0.67	0.45				2.2
	p11	D5S1968	8	0.55	−0.33	**1.03**	0.23	0.62	0.13				2.0
	q11.1	D5S427	15	0.50	0.56	0.18	0.70	0.51	0.72	2.5			
	q12.1	D5S424	13[a]	0.51	**1.00**	−0.19	0.13	**0.02**	0.76	1.3		0.42	
	q14.1	D5S428	4	−1.27	0.11	−1.80	0.62	0.29	0.84	1.0			
	q14.3	D5S815	47[a]	−0.55	**1.15**	−1.75	0.78	0.16	0.98		1.14		
	q31.1	IL9	1.5	0.33	0.70	−0.16	0.24	0.45	0.24				
	q31.1	D5S393	2	−0.13	0.10	−0.27	0.94	0.91	0.85				
6	p21.31	HLA-DRB1		**1.26**	0.56	**1.17**	0.36	0.54	0.31				
7	p15.3	D7S629	19	−0.14	0.36	−0.48	0.35	0.25	0.50	1.6			
	p15.2	D7S484	>50[b]	**1.66**	**1.60**	0.84	0.21	0.44	0.09	1.2		0.38	
	q11.23	D7S524	10	0.50	−0.37	0.95	0.82	0.77	0.70	0.5		0.7	
	q22.1	D7S554	13	0.13	0.39	−0.15	0.13	0.18	0.22		2.86		
	q31.31	D7S523		0.14	0.76	−0.44	0.33	0.21	0.69		1.11		
12	p13.31	D12S374	10	−0.19	−0.04	−0.22	0.69	0.63	0.60				
	p13.31	D12S77	8.8	−0.38	0.11	−0.61	0.73	0.66	0.63	0.8			
	p12.3	D12S364	5.6	−0.32	0.06	−0.50	0.47	0.28	0.64	1.0			
	p12.3	D12S62	3	−1.82	−1.41	−1.19	0.77	0.78	0.93	2.2			
	p12.3	D12S310	10	−1.91	−0.81	−1.84	1.00	0.88	1.00	1.8			
	p11.1	D12S87	>50[b]	−0.72	−0.28	−0.72	0.90	0.90	0.62	0.6			
	q23.2	PAH	>50[b]	−1.28	−0.64	−1.14	0.96	0.68	0.96		1.56		
	q24.33	D12S392		0.33	−0.01	0.45	0.17	0.14	0.40		1.71		
14	q32.33	D14S292		−0.06	−0.04	−0.04	0.48	0.67	0.26	1.4			
17	p11.2	D17S953	17	0.36	0.25	0.26	0.31	0.75	0.13	1.5			
	p11.1	D17S798	15	0.60	−0.07	0.86	0.43	0.27	0.66	1.5			
	q12	D17S250	26	0.55	**1.16**	−0.28	0.83	0.61	0.85	0.3			
	q21.33	D17S807		0.93	0.50	0.80	0.19	0.29	0.14	2.7			2.8

