

# 1 The Genetics of Multiple Sclerosis

ALESSANDRO DIDONNA • JORGE R. OKSENBORG

Department of Neurology, University of California at San Francisco, San Francisco, CA, USA

**Author for correspondence:** Alessandro Didonna, Department of Neurology, University of California at San Francisco, 675 Nelson Rising Lane, San Francisco, CA 94158, USA. Email: Alessandro.Didonna@ucsf.edu

Doi: <http://dx.doi.org/10.15586/codon.multiplesclerosis.2017.ch1>

**Abstract:** Multiple sclerosis (MS) is an autoimmune disease of the central nervous system, characterized by focal inflammation, demyelination, and axonal injury. The etiology of MS is still uncertain, but the most updated working model for disease pathogenesis proposes the interplay between genetic and environmental factors as necessary for MS manifestation. With the notable exception of the major histocompatibility complex (*MHC*), the identity of MS genetic determinants has been elusive for decades. In recent years, the advent of genome-wide association studies (GWAS) and collaborative efforts among international centers have fueled the characterization of several non-*MHC* loci associated with MS susceptibility. To date, after a number of GWAS screenings, 110 MS risk variants have been discovered outside the *MHC* locus in European populations. In the future, functional studies will be required to define the biological pathways and cellular activities connected to these variants.

**Key words:** Autoimmunity; Genome-wide association studies; Human leukocyte antigen; Multiple sclerosis; Single-nucleotide polymorphism

---

In: *Multiple Sclerosis: Perspectives in Treatment and Pathogenesis*. Ian S. Zagon and Patricia J. McLaughlin (Editors), Codon Publications, Brisbane, Australia. ISBN: 978-0-9944381-3-3; Doi: <http://dx.doi.org/10.15586/codon.multiplesclerosis.2017>

**Copyright:** The Authors.

**Licence:** This open access article is licenced under Creative Commons Attribution 4.0 International (CC BY-NC 4.0). <https://creativecommons.org/licenses/by-nc/4.0/>

---

## Introduction

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS), characterized by focal lymphocytic infiltrates, the breakdown of myelin sheaths wrapping axons, astrogliosis, microglia activation, and diffuse neurodegeneration (1). Clinical manifestation is heterogeneous, ranging from relatively mild neurological symptoms to a rapidly evolving and debilitating disease. MS typically begins with a relapsing-remitting clinical phase (RR-MS), dominated by inflammatory events, both in the periphery and CNS, and full or partial recovery. In the majority of affected individuals, this initial relapsing-remitting course evolves years later into a secondary progressive MS (SP-MS), characterized by the irreversible accumulation of neurological disabilities as a result of axonal injury and neuronal loss. However, a proportion of MS patients (up to 15%) enter directly into the progressive phase after clinical onset, without experiencing initial relapses (2). This disease subtype is known as primary progressive MS (PP-MS) and is associated with an irreversible and progressive severe clinical phenotype. Significantly, the mean age of onset of SP-MS and PP-MS is similar, approximately 40 years (3). A total of 14 FDA-approved treatments for RR-MS are now available as disease modifiers to control inflammatory lesions and clinical relapsing activity. However, their long-term effects on disease progression remain largely unknown.

With the age of onset ranging between 20 and 40 years, MS represents the most common cause of acquired neurological disability among young adults, affecting over 2.5 million people worldwide. MS affects women more often than men (3:1 ratio), but its incidence also varies according to ethnicity and geographical location, with northern Europeans and their descendants being more susceptible to develop the disease (4). MS etiology is still elusive but there is a growing body of experimental evidence, suggesting that both genetic determinants and environmental factors converge to determine disease susceptibility and clinical trajectory. This chapter will review key milestones in MS genetic research with an emphasis on the technological and conceptual advances that have fueled the identification of discrete genomic loci associated with MS risk.

---

## Multiple Sclerosis Holds a Genetic Component

The discovery of family aggregation in the second half of the 19th century shed light for the first time on the genetic component of the disease. Compared to a lifetime risk of 0.2% in the general population, siblings of affected individuals have a 10- to 20-fold higher risk of developing the disease (2–4%), with monozygotic twins having an even higher risk (30%) (5, 6). In contrast, spouses and adoptees hold a risk comparable to that of the general population (or their original nuclear families), consistent with genetic sharing being the driver of familial aggregation (7). On the other hand, the fact that the relative risk does not reach 100% even in identical twins suggests that other factors beyond DNA sequence identity must concur to create the conditions that cause or allow the dysregulation of the immune response associated with MS. A broad range of determinants lie in this category; they include environmental exposures (e.g., smoking, viral infections,

vitamin D intake, diet, and microbiome) as well as epigenetic signatures (e.g., DNA methylation patterns, histone modifications, and non-coding RNAs) (8).

Another factor supporting MS heritability consists in the distinctive worldwide prevalence of the disease. People living in northern Europe and North America exhibit a higher disease incidence (1–2 in 1000) when compared with southern Europeans. Moreover, MS is uncommon in some ethnic groups such as Uzbeks, Samis, Turkmen, Kyrgyzis, Kazakhs, native Siberians, North and South Amerindians, Japanese, Chinese, African blacks, and New Zealand Maori (9). Although these differences could be partially explained by differential exposure to specific environmental factors (such as certain nonubiquitous pathogens), the presence of MS-resistant or low-incidence ancestral groups suggests that the history and genetic architecture of a population influence its own risk of developing MS.

