Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic analysis of multiple sclerosis severity identifies a novel locus and implicates CNS resilience as a major determinant of outcome

Sergio Baranzini (Sergio.Baranzini@ucsf.edu)
University of California, San Francisco https://orcid.org/0000-0003-0067-194X
Stephen Sawcer
University of Cambridge https://orcid.org/0000-0001-7685-0974
International Multiple Sclerosis Genetics Consortium
MultipleMS Consortium
*

Biological Sciences - Article

Keywords:

Posted Date: June 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1723574/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Nature on June 28th, 2023. See the published version at https://doi.org/10.1038/s41586-023-06250-x.

Abstract

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS) that results in significant neurodegeneration in the majority of those affected and is a common cause of chronic neurological disability in young adults. To provide insight into the mechanisms determining progression, we conducted a genome-wide association study of the age-related MS severity score in 12,584 cases and replicated our findings in a further 9,805 cases. We identified a significant association with rs10191329 in the DYSF-ZNF638 locus (P=3.6×10-9), the risk allele shortening the median time to require a walking aid by up to 3.7 years. We also identified suggestive association with rs149097173 in the DNM3-PIGC locus (P=2.3×10-7) and significant enrichment for expression in CNS tissues. Mendelian randomization analyses indicated a protective role for higher educational attainment. In contrast to immune-driven susceptibility, these findings indicate a key role of CNS resilience and neurocognitive reserve in determining outcome in MS.

Introduction

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS)[1] affecting more than 2.8 million individuals worldwide[2] and profoundly reducing quality of life for the majority of affected individuals[3]. Clinically, the disease is characterized by recurrent episodes of largely reversible neurological dysfunction, known as relapses, together with steady and unrelenting accumulation of chronic neurological disability, referred to as progression[4]. The relative impact of these largely independent features varies between patients and during the course of illness within individuals[4]. Over the last few decades the introduction of a range of immunological treatments has transformed the ability to control relapse activity in the disease, leaving therapy capable of controlling progression as the greatest currently unmet clinical need[5].

Case-control genome-wide association studies (GWAS) have identified over 200 variants influencing susceptibility to the disease, with the strongest effects coming from the major histocompatibility complex (MHC)[6] and the implicated genes being overwhelmingly enriched for immune relevance. Although these risk variants have been found to reduce the age at onset[7–11], it is notable that they do not appear to have any impact on disease severity[11–17]. These findings, together with the concordance for outcome within families[18–21], suggest that an independent genetic architecture determines the clinical course of the disease, as has been seen in other autoimmune[22] and neurological conditions[23,24]. However, published efforts to systematically interrogate severity have to date only involved modest numbers of cases, and unanimously fall short of identifying any convincingly associated genetic variants[10,17,25–27].

Through long-standing international collaborations, we have completed the largest in-depth effort to date aimed at characterizing the genetic architecture underlying MS severity. In this study, we combined cross-sectional and longitudinal analyses of MS-specific disability outcomes, and correlated findings with tissue-specific expression patterns. We contrasted the genetic determinants of susceptibility and severity, and examined potential modifiable risk factors for MS progression. Given the significantly increased potential for the development of rational therapies attached to drug targets with genetic support[28], our work will likely help to advance patient priorities with regard to treatment and prognosis.

Results

Here we describe a genetic analysis of disease severity performed in 12,584 people with MS of European ancestry. After imputation to the Haplotype Reference Consortium and rigorous quality control (Methods), a total of 7.8 million autosomal single nucleotide variants with a minor allele frequency (MAF) > 0.01 were analyzed. The discovery cohort was collected from 21 centers across North America, Europe and Australia (**Extended Data Fig. 1** and **Supplementary Table 1**). In line with standard practice, neurological disability was measured using the Expanded Disability Status Scale (EDSS)[29], an ordinal numerical scale that increases as neurodegeneration progresses. To control for the effects of aging, individual EDSS measures were converted to the age-related MS severity (ARMSS) score by ranking disability within age-specific strata[30] (Methods). To ensure that residuals were normally distributed, we based our analyses on the rank-based inverse-normal transformation (RINT) of the ARMSS score, unless otherwise indicated. To reduce the influence of disability fluctuation related to relapses and lessen the imprecision of attempting to predict outcome in patients early in the disease, we focused recruitment on older individuals with longer duration of disease who had effectively declared their outcome. Consequently, mean age at last follow-up and disease duration were 51.7 and 18.2 years, respectively (**Supplementary Table 2** and **Extended Data Fig. 2**). Replication of variant associations was tested in existing data from an independent cohort of 9,805 cases (**Supplementary Tables 1 and 2, Extended Data Fig. 2**). The replication population was organized into four strata matched by genotyping platform and was subjected to equivalent quality control procedures (**Extended Data Fig. 1** and **Supplementary Tables 3 and 4**).

Heritability and tissue enrichment. The SNP-based heritability estimate (h^2_{SNP}) for variants with a MAF > 0.01 was 0.10 (s.e. 0.03). After partitioning the data into 10 MAF and linkage disequilibrium (LD) score bins, an approach which is generally regarded as more robust[31], the estimated h^2_{SNP} was slightly higher (0.13, s.e. 0.04; **Supplementary Table 5**). Partitioned heritability analysis by functional annotation with 96 categories[32,33] did not identify strong enrichment in any category after correction for multiple testing (**Supplementary Table 6**). To uncover disease-relevant tissues, we combined variant association statistics with gene expression profiles from 205 tissues and cell types in a heritability enrichment analysis using

stratified LD score regression (LDSC)[34]. We observed a significant enrichment (adjusted for multiple testing) exclusively in CNS tissues across multiple brain regions and the C1 segment of the cervical spinal cord (**Fig. 1** and **Supplementary Table 7**). In contrast, repeating the same analysis for MS susceptibility revealed strong enrichment in lymphoid organs, immune lymphoid and myeloid cells, as well as in tissues with recognized immunological functions and microbiota interactions (pharynx, lung, terminal ileum and endocervix; **Fig. 1 and Supplementary Table 8**)[6]. This pattern faithfully recapitulates the immune-related nature of susceptibility associations, further highlighting the striking difference from the heritability pattern observed for disease severity.

Discovery and replication of a disease severity locus for MS. To identify genetic variants associated with MS severity, we first performed a cross-sectional GWAS using ARMSS scores with the entire discovery cohort, adjusting for age, sex, date of birth, EDSS source, center, genotyping batch and the first ten principal components. Use of MS disease modifying therapy was not included as a covariate given the potential for collider bias[35]. We observed only modest inflation of the test statistics (λ_{GC} = 1.016; **Supplementary Fig. 3**) and LDSC yielded an intercept not significantly different from 1 (1.009, s.e. 0.007, 95% confidence interval [CI] 0.996 to 1.022), consistent with polygenicity driving inflation[36]. An association signal in the *DYSF–ZNF638* locus reached genome-wide significance ($P < 5\times10^{-8}$) (**Fig. 2** and **Table 1**). The lead variant rs10191329^{A} was not close to (> 3 Mb) or in LD with ($r^2 \leq 0.006$) any of the lead MS susceptibility variants[6]. Eleven additional loci showed suggestive association with ARMSS score ($P < 5\times10^{-6}$; **Fig. 2**), thereby identifying 12 independent loci that were brought forward for replication (**Supplementary Table 9**). Conditional and joint analysis did not identify secondary signals.

The *DYSF–ZNF638* locus was confirmed in the replication population and retained genome-wide significance in fixed-effects meta-analysis (**Table 1**). The direction of effect was consistent across all replication centers without evidence of heterogeneity (Q-statistic = 1.5, P = 0.99; I^2 = 0%; **Extended Data Fig. 4**). A suggestive association signal in the *DNM3-PIGC* locus replicated but did not reach genome-wide significance in the combined analysis (**Table 1**). The lead variant in this locus (rs149097173^{T}) did not overlap with any of the known MS susceptibility loci. The ten other suggestive loci were not replicated. Statistical fine-mapping supported the replicated lead variants to be causal at their respective loci (rs10191329 posterior inclusion probability (PIP) = 0.75, rs149097173 PIP = 0.95; **Extended Data Fig. 5**).

Genetic modifiers of longitudinal disability outcomes in MS. We next investigated whether the associations identified using the cross-sectional ARMSS score-based GWAS could be confirmed using additional MS specific disability outcomes from patients who had been assessed longitudinally. For this analysis, we identified 8,325 patients in our study with EDSS documented at three or more timepoints.

Cumulatively, these patients were evaluated over 54,113 study visits spanning up to 13.9 years (Methods). Adjusted Cox proportional hazards analyses showed that the lead *DYSF–ZNF638* variant (rs10191329^{A}) was associated with faster 24-week confirmed disability worsening (hazard ratio [HR] = 1.1 per unit increase in allele dosage, 95% CI 1.02-1.18, $P = 7.9\times10^{-3}$; **Fig. 3a**), a metric used as the primary outcome in progressive MS therapeutic trials[37]. In homozygous carriers, the lead variant also conferred a 3.7-year shorter median time to using a walking aid (HR = 1.22, 95% CI 1.09-1.38, $P = 9.3\times10^{-4}$; **Fig. 3b**), a clinically relevant MS disability milestone that typically tracks with the progressive phase of the disease and fixed neurological disability[38]. Moreover, a generalized linear mixed model analysis of serial EDSS across all visits confirmed that *DYSF–ZNF638* risk allele carriers displayed faster disability progression ($P = 0.002$; **Fig. 3c**).

Although less frequent (MAF 0.01), carrier status at rs149097173^{T} in the *DNM3* locus was similarly associated with faster 24-week confirmed disability worsening (HR = 1.29, 95% CI 1.02-1.65, $P = 0.037$), shorter time to EDSS 6.0 (HR = 1.56, 95% CI 1.05-2.34, $P = 0.029$), and accelerated rate of disability accrual ($P = 0.041$; **Fig. 3d-f**). The median time to require a walking aid was 2.2 years less for risk allele carriers than for non-carriers.

Gene prioritization and associations with other traits. To identify possible biological mechanisms at the discovered loci, we applied several approaches to prioritize putative causal genes (Methods, **Supplementary Table 10**). The intergenic MS severity variant rs10191329 is nearest to *DYSF* (3,692 base pairs to the transcription start site), and this gene was prioritized by the combined SNP-to-gene (cS2G)[39] strategy based on enhancer-gene linking. This variant also displayed a methylation quantitative trait locus (QTL) effect in the promoter region of *DYSF* (ENSR00001922663) in the dorsolateral prefrontal cerebral cortex[40] (**Supplementary Table 11**). In addition, rs10191329 showed correlation ($r^2 > 0.6$) with fine-mapped expression QTLs for the upstream gene *ZNF638* (**Supplementary Table 12**) and weaker correlation with splicing QTLs for the same gene in brain (r^2 0.3 to 0.4). Among other traits, rs10191329^{A} has been negatively associated with intelligence (**Supplementary Table 13**). Both these genes are highly expressed in neuronal and glial cells in the CNS with shared specificity for oligodendrocytes (**Extended Data Fig. 6 and 7**) and are important in biological processes of potential relevance. *DYSF* is implicated in membrane repair[41]; *ZNF638* mediates the silencing of unintegrated viral DNA[42] and regulates adipogenesis[43]. The suggestive variant rs149097173 is intronic to *DNM3* and *PIGC*, the latter also being nominated by cS2G. Reported trait associations for this second variant are limited to height (**Supplementary Table 14**), but *DNM3* is known to participate in the morphogenesis of the postsynaptic density and excitatory synaptic transmission[44] and demonstrates preferential expression in the CNS, specifically in neurons and oligodendrocyte lineage cells (**Extended Data Fig. 6 and 7**). *PIGC* initiates biosynthesis of the glycosylphosphatidylinositol anchor (**Extended Data Fig. 8**)[45].

