

The genetic aspects of multiple sclerosis

Stephen Sawcer

University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK

Abstract

The epidemiology of multiple sclerosis has been extensively investigated and two features have consistently emerged: marked geographical variation in prevalence and substantial familial clustering. At first sight, geographic variation would seem to imply an environmental cause for the disease, while familial clustering would seem to suggest that genetic factors have the predominant etiological effect. However, given that geographic variation in prevalence could result from variation in the frequency of genetic risk alleles and that familial clustering might result from shared environmental exposure rather than shared genetic risk alleles, it is clear that these crude inferences are unreliable. Epidemiologists have been resourceful in their attempts to resolve this apparent conflict between "nurture and nature" and have employed a whole variety of sophisticated methods to try and untangle the etiology of multiple sclerosis. The body of evidence that has emerged from these efforts has formed the foundation for decades of research seeking to identify relevant genes and this is the obvious place to start any consideration of the genetics of multiple sclerosis.

Key Words

Genetics, genome-wide association study, multiple sclerosis

For correspondence:

Stephen Sawcer, University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK. E-mail: sjs1016@mole.bio.cam.ac.uk

Ann Indian Acad Neurol 2009;12:206-14

Epidemiology

Although prevalence studies are difficult to perform and have an inherent tendency to underestimate the true frequency of disease, many such studies have been completed in multiple sclerosis and, together, they provide an almost global map of the distribution of the disease.^[1] The core feature of this distribution is often summarized as a latitudinal gradient, with the disease being common in temperate regions and rare in tropical climes [Figure 1].^[2] Some authors have pointed out that the observed distribution also reflects the migration pattern of Northern Europeans, with the disease being common in those parts of the world in which Northern Europeans have settled.^[3] In considering the global distribution of multiple sclerosis it is also important to note that there are exceptions to this general latitudinal gradient: some populations having rates of disease that are significantly different from that seen in their geographically nearby neighbors.^[4] Studies looking at the risk of disease in migrant individuals moving from regions of low risk to regions of high risk, and vice versa^[5-9] [Figure 1], suggest that risk changes with migration. These studies provide some of the clearest evidence supporting the role of environmental factors in the etiology of the disease. In any given part of

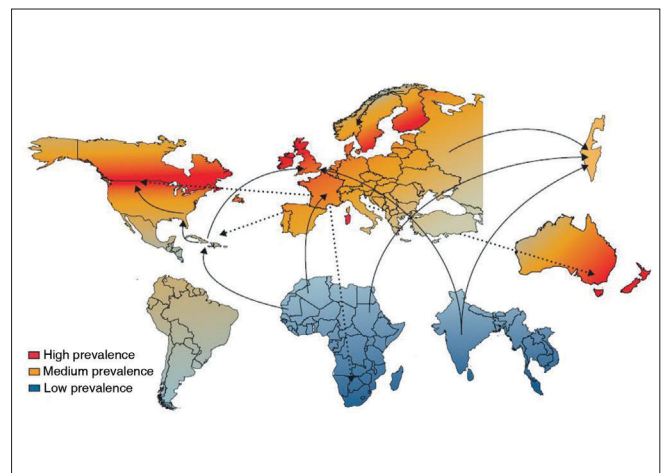


Figure 1: Global distribution of multiple sclerosis and migrations
The five continents are depicted, showing areas of medium prevalence of multiple sclerosis (orange), areas of exceptionally high frequency (red), and areas with low rates (grey-blue). Some regions are largely uncharted so these colors are only intended to provide an impression of the geographical trends. Major routes of migration studied from high-risk zone of northern Europe are shown as dotted arrows. Studies involving migrants from low-risk to high-risk zones are shown as solid arrows. Reprinted from *The Lancet*, Vol. 372, Compston and Coles, *Multiple sclerosis, 1502-1517*, Copyright (2008), with permission from Elsevier.^[2]

the world, the risk of multiple sclerosis seems to be greatest in white individuals of European descent.

Amongst white individuals living in temperate regions such as Europe, Canada, and North America, 15–20% of those affected by multiple sclerosis report a family history of the disease—a rate which is significantly greater than would be expected given the prevalence of multiple sclerosis in these regions (typically 1 per 1,000).^[1] Population-based studies of familial recurrence risk have confirmed and quantified the increased risk of the disease in the relatives of affected individuals.^[10–12] This familial clustering can be usefully summarized as λ_s – the relative risk of the disease seen in the siblings of affected individuals as compared to the risk seen in the general population.^[13] In multiple sclerosis this risk ratio takes a value of approximately 15^[14] and, of course, reflects the combined effects of all shared etiological influences, both genetic and environmental.^[15] In order to tease these influences apart, investigators have studied a variety of informative subgroups, including twins,^[16–20] adoptees,^[21] conjugal pairs,^[22,23] half-siblings,^[24] and step-siblings.^[25] Taken together these data suggest that living with someone who has, or who will eventually develop, multiple sclerosis has little or no effect on one’s risk of developing the disease unless one is genetically related to that person, in which case one’s risk of having the disease increases with the degree of relatedness.^[25] This is not to imply that environmental factors have no role, only that they seem to exert their effects mainly at a population level, with the micro-environmental differences between

families within a given population seeming to be of relatively little importance.^[25]

The epidemiology of multiple sclerosis continues to be studied, and interesting twists suggesting unsuspected possibilities will no doubt emerge in the future. However, as these studies are difficult to perform, are frequently subject to confounding and bias, and are also generally underpowered, it seems unlikely that they will ever provide any major insights. That said, it has been pointed out that the variation in recurrence risk with the degree of genetic relatedness can provide a useful indication of the genetic architecture underlying susceptibility to a complex disease.^[26] In multiple sclerosis, these data [Figure 2] indicate that perhaps 100 common variants (those with a minor allele frequency of more than a few percent), each exerting only a modest effect on risk (increasing the odds of developing the disease by a factor of approximately 1.2–1.3) are likely to be involved.^[27,28]

Major Histocompatibility Complex

Early attempts to identify susceptibility genes in multiple sclerosis were highly successful and quickly identified the now well-established relevance of the Major Histocompatibility Complex (MHC) on chromosome 6p21. The first successes emerged in the early 1970s, when investigators showed that multiple sclerosis was