Table 2 continued

Chr.	Cytogenetic band	Marker	Distance* cM	*GENE HUNTER (NPL)* Total	Sard.	Cont.	*SIMIBD (P)* Total	Sard.	Cont.	*UK* MLS	*US/FR* lod score	*C* MLS	*F* lod score
18	q21.3	BCL-2		0.27	–0.19	0.57	0.46	0.78	0.19				
19	q13.11	D19S49	7	–0.98	–0.26	0.17	0.69	0.72	0.52				
	q13.12	D19S251	6	0.28	–0.01	0.38	0.38	0.77	0.06	1.1			
	q13.12	D19S225	9	–0.02	–0.04	0.00	0.20	0.29	0.25	1.1			
	q13.113	D19S220	7	0.29	0.31	0.13	0.51	0.40	0.61	1.2			
	q13.2	D19S217	4	–0.44	–0.04	–0.55	0.66	0.34	0.63				
	q13.32	D19S219	7	–0.01	–0.59	0.48	0.49	0.90	0.15		1.13		
	q13.32	D19S606	7	–0.42	–0.04	–0.50	0.61	0.43	0.68				
	q13.32	APOC2	5b	–0.37	–0.19	–0.32	0.60	0.80	0.67	1.4	1.47		
	q13.32	APOE	3b	–0.27	–0.37	–0.65	0.56	0.24	0.58				
	q13.33	D19S246		–0.57	0.84	–1.43	0.34	0.10	0.74	1.5			
22	q13.1	D22S283	12	0.92	**1.01**	0.36	0.11	**0.04**	0.33	1.1			
	q13.1	PDGFB	14	–0.02	0.18	–0.18	0.40	0.50	0.36	1.4			
	q13.1	D22S274		0.01	–0.86	0.73	0.56	0.86	0.22	1.2			
X	p11.4	DXS1068	13[c]	–0.84	0.07	–1.19	0.97	0.73	0.97	0.3		1.85	
	p11.3	MAOB	18[c]	–0.33	**1.04**	–1.37	0.89	0.59	0.88				
	p11.21	DXS991	17[c]	–0.09	–0.85	0.63	0.24	0.61	0.10	1.8			
	q21.33	DXS990	19[c]	0.01	–0.83	0.75	0.59	0.73	0.47	1.2			
	q23	DXS1059		–0.06	0.28	–0.33	0.54	0.57	0.45	1.7			

*calculated distances except for [a]Genethon map distances; [b]LDB;[26] [c]Davies *et al*;[27] see Materials and Methods.

linkage analysis is unlikely to provide significant lod scores with any practical number of multiplex families[15] and any low level lod score that may be obtained is unlikely to be replicated in independent studies or sets of patients. However, if any of the genes has a somewhat higher effect, eg a λs close to 2, similar to that of *APOE* in Alzheimer's disease,[38] it is still possible that it can be detected by a sample size of the order of a few hundreds, which is already available by combining the multiplex families so far tested (Table 3). In the present study, the same markers used by the other groups were tested also in view of a possible meta-analysis. It is desirable that such a combined analysis be performed.

The limitation due to the sample size in any single study and population is apparent considering the multipoint analysis of chromosome 5 in the Sardinian families (Figure 1). The NPL score peak around 5q14, which is absent in Continental Italians, may point to a favourable situation in the Sardinians in terms of frequency of a particular disease allele. However, further linkage analysis is discouraged by the fact that the 28 families tested already represent about 50% of the total multiplex families in Sardinia.

Why is it important to record 'weak linkage data' despite the seemingly disproportionate effort required? The hopes of finding 'major genes' have disappeared in MS as well as in other complex diseases. The underlying reasons, beside the low sensitivity of the analytical tools, may also be the presence of strong and difficult to unravel epistatic interactions between different genes. However, there is a recognisable trend towards accumulation of weak linkage evidence on some regions and not on others, with some sharing between different autoimmune diseases.[39] These 'tips of icebergs' that in some case coincide with '*eae* homology regions' may be important pointers to regions where association studies could be undertaken. The low-degree *HLA* linkage provides a good example. Historically, since HLA is an obvious candidate locus, linkage studies were preceded by association studies. The rather low and scattered linkage scores obtained in the different sets of families, including ours do not represent a proof of involvement of HLA in the disease. However, retrospectively, if they had prompted an association study, the latter would have been successful. One can hope that similar weak evidence of linkage in other regions underlies the presence of some disease gene that can be detected by association studies. The latter are *per se* more sensitive and experimentally easier since they require single-case families. In this regard, some regions stand out from the present work as good candidates, ie 2p11, marked by