Altogether, these epidemiological observations—in particular the nonlinear relationship between genetic distance from a proband and the lifetime risk to develop MS—support a polygenic etiology for MS following the “common variant-common disease” paradigm of genetic influences and inheritance. According to this model, the overall MS risk is the result of the contributions of multiple polymorphic genes with risk alleles common in the population, each one determining a moderate portion of the risk (10, 11). This non-Mendelian pattern of transmission is not exclusive of MS but is shared with other autoimmune diseases and chronic disorders such as type II diabetes and obesity. These conditions are collectively known as complex genetic disorders, which are characterized primarily by polygenic risk and multifaceted gene–environment interactions.

## THE HUMAN LEUKOCYTE ANTIGEN LOCUS IN MS

The strongest genetic association signal in MS resides within the major histocompatibility complex (MHC) in chromosome 6p21.3. This 4-megabase region contains approximately 160 closely linked genes. About half of these genes have important roles in the regulation of the immune system, and include the six classical transplantation human leukocyte antigen (HLA) genes—the class I genes *HLA-A*, *HLA-B*, and *HLA-C*, and the class II genes *HLA-DPB1*, *HLA-DQB1*, and *HLA-DRB1* (12). *HLA* genes are highly polymorphic, with over 15,000 alleles identified to date (<http://hla.alleles.org/nomenclature/index.html>). The first evidence of association between *HLA* and MS risk dates back to 1972, when the frequencies of surface glycoproteins encoded by the *HLA-A3* and *HLA-B7* class I alleles were found enriched in MS patients using serological reagents (13, 14). In the following years, numerous investigations, regardless of sample size and the resolution, have independently replicated the association of the *HLA* locus with MS risk across all populations studied, in both primary progressive and relapsing-remitting patients. Although the initial association was to class I *HLA-A* and *HLA-B* alleles, better powered studies, including genome-wide association studies (GWAS), have shown that the main MS susceptibility signal genome-wide maps to the *HLA-DRB1* locus in the class II region of the MHC. The *HLA-DRB1*\*15:01 allele has the strongest effect, with an average odds ratio (OR, a frequently used measure of effect size) of 3.08 and a clear dose response to 0, 1, or 2 allele copies the individual carries (15). However, complex allelic hierarchical lineages, cis/trans-epistatic and haplotypic effects, and independent protective signals, specifically in the class I region of the locus, have been documented as well.

Using GWAS single-nucleotide polymorphism (SNP) data (5091 cases/9595 controls), the International Multiple Sclerosis Genetics Consortium (IMSGC) reported in 2013 the isolation of 11 statistically independent effects in the *MHC* region: six *HLA-DRB1* and one *HLA-DPB1* alleles in the centromeric class II region of the locus; one *HLA-A* and two *HLA-B* alleles in the telomeric class I region; and one in the class III region between *MHC* class I polypeptide-related sequence B (*MICB*) and leukocyte-specific transcript 1 (*LST1*) (16). More recently, the analysis of independent high-density *MHC* region SNP data from multiple cohorts of European ancestry has provided, in addition to novel and previously identified *HLA* class II risk alleles (*DRB1\*15:01*, *DRB1\*13:03*, *DRB1\*03:01*, *DRB1\*08:01*, and *DQB1\*03:02*) and independent *HLA* class I protective alleles (*A\*02:01*, *B\*44:02*, *B\*38:01*, and *B\*55:01*), evidence for two interactions involving pairs of class II alleles: *DQA1\*01:01-DRB1\*15:01* and *DQB1\*03:01-DQB1\*03:02* (17). Larger ongoing studies hold the potential for discovering additional independent and interactive effects.

