

## Research Article

# What is Next for the Genetics of Multiple Sclerosis?

Sreeram V. Ramagopalan<sup>1,2,3</sup> and David A. Dymant<sup>4</sup>

<sup>1</sup> Wellcome Trust Centre for Human Genetics, University of Oxford, OX3 7BN Oxford, UK

<sup>2</sup> Department of Clinical Neurology, University of Oxford, OX3 9DU Oxford, UK

<sup>3</sup> Blizard Institute of Cell and Molecular Science, Queen Mary University of London, Barts and The London School of Medicine and Dentistry, E1 2AT London, UK

<sup>4</sup> The Department of Medical Genetics, University of Calgary, Calgary Alberta, Canada T2N 7A1

Correspondence should be addressed to David A. Dymant, david.dymant@calgaryhealthregion.ca

Received 12 January 2011; Accepted 26 January 2011

Academic Editor: Noriko Isobe

Copyright © 2011 S. V. Ramagopalan and D. A. Dymant. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We review here our current understanding of the genetic aetiology of the common complex neurological disease multiple sclerosis (MS). The strongest genetic risk factor for MS is the major histocompatibility complex which was identified in the 1970s. In 2011, after a number of genome-wide association studies have been completed and have identified approximately 20 new genes for MS, we ask the question—what is next for the genetics of MS?

## 1. Introduction

Hermann Eichorst first recognized the familial clustering of multiple sclerosis (MS) late in the 19th century [1] but it was not until almost a century had passed that it was firmly established by population-based studies that relatives of MS patients have an increased risk for developing the disease [2]. In keeping with this observation, Davenport had already made it clear that northern Europeans had a higher frequency of MS, whereas the disease was much less common in Asians and Africans [1]. Twin studies and later an adoption study showed that the observed familial aggregation was due to shared genes versus any shared familial exposures [3].

Progress in identifying the genes responsible has, historically, progressed at a glacial pace, despite the early success with the major histocompatibility complex (MHC) association. In 1972, MS was shown to be associated with the MHC [4], later pinpointed to a specific allele, *HLA-DRB1\*15*, of the class II gene *HLA-DRB1* [5]. The success in identifying this association was primarily due to the large effect size of this region (odds ratio (OR) >2). The next significant step forward occurred in 2007 when new genes were identified for MS by genomewide association studies (GWAS) [6].

Now at the dawn of 2011, we have approximately 20 genes conferring a mild to modest effect (OR < 1.3) on risk that are robustly associated to MS [7], as well as the long known HLA association. The question can therefore be posed, is this “job done” for MS genetics?

## 2. Discussion

Unfortunately far from it; the associated variants so far identified explain about 50% of the inherited risk of MS. There are several possible explanations as to the “missing” genetic basis of MS.

It is possible that the immune-related disease loci identified to date have more of an overall impact on MS risk than currently estimated. This can result when the marker SNP is an imperfect proxy for the actual causal mutation that led to the association signal. There is some evidence to support this hypothesis in complex disease. Recent resequencing of 63 GWAS-identified positional candidate genes in Crohn's disease identified three novel low-frequency coding variants in the *IL23R* gene [8]. The odds ratios conferred by these newly detected low frequency variants was approximately 2.4 on average [8]. Although this odds ratio is a possible

overestimation due to winner's curse, the value appears larger than the approximate 20% risk increase conferred by most common variants detected by GWAS. However, the fact that only 1 out of 63 genes had robustly associated rare variants and that these newly detected variants jointly explained only an extra 0.44% of the variance of Crohn's disease suggests that rare variants in GWAS-associated genes are not likely to make a large contribution to inherited predisposition to complex disease [8].

Perhaps there are additional disease loci than the roughly 20 or so associated genes? These other susceptibility genes could be identified by even larger scale GWAS involving tens or hundreds of thousands of MS patients and controls. The GWAS published so far in MS have not exceeded 2000 patients in the initial (screening) phase. Statistical modelling has suggested that 12,627 SNPs explains approximately 3% of the variance in MS risk [9]. We await the results of the MS GWAS funded by the Wellcome Trust which involves tens of thousands of participants to see if the vast resources expended in such a project are translated into novel insights into MS aetiology. It should be remembered that MS is phenotypically a heterogeneous disease [1], and while current GWAS have used unselected patient populations to identify disease associations, this may miss variants more important to certain patient groups.

