

ARTICLE

Sharing of a conserved haplotype suggests a susceptibility gene for multiple sclerosis at chromosome 17p11

Bing He^{1,2}, Vilmantas Giedraitis^{*1}, Arturs Ligers¹, Michael Binzer³, Peter M Andersen³, Lars Forsgren³, Lodewijk A Sandkuijl⁴ and Jan Hillert¹

¹Division of Neurology, Huddinge University Hospital, Karolinska Institute, Huddinge, Sweden; ²Division of Matrix Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden; ³Department of Neurology, University hospital of Umeå, Umeå, Sweden; ⁴Institute of Clinical Genetics, Erasmus University, Rotterdam and Department of Human Genetics, Leiden University, The Netherlands

Multiple sclerosis (MS) is a chronic inflammatory disease resulting in demyelination in the central nervous system (CNS). Increasing evidence supports that genetic factors confer susceptibility to MS. One locus, the HLA complex (6p21), has been identified as important in MS, but no other loci have been clearly implicated, neither by a candidate gene approach, nor by a genomic screen strategy. Here, we studied a genetically isolated population in the northern-most part of Sweden, which demonstrates a high incidence of MS, using haplotype sharing analysis. Genealogical analysis demonstrated that 22 MS patients originate from a single common ancestral couple in the eighteenth century. Five affected individuals from four nuclear families were selected for an initial genomic screen with 390 microsatellite markers. Seven shared haplotypes in six different chromosomal regions were observed. After genotyping for these haplotypes with the same and additional markers in 15 MS patients and healthy relatives, some portion of a conserved haplotype spanning 10 cM at 17p11 was found to be shared by 12 of 15 affected individuals. The statistical analysis revealed a significant excess of transmission of alleles of three markers to affected individuals ($P < 0.05$) by the transmission/disequilibrium test (TDT). An identical four-marker haplotype was shared by six of 15 patients (40%; $P < 0.01$). Surprisingly, DR-typing revealed no significant sharing of the HLA region. In conclusion, our data suggests a novel susceptibility gene for MS in chromosome 17p11.

European Journal of Human Genetics (2002) 10, 271–275. DOI: 10.1038/sj/ejhg/5200802

Keywords: chromosome 17p; genomic screen; haplotype sharing; multiple sclerosis; susceptibility gene

Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS), characterised by relapsing or progressive demyelination of the CNS. The aetiology of MS is poorly understood, although both genetic and environ-

mental factors are considered to be of importance. Epidemiological analyses of familial clustering, twin concordance and adoptee studies^{1–3} supports the notion that several genetic factors are likely to confer susceptibility to MS as a complex trait.

Studies show that the HLA class II region on 6p21 confers susceptibility to MS, probably as a major genetic determinant,^{4,5} and also gives indirect support for the autoimmune mechanisms thought to be involved in the pathogenesis of MS. Non-HLA genes, however, have not been conclusively identified, in spite of rather large scale analyses of candidate genes, mostly focusing on components involved in immune

*Correspondence: V Giedraitis, Division of Neurology, Huddinge University Hospital, Karolinska Institutet, S-141 86 Huddinge, Sweden. Tel: +46 8 585 81927; Fax: +46 8 585 87080; E-mail: Vilmantas.Giedraitis@neurotec.ki.se
Received 15 November 2001; revised 25 February 2002; accepted 1 March 2002

regulation and possible autoimmune targets.^{6–8} In addition, four published genomic screens^{9–12} have indicated the possible importance of a few different loci, in particular, in 5p and 17q. A recent meta-analysis of published genomic screens identified several regions with nonparametric linkage scores (NPL) higher than 2, including 17q11 and 6p21.¹³ However, neither of these regions reached genome-wide significance for linkage, and most other potential loci differed between screens. These findings could result from a true but modest contribution of each involved gene, which thus are inefficiently detected by conventional linkage methods.