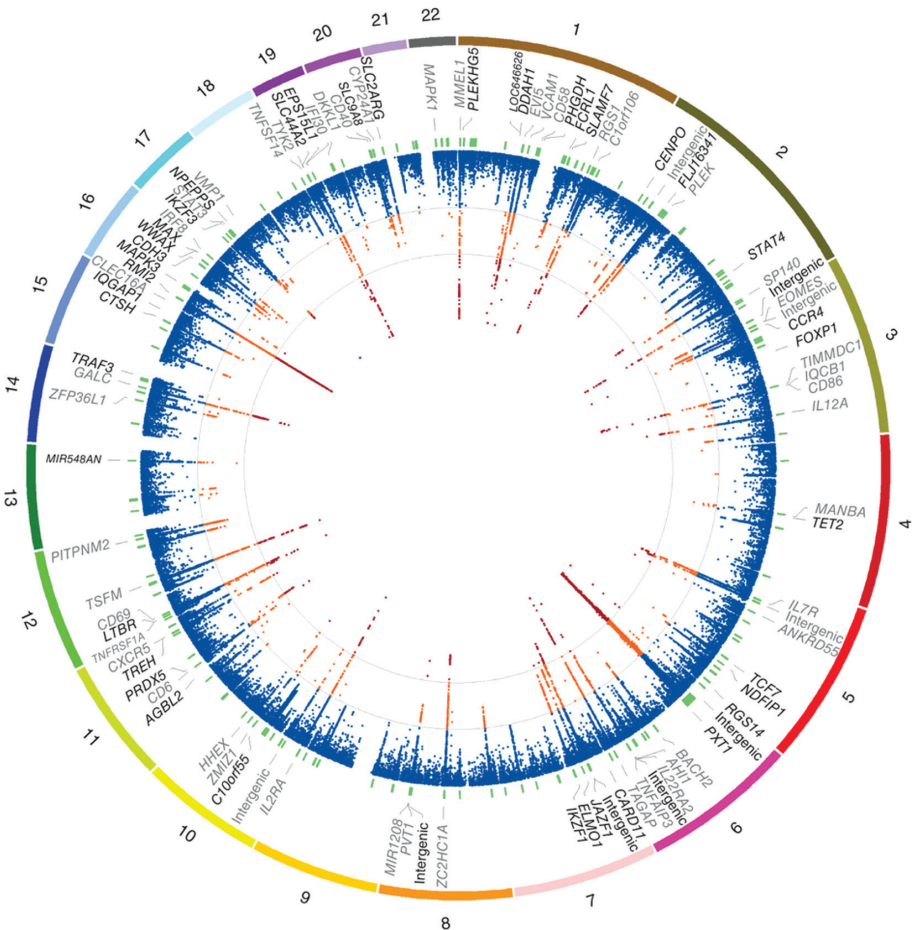
## THE ADVENT OF GENOME-WIDE ASSOCIATION STUDIES IN MS RESEARCH

In the early 2000s, the introduction of chip-based technologies with the capacity to genotype simultaneously hundreds of thousands of SNPs allowed the development of a new analytical methodology known as genome-wide association studies or GWAS—a hypothesis-free method in which SNPs spaced across the entire genome are screened for association with a particular trait in case–control datasets composed of genetically unrelated individuals (18). Compared to classic linkage studies that rely on extended families, the possibility to test unrelated individuals allows collecting much larger datasets, substantially increasing the statistical power of gene-discovery studies. GWA studies have been a determinant to deconstruct the genetics of many multifactorial disorders, characterized by common genetic variants conferring moderate risk to disease susceptibility.

The first MS GWAS was reported in 2007 by the IMSGC employing 931 family trios (one affected child and both parents). The screening confirmed with genome-wide significance the association of the previously identified locus containing the interleukin-7 receptor  $\alpha$  (*IL7R $\alpha$* ) gene, and detected a novel non-*HLA* disease-risk locus, defined by the presence of the interleukin-2 receptor  $\alpha$  (*IL2R $\alpha$* ) gene (19). In the following years, between 2007 and 2011, seven additional GWA studies of comparable size and one meta-analysis were performed, adding 21 new loci to the roster of MS risk variants. However, theoretical power estimations showed that all the studies conducted at that time were substantially underpowered to capture risk variants with odd ratios less than 1.2, which were the values expected for most of the MS risk variants (20). For that reason, the IMSGC decided in 2011 to embark on the largest MS GWAS with the collaborative effort of the Wellcome Trust Case Control Consortium 2 (WTCCC2). This new study employed nearly 10,000 MS cases and 20,000 healthy controls of European ancestry and was able to extend the list of genome-wide significant MS loci to 52, of which 29 were never reported before (21). Remarkably, most of the associated variants were found located in proximity to genes with documented immune functions, corroborating the hypothesis that the dysregulation of physiological immune response most likely represents the driving factor of MS. Two years later,

MS genetic association was further refined through a novel multicenter study based on a custom high-density genotyping array named ImmunoChip. Over 80,000 individuals of European descent were analyzed and 48 new susceptibility variants were identified as genome-wide significant (22).

After a decade of GWAS screenings in European populations, the MS genetic atlas currently includes 110 non-*MHC* risk variants belonging to 103 genetic loci (Figure 1). In aggregate, the proportion of the genetic variance accounting for disease risk explained by these polymorphisms has been estimated as roughly 30%,



**Figure 1** Genetic atlas of multiple sclerosis. The circus plot summarizes all the known MS-associated risk loci. The outer most track indicates the numbered autosomal chromosomes, while the second track shows the closest gene to the top hit within each locus (previously identified associations are in gray). The third track indicates the physical position of the 184 fine-mapped intervals (in green). The inner most track indicates  $-\log(p)$  for each SNP (scaled from 0 to 12 which truncates the signal in several regions). Also, contour lines are given at the *a priori* discovery ( $-\log(p) = 4$ ) and genome-wide significance ( $-\log(p) = 7.3$ ) thresholds. Orange indicates  $-\log(p) \geq 4$  and  $< 7.3$ , while red indicates  $-\log(p) \geq 7.3$ . (Reproduced from Ref. (22)).

but the mapping of additional risk variants has been proceeding rapidly through ongoing multicenter initiatives utilizing dense, specialized arrays and very large sample collections. In this regard, a recent report anticipated that over 200 risk variants have been identified through the meta-analysis of all previous GWA studies conducted in MS (23). It is not inconceivable, however, that the potential for the discovery of additive risk variance extractable from large genomic screens will be quickly exhausted. The remaining fraction of the risk commonly known as “missing heritability” is likely due to still unknown common variants characterized by much smaller effects, below the detection limits of the GWA studies conducted so far. Some authors have proposed that a substantial portion of the missing heritability lies in genetic interactions between known variants, the so-called phantom heritability (24). Also, likewise gene by environment interactions, cis/trans-regulators of allelic expression, unidentified rare and penetrant semi-private variants, population and/or disease heterogeneity, neglecting the analysis of sex chromosomes, and hidden epigenetic effects may all contribute to the missing heritability.

