Journal of NeuroVirology (2000) 6, Suppl 2, S5–S9

www.jneurovirol.com

The genetics of multiple sclerosis

Alastair Compston*,[1]

[1]*University of Cambridge Neurology Unit, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK*

Epidemiological studies implicate an interplay between genetic and environmental factors in the aetiology of multiple sclerosis. The classical genetic observations suggest that multiple sclerosis is a complex trait in which susceptibility is determined by several genes acting independently or epistatically. The main dividend from understanding the genetic basis of susceptibility in multiple sclerosis will be an improved understanding of the pathogenesis. To date, candidate gene approaches have proved relatively unrewarding other than in establishing the association with alleles of the major histocompatibility complex (MHC). In common with most other complex traits, no major susceptibility gene has been identified through full genome screens but regions of interest have provisionally been identified. An important part of future studies in the genetics of multiple sclerosis will be to resolve the question of disease heterogeneity. *Journal of NeuroVirology* (2000) **6,** S5–S9.

Keywords: genetics; multiple sclerosis; candidates; genome search

Introduction

Progress in characterising the extent to which genetic factors contribute to the aetiology of multiple sclerosis has depended on classical studies involving the analysis of pedigrees and recurrence risk within families, and molecular investigation either of candidate genes or the whole genome using microsatellite and single nucleotide polymorphisms.

Pedigree analysis

Families can be used in several ways to explore the aetiology of multiple sclerosis. In planning a genetic study in multiple sclerosis, the usual practice is to retrieve potentially informative cases from those already known to be affected. This requires the co-ordination of lists from various sources. For surveys testing biological features, such as genetic susceptibility in families, the trend is towards inclusion of individuals who have the disease process even if this does not yet meet standard criteria for clinically definite diagnosis. For example, in surveying the concordance of multiple sclerosis in twins the aim is to identify all individuals having the disease process irrespective of whether this has reached some arbitrary point of clinical expression. However, if the criteria for inclusion are too relaxed, errors of inclusion will contaminate the results. It is necessary to be aware of potential ethnic differences within or between populations when planning or comparing studies of genetic susceptibility.

Genetic epidemiological studies providing morbidity statistics for multiple sclerosis have been used over many years to generate hypotheses for the aetiology of the disease (for review, see Compston, 1998). Multiple sclerosis is common in northern Europeans but not Orientals, Asians or Africans. There is a familial recurrence rate of approximately 15%. A meta-analysis of recurrence risk amongst 44 177 relatives of 2163 probands from three population based series (Sadovnick *et al*, 1988; Robertson *et al*, 1996; Carton *et al*, 1997) shows that the age-adjusted risk is highest for siblings (3%), then parents (2%) and children (2%) with lower rates in second and third degree relatives. Recurrence in monozygotic twins is around 35% (Sadovnick *et al*, 1993; French Research Group on Multiple Sclerosis, 1992; Mumford *et al*, 1994). Conversely, the frequency of multiple sclerosis in non-biological parents, siblings and children is more or less identical to the population lifetime risks for Europeans, and significantly lower than those expected from the study of recurrence in the biological relatives of index cases (Ebers *et al*, 1995). The age-adjusted risk for half-siblings is also less than for full siblings and no difference is seen between half-siblings reared together and apart (Sadovnick *et al*, 1996). Recurrence is higher in the children of conjugal pairs with multiple sclerosis (6%; age-adjusted 20%) than the offspring of single parents with multiple sclerosis (2%: Robertson *et al*, 1997).

*Correspondence: A Compston

Strategies for identifying susceptibility genes

Originally, the search for genes determining susceptibility to complex disease was organised by comparing the frequency of polymorphisms at a given genetic locus in groups of unrelated individuals with and without the trait. A change in strategy for studying the genetics of complex traits became routine over the last decade with the re-application of previously described family based methods (Risch, 1990). The biased inheritance of genomic regions presumed to encode susceptibility genes for multiple sclerosis in affected individuals forms the basis for family-based linkage and association studies.

Family based linkage methods have been applied most energetically to sibling pairs. The probability that, for a given locus, each affected sibling will share neither, one or both parental alleles is 0.25, 0.50 and 0.25, respectively. Linkage is present if, in a sufficient sample size, that distribution is distorted in favour of one or two sharing. The evidence relating to a given locus can be expressed as the logarithm of the odds (lod) for a gene being encoded at that site (the maximum lod score) or, conversely, as the probability that such a gene is not there encoded. Linkage is most efficient for detecting genes exerting major effects.