Searching for haplotype sharing among affected individuals in an isolated population is a potentially powerful alternative method for mapping disease genes.^{14,15} This method identifies regions shared identical-by-descent (IBD) by affected individuals through searching for conserved haplotypes. Therefore, it does not require knowledge of mode of inheritance and it is independent of accurate pedigree information. Usually, pedigree structures in isolated populations are too complicated to enable accurate evaluation of all possible routes of transmission of a phenotype and marker alleles.^{15–17} To some extent, this method is known to overcome certain drawbacks of classic lod score analysis, which is sometimes not capable of producing adequate proof for linkage. Recently, a kindred of 22 MS patients, originating from a single common ancestral couple in the 18th century, was reported from the community of Överkalix in the northern-most part of Sweden.¹⁸ Due to the geographic location, the population in this rural area is thought to have developed in relative isolation for some 300 years. Such a population is theoretically ideal for the identification of conserved disease-associated haplotypes, which includes gene mutations introduced by a few founders. In this report, we present a genetic analysis in this population by using a conserved haplotype sharing method.

Materials and methods

Pedigrees

A high incidence of multiple sclerosis (33 cases/about 4744 population) in the community of Överkalix northern Sweden previously has been described in detail.¹⁸ This isolated population is ethnically homogeneous and was probably founded in the 17th century by a few couples. In the present study, we collected DNA samples from 79 individuals, including 31 persons with MS and 48 healthy relatives, from the Överkalix area. Using public archives, a genealogical analysis established that 19 of the MS patients and their relatives originate from a single common ancestral couple. Among them, 15 affected individuals provided samples from nuclear family members that permitted the construction of haplotypes. Twelve patients have so far not been identified as members of the pedigree.

Genotyping

Genomic DNA was purified from peripheral blood by a conventional method. The microsatellite markers used for genotyping were the human linkage mapping set from the Cooperative Human Genetic Center (CHLC)¹⁹ and published maps of Genethon.²⁰ The genotyping procedures, based on the ABI 377 system for semiautomated genotyping, were as previously described.⁷ Most markers from the CHLC database used for the initial genome screen were tetranucleotide repeats, thus minimising potential errors in allele determination. In addition, all individuals were genotyped for the HLA-DRB, -DQA and -DQB genes by the restriction fragment length polymorphism (RFLP) analysis.²¹

Haplotype construction

Haplotypes defined by alleles of adjacent marker pairs were constructed by tracing the parental origin of the alleles. When parents were not available for genotyping, haplotypes were, when possible, inferred from siblings or children.

Statistical analysis

In order to test for linkage disequilibrium, the transmission/disequilibrium test (TDT) was used to determine whether alleles of a conserved haplotype demonstrated a higher frequency on disease-bearing chromosomes *vs* presumed normal, nontransmitted chromosomes.²² Within each nuclear family, the parental chromosomes transmitted to affected individuals were assumed to harbour a susceptibility allele, while those nontransmitted were assumed to harbour a normal allele. If no linkage disequilibrium exists, the expected frequency of transmitted and nontransmitted marker alleles should be 50%. The allele frequency differences between affected individuals and nontransmitted parental alleles at each locus were compared to the expected 1:1 ratio using the χ^2 test with one degree of freedom (1 df). Subsequently, haplotypes were constructed and haplotype frequencies were compared between affected individuals and nontransmitted haplotypes from their parents.

Results

Five affected individuals and their parents or siblings in four nuclear families (Figure 1) were selected for genome-wide genotyping with all 390 markers, spanning the genome at an average of 10 cM intervals. In this group, one sib-pair was included, since we considered regions unlikely to harbour susceptibility genes unless they were shared IBD in a sibling pair. Approximately 380 markers generated genotyping data. Using each pair of flanking markers, totally 358 haplotypes transmitted to patients were constructed. Next, we searched for segments shared by at least four out of eight chromosomes (50% or more) and also shared in the affected sibling pair. Seven segments on six different chromosomes were identified (Table 1). Surprisingly, haplotypes covering chro-