Limited influence for genetic susceptibility to MS on disease outcomes. We undertook multiple approaches to determine whether previously described MS susceptibility variants[6] also drive disease severity. First, in an LD score regression analysis we observed only weak non-significant genetic correlation between MS severity and susceptibility (r_g = 0.17, p = 0.25). Next, the proportion of susceptibility variants showing concordant direction of effect in the severity GWAS was not different from that expected by chance (P_{binom} = 0.097). We then aggregated the effect of the genome-wide significant MS susceptibility variants into a polygenic risk score (PRS) and evaluated the gain in coefficient of determinant (incremental R^2) when the PRS is added as a variable to a regression of the phenotype on a set of baseline covariates (Methods). We found a weak but statistically significant positive correlation with ARMSS score (incremental R^2 = 0.001, P = 7.1×10^{-5}) across MHC and non-MHC regions (**Supplementary Fig. 5**). However, higher genetic susceptibility for MS leads to earlier age at onset, which in turn is associated with increased MS severity (**Supplementary Fig. 4**). Therefore,we repeated this analysis adjusting for age at onset and observed that the effect of the susceptibility PRS on ARMSS score was substantially attenuated (incremental R^2 = 3.9×10^{-4}, P = 0.014; **Supplementary Fig. 5**). In addition, we interrogated the association of susceptibility variants with longitudinal disability outcomes. Individually, none of the variants influenced these outcomes after adjusting for the number tested (**Extended Data Fig. 9a-c** and **Supplementary Table 15**). Furthermore, none showed consistent nominal association (P < 0.05) across outcomes (**Extended Data Fig. 9d** and **Supplementary Table 15**). Comparing individuals in the highest susceptibility PRS quartile to those in the lowest, we detected no significant differences in longitudinal outcomes in the adjusted survival and linear mixed model analyses (**Extended Data Fig. 10**). In short, we found no evidence that susceptibility variants exert a meaningful effect on the outcome of the disease.

Mendelian randomization (MR) highlights an association between educational attainment and MS severity. We investigated putative causal and modifiable risk factors for MS severity using two-sample MR. We focused our analyses on traits with prior evidence for association with MS outcomes and suitable genetic instruments, namely 25-hydroxyvitamin D (25OHD) levels[46,47], body mass index (BMI)[48,49] and educational attainment[50–52] (**Supplementary Table 16**). The latter was further motivated by the implication of brain reserve in MS disability progression[53] and our finding of CNS heritability enrichment. MR analyses did not indicate a causal role for either 25OHD levels or BMI (**Fig. 4**). In contrast, the main inverse-variance weighted MR estimate provided support for an association between higher years of education and milder MS severity, at two p-value thresholds for genetic instrument selection (β = -0.16, P_{IVW} = 0.014 based on 263 education-associated variants; β = -0.16, P_{IVW} = 9.7×10^{-4} based on 610 education-associated variants). This result was substantiated by pleiotropy-robust MR sensitivity analyses (**Fig. 4**; Methods). Additionally, the MR-Egger intercept revealed little evidence of directional pleiotropy and MR-PRESSO found no outliers (**Supplementary Table 17**). We observed no significant

heterogeneity based on Cochran's Q-statistic and MR-PRESSO global test. Reverse analysis did not support an effect of genetic liability to MS severity on 25OHD levels, BMI or years of education (**Supplementary Table 17**).

Discussion

In summary, this GWAS, which included over 22,000 people with MS, suggests that outcome in the disease is at least in part influenced by the resilience of the CNS to injury. We have identified the first genome-wide significant modifier of long-term outcome in MS, and have thereby identified high value targets for drug discovery[28]. The lead variant, and an additional suggestive association, replicated and showed concordant significant effects in a range of MS-specific longitudinal disability outcomes across tens of thousands of patient visits. These severity variants were not associated with susceptibility. Furthermore, we show that genetic susceptibility burden has little influence on cross-sectional and longitudinal outcomes outside of its effect on age at onset. Finally, MR analyses provide evidence for educational attainment as a potential modifiable risk factor for MS progression. Our observations concord with the proposed enhanced penetrance of monogenic causes of neurological disease reported to result from comorbidity with MS[54–56].

Our findings demonstrate that approximately 13% of the variance in long-term MS severity (by heritability analysis) can be attributed to common and low frequency single nucleotide variation, explaining some of the considerable variability in MS outcome. Notably, this GWAS revealed enrichment for this heritability in components of the brain and spinal cord, in marked contrast to the pronounced immune signal seen for MS susceptibility. Although divergent genetic determinants of susceptibility and progression have been noted in other conditions[22–24], the observation of distinct tissue enrichment is to our knowledge unique to MS. This result has potentially significant clinical implications. A persistent challenge in understanding MS progression has been determining the relative contributions of inflammatory activity (including CNS-compartmentalized immune responses) and neurodegeneration[5]. Here, we show that genes preferentially expressed within the CNS in controls likely contribute to MS severity. This strongly implicates neuronal and glial mechanisms as key determinants of MS progression, and provides genetic evidence to support the search for new therapeutic targets focused on neuroprotection and brain repair. It may also partly explain why immunosuppressive therapies have thus far had little or no effect on disability accumulation in progressive MS trials[5].

The two main identified MS severity variants had a clinically meaningful impact on time to needing a walking aid, with the median interval from onset shortened by 3.7 years for homozygous carriers of the common *DYSF-ZNF638* variant (rs10191329^{A}) and 2.2 years for carriers of the *DNM3-PIGC* variant (rs149097173^{T}). Although not comparable in terms of likely mechanism, the magnitude of this effect is

comparable to the impact of treatment with a first line disease modifying agent such as beta-interferon[57]. This key MS disability milestone is associated with unemployment[58], reduced quality of life[59] and irreversible neurological disability[38]. In principle, relapses and progression could both influence the MS severity outcomes used in this study. However, although relapses typically lead to transient increase in disability, it is recognized that their contribution to long-term disability and confirmed disability progression is limited, especially after the first few years post diagnosis[60]. In addition, relapse frequency spontaneously diminishes over time[61]. In this context, a recent study of relapse activity in MS reported distinct genetic association signals and alternate pathways[62]. Given the average age and disease duration of our population (respectively 51.7 and 18.2 years), as well as the associations with time to EDSS 6.0, our findings are likely to reflect independent mechanisms underlying MS progression. Nevertheless, additional longitudinal analyses will be required to further refine the effects of these severity variants on MS phenotypes, including molecular, imaging and pathology.

Our gene prioritization analyses implicated four biologically plausible genes at the identified loci, including *ZNF638* upstream of the intergenic variant rs10191329. *ZNF638* encodes the DNA-binding zinc finger protein 638 (also known as NP220), which mediates the transcriptional repression of unintegrated retroviral DNA through recruitment of the human silencing hub (HUSH) complex and the histone methyltransferase SETDB1[42]. The same chromatin repressors are involved in epigenetic silencing of endogenous retroviruses[63,64]. Several exogenous and endogenous viruses have been considered in MS pathogenesis, with the most compelling evidence implicating respectively Epstein-Barr virus (EBV)[65–67] and human endogenous retrovirus type-W (HERV-W)[68,69]. The possibility of *ZNF638* silencing EBV or HERV-W could have therapeutic implications in MS, as demonstrated by the ongoing development of EBV T-cell therapy (NCT03283826) and HERV-W envelope protein-binding monoclonal antibody[70]. Furthermore, convergent evidence supports a role, still to be determined, for *ZNF638* in the CNS, including in the context of MS. The gene is highly expressed in the brain, particularly in oligodendrocytes and their precursor cells, and has been implicated repeatedly in large-scale genetic studies of intelligence and general cognitive ability[71–73]. In single-nucleus RNA sequencing from brain white matter areas in MS patients and controls, *ZNF638* was preferentially expressed in an oligodendrocyte cluster with a predicted actively myelinating phenotype[74]. Moreover, cell expression of *ZNF638* was proportionally enriched in control brain tissue and chronic inactive MS lesions compared to other MS lesions[74].

DYSF, the nearest gene to rs10191329, encodes dysferlin, a type II transmembrane protein. Although widely expressed, its functions are mainly characterized in skeletal muscle where it participates in calcium-mediated membrane repair and regeneration[41]. Recessive pathogenic variants lead to muscular dystrophies (OMIM 254130, 253601, 606768). *DYSF* is also specifically expressed in oligodendrocytes and excitatory neurons, and the protein has been found to accumulate in Aβ-containing extracellular

neuritic plaques, in proportion to Alzheimer disease severity[75]. Although its role in the CNS has yet to be determined, participation in membrane maintenance of neurons or glia could influence neuronal and axonal survival (such as in response to axonal swelling[76]) or subsequent remyelination.

The suggestive variant rs149097173 is located in intron 20 of *DNM3*, which encodes dynamin-3 and mediates synaptic vesicle endocytosis. As with other prioritized genes, expression is preferentially in oligodendrocytes lineage cells and neurons. An independent variant in *DNM3* was reported to associate with age of onset in *LRRK2* parkinsonism[77], although this did not replicate in a follow-up study[78]. Interestingly, the paralog dynamin-2 participates in membrane repair by wound-induced endocytosis in skeletal muscle[79], which may point to a convergence of mechanisms with *DYSF*. Variant rs149097173 is also intronic to *PIGC*, mutations in which can lead to intellectual disability and epilepsy[45].

Our MR results do not support a causal role for serum 25OHD levels or BMI on MS severity, which may potentially indicate confounding or reverse causality in the reported observational associations. This agrees with the inconclusive results of randomized trials of vitamin D supplementation in MS[80], and with a recent prospective study that found no association between BMI and clinical disability[81]. We note that these MR analyses assume linearity and may not be applicable to individuals at the extremes of trait distributions. Additionally, as obesity and vitamin D deficiency are risk factors for the development of MS[1], collider bias may occur, although its effect is likely to be small[82] On the other hand, a few observational studies have documented an inverse association between educational attainment and subsequent MS disability[50–52] as well as retinal neurodegeneration[83]. In accordance with these data, we have found genetic support for educational attainment having a causal effect on reducing long-term MS severity, with little evidence of horizontal pleiotropy. The effect size was substantial, with 4 years of additional education (equivalent to an undergraduate degree) predicted to reduce the rank of disability by a quintile. This finding would be consistent with education promoting neurocognitive reserve[84], and thereby increasing resilience to neuronal degeneration resulting from MS injury and aging. Similar protective effects of education have also been observed in Alzheimer's disease and frontotemporal dementia[84,85], indicating some commonality with other neurodegenerative conditions. In addition, our results support the study of modifiable lifestyle factors that have been proposed to influence neurocognitive reserve and maintenance[84], such as social engagement, diet and physical activity, as potential approaches to slow MS progression.