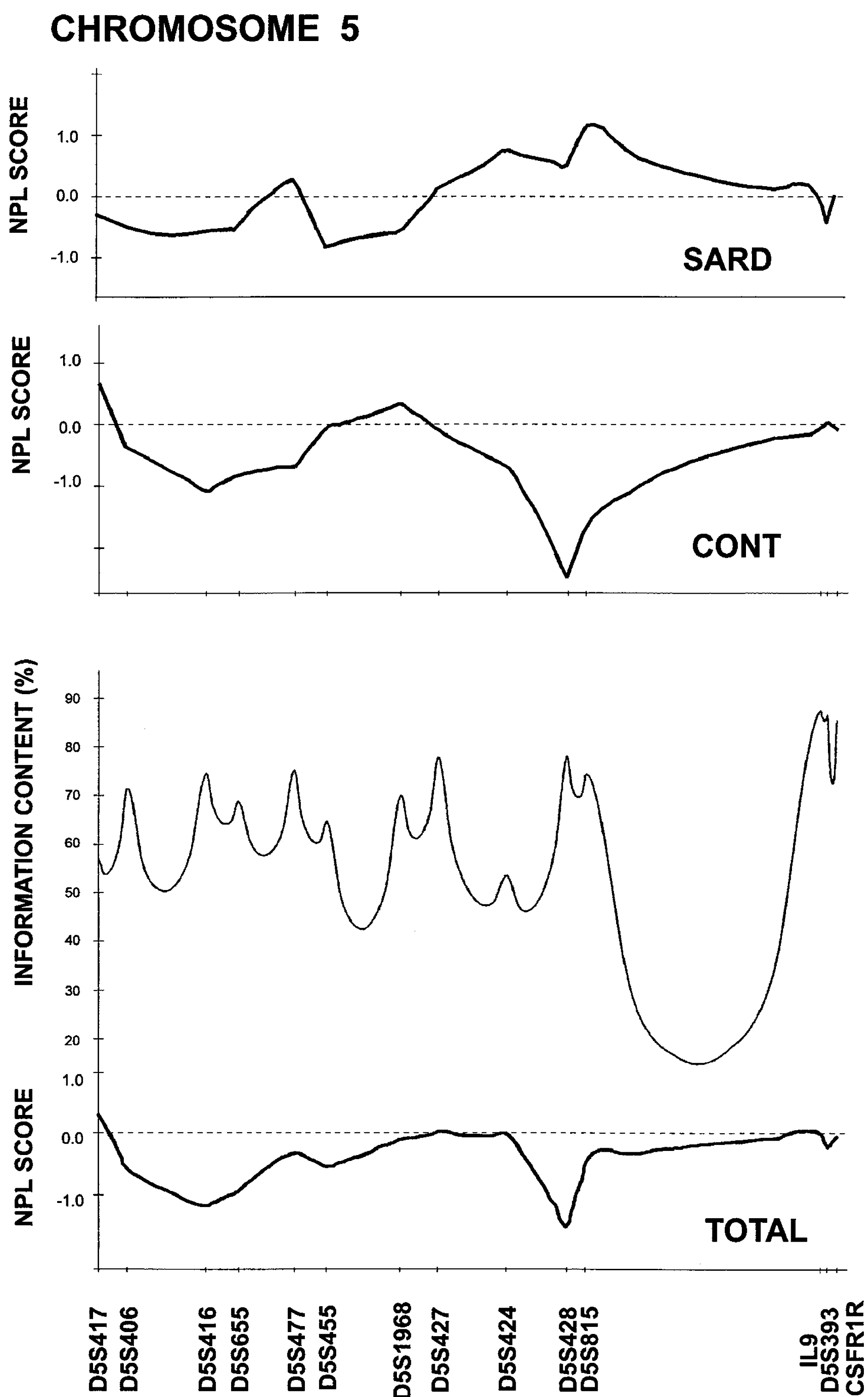

Figure 1 *Multipoint linkage analysis of chromosome 5 markers by GENEHUNTER. Sardinians (Sard), Continental Italians (Cont) and the two sets combined (Total) are shown. The thick line graphs show multipoint* NPL_{all} *scores (basal 0 score is indicated by the broken line). The thin line graph shows the percentage of information extracted from the total set of families. The information curves of the two subsets (not shown) were very similar.*