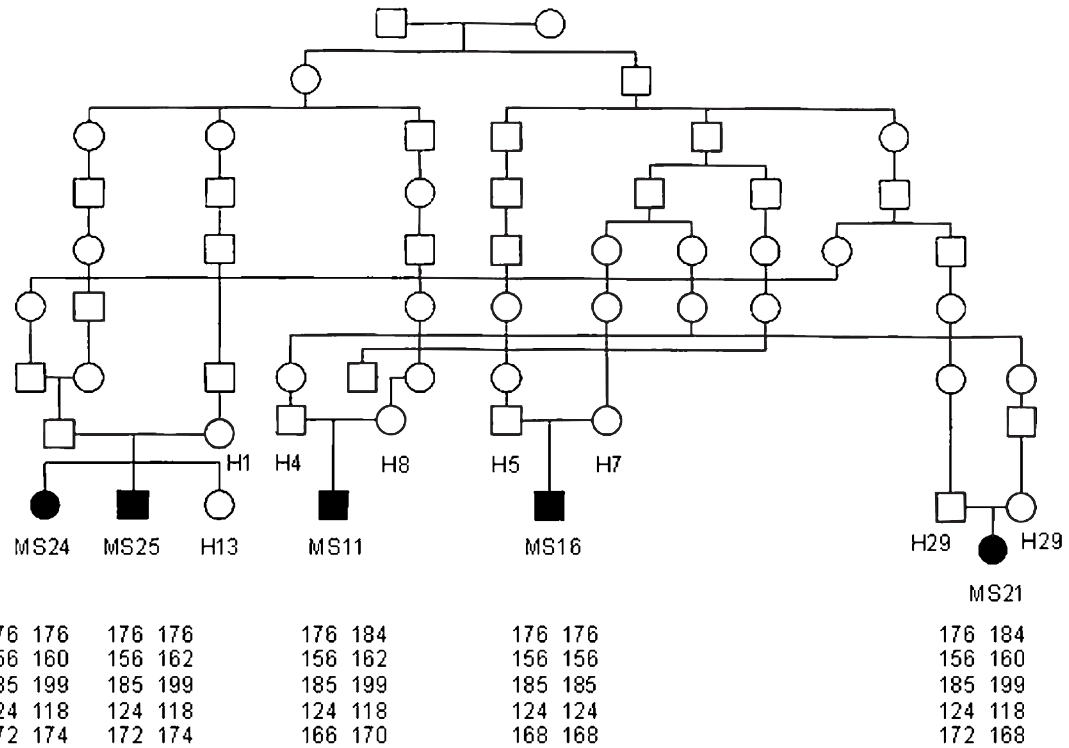


Figure 1 The pedigree of five MS patients selected for initial genomic screen and identified haplotypes expressed as length in base pairs of marker amplification product.

Table 1 Constructed flanking marker haplotypes shared by at least four of eight 'disease' chromosomes in the genomic screen. Only one of the suggested haplotypes was confirmed to be identical-by-descent (IBD) after analysis of additional markers

Flanking marker	Distance (cM)	Number shared	IBD confirmed
D4S408–D4S1652	12	5/8	no
D10S1435–D10S189	15	5/8	no
D11S912–D11S2359	20	4/8	no
D13S793–D13S779	10	5/8	no
D14S53–D14S606	5	4/8	no
D14S749–D14S611	9	5/8	no
D17S921–D17S953	7	5/8	yes

mosome 6p21, harbouring the HLA gene complex, did not reach the selected threshold for sharing, although HLA class II genes have been identified as a major locus for MS in the Swedish population.^{4,23}

Subsequently, we genotyped all the affected individuals and healthy relatives for these seven segments, also using additional markers, mainly dinucleotide repeats from the Genethon database. In one chromosomal region, in 17p11, 12 out of 15 patients belonging to the large kindred shared some portion of the conserved haplotype. Assuming that the chromosomal region defined by the shared marker alleles

might harbour a founder mutation, we subjected this region to testing for LD.

