PNAS RESEARCH ARTICLE | IMMUNOLOGY AND INFLAMMATION OPEN ACCESS

A disease-specific convergence of host and Epstein–Barr virus genetics in multiple sclerosis

Rosella Mechelli[a,b,1], Renato Umeton[c,d,e], Gianmarco Bellucci[f], Rachele Bigi[f], Virginia Rinaldi[f], Daniela F. Angelini[g], Gisella Guerrera[g], Francesca C. Pignalosa[a,b], Sara Ilari[b], Marco Patrone[h], Sundararajan Srinivasan[i], Gabriel Cerono[j], Silvia Romano[f], Maria C. Buscarinu[f], Serena Martire[k,l,m], Simona Malucchi[n], Doriana Landi[o], Lorena Lorefice[p], Raffaella Pizzolato Umeton[q,r,s], Eleni Anastasiadou[t], Pankaj Trivedi[u], Arianna Fornasiero[f], Michela Ferraldeschi[f], IMSGC WTCCC2[v,2], Alessia Di Sapio[k,l,n], Gerolama Marfia[o], Eleonora Cocco[p], Diego Centonze[w,x], Antonio Uccelli[y,z], Dario Di Silvestre[aa], Pierluigi Mauri[aa], Paola de Candia[bb], Sandra D'Alfonso[cc], Luca Battistini[g], Cinthia Farina[i], Roberta Magliozzi[dd,ee], Richard Reynolds[ee], Sergio E. Baranzini[j], Giuseppe Matarese[bb,ff], Marco Salvetti[f,x,1], and Giovanni Ristori[f]

Affiliations are included on p. 10.

Edited by Lawrence Steinman, Stanford University Beckman Center for Molecular and Genetic Medicine, Stanford, CA; received September 14, 2024; accepted January 27, 2025

Recent sero-epidemiological studies have strengthened the hypothesis that Epstein–Barr virus (EBV) may be a causal factor in multiple sclerosis (MS). Given the complexity of the EBV–host interaction, various mechanisms may be responsible for the disease pathogenesis. Furthermore, it remains unclear whether this is a disease-specific process. Here, we showed that genes encoding EBV interactors are enriched in loci associated with MS but not with other diseases and in prioritized therapeutic targets. Analyses of MS blood and brain transcriptomes confirmed a dysregulation of MS-associated EBV interactors affecting the CD40 pathway. Such interactors were strongly enriched in binding sites for the EBV nuclear antigen 2 (EBNA2) viral transcriptional regulator, often in colocalization with CCCTC binding factor (CTCF) and RNA Polymerase II Subunit A (POLR2A). EBNA2 was expressed in the MS brain. The 1.2 EBNA2 allele downregulated the expression of the CD40 MS-associated gene analogously to the CD40 MS-risk variant. Finally, we showed that the 1.2 EBNA2 allele associates with the risk of MS. This study delineates how host and viral genetic variability converge in MS-specific pathogenetic mechanisms.

autoimmunity | Epstein–Barr virus | multiple sclerosis

Significance

The etiology of multiple sclerosis remains unclear. As a result, current therapies, though effective, target disease mechanisms rather than the underlying causes. Hence, there is a need to prolong treatments for years, if not decades. Recent evidence highlights Epstein–Barr virus as a key factor in disease development, opening new opportunities for the development of therapies that target one of the root causes of the disease. By integrating diverse experimental approaches, we uncover disease-specific host–virus interactions and show that at-risk human and viral genetic factors converge on similar mechanisms. Such mechanisms, and the identified at-risk viral subtypes, may inform highly targeted preventive or therapeutic strategies.

As for other major autoimmune conditions, our knowledge about the causes of multiple sclerosis (MS) remains incomplete. In addition to an undeniable genetic susceptibility (1), very recent progress supports the causal role of Epstein–Barr virus (EBV) in MS (2). However, key questions remain open. Concerning the epidemiologically defined role of EBV, confirmation and mechanistic explanations from other fields are needed (3). Immunological studies highlight a dysregulated antibody response to the EBV-peptidome (4) but also a pathogenetic B and T cell cross-reactivity between portions of the EBV nuclear antigen 1 (EBNA1) protein and different central nervous system (CNS) antigens such as Anoctamin 2, Glial Cell Adhesion Molecule (GlialCAM), and alpha-crystallin B (5–7). Moreover, recent work shows that healthy individuals at high risk of pathogenetic cross-reactivity between EBNA1 and GlialCAM are protected by cytotoxic immune cells targeting the autoreactive cells (8). However, we still do not fully capture the complex interaction between a sophisticated virus (9)—with its genetic variability—and human predisposing genetic variability (10), how it leads to the disease itself, why it occurs in a tiny fraction of the EBV-infected population, and whether this is a disease-specific process. In fact, EBV is potentially implicated in other autoimmune conditions. In some of these diseases, the same viral transcriptional regulator—Epstein–Barr nuclear antigen 2 (EBNA2)—may operate across risk loci in various conditions, including MS. EBV is also a causative agent of infectious mononucleosis (IM), a recognized risk factor for MS. Recent studies on the genetic control of IM suggest that the risk association between anti-EBNA-1 IgG levels and MS may partly stem from overlapping HLA associations, with *DRB1*15:01* being a poor class II antigen in the immune defense against EBV (11). In all, what determines the distinct outcomes of a shared exposure to EBV remains unclear. In this study, we started from human genetic data to understand EBV causality in MS and what distinguishes it from other diseases.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: rosella.mechelli@uniroma5.it or marco.salvetti@uniroma1.it.

