Journal of NeuroVirology (2000) 6, Suppl 2, S18–S22

www.jneurovirol.com

An attempt of identifying MS-associated loci as a follow-up of a genomic linkage study in the Italian population

Sandra D'Alfonso[1], Lorenza Nisticò[2], Daniela Bocchio[1], Roberto Bomprezzi[3], Maria Giovanna Marrosu[4], Maria Rita Murru[4], Marina Lai[4], Luca Massacesi[5], Clara Ballerini[5], Anna Repice[5], Marco Salvetti[3], Chiara Montesperelli[3], Giovanni Ristori[3], Maria Trojano[6], Maria Liguori[6], Domenico Gambi[7], Aldo Quattrone[8], Roberto Tosi[2] and Patricia Momigliano-Richiardi*,[1]

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

Subsequent to a genomic linkage study on Sardinian and Continental Italian families, we considered the possibility that some of the tested microsatellite markers showed association to MS. Markers selected on the basis of the data obtained in the original set of 70 multiplex families were tested for MS association in an additional set of 154 simplex families. A limited set of markers were further tested on an additional set of 100 simplex families. The results indicate the presence of a putative MS gene in 19q13.13. *Journal of NeuroVirology* (2000) **6,** S18–S22.

Keywords: multiple sclerosis; association study; candidate regions; Italian families; microsatellites

Introduction

The problem of identifying MS loci has been approached by studies of linkage of microsatellite markers throughout the genome. For complex diseases such as MS, linkage analysis is usually performed by a non parametric approach utilising families with more than one affected sib (multiplex families). Linkage is revealed by an increased allele sharing between the two affected sibs above the expected 50%. However, the linkage approach does not permit us to pinpoint restricted genome regions or to look for specific candidate genes. This is mainly due to the inherent low power of the affected sib pair method and to the influence on lod scores of markers at considerable distances. Genetic association is more sensitive and precise in terms of localisation and can be performed on more easily available simplex families (with a single affected offspring) by testing the preferential transmission of disease-associated alleles from heterozygous parents to affected offspring. However a systemic genome association screening requires a very high marker density and is not feasible with the existing technology.

We decided to try an intermediate approach. Starting from the 67 microsatellite markers that we utilised for our previous linkage study (D'Alfonso *et al*, 1999) and which were selected in genome regions showing some evidence of linkage to MS in other genome screens (Sawcer *et al*, 1996; Ebers *et al*, 1996; Haines *et al*, 1996), we tested whether some of them showed evidence of association. We were aware that such an approach had a poor chance of success being the density of markers much lower than theoretically required. Nevertheless, we considered this attempt worth pursuing in view of the importance of the issue.

We started from our original set of Sardinian and Continenal Italian multiplex families used for linkage analysis and applied to them the transmission-disequilibrium test (TDT). Under these conditions, TDT becomes a test of linkage only in the presence of association. The increased transmission of a marker raises the possibility of the presence of a disease gene in very close proximity (<1 cM) to it. Markers showing increased transmission in the set

*Correspondence: P Momigliano-Richiardi

of multiplex families were then re-tested for true association by TDT on independent sets of simplex families.

The data obtained are indeed suggestive of the presence of an MS gene in a restricted region of chromosome 19.