Table 3 Number of families utilized in MS linkage studies

Study	*Ref*	*Set*	*No. fam's*	*No. pairs*	*Set*	*No. fam's*	*No. pairs*	*Set*	*No. fam's*	*No. pairs*
UK	Sawcer *et al*[34]	1	129	143	2	98	128			
US/FR	Haines *et al*[32]	1	52	81	2	23	45			
Can	Ebers *et al*[35]	1	61	100	2	42	44	3	72	78
Fin	Kuokkanen *et al*[11]		21	24						
It	present study	Cont	41	39	Sard	28	28			
Total		All sets	567	710						

microsatellites D2S169 and D2S139, 3q21 (microsatellite D3S1278), 7p15 (microsatellite D7S484) and 22q13 (microsatellite D22S283). 7p15 is especially appealing since it is homologous to a region harbouring a rat *eae* locus[14] and coincides with a region where markers linked to several autoimmune diseases (MS, Crohn disease and asthma) are clustered.[39] Moreover, our data support previous results indicating that there might be a candidate susceptibility locus in 17q21.3 (microsatellite D17S807).[11]

In conclusion, we think that the present data provide some additional evidence for the presence of linkage to MS in selected genomic regions and also provide sufficient ground for the collection of a large number of simplex families in Sardinia and in Continental Italy in which to test for genetic association by a dense array of markers.

Acknowledgements

This work was supported by grants from the Italian Multiple Sclerosis Foundation (FISM), by the National Project on Multiple Sclerosis of the Istituto Superiore di Sanità and by MURST. The help of Drs Richard Butler and Walter Rolando in the implementation of the statistical programs is gratefully acknowledged.

References

1 Poser CM: The epidemiology of multiple sclerosis: a general overview. *Ann Neurol* 1994; **36 (S2)**: S180–S193.
2 Reder AT, Arnason BGV: Immunology of multiple sclerosis. In: Vinken PJ, Bruyn GW, Klawans HC (eds). *Handbook of Clinical Neurology.* Elsevier Science: Amsterdam, 1985, vol. 47, pp 337–396.
3 Martin R, McFarland H, McFarlin D: Immunological aspects of demyelinating disease. *Annu Rev Immunol* 1992; **10**: 153–187.
4 Hohlfeld R, Londei M, Massacesi L, Salvetti M: T-cell autoimmunity in multiple sclerosis. *Immunology Today* 1995; **16**: 259–261.
5 Kurtzke JF: MS epidemiology world wide. One view of current status. *Acta Neurol Scand* 1995; **S161**: 23–33.
6 Ebers GC, Sadovnick AD: The geographic distribution of multiple sclerosis: a review. *Neuroepidemiology* 1993; **12**: 1–5.
7 Robertson NP, Fraser M, Deans J, Clayton D, Walker N, Compston DAS: Age-adjusted recurrence risks for relatives of patients with MS. *Brain* 1996; **119**: 449–455.
8 Ebers GC, Sadovnick AD, Risch NJ and the Canadian Collaborative Study Group: Familial aggregation in MS is genetic. *Nature* 1995; **377**: 150–151.
9 Sadovnick AD, Ebers GC, Dyment DA, Risch NJ and the Canadian Collaborative Study Group: Evidence for genetic basis of MS. *Lancet* 1996; **347**: 1728–1730.
10 Dyment DA, Sadnovich AD, Ebers GC: Genetics of MS. *Hum Mol Genet* 1997; **6**: 1693–1698.
11 Kuokkanen S, Gschwend M, Rioux JD *et al*: Genomewide scan of MS in Finnish multiplex families. *Am J Hum Genet* 1997; **61**: 1379–1387.
12 Lander ES, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
13 Sundvall M, Jirholt J, Yang HT *et al*: Identification of murine loci associated with susceptibility to chronic experimental autoimmune encephalomyelitis. *Nat Genet* 1995; **10**: 313–317.
14 Roth MP, Viratelle C, Dolbois L *et al*: Genome wide search identifies two susceptibility loci for experimental autoimmune encephalomyelitis on rat chromosomes 4 and 10. *J Immunol* in press.
15 Risch N, Merikangas K: The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
16 Nisticò L, Buzzetti R, Pritchard LE *et al*: The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. *Hum Mol Genet* 1996; **5**: 1075–1080.
17 Marron MP, Raffel LJ, Garchon HJ *et al*: Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA-4 polymorphisms in multiple ethnic groups. *Hum Mol Genet* 1997; **6**: 1275–1282.
18 Piazza A, Mayr WR, Contu L *et al*: Genetic and population structure of four Sardinian villages. *Ann Hum Genet* 1985; **49**: 47–63.
19 Granieri E, Casetta I, Tola MR: Epidemiology of MS in Italy and in Southern Europe. *Acta Neurol Scand* 1995; **S161**: 60–70.
20 Rosati G, Aiello I, Pirastu MI *et al*: Epidemiology of MS in north-western Sardinia: further evidence for higher frequency in Sardinians compared to other Italians. *Neuroepidemiology* 1996; **15**: 10–19.

