

Review

Reworking GWAS Data to Understand the Role of Nongenetic Factors in MS Etiopathogenesis

Rosella Mechelli ^{1,*} , Renato Umeton ^{2,3,4}, Grazia Manfrè ¹, Silvia Romano ⁵,
Maria Chiara Buscarinu ⁵ , Virginia Rinaldi ⁵, Gianmarco Bellucci ⁵, Rachele Bigi ⁵,
Michela Ferraldeschi ⁵, Marco Salvetti ^{5,6,*}  and Giovanni Ristori ⁵

¹ IRCCS San Raffaele Pisana, San Raffaele Roma Open University, 00166 Rome, Italy; grazia.manfre@gmail.com

² Department of Informatics and Analytics, Dana-Farber Cancer Institute, Boston, MA 02215, USA; renato83@gmail.com

³ Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁴ Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA

⁵ Centre for Experimental Neurological Therapies (CENTERS), Department of Neurosciences, Mental Health and Sensory Organs, Sapienza University of Rome, 00189 Rome, Italy; silvia.romano@uniroma1.it (S.R.); mchiara.buscarinu@gmail.com (M.C.B.); virginia.rinaldi@mail.com (V.R.); bellucci.1638116@studenti.uniroma1.it (G.B.); rakele83@gmail.com (R.B.); michela.ferraldeschi@gmail.com (M.F.); giovanni.ristori@uniroma1.it (G.R.)

⁶ IRCCS Istituto Neurologico Mediterraneo (INM) Neuromed, 86077 Pozzilli, Italy

* Correspondence: rosella.mechelli@uniroma5.it (R.M.); marco.salvetti@uniroma1.it (M.S.)

Received: 7 November 2019; Accepted: 10 January 2020; Published: 14 January 2020



Abstract: Genome-wide association studies have identified more than 200 multiple sclerosis (MS)-associated loci across the human genome over the last decade, suggesting complexity in the disease etiology. This complexity poses at least two challenges: the definition of an etiological model including the impact of nongenetic factors, and the clinical translation of genomic data that may be drivers for new druggable targets. We reviewed studies dealing with single genes of interest, to understand how MS-associated single nucleotide polymorphism (SNP) variants affect the expression and the function of those genes. We then surveyed studies on the bioinformatic reworking of genome-wide association studies (GWAS) data, with aggregate analyses of many GWAS loci, each contributing with a small effect to the overall disease predisposition. These investigations uncovered new information, especially when combined with nongenetic factors having possible roles in the disease etiology. In this context, the interactome approach, defined as “modules of genes whose products are known to physically interact with environmental or human factors with plausible relevance for MS pathogenesis”, will be reported in detail. For a future perspective, a polygenic risk score, defined as a cumulative risk derived from aggregating the contributions of many DNA variants associated with a complex trait, may be integrated with data on environmental factors affecting the disease risk or protection.

Keywords: genome wide association studies; multiple sclerosis; gene-environment interaction; expression quantitative trait loci (eQTL); protein–protein interaction; pathway analysis; polygenic risk score

1. Introduction

Multiple sclerosis (MS) is the most common chronic inflammatory disease of the central nervous system causing neurological disability. There is a strong need to understand the still elusive MS cause(s) and use this knowledge to develop safe drugs that specifically target disease mechanisms. Like other complex traits, MS results from the interaction of genetic and environmental factors [1,2].

While genetic studies emphasize the importance of polymorphisms in several genes regulating the immune response, the role of environmental factors is less clear.

Current multiple sclerosis (MS) research has been significantly affected by the increasing number of disease-modifying therapies (DMTs) that are now available for its relapsing-remitting phase. Nonetheless, some aspects remain the focus of considerable interest, e.g., the treatment of progressive MS, that is just coming out of its infancy [3]; the identification of a ‘therapeutic window’ to potentially treat the prodromal, reversible phase of the disease, a topic related to MS endophenotypes (coming from magnetic resonance imaging and/or predictive biomarkers; [4]); and an etiologic therapeutic approach, that is still out of reach, since molecular models and the complex interactions between the heritable, nonheritable, and stochastic events underlying the disease need to be detailed [5–7].

Contributions with respect to the last point may be made through genome-wide association studies (GWAS) that provide an unequalled tool by which to construct etiologic patterns, as well as to deepen our understanding of the complex interactions between heritable and nonheritable causal factors underpinning neuroinflammation. GWAS are population-based association studies, comparing disease cases and controls for common genetic variants at variable frequencies in the general population. They have identified more than 200 MS-associated loci across the human genome over the last decade [8–11]. Methodological advances, increased sample sizes, and improved bio-statistical approaches have all contributed to important progress in the definition of the genetic architecture of MS; complexity is now evident for disease genetics, which had, until 15 years ago, essentially been limited to the role of human leukocyte antigen (HLA). GWAS results are derived from each common variant (signaled by a single nucleotide polymorphism (SNP)) that explains a small fraction of the risk/protection in a given population. The overall genetic risk is largely due to many common variants of small effect spread throughout the genome, except for loci lying in the HLA complex and a handful of rare gene variants that have recently been associated with MS [12–14].

This complexity poses at least two challenges: i) understanding the plausible causal effects of the gene regions identified, as well as the impact of nongenetic factors, that, as demonstrated by the twin studies, is substantial and different in diverse populations [15,16]; and ii) translating genomic data in clinical practice through the repurposing of available drugs or the discovery of new druggable targets.

Concerning the first point, several works have dealt with single genes of interest, trying to understand whether and how MS-associated SNP variants affect the expression and the function of those genes. There are a few examples where, among the many signals resulting from GWAS, the causal variant and the underlying mechanism has been defined, such as the following. SNPs in the *interleukin 2 receptor* gene seem to differently influence various autoimmune diseases [17–19]. Likewise, variants of *TYK 2* locus resulting from GWAS in different immune-mediated conditions normally mediate a balanced genetic effect, warranting a trade-off between autoimmunity and immunodeficiency [20]. Next, SNPs in the *interleukin 7 receptor* gene affect MS genetic risk through an epistatic interaction, that, altering the regulation of exon 6 splicing, impacts on MS risk through its soluble form [21]. A brilliant example of how clinical practice can be informed by reworking GWAS data was given in a work on an MS-associated SNP in the *tumor necrosis factor (TNF) receptor* gene region, where the authors demonstrated that the genic variant mimicked the effect of TNF-blocking drugs in increasing the risk of demyelinating disease [22]. Finally, another recent finding (overexpression of the B cell-promoting cytokine BAFF due to an MS-associated genetic variant; [23]) seems to be in line with the substantial impact of B cell-depleting therapies on MS course [24], but is yet to be reconciled with the worsening effects resulting from an anti-BAFF trial on disease activity [24,25]. Overall, these studies have shown that GWAS reworking allows us to open some instructive windows on MS mechanisms, but they contribute only in part to the construction of a comprehensive etiologic model.