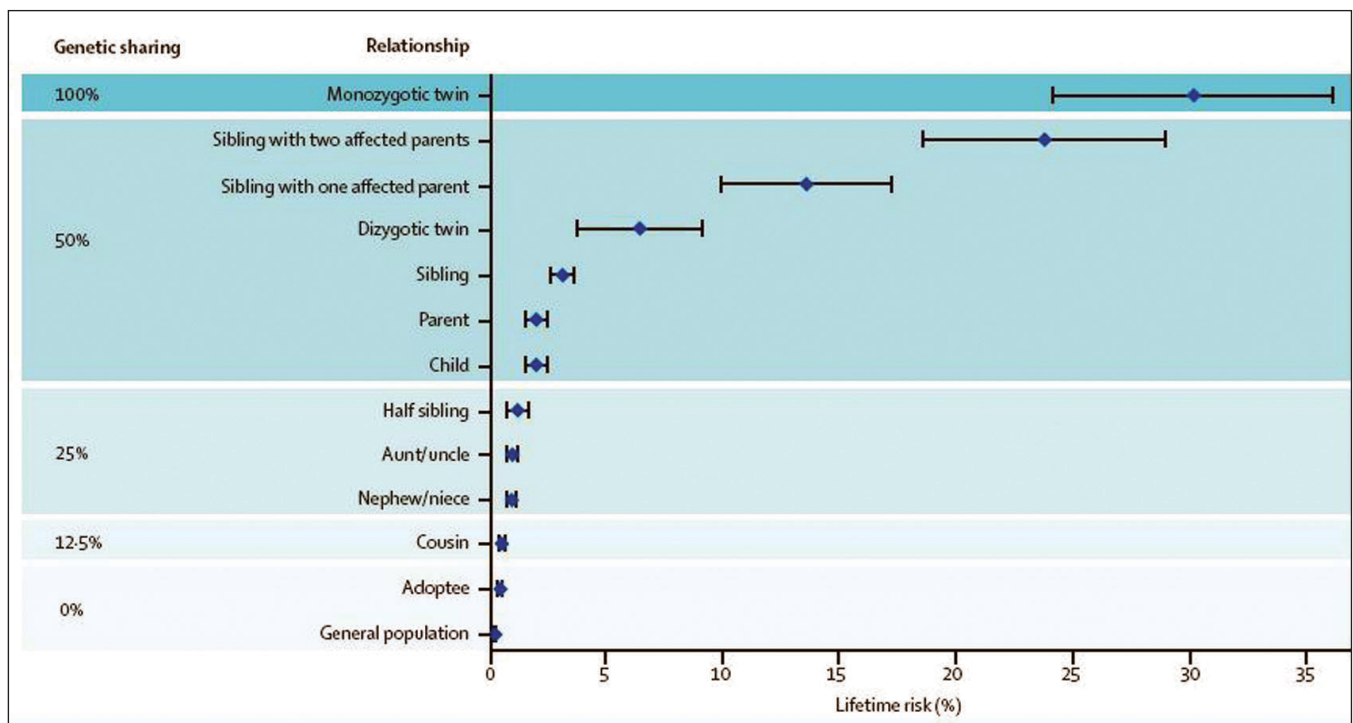


Figure 2: Recurrence risks for multiple sclerosis in families
 Age-adjusted recurrence risks for different relatives of probands with multiple sclerosis and the degree of genetic sharing between relatives and the proband. Pooled data from population-based surveys. Error bars indicate 95% confidence intervals. Reprinted from *The Lancet*, Vol. 372, Compston and Coles, *Multiple sclerosis*, 1502-1517, Copyright (2008), with permission from Elsevier.^[2]

associated with various Human Leukocyte Antigens (HLA), in particular HL-A3,^[29] HL-A7,^[30] and LD-7a^[31] (in today's nomenclature HLA-A3, HLA-B7, and DR2, respectively). It was quickly realized that these associations were not independent but were, rather, a reflection of linkage disequilibrium (LD; the tendency for certain alleles from linked loci to occur together on the same chromosome more often than you would expect by chance alone) between the corresponding alleles.^[32,33] Over the years, technology has improved and the resolution of the associated alleles has been refined.^[34-36]

The MHC is a gene-dense region of the genome that is characterized by extensive LD and extreme levels of polymorphism.^[37] In light of these features, it is not surprising to find that many variants from this region show association with multiple sclerosis as a result of LD (correlation) between these various associated alleles.^[38,39] The modest levels of LD between the class I region (containing the *HLA-A* and *HLA-B* genes) and the class II region (containing the *DRB1* and *DQB1* genes) enabled researchers to quickly establish that association primarily derived from the class II region.^[32,33] However, the more extensive LD between *DRB1* and *DQB1* made it much more difficult to refine which of these genes was primarily responsible for the association. Studying African American patients, who have less intense LD between *DRB1**1501 and *DQB1**0602, Oksenberg *et al.*^[40] provided the first convincing evidence that the primary association was with the *DRB1* gene, an observation which has been confirmed in subsequent studies in large cohorts of patients of European descent.^[39]

Our understanding of the association between multiple sclerosis and the MHC continues to improve. It is now well established that MHC haplotypes other than just those carrying the *DRB1**1501 allele also exert influence on susceptibility, either independently^[41] or by modifying the risk associated with *1501.^[42-44] What is unclear is whether or not these additional signals stem primarily from the *DRB1* gene or from the effects of other MHC loci. Many researchers have found evidence supporting the existence of an independent signal from the class I region,^[39,41,45-48] while others have not.^[38,49] Establishing the presence of additional susceptibility loci located close to a primary locus is a complex problem, especially in the presence of prominent LD and likely allelic heterogeneity.^[50] The necessity to correct for the effects of LD with the *DRB1**1501 allele limit the power of studies based on white European populations in which this allele is common.^[38,39,42,43] Despite these difficulties, a role for the *DRB1**03 haplotype is now well established^[39,42,43] and Sardinian data would suggest that this association, at least, most likely stems from the *DRB1* gene.^[41] The nature and origin of the remaining signals from the MHC region remain unresolved.

The role of HLA genes has also been studied in Asian multiple sclerosis. Serological studies performed in the 1980s found association with HLA-B12^[51,52] rather than with B7, while molecular genetic analysis of the class II HLA genes, *DRB1*, *DQA1*, and *DQB1*, revealed the expected association with the European susceptibility haplotype *DRB1**1501 –*DQB1**0602.^[53] In Asians, the *DRB1**1502 and *DRB1**16 subtypes of DR2 are more common than in Europeans and the extent of LD between *DRB1**1501 and *DQB1**0602 is significantly less intense.^[54] In an analysis of Indian migrants living in the US, Rosenberg *et al.*^[55] found that there was surprisingly little variation in ancestry across India, although there was a clear difference between Indian Asians and other ethnic groups.