Results

In a previous linkage study (D'Alfonso *et al*, 1999) we analysed 67 microsatellite markers in multiplex Sardinian (SAR1) and Continental Italian families (ITC1). In the present study the same data were systematically tested by TDT. In most cases no significant deviation from the expected 50/50 transmission/non transmission ratio was detected. Ten loci had an allele (two alleles in the case of D7S524) showing a significantly increased transmission ratio to affected sibs while the transmission to unaffected sibs was either non significantly distorted or was decreased (Table 1). Evidence of association for at least four loci was strengthened by a significant ($P<0.05$) Tsp test. The latter, which contains a correction factor for multiple affected sibs, offers a weighted measure of association based on the whole allelic distribution of a given locus (Martin *et al*, 1997). No attempt of grouping low frequency microsatellite alleles was made and *P* values were not corrected for number of comparison. The likelihood that most deviations are due to chance is therefore high. However we intended at this stage to avoid discarding true positives (i.e. type II errors). In order to test whether the markers indicated in Table 1 were truly associated, they were tested on an independent set of 100 Sardinian (SAR2) and 54 Continental (ITC2) simplex families. On simplex families TDT is a true association test (Spielman *et al*, 1993). Ten of the 11 alleles tested failed to maintain a signifcant positive association in either population and also in the combined set (Table 2). Only allele 10 at the D19S220 locus, located in 19q13.13, showed a significantly increased transmission ($P=0.029$) in the Sardinian simplex families. A borderline significance ($P=0.053$) was maintained also in the combined set.

Based on this replicated result, we concentrated on the D19S220 region. Four other microsatellites were selected in close proximity of D19S220 such as to span the region where the presence of a putative disease gene may cause association due to linkage disequilibrium. Since linkage disequilibrium is normally expected within a distance of less than 0.5 cM (about 0.5 Mb) (Thompson and Neel, 1997), the microsatellites were chosen in a range of roughly 0.5 cM centromeric and telomeric to D19S220. The genetic distances, based on Genethon database (Dib *et al*, 1996) are shown in Table 3. They must be considered rather tentative in this close range. Physical distances are known between D19S224 and D19S220 (1.2 Mb) and between D19S220 and D19S421 (about 0.3 Mb) (Ashworth *et al*, 1995). The three loci D19S220, D19S876 and D19S909 are possibly at a very short distance from each other. Significant pairwise linkage disequilibria were detected (using the Arlequin program) in the Sardinian population between microsatellites D19S220 and D19S876 and the other four closely linked loci, while the more distant D19S224 and D19S421 were not in linkage disequilibrium with each other (data not shown).

The new markers were tested with all simplex family sets. Moreover, since the association with D19S220 appeared to be stronger in the Sardinian population (Table 2) a further set of 100 Sardinian simplex families (SAR3) was tested. The TDT results are shown in Table 3. The telomeric marker D19S421 did not show any trend toward positive association and can be confidently ruled out. Also D19S909 is unlikely to be associated in spite of a

Table 1 Microsatellites showing a significant TDT and Tsp test in multiplex families

				TDT						
				Affected sibs			Unaffected sibs			TSP
Chr	*Cytogenetic band*	*Locus*	*Allele*	*T*	*NT*	P	*T*	*NT*	P	P
2	p11.2	D2S169	3	17	5	0.010	2	5	ns	0.053
	p11.1	D2S139	4	35	20	0.043	13	16	ns	ns
3	p14.3	D3S1300	11	45	28	0.047	21	20	ns	0.034
7	p15.2	D7S484	4	34	19	0.039	8	18	0.050[a]	ns
	q11.23	D7S524	2	47	27	0.020	15	22	ns	0.0095
			3 (sard)	18	8	0.050	6	17	0.020[a]	
	q31.31	D7S523	7	42	20	0.0052	18	16	ns	0.036
12	q23.2	PAH	3	38	22	0.039	17	16	ns	ns
17	q12	D17S250	3	42	23	0.018	19	10	ns	0.057
19	q13.13	D19S220	10	29	15	0.035	14	12	ns	0.031
	q13.33	D19S246	10 (sard)	9	1	0.011	5	3	ns	ns

Notes: Sard: results concerning only Sardinian families; ns: not significant; T, NT: number of transmitted, non transmitted, alleles (from heterozygous parents); [a]decreased transmission.