Concerning the clinical translation of genomic data with a focus on therapeutic targets, some attempts based on MS GWAS have emerged [26,27]. An interesting recent resource, reported by Fang and colleagues, developed a priority index (PI) pipeline, based on GWAS variants for several immunopathological traits, including MS [28]. By integrating functional genomic, immune-related

annotations, and knowledge of network connectivity, the authors were able to maximize the genetic information for target validation of 30 immune traits. The added value of this approach was the incorporation of network connectivity information that increases enrichment for established therapeutic targets and helps overcome the possibility that many potential targets do not contain naturally-occurring variants that disrupt gene function and are associated with a relevant trait. A possible channel for the future development of similar approaches may lie in the bioinformatic reworking of GWAS data, considering the components of PI together with supposedly active, nonheritable factors which are known to interact with the genetic signals resulting from genome-scale data.

2. Bioinformatic Reworking of GWAS Data

The analysis performed on GWAS data considered SNPs exceeding a p -value threshold of 5×10^{-8} as being relevant for association with the disease. This approach stems from the assumption that genetic markers independently contribute to disease development. Other approaches took into consideration a looser cutoff (p -values of less than 0,05), with the assumption that a combined effect of many loci, albeit with a modest contribution, may account for overall disease susceptibility (Table 1).

Table 1. Main studies reworking genome-wide association studies (GWAS) data by considering multiple sclerosis (MS)-associated loci with different p -value cut-off.

Source	GWAS Datasets	p -Value Cut-Off of MS-Associated SNPs	Analysis	Results
Baranzini et al. 2009 [29]	IMSGC, 2007 [30]; Baranzini et al. 2009 [31]	$p < 0,05$	Pathway analysis and protein–protein interaction (PPI)-network	Identification of immunological and neural pathways enriched in MS
IMSGC et al. 2013 [32]	IMSGC, 2011 [8]; Patsopoulos et al. 2011 [33]	$p < 0,05$	protein-interaction-network-based pathway analysis (PINBPA)	Identification of new MS susceptibility loci association blocks (groups of contiguous genes with a p -value < 0.05) including <i>BCL10</i> , <i>CD48</i> , <i>REL</i> , <i>TRAF3</i> , and <i>TEC</i> .
Mechelli et al. 2013 [34]	IMSGC, 2011 [8]	$p < 0,05$	Candidate-interactome approach	EBV is the most MS-associated environmental risk factor interacting with MS-associated SNPs.
Ricigliano et al. 2015 [35]	IMSGC 2013 [9]; Cree et al. 2010 [36]	$p < 5 \times 10^8$	Transcription factor binding	EBNA2 and VDR binding sites are enriched in MS-associated loci.
Harley et al. 2018 [37]	NHGRI GWAS catalog	$p < 5 \times 10^8$	Transcription factor binding	EBNA2 binding sites are enriched in autoimmune-associated loci, including MS.
IMSGC et al. 2019 [38]	IMSGC 2019 [11]	$p < 5 \times 10^8$	Prioritization of cell specific gene/protein networks	Explanation of the potential role of GAWS signals in a tissue/cell-specific manner: identification of cell-specific susceptibility pathways.
IMSGC 2019 [11]	IMSGC 2019 [11]	$p < 5 \times 10^8$	Multiple approaches: cell-specific eQTL, pathway analyses; PPI	Prioritization of genes putatively associated with the disease, and identification of possible major implications for resident microglia and the B cell in MS.

Pathway analysis was one of the approaches selected for GWAS data interpretation, based on the identification of molecular function and biological processes that can be targeted by disease-associated SNPs [39]. An interpretation of nominally MS-associated SNPs ($p < 0,05$), obtained from two studies [30,31] and combining pathway analysis and protein–protein interaction networks (PINs), identified subnetworks of genes involved in several immunological and neural pathways which are enriched in MS [29]. In another study, this approach highlighted the oxidative stress and immune dysfunction pathways as being relevant in primary progressive MS, despite the fact that conventional GWAS analysis has not shown any association at the single SNP level [40]. To further unravel the heritable factors that are potentially involved in MS, a PIN-based pathway analysis (PINBPA, [41]) was conducted. Specifically, significant MS-associated SNPs ($p < 0,05$) of two independent GWAS datasets [8,33] were matched with data of protein–protein interaction networks that were subsequently gathered in a pathway analysis. This combined approach

showed that proteins encoded by genes carrying risk variants are more likely to interact and share the same or related pathways. Moreover, the integration with protein–protein interaction pathways suggested new MS-susceptibility loci, including TNF-receptor-associated factor 3 (TRAF3), B cell membrane protein (CD48), B cell lymphoma 10 (BCL10), v-rel reticuloendotheliosis viral oncogene homolog (REL), and TEC protein tyrosine kinase (TEC) [32]. This approach proved to be informative. In fact, two *TRAF3* MS risk SNPs (rs12147246 and rs12588969; [11]) were then identified as being of genome-wide significance, while other studies have shown their involvement in a dysregulated response to EBV infection [42], i.e., one of the main environmental factors associated with MS [43,44].