Many markers at adjacent loci across the genome are still in linkage disequilibrium. This means that association studies can be informative even when the marker is up to 0.5 centiMorgans (cM) from a disease susceptibility locus. Transmission disequilibrium testing (TDT) uses trios consisting of a single person with multiple sclerosis and both parents (who usually are but do not need to be unaffected). The aim is to show that in a sufficient sample size, alleles encoded at a particular locus are transmitted to the affected individual more (or less) often than expected by chance. The efficiency of TDT is determined by the extent of linkage disequilibrium.

Linkage and association provide different types of information. Each has its limitations and advantages. Assuming a reasonable sample size, any marker which is well distant from the true susceptibility locus will be linked but not associated. A marker which is close to the susceptibility locus will not be linked in sibling pairs if the sample size is inadequate (and, to date, most sibling pair studies are not) but it may be associated. Only the combination of a close marker tested in an adequate sample will demonstrate both linkage and association. One strategy is first to screen the genome with linkage to establish regions of interest and then to refine the map with transmission disequilibrium testing.

Candidate genes

Population studies demonstrate an association between the class II MHC alleles DR15 and DQ6 and their corresponding genotypes DRB1*1501, DRB5*0101 and DQA1*0102, DQB2*0602 (Olerup and Hillert, 1991) and the gene for TNF-α is usually also associated (Coraddu *et al*, 1998) since this is encoded within the same linkage group. A specifically different association is seen in Mediterranean populations. For example, in Sardinians, multiple sclerosis is associated with DR4 [DRB1*0405-DQA1*0301-DQB1*0302] (Marrosu *et al*, 1992) although the DR15 association is seen amongst patients from all other parts of Italy in which adequate studies have been performed (La Mantia *et al*, 1990). In the Canaries, the disease is primarily associated with DR15 and DQ6 but a secondary association exists with DR4 (Coraddu *et al*, 1999). In Turkey, there is also an allelic association with both DR2 (DBR1*1501, DQA1*0102, DQB1*0602) and DR4 (DRB1*04, DQA1*03, DQB1*0302: Saruhan-Direskeneli *et al*, 1997).

Extensive searches, using association and linkage studies over many years, have only yielded additional putative candidate genes in the VH2-5 immunoglobulin heavy chain and the T cell receptor β chain variable regions. Recent contributions to this aspect of the genetics of multiple sclerosis have tended to illustrate the difficulty which exists in confirming weak associations (VH2-5; Wansen *et al*, 1997; Ligers *et al*, 1997: and TCR-β; Droogan *et al*, 1996; Wansen *et al*, 1997). Studies of linkage or association with the genes encoding ICAM-1 (Mycko *et al*, 1998), IL-1ra (De La Concha *et al*, 1997; Wansen *et al*, 1997), IL-1β and IFN-gamma (Wansen *et al*, 1997), IL-4 (Vanderbroeck *et al*, 1997) and CCR5 (Bennets *et al*, 1997) have all been negative. Except in the isolated population of Finns where there is both an association and linkage to the gene for myelin basic protein (Tienari *et al*, 1992, 1998), studies of structural genes of myelin have also been uninformative (Nellemann *et al*, 1995; Thompson *et al*, 1996; Price *et al*, 1997; Rodriguez *et al*, 1997; He *et al*, 1998a). Some investigators have explored new candidates. Using a variety of individuals (34 families, 147 cases and 95 controls), a Scandinavian group failed to implicate any one of several candidates encoding cytokines (IFN-gamma, IL-2, IL-4, IL-10, TGF-β1 and -β2, and IL-4R) or myelin proteins (PLP, MAG, OMGP, CNPase) – although IFN-gamma (12q24.1), IL-4R (16p12.1), and TGF-β2 (1q41) could not be fully excluded (He *et al*, 1998b). In France, efforts have been directed at growth factors which determine oligodendrocyte development, and their receptors (Reboul *et al*, 1999). Again, no evidence for linkage emerges but these may also be type 2 errors due to the small sample size.