Table 2 Replication of the association study in two independent sets of simplex families

		Sardinia			*Continental Italy*					
			SAR2			*ITC2*			*Combined set*[b]	
Locus	*Allele*	*%T*	*T/NT*	P	*%T*	*T/NT*	P	*%T*	*T/NT*	P
D2S169	3	31	9/20	0.041[a]	59	10/7	ns	48	27/29	ns
D2S139	4	45	34/41	ns	58	25/18	ns	52	75/69	ns
D3S1300	11	53	35/31	ns	61	14/9	ns	56	69/54	ns
D7S484	4	52	22/20	ns	33	8/16	ns	49	45/47	ns
D7S524	2	54	38/32	ns	54	21/18	ns	57	83/62	ns
	3	59	30/21	ns	56	20/16	ns	57	65/50	ns
D7S523	7	50	26/26	ns	33	9/18	ns	52	58/53	ns
PAH	3	51	39/37	ns	39	14/22	ns	52	74/68	ns
D17S250	3	51	35/34	ns	43	13/17	ns	51	67/64	ns
D19S220	10	65	35/19	0.029	43	9/12	ns	59	58/39	0.053
D19S246	10	54	19/16	ns	50	10/10	ns	55	42/34	ns

Notes: %T: transmission ratio; T, NT: number of transmitted, non transmitted, alleles (from heterozygous parents); ns: not significant; [a]decreased transmission; [b]the combined set includes Sardinian and Continental Italian multiplex (SAR1, ITC1) and simplex (SAR2, ITC2) families. Transmission to the index case only was considered for multiplex families.

Table 3 Association analysis in the 19q13.13 region

			Sardinian families									*Continental families*			*Combined*					
			SAR2			*SAR3*			*SAR2+SAR3*			*ITC2*			*all sets*[b]			*Unaffected sibs*		
Locus	*cM*	*Allele*	*%T*	*T/NT*	P	*%T*	*T/NT*	P	*%T*	*T/NT*	P	*%T*	*T/NT*	P	*%T*	*T/NT*	P	*%T*	*T/NT*	P
D19S224	0.5	6	71	17/7	0.04	nd			71	17/7	0.041	50	5/5	ns	65	22/12	ns	45	10/12	ns
D19S220	0	10	65	35/19	0.03	56	24/19	ns	60	64/43	0.040	43	9/12	ns	59	82/58	0.042	43	35/36	ns
D19S876	0	1	77	10/3	0.05	64	9/5	ns	70	19/8	0.034	33	1/2	ns	67	20/10	0.059	40	4/6	ns
		6	58	33/24	ns	58	42/30	ns	58	75/54	0.059	63	19/11	ns	59	94/65	0.021	57	30/23	ns
D19S909	0.6	8	76	31/10	0.001	45	29/35	ns	57	60/45	ns	24	6/19	0.01[a]	51	66/64	ns	66	25/13	0.05
D19S421		2	57	29/22	ns	41	27/39	ns	48	56/61	ns	58	11/8	ns	49	67/69	ns	44	23/29	ns

Notes: ns: not significant; nd: not determined; cM: genetic distance between the corresponding microsatellite locus and that indicated on the row immediately below, from Genethon database (Dib *et al*, 1996). %T: transmission ratio; T, NT: number of transmitted, non transmitted, alleles (from heterozygous parents); [a]decreased transmission; [b]all simplex families (SAR2, SAR3, ITC2). Microsatellite D19S220 was also tested with multiplex families (SAR1, ITC1). Transmission to the index case only was considered for multiplex families.

significant positive deviation (31/10) in set SAR2 since a negative deviation is seen in the other sets. The other data, involving the closely linked loci D19S220 and D19S876 and, possibly, also the more centromeric D19S224, show a consistent trend of increased transmission mostly in the Sardinian population. The transmission of the corresponding alleles to unaffected sibs was not significantly distorted.

At the D19S876 locus two alleles showed some association. Notably, when the Sardinian data were stratified according to HLA genotypes, by subdividing transmissions to patients either positive or negative for MS-associated HLA alleles (DR3 and DR4; Marrosu *et al*, 1997) allele 1 was mostly transmitted to DR3, DR4 negative patients (%T=86, T/NT=12/2, P=0.0075) whereas allele 6 was mostly transmitted to DR3 and/or DR4 positive patients (%T=61, T/NT=49/31, P=0.044).