Another approach of GWAS reworking is based on the wealth of information on regulatory elements which is available thanks to the efforts of Encyclopedia of DNA elements (ENCODE) and Regulome Epigenomics Consortia [45,46]. These repositories can help to identify disease-associated functional variants that could have an impact on specific cell types. Epigenetic changes, including DNA methylation, posttranslational modification of histones, and the synthesis of noncoding RNAs, are representative of molecular mechanisms through which environmental signals are translated within the cells to change their gene expression [47,48]. In MS pathogenesis, a great deal of evidence suggests the integration of the risk related to genetic predisposition with cell-type-specific epigenetic changes occurring in the immune system and in the brain in response to environmental stimuli [49–51]. A reworking of GWAS data, aimed at identifying the cellular type where the MS-associated variants might exert functional effects, was recently performed in an association study that analyzed a total of 47,351 cases and 68,284 healthy controls [38]. Taking into consideration the 200 autosomal susceptibility variants outside the major histocompatibility complex (MHC), the authors considered all the regulatory elements that could be affected by the presence of these variants in a cell-specific manner, and created a cell-specific protein network. This paradigmatic approach to decoding the disease risk in a cell/tissue specific context suggested MS-associated variants operative in CNS resident microglia as important contributors (besides those of peripheral immune cells) to disease development [11,38].

A further level of complexity stems from the fact that many disease-associated polymorphisms identified by GWAS lie within regulatory regions of the genome. In fact, many MS-associated variants are located within noncoding regions [8,50] with a potential impact on gene expression as expression quantitative trait loci (eQTL). Currently available data on eQTL suggest a relationship between gene variants and the gene expression of certain cell types; these effects can be disease-specific, and possibly depend on external stimuli (i.e. signaling pathways activated by cytokines affecting a specific cell type) [50,52]. Specifically, most of eQTL have been studied in healthy volunteers, with the assumption that the effect of a single nucleotide variant may be independent from the diseased condition. Conversely, recent data has revealed a specificity of the eQTL effects in diseased subjects. A recent work based on RNA-Seq in PBMC from MS patients to identify eQTLs in regions centered on at-risk SNPs showed 77 statistically-relevant eQTL associations, 40% of which were more pronounced in MS patients compared with noninflammatory neurological disease patients [52]. Another interesting approach of eQTL analysis was based on public RNA-sequencing and microarray data of blood-derived cells. A group investigated the role of SNP rs1414273, which is located within the microRNA-548ac stem-loop sequence in the first intron of the CD58 gene. They provided evidence that this MS-associated SNP might alter Droscha cleavage activity, thus modifying CD58 and microRNA-548ac gene expression in immune cells, a change that has already been reported to be relevant for MS development [53–55].

3. Interactome-Based Approach

To understand a complex disease such as MS, the impact of environmental factors should be included in investigations. The influence of genetic variants on environmental exposures that associate with multifactorial diseases is far from being fully understood, particularly at a genome-wide level. The complexity of interactions between genes and environment needs to be explored using analytical approaches which are capable of considering many variables simultaneously; this may account for the so called ‘missing heritability’. New investigations on the aggregate analysis of many GWAS loci,

each contributing with a small effect to overall disease predisposition, might uncover new information when combined with nongenetic factors that can play a role in disease etiology. We tried to implement this concept, using available actionable data on interactions at the protein level between human gene products and exposures. The possible causal significance of environmental exposures was investigated by measuring, at a genome-wide level, the enrichment of MS-associated genetic variants in genomic regions coding for proteins interacting with the exposure (i.e. influencing the exposure). Being aware that most SNPs associated with complex traits fall within regulatory genomic regions with distal effects or even in trans effects, we operationally chose to consider a 20kbp distance between MS variants and the nearest genes. This analysis was performed using Association LIst Go AnnoTatOR (ALIGATOR) bioinformatic tool [34] to search for statistical enrichment in associations between interactome's genes and MS genome-wide association data published by IMISGC in 2011 [8].

The interactomes can be defined as “modules of genes whose products are known to physically interact with environmental or human factors with plausible, uncertain, or unlikely relevance for MS pathogenesis” [56]. The analysis was centered on viral interactomes, based on the classical hypothesis of a viral etiology of MS, examining only direct interactions between viral and human proteins, i.e., those of primary importance for the phenotypic impact of the environment within the host physiology [57,58]. This approach took into consideration MS-associated SNPs contributing with a small effect to overall disease susceptibility (p -value cut-off of association less than 0.05), thus starting from GWAS signal values which were well beyond those considered to be at the genome-wide significance cutoff. This candidate-interactome analysis allowed us to determine a relative low number of known or new nongenetic factors which are associated with MS risk/protection, and to formalize interplays between the heritable and nonheritable elements of possible causal nature. Specifically, the relevance and complexity of interactions between host genotype and EBV were disclosed, highlighting, through a pathway analysis, some cellular functions that may be affected by such interactions (Figure 1). Moreover, this approach has the power to identify other viruses, and related interacting proteins, which are potentially relevant for MS etiology [56].

Other complementary approaches may unveil the relevance of interactions between environmental factors and genetic predisposition. MS susceptibility regions may be preferentially targeted by both the viral and cellular proteins which are directly involved in molecular mechanisms that are able to translate environmental signals into cellular perturbations. Epstein-Barr nuclear 2 (EBNA2, a viral transactivator of viral and cellular genes) binding motifs were shown to be significantly enriched in genomic intervals associated with MS [35,37]. Another layer of complexity is represented by a striking overlap between MS-associated loci, EBNA2 binding motifs, and Vitamin D receptor binding sites [35], suggesting that a complex interplay between host genetic variants and known associated environmental factors [5] may contribute to disease development. These results indicate that looking at genetic predisposition through the lens of nonheritable risk factors represents an advancement in studies aiming at disclosing functional meaning and prioritizing genetic variants coming from GWAS data.