Genome screens

Four groups of investigators have undertaken a systematic search of the genome in an attempt to locate additional susceptibility genes using affected family members – using identity by descent analysis in sibling pairs. Genotyping was completed on cohorts each of between 21 – 225 families, together

involving in excess of 1000 individuals, for each of between 257–443 microsatellite markers. These were chosen to have an average spacing of around 10 centiMorgans, giving enough power to identify regions encoding a major susceptibility gene; and they are sufficiently polymorphic to make a high proportion of the available families fully informative. Although linkage analysis has revealed several new genomic regions which may encode genes conferring susceptibility to multiple sclerosis, some of these will be true and others false positives. Superficially, the results show a disappointing lack of overlap. In common with most other complex traits, no major susceptibility gene has been identified: the possible reasons are that no such gene exists, it has been missed by all three groups, or genetic complexity (i.e. heterogeneity) has obscured the picture. The importance of the major histocompatability complex is confirmed but of the other new putative susceptibility loci, several are clearly unique to each screen and so may be false positives. The regions of interest emerging from the UK genome screen (Sawcer *et al*, 1996; Chataway *et al*, 1998) were 1cen, 5cen, 6p, 7p, 14q, 17q, 19q, Xp. They were 2p, 3p, 5p, 11q and Xp in the Canadian series (Ebers *et al*, 1996). The US/French consortium identified 6p, 7q, 11p, 12q and 19q (Multiple Sclerosis Genetics Group, 1996). There were no statistically significant regions of interest in the relatively small Finnish screen (Kuokkanen *et al*, 1997) although positive lod scores were obtained for 6p21 (MHC) and 5p14-p12; increasing the density of markers raised the lod scores in several other regions (4cen, 11tel and 17q) whereas others (2q32 and 10q21) were unchanged; when all 21 families were typed across the regions of interest, the highest lod score (2.8) was at 17q22-q24 – as in the previously reported UK screen.

In a subsequent study from Scandinavia (Xu *et al*, 1999), 15 markers from regions of interest identified in one or more full genome screens were assessed using affected relative pairs (from 46 multiplex families, 28 trios, and 190 cases and 148 unrelated controls): depending on the method of analysis positive results were obtained for 5p15, 6p21, 7ptr15 and 12q23; the latter two also showed transmission distortion in families and an association in the case-control series. The remaining markers (2p23, 5p, 5q, 6q22-25, 7q, 11q21-23, 12q24-qtr, 13q33-34, 16p12, 18p11 and Xp21) did not identify regions of interest in the Scandinavian series. An Italian group used 67 markers covering regions of interest from the three full genome screens and candidates (HLA-DRB1, CTLA-4, IL-9, CSF-1R, APOE, Bcl-2 and TNF-R2) in 69 multiplex families (28 from Sardinia and 41 from mainland Italy). Some positive scores were obtained especially for the region at 5q14 in the Sardinian subset. Chromosome 5 has now featured in all the family studies but the region is broad and it may be difficult to assess in more detail. Three of these regions (2p11.2, 7p15.2 and 17q12) also showed positive TDT scores (D'Alfonso *et al*, 1999).

These genetic analyses are predicated on the assumption that multiple sclerosis is one disease. Mutations of mitochondrial DNA are responsible for a multiple sclerosis-like illness characterised by disproportionate involvement of the anterior visual pathway (Harding *et al*, 1992; Riordan Eva *et al*, 1995) although mitochondrial genes do not contribute generally to susceptibility in multiple sclerosis (Kellar Wood *et al*, 1994; Kalman *et al*, 1995; Nishimura *et al*, 1995). Conditioning the United Kingdom genome screen DR15 (or an extended DR15 linked haplotype also encoding alleles of TNF and the DQ locus) shows that the regions of interest on 1p, 17p, 17q and X cluster in families which are identical by state for DR15 whereas the non-sharing group associated with 1cen, 3p, 5cen, 7p, 14q and 22q (Coraddu *et al*, 1999); in addition new regions of interest are found at 5q and 13p (DR15 sharing) and 16p and 20p (DR15 non-sharing). A major part of future studies in the genetics of multiple sclerosis will be to resolve the question of disease heterogeneity.

Conclusion

The contribution to susceptibility made by the genes which have provisionally been identified only accounts for a proportion of the increased risk of multiple sclerosis implicated by family studies; clearly there is more to be learned concerning susceptibility to multiple sclerosis and the interplay between genes and environmental factors. The application of this knowledge for improved understanding of the pathogenesis, in counselling and in designing novel treatments is potentially considerable.

References

Bennets BH, Teutsch SM, Buhler MM *et al* (1997). The CCR5 deletion mutation fails to protect against multiple sclerosis. *Human Immunol* **58:** 52–59.