By cataloguing variation in the MHC through the resequencing of specific haplotypes^[56] and empirically establishing the complex patterns of LD across the region,^[57] it has been possible to establish a comprehensive panel of haplotype-tagging Single-Nucleotide Polymorphisms (SNPs).^[58] These SNPs are currently being typed in multiple sclerosis and a number of other autoimmune diseases as part of the International MHC and Autoimmunity Genetics Network (IMAGEN) project. Hopefully these systematic fine-mapping efforts will help to unravel this complex association, although it can be anticipated that large sample sizes will be needed to confirm the findings emerging from IMAGEN.

Other than the findings related to the MHC, the genetic analysis of multiple sclerosis has only very recently started to yield to the efforts of researchers. Prior to 2007, the lack of any convincing progress was a source of great frustration, and the inconsistency in early claims was rightly criticized.^[59] It is now clear that two main issues have confounded the identification of relevant genes – the modest size of effects attributable to individual loci^[60] and the failure to correctly allow for the statistical consequences that result from the enormous size of the genome.^[61] The search for the genes of relevance in multiple sclerosis has been likened to searching for a handful of rather small needles in a very large haystack.^[62,63]

Linkage studies

Although available epidemiological data confirms that genetic factors are unequivocally relevant in multiple sclerosis, large extended families with multiple affected individuals in multiple generations are extremely uncommon. Most families contain no more than two or three affected individuals and no clear mode of inheritance can be inferred.^[1] Parametric linkage analysis of the few larger families that have been described^[64-68] has failed to identify any rare penetrant alleles that influence the risk of developing multiple sclerosis.

Since large extended families are not available for study, researchers have relied on nonparametric methods of linkage analysis based on looking for excess sharing of alleles amongst related individuals, most typically affected sibling pairs. In 1996, the results from the first attempts to systematically screen the genome for linkage to multiple sclerosis were reported in back-to-back publications from three populations – the UK,^[69] the US,^[70] and Canada.^[71] Each of these studies was based on approximately 100 affected sib pairs and employed 300–400 microsatellite markers. Subsequently, similar studies were reported from Finland,^[72] Sardinia,^[73] Italy,^[74] Scandinavia,^[75] Australia,^[76] and Turkey^[77] and, in addition, each of the original three groups extended their analysis using further families and more microsatellite markers.^[78–80] Disappointingly, none of these studies identified any statistically significant linkage, not even in the region of the MHC. Attempts at meta-analysis of these linkage data were no more successful, although linkage to the MHC just reached genome-wide significance in some of these studies.^[81,82] In order to compensate for the inadequacies inherent in low-resolution microsatellite-based studies^[83] the International Multiple Sclerosis Genetics Consortium (IMSGC) typed 4,506 SNPs in 730 multiplex families from Australia, Scandinavia, the US, and the UK. The enhanced power provided by this SNP-based screen is evident from the overwhelming evidence for linkage found in the MHC region, where a lod score of 11.7 was observed.^[84] Once again, however, no other region of statistically significant linkage was identified. The comprehensive marker map used in this study makes it virtually impossible that any signals of a magnitude similar to that attributable to the MHC could have been missed. As with the previous studies, the number of suggestive linkage peaks was significantly greater than would have been expected to occur by chance alone,^[84] indicating that there is excess allele sharing but providing no clear guide as to the location of the relevant genes.

Although these linkage data provide no useful information concerning the location of non-MHC susceptibility loci, the observed allele sharing does provide useful guidance concerning the size of effects attributable to such loci.^[13] Employing the approach suggested by Risch and Merikangas,^[85] and remembering that the observed allele sharing is expected to provide a significantly inflated estimate of effect size,^[86] it is a straightforward matter to show that common non-MHC risk alleles are highly unlikely to increase the risk of the disease by more than a factor of 2. Under these circumstances it is clear that further linkage analysis is almost certain to be unrewarding since the number of sib pair families necessary to demonstrate significant linkage is impractically large.^[85] These extensive linkage efforts are consistent with the findings from segregation analysis

in suggesting that common risk alleles in multiple sclerosis are highly unlikely to exert more than a modest individual effect on risk. Fortunately, association-based studies are significantly more powerful and thus provide a means to identify genes exerting effects that fall below the resolution of linkage analysis.^[85] However until recently studies looking at candidate genes in multiple sclerosis generally only considered modest numbers of cases, with the result that most of these studies were seriously underpowered to detect effects of the size now realized to be relevant. On the other hand this also means that there are no loci investigated prior to 2007, with the possible exception of APOE,^[87] where published studies have been adequately powered to confidently exclude the possibility of a meaningful effect. It seems highly likely that many of the entirely plausible candidates considered to have been excluded on the basis of the absence of any consistent evidence to date will eventually emerge as genuinely relevant in the disease. It is surely the virtual absence of any power that is responsible for nearly all the apparent inconsistency in the literature concerning the genetics of multiple sclerosis.^[88]

Association studies

In the human population as a whole the total number of genetic variants runs into billions. However, most of these are exceedingly rare^[89] and only around 10 million have a minor allele frequency (MAF) of greater than a few percent.^[89,90] Although these common variants represent only a small fraction of the total number, they are responsible for 90% of the genetic difference between any two individuals.^[26] Given that segregation analysis suggests that around 100 of these 10 million common variants might influence susceptibility to multiple sclerosis (see above) it is clear that the odds that any randomly chosen common variant is relevant in multiple sclerosis are approximately 100,000 to 1 against. These very low prior odds have a profound influence upon the odds that any nominally significant result identified in an association study is a true positive rather than false positive (the posterior odds).^[91]

From Figure 3 it is obvious why *P*-values in the range of 5% to 0.1% will nearly always be false positives, even in well-powered studies. Importantly, Figure 3 also shows that the lower the power of a study, the more extreme the *P*-value must be before any result becomes more likely to be true rather than false.^[93] The Wellcome Trust Case Control Consortium (WTCCC) proposed this Bayesian approach and through similar reasoning suggested that association studies in complex diseases should employ a minimum of 2,000 cases and that in this setting *P*-values of $< 5 \times 10^{-7}$ would more likely be true than false.^[93]