---

## From Genes to Function: Understanding the Molecular Basis of MS

The translation of GWAS data into biological functions has been challenging. The principal reason for this shortcoming consists in the pervasive linkage disequilibrium (LD) along the human genome, which hinders the identification of true causative variants. LD refers to the tendency of genetic loci in physical proximity to segregate together during meiosis, leading DNA to be inherited in large blocks through generations. This peculiarity of genome architecture substantially impairs GWAS resolution since SNPs in the same LD block are inherited together as well. Thus, statistically significant GWAS risk variants are usually proxy for the real causative variants, which can be located up to several megabases away within the same LD block. In addition, the identification of the causative variants is further complicated by the fact that most of them are not translated but rather map to regulatory elements (promoters, enhancers, silencers, and other transcription factor-binding sites). Nevertheless, substantial effort has been directed in this post-genomic era toward the functional characterization of the huge amount of genetic data generated by GWAS screenings, using either wet lab approaches or *in silico* analyses (or a combination of both).

### FUNCTIONAL STUDIES IN MS

A variety of experimental systems have been employed to study the biological functions associated with MS risk variants, ranging from patients-derived primary blood cells to animal models of disease. The first putative causal variant identified in MS was the SNP rs6897932 located within the exon 6 of the *IL7R* gene, coding for the trans-membrane segment of the receptor. This SNP was shown to disrupt an exonic splicing silencer, affecting the relative amounts of soluble and membrane-bound isoforms of the protein (25). Recent evidence has shown that the RNA helicase DEAD box polypeptide 39B (DDX39B) is also a potent activator of

*IL7R* exon 6, and the SNP rs2523506 located in the *DDX39B* 5'UTR increases MS risk by reducing *DDX39B* mRNA translation (26). A similar effect was described for the intronic SNP rs2104286 in the *IL2RA* gene as well. In fact, this risk variant was also found to alter the soluble/membrane-bound ratio of *IL2RA* protein by driving the expression of higher levels of its soluble form (27).

Another well-characterized example is the intronic SNP rs1800693 in the *TNFRF1A* gene. In this case, the risk allele promotes the skipping of exon 6 with the production of a novel soluble form of the tumor necrosis factor (TNF) receptor which is able to inhibit TNF signaling inside the cells, mirroring somehow, the exacerbating effects of TNF-blocking drugs on MS course (28). More recently, our group has reported that the nonsynonymous exonic SNP rs11808092 in the ecotropic viral integration site 5 (*EVI5*) gene induces changes in superficial hydrophobicity patterns of the coiled-coil domain of *EVI5* protein, which, in turns, affects the *EVI5* interactome. In particular, we demonstrated that *EVI5* protein bearing the risk allele selectively interacts with sphingosine 1-phosphate lyase (*SGPL1*), an enzyme important for the creation of the *S1P* gradient—which is relevant to adaptive immune response and the therapeutic management of MS (29).

Altogether, available functional data pinpoint at a “transcriptional hypothesis” where risk variants increase the propensity to develop MS by affecting primarily the expression of the associated genes. To this extent, recent advances in bioinformatics and computer-based methods of analysis have greatly helped toward the identification of the cellular pathways dysregulated upon disease.

## PATHWAY ANALYSIS AND SYSTEMS BIOLOGY APPROACHES

The advent of “big data” in genetic research has been paralleled by the development of computational methods that could handle the size and complexity of this new type of information. In particular, different *in silico* approaches have been optimized to extract biologically meaningful associations from large genomic, transcriptomic, and proteomic datasets. These methods usually rely on the computation of overrepresentation of the input genes in specific gene ontology (GO) categories or biological pathways. More elaborated algorithms instead take advantage of gene interaction networks and search for possible sub-networks (modules) enriched in the input genes. Cell specificity and epigenomic reference datasets add additional layers of complexity to the analysis.

An early application of network-based methods in the context of MS was reported in 2011 by the IMSGC, which analyzed the results of the 2011 large GWAS and a following meta-analysis, comprising together a total of 15,317 cases and 29,529 controls. A large protein network encompassing more than 400,000 interactions among ~25,000 human proteins was created for the analysis. Notably, the intersection network between the two independent studies resulted in 88 genes arranged in 13 sub-networks. Furthermore, GO analysis on the 79 MS risk genes arranged in networks in at least one of the two studies highlighted the categories “leukocyte activation,” “apoptosis,” and “positive regulation of macromolecule metabolic process” as well as the KEGG pathways “JAK-STAT signaling pathway,” “acute myeloid leukemia,” and “T cell receptor signaling” (30). Extending pathway analysis to all the 110 non-*MHC* variants identified after the ImmunoChip study also detected the NF- $\kappa$ B cascade to be significantly associated with MS risk genes (22, 31).

In a recent paper, a gene network candidate approach has highlighted the putative role of cellular adhesion molecules (CAMs) in MS pathology (32). By using eight GWAS datasets and considering all the genes interacting in the CAM pathway, five sub-networks were found associated with MS susceptibility, possibly connecting the risk to the regulation of blood–brain barrier (BBB) crossing by T cells.