Carton H, Vlietinck R, Debruyne J *et al* (1997). Recurrence risks of multiple sclerosis in relatives of patients in Flanders, Belgium. *J Neurol Neurosurg Psychiatry* **62:** 329–333.

Chataway J, Feakes R, Coraddu F *et al* (1998). The genetics of multiple sclerosis: principles, background and updated results of the United Kingdom systematic genome screen. *Brain* **121:** 1869–1887.

Compston DAS, Ebers GC, Lassmann H, McDonald WI, Matthews WB, Wekerle H (1998). *McAlpine's Multiple Sclerosis.* 3rd edn. Compston DAS (ed), WB Saunders: London, pp. 592.

Coraddu F, Reyes-Yanez MP, Aladro Y *et al* (1998). HLA associations with multiple sclerosis in the Canary Islands. *J Neiroimmunol* **87:** 130–135.

Coraddu F, Sawcer S, Feakes R *et al* (1999). HLA typing in the United Kingdom multiple sclerosis genome screen. *Neurogenetics* **2:** 24–33.

D'Alfonso S, Nistico L, Zavattari P *et al* (1999). A linkage analysis of multiple sclerosis with candidate region markers in Sardinian and continental Italian families. *Eur J Human Genet* **7:** 377–385.

De La Concha EG, Arroyo R, Crusius JB *et al* (1997). Combined effect of HLA-DRB1*1501 and interleukin-1 receptor antagonist gene allele 2 in susceptibility to relapsing/remitting multiple sclerosis. *J Neuroimmunol* **80:** 172–178.

Droogan AG, Kirk CW, Hawkins SA *et al* (1996). T cell receptor alpha, beta, gamma and delta chain microsatellites show no association with multiple sclerosis. *Neurology* **47:** 1049–1053.

Ebers GC, Sadovnick AD, Risch NJ (1995). A genetic basis for familial aggregation in multiple sclerosis. *Nature* **377:** 150–151.

Ebers GC, Kukay K, Bulman D *et al* (1996). A full genome search in multiple sclerosis. *Nature Genet* **13:** 472–476.

French Research Group on Multiple Sclerosis (1992). Multiple Sclerosis in 54 twinships: concordance rate is independent of zygosity. *Ann Neurol* **32:** 724–727.

Harding AE, Sweeney MG, Brockington M *et al* (1992). Occurrence of a multiple sclerosis-like illness in women who have a Leber's hereditary optic neuropathy mitochondrial DNA mutation. *Brain* **115:** 989.

He B, Yang B, Lundahl J *et al* (1998a). The myelin basic protein gene in multiple sclerosis: identification of discrete alleles of a 1.3 kb tetranucleotide repeat sequence. *Acta Neurologica Scandinavica* **97:** 46–51.

He B, Xu C, Yang B, Landtblom A-M, Fredrikson S, Hillert J (1998b). Linkage and association analysis of genes encoding cytokines and myelin proteins in multiple sclerosis. *J Neuroimmunol* **86:** 13–19.

Kalman B, Lublin FD, Alder H (1995). Mitochondrial DNA mutations in multiple sclerosis. *Multiple Sclerosis* **1:** 32–36.

Kellar Wood H, Robertson N, Govan GG *et al* (1994). Leber's hereditary optic neuropathy mitochondrial DNA mutations in multiple sclerosis. *Ann Neurol* **36:** 109–112.

Kuokkanen S, Gschwend M, Rioux JD *et al* (1997). Genomewide scan of multiple sclerosis in Finnish multiplex families. *Am J Human Genet* **61:** 1379–1387.

La Mantia L, Illeni MT, Milanese C *et al* (1990). HLA and multiple sclerosis in Italy: a review of the literature. *J Neurol* **237:** 441–444.

Ligers A, He B, Fogdell-Hahn A, Olerup O, Hillert J (1997). No linkage or association of a VNTR marker in the junction region of the immunoglobulin heavy chain genes in multiple sclerosis. *Eur J Immunogenet* **24:** 259–264.

Marrosu MG, Muntoni F, Murru MR *et al* (1992). HLA-DQB1 genotype in Sardinian multiple sclerosis: evidence for a key role of DQB1.0201 and DQB1.0302 alleles. *Neurology* **42:** 883–886.

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

Mumford CJ, Wood NW, Kellar-Wood HF *et al* (1994). The British Isles survey of multiple sclerosis in twins. *Neurology* **44:** 11–15.