In order to evaluate the statistical evidence for this observation, we used TDT to detect LD for the markers based on the separate nuclear families within the large pedigree. The identified region at 17p11 consisted of five markers including two from the original set. Three of these five markers (*D17S921*, *D17S1843* and *D17S953*) showed a significant deviation of transmission ($P < 0.05$) in favour of transmission to affected individuals (Table 2). Furthermore, an identical four-marker haplotype (*D17S921-D17S122-D17S1843-D17S953*) was transmitted to six of 15 affected individuals, (40%) ($\chi^2 = 7.0$; $P < 0.01$) (Table 3).

In addition, we genotyped all patients and healthy relatives for HLA DR and DQ. No increased sharing of haplotypes was observed. Likewise, a comparison of transmitted and non-transmitted DR15-containing haplotypes, failed to show statistically significant transmission distortion ($\chi^2 = 0.7$, $P > 0.05$). 56% of the patients were carriers of DR15 allele in comparison with 50% of controls.

Discussion

We have used a haplotype sharing method to study a genetically isolated kindred in a community with a high prevalence of multiple sclerosis. Our results suggest the

Table 2 Transmission analysis in 14 MS patients with informative parental alleles (in base-pairs) of a conserved, shared haplotype in 17p11. The number of transmitted (T) (to MS-patients) and nontransmitted (N) parental chromosomes are given, as well as results of the χ^2 analysis

Marker	Allele	T	N	χ^2	P
D17S921	176 bp	22	10	4.5	<0.05
D17S122	156 bp	12	7	1.3	>0.05
D17S1843	185 bp	7	0	7.0	<0.01
D17S953	124 bp	18	7	4.8	<0.05
D17S1871	172 bp	11	7	0.9	>0.05

Table 3 Transmission analysis of related haplotypes on chromosome 17p11 defined by alleles of four adjacent microsatellite markers in 15 MS patients, where T signifies transmission and N nontransmission of a parental haplotype to an individual with MS

Marker	D17S921	D17S122	D17S1843	D17S953	T	N	χ^2
Interlocus distance (cM)	5	0	2				
	176	156	185	124	7	0	7.0
	176	156	189	124	1	1	0.0
	176	160	189	124	5	2	1.2
	176	162	189	124	2	0	2.0
	176	162	183	124	1	1	1.0

presence of a susceptibility gene for MS in 17p11. This conserved haplotype overlaps with a modest linkage peak observed in the British genome-wide screen.¹⁰ In addition, the MS genomic screen meta-analysis indicated an NPL score higher than 2 in this approximate region although the highest peak was found close to the neurofibromatosis-1 gene (17q11).¹³ The importance of this chromosomal segment was supported by several lines of evidence: (1) some portion of the conserved haplotypes, defined by four markers spanning ~10 cM, was shared by 12 of the 15 patients; (2) particular alleles of the three markers showed linkage disequilibrium with MS by the TDT analysis; (3) six of 15 patients shared an identical four-marker haplotype (40%) with significant excess of transmission to patients.

The first conclusion of the present findings is that even in this large pedigree, in fact the largest MS pedigree to be described so far,¹⁸ there seems to be no single genetic factor that is both necessary and sufficient for the development of MS. Thus, even in this case, MS seems to fulfil the criteria of a genetically complex disorder.

Polygenes, and even other factors, such as environmental or stochastic events, are likely to influence the susceptibility to MS, which, as a consequence, results in a non-mendelian inheritance in MS. The contribution of each involved gene is likely to be small, as Ebers *et al.*⁹ have excluded 88% of the genome for a locus having an effect corresponding to $\lambda_s > 3$, while Sawcer *et al.*¹⁰ have excluded 93% of the genome for a

locus that contributes a λ_s of >5 . The modest effect of each gene limits genetic mapping by conventional methods.²⁴ The haplotype sharing analysis may have greater power for gene localisation, as this method is based on the principle of LD mapping which differs from classic lod score and affected sibpair analysis.¹⁵ LD mapping was first proposed by Landers and Bostein in 1986, and it has been applied to map genes for simple and even complex disorders in general populations.²⁵⁻²⁷ However, it was seriously questioned whether sufficient LD exists in the general population to allow gene identification at the genomic level.²⁸ In contrast, genome-wide LD mapping in isolated populations has been proven efficient for rare diseases,¹⁵⁻¹⁷ but remains to be shown useful for complex traits such as MS.