One way to improve the likelihood of success in an association study is to use additional sources of

information, such as animal models or expression data, to guide the selection of variants and thereby improve the prior odds. This candidate gene approach formed the bedrock of association studies in multiple sclerosis prior to the advent of Genome-Wide Association Studies (GWAS). However, even when multiple sources of alternate information are employed in a systematic way (so-called genomic convergence^[94]) it is unlikely that the prior odds can be reduced below 1000:1.^[91] In 2005, Fernald *et al.*^[95] applied genomic convergence to multiple sclerosis and identified a short list of candidate genes, including the interleukin 7 receptor (*IL7R*), a locus that had previously been considered by Australian^[96] and Swedish researchers.^[97] In a follow-up study involving collaborators from the UK and Belgium, these US researchers studied 197 multiplex families, 1,901 trio families, 1,515 cases and 3,204 unrelated controls and identified a highly significant association ($P = 2.9 \times 10^{-7}$) with rs6897932 a non-synonymous SNP (nsSNP)

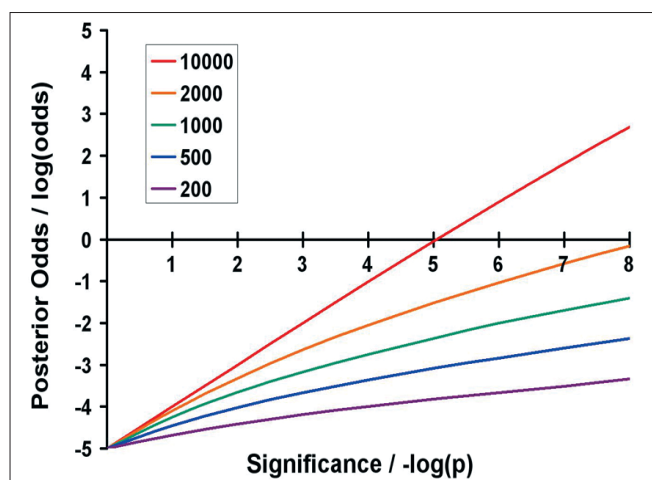


Figure 3: Posterior odds that a result is true, assuming risk alleles with a frequency of 10% and a genotype relative risk (GRR) of 1.2 and a multiplicative model

This figure indicates the posterior odds that a result is true (plotted on a log scale on the y-axis) against the significance of the result (plotted as the negative log of the P-value on the x-axis). Five sample sizes are listed in the legend; in each, the number of cases and controls are equal. The 200 line thus indicates the posterior odds for a study involving 200 cases and 200 controls, and so on. Power was calculated using the online genetic power calculator.^[92] Reprinted from *Brain*, Vol. 131, 3118-31, Copyright (2008), with permission from Oxford University Press.^[63]

Table 1: Established multiple sclerosis non-MHC risk alleles

Implicated gene	Associated SNP	RAF/%	Odds ratio
<i>IL7R</i>	rs6897932	72	1.2
<i>IL2R</i>	rs2104286	72	1.3
<i>CLEC16A</i>	rs12708716	65	1.2
<i>CD226</i>	rs763361	47	1.1
<i>CD58</i>	rs2300747	89	1.3
<i>GPC5</i>	rs9523762	35	1.3

RAF = Risk Allele Frequency / %

from exon 6 of the *IL7R* gene,^[98] a result which was replicated in an accompanying paper from Swedish researchers.^[99] This locus was thus the first non-MHC susceptibility gene to be identified in multiple sclerosis.

It is worth pausing to consider the nature of this association. The multiple sclerosis-associated allele of rs6897932 has a frequency of 72%, which means that approximately 9 out of every 10 white Europeans carry this risk allele. The allele is estimated to increase the risk of the disease by a factor of just 1.2. Using these parameters, we can calculate the P-value that would be expected in studies attempting to replicate this finding as a function of the sample size considered, as shown in Figure 4.

It is clear from Figure 4 that studies attempting to replicate association with effects of this size need to involve at least 2,000 cases and 2,000 controls if they are to have > 95% power to demonstrate nominally significant P-values (5%). Most attempts at replication involving more than 600 cases and 600 controls can be expected to yield a P-value of < 5% but not all. Studies with less than 600 cases and 600 controls are unlikely to identify even nominally significant association. It is important to keep these values in mind when interpreting replication studies. If a study involving just 400 cases and 400 controls fails to identify nominally significant association, this should not be interpreted as evidence that the association is not relevant in the tested population; in fact, this is perhaps the least likely explanation. Given the effect size and allele frequency, we

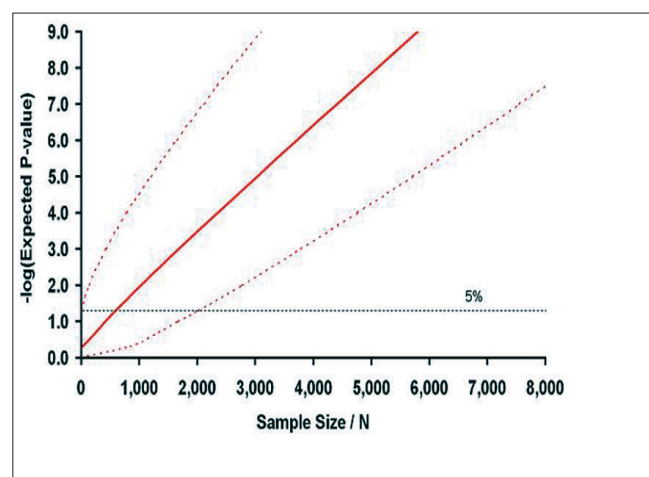


Figure 4: Expected P-value in follow-up studies of rs6897932, the *IL7R*-associated SNP

The red line indicates the expected P-value and the dotted lines the 95% confidence intervals on this estimate (plotted as the negative log). It can thus be expected that 95% of the time the observed P-value will fall within this space. The horizontal dotted line indicates the nominal 5% significance level. Reprinted from *Brain*, Vol. 131, 3118-31, Copyright (2008), with permission from Oxford University Press.^[63]

can also calculate that the lod score that rs6897932 would be expected to generate in a set of 100 sib pairs is just 0.01. It is therefore clear that loci such as rs6897932 would not be expected to generate any linkage signals discernable in previously published linkage screens, and that any apparent concordance between identified susceptibility loci such as *IL7R* and previously reported linkage peaks is entirely coincidental.