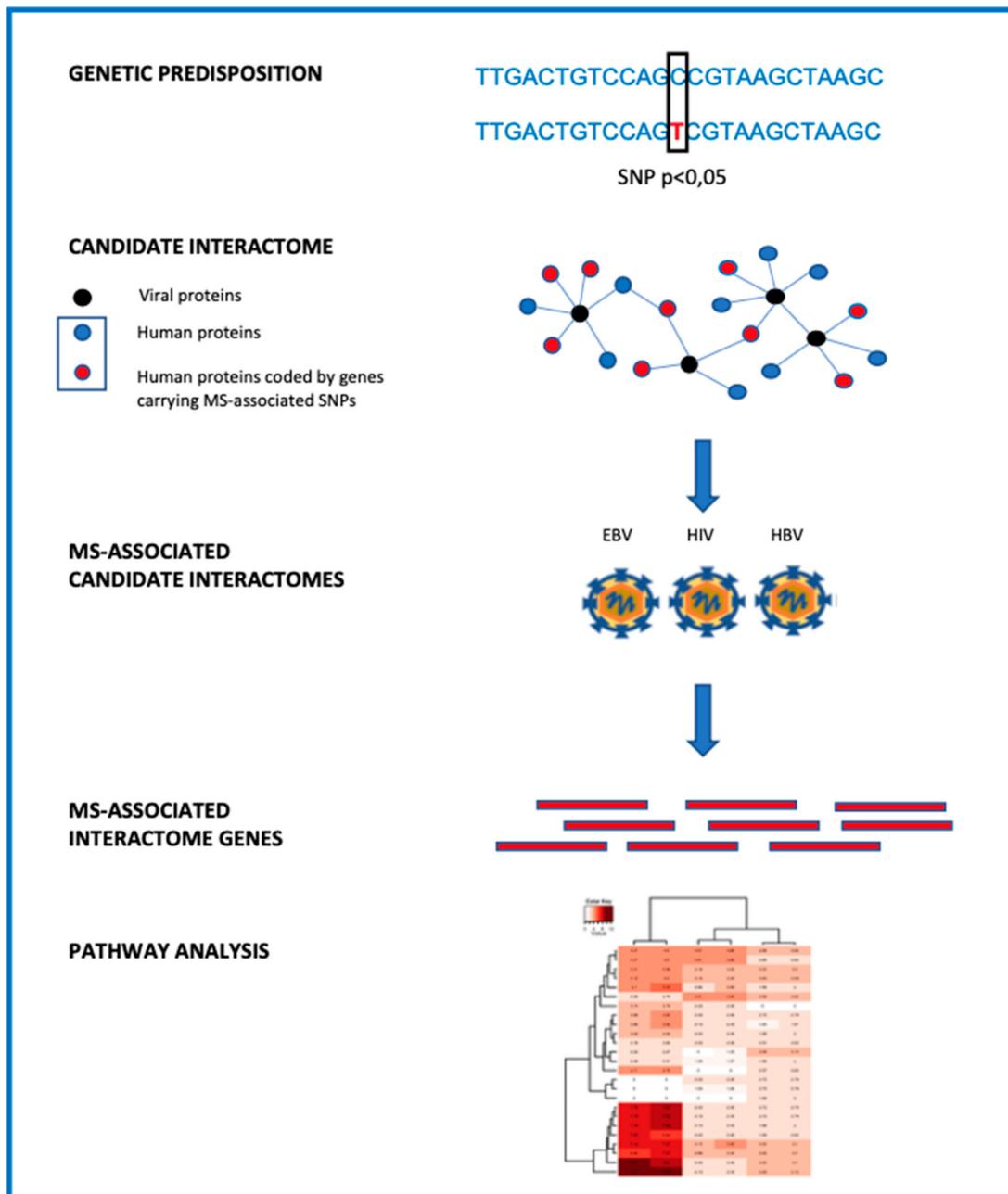


Figure 1. Schematic representation of the candidate-interactome approach.

4. Future Perspectives

The above studies have highlighted the need to integrate different analytical approaches with the aim of deepening our understanding of MS pathophysiology and etiology. New approaches, still experimental, could have future applications in quantifying the overall burden of genetic risk factors or serving as a stratification biomarker for treatment optimization. A promising approach in this field is the polygenic risk score (PRS), also known as ‘genetic risk score’. PRS can be defined as a cumulative risk derived from aggregating the contributions of many DNA variants associated with a complex trait or disease. Existing research using PRS mainly focuses on two problems: association analysis and outcome prediction.

Although the use of PRS has not yet achieved clinical accuracy levels, interesting potential perspectives have emerged in diseases like cancer [59], psoriasis [60], rheumatoid arthritis [61], mental

disorders [62,63], atherosclerosis [64], Type 2 diabetes [65,66], asthma [67], Parkinson's disease [68,69], and cardiovascular diseases (CVD) [70], including coronary heart disease (CHD) [71]. Polygenic risk scores can help to select a therapy for disease prevention. For example, statin therapy was shown to lead to a greater relative risk reduction for coronary heart disease among patients at high genetic risk score compared with patients at low genetic risk [63]. Polygenic risk scores have also been used to explore the genetic overlap between different diseases (e.g., application of schizophrenia-specific PRS to bipolar disorder), where the PRS derived from one disease is evaluated in another.

In a complex disease such as MS, the PRS could be correlated with phenotypic data (the 'classical' clinical and neuradiological parameters, but also to some emerging biomarkers that are being intensively investigated in biological fluids) as well as with MS endophenotypes (such as the radiologically-isolated syndrome; [72]) to generate a complex risk model which is able to predict the cumulative effects leading to overt disease onset. Moreover, the PRS may be integrated once again by data on environmental factors affecting disease risk or protection. In this context, it is possible to speculate on the calculation of an individual's risk of disease based on the "candidate interactome" approach, integrating genetic-environmental interplay as reported in the previous paragraph. Also, the recently reported genomic variants of EBV [73,74], the main environmental risk factor for MS development, may prompt us to define a PRS combining the MS-associated loci of the host with the risky or protecting genomic variants of the virus. All these approaches would require a model that was calibrated to proportionately include genetic variants (G) and environmental exposures (E), i.e., taking into account G, E, and their interaction Gx E. Such a tool does not currently seem to be within reach for complex diseases, where problems such as pleiotropy, causal effect estimation, and other questions of etiologic epidemiology remain. Overall, the etiology of multifactorial diseases seems to be more complex than anticipated [75]; therefore, their inherently multifaceted nature requires new models. Nonetheless, emerging bio-statistical approaches seem poised to start a post-GWAS era [76].