For none of the other markers considered in this study HLA stratification produced any marked deviation.

Discussion

Data of linkage and association of complex diseases must be considered with caution. Methods like the affected sib pair linkage analysis and TDT have inherently low sensitivity and distinction of relevant results from background noise is always a problem, especially in screenings involving a high number of markers. Testing many loci and for each locus several alleles implies an impossibly high correction factor for single probability levels. This problem can be circumvented by testing a limited number of markers on a second set of samples. Associations due to chance in the first sample will almost certainly disappear and only true associations will survive. Test of additional sets of samples and of markers located very near to the putatively associated marker may provide a further confirmation. We applied this stepwise procedure to a large number of MS families. Starting from a set of microsatellites selected in 'MS candidate regions' we have collected data suggesting MS association

for markers located in a restricted region (about 1 cM) in 19q13.13. The significance levels are mostly borderline and replication was not seen in all family sets. However the trend toward an increased transmission ratio to the patients was consistent in repeated tests especially in Sardinian families, whereas transmission to the unaffected siblings was always decreased or flat.

Even keeping in mind the above cautionary considerations, we think that there is sufficient ground for considering candidate genes in this region. Among possible candidates, APOC1, APOC2 and APOE are located at a physical distance of about 5 Mb and TGFB1 of about 2.5 Mb, telomeric to D19S220 (Ashworth *et al*, 1995). These distances are likely too high to justify linkage disequilibrium. We have already tested APOC2 and APOE in our families but neither linkage nor association was detected (data not shown). However haplotypic combinations including an APOC2 intragenic marker and closely linked microsatellite loci were significantly associated in a case-control study performed in Caucasoid Americans (Barcellos *et al*, 1997). Another appealing candidate is MAG (myelin associated glycoprotein), thought to be involved in the process of myelination. This gene is located 0.7 Mb centromeric to D19S224, i.e. at a distance range compatible with linkage disequilibrium. No association studies with MAG have been reported to date. Further candidate genes mapping between MAG and D19S224 could be: GPI (glucose-6-phosphate isomerase) which can function as a neurotrophic factor for spinal and sensory neurons; APLP1 (Amyloid-precursor-like protein 1) a membrane-associated glycoprotein predominantly expressed in brain, whose gene is homologous to the APP gene involved in the pathogenesis of Alzheimer's disease and, interestingly, RDRC (RD114 virus receptor) a cell surface molecule used for cell entry by the RD114/simian type D retroviruses and baboon endogenous viruses. This gene has been recently shown to correspond to that encoding for a neutral amino acid transporter-like protein (Rasko *et al*, 1999). Another viral receptor gene, namely E11S (echo virus sensitivity), has also been mapped in this region (Kaneda *et al*, 1987). Viral receptor genes are appealing candidates in view of the many data indicating the involvement of a viral infection in the pathogenesis of MS.

Materials and methods

Families
Multiplex families from Sardinia (SAR1=28) and from Continental Italy (ITC1=42) have already been described in D'Alfonso *et al*, 1999.

Simplex families were composed by different sets: SAR2 (n=100) and SAR3 (n=100) from Sardinia and ITC2 (n=54) from Continental Italy (Total=324 families).

Index patients, with a definite diagnosis of MS, and their relatives were recruited in the Universities of Cagliari (SAR1, SAR2, SAR3) Bari, Catanzaro, Chieti, Florence and Rome (ITC1, ITC2).

Genomic typing
Following informed consent, DNA was extracted from peripheral blood by standard techniques.

PCR products of the 10 selected microsatellites were run on 6% polyacrylamide denaturing gel as already described (D'Alfonso *et al*, 1999). 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.