Another explanation may be that some disease loci may contain only rare variants. In order to identify these genes a sequencing-based approach would be required. In the past this was not possible given the cost and technology available; however recent advances in next-generation sequencing technologies (whole exome and whole genome sequencing) could rapidly facilitate the identification of these variants that would be too rare to be picked up by GWAS. These rare variants would be expected to be causal and have a relatively large effect on risk (i.e.,  $OR > 3$ ). The 1000-Genomes project has highlighted the fact that each of us has 250 to 300 loss-of-function variants in our genes [10]. However, for complex diseases, power considerations will be an issue to cope with the wealth of data generated by whole genome sequencing. It has been suggested that whole-exome sequencing will be most fruitful in identifying rare disease causative variants in families that have multiple affected individuals [11].

SNPs are only one type of genetic variation. It has been observed that individual copies of the human genome contain large regions (tens to hundreds of kilobases in size) that are deleted, duplicated, or inverted relative to the reference sequence. These structural variants may contribute to MS aetiology but have not yet been adequately tested. However, a study by the Wellcome Trust Case Control Consortium observed that most common structural variation are well tagged by SNPs and so have been indirectly explored through genome-wide SNP studies and therefore concluded that common structural variants are unlikely to contribute greatly to the genetic basis of common human diseases [12].

Moving on from single locus associations to consider biological systems, it may be that gene-gene and gene-environment interactions may play an important role in disease. Once patterns of association and interaction are

better understood, the effects of specific gene and environmental exposures on developing MS may be significant. Indeed epistatic interactions exist between MHC haplotypes [13] and can greatly alter risk. For example, *HLA-DRB1\*08*, interacts with *HLA-DRB1\*15* to more than double the risk associated with a single copy of *HLA-DRB1\*15* [13]. On its own, *HLA-DRB1\*08* increases the risk of MS modestly, highlighting that a variant with a small marginal effect is not necessarily clinically insignificant; it may turn out to have a strong effect in certain genetic backgrounds. As yet, no functional explanation can be given for these interactions; understanding the mechanism of these interactions will be critical to further understanding MS aetiology.

Epigenetic contributions may also play an important role in MS. Epidemiological data strongly hints at a parent-of-origin effect in MS [14]. For example, maternal half-siblings have double the risk for MS as compared to paternal half-siblings (2.35% versus 1.31%) [14]. Risk for MS in maternal half-siblings compared to their full siblings does not differ significantly [14] suggesting that this maternal effect is a major component of familial aggregation of the disease. The mechanism of the increased risk conferred maternally remains to be elucidated but epigenetic mechanisms such as DNA methylation and histone modification may regulate genomic function in such a way to increase MS risk [15]. A recent study investigating these effects utilized next-generation sequencing in discordant twins. The investigators could not find evidence for any epigenetic differences between the twins to explain the MS discordance [16]. However, there were a number of limitations to the study design used, and it is of interest that DNA methylation differences have been shown to exist between twins discordant for systemic lupus erythematosus [17] and for parent of origin effects in type 2 diabetes [18].

### 3. Conclusion

As with all complex diseases, the genetics of MS has not yet been fully elucidated. While GWAS have been responsible for a wealth of new information these association studies have not provided all the answers for MS risk. We are now in an era of very exciting potential applications of sequencing technology. Next-generation sequencing platforms allow us to survey multiple levels of natural variation at unprecedented resolution and depth. As sequencing costs continue to decrease, and both laboratory and computational protocols improve, we will see ever increasing use of this technology, hopefully enabling us to completely unlock the complex genetic basis of MS. There is unlikely to have a single answer, with interactions, rare variants, epigenetic factors all likely to be contributing. Ultimately, well-performed functional studies will be required to understand how all these risk factors interact to predispose to MS. Against this it will be debated whether further genetic research will actually advance our understanding of MS. However, the motivation for future work is the need to understand disease mechanisms to derive safe and effective treatments and ultimately to prevent the disease.