In conclusion, this study presents conclusive evidence for the role of genetic variation in influencing MS progression. MS has undergone a therapeutic revolution in the past few decades, with the emergence of ever more effective immune therapies that reduce and even halt relapses. Despite this, treatment of

progression remains an unmet need. We have identified genetic drivers of disability in MS, providing new directions for functional characterization and drug development targeted on the neurodegenerative component of the disease. Successful unraveling of the genetic basis for disease susceptibility has implicated dysregulation across immune cells as a driver of MS onset. Our findings establish CNS resilience and reserve as key determinants of MS progression, and may have broader implications for other neurodegenerative diseases.

Materials And Methods

Study participants and GWAS outcome. The discovery population consisted of patients with MS recruited through 21 centers from North America, Europe and Australia. A total of 15,072 patients were genotyped on a common platform (Illumina Global Screening Array) in five cohorts. Samples from patients with longer disease duration, older age, and availability of longitudinal outcome measures were preferentially submitted for genotyping. A primary progressive onset was reported in 8.6% of patients with a documented disease phenotype. **Supplementary Tables 1 and 2** respectively describe the case counts per center and additional demographic characteristics. The replication population consisted of a combination of already genotyped MS patients and controls with available clinical information assembled through 9 European centers and genotyped on various Illumina arrays, resulting in 17 cohorts (**Supplementary Table 3**). Patients that passed sample quality control and had at least one disability measure were included in the analysis (**Extended Data Fig. 1**). All participants gave written informed consent in accordance with approval from the relevant local ethical committees or institutional review boards. Patients with MS were ascertained and diagnosed by a neurologist locally according to established criteria. Neurological disability was measured using the EDSS[29], an ordinal scale which incorporates a range of neurological functions relevant to MS. EDSS was scored by neurologist assessment in all but 1,040 cases (4.6%), where it was approximated via questionnaire. For each individual, the last recorded EDSS was converted to an ARMSS score by ranking disability against participants with the same age (±2 years) from the same cohort and from an additional 26,058 patients with MS[30].

Quality control and imputation. For each cohort, we performed individual- and variant-level quality control, after which cohorts were merged into strata based on genotyping platform and submitted to additional stratum-level quality control (**Supplementary Note**). Sample overlap across strata and between the discovery and replication populations was assessed, and duplicates removed. Imputation to the Haplotype Reference Consortium panel (release 1.1)[86] was performed using Minimac4 (v1.0.2)[87] and in-house scripts. The resulting variant counts and imputation quality metrics are described in **Supplementary Table 4** and **Supplementary Fig. 2**, respectively.

GWAS and replication. To identify genetic variants associated with MS severity, we performed a linear regression model implemented in fastGWA[88] using genotype dosages. We applied a rank-based inverse-normal transformation (RINT) to the ARMSS scores and fit as covariates in the model age, sex, date of birth, EDSS source (neurologist assessment vs. questionnaire), center, genotyping batch and the first ten principal components. Disease modifying therapy was not included as it is not a confounder (i.e. does not influence genotype) and may instead introduce collider bias[35]. To assess any residual confounding due to population stratification or cryptic relatedness, we calculated the genomic inflation factor and LDSC intercept using HapMap3 variants and LD scores from 1000 Genomes phase 3[36]. Conditional and joint (COJO) analysis[89] was performed to identify potential secondary association signals. Lead variants with association $P \leq 5\times10^{-8}$ were considered genome-wide significant and were tested in the replication population, together with those with suggestive association $P \leq 5\times10^{-6}$.

As above, linear regression of ARMSS scores was performed in the replication population using the same covariates. Individual-level imputed genotypes were merged across strata prior to joint analysis. Principal components were calculated on a set of hard-called high-quality (imputation $R^2 \geq 0.9$, genotype missingness < 0.01, MAF > 0.05) and LD-pruned genotypes. To examine for heterogeneity, we recalculated the association between lead variants and MS severity in the replication stratified by center (n=9) and computed Q-statistics and I^2 tests. Finally, association statistics from the discovery and replication were combined using fixed-effects meta-analysis.

Heritability estimation. To estimate SNP-based heritability, we constructed a genomic relationship matrix (GRM) from all variants and used it to remove individuals (n = 848) with a coefficient of relationship > 0.025. The resulting GRM was used to estimate SNP-heritability with restricted maximum likelihood (single-component GREML)[90]. As SNP-heritability can be sensitive to LD and allele frequency assumptions[31], we also fitted a model with ten GRMs (GREML-LDMS) constructed from variants assigned to five MAF bins (0.01-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4 and 0.4-0.5) each divided into two by the median LD score in each bin. To calculate LD scores, variants were first hard-called (PLINK2 –hard-call-threshold 0.1) then filtered for missingness < 0.05, MAF > 0.01 and HWE $P > 10^{-6}$. Heritability analyses were adjusted for the same set of covariates as the GWAS.

Heritability enrichment analyses. We used stratified LDSC (version 1.0.1) to calculate SNP-based heritability enrichment for 96 categories (baseline-LD model version 2.2)[32,33], including functional, MAF-related and LD-related annotations. Next, we assessed the SNP-based heritability associated with different tissues by applying stratified LDSC to our GWAS summary statistics using a gene expression dataset consisting of 205 tissues and cell types (as provided in the LDSC software)[34]. Tissues and cell

types were grouped into nine categories for visualization (**Supplementary Tables 7 and 8**). The same analysis was repeated with the summary statistics from the discovery phase of our previous GWAS meta-analysis of MS susceptibility[6] to compare the enrichment patterns. We applied FDR correction for multiple testing within each enrichment analysis, and FDR-corrected $P < 0.05$ were considered statistically significant.

Longitudinal analysis of MS disability outcomes. We identified a subset of 8,325 MS patients from our study population with a minimum of 3 visits separated by at least 6 months (5,565 from the discovery cohort and 2,760 from the replication cohort). These patients contributed a total of 56,966 visits, of which 54,113 (95%) occurred within 13.9 years of follow-up from the first study visit (mean 5.2 years). Two key MS-specific disability outcomes were examined in survival analyses. First, we estimated the influence of MS severity variants on time to a clinically meaningful increase in neurological disability. Similar to MS clinical trials[37], worsening was defined as an increase in EDSS by 1.0 if the baseline score was < 5.5 and by 0.5 if the baseline was ≥ 5.5. To increase specificity, the endpoint also required this EDSS increase to be maintained on a subsequent visit and for at least 24 weeks. Second, we examined the influence of genotype on time (from disease onset) to reaching EDSS 6.0 (defined as requiring unilateral assistance to walk more than 100 meters). Following left-censoring, 7,832 patients and 51,189 study visits remained, extending to 28.3 years from disease onset. Cox proportional hazards analyses were carried out using the coxph function in the 'survival' package (version 3.2-11) in R, with Efron approximation for tie handling. Sex, age at onset, date of birth, center, genotyping platform and the first ten principal components were included as covariates. Adjustment for baseline EDSS was included in the 24-week confirmed disability worsening analysis to account for the non-linear nature of this scale; this was not applicable for the time to EDSS 6.0 analysis. The proportional hazards assumption was examined by inspection of scaled Schoenfeld residuals. Hazard ratios were calculated using dosages for rs10191329 and carrier status for rs149097173 given its low frequency.

To assess the influence of MS severity variants on the rate of disability progression, we constructed a generalized linear mixed model with serial EDSS scores as the dependent variable. The primary predictor was the interaction term between genotype (dosage or carrier status) and time in the study (years), with individuals and centers as random terms. Subject-level fixed covariates were sex, age at onset and study entry, date of birth and the first ten principal components. This analysis was performed using penalized quasi-likelihood estimation as implemented in the glmmPQL function from the 'MASS' package (version 7.3-54) in R to address the non-normal distribution of EDSS.

Fine-mapping. For each lead variant, effect estimates on MS severity in a 250 kb region centered on the variant were extracted. A variant correlation matrix was computed with LDstore2 (version 2.0)[91] from the

same genotype dosage used to generate the GWAS summary statistics. Fine-mapping with shotgun stochastic search was performed using FINEMAP (version 1.4)[92] with equal prior probabilities.

Gene prioritization and associations with other traits. To prioritize putative causal genes, we applied a combination of functional and non-functional strategies: (1) the closest gene(s), defined as genes with overlapping bodies or closest transcription start site; (2) genes that overlap with a genomic range of 200 kb centered around the variant; (3) genes with missense or loss of function coding variants in LD (r^2 > 0.6) with the lead variant; (4) genes with fine-mapped (PIP > 0.1) *cis*-eQTL or splicing QTL in LD (r^2 > 0.6) with the lead variant; (5) genes prioritized by Open Targets Genetics using a V2G[93] threshold of 0.5; (6) genes prioritized by the combined SNP-to-gene (cS2G) strategy[39]. We retrieved fine-mapped QTLs from GTEx[94] (version 8) and the eQTL catalogue[95]. The V2G aggregates weighted evidence from variant functional prediction, colocalization with molecular QTLs, chromatic interaction and gene distance. The cS2G strategy consists of seven components, with gene assignments most often driven by a single feature. Moreover, we evaluated the influence of MS severity variants on brain dorsolateral prefrontal cortex methylation based on 543 individuals from ROSMAP (Bonferroni-corrected $P < 5\times10^{-9}$)[40].

To investigate the effects of the MS severity variants on previously reported phenotypes, we retrieved phenome-wide associations in the Open Target Genetics portal[96] obtained from the GWAS Catalog, UK Biobank and FinnGen.

Gene expression profiles. Gene expression values in human tissues for the prioritized genes at the two MS severity loci were obtained from GTEx[94] (version 8). Cell type expression profiles for the same genes were evaluated using single cell RNA sequencing data in 76 cell types from the Human Protein Atlas[97]. We examined genes for cell type specificity, defined as expression that is at least fourfold higher in a cell type compared to the mean of all others (cell type enhanced)[97]. Since *PIGC* expression in brain neuronal and glial cell types was missing, we obtained it from a study of 4 progressive MS patients and 5 non neurological controls with single nuclear RNA expression in white matter tissues[74].

MS susceptibility variants. To compare the genetic architecture of MS susceptibility and severity, we calculated the genome-wide genetic correlation excluding the MHC region using bivariate LDSC (version 1.0.1)[36]. A free intercept was modeled to allow for sample overlap. We then focused our analyses on the 232 autosomal MS susceptibility associations we previously reported[6]. For non-MHC variants, we included the association statistics from the joint analysis and labeled them using the discovery variant

('SNP discovery'). We excluded variants that were palindromic (n=1), missing from the current study (n=1) or with a joint $P > 5\times10^{-8}$ (n=2). For MHC associations, we included those reported as non-palindromic single nucleotide variants (as opposed to HLA alleles) and added rs3135388 to tag *HLA-DRB1*1501*[98]. In total, 209 variants (197 non-MHC and 12 MHC) were examined (**Supplementary Table 15**). A two-sided exact binomial test was used to assess concordance of direction of effect on MS susceptibility and severity. The same variants were tested for association with longitudinal outcomes using a Bonferroni-corrected significance threshold ($P < 0.05/209$ or 2.4×10^{-4}) and evaluated for concordance of nominal association ($P < 0.05$) across four disability outcomes (ARMSS score, 24-week confirmed disability worsening, time to EDSS 6.0 and rate of EDSS change).