[2]Complete lists of IMSGC and WTCCC2 can be found in the *SI Appendix*.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2418783122/-/DCSupplemental.

Published April 4, 2025.

Results

Genes Coding for Herpesvirus Interactors Are Enriched in MS Risk Loci. We identified modules of genes whose products are known to physically interact ("interactomes") with environmental exposures or participate in biological processes of plausible, uncertain, or unlikely relevance for several complex disorders. In detail, we investigated MS, type 1 and type 2 diabetes, rheumatoid arthritis, Crohn's disease, celiac disease as autoimmune diseases, and bipolar disorder, hypertension, coronary artery disease as nonautoimmune conditions. We considered 20 interactomes, 10 related to environmental exposures, 10 to biological processes (*SI Appendix,* Table S1) and 1 Human-microRNA targets (human miRNA-mRNA) network.

To verify the reproducibility of the results, we set a discovery stage with MS-GWAS data from the dataset published in 2011 (12) and from Immunochip data (13). To evaluate the contribution of all MS-associated SNPs, we constructed two combinations of the MS-GWAS 2011 and MS-GWAS 2013: In the first combination, METACHIP1, MS-GWAS 2011 data were considered in case of overlap (i.e., if both chips had a SNP in the same position, the MS-GWAS 2011 *P*-value was preferred); in the second combination, METACHIP2, MS-GWAS 2013 data were preferred in case of overlap (*SI Appendix,* Fig. S1). The GWAS data in the other diseases were from the Wellcome Trust Case Control Consortium 2 (14).

We selected significant SNPs associated with MS and the other complex diseases from GWAS datasets, considering a *P*-value cut-off of association with each disease at 0.05. We investigated their statistical enrichment of association within each one of the 20 selected interactomes using Association List Go AnnoTatOR (ALIGATOR).

When analyzing the frequency of disease-associated genes from distinct GWAS studies in the 20 interactomes, viral interactomes were consistently associated with MS (Fig. 1*A*). The three herpes viruses studied showed statistical significance with good consistency throughout the analyses (levels of significance: EBV > HHV8 > CMV), except for MS GWAS 2013. Given the design of GWAS 2013 (SNPs selected from immune system-related loci shared by autoimmune disorders), this result may suggest that herpesviruses do not associate with MS through genetic traits common to other autoimmune diseases. In fact, the herpesvirus association was characteristic of MS but not of any other condition studied, including immune-mediated disorders (Fig. 1*A*). Other isolated associations in MS and non-MS conditions are visible in Fig. 1*A*. By calculating Spearman's correlations, we excluded linear correlations between the size of single interactomes and their cumulative *P*-value of association with the diseases (*SI Appendix,* Figs. S2 and S3)

To assess the consistency of the MS-EBV association, we verified whether the result could be confirmed in the latest 2019 MS GWAS (15), in which a proportion of the genotyped population was independent of the previous GWAS, and was robust with respect to another enrichment testing approach. Moreover, for the SNP-to-gene mapping, we switched from ALIGATOR to the Open Targets machine-learning model, a more recent approach that integrates fine mapping and several functional genomic features (16). At the same SNP *P*-value cut-off set at 0.05, the results showed a remarkable macroscopic coherence with the 2011 and 2013 GWAS analysis, confirming a significant association of herpesviruses and HIV interactomes. Concerning more "sporadic" signals, we could confirm those of the VIRORF, HBV, AHR, and SIRT7 interactomes but not those of Polyomavirus, *Chlamydia*, and AIRE. Other associated interactomes emerged (HCV, SIRT7, HDACs, huIFN, inflammasome, SIRT1, and VDR) with HCV appearing to be more consistent. Interestingly, when we verified the sensitivity of the results concerning different SNP *P*-value cut-offs ($P < 0.03$, 0.01, 5×10^{-5}, and 5×10^{-8}), we witnessed a progressive clarification of the results with the EBV interactome becoming increasingly significant, followed by those of CMV and, to a lesser extent, of HIV (Fig. 1*B* and *SI Appendix,* Fig. S4 *A–E*). We performed the same analysis for Systemic Lupus Erythematosus (SLE), hereby considered as a positive control given its robust association with EBV. Consistently, results showed an enrichment of EBV interactors (*SI Appendix,* Fig. S4*F*), confirming the reliability of our approach.