Author Contributions: All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication. R.M., R.U., G.M., V.R. and M.C.B. wrote the draft. S.R. and M.F. curated the table and the figure. R.B. and G.B. selected the bibliography. R.M., G.R. and M.S. wrote the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Italian Foundation of Multiple Sclerosis (FISM, grant number 2014/R/12 to RM).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Olsson, T.; Barcellos, L.F.; Alfredsson, L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nat. Rev. Neurol.* **2017**, *13*, 25–36. [[CrossRef](#)] [[PubMed](#)]
2. Kakalacheva, K.; Lunemann, J.D. Environmental triggers of multiple sclerosis. *FEBS Lett.* **2011**, *585*, 3724–3729. [[CrossRef](#)] [[PubMed](#)]
3. Ontaneda, D.; Thompson, A.J.; Fox, R.J.; Cohen, J.A. Progressive multiple sclerosis: Prospects for disease therapy, repair, and restoration of function. *Lancet* **2017**, *389*, 1357–1366. [[CrossRef](#)]
4. Ramagopalan, S.V.; Dobson, R.; Meier, U.C.; Giovannoni, G. Multiple sclerosis: Risk factors, prodromes, and potential causal pathways. *Lancet Neurol.* **2010**, *9*, 727–739. [[CrossRef](#)]
5. Dobson, R.; Giovannoni, G. Multiple sclerosis—A review. *Eur. J. Neurol.* **2019**, *26*, 27–40. [[CrossRef](#)] [[PubMed](#)]
6. Bordi, I.; Umeton, R.; Ricigliano, V.A.; Annibali, V.; Mechelli, R.; Ristori, G.; Grassi, F.; Salvetti, M.; Sutera, A. A mechanistic, stochastic model helps understand multiple sclerosis course and pathogenesis. *Int. J. Genom.* **2013**, *2013*, 910321. [[CrossRef](#)] [[PubMed](#)]
7. Bordi, I.; Ricigliano, V.A.; Umeton, R.; Ristori, G.; Grassi, F.; Crisanti, A.; Sutera, A.; Salvetti, M. Noise in multiple sclerosis: Unwanted and necessary. *Ann. Clin. Transl. Neurol.* **2014**, *1*, 502–511. [[CrossRef](#)]
8. 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*, 214–219. [[CrossRef](#)]

9. IMSSGC. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **2013**, *45*, 1353–1360. [[CrossRef](#)]
10. Cotsapas, C.; Mitrovic, M. Genome-wide association studies of multiple sclerosis. *Clin. Trans. Immunol.* **2018**, *7*, e1018. [[CrossRef](#)]
11. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **2019**, *365*, 6460.
12. International Multiple Sclerosis Genetics Consortium. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. *Cell* **2018**, *175*, 1679–1687. [[CrossRef](#)] [[PubMed](#)]
13. Vilariño-Güell, C.; Zimprich, A.; Martinelli-Boneschi, F.; Herculano, B.; Wang, Z.; Matesanz, F.; Urcelay, E.; Vandebroek, K.; Leyva, L.; Gris, D.; et al. Exome sequencing in multiple sclerosis families identifies 12 candidate genes and nominates biological pathways for the genesis of disease. *PLoS Genet.* **2019**, *15*, e1008180. [[CrossRef](#)] [[PubMed](#)]
14. Vidmar, L.; Maver, A.; Drulović, J.; Sepčić, J.; Novaković, I.; Ristić, S.; Šega, S.; Peterlin, B. Multiple Sclerosis patients carry an increased burden of exceedingly rare genetic variants in the inflammasome regulatory genes. *Sci. Rep.* **2019**, *9*, 9171. [[CrossRef](#)] [[PubMed](#)]
15. Ristori, G.; Cannoni, S.; Stazi, M.A.; Vanacore, N.; Cotichini, R.; Alfò, M.; Pugliatti, M.; Sotgiu, S.; Solaro, C.; Bompreszi, R.; et al. Multiple sclerosis in twins from continental Italy and Sardinia: A nationwide study. *Ann. Neurol.* **2006**, *59*, 27–34. [[CrossRef](#)] [[PubMed](#)]
16. Fagnani, C.; Neale, M.C.; Nisticò, L.; Stazi, M.A.; Ricigliano, V.A.; Buscarinu, M.C.; Salvetti, M.; Ristori, G. Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. *Mult. Scler.* **2015**, *21*, 1404–1413. [[CrossRef](#)]
17. Maier, L.M.; Lowe, C.E.; Cooper, J.; Downes, K.; Anderson, D.E.; Severson, C.; Clark, P.M.; Healy, B.; Walker, N.; Aubin, C.; et al. IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production. *PLoS Genet.* **2009**, *5*, e1000322. [[CrossRef](#)]
18. Wallace, C.; Cutler, A.J.; Pontikos, N.; Pekalski, M.L.; Burren, O.S.; Cooper, J.D.; García, A.R.; Ferreira, R.C.; Guo, H.; Walker, N.M.; et al. Dissection of a Complex Disease Susceptibility Region Using a Bayesian Stochastic Search Approach to Fine Mapping. *PLoS Genet.* **2015**, *11*, e1005272. [[CrossRef](#)]
19. Afanasyeva, M.A.; Putlyaeva, L.V.; Demin, D.E.; Kulakovskiy, I.V.; Vorontsov, I.E.; Fridman, M.V.; Makeev, V.J.; Kuprash, D.V.; Schwartz, A.M. The single nucleotide variant rs12722489 determines differential estrogen receptor binding and enhancer properties of an IL2RA intronic region. *PLoS ONE* **2017**, *12*, e0172681. [[CrossRef](#)]
20. Dendrou, C.A.; Cortes, A.; Shipman, L.; Evans, H.G.; Attfield, K.E.; Jostins, L.; Barber, T.; Kaur, G.; Kuttikkatte, S.B.; Leach, O.A.; et al. Resolving TYK2 locus genotype-to-phenotype differences in autoimmunity. *Sci. Transl. Med.* **2016**, *8*, 363ra149. [[CrossRef](#)]
21. Galarza-Muñoz, G.; Briggs, F.B.S.; Evsyukova, I.; Schott-Lerner, G.; Kennedy, E.M.; Nyanhete, T.; Wang, L.; Bergamaschi, L.; Widen, S.G.; Tomaras, G.D.; et al. Human Epistatic Interaction Controls IL7R Splicing and Increases Multiple Sclerosis Risk. *Cell* **2017**, *169*, 72–84. [[CrossRef](#)] [[PubMed](#)]
22. Gregory, A.P.; Dendrou, C.A.; Attfield, K.E.; Haghikia, A.; Xifara, D.K.; Butter, F.; Poschmann, G.; Kaur, G.; Lambert, L.; Leach, O.A.; et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* **2012**, *488*, 508–511. [[CrossRef](#)] [[PubMed](#)]
23. Steri, M.; Orrù, V.; Idda, M.L.; Pitzalis, M.; Pala, M.; Zara, I.; Sidore, C.; Faà, V.; Floris, M.; Deiana, M.; et al. Overexpression of the Cytokine BAFF and Autoimmunity Risk. *N. Engl. J. Med.* **2017**, *376*, 1615–1626. [[CrossRef](#)] [[PubMed](#)]
24. Milo, R. Therapies for multiple sclerosis targeting B cells. *Croat. Med. J.* **2019**, *60*, 87–98. [[CrossRef](#)] [[PubMed](#)]
25. Kappos, L.; Hartung, H.P.; Freedman, M.S.; Boyko, A.; Radü, E.W.; Mikol, D.D.; Lamarine, M.; Hyvert, Y.; Freudensprung, U.; Plitz, T.; et al. Atacept in multiple sclerosis (ATAMS): A randomised, placebo-controlled, double-blind, phase 2 trial. *Lancet Neurol.* **2014**, *13*, 353–363. [[CrossRef](#)]
26. Damotte, V.; Guillot-Noel, L.; Patsopoulos, N.A.; Madireddy, L.; El Behi, M.; International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2; De Jager, P.L.; Baranzini, S.E.; Cournu-Rebeix, I.; et al. A gene pathway analysis highlights the role of cellular adhesion molecules in multiple sclerosis susceptibility. *Genes Immun.* **2014**, *15*, 126–132. [[CrossRef](#)] [[PubMed](#)]