To determine the aggregate effect of MS susceptibility on disability outcomes, we constructed a PRS using 178 variants retained following LD clumping ($r^2 < 0.01$) of the 209 susceptibility associations. Variants were weighted by the natural log of their joint odds ratio. We then regressed the ARMSS scores on the PRS adjusting for the same covariates as in the GWAS. We also regressed the phenotype on the covariates alone and measured the difference in R^2 with and without the PRS, reported as the incremental R^2. We performed similar analyses using age at onset, as well as ARMSS scores adjusted for age at onset. Next, we compared individuals in the highest and lower quartile of PRS based on the same survival and linear mixed model analyses as previously described for the MS severity variants.

Mendelian randomization. We applied MR analysis to investigate the effects of 3 exposures with robust genetic associations and strong prior evidence of association with MS severity. In the case of body mass index and 25-hydroxyvitamin D, previous MR studies additionally provided support for a causal role in the development of MS[99]. A description of the GWAS used to proxy the exposures is provided in **Supplementary Table 16**. For each of these, variants were selected at two different association thresholds ($P < 5\times10^{-8}$ and $P < 5\times10^{-5}$), as in previous studies[24], and LD clumped ($r^2 < 0.001$) to ensure independence. Palindromic variants were excluded. For variants absent from our MS severity GWAS, we selected a strong LD proxy ($r^2 > 0.8$) when possible. The variants included were examined for instrument strength[100] (mean *F*-statistic > 10; **Supplementary Table 16**).

The main analysis was performed using the inverse-variance weighted MR approach with a random-effects model. We also tested for heterogeneity across the genetic variants as a potential indicator of horizontal pleiotropy, using the Cochran's Q-statistic and MR-pleiotropy residual sum and outlier (PRESSO) global test[101,102]. To further examine the assumption of no horizontal pleiotropy, we applied four additional MR methods: robust adjusted profile score, weighted median, MR-PRESSO and MR-Egger regression (reviewed in ref[102]). Consistent results across these methods reduce the likelihood of bias. For

the MR-Egger regression, we focused on the intercept as a test for unbalanced pleiotropy given that the association estimate is considerably underpowered[103], although beta-coefficients are reported in **Supplementary Table 17**.

To determine the direction of effect, we also conducted a reverse analysis examining the effect of genetic liability to MS severity on each of the traits considered. As a single variant was available at the instrument selection threshold of $P < 5\times10^{-8}$, we applied a Wald ratio test in place of the inverse-variance weighted MR.

Finally, to provide an interpretable estimate of the effect size of education on MS severity, we conducted a GWAS of untransformed ARMSS scores and repeated the educational attainment MR analysis with estimates on the absolute scale.

MR analysis was conducted in R using in-house scripts, as well as the 'MendelianRandomization' and 'TwoSampleMR' packages.

Declarations

Data availability. The GWAS summary statistics generated in this study can be accessed through the International Multiple Sclerosis Genetics Consortium website (https://imsgc.net/). Individual-level genetic and phenotype data necessary to replicate the main analysis will be deposited in the European Genome-phenome Archive (EGA) for European centers, and in dbGAP (accession number phs002929.v1.p1) for all other centers. The gene expression profiles of human tissues used in this study can be downloaded from the GTEx Portal v8 (https://gtexportal.org/). The single-cell type expression profiles in human tissues can be downloaded from the Human Protein Atlas (https://www.proteinatlas.org/). We used publicly available data from the eQTL Catalogue release 4 (https://www.ebi.ac.uk/eqtl/), the LDSC GitHub repository (https://github.com/bulik/ldsc/) and the Gonçalo Castelo-Branco Group (https://ki.se/en/mbb/oligointernode/). Detailed information on the GWAS summary statistics used in the Mendelian randomization analysis is provided in **Supplementary Table 16**.

Code availability. The following software packages were used for data analyses: R version 4.0.5 (https://www.r-project.org/) with additional packages ms.sev version 1.0.4, aberrant version 1.0, survminer version 0.4.9, survival version 3.2-11, metafor version 3.0-2, MASS version 7.3-54, lme4 version 1.1-27.1, lmerTest version 3.1-3, bootpredictlme4 version 0.1, gwasglue version 0.0.0.9000, MendelianRandomization version 0.5.1, TwoSampleMR version 0.5.6, mr.raps version 0.4, MRPRESSO

version 1.0, data.table version 1.14.0, tidyverse version 1.3.1, ggplot2 version 3.3.5, ggpubr version 0.4.0, ggvenn version 0.1.9, scattermore version 0.7; GenomeStudio version 2.0 (https://support.illumina.com/downloads/genomestudio-2-0.html), GCTA version 1.93.2beta (https://yanglab.westlake.edu.cn/software/gcta/), EIGENSOFT version 6.1.4 (https://github.com/DreichLab/EIG), PLINK version 1.90beta (https://www.cog-genomics.org/plink/1.9/) and version 2.00 (https://www.cog-genomics.org/plink/2.0/), bcftools version 1.12 (https://samtools.github.io/bcftools/), qctool version 2.0.6 (https://www.well.ox.ac.uk/~gav/qctool_v2/), FINEMAP version 1.4 and LDstore version 2.0 (http://www.christianbenner.com/), EAGLE version 2.4.1 (https://alkesgroup.broadinstitute.org/Eagle/), Minimac4 version 1.0.2 (https://genome.sph.umich.edu/wiki/Minimac4), GWAMA version 2.2.2 (https://manpages.ubuntu.com/manpages/xenial/man1/GWAMA.1.html), LDSC version 1.0.1 (https://github.com/bulik/ldsc), KING version 2.2.5 (https://www.kingrelatedness.com/).

Acknowledgments. We thank all study participants for their support and for making this work possible. This work was supported by funding from the NIH/NINDS (R01NS099240) to S.E.B. and S.J.S., and the European Union's Horizon 2020 Research and Innovation Funding Programme (EU RIA 733161) to MultipleMS. We acknowledge support from the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. A.H. is supported by the NMSS-ABF Clinician Scientist Development Award (FAN-1808-32256) funded by the National Multiple Sclerosis Society (NMSS) and the Multiple Sclerosis Society of Canada (MSSC). P.S. is supported by the Magretha af Ugglas foundation and Horizon 2020 EU grant (MultipleMS, 733161). S.E.B holds the Professorship in Neurology I and the Heidrich Family and Friends Endowed Chair in Neurology. The UCSF DNA biorepository is supported by the NMSS (Si-2001-35701). J.L.M. acknowledges funding support from the NIH/NINDS (R01NS096212). L.A. has received academic grant support from the Swedish Research Council, the Swedish Research Council for Health, Working Life and Welfare and the Swedish Brain foundation. S.R.D. has received institutional research grant funding from the NMSS and the NIH/NINDS. T.O. has received academic grant support from the Swedish Research Council, the Swedish Brain foundation, Knut and Alice Wallenberg foundation and Margaretha af Ugglas foundation. M.J.F.-P. has received grant support from the Multiple Sclerosis Society of Western Australia (MSWA). M.V. is a PhD fellow (11ZZZ21N) and B.D. is a Clinical Investigator of the Research Foundation-Flanders (FWO-Vlaanderen). B.D. and A.G. have received academic grant support from the Research Fund KU Leuven (C24/16/045) and the Research Foundation Flanders (FWO G.07334.15). S.L. holds research support from the Spanish Government (PI21/010189, PI18/01030, PI15/00587), funded by the Instituto de Salud Carlos III-Subdirección General de Evaluación and co-funded by the European Union, and the Red Española de Esclerosis Múltiple (REEM: RD16/0015/0002, RD16/0015/0003). S.B. and F.Z. have received funding from the German Research Foundation (CRC-TR-128). F.Z. also acknowledges support from the Progressive MS Alliance (BRAVEinMS PA-1604-08492) and the Federal Ministry of Education and Research (VIP+ HaltMS-03VP07030). A.M. is supported by Margaretha af Ugglas foundation. B.H. is associated with DIFUTURE (Data Integration for Future Medicine) [BMBF 01ZZ1804[A-I]]. He received funding for the study by the Deutsche

Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology [EXC 2145 SyNergy – ID 390857198]. The study was supported by the Italian Foundation of Multiple Sclerosis (FISM, 2011/R/14 2015/R/10, 2019/R-Multi/033, grants), Ricerca finalizzata, Italian Ministry of Health (RF-2016-02361294 grant), the AGING Project for Department of Excellence at the Department of Translational Medicine (DIMET), Università del Piemonte Orientale, Novara, Italy. N.B. is partly supported by the MultipleMS project (Horizon 2020 European, Grant N. 733161). N.A.P. was supported in part by the NMSS (grants JF-1808-32223 and RG-1707-28657). In.K. was partly supported by the MultipleMS project (Horizon 2020 European, Grant N. 733161), the Swedish Research Council (Grant N. 2020-01638) and the Swedish Brain foundation.

This manuscript is dedicated to the memory of Rogier Q. Hintzen, a member of the International Multiple Sclerosis Genetics Consortium, in recognition of his contributions to human genetics.

Competing interests. T.O. has received compensation for advisory boards/lectures from Biogen, Novartis, Merck and Sanofi, as well as unrestricted MS research grants from the same companies, none of which are related to the current article. A.B. and his institution have received compensation for consultancy, lectures and participation in clinical trials from Alexion, Biogen, Celgene, Merck, Novartis, Sandoz/Hexal, Sanofi and Roche, all outside the current work. S.R.D. has received compensation for serving on advisory boards from Novartis, and institutional research grant funding from EMD Serono and Novartis, all outside the current work. M.F. is Editor-in-Chief of the Journal of Neurology, Associate Editor of Human Brain Mapping, Associate Editor of Radiology, and Associate Editor of Neurological Sciences; received compensation for consulting services and/or speaking activities from Alexion, Almirall, Bayer, Biogen, Celgene, Eli Lilly, Genzyme, Merck-Serono, Neopharmed Gentili, Novartis, Roche, Sanofi, Takeda, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Roche, Teva Pharmaceutical Industries, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). J.L.-S. received travel compensation from Biogen, Merck, Novartis; has been involved in clinical trials with Biogen, Novartis, Roche; her institution has received honoraria for talks and advisory board service from Biogen, Merck, Novartis, Roche, all outside the current work. M.J.F.-P. has received travel compensation from Merck outside the current work. A.G.K. has received speaker honoraria and Scientific Advisory Board fees from Bayer, BioCSL, Biogen-Idec, Lgpharma, Merck, Novartis, Roche, Sanofi-Aventis, Sanofi-Genzyme, Teva, NeuroScientific Biopharmaceuticals, Innate Immunotherapeutics, and Mitsubishi Tanabe Pharma, all outside of the current work. F.Z. has recently received research grants and/or consultation funds from Biogen, Ministry of Education and Research (BMBF), Bristol-Meyers-Squibb, Celgene, German Research Foundation (DFG), Janssen, Max-Planck-Society (MPG), Merck Serono, Novartis, Progressive MS Alliance (PMSA), Roche, Sanofi Genzyme, and Sandoz, all outside of the current work. B.D. has received consulting fees and/or

funding from Biogen Idec, BMS, Sanofi-Aventis and Teva. B.D. and A.G. have received consulting/travel fees and/or research funding from Novartis, Roche and Merck, all outside the current work. SL received compensation for consulting services and speaker honoraria from Biogen Idec, Novartis, TEVA, Genzyme, Sanofi and Merck, all outside the current work. S.B. has received honoraria from Biogen Idec, Bristol Meyer Squibbs, Merck Healthcare, Novartis, Roche, Sanofi Genzyme and TEVA; his research is funded by the German Research Foundation (DFG), Hertie Foundation and the Hermann and Lilly-Schilling Foundation. F.E. received compensation for consulting services and speaker honoraria from Novartis, Sanofi Genzyme, Almirall, Teva, and Merck-Serono. Jo.S. received consultancy and/or lecture fee from Biogen, Merck, Novartis and Sanofi Genzyme, his institution received research funding by Biogen, GSK, Idorsia, and Merck, all outside the current work. B.H. has served on scientific advisory boards for Novartis; he has served as DMSC member for AllergyCare, Polpharma and TG therapeutics; he or his institution have received speaker honoraria from Desitin; his institution received research grants from Regeneron for multiple sclerosis research. He holds part of two patents; one for the detection of antibodies against KIR4.1 in a subpopulation of patients with multiple sclerosis and one for genetic determinants of neutralizing antibodies to interferon. Ja.S. received speaker honoraria and a research grant for rare diseases from Sanofi Genzyme, and is a founder and minority shareholder of the University of Helsinki spin-off company VEIL.AI. J.L.M. has participated in advisory board meetings for Sanofi-Genzyme and received research funding from Genentech, Biogen Idec, and the Bristol-Myers Squibb Foundation. N.A.P. is currently an employee of Novartis Institutes for BioMedical Research (NIBR).