Another way to improve the chances of identifying susceptibility genes would be to consider all common variation rather than just a single randomly selected or candidate variant. If all common variation were to be typed in a study, then this study would be sure to include an analysis of the relevant variants. In this situation, concerns about prior odds might be ignored and tests simply interpreted after some correction for multiple testing. However, rather predictably, nothing is gained by adopting this approach to analysis since the statistical penalty required to correct for multiple testing is equivalent to that incurred by allowing for the prior odds.^[100] This is not surprising since both are simply a reflection of the size of the genome. This type of comprehensive (direct) GWAS would seem to be ideal, but in fact the extent of LD between common variants is so extensive that an indirect screen involving just a fraction of the markers relying on LD between tested and untested variants enables a large proportion (typically >80%) of common variation to be screened in a highly efficient manner.^[101] Direct GWAS remains unaffordable and impractical at this time, whereas available technology means that indirect GWAS are possible and have proven to be a highly successful means of identifying the common variants influencing susceptibility to complex diseases.^[90]

GWAS in multiple sclerosis

In 2007, the IMSSC published the first GWAS completed in multiple sclerosis. In this study, 931 trio families (half from the US and half from the UK) were screened using the Affymetrix 500K chip, this yielded usable data from 334,923 SNPs.^[102] The limited power provided by this modest number of trios meant that no unequivocal associations were identified in the screening phase, outside of the expected signals from the MHC. However, by utilizing additional control data from the WTCCC ($n = 1,475$) and the National Institutes of Mental Health (NIMH) ($n = 956$) along with candidate gene information, 110 SNPs were identified and followed up in an additional 609 trio families, 2,322 cases and 2,987 controls. In the final analysis (employing a total of 12,360 individuals), association with rs6897932 (from *IL7R*) was confirmed and significant association was also established with two SNPs from the interleukin-2 receptor (*IL2R*) gene, rs12722489 ($P = 3.0 \times 10^{-8}$), and rs2104286 ($P = 2.2 \times 10^{-7}$).^[102] Studying these three SNPs in additional cohorts from Australia, Belgium,

Denmark, Finland, Germany, Ireland, Italy, Netherlands, Norway, Sardinia, Spain, and Sweden, the IMSSC went on to extensively replicate these findings and were also able to show that the association with rs12722489 was entirely secondary to LD with rs2104286.^[103]

Several investigators have observed an increased risk of a second autoimmune disease in the families of individuals affected by multiple sclerosis.^[104,105] This familial clustering of autoimmune diseases immediately suggests the possibility that some genetic variants might influence the risk of autoimmunity in general and, therefore, that risk alleles for one autoimmune disease might make logical candidates for multiple sclerosis. Following this logic, IMSSC typed seven SNPs established as being relevant in type 1 diabetes in 2,369 trio families, 5,737 cases and 10,296 controls. This analysis confirmed highly significant association with rs12708716 from the C-type lectin domain family 16, member A (*CLEC16A*) gene ($P = 1.6 \times 10^{-15}$) and rs763361 from the *CD226* gene ($P = 5.4 \times 10^{-8}$).^[106] Both these associations were replicated in accompanying papers.^[107,108]

The *CLEC16A* gene (aka KIAA0350) had been identified as potentially associated in the original IMSSC GWAS,^[102] as had *CD58*. Follow-up efforts in *CD58* have also confirmed this association.^[109]

In 2007, the WTCCC also produced a multiple sclerosis GWAS^[110] that was based on 975 cases and 1,466 controls screened with 12,374 nsSNPs. Again, the limited power provided by the cohort size meant that the screen failed to identify any unequivocally associated markers. However, it is relevant to note that rs6897932 was the eighth most associated marker identified, further confirming that a GWAS-based approach would have identified this association had it not already been established through genomic convergence and the candidate gene approach.

The recent publication of the 'Gene MSA' (Multiple Sclerosis Association) GWAS,^[111] the product of a collaboration between US, Dutch, and Swiss researchers (and supported by GlaxoSmithKline), brings the total of multiple sclerosis GWAS to three. This study typed the Illumina HumanHap550 chip in 978 cases and 883 controls and considered extensive phenotyping data, including that from magnetic resonance imaging (MRI). As would be expected, there were no unequivocal associations outside of the MHC, but association with the Glypican-5 (*GPC5*) locus was suggested, and further supported in an additional 974 US cases. Table 1 shows the current list of established non-MHC risk alleles for multiple sclerosis.

The future

Further GWAS in multiple sclerosis are expected to emerge in the near future and will no doubt add to the growing list of established candidate susceptibility genes.

Refining these associations and understanding how these variants exert their effect will engage researchers for some time but seems likely to shed new light on our understanding of this enigmatic disease.^[90]

Note added in proofs a further multiple sclerosis GWAS has been reported since this review was constructed - Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet.* 2009 41:824-8.