27. La Starza, S.; Ferraldeschi, M.; Buscarinu, M.C.; Romano, S.; Fornasiero, A.; Mechelli, R.; Umeton, R.; Ristori, G.; Salvetti, M. Genome-Wide Multiple Sclerosis Association Data and Coagulation. *Front. Neurol.* **2019**, *10*, 95. [[CrossRef](#)]
28. Fang, H.; ULTRA-DD Consortium; De Wolf, H.; Knezevic, B.; Burnham, K.L.; Osgood, J.; Sanniti, A.; Lledó Lara, A.; Kasela, S.; De Cesco, S.; et al. A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat. Genet.* **2019**, *51*, 1082–1091. [[CrossRef](#)]
29. Baranzini, S.E.; Galwey, N.W.; Wang, J.; Khankhanian, P.; Lindberg, R.; Pelletier, D.; Wu, W.; Uitdehaag, B.M.; Kappos, L.; Pol-Man, C.H.; et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum. Mol. Genet.* **2009**, *18*, 2078–2090. [[CrossRef](#)]
30. IMSGC; Hafler, D.A.; Compston, A.; Sawcer, S.J.; Lander, E.S.; Daly, M.J.; De Jager, P.L.; de Bakker, P.I.; Gabriel, S.B.; Mirel, D.B.; et al. Risk alleles for multiple sclerosis identified by a Genome Wide Study. *N. Engl. J. Med.* **2007**, *357*, 851–862.
31. Baranzini, S.E.; Wang, J.; Gibson, R.A.; Galwey, N.; Naegelin, Y.; Barkhof, F.; Radue, E.W.; Lindberg, R.L.; Uitdehaag, B.M.; Johnson, M.R.; et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum. Mol. Genet.* **2009**, *18*, 767–778. [[CrossRef](#)] [[PubMed](#)]
32. 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*, 854–865. [[CrossRef](#)] [[PubMed](#)]
33. Patsopoulos, N.A.; Bayer Pharma MS Genetics Working Group; Steering Committees of Studies Evaluating IFN β -1b and a CCR1- Antagonist; ANZgene Consortium; GeneMSA; International Multiple Sclerosis Genetics Consortium; Reischl, J.; Lehr, S.; Bauer, D.; Heubach, J.; et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann. Neurol.* **2011**, *70*, 897–912. [[CrossRef](#)] [[PubMed](#)]
34. Holmans, P.; Green, E.K.; Pahwa, J.S.; Ferreira, M.A.; Purcell, S.M.; Sklar, P.; Owen, M.J.; O'Donovan, M.C.; Craddock, N.; Consortium, W.T.C.C. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am. J. Hum. Genet.* **2009**, *85*, 13–24. [[CrossRef](#)] [[PubMed](#)]
35. Ricigliano, V.A.; Handel, A.E.; Sandve, G.K.; Annibaldi, V.; Ristori, G.; Mechelli, R.; Cader, M.Z.; Salvetti, M. EBNA2 binds to genomic intervals associated with multiple sclerosis and overlaps with vitamin D receptor occupancy. *PLoS ONE* **2015**, *10*, e0119605. [[CrossRef](#)] [[PubMed](#)]
36. Cree, B.A.; Rioux, J.D.; McCauley, J.L.; Gourraud, P.F.D.; Goyette, P.; McElroy, J.; De Jager, P.; Santaniello, A.; Vyse, T.J.; Gregersen, P.K.; et al. A major histocompatibility Class I locus contributes to multiple sclerosis susceptibility independently from HLA-DRB1*15:01. *PLoS ONE* **2010**, *5*, e11296. [[CrossRef](#)] [[PubMed](#)]
37. Harley, J.B.; Chen, X.; Pujato, M.; Miller, D.; Maddox, A.; Forney, C.; Magnussen, A.F.; Lynch, A.; Chetal, K.; Yukawa, M.; et al. Transcription factors operate across disease loci with EBNA2 implicated in autoimmunity. *Nat. Genet.* **2018**, *50*, 699–707. [[CrossRef](#)]
38. International Multiple Sclerosis Genetics Consortium. A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. *Nat. Commun.* **2019**, *10*, 2236. [[CrossRef](#)]
39. Cirillo, E.; Parnell, L.D.; Evelo, C.T. A Review of Pathway-Based Analysis Tools That Visualize Genetic Variants. *Front. Genet.* **2017**, *8*, 174. [[CrossRef](#)]
40. Giacalone, G.; Clarelli, F.; Osiceanu, A.M.; Guaschino, C.; Brambilla, P.; Sorosina, M.; Liberatore, G.; Zauli, A.; Esposito, F.; Rodegher, M.; et al. Analysis of genes, pathways and networks involved in disease severity and age at onset in primary-progressive multiple sclerosis. *Mult. Scler.* **2015**, *21*, 1431–1442. [[CrossRef](#)]
41. Wang, L.; Mousavi, P.; Baranzini, S.E. iPINBPA: An integrative network-based functional module discovery tool for genome-wide association studies. *Pac. Symp. Biocomput.* **2015**, 255–266.
42. Afrasiabi, A.; Parnell, G.P.; Fewings, N.; Schibeci, S.D.; Basuki, M.A.; Chandramohan, R.; Zhou, Y.; Taylor, B.; Brown, D.A.; Swaminathan, S.; et al. Evidence from genome wide association studies implicates reduced control of Epstein-Barr virus infection in multiple sclerosis susceptibility. *Genome Med.* **2019**, *11*, 26. [[CrossRef](#)]
43. Salvetti, M.; Giovannoni, G.; Aloisi, F. Epstein-Barr virus and multiple sclerosis. *Curr. Opin. Neurol.* **2009**, *22*, 201–206. [[CrossRef](#)] [[PubMed](#)]
44. Levin, L.I.; Munger, K.L.; O'Reilly, E.J.; Falk, K.I.; Ascherio, A. Primary infection with the Epstein-Barr virus and risk of multiple sclerosis. *Ann. Neurol.* **2010**, *67*, 824–830. [[CrossRef](#)] [[PubMed](#)]
45. ENCODE Project Consortium. The ENCODE (ENCyclopedia of DNA Elements) Project. *Science* **2004**, *306*, 636–640. [[CrossRef](#)]