The remaining authors declare no competing interests related to this work.

The International Multiple Sclerosis Genetics Consortium and MultipleMS Consortium

Adil Harroud[1], Pernilla Stridh[2], Jacob L. McCauley[3,4], Janna Saarela[5,6], Ingileif Jónsdóttir[7,8], Lars Alfredsson[2], Katayoun Alikhani[9], Till F. M. Andlauer[10], Maria Ban[11], Lisa F. Barcellos[12], Nadia Barizzone[13], Ashley H. Beecham[3], Tone Berge[14,15], Achim Berthele[10], Stefan Bittner[16], Yolanda Blanco[17], Steffan D. Bos[18,19], Farren B. S. Briggs[20], Stacy J. Caillier[1], Domenico Caputo[21], Paola Cavalla[22], Elisabeth G. Celius[18,19], Tanuja Chitnis[23,24], Ferdinando Clarelli[25], Manuel Comabella[26], Giancarlo Comi[27,28], Chris Cotsapas[29,30], Bruce C. A. Cree[1], Sandra D'Alfonso[13], Efthimios Dardiotis[31], Philip L. De Jager[32], Silvia R. Delgado[33], Bénédicte Dubois[34,35], Sinah Engel[16], Federica Esposito[36], Marzena J. Fabis-Pedrini[37,38], Massimo Filippi[39,28], Christiane Gasperi[10], Lissette Gomez[3], Refujia Gomez[1], Georgios Hadjigeorgiou[40], Friederike Held[10], Roland G. Henry[1], Jan Hillert[2], Noriko Isobe[41], Maja Jagodic[2], Allan G. Kermode[42,38], Michael Khalil[43], Trevor J. Kilpatrick[44,45,46], Ioanna Konidari[3], Karim L. Kreft[47], Jeannette Lechner-Scott[48,49], Maurizio Leone[50], Sara Llufriu[17], Felix Luessi[16], Lohith Madireddy[1], Sunny Malhotra[26], Ali

Manouchehrinia[2], Clara P. Manrique[3], Filippo Martinelli-Boneschi[51,52], Elisabetta Mascia[25], Luanne M. Metz[9], Luciana Midaglia[26], Xavier Montalban[26], Jorge R. Oksenberg[1], Tomas Olsson[2], Annette Oturai[53], Kimmo Pääkkönen[6], Grant P. Parnell[54,55], Nikolaos A. Patsopoulos[56,57,58], Margaret A. Pericak-Vance[3,4], Fredrik Piehl[2], Justin P. Rubio[45,46], Albert Saiz[17], Adam Santaniello[1], Silvia Santoro[25], Catherine Schaefer[59], Finn Sellebjerg[53,60], Hengameh Shams[1], Klementy Shchetynsky[2,61], Claudia Silva[9], Vasileios Siokas[31], Joost Smolders[62,63], Helle B. Søndergaard[53], Melissa Sorosina[25], Bruce Taylor[64], Marijne Vandebergh[35], Domizia Vecchio[65], Pablo Villoslada[17,66], Margarete M. Voortman[43], Howard L. Weiner[23,24], V. Wee Yong[9], Kári Stefánsson[7,8], David A. Hafler[56,67], Graeme J. Stewart[68,69], Alastair Compston[11], Frauke Zipp[16], Hanne F. Harbo[18,19], Bernhard Hemmer[10,70], An Goris[35], Stephen L. Hauser[1], Ingrid Kockum[2], Stephen J. Sawcer[11,71], Sergio E. Baranzini[1,71].

[1]UCSF Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, San Francisco, CA, USA. [2]Department of Clinical Neuroscience, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. [3]John P Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL, USA. [4]The Dr. John T. Macdonald Foundation Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, FL, USA. [5]Centre for Molecular Medicine Norway, University of Oslo, Oslo, Norway. [6]Institute for Molecular Medicine Finland, Helsinki Institute for Life Sciences, University of Helsinki, Helsinki, Finland. [7]deCODE Genetics/Amgen, Inc., Reykjavik, Iceland. [8]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland.. [9]Department of Clinical Neurosciences and the Hotchkiss Brain Institute, University of Calgary, Calgary, Canada. [10]Department of Neurology, School of Medicine, Technical University of Munich, Munich, Germany. [11]Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK. [12]Genetic Epidemiology and Genomics Laboratory, Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, CA, USA. [13]Department of Health Sciences and Center on Auto-immune and Allergic Diseases (CAAD), University of Eastern Piedmont, Novara, Italy. [14]Department of Research, Innovation and Education, Oslo University Hospital, Oslo, Norway. [15]Institute of Mechanical, Electronics and Chemical Engineering, Faculty of Technology, Art and Design, Oslo Metropolitan University, Oslo, Norway.[16]Department of Neurology, Focus Program Translational Neuroscience (FTN) and Immunotherapy (FZI), University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany. [17]Department of Neurology, Hospital Clinic Barcelona, Institut d'Investigacions Biomediques August Pi Sunyer (IDIBAPS) and Universitat de Barcelona, Barcelona, Spain. [18]Department of Neurology, Oslo University Hospital, Oslo, Norway. [19]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [20]Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, OH, USA. [21]IRCCS Fondazione Don Gnocchi ONLUS, Milano, Italy. [22]Department Neuroscience and Mental Health, City of Health and Science University Hospital of Turín, Turín, Italy. [23]Ann Romney Center for Neurologic Diseases, Brigham and

Women's Hospital, Boston, MA, USA. [24]Brigham Multiple Sclerosis Center, Brigham and Women's Hospital, Boston, MA, USA. [25]Laboratory of Human Genetics of Neurological Disorders, IRCCS San Raffaele Scientific Institute, Milan, Italy. [26]Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Vall d'Hebron Institut de Recerca, Vall d'Hebron Hospital Universitari, Barcelona, Spain. [27]Casa di Cura Privata del Policlinico, Milan, Italy. [28]Vita-Salute San Raffaele University, Milan, Italy. [29]Departments of Neurology and Genetics, Yale School of Medicine, New Haven, CT, USA. [30]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [31]Department of Neurology, University General Hospital of Larissa, Faculty of Medicine, School of Health Sciences, University of Thessaly, Larissa, Greece. [32]Center For Translational & Computational Neuroimmunology and the Multiple Sclerosis Center, Department of Neurology, Columbia University Irving Medical Center, New York, NY, USA. [33]Multiple Sclerosis Division, Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA. [34]Department of Neurology, University Hospitals Leuven, Leuven, Belgium. [35]KU Leuven, Leuven Brain Institute, Department of Neurosciences, Leuven, Belgium. [36]Neurology Unit and Laboratory of Human Genetics of Neurological Disorders, IRCCS San Raffaele Scientific Institute, Milan, Italy. [37]Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Australia. [38]Perron Institute for Neurological and Translational Science, University of Western Australia, Perth, Australia. [39]Neurology Unit, Neurorehabilitation Unit, Neurophysiology Service and Neuroimaging Research Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy. [40]Medical School, University of Cyprus, Nicosia, Cyprus. [41]Department of Neurology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan. [42]Institute for Immunology and Infectious Diseases, Murdoch University, Perth, Australia. [43]Department of Neurology, Medical University of Graz, Graz, Austria. [44]Department of Neurology, Royal Melbourne Hospital, Melbourne, Australia. [45]Florey Department of Neuroscience and Mental Health, University of Melbourne, Melbourne, Australia. [46]Florey institute of Neuroscience and Mental Health, Melbourne, Australia. [47]Department of Neurology, MS center ErasMS, Erasmus University Medical Center, Rotterdam, Netherlands. [48]Department of Neurology, John Hunter Hospital, Hunter New England Health District, Newcastle, Australia. [49]Hunter Medical Research Institute, University of Newcastle, Newcastle, Australia. [50]SC Neurologia, Dipartimento di Scienze Mediche, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy. [51]Dino Ferrari Center, Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy. [52]IRCCS Fondazione Ca' Granda Ospedale Maggiore Policlinico, Neurology Unit, Milan, Italy. [53]Danish Multiple Sclerosis Center, Department of Neurology, Copenhagen University Hospital - Rigshospitalet, Glostrup, Denmark. [54]Centre for Immunology and Allergy Research, The Westmead Institute for Medical Research, Westmead, Australia. [55]School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia. [56]Broad Institute of MIT and Harvard University, Cambridge, MA, USA. [57]Division of Genetics, Department of Medicine, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA. [58]Systems Biology and Computer Science Program, Ann Romney Center for Neurological Diseases, Department of Neurology, Brigham & Women's Hospital, Boston, MA, USA.

[59]Kaiser Permanente Division of Research, Oakland, CA, USA. [60]Department of Clinical Medicine, Faculty of Healthy and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [61]Department of Neurology, Yale School of Medicine, New Haven, CT, USA. [62]Departments of Neurology and Immunology, MS center ErasMS, Erasmus University Medical Center, Rotterdam, Netherlands. [63]Neuroimmunology Research group, Netherlands Institute for Neuroscience, Amsterdam, Netherlands. [64]Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia. [65]Department of Translational Medicine and Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), University of Eastern Piedmont, Novara, Italy. [66]Stanford University, Stanford, CA, USA. [67]Departments of Neurology and Immunobiology, Yale School of Medicine, New Haven, CT, USA. [68]University of Sydney, Sydney, Australia. [69]Westmead Institute for Medical Research, Sydney, Australia. [70]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. [71]These authors jointly supervised this work: Stephen J. Sawcer and Sergio E. Baranzini.