References

- Compston A, Confavreux C, Lassmann H, McDonald I, Miller D, Noseworthy J, *et al.* *McAlpine's multiple sclerosis*. 4th ed. London: Churchill Livingstone; 2006.
- Compston A, Coles A. Multiple sclerosis. *Lancet* 2008;372:1502-17.
- Poser CM. The dissemination of multiple sclerosis: A viking saga? A historical essay. *Ann Neurol* 1994;36:S231-43.
- Pugliatti M, Sotgiu S, Rosati G. The worldwide prevalence of multiple sclerosis. *Clin Neurol Neurosurg* 2002;104:182-91.
- Dean G, Kurtzke JF. On the risk of multiple sclerosis according to age at immigration to South Africa. *Br Med J* 1971;3:725-9.
- Alter M, Halpern L, Kurland LT, Bornstein B, Leibowitz U, Silberstein J. Multiple sclerosis in Israel: Prevalence among immigrants and native inhabitants. *Arch Neurol* 1962;7:253-63.
- Detels R, Visscher BR, Malmgren RM, Coulson AH, Lucia MV, Dudley JP. Evidence for lower susceptibility to multiple sclerosis in Japanese-Americans. *Am J Epidemiol* 1977;105:303-10.
- Elian M, Nightingale S, Dean G. Multiple sclerosis among United Kingdom-born children of immigrants from the Indian subcontinent, Africa and the West Indies. *J Neurol Neurosurg Psychiatry* 1990;53:906-11.
- Hammond SR, McLeod JG, Millingen KS, Stewart-Wynne EG, English D, Holland JT, *et al.* The epidemiology of multiple sclerosis in three Australian cities: Perth, Newcastle and Hobart. *Brain* 1988;111:1-25.
- Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: Updated risks for relatives. *Am J Med Genet* 1988;29:533-41.
- Robertson NP, Fraser M, Deans J, Clayton D, Walker N, Compston DA. Age-adjusted recurrence risks for relatives of patients with multiple sclerosis. *Brain* 1996;119:449-55.
- Carton H, Vlietinck R, Debryne J, De Keyser J, D'Hooghe MB, Loos R, *et al.* Risks of multiple sclerosis in relatives of patients in Flanders, Belgium. *J Neurol Neurosurg Psychiatry* 1997;62:329-33.
- Risch N. Linkage strategies for genetically complex traits: I: Multilocus models. *Am J Hum Genet* 1990;46:222-8.
- Sawcer S. A new era in the genetic analysis of multiple sclerosis. *Curr Opin Neurol* 2006;19:237-41.
- Guo SW. Sibling recurrence risk ratio as a measure of genetic effect: Caveat emptor. *Am J Hum Genet* 2002;70:818-9.
- Mumford CJ, Wood NW, Kellar-Wood H, Thorpe JW, Miller DH, Compston DA. The British Isles survey of multiple sclerosis in twins. *Neurology* 1994;44:11-5.
- Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A* 2003;100:12877-82.
- Hansen T, Skytthe A, Stenager E, Petersen HC, Bronnum-Hansen H, Kyvik KO. Concordance for multiple sclerosis in Danish twins: An update of a nationwide study. *Mult Scler* 2005;11:504-10.
- Ristori G, Cannoni S, Stazi MA, Vanacore N, Cotichini R, Alfo M, *et al.* Multiple sclerosis in twins from continental Italy and Sardinia: A nationwide study. *Ann Neurol* 2006;59:27-34.
- Islam T, Gauderman WJ, Cozen W, Hamilton AS, Burnett ME, Mack TM. Differential twin concordance for multiple sclerosis by latitude of birthplace. *Ann Neurol* 2006;60:56-64.
- Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. *Nature* 1995;377:150-1.
- Robertson NP, O'Riordan JI, Chataway J, Kingsley DP, Miller DH, Clayton D, *et al.* Offspring recurrence rates and clinical characteristics of conjugal multiple sclerosis. *Lancet* 1997;349:1587-90.
- Ebers GC, Yee IM, Sadovnick AD, Duquette P. Conjugal multiple sclerosis: Population-based prevalence and recurrence risks in offspring. *Ann Neurol* 2000;48:927-31.
- Ebers GC, Sadovnick AD, Dyment DA, Yee IM, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: Observations in half-siblings. *Lancet* 2004;363:1773-4.
- Dyment DA, Yee IM, Ebers GC, Sadovnick AD. Multiple sclerosis in stepsiblings: Recurrence risk and ascertainment. *J Neurol Neurosurg Psychiatry* 2006;77:258-9.
- Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* 2005;6:109-18.
- Lindsey JW. Familial recurrence rates and genetic models of multiple sclerosis. *Am J Med Genet A* 2005;135:53-8.
- Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? *Int J Epidemiol* 2005;34:1129-37.
- Naito S, Namerow N, Mickey MR, Terasaki PI. Multiple sclerosis: Association with HL-A3. *Tissue Antigens* 1972;2:1-4.
- Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. *Lancet* 1972;1:1240-1.
- Jersild C, Fog T, Hansen GS, Thomsen M, Svejgaard A, Dupont B. Histocompatibility determinants in multiple sclerosis, with special reference to clinical course. *Lancet* 1973;2:1221-5.
- Compston DA, Batchelor JR, McDonald WI. B-lymphocyte alloantigens associated with multiple sclerosis. *Lancet* 1976;2:1261-5.
- Terasaki PI, Park MS, Opelz G, Ting A. Multiple sclerosis and high incidence of a B lymphocyte antigen. *Science* 1976;193:1245-7.
- Cohen D, Cohen O, Marcadet A, Massart C, Lathrop M, Deschamps I, *et al.* Class II HLA-DC beta-chain DNA restriction fragments differentiate among HLA-DR2 individuals in insulin-dependent diabetes and multiple sclerosis. *Proc Natl Acad Sci U S A* 1984;81:1774-8.
- Vartdal F, Sollid LM, Vandvik B, Markussen G, Thorsby E. Patients with multiple sclerosis carry DQB1 genes which encode shared polymorphic amino acid sequences. *Hum Immunol* 1989;25:103-10.
- Olerup O, Hillert J. HLA class II-associated genetic susceptibility in multiple sclerosis: A critical evaluation. *Tissue Antigens* 1991;38:1-15.
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, *et al.* Gene map of the extended human MHC. *Nat Rev Genet* 2004;5:889-99.
- Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dyment DA, Tiislar M, *et al.* A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 2005;37:1108-12.
- Yeo TW, De Jager PL, Gregory SG, Barcellos LF, Walton A, Goris A, *et al.* A second major histocompatibility complex susceptibility locus for multiple sclerosis. *Ann Neurol* 2007;61:228-36.
- Oksenberg JR, Barcellos LF, Cree BA, Baranzini SE, Bugawan TL, Khan O, *et al.* Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. *Am J Hum Genet* 2004;74:160-7.
- Marrosu MG, Murru R, Murru MR, Costa G, Zavattari P, Whalen M, *et al.* Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. *Hum Mol Genet* 2001;10:2907-16.
- Dyment DA, Herrera BM, Cader MZ, Willer CJ, Lincoln MR, Sadovnick AD, *et al.* Complex interactions among MHC haplotypes in multiple sclerosis: Susceptibility and resistance. *Hum Mol Genet*