Statistical analysis
Transmission Disequilibrium Test (TDT, Spielman *et al*, 1993) was used as a test of linkage when transmissions to all affected siblings of multiplex families were considered. As a test of association, transmission to index cases only was scored. Allelic association in multiplex families was also evaluated by Tsp (Martin *et al*, 1997). This statistic is more powerful than TDT to probands only, as it allows to use all affected individuals. Calculations were made using software written by Frank Dudbridge (available by anonymous ftp from ftp-gene.cimr.cam.ac.uk, directory/pub/software).

Arlequin (freely available from http://anthropologie.unige.ch/arlequin was used to assess pairwise linkage disequilibrium between markers d19s224, d19s220, d19s876, d19s909 and d19s421.

Acknowledgements

This work was supported by ISS (Istituto Superiore di Sanità) MS project, by FISM (Fondazione Italiana Sclerosi Multipla) and by MURST (Italian Ministry of University and Scientific Research).

References

Ashworth LK, Batzer MA, Brandriff B, Branscomb E, de Jong P, Garcia E, Garnes JA, Gordon LA, Lamerdin JE, Lennon G, Mohrenweiser, AA Olsen, Slezak T, Carrano AV. (1995). An integrated metric physical map of human chromosome 19. *Nat Genet* **11:** 422–427.

Barcellos LF, Thomson G, Carrington M, Schafer J, Begovich AB, Lin P, Xu XH, Min BQ, Marti D, Klitz W (1997). Chromosome 19 single-locus and multilocus haplotype associations with multiple sclerosis. Evidence of a new susceptibility locus in Caucasian and Chinese patients. *JAMA* **278:** 1256–1261.

D'Alfonso S, Nistico' L, Zavattari P, Marrosu M, Murru R, Lai M, Massacesi L, Ballerini C, Gestri D, Salvetti M, Ristori G, Bomprezzi R, Trojano M, Liguori M, Gambi D, Quattrone A, Fruci D, Cucca F, Momigliano Richiardi P, Tosi R (1999). Linkage analysis of Multiple Sclerosis with candidate region markers in Sardinian and Continental Italian families. *Eur J Hum Genet* **7:** 377–385.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380:** 152–154.

Ebers GC, Kukay K, Bulman DE, Sadovnick AD, Rice G, Anderson A, Armstrong H, Cousin K, Bell RB, Hader W, Paty DW, Hashimoto S, Oger J, Duquette P, Warren S, Gray T, O'Connor P, Nath A, Auty A, Metz L, Francis G, Paulseth JE, Murray TJ, Pryse-Phillips W, Nelson R, Freedman M, Brunet D, Bouchard JP, Hinds D, Risch N. (1996). A full genome screen in multiple sclerosis. *Nat Genet* **13:** 472–476.

Haines JL for The Multiple Sclerosis Genetics Group (1996). A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. *Nat Genet* **13:** 469–471.

Kaneda Y, Hayes H, Uchida T, Yoshida MC, Okada Y (1987). Regional assignment of five genes on human chromosome 19. *Chromosome* **95:** 8–12.

Marrosu MG, Murru MR, Costa G, Cucca F, Sotgiu S, Rosati G, Muntoni F (1997). Multiple sclerosis in Sardinia is associated and in linkage disequilibrium with HLA-DR3 and DR4 alleles. *Am J Hum Genet* **61:** 454–457.

Martin ER, Kaplan NL, Weir BS (1997). Tests for linkage and association in nuclear families. *Am J Hum Genet* **61:** 439–448.

Rasko JEJ, Battini J-L, Gottschalk RJ, Mazo I, Miller AD (1999). The RD114/simian type D retrovirus receptor is a neutral amino acid transporter. *Proc Natl Acad Sci* **96:** 2129–2134.

Sawcer S, Jones HB, Feakes R *et al*. (1996). A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* **13:** 464–468.

Spielman RS, McGinnis RE, Ewens WJ (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* **52:** 506–516.

Thompson EA, Neel JV (1997). Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. *Am J Hum Genet* **60:** 197–204.