Author contributions. Conceived and designed the study: A.H., Ja.S., D.A.H., G.J.S., A.C., F.Z., H.F.H., A.G., S.L.H., In.K., S.J.S., S.E.B. Collected the data: A.H., J.L.M., L.A., K.A., T.F.M.A., M.B., L.F.B., N.B., T.B., A.B., S.B., Y.B., S.D.B., S.J.C., D.C., P.C., E.G.C., T.C., F.C., M.C., G.C., C.C., B.C.A.C., S.D., E.D., P.L.D., S.R.D., B.D., S.E., F.E., M.F.-P., M.F., C.G., R.G., G.H., F.H., J.H., N.I., A.G.K., M.K., T.J.K., Io.K., K.L.K., J.L.-S., M.L., S.L., F.L., L.M., S.M., C.P.M., F.M.-B., E.M., L.M.M., Lu.M., X.M., J.R.O., T.O., A.O., K.P., G.P.P., N.A.P., M.P.-V., F.P., J.P.R., Al.S., Ad.S., S.S., Ca.S., F.S., H.S., Kl.S., Cl.S., V.S., Jo.S., H.B.S., M.S., B.T., M.V., D.V., P.V., M.M.V., H.L.W., V.Y., D.A.H., G.J.S., A.C., F.Z., H.F.H., B.H., A.G., S.L.H., In.K., S.J.S., S.E.B. Performed genotyping and/or quality control: A.H., J.L.M., Ja.S., I.J., A.H.B., L.G., Io.K., K.P., Kl.S., Ká.S. Analyzed the data: A.H., P.S., S.J.S., S.E.B. Supervised the study: In.K., S.J.S., S.E.B. Drafted the manuscript: A.H., S.J.S., S.E.B. Revised and edited the manuscript: A.H., P.S., J.L.M., Ja.S., I.J., K.A., T.F.M.A., M.B., L.F.B., A.H.B., T.B., S.B., S.D.B., F.B.S.B., E.G.C., F.C., C.C., B.C.A.C., S.D., P.L.D., B.D., S.E., F.E., M.F.-P., M.F., C.G., R.G.H., N.I., M.J., A.G.K., M.K., T.J.K., K.L.K., J.L.-S., F.L., A.M., F.M.-B., L.M.M., J.R.O., G.P.P., J.P.R., H.S., Cl.S., Jo.S., M.S., B.T., M.V., M.M.V., V.Y., Ká.S., D.A.H., G.J.S., A.C., F.Z., H.F.H., B.H., A.G., S.L.H., In.K., S.J.S., S.E.B.

Tables

Table 1 | Variants associated with MS severity.

Chr.	Position (bp)	ID	Risk allele	RAF	Effect (s.e.)	*P* discovery	*P* replication	*P* combined	Genes
2	**71676999**	**rs10191329**	**A**	**0.17**	**0.089 (0.015)**	**9.7×10^{-9}**	**0.021**	**3.6×10^{-9}**	***DYSF–ZNF638***
1	172370873	rs149097173	T	0.01	0.256 (0.056)	4.1×10^{-6}	0.010	2.3×10^{-7}	*DNM3–PIGC*

Effect on ARMSS score in patients with MS. Two variants were genome-wide significant (bold) or suggestive in the discovery GWAS and confirmed in the replication population. Chr., chromosome; bp, base pair (GRCh37); RAF, risk

allele frequency; s.e., standard error.

References

1. Thompson, A. J., Baranzini, S. E., Geurts, J., Hemmer, B. & Ciccarelli, O. Multiple sclerosis. *Lancet* **391**, 1622–1636 (2018).

2. Walton, C. *et al.* Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Mult. Scler.* **26**, 1816–1821 (2020).

3. GBD 2016 Multiple Sclerosis Collaborators. Global, regional, and national burden of multiple sclerosis 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **18**, 269–285 (2019).

4. Lublin, F. D. *et al.* Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* **83**, 278–286 (2014).

5. Hauser, S. L. & Cree, B. A. C. Treatment of Multiple Sclerosis: A Review. *Am. J. Med.* **133**, 1380–1390.e2 (2020).

6. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, (2019).

7. Misicka, E. *et al.* A higher burden of multiple sclerosis genetic risk confers an earlier onset. *Mult. Scler.* 13524585211053155 (2021).

8. Isobe, N. *et al.* Association of HLA Genetic Risk Burden With Disease Phenotypes in Multiple Sclerosis. *JAMA Neurol.* **73**, 795–802 (2016).

9. Masterman, T. *et al.* HLA-DR15 is associated with lower age at onset in multiple sclerosis. *Ann. Neurol.* **48**, 211–219 (2000).

10. International Multiple Sclerosis Genetics Consortium *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219 (2011).

11. Harbo, H. F. *et al.* Oligoclonal bands and age at onset correlate with genetic risk score in multiple sclerosis. *Mult. Scler.* **20**, 660–668 (2014).

12. Hilven, K., Patsopoulos, N. A., Dubois, B. & Goris, A. Burden of risk variants correlates with phenotype of multiple sclerosis. *Mult. Scler.* **21**, 1670–1680 (2015).

13. Gourraud, P.-A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. *Ann. Neurol.* **69**, 65–74 (2011).

14. Mühlau, M., Andlauer, T. F. M. & Hemmer, B. HLA Genetic Risk Burden in Multiple Sclerosis. *JAMA neurology* vol. 73 1500–1501 (2016).

15. George, M. F. *et al.* Multiple sclerosis risk loci and disease severity in 7,125 individuals from 10 studies. *Neurol Genet* **2**, e87 (2016).

16. Moutsianas, L. *et al.* Class II HLA interactions modulate genetic risk for multiple sclerosis. *Nat. Genet.* **47**, 1107–1113 (2015).

17. International Multiple Sclerosis Genetics Consortium. Genome-wide association study of severity in multiple sclerosis. *Genes Immun.* **12**, 615–625 (2011).

18. Brassat, D. *et al.* Familial factors influence disability in MS multiplex families. French Multiple Sclerosis Genetics Group. *Neurology* **52**, 1632–1636 (1999).

19. Weinshenker, B. G., Bulman, D., Carriere, W., Baskerville, J. & Ebers, G. C. A comparison of sporadic and familial multiple sclerosis. *Neurology* **40**, 1354–1358 (1990).

20. Chataway, J. *et al.* Multiple sclerosis in sibling pairs: an analysis of 250 families. *J. Neurol. Neurosurg. Psychiatry* **71**, 757–761 (2001).

21. Hensiek, A. E. *et al.* Familial effects on the clinical course of multiple sclerosis. *Neurology* **68**, 376–383 (2007).

22. Lee, J. C. *et al.* Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat. Genet.* **49**, 262–268 (2017).

23. Liu, G. *et al.* Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease. *Nat. Genet.* **53**, 787–793 (2021).

24. van Rheenen, W. *et al.* Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. *Nat. Genet.* **53**, 1636–1648 (2021).

25. Baranzini, S. E. *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum. Mol. Genet.* **18**, 767–778 (2009).

26. Brynedal, B. *et al.* MGAT5 alters the severity of multiple sclerosis. *J. Neuroimmunol.* **220**, 120–124 (2010).

27. International Multiple Sclerosis Genetics Consortium (IMSGC) *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **45**, 1353–1360 (2013).

28. King, E. A., Davis, J. W. & Degner, J. F. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of

drug approval. *PLoS Genet.* **15**, e1008489 (2019).

29. Kurtzke, J. F. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* **33**, 1444–1452 (1983).

30. Manouchehrinia, A. *et al.* Age Related Multiple Sclerosis Severity Score: Disability ranked by age. *Mult. Scler.* **23**, 1938–1946 (2017).

31. Evans, L. M. *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat. Genet.* **50**, 737–745 (2018).

32. Hujoel, M. L. A., Gazal, S., Hormozdiari, F., van de Geijn, B. & Price, A. L. Disease Heritability Enrichment of Regulatory Elements Is Concentrated in Elements with Ancient Sequence Age and Conserved Function across Species. *Am. J. Hum. Genet.* **104**, 611–624 (2019).

33. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

34. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

35. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).

36. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

37. Kappos, L. *et al.* Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. *Lancet* **391**, 1263–1273 (2018).

38. Tremlett, H. *et al.* Impact of multiple sclerosis relapses on progression diminishes with time. *Neurology* **73**, 1616–1623 (2009).

39. Gazal, S. *et al.* Combining SNP-to-gene linking strategies to pinpoint disease genes and assess disease omnigenicity. *bioRxiv* (2021) doi:10.1101/2021.08.02.21261488.

40. Ng, B. *et al.* An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nature Neuroscience* vol. 20 1418–1426 (2017).

41. Bansal, D. *et al.* Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature* **423**, 168–172 (2003).

42. Zhu, Y., Wang, G. Z., Cingöz, O. & Goff, S. P. NP220 mediates silencing of unintegrated retroviral DNA. *Nature* **564**, 278–282 (2018).

43. Meruvu, S., Hugendubler, L. & Mueller, E. Regulation of Adipocyte Differentiation by the Zinc Finger Protein ZNF638. *Journal of Biological Chemistry* vol. 286 26516–26523 (2011).

44. Lu, J. *et al.* Postsynaptic positioning of endocytic zones and AMPA receptor cycling by physical coupling of dynamin-3 to Homer. *Neuron* **55**, 874–889 (2007).

45. Edvardson, S. *et al.* Mutations in the phosphatidylinositol glycan C () gene are associated with epilepsy and intellectual disability. *J. Med. Genet.* **54**, 196–201 (2017).

46. Ascherio, A. *et al.* Vitamin D as an early predictor of multiple sclerosis activity and progression. *JAMA Neurol.* **71**, 306–314 (2014).

47. Fitzgerald, K. C. *et al.* Association of Vitamin D Levels With Multiple Sclerosis Activity and Progression in Patients Receiving Interferon Beta-1b. *JAMA Neurol.* **72**, 1458–1465 (2015).

48. Mowry, E. M. *et al.* Body mass index, but not vitamin D status, is associated with brain volume change in MS. *Neurology* **91**, e2256–e2264 (2018).

49. Tettey, P. *et al.* An adverse lipid profile and increased levels of adiposity significantly predict clinical course after a first demyelinating event. *J. Neurol. Neurosurg. Psychiatry* **88**, 395–401 (2017).

50. Dekker, I. *et al.* Predicting clinical progression in multiple sclerosis after 6 and 12 years. *European Journal of Neurology* vol. 26 893–902 (2019).

51. D'hooghe, M. B., Haentjens, P., Van Remoortel, A., De Keyser, J. & Nagels, G. Self-reported levels of education and disability progression in multiple sclerosis. *Acta Neurol. Scand.* **134**, 414–419 (2016).

52. Harding, K. E. *et al.* Socioeconomic status and disability progression in multiple sclerosis: A multinational study. *Neurology* **92**, e1497–e1506 (2019).

53. Sumowski, J. F. *et al.* Brain reserve against physical disability progression over 5 years in multiple sclerosis. *Neurology* **86**, 2006–2009 (2016).

54. Kellar-Wood, H., Robertson, N., Govan, G. G., Compston, D. A. & Harding, A. E. Leber's hereditary optic neuropathy mitochondrial DNA mutations in multiple sclerosis. *Ann. Neurol.* **36**, 109–112 (1994).

55. Ismail, A. *et al.* Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of C9ORF72. *J. Neurol. Neurosurg. Psychiatry* **84**, 79–87 (2013).

56. Yu-Wai-Man, P., Spyropoulos, A., Duncan, H. J., Guadagno, J. V. & Chinnery, P. F. A multiple sclerosis-like disorder in patients with OPA1 mutations. *Ann Clin Transl Neurol* **3**, 723–729 (2016).

57. Palace, J. *et al.* Assessing the long-term effectiveness of interferon-beta and glatiramer acetate in multiple sclerosis: final 10-year results from the UK multiple sclerosis risk-sharing scheme. *J. Neurol. Neurosurg. Psychiatry* **90**, 251–260 (2019).

58. Smith, M. M. & Arnett, P. A. Factors related to employment status changes in individuals with multiple sclerosis. *Mult. Scler.* **11**, 602–609 (2005).