- 2005;14:2019-26.
43. Barcellos LF, Sawcer S, Ramsay PP, Baranzini SE, Thomson G, Briggs F, *et al.* Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. *Hum Mol Genet* 2006;15:2813-24.
 44. Ramagopalan SV, Morris AP, Dymont DA, Herrera BM, DeLuca GC, Lincoln MR, *et al.* The inheritance of resistance alleles in multiple sclerosis. *PLoS Genet* 2007;3:1607-13.
 45. Fogdell-Hahn A, Ligens A, Gronning M, Hillert J, Olerup O. Multiple sclerosis: A modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens* 2000;55:140-8.
 46. Rubio JP, Bahlo M, Butzkueven H, van Der Mei IA, Sale MM, Dickinson JL, *et al.* Genetic dissection of the human leukocyte antigen region by use of haplotypes of Tasmanians with multiple sclerosis. *Am J Hum Genet* 2002;70:1125-37.
 47. de Jong BA, Huizinga TW, Zanelli E, Giphart MJ, Bollen EL, Uitdehaag BM, *et al.* Evidence for additional genetic risk indicators of relapse-onset MS within the HLA region. *Neurology* 2002;59:549-55.
 48. Harbo HF, Lie BA, Sawcer S, Celius EG, Dai KZ, Oturai A, *et al.* Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis. *Tissue Antigens* 2004;63:237-47.
 49. Chao MJ, Barnardo MC, Lui GZ, Lincoln MR, Ramagopalan SV, Herrera BM, *et al.* Transmission of class I/II multi-locus MHC haplotypes and multiple sclerosis susceptibility: Accounting for linkage disequilibrium. *Hum Mol Genet* 2007;16:1951-8.
 50. Koeleman BP, Herr MH, Zavattari P, Dudbridge F, March R, Campbell D, *et al.* Conditional ETD analysis of the human leukocyte antigen region in type 1 diabetes. *Ann Hum Genet* 2000;64:215-21.
 51. Wadia NH, Trikannad VS, Krishnaswamy PR. HLA antigens in multiple sclerosis amongst Indians. *J Neurol Neurosurg Psychiatry* 1981;44:849-51.
 52. Trikannad VS, Wadia NH, Krishnaswamy PR. Multiple sclerosis and HLA-B12 in Parsi and non-Parsi Indians: A clarification. *Tissue Antigens* 1982;19:155-7.
 53. Kelly MA, Jacobs KH, Penny MA, Mijovic CH, Nightingale S, Barnett AH, *et al.* An investigation of HLA-encoded genetic susceptibility to multiple sclerosis in subjects of Asian Indian and Afro-Caribbean ethnic origin. *Tissue Antigens* 1995;45:197-202.
 54. Rani R, Fernandez-Vina MA, Stastny P. Associations between HLA class II alleles in a North Indian population. *Tissue Antigens* 1998;52:37-43.
 55. Rosenberg NA, Mahajan S, Gonzalez-Quevedo C, Blum MG, Nino-Rosales L, Ninis V, *et al.* Low levels of genetic divergence across geographically and linguistically diverse populations from India. *PLoS Genet* 2006;2:e215.
 56. Horton R, Gibson R, Coggill P, Miretti M, Allcock RJ, Almeida J, *et al.* Variation analysis and gene annotation of eight MHC haplotypes: The MHC Haplotype Project. *Immunogenetics* 2008;60:1-18.
 57. Miretti MM, Walsh EC, Ke X, Delgado M, Griffiths M, Hunt S, *et al.* A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet* 2005;76:634-46.
 58. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet* 2006;38:1166-72.
 59. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002;4:45-61.
 60. Ioannidis JP, Trikalinos TA, Khoury MJ. Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. *Am J Epidemiol* 2006;164:609-14.
 61. Ioannidis JP. Genetic associations: False or true? *Trends Mol Med* 2003;9:135-8.
 62. Hensiek AE, Sawcer SJ, Compston DA. Searching for needles in haystacks-the genetics of multiple sclerosis and other common neurological diseases. *Brain Res Bull* 2003;61:229-34.
 63. Sawcer S. The complex genetics of multiple sclerosis: Pitfalls and prospects. *Brain* 2008;131:3118-31.
 64. Vitale E, Cook S, Sun R, Specchia C, Subramanian K, Rocchi M, *et al.* Linkage analysis conditional on HLA status in a large North American pedigree supports the presence of a multiple sclerosis susceptibility locus on chromosome 12p12. *Hum Mol Genet* 2002;11:295-300.
 65. Modin H, Masterman T, Thorlacius T, Stefansson M, Jonasdottir A, Stefansson K, *et al.* Genome-wide linkage screen of a consanguineous multiple sclerosis kinship. *Mult Scler* 2003;9:128-34.
 66. Haghighi S, Andersen O, Nilsson S, Rydberg L, Wahlstrom J. A linkage study in two families with multiple sclerosis and healthy members with oligoclonal CSF immunopathy. *Mult Scler* 2006;12:723-30.
 67. Willer CJ, Dymont DA, Cherny S, Ramagopalan SV, Herrera BM, Morrison KM, *et al.* A genome-wide scan in forty large pedigrees with multiple sclerosis. *J Hum Genet* 2007;52:955-62.
 68. Dymont DA, Cader MZ, Herrera BM, Ramagopalan SV, Orton SM, Chao M, *et al.* A genome scan in a single pedigree with a high prevalence of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2008;79:158-62.
 69. Sawcer S, Jones HB, Feakes R, Gray J, Smaldon N, Chataway J, *et al.* A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* 1996;13:464-8.
 70. Haines JL, Ter-Minassian M, Bazyk A, Gusella JF, Kim DJ, Terwedow H, *et al.* A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. *Nat Genet* 1996;13:469-71.
 71. Ebers GC, Kukay K, Bulman DE, Sadovnick AD, Rice G, Anderson C, *et al.* A full genome search in multiple sclerosis. *Nat Genet* 1996;13:472-6.
 72. Kuokkanen S, Gschwend M, Rioux JD, Daly MJ, Terwilliger JD, Tienari PJ, *et al.* Genomewide scan of multiple sclerosis in Finnish multiplex families. *Am J Hum Genet* 1997;61:1379-87.
 73. Coraddu F, Sawcer S, D'Alfonso S, Lai M, Hensiek A, Solla E, *et al.* A genome screen for multiple sclerosis in Sardinian multiplex families. *Eur J Hum Genet* 2001;9:621-6.
 74. Broadley S, Sawcer S, D'Alfonso S, Hensiek A, Coraddu F, Gray J, *et al.* A genome screen for multiple sclerosis in Italian families. *Genes Immun* 2001;2:205-10.
 75. Akesson E, Oturai A, Berg J, Fredrikson S, Andersen O, Harbo HF, *et al.* A genome-wide screen for linkage in Nordic sib-pairs with multiple sclerosis. *Genes Immun* 2002;3:279-85.
 76. Ban M, Stewart GJ, Bennetts BH, Heard R, Simmons R, Maranian M, *et al.* A genome screen for linkage in Australian sibling-pairs with multiple sclerosis. *Genes Immun* 2002;3:464-9.
 77. Eraksoy M, Kurtuncu M, Akman-Demir G, Kilinc M, Gedizlioglu M, Mirza M, *et al.* A whole genome screen for linkage in Turkish multiple sclerosis. *J Neuroimmunol* 2003;143:17-24.
 78. Hensiek AE, Roxburgh R, Smilie B, Coraddu F, Akesson E, Holmans P, *et al.* Updated results of the United Kingdom linkage-based genome screen in multiple sclerosis. *J Neuroimmunol* 2003;143:25-30.
 79. Kenealy SJ, Babron MC, Bradford Y, Schnetz-Boutaud N, Haines JL, Rimmler JB, *et al.* A second-generation genomic screen for multiple sclerosis. *Am J Hum Genet* 2004;75:1070-8.
 80. Dymont DA, Sadovnick AD, Willer CJ, Armstrong H, Cader ZM, Wiltshire S, *et al.* An extended genome scan in 442 Canadian multiple sclerosis-affected sibships: A report from the Canadian Collaborative Study Group. *Hum Mol Genet* 2004;13:1005-15.
 81. Ligens A, Dymont DA, Willer CJ, Sadovnick AD, Ebers G, Risch N, *et al.* Evidence of linkage with HLA-DR in DRB1*15-negative families with multiple sclerosis. *Am J Hum Genet* 2001;69:900-3.
 82. Genetic Analysis of Multiple Sclerosis in EuropeanS (GAMES) and the Transatlantic Multiple Sclerosis Genetics Cooperative. A meta-analysis of whole genome linkage screens in multiple sclerosis. *J Neuroimmunol* 2003;143:39-46.