46. Roadmap Epigenomics Consortium; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330. [[CrossRef](#)]
47. Jirtle, R.L.; Skinner, M.K. Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* **2007**, *8*, 253–262. [[CrossRef](#)]
48. Mechelli, R.; Annibaldi, V.; Ristori, G.; Vittori, D.; Coarelli, G.; Salvetti, M. Multiple sclerosis etiology: Beyond genes and environment. *Expert Rev. Clin. Immunol.* **2010**, *6*, 481–490. [[CrossRef](#)]
49. Pedre, X.; Mastronardi, F.; Bruck, W.; López-Rodas, G.; Kuhlmann, T.; Casaccia, P. Changed histone acetylation patterns in normal-appearing white matter and early multiple sclerosis lesions. *J. Neurosci.* **2011**, *31*, 3435. [[CrossRef](#)]
50. Farh, K.K.H.; Marson, A.; Zhu, J.; Kleinewietfeld, M.; Housley, W.J.; Beik, S.; Shores, N.; Whitton, H.; Ryan, R.J.H.; Shishkin, A.A.; et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **2015**, *518*, 337–343. [[CrossRef](#)]
51. Castro, K.; Casaccia, P. Epigenetic modifications in brain and immune cells of multiple sclerosis patients. *Mult. Scler.* **2018**, *24*, 69–74. [[CrossRef](#)] [[PubMed](#)]
52. James, T.; Lindén, M.; Morikawa, H.; Fernandes, S.J.; Ruhrmann, S.; Huss, M.; Brandi, M.; Piehl, F.; Jagodic, M.; Tegnér, J.; et al. Impact of genetic risk loci for multiple sclerosis on expression of proximal genes in patients. *Hum. Mol. Genet.* **2018**, *27*, 912–928. [[CrossRef](#)] [[PubMed](#)]
53. De Jager, P.L.; Baecher-Allan, C.; Maier, L.M.; Arthur, A.T.; Ottoboni, L.; Barcellos, L.; McCauley, J.L.; Sawcer, S.; Goris, A.; Saarela, J.; et al. The role of the CD58 locus in multiple sclerosis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5264–5269. [[CrossRef](#)] [[PubMed](#)]
54. Hecker, M.; Fitzner, B.; Blaschke, J.; Blaschke, P.; Zettl, U.K. Susceptibility variants in the CD58 gene locus point to a role of microRNA-548ac in the pathogenesis of multiple sclerosis. *Mutat. Res. Rev. Mutat. Res.* **2015**, *763*, 161–167. [[CrossRef](#)]
55. Hecker, M.; Boxberger, N.; Illner, N.; Fitzner, B.; Schroder, I.; Winkelmann, A.; Dudesek, A.; Meister, S.; Koczan, D.; Lorenz, P.; et al. A genetic variant associated with multiple sclerosis inversely affects the expression of CD58 and microRNA-548ac from the same gene. *PLoS Genet.* **2019**, *15*, e1007961. [[CrossRef](#)]
56. Mechelli, R.; Umeton, R.; Policano, C.; Annibaldi, V.; Coarelli, G.; Ricigliano, V.A.; Vittori, D.; Fornasiero, A.; Buscarinu, M.C.; International Multiple Sclerosis Genetics Consortium; et al. A “candidate-interactome” aggregate analysis of genome-wide association data in multiple sclerosis. *PLoS ONE* **2013**, *8*, e63300. [[CrossRef](#)]
57. Menche, J.; Sharma, A.; Kitsak, M.; Ghiassian, S.D.; Videl, M.; Loscalzo, J.; Barabasi, A.L. Uncovering disease-disease relationships through the incomplete interactome. *Science* **2015**, *347*, 841–849. [[CrossRef](#)]
58. Huttlin, E.L.; Bruckner, R.J.; Paulo, J.A.; Cannon, J.R.; Ting, L.; Baltier, K.; Colby, G.; Gebreab, F.; Gygi, M.P.; Parzen, H.; et al. Architecture of the human interactome defines protein communities and disease networks. *Nature* **2017**, *545*, 505–509. [[CrossRef](#)]
59. Li, J.; Holm, J.; Bergh, J.; Eriksson, M.; Darabi, H.; Lindström, L.S.; Törnberg, S.; Hall, P.; Czene, K. Breast cancer genetic risk profile is differentially associated with interval and screen-detected breast cancers. *Ann. Oncol.* **2015**, *26*, 517–522. [[CrossRef](#)]
60. Khera, A.V.; Chaffin, M.; Aragam, K.G.; Haas, M.E.; Roselli, C.; Choi, S.H.; Natarajan, P.; Lander, E.S.; Lubitz, S.A.; Ellinor, P.T.; et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **2018**, *50*, 1219–1224. [[CrossRef](#)]
61. Karlson, E.W.; Chibnik, L.B.; Kraft, P.; Cui, J.; Keenan, B.T.; Ding, B.; Raychaudhuri, S.; Klareskog, L.; Alfredsson, L.; Plenge, R.M. Cumulative association of 22 genetic variants with sero positive rheumatoid arthritis risk. *Ann. Rheum. Dis.* **2010**, *69*, 1077–1085. [[CrossRef](#)] [[PubMed](#)]
62. Sengupta, S.M.; MacDonald, K.; Fathalli, F.; Yim, A.; Lepage, M.; Iyer, S.; Malla, A.; Joobor, R. Polygenic risk score associated with specific symptom dimension in first-episode psychosis. *Schizophr. Res.* **2017**, *184*, 116–121. [[CrossRef](#)] [[PubMed](#)]
63. Vassos, E.; Di Forti, M.; Coleman, J.; Iyegbe, C.; Prata, D.; Euesden, J.; O’Reilly, P.; Curtis, C.; Kolliakou, A.; Patel, H.; et al. An examination of polygenic score risk prediction in individuals with first-episode psychosis. *Biol. Psychiatry* **2017**, *81*, 470–477. [[CrossRef](#)] [[PubMed](#)]