---

## Genotype-Phenotype Correlations in MS

In addition to genetic factors contributing to MS susceptibility, specific variants also affect the clinical manifestation and the course of disease. Since the *HLA* locus is the first MS risk genetic determinant to be discovered and exerts the strongest influence on MS susceptibility, most of the genotype–phenotype studies are focused on *HLA* alleles. For instance, *HLA-DRB1\*15:01* carriage has been found to be consistently associated with lower age at the onset of disease (33). Furthermore, *HLA-DRB1\*15:01* seems to modulate the response toward glatiramer acetate, an immunomodulatory drug whose mechanism of action involves its binding to MHC class II molecules as an initial step (34). In addition, this allele was shown to increase the progression of MS brain pathology in terms of decline in brain magnetization transfer and T2 lesion load, as assessed by magnetic resonance imaging (MRI) (35). In contrast, the protective allele *HLA-B\*44:02* appears to preserve brain volume and reduce the burden of T2 hyper-intense lesions (36). In a recent work by our group, we carried out an analysis of the global contribution of the *HLA* locus to a number of clinical and MRI outcomes. We calculated the cumulative *HLA* genetic burden (HLAGB) resulting from carrying different alleles in different *HLA* genes in 652 MS patients who had comprehensive phenotypic information and 455 controls of European descent. As suggested by previous studies, we found that higher HLAGB scores are associated with younger age at onset and the atrophy of subcortical gray matter fraction in women with RR-MS. Conversely, *HLA-B\*44:02* showed a nominally protective effect for subcortical gray matter atrophy (37).

---

## Genetics of MS Animal Models

Although MS naturally occurs only in humans, different animal models have been developed in which a disease mimicking MS is induced artificially. According to the nature of the inducing agent, the current models can be grouped into three categories: autoimmune, viral, and neurotoxic (38). Among them, the most widely used model is experimental autoimmune encephalomyelitis (EAE), which falls in the first category. EAE is an experimental disease that can be induced in several species (e.g., rodents, primates, cats, dogs, and chickens) via immunization with spinal cord homogenates or, more often, with purified peptides containing specific sequences of myelin proteins such as myelin oligodendrocyte glycoprotein (MOG), myelin basic protein (MBP), and myelin proteolipid protein (PLP). EAE recapitulates several features of MS, including the influence of genetic



and environmental factors. This evidence has led to the search for the genetic determinants modulating EAE susceptibility with the intention of getting insights into the human counterpart.

Like MS, the *MHC* locus displays the biggest contribution to EAE susceptibility and manifestation, confirming the important role of T cells and antigen presentation in disease pathogenesis (39). In addition, at least 27 non-*MHC* loci (*Eae1-Eae27*) have been found to be associated with different traits of the disease, including incidence, onset, severity, and histopathology (40–42). Interestingly, a large part of them show sex specificity, possibly mimicking differences between genders in MS susceptibility. Most of these quantitative trait loci (QTL) have been mapped through genetic linkage studies in backcross mice derived from SJL/J and B10.S strains. The choice of these two specific strains is due to the fact that the former is highly susceptible to EAE induction, whereas the latter is characterized by poor encephalitogenic responses. More sophisticated approaches rely on the generation of congenic lines between these two strains, in order to fine-map the loci of interest. A recent study combining phenotype-selected congenic mice and gene interaction network analysis was able to identify candidate genes shared between EAE and MS within several *Eae* loci. Interestingly, most of these genes belong to evolutionary conserved pathways important for CD4<sup>+</sup> T helper-cell differentiation (43). Following a similar approach in a panel of consomic lines from the wild-derived PWD strain, the same group has also identified candidate genes associated with sexual dimorphism in CNS autoimmunity, highlighting the possible involvement of the mitogen-activated protein kinase (MAPK) pathway in driving gender-related EAE differences (44).

The EAE model offers an additional advantage through the option to easily engineer the mouse genome and test candidate genes for their putative effects on disease expression. Such an approach encompasses either the knockout of endogenous mouse genes evolutionarily related to the human genes of interest or the introduction of human alleles into the mouse genome. As a paradigmatic example of the first scenario, knockout mice lacking the orthologue of the human *IL7R $\alpha$*  gene were shown to be refractory to EAE induction, confirming the GWAS statistical association at the experimental level (45). The generation of transgenic mice carrying MS-relevant *HLA* alleles is instead the most common application of the second methodology. For instance, humanized mice expressing *HLA-DRB1\*15:01* and *HLA-DRB5\*01:01* alone or in combination, along with the human T cell receptor (TCR) specific for the MBP<sub>85–99</sub> peptide, have been instrumental in demonstrating the functional epistasis between the two alleles. Mice expressing both alleles indeed develop a milder form of a spontaneous MS-like disease as compared to mice expressing *DRB1\*15:01* only (46).

---

## Conclusion

GWA studies have undoubtedly energized and changed the field of MS genetics, allowing the discovery of more than a hundred risk loci following decades of unsuccessful attempts. A pressing challenge for the MS research community lies in the organization of the vast amount of genetic data finally available in a coherent

biological frame, which could explain the primary causes of the disease and its pathogenic processes. Considering the heterogeneity of MS and the intrinsic complexity of the human genome, a number of rational approaches can be envisioned to characterize the biological functions connected to MS susceptibility and pathophysiology.