Mycko MP, Kwinkowoski M, Tronczynska E *et al* (1998). Multiple sclerosis: increased frequency of the ICAM-1 exon 6 gene point mutation genetic type K469. *Ann Neurol* **44:** 70–75.

Nellemann LJ, Frederiksen J, Morling N (1995). PCR typing of two short tandem repeat (STR) structures upstream of the human myelin basic protein (MBP) gene; genetic susceptibility in multiple sclerosis and monosymptomatic idiopathic optic neuritis in Danes. *Multiple Sclerosis* **1:** 186–189.

Nishimura M, Obayashi H, Ohta M *et al* (1995). No association of the 11778 mitochondrial mutation and multiple sclerosis in Japan. *Neurology* **45:** 1333–1334.

Olerup O, Hillert J (1991). HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens* **38:** 1–15.

Price SE, Sharpe G, Boots A *et al* (1997). Role of myelin basic protein and proteolipid protein genes in multiple sclerosis: single strand conformation polymorphism analysis of the human sequences. *Neuropathol Applied Neurobiol* **23:** 457–467.

Reboul J, Mertens C, Levillayer F *et al* (2000). Cytokines in genetic susceptibility to multiple sclerosis: a candidate gene approach. *J Neuroimmunol* **102:** 107–112.

Riordan Eva P, Sanders MD, Govan GG *et al* (1995). The clinical features of Leber's hereditary optic neuropathy defined by the presence of pathogenic mitochondrial DNA mutation. *Brain* **118:** 319–337.

Risch N (1990). Linkage strategies for genetically complex traits. *Am J Human Genet* **46:** 222–253.

Robertson NP, Fraser M, Deans J, Clayton D, Compston DAS (1996). Age adjusted recurrence risks for relatives of patients with multiple sclerosis. *Brain* **119:** 449–455.

Robertson NP, O'Riordan JI, Chataway J *et al* (1997). Clinical characteristics and offspring recurrence rates of conjugal multiple sclerosis. *Lancet* **349:** 1587–1590.

Rodriguez D, Della Gaspera B, Zalc B *et al* (1997). Identification of a Val 145 Ile substitution in the human myelin oligodendrocyte glycoprotein: lack of association with multiple sclerosis. *Multiple Sclerosis* **3:** 377–382.

Sadovnick AD, Baird PA, Ward RH (1988). Multiple sclerosis; updated risks for relatives. *Am J Medical Genet* **29:** 533–541.

Sadovnick AD, Armstrong H, Rice GPA *et al* (1993). A population-based study of multiple sclerosis in twins: update. *Ann Neurol* **33:** 281–285.

Sadovnick AD, Ebers GC, Dyment DA, Risch N, the Canadian Collaborative Study Group (1996). Evidence for genetic basis of multiple sclerosis. *Lancet* **347:** 1728–1730.

Saruhan-Direskeneli G, Esin S, Baykan-Kurt B *et al* (1997). HLA-DR and -DQ associations with multiple sclerosis in Turkey. *Human Immunol* **55:** 59–65.

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

Thompson RJ, Mason CR, Douglas AJ *et al* (1996). Analysis of polymorphisms of the 2′3′-cyclic nucleotide-3′-phosphodiesterase gene in patients with multiple sclerosis. *Multiple Sclerosis* **2:** 215–221.

Tienari P, Wikstrom J, Sajantila A, Palo J, Peltonen L (1992). Genetic susceptibility to multiple sclerosis linked to myelin basic protein gene. *Lancet* **340:** 987–991.

Tienari P, Kuokkanen S, Pastinen T *et al* (1998). Golli-MBP gene in multiple sclerosis. *J Neuroimmunol* **81:** 158–167.

Vanderbroeck K, Martino G, Marrosu M *et al* (1997). Occurrence and clinical relevance of an interleukin-4 gene polymorphism in patients with multiple sclerosis. *J Neuroimmunol* **76:** 189–192.

Wansen K, Pastinen T, Kuokkanen S, Wokstrom J *et al* (1997). Immune system genes in multiple sclerosis: genetic association and linkage analyses on TCR-β, IgH, IFN-gamma and IL-12ra/IL-1β loci. *J Neurol* **79:** 29–36.

Xu C, Dai Y, Fredrikson S, Hillert J (1999). Association and linkage analysis of candidate chromosomal regions in multiple sclerosis: indication of disease genes in 12q23. *Eur J Human Genet* **7:** 110–116.