The population of Överkalix has remained relatively isolated until recently, and the majority of the population is likely to be descendants of a small number of founders.¹⁸ We can assume that the chromosomal segments containing the disease genes brought into the population by a few ancestors are likely to be shared by most patients. The sizes of these segments are likely to remain large enough to be detected by the sharing search and further LD analysis, because of the low rate of recombination due to a limited number of meioses, and rare foreign genome interference. As MS is presumably polygenic, it is not to be expected that a possible single gene at 17p11 should confer the full genetic susceptibility to MS. The failure to detect other haplotypes could depend on insufficient density of the present marker map, or on the use of a limited number of affected individuals in the first screen in which only five patients were genotyped extensively. Durham and Freingold,²⁹ using a simulation analysis, proposed that the median size of a segment shared IBD by five chromosomes, with a separation distance of six meioses, is only ~5.5 cM. Therefore, an optimal analysis of the present material might require an increased number of nuclear families in an initial screen and a denser marker map.

The chance of a false positive finding in a genome-wide analysis such as the present must be taken into account, although the theoretical risk has been evaluated using simulation analysis.^{15,29} In general, the shared haplotype method is considered to give a low rate of false positives. To minimise false positive findings in the initial screen, we set up the strict criteria outlined above. Indeed, a haplotype (D17S921 -D17S953) at 17p11 was shared by five of eight chromosomes among four patients and was shared by two affected sibs. In the continuation of this project, additional markers will be included to allow a denser marker map, which may permit the identification of smaller shared chromosomal segments. Further analysis of the identified segment in independent patient materials could help to delineate the region of interest.

The chromosome 17p11 region is rich in genes and a number of potential candidates may be selected. Possible candidate genes include peripheral myelin protein 22 and

neurofibromin-1, both well known to cause neurological disorders.

In contrast, the transmission of DR15 to MS patients in this material was not more common than nontransmission. This implies that the HLA class II locus is less important in this kindred, compared to the general Swedish population of unrelated MS patients where its importance has been confirmed in several case–controls studies.³⁰ In conclusion, our findings suggest the existence of a susceptibility gene for MS on chromosome 17p11. Further work is necessary to fine-map the 10 cM shared chromosomal segment.

Acknowledgments

This study was supported from the Swedish Medical Research Council (project numbers:11023 and 11220), the Society for the Neurologically Disabled, the Sigurd and Elsa Goljes Minne foundation, the Karolinska Institute, the Magn Bergvalls foundation, the Åke Wiberg and the Marcus Borgströms foundation.