First, fine-mapping projects will be required to refine the association in previously identified genomic loci and prioritize the candidate variants for further studies. This could be done by employing batteries of genetic markers saturating the region of interest as well as by analyzing populations with different LD patterns. In this regard, we recently reported the analysis resulting from genotyping an African American MS dataset with the ImmunoChip platform (47). African American genomes possess shorter LD, reflecting their unique ancestral history, a characteristic that facilitated narrowing down the association to tumor necrosis factor receptor superfamily member 14 (*TNFRSF14*) in a confirmed locus that included tetratricopeptide repeat domain 34 (*TTC34*), *LOC115110*, membrane metalloendopeptidase like 1 (*MMEL1*), *TNFRSF14*, and family with sequence similarity 213 member B (*FAM213B*) as candidate genes. These results support the utility of transancestral studies to better map the relevant variants within MS loci and suggest that common genetic basis underlies susceptibility across different ethnic groups.

Second, the increasing availability in public databases of gene expression datasets with relative genotype annotation can greatly facilitate the assessment of expression quantitative trait locus (eQTL) effects associated with the carriage of genetic variants relevant for MS. In this regard, computational strategies integrating gene expression measurements with summary GWAS data have been recently developed to identify genes whose cis-regulated expression is associated with complex traits, an approach called transcriptome-wide association study (TWAS) (48, 49). In addition, transcriptomic studies in relevant tissue samples from MS patients can also help identifying specific genetic signatures associated with disease susceptibility or progression. For example, following this approach, our group has shown that low levels of transducer of ERBB2-1 (*TOBI*) transcript in CD4<sup>+</sup> T cells are strongly associated with a higher risk of early conversion to clinically defined MS in patients experiencing a first demyelinating event in the CNS (50, 51).

Finally, recent remarkable innovations in genomic editing, such as the CRISPR-Cas9 or the TALEN systems (52), promise to reshape the next generation of functional studies aiming at translating genetic observation into mechanistic insights. These tools afford the modification of the genome at the single nucleotide level in a mono-allelic or bi-allelic fashion. Compared with classical methods of transgenesis, these new methodologies allow assessing the functional impact of genetic variants in physiological conditions via direct modification of the host genome in cell or animal models. These systems will be particularly relevant to efficiently screen regulatory variants mapping outside genes, whose function is less intuitive as compared to variants inducing amino acid substitutions. Furthermore, the possibility to simultaneously introduce multiple modifications in different genomic regions makes these systems suitable to explore possible epistatic effects between two or more variants (53).

In summary, an integrated approach involving multiple disciplines and technologies is likely to be the most effective way to address the complexity of MS

genetics and identify biologically meaningful correlations between risk variants and specific molecular functions.

**Acknowledgment:** This work was supported by FISM-Fondazione Italiana Sclerosi Multipla Senior Research Fellowship Cod. 2014/B/1 to AD.

### Conflict of interest

The authors declare no potential conflicts of interest with respect to research, authorship, and/or publication of this chapter.

### Copyright and permission statement

To the best of our knowledge, the materials included in this chapter do not violate copyright laws. All original sources have been appropriately acknowledged and/or referenced. Where relevant, appropriate permissions have been obtained from the original copyright holder(s).

---

## References

1. Hauser SL, Goodin DS. Multiple sclerosis and other demyelinating diseases. *Harrison's Principle of Internal Medicine*. 18th ed. McGraw-Hill; 2012, New York, USA
2. Compston A, Coles A. Multiple sclerosis. *Lancet*. 2008;372(9648):1502–17. [http://dx.doi.org/10.1016/S0140-6736\(08\)61620-7](http://dx.doi.org/10.1016/S0140-6736(08)61620-7)
3. Koch M, Kingwell E, Rieckmann P, Tremlett H. The natural history of primary progressive multiple sclerosis. *Neurology*. 2009;73(23):1996–2002. <http://dx.doi.org/10.1212/WNL.0b013e3181c5b47f>
4. Koch-Henriksen N, Sorensen PS. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet Neurol*. 2010;9(5):520–32. [http://dx.doi.org/10.1016/S1474-4422\(10\)70064-8](http://dx.doi.org/10.1016/S1474-4422(10)70064-8)
5. Sadovnick AD, Baird PA. The familial nature of multiple sclerosis: Age-corrected empiric recurrence risks for children and siblings of patients. *Neurology*. 1988;38(6):990–1. <http://dx.doi.org/10.1212/WNL.38.6.990>
6. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC, Canadian Collaborative Study G. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A*. 2003;100(22):12877–82. <http://dx.doi.org/10.1073/pnas.1932604100>
7. 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(Pt 2):449–55. <http://dx.doi.org/10.1093/brain/119.2.449>
8. Olsson T, Barcellos LF, Alfredsson L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nat Rev Neurol*. 2017;13(1):25–36. <http://dx.doi.org/10.1038/nrneurol.2016.187>
9. Rosati G. The prevalence of multiple sclerosis in the world: An update. *Neurol Sci*. 2001;22(2):117–39. <http://dx.doi.org/10.1007/s100720170011>
10. Sawcer S, Franklin RJ, Ban M. Multiple sclerosis genetics. *Lancet Neurol*. 2014;13(7):700–9. [http://dx.doi.org/10.1016/S1474-4422\(14\)70041-9](http://dx.doi.org/10.1016/S1474-4422(14)70041-9)
11. Didonna A, Oksenberg JR. Genetic determinants of risk and progression in multiple sclerosis. *Clin Chim Acta*. 2015;449:16–22. <http://dx.doi.org/10.1016/j.cca.2015.01.034>
12. 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(12):889–99. <http://dx.doi.org/10.1038/nrg1489>
13. Naito S, Namerow N, Mickey MR, Terasaki PI. Multiple sclerosis: Association with HL-A3. *Tissue Antigens*. 1972;2(1):1–4. <http://dx.doi.org/10.1111/j.1399-0039.1972.tb00111.x>
14. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. *Lancet*. 1972;1(7762):1240–1. [http://dx.doi.org/10.1016/S0140-6736\(72\)90962-2](http://dx.doi.org/10.1016/S0140-6736(72)90962-2)