59. McKay, K. A., Ernstsson, O., Manouchehrinia, A., Olsson, T. & Hillert, J. Determinants of quality of life in pediatric- and adult-onset multiple sclerosis. *Neurology* **94**, e932–e941 (2020).

60. University of California, San Francisco MS-EPIC Team *et al.* Silent progression in disease activity-free relapsing multiple sclerosis. *Ann. Neurol.* **85**, 653–666 (2019).

61. Tremlett, H., Zhao, Y., Joseph, J., Devonshire, V. & UBCMS Clinic Neurologists. Relapses in multiple sclerosis are age- and time-dependent. *J. Neurol. Neurosurg. Psychiatry* **79**, 1368–1374 (2008).

62. Vandebergh, M. *et al.* Genetic Variation in WNT9B Increases Relapse Hazard in Multiple Sclerosis. *Ann. Neurol.* **89**, 884–894 (2021).

63. Robbez-Masson, L. *et al.* The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res.* **28**, 836–845 (2018).

64. Collins, P. L., Kyle, K. E., Egawa, T., Shinkai, Y. & Oltz, E. M. The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8367–8372 (2015).

65. Bjornevik, K. *et al.* Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* **375**, 296–301 (2022).

66. Lanz, T. V. *et al.* Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature* **603**, 321–327 (2022).

67. Martyn, C. N., Cruddas, M. & Compston, D. A. Symptomatic Epstein-Barr virus infection and multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **56**, 167–168 (1993).

68. Antony, J. M. *et al.* Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. *Nat. Neurosci.* **7**, 1088–1095 (2004).

69. Kremer, D. *et al.* Human endogenous retrovirus type W envelope protein inhibits oligodendroglial precursor cell differentiation. *Ann. Neurol.* **74**, 721–732 (2013).

70. Hartung, H.-P. *et al.* Efficacy and safety of temelimab in multiple sclerosis: Results of a randomized phase 2b and extension study. *Mult. Scler.* **28**, 429–440 (2022).

71. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).

72. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat. Genet.* **49**, 1107–1112 (2017).

73. Lam, M. *et al.* Large-Scale Cognitive GWAS Meta-Analysis Reveals Tissue-Specific Neural Expression and Potential Nootropic Drug Targets. *Cell Rep.* **21**, 2597–2613 (2017).

74. Jäkel, S. *et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).

75. Galvin, J. E., Palamand, D., Strider, J., Milone, M. & Pestronk, A. The muscle protein dysferlin accumulates in the Alzheimer brain. *Acta Neuropathol.* **112**, 665–671 (2006).

76. Trapp, B. D. *et al.* Axonal Transection in the Lesions of Multiple Sclerosis. *New England Journal of Medicine* vol. 338 278–285 (1998).

77. Trinh, J. *et al.* DNM3 and genetic modifiers of age of onset in LRRK2 Gly2019Ser parkinsonism: a genome-wide linkage and association study. *The Lancet Neurology* vol. 15 1248–1256 (2016).

78. Brown, E. E. *et al.* Analysis of DNM3 and VAMP4 as genetic modifiers of LRRK2 Parkinson's disease. *Neurobiol. Aging* **97**, 148.e17–148.e24 (2021).

79. McDade, J. R., Naylor, M. T. & Michele, D. E. Sarcolemma wounding activates dynamin-dependent endocytosis in striated muscle. *FEBS J.* **288**, 160–174 (2021).

80. Bhargava, P. *et al.* The vitamin D to ameliorate multiple sclerosis (VIDAMS) trial: study design for a multicenter, randomized, double-blind controlled trial of vitamin D in multiple sclerosis. *Contemp. Clin. Trials* **39**, 288–293 (2014).

81. Manuel Escobar, J. *et al.* Body mass index as a predictor of MS activity and progression among participants in BENEFIT. *Mult. Scler.* 13524585211061861 (2022).

82. Gkatzionis, A. & Burgess, S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int. J. Epidemiol.* **48**, 691–701 (2019).

83. Vasileiou, E. S. *et al.* Socioeconomic disparity is associated with faster retinal neurodegeneration in multiple sclerosis. *Brain* **144**, 3664–3673 (2021).

84. Cabeza, R. *et al.* Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat. Rev. Neurosci.* **19**, 701–710 (2018).

85. Gazzina, S. *et al.* Education modulates brain maintenance in presymptomatic frontotemporal dementia. *J. Neurol. Neurosurg. Psychiatry* **90**, 1124–1130 (2019).

86. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).

87. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).

88. Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nat. Genet.* **53**, 1616–1621 (2021).

89. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–75, S1–3 (2012).

90. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).

91. Benner, C. *et al.* Prospects of Fine-Mapping Trait-Associated Genomic Regions by Using Summary Statistics from Genome-wide Association Studies. *Am. J. Hum. Genet.* **101**, 539–551 (2017).

92. Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).

93. Mountjoy, E. *et al.* An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat. Genet.* **53**, 1527–1533 (2021).

94. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

95. Kerimov, N. *et al.* A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat. Genet.* **53**, 1290–1299 (2021).

96. Ghoussaini, M. *et al.* Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* **49**, D1311–D1320 (2021).

97. Karlsson, M. *et al.* A single-cell type transcriptomics map of human tissues. *Sci Adv* **7**, (2021).

98. de Bakker, P. I. W. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* **38**, 1166–1172 (2006).

99. Harroud, A. *et al.* The relative contributions of obesity, vitamin D, leptin, and adiponectin to multiple sclerosis risk: A Mendelian randomization mediation analysis. *Mult. Scler.* **27**, 1994–2000 (2021).

100. Bowden, J. *et al.* Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int. J. Epidemiol.* **48**, 728–742 (2019).

101. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).

102. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).

103. Burgess, S., Foley, C. N., Allara, E., Staley, J. R. & Howson, J. M. M. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat. Commun.* **11**, 376 (2020).

Figures

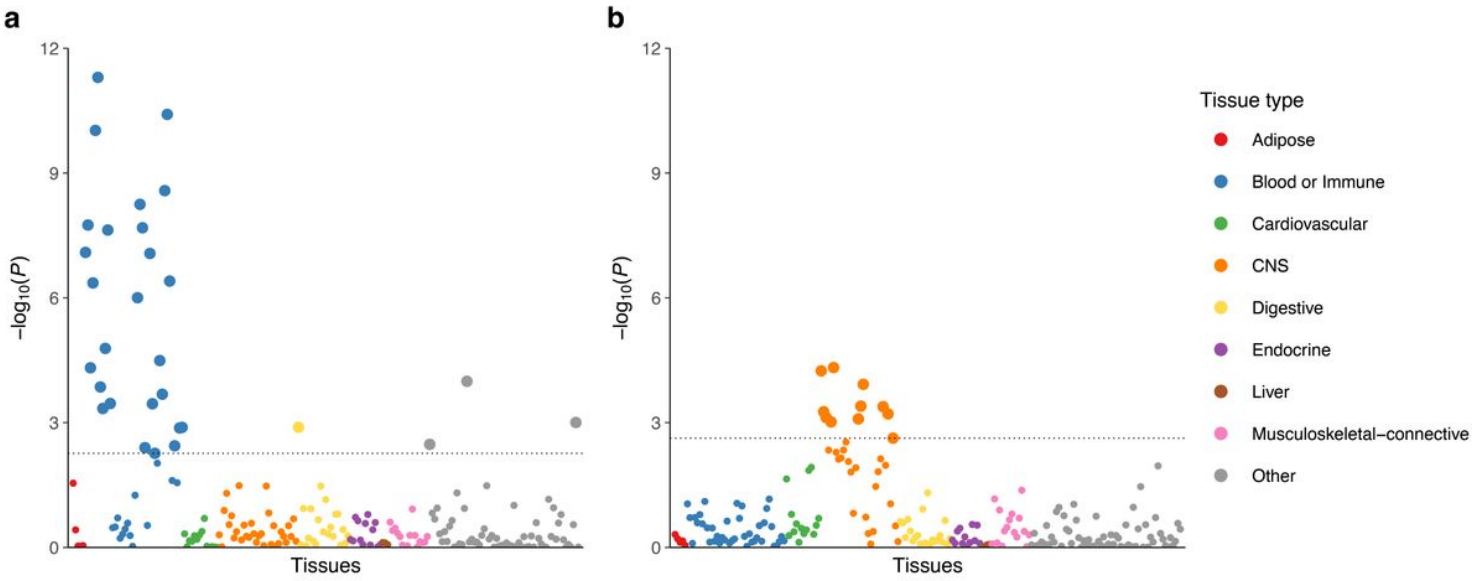

Figure 1

Tissue and cell type heritability enrichment. a, MS susceptibility from previous meta-analysis6. **b,** MS severity from this study. While susceptibility associations display strong immunological lymphoid and myeloid enrichment, our analysis of MS severity uncovered significant enrichment exclusively in CNS tissues. Each point represents one of 205 tissues and cell types, grouped by color into 9 categories. Large circles are significant at a false discovery rate cutoff of 0.05 (dotted line). Full results including tissue and cell type labels are provided in **Supplementary Tables 7 and 8**.