83. International Multiple Sclerosis Genetics Consortium (IMSGC). Enhancing linkage analysis of complex disorders: An evaluation of high-density genotyping. *Hum Mol Genet* 2004;13:1943-9.
84. International Multiple Sclerosis Genetics Consortium (IMSGC). A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet* 2005;77:454-67.
85. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516-7.
86. Goring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 2001;69:1357-69.
87. Burwick RM, Ramsay PP, Haines JL, Hauser SL, Oksenberg JR, Pericak-Vance MA, *et al.* APOE epsilon variation in multiple sclerosis susceptibility and disease severity: Some answers. *Neurology* 2006;66:1373-83.
88. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177-82.
89. Kruglyak L, Nickerson DA. Variation is the spice of life. *Nat Genet* 2001;27:234-6.
90. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science* 2008;322:881-8.
91. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: An approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004;96:434-42.
92. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19:149-50.
93. Wellcome Trust Case Control Consortium (WTCCC). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-78.
94. Hauser MA, Li YJ, Takeuchi S, Walters R, Noureddine M, Maready M, *et al.* Genomic convergence: Identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. *Hum Mol Genet* 2003;12:671-7.
95. Fernald GH, Yeh RF, Hauser SL, Oksenberg JR, Baranzini SE. Mapping gene activity in complex disorders: Integration of expression and genomic scans for multiple sclerosis. *J Neuroimmunol* 2005;167:157-69.
96. Teutsch SM, Booth DR, Bennetts BH, Heard RN, Stewart GJ. Identification of 11 novel and common single nucleotide polymorphisms in the interleukin-7 receptor-alpha gene and their associations with multiple sclerosis. *Eur J Hum Genet* 2003;11:509-15.
97. Zhang Z, Duvefelt K, Svensson F, Masterman T, Jonasdottir G, Salter H, *et al.* Two genes encoding immune-regulatory molecules (LAG3 and IL7R) confer susceptibility to multiple sclerosis. *Genes Immun* 2005;6:145-52.
98. Gregory SG, Schmidt S, Seth P, Oksenberg JR, Hart J, Prokop A, *et al.* Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 2007;39:1083-91.
99. Lundmark F, Duvefelt K, Iacobaeus E, Kockum I, Wallstrom E, Khademi M, *et al.* Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 2007;39:1108-13.
100. Freimer N, Sabatti C. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet* 2004;36:1045-51.
101. Pe'er I, de Bakker PI, Maller J, Yelensky R, Altshuler D, Daly MJ. Evaluating and improving power in whole-genome association studies using fixed marker sets. *Nat Genet* 2006;38:663-7.
102. International Multiple Sclerosis Genetics Consortium (IMSGC). Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *N Engl J Med* 2007;357:851-62.
103. International Multiple Sclerosis Genetics Consortium (IMSGC). Refining genetic associations in multiple sclerosis. *Lancet Neurol* 2008;7:567-9.
104. Broadley SA, Deans J, Sawcer SJ, Clayton D, Compston DA. Autoimmune disease in first-degree relatives of patients with multiple sclerosis: A UK survey. *Brain* 2000;123:1102-11.
105. Barcellos LF, Kamdar BB, Ramsay PP, DeLoa C, Lincoln RR, Caillier S, *et al.* Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: A descriptive study. *Lancet Neurol* 2006;5:924-31.
106. International Multiple Sclerosis Genetics Consortium (IMSGC). The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes Immun* 2009;10:11-4.
107. Hafler JP, Maier LM, Cooper JD, Plagnol V, Hinks A, Simmonds MJ, *et al.* CD226 Gly307Ser association with multiple autoimmune diseases. *Genes Immun* 2009;10:5-10.
108. Zoledziwska M, Costa G, Pitzalis M, Cocco E, Melis C, Moi L, *et al.* Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. *Genes Immun* 2009;10:15-7.
109. De Jager PL, Baecher-Allan C, Maier LM, Arthur AT, Ottoboni L, Barcellos L, *et al.* The role of the CD58 locus in multiple sclerosis. *Proc Natl Acad Sci U S A.* 2009 31;106:5264-9.
110. Wellcome Trust Case Control Consortium (WTCCC). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329-37.
111. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009;18:767-78.

Received: 08-02-09, **Accepted:** 06-07-09

Source of Support: Nil, **Conflict of Interest:** Nil