15. Hollenbach JA, Pando MJ, Caillier SJ, Gourraud PA, Oksenberg JR. The killer immunoglobulin-like receptor KIR3DL1 in combination with HLA-Bw4 is protective against multiple sclerosis in African Americans. *Genes Immun.* 2016;17(3):199–202. <http://dx.doi.org/10.1038/gene.2016.5>
16. Patsopoulos NA, Barcellos LF, Hintzen RQ, Schaefer C, van Duijn CM, Noble JA, et al. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. *PLoS Genet.* 2013;9(11):e1003926. <http://dx.doi.org/10.1371/journal.pgen.1003926>
17. Moutsianas L, Jostins L, Beecham AH, Dilthey AT, Xifara DK, Ban M, et al. Class II HLA interactions modulate genetic risk for multiple sclerosis. *Nat Genet.* 2015;47(10):1107–13. <http://dx.doi.org/10.1038/ng.3395>
18. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med.* 2010;363(2):166–76. <http://dx.doi.org/10.1056/NEJMra0905980>
19. International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med.* 2007;357(9):851–62. <http://dx.doi.org/10.1056/NEJMo a073493>
20. Sawcer S, Ban M, Wason J, Dudbridge F. What role for genetics in the prediction of multiple sclerosis? *Ann Neurol.* 2010;67(1):3–10. <http://dx.doi.org/10.1002/ana.21911>
21. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476(7359):214–19. <http://dx.doi.org/10.1038/nature10251>
22. International Multiple Sclerosis Genetics Consortium. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet.* 2013;45(11):1353–60. <http://dx.doi.org/10.1038/ng.2770>
23. Patsopoulos NA, International Multiple Sclerosis Genetics Consortium. 200 loci complete the genetic puzzle of multiple sclerosis. American Society of Human Genetics (ASHG) 2016 Annual Meeting; 2016, Vancouver, BC, Canada.
24. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A.* 2012;109(4):1193–8. <http://dx.doi.org/10.1073/pnas.1119675109>
25. 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(9):1083–91. <http://dx.doi.org/10.1038/ng2103>
26. Galarza-Munoz G, Briggs FB, Evsyukova I, Schott-Lerner G, Kennedy EM, Nyanhete T, et al. Human epistatic interaction controls IL7R splicing and increases multiple sclerosis risk. *Cell.* 2017;169(1):72–84 e13. <http://dx.doi.org/10.1371/journal.pgen.1000322>
27. Maier LM, Lowe CE, Cooper J, Downes K, Anderson DE, Severson C, et al. IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production. *PLoS Genet.* 2009;5(1):e1000322.
28. Gregory AP, Dendrou CA, Attfield KE, Haghikia A, Xifara DK, Butter F, et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature.* 2012;488(7412):508–11. <http://dx.doi.org/10.1038/nature11307>
29. Didonna A, Isobe N, Caillier SJ, Li KH, Burlingame AL, Hauser SL, et al. A non-synonymous single-nucleotide polymorphism associated with multiple sclerosis risk affects the EVI5 interactome. *Hum Mol Genet.* 2015;24(24):7151–8. <http://dx.doi.org/10.1093/hmg/ddv412>
30. International Multiple Sclerosis Genetics Consortium. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet.* 2013;92(6):854–65. <http://dx.doi.org/10.1016/j.ajhg.2013.04.019>
31. Hussman JP, Beecham AH, Schmidt M, Martin ER, McCauley JL, Vance JM, et al. GWAS analysis implicates NF-kappaB-mediated induction of inflammatory T cells in multiple sclerosis. *Genes Immun.* 2016;17(5):305–12. <http://dx.doi.org/10.1038/gene.2016.23>
32. Damotte V, Guillot-Noel L, Patsopoulos NA, Madireddy L, El Behi M, International Multiple Sclerosis Genetics Consortium, et al. A gene pathway analysis highlights the role of cellular adhesion molecules in multiple sclerosis susceptibility. *Genes Immun.* 2014;15(2):126–32. <http://dx.doi.org/10.1038/gene.2013.70>