References

- 1 Ebers GC, Sadovnick AD: The role of genetic factors in multiple sclerosis susceptibility. *J Neuroimmunol* 1994; **54**: 1–17.
- 2 Ebers GC, Sadovnick AD, Risch NJ: A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature* 1995; **377**: 150–151.
- 3 Sawcer S, Goodfellow PN: Inheritance of susceptibility to multiple sclerosis. *Curr Opin Immunol* 1998; **10**: 697–703.
- 4 Olerup O, Hillert J: HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens* 1991; **38**: 1–15.
- 5 Ligiers A, Dyment DA, Willer CJ *et al*: Evidence of linkage with HLA-DR in DRB1*15-negative families with multiple sclerosis. *Am J Hum Genet* 2001; **69**: 900–903.
- 6 Luomala M, Lehtimäki T, Elovaara I *et al*: A study of interleukin-1 cluster genes in susceptibility to and severity of multiple sclerosis. *J Neurol Sci* 2001; **185**: 123–127.
- 7 He B, Xu C, Yang B, Landtblom AM, Fredrikson S, Hillert J: Linkage and association analysis of genes encoding cytokines and myelin proteins in multiple sclerosis. *J Neuroimmunol* 1998; **86**: 13–19.
- 8 Wansén K, Pastinen T, Kuokkanen S *et al*: Immune system genes in multiple sclerosis: genetic association and linkage analyses on TCR beta, IGH, IFN-gamma and IL-1ra/IL-1 beta loci. *J Neuroimmunol* 1997; **79**: 29–36.
- 9 Ebers GC, Kukay K, Bulman DE *et al*: A full genome search in multiple sclerosis. *Nat Genet* 1996; **13**: 472–476.
- 10 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.
- 11 Kuokkanen S, Gschwend M, Rioux JD *et al*: Genomewide scan of multiple sclerosis in Finnish multiplex families. *Am J Hum Genet* 1997; **61**: 1379–1387.
- 12 Haines JL, Ter-Minassian M, Bazyk A *et al*: A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. The Multiple Sclerosis Genetics Group. *Nat Genet* 1996; **13**: 469–471.
- 13 Meta-analysis of genomic screens in multiple sclerosis. The Transatlantic Multiple Sclerosis Genetics Cooperative. *Mult Scler* 2001; **7**: 3–11.
- 14 de la Chapelle A, Wright FA: Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. *Proc Natl Acad Sci USA* 1998; **95**: 12416–12423.
- 15 Houwen RH, Baharloo S, Blankenship K *et al*: Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet* 1994; **8**: 380–386.
- 16 Puffenberger EG, Kauffman ER, Bolk S *et al*: Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. *Hum Mol Genet* 1994; **3**: 1217–1225.
- 17 Freimer NB, Reus VI, Escamilla MA *et al*: Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. *Nat Genet* 1996; **12**: 436–441.
- 18 Binzer M, Forsgren L, Holmgren G, Drugge U, Fredrikson S: Familial clustering of multiple sclerosis in a northern Swedish rural district. *J Neurol Neurosurg Psychiatry* 1994; **57**: 497–499.
- 19 Dubovsky J, Sheffield VC, Duyk GM, Weber JL: Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. *Hum Mol Genet* 1995; **4**: 449–452.
- 20 Dib C, Faure S, Fizames C *et al*: A comprehensive genetic map of the human genome based on 5264 microsatellites. *Nature* 1996; **380**: 152–154.
- 21 Carlsson B, Wallin J, Bohme J, Moller E: HLA-DR-DQ haplotypes defined by restriction fragment analysis. Correlation to serology. *Hum Immunol* 1987; **20**: 95–113.
- 22 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
- 23 Fogdell A, Olerup O, Fredrikson S, Vrethem M, Hillert J: Linkage analysis of HLA class II genes in Swedish multiplex families with multiple sclerosis. *Neurology* 1997; **48**: 758–762.
- 24 Risch N: Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 1990; **46**: 229–241.
- 25 Hastbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E: Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nat Genet* 1992; **2**: 204–211.
- 26 Lehesjoki AE, Koskineniemi M, Norio R *et al*: Localization of the EPM1 gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. *Hum Mol Genet* 1993; **2**: 1229–1234.
- 27 Copeman JB, Cucca F, Hearne CM *et al*: Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (IDDM7) to chromosome 2q31-q33. *Nat Genet* 1995; **9**: 80–85.
- 28 Terwilliger JD, Weiss KM: Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol* 1998; **9**: 578–594.
- 29 Durham LK, Feingold E: Genome scanning for segments shared identical by descent among distant relatives in isolated populations. *Am J Hum Genet* 1997; **61**: 830–842.
- 30 Hillert J: Human leukocyte antigen studies in multiple sclerosis. *Ann Neurol* 1994; **36** (Suppl.): S15–S17.