## References

- [1] D. W. Paty and G. C. Ebers, *Multiple Sclerosis*, Davis, Philadelphia, Pa, USA, 1998.
- [2] A. D. Sadovnick, P. A. Baird, and R. H. Ward, "Multiple sclerosis: updated risks for relatives," *American Journal of Medical Genetics*, vol. 29, no. 3, pp. 533–541, 1988.
- [3] G. C. Ebers, D. E. Bulman, A. D. Sadovnick et al., "A population-based study of multiple sclerosis in twins," *New England Journal of Medicine*, vol. 315, no. 26, pp. 1638–1642, 1986.
- [4] C. Jersild, A. Svejgaard, and T. Fog, "HL-A antigens and multiple sclerosis," *Lancet*, vol. 1, no. 7762, pp. 1240–1241, 1972.
- [5] A. Fogdell, J. Hillert, C. Sachs, and O. Olerup, "The multiple sclerosis- and narcolepsy-associated HLA class II haplotype includes the DRB5\*0101 allele," *Tissue Antigens*, vol. 46, no. 4, pp. 333–336, 1995.
- [6] D. A. Hafler, A. Compston, S. Sawcer et al., "Risk alleles for multiple sclerosis identified by a genomewide study," *New England Journal of Medicine*, vol. 357, no. 9, pp. 851–862, 2007.
- [7] P. L. De Jager, X. Jia, J. Wang et al., "Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci," *Nature Genetics*, vol. 41, no. 7, pp. 776–782, 2009.
- [8] Y. Momozawa, M. Mni, K. Nakamura et al., "Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease," *Nature Genetics*, vol. 43, no. 1, pp. 43–47, 2011.
- [9] The International Multiple Sclerosis Genetics Consortium (IMSGC), "Evidence for polygenic susceptibility to multiple sclerosis—the shape of things to come," *American Journal of Human Genetics*, vol. 86, no. 4, pp. 621–625, 2010.
- [10] R. M. Durbin, G. R. Abecasis, D. L. Altshuler et al., "A map of human genome variation from population-scale sequencing," *Nature*, vol. 467, pp. 1061–1073, 2010.
- [11] E. T. Cirulli and D. B. Goldstein, "Uncovering the roles of rare variants in common disease through whole-genome sequencing," *Nature Reviews Genetics*, vol. 11, no. 6, pp. 415–425, 2010.
- [12] N. Craddock, M. E. Hurles, N. Cardin et al., "Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls," *Nature*, vol. 464, no. 7289, pp. 713–720, 2010.
- [13] S. V. Ramagopalan and G. C. Ebers, "Epistasis: multiple sclerosis and the major histocompatibility complex," *Neurology*, vol. 72, no. 6, pp. 566–567, 2009.
- [14] G. C. Ebers, A. D. Sadovnick, D. A. Dymont, I. M. L. Yee, C. J. Willer, and N. Risch, "Parent-of-origin effect in multiple sclerosis: observations in half-siblings," *Lancet*, vol. 363, no. 9423, pp. 1773–1774, 2004.
- [15] A. E. Handel, G. C. Ebers, and S. V. Ramagopalan, "Epigenetics: molecular mechanisms and implications for disease," *Trends in Molecular Medicine*, vol. 16, no. 1, pp. 7–16, 2010.
- [16] S. E. Baranzini, J. Mudge, J. C. Van Velkinburgh et al., "Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis," *Nature*, vol. 464, no. 7293, pp. 1351–1356, 2010.
- [17] B. M. Javierre, A. F. Fernandez, J. Richter et al., "Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus," *Genome Research*, vol. 20, no. 2, pp. 170–179, 2010.
- [18] A. Kong, V. Steinthorsdottir, G. Masson et al., "Parental origin of sequence variants associated with complex diseases," *Nature*, vol. 462, no. 7275, pp. 868–874, 2009.