33. Masterman T, Ligers A, Olsson T, Andersson M, Olerup O, Hillert J. HLA-DR15 is associated with lower age at onset in multiple sclerosis. *Ann Neurol*. 2000;48(2):211–19. [http://dx.doi.org/10.1002/1531-8249\(200008\)48:2<211::AID-ANA11>3.0.CO;2-R](http://dx.doi.org/10.1002/1531-8249(200008)48:2<211::AID-ANA11>3.0.CO;2-R)
34. Fusco C, Andreone V, Coppola G, Luongo V, Guerini F, Pace E, et al. HLA-DRB1\*1501 and response to copolymer-1 therapy in relapsing-remitting multiple sclerosis. *Neurology*. 2001;57(11):1976–9. <http://dx.doi.org/10.1212/WNL.57.11.1976>
35. Tur C, Ramagopalan S, Altmann DR, Bodini B, Cercignani M, Khaleeli Z, et al. HLA-DRB1\*15 influences the development of brain tissue damage in early PPMS. *Neurology*. 2014;83(19):1712–18. <http://dx.doi.org/10.1212/WNL.0000000000000959>
36. Healy BC, Liguori M, Tran D, Chitnis T, Glanz B, Wolfish C, et al. HLA B\*44: Protective effects in MS susceptibility and MRI outcome measures. *Neurology*. 2010;75(7):634–40. <http://dx.doi.org/10.1212/WNL.0b013e3181ed9c9c>
37. Isobe N, Keshavan A, Gourraud PA, Zhu AH, Datta E, Schlaefer R, et al. Association of HLA genetic risk burden with disease phenotypes in multiple sclerosis. *JAMA Neurol*. 2016;73(7):795–802. <http://dx.doi.org/10.1001/jamaneuro.2016.0980>
38. Didonna A. Preclinical models of multiple sclerosis: Advantages and limitations towards better therapies. *Curr Med Chem*. 2016;23(14):1442–59. <http://dx.doi.org/10.2174/0929867323666160406121218>
39. Weissert R, Wallstrom E, Storch MK, Steffler A, Lorentzen J, Lassmann H, et al. MHC haplotype-dependent regulation of MOG-induced EAE in rats. *J Clin Invest*. 1998;102(6):1265–73. <http://dx.doi.org/10.1172/JCI3022>
40. Butterfield RJ, Sudweeks JD, Blankenhorn EP, Korngold R, Marini JC, Todd JA, et al. New genetic loci that control susceptibility and symptoms of experimental allergic encephalomyelitis in inbred mice. *J Immunol*. 1998;161(4):1860–7.
41. Butterfield RJ, Blankenhorn EP, Roper RJ, Zachary JF, Doerge RW, Sudweeks J, et al. Genetic analysis of disease subtypes and sexual dimorphisms in mouse experimental allergic encephalomyelitis (EAE): Relapsing/remitting and monophasic remitting/nonrelapsing EAE are immunogenetically distinct. *J Immunol*. 1999;162(5):3096–102.
42. Butterfield RJ, Blankenhorn EP, Roper RJ, Zachary JF, Doerge RW, Teuscher C. Identification of genetic loci controlling the characteristics and severity of brain and spinal cord lesions in experimental allergic encephalomyelitis. *Am J Pathol*. 2000;157(2):637–45. [http://dx.doi.org/10.1016/S0002-9440\(10\)64574-9](http://dx.doi.org/10.1016/S0002-9440(10)64574-9)
43. Blankenhorn EP, Butterfield R, Case LK, Wall EH, del Rio R, Diehl SA, et al. Genetics of experimental allergic encephalomyelitis supports the role of T helper cells in multiple sclerosis pathogenesis. *Ann Neurol*. 2011;70(6):887–96. <http://dx.doi.org/10.1002/ana.22642>
44. Bearoff F, Case LK, Kremensov DN, Wall EH, Saligrama N, Blankenhorn EP, et al. Identification of genetic determinants of the sexual dimorphism in CNS autoimmunity. *PLoS One*. 2015;10(2):e0117993. <http://dx.doi.org/10.1371/journal.pone.0117993>
45. Ashbaugh JJ, Brambilla R, Karmally SA, Cabello C, Malek TR, Bethea JR. IL7Ralpha contributes to experimental autoimmune encephalomyelitis through altered T cell responses and nonhematopoietic cell lineages. *J Immunol*. 2013;190(9):4525–34. <http://dx.doi.org/10.4049/jimmunol.1203214>
46. Gregersen JW, Kranc KR, Ke X, Svendsen P, Madsen LS, Thomsen AR, et al. Functional epistasis on a common MHC haplotype associated with multiple sclerosis. *Nature*. 2006;443(7111):574–7. <http://dx.doi.org/10.1038/nature05133>
47. Isobe N, Madireddy L, Khankhanian P, Matsushita T, Caillier SJ, More JM, et al. An ImmunoChip study of multiple sclerosis risk in African Americans. *Brain*. 2015;138(Pt 6):1518–30. <http://dx.doi.org/10.1093/brain/awv078>
48. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245–52. <http://dx.doi.org/10.1038/ng.3506>
49. Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating gene expression with summary association statistics to identify genes associated with 30 complex traits. *Am J Hum Genet*. 2017;100(3):473–87. <http://dx.doi.org/10.1016/j.ajhg.2017.01.031>

50. Corvol JC, Pelletier D, Henry RG, Caillier SJ, Wang J, Pappas D, et al. Abrogation of T cell quiescence characterizes patients at high risk for multiple sclerosis after the initial neurological event. *Proc Natl Acad Sci U S A*. 2008;105(33):11839–44. <http://dx.doi.org/10.1073/pnas.0805065105>
51. Didonna A, Cekanaviciute E, Oksenberg JR, Baranzini SE. Immune cell-specific transcriptional profiling highlights distinct molecular pathways controlled by Tob1 upon experimental autoimmune encephalomyelitis. *Sci Rep*. 2016;6:31603. <http://dx.doi.org/10.1038/srep31603>
52. Gaj T, Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013;31(7):397–405. <http://dx.doi.org/10.1016/j.tibtech.2013.04.004>
53. Schumann K, Lin S, Boyer E, Simeonov DR, Subramaniam M, Gate RE, et al. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc Natl Acad Sci U S A*. 2015;112(33):10437–42. <http://dx.doi.org/10.1073/pnas.1512503112>