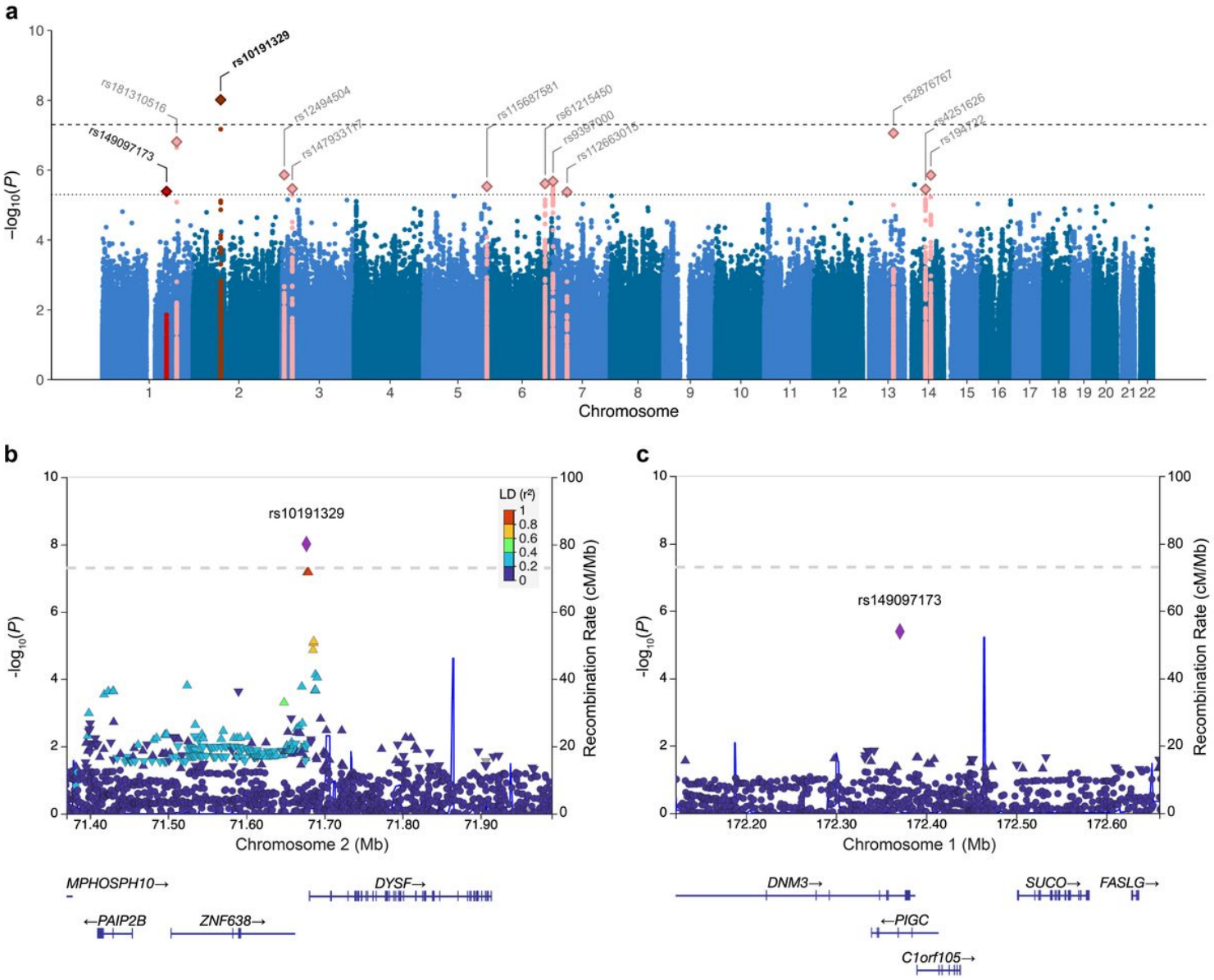

Figure 2

Within-cases GWAS identifies a novel locus associated with MS severity. a, Genome-wide association statistics obtained by linear regression of ARMSS scores. The -log10(*P*) are plotted against chromosomal position. The horizontal dashed line corresponds to the genome-wide significant threshold (*P* < 5×10-8) and the horizontal dotted line reflects the threshold for suggestive association (*P* < 5×10-6). The bold label indicates the lead genome-wide significant and replicated variant. Variants labeled in gray were not replicated. **b**, Locus Zoom plot for rs10191329 (*DYSF-ZNF638* locus). **c**, Locus Zoom plot for rs149097173 (*DNM3-PIGC* locus). Top, -log10(*P*) for variants at each locus (left y-axis) with the recombination rate indicated by the blue line (right y-axis); bottom, gene positions. Colors represent LD (*r2* values) with the lead variant. LD, linkage disequilibrium.

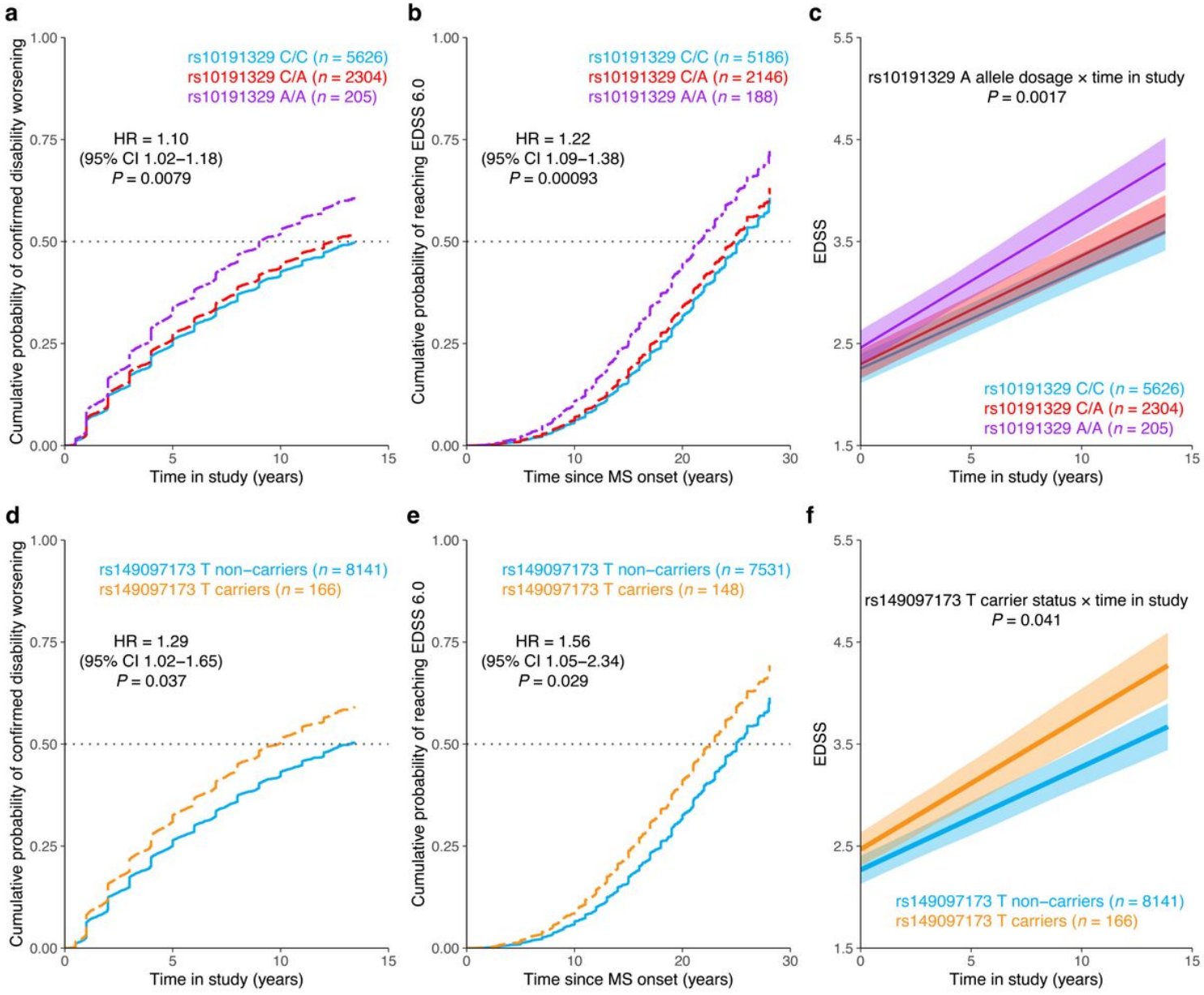

Figure 3

MS severity variants accelerate disability accumulation in longitudinal analysis. a, Covariate-adjusted cumulative incidence of 24-week confirmed disability worsening in MS patients based on rs10191329 genotype. Similar to MS clinical trials37, worsening was defined as an increase in EDSS by 1.0 if the baseline score was < 5.5 and by 0.5 if the baseline was ≥ 5.5. **b**, Covariate-adjusted cumulative incidence of requiring a walking aid for the same lead variant. Homozygous carriers had a 3.7-year shorter median time to require a walking aid. **c**, Adjusted mean EDSS scores over time predicted from linear mixed model analysis showed faster disability worsening in risk allele carriers. Shaded ribbons indicate the standard error of the mean over time. The same analyses were repeated for the low-frequency variant rs149097173 (**d–f**); carriers had a 2.2-year shorter median time to require a walking aid. HR and *P* values were obtained from Cox proportional hazards models using imputed allele dosage for rs10191329, and carrier status for rs149097173 (**a–b,d–e**; Methods). CI, confidence interval; HR, hazard ratio.

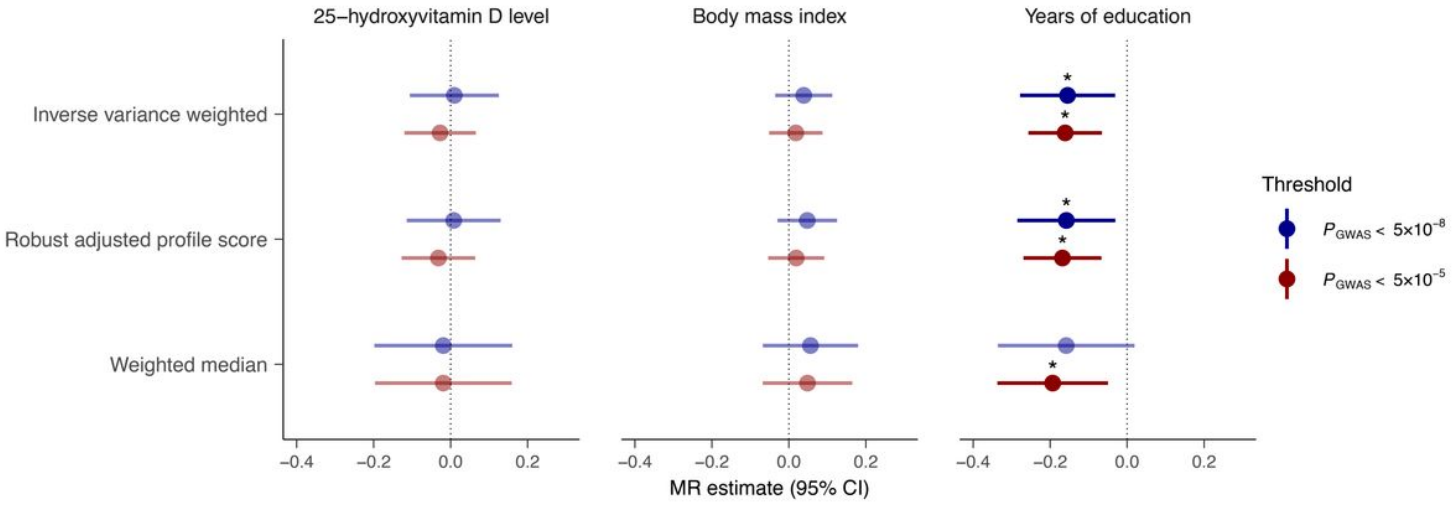

Figure 4

Mendelian randomization analysis estimates from MS severity and a priori selected phenotypes. MR results for the effect of 25OHD, BMI and years of education on RINT(ARMSS). The inverse variance weighted analysis and sensitivity analyses consistently demonstrated reduced MS severity with higher years of education. Additional MR methods (MR-Egger intercept and MR-PRESSO) showed no evidence of pleiotropy (**Supplementary Table 17**). Results for 25OHD and BMI were not significant. Point estimates are presented for two *p*-value instrument selection thresholds, with error bars reflecting 95% confidence intervals. Significant results are marked with an asterisk. 25OHD, 25-hydroxyvitamin D; ARMSS, age-related multiple sclerosis severity score; BMI, body mass index; RINT, rank-based inverse-normal transformation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplTableindex.pdf
- SupplTable1.pdf
- SupplTable2.pdf
- SupplTable3.pdf
- SupplTable4.pdf
- SupplTable5.pdf
- SupplTable6.pdf
- SupplTable7.pdf
- SupplTable8.pdf
- SupplTable9.pdf

- SupplTable10.pdf
- SupplTable11.pdf
- SupplTable12.pdf
- SupplTable13.pdf
- SupplTable14.pdf
- SupplTable15.pdf
- SupplTable16.pdf
- SupplTable17.pdf
- fig1ext.pdf
- fig2ext.pdf
- fig3ext.pdf
- fig4ext.pdf
- fig5ext.pdf
- fig6ext.pdf
- fig7ext.pdf
- fig8ext.pdf
- fig9ext.pdf
- fig10ext.pdf
- harroudpgwassupplementnaturesubmit.docx
- EXTENDEDFIGURES.docx