21 Kotsa K, Watson PF, Weetman AP: A CTLA-4 gene polymorphism is associated with both Graves' disease and autoimmune hypothyroidism. *Clin Endocrinol* 1997; **46**: 551–554.
22 Poser CM, Paty DW, Scheinberg L: New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; **13**: 227–231.
23 Dib C, Faure S, Fizames C *et al*: A comprehensive genetic map of the human genoma based on 5264 microsatellites. *Nature* 1966; **380**: 152–154.
24 Murray JC, Buetow KH, Weber JL *et al*: A comprehensive human linkage map with centimorgan density: Cooperative Human Linkage Center (CHLC). *Science* 1994; **265**: 2049–2054.
25 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and non parametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
26 Collins A, Frezal J, Teague J, Morton NE: A metric map of humans: 23,500 loci in 850 bands. *Proc Natl Acad Sci USA* 1996; **93**: 14771–14775; (http://cedar.genetics.soton.ac.uk/publ.ic — html/gmap.html).
27 Davies JL, Kawaguchi Y, Bennett ST *et al*: A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 1994; **371**: 130–136.
28 Davis S, Schroeder M, Goldin LR, Weeks DE: Nonparametric simulation-based statistics for detecting linkage in general pedigrees. *Am J Hum Genet* 1996; **58**: 867–880.
29 Davis S, Goldin LR, Weeks DE: SimIBD: A powerful robust non parametric method for detecting linkage in general pedigrees. In: Pawlowitzki I-H, Edwards JH, Thompson EA (eds). *Genetic Mapping of Disease Genes*. Academic Press: San Diego, CA, 1997, pp 189–204.
30 Haines JL for The Multiple Sclerosis Genetics Group: A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. *Nat Genet* 1996; **13**: 469–471.
31 Sawcer S, Jones HB, Feakes R *et al*: A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* 1996; **13**: 464–468.
32 Ebers GC, Kukay K, Bulman DE *et al*: A full genome screen in multiple sclerosis. *Nat Genet* 1996; **13**: 472–476.
33 Kuokkanen S, Sundvall M, Terwilliger JD *et al*: A putative vulnerability locus to multiple sclerosis maps to 5p14–p12 in a region syntenic to the murine locus Eae2. *Nat Genet* 1996; **13**: 477–480.
34 Marrosu MG, Murru MR, Costa G *et al*: Multiple sclerosis in Sardinia is associated and in linkage disequilibrium with HLA-DR3 and -DR4 alleles. *Am J Hum Genet* 1997; **61**: 454–457.
35 Risch N: Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987; **40**: 1–14.
36 Todd JA: Genetic analysis of type 1 diabetes using whole genome approaches. *Proc Natl Acad Sci USA* 1995; **92**: 8560–8565.
37 Haines JL, Terwedow HA, Burgess K *et al*: Linkage of the MHC to familial multiple sclerosis suggests genetic heterogeneity. *Hum Mol Genet* 1998; **7**: 1229–1234.
38 Roses AD, Devlin D, Conneally PM *et al*: Measuring the genetic contribution of ApoE in late-onset Alzheimer disease. *Am J Hum Genet* 1995; **57: (suppl.)**: A202.
39 Becker KG, Simon RM, Bailey-Wilson JE *et al*: Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci USA* 1998; **95**: 9979–9984.