64. Natarajan, P.; Young, R.; Stitzel, N.O.; Padmanabhan, S.; Baber, U.; Mehran, R.; Sartori, S.; Fuster, V.; Reilly, D.F.; Butterworth, A.; et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* **2017**, *135*, 2091–2101. [[CrossRef](#)]
65. Raynor, L.A.; Pankow, J.S.; Rasmussen-Torvik, L.J.; Tang, W.; Prizment, A.; Couper, D.J. Pleiotropy and pathway analyses of genetic variants associated with both type 2 diabetes and prostate cancer. *Int. J. Mol. Epidemiol. Genet.* **2013**, *4*, 49–60.
66. Layton, J.; Li, X.; Shen, C.; De Groot, M.; Lange, L.; Correa, A.; Wessel, J. Type 2 diabetes genetic risk scores are associated with increased type 2 diabetes risk among African Americans by cardiometabolic status. *Clin. Med. Insights Endocrinol. Diabetes* **2018**, *11*, 1179551417748942. [[CrossRef](#)]
67. Belsky, D.W.; Sears, M.R.; Hancox, R.J.; Harrington, H.; Houts, R.; Moffitt, T.E.; Sugden, K.; Williams, B.; Poulton, R.; Caspi, A. Polygenic risk and the development and course of asthma: An analysis of data from a four-decade longitudinal study. *Lancet Respir. Med.* **2013**, *1*, 453–461. [[CrossRef](#)]
68. Pihlstrom, L.; Morset, K.R.; Grimstad, E.; Vitelli, V.; Toft, M. A cumulative genetic risk score predicts progression in Parkinson's disease. *Mov. Disord.* **2016**, *31*, 487–490. [[CrossRef](#)]
69. Nalls, M.A.; Escott-Price, V.; Williams, N.M.; Lubbe, S.; Keller, M.F.; Morris, H.R.; Singleton, A.B.; International Parkinson's Disease Genomics Consortium (IPDGC). Genetic risk and age in Parkinson's disease: Continuum not stratum. *Mov. Disord.* **2015**, *30*, 850–854. [[CrossRef](#)]
70. Khera, A.V.; Emdin, C.A.; Drake, I.; Natarajan, P.; Bick, A.G.; Cook, N.R.; Chasman, D.I.; Baber, U.; Mehran, R.; Rader, D.J.; et al. Genetic risk: Adherence to a healthy life style, and coronary disease. *N. Engl. J. Med.* **2016**, *375*, 2349–2358. [[CrossRef](#)]
71. Ripatti, S.; Tikkanen, E.; Orho-Melander, M.; Havulinna, A.S.; Silander, K.; Sharma, A.; Guiducci, C.; Perola, M.; Jula, A.; Sinisalo, J.; et al. A multi locus genetic risks core for coronary heart disease: Case-control and prospective cohort analyses. *Lancet* **2010**, *376*, 1393–1400. [[CrossRef](#)]
72. Okuda, D.T.; Mowry, E.M.; Beheshtian, A.; Waubant, E.; Baranzini, S.E.; Goodin, D.S.; Hauser, S.L.; Pelletier, D. Incidental MRI anomalies suggestive of multiple sclerosis: The radiologically isolated syndrome. *Neurology* **2009**, *72*, 800–805. [[CrossRef](#)] [[PubMed](#)]
73. Yea, C.; Tellier, R.; Chong, P.; Westmacott, G.; Marrie, R.A.; Bar-Or, A.; Banwell, B.; Canadian Pediatric Demyelinating Disease Network. Epstein-Barr virus in oral shedding of children with multiple sclerosis. *Neurology* **2013**, *81*, 1392–1399. [[CrossRef](#)] [[PubMed](#)]
74. Mechelli, R.; Manzari, C.; Policano, C.; Annese, A.; Picardi, E.; Umeton, R.; Fornasiero, A.; D'Erchia, A.M.; Buscarinu, M.C.; Agliardi, C.; et al. Epstein-Barr virus genetic variants are associated with multiple sclerosis. *Neurology* **2015**, *84*, 1362–1368. [[CrossRef](#)] [[PubMed](#)]
75. Wray, N.R.; Wijmenga, C.; Sullivan, P.F.; Yang, J.; Visscher, P.M. Common disease is more complex than implied by the core omnigenic model. *Cell* **2018**, *173*, 1573–1580. [[CrossRef](#)] [[PubMed](#)]
76. Gallagher, M.D.; Chen-Plotkin, A.S. The post-GWAS era: From association to function. *Am. J. Hum. Genet.* **2018**, *102*, 717–730. [[CrossRef](#)]