EBV and Human Transcriptional Regulators Bind to Genomic Intervals Associated with MS-Associated Interactome Genes (MS-AIG). To evaluate the existence of common regulatory mechanisms, we mapped the 10,788 nominally significant SNPs, resulting from ALIGATOR analysis, and related corresponding 741MS-AIG (Fig. 2*A* and *SI Appendix,* Table S2) in the human genome. We unveiled a wide distribution on autosomal chromosomes (Fig. 2*B*) and a prevalent localization in intronic and upstream/downstream regions (Fig. 2*C*), in line with previous studies (17–19). We then used the RegulomeDB database (20) to look for proteins with binding sites close to SNPs extracted from ALIGATOR. We overlapped the SNPs with the protein binding map generated by the Encyclopedia of DNA Elements (ENCODE) and RoadMap Epigenomics (20), obtaining a list of enriched proteins. RNA Polymerase II Subunit A (POLR2A) and CCCTC-Binding Factor (CTCF) displayed the highest regulatory potential (Fig. 2*D*). We also carried out an analysis of "co-occurrence" for POLR2A and CTCF using Fisher's exact test. After controlling *P*-values for multiple testing (FDR, Benjamini–Hochberg), co-occurrence of POLR2A and CTCF was significant ($q < 0.05$). This is in line with data showing that SNPs in CTCF binding sites could alter chromatin topology and affect gene expression in the presence of RNA polymerase II (21).

Concerning EBV transcriptional regulators, we chose EBNA2, a viral transactivator controlling both cellular and viral gene expression. Previous work from our and other groups showed that EBNA2 can bind to a substantial fraction of risk loci in MS and in other autoimmune diseases (22–25). On these bases, we evaluated whether genetic regions containing the 10,788 nominally significant SNPs selected through the interactome approach were enriched for EBNA2 binding sites. We used a colocalization approach considering our region-of-interest (ROI) as 20 Kbp upstream and downstream of the SNP positions, according to the ALIGATOR setting. We matched our ROI with the list of EBNA2 binding regions from a ChIP-Seq performed on GM12878 EBV-transformed LCLs (25) and found a statistically significant enrichment with harmonic score (HS) = 409.99 [a combination of –*P*-value = 10 to 50; Odds Ratio (OR) = 13.73; support regions = 3018; see *Materials and Methods*]. This result confirmed EBNA2 as strongly associated with MS risk loci, extending the result to nominally significant SNPs. It also implied EBNA2 in the control of the expression of susceptibility genes that are possibly part of a pathological interaction with the "environmental" (mainly viral) risk. This result extends previous data on five out of six preselected loci (22). To compare results and identify possible cooperative regulatory mechanisms between EBNA2, CTCF, and POLR2A, we replicated the RegulomeDB enrichment analysis using our colocalization approach. We confirmed a strong enrichment of POLR2A (Harmonic Score = 832.76, *P*-value = 10 to 50, OR = 27.95, support = 4,139) and CTCF binding sites (HS = 99.36, *P*-value = 10 to 50, OR = 3.31, support = 7,874) (Fig. 2*E*). We were able to identify 1,044 colocalization hotspots bound by all the three transcription factors (TFs), 2,253 by POLR2A-CTCF, 242 by POLR2A-EBNA2, 1,373 by

A

	MULTIPLE SCLEROSIS				OTHER COMPLEX DISEASES							
	GWAS 2011	GWAS 2013	METACHIP1	METACHIP2	ChD	RA	CD	T1D	T2D	HT	BD	CAD
EBV	**0,0012**	0,0828	**0,0016**	**0,0004**	0.4578	0.523	0.4716	0.9968	0.5444	0.1674	0.539	0.4784
HHV8	**0,0046**	0,2094	**0,0052**	**0,0022**	0.8732	0.6424	0.5574	0.3588	**0.0214**	0.7842	0.6578	0.3154
HIV	**0,0026**	0,3970	**0,0216**	**0,0256**	0.7636	0.903	0.6046	0.4518	0.0832	0.5958	0.7344	0.7496
CMV	**0,0444**	0,6720	**0,0482**	**0,0298**	0.2198	0.5126	0.0704	0.9056	0.3438	0.6956	0.5378	**0.0288**
VIRORF	0,0610	**0,0152**	0,0890	0,0704	0.1976	0.7448	0.4766	0.8866	**0.0372**	0.7168	0.3362	0.2036
HBV	**0,0170**	0,9636	0,1504	0,1668	0.6862	0.7642	0.3258	0.5804	0.5576	0.8778	0.3766	0.7486
Polyomavirus	0,2536	**0,0418**	0,1912	0,1996	0.8052	0.8536	0.1856	0.7474	0.1752	0.1268	0.9238	0.2774
H1N1	0,9572	0,4228	0,9738	0,9516	0.7792	0.6486	0.9022	0.5498	0.1316	0.688	0.693	0.7376
HCV	0,4244	0,1758	0,3216	0,2818	0.7698	0.5986	0.7654	0.8968	0.5666	0.9428	0.8824	0.9596
JCV	1,0000	1,0000	1,0000	1,0000	0.4108	1	0.1758	0.8256	0.1758	0.3734	0.836	0.0912
AHR	0,5428	**0,0260**	0,3744	0,5332	0.8734	0.9176	0.3814	0.8262	0.3552	0.8256	0.892	0.8538
Chlamydia	0,9910	**0,0010**	0,9864	0,9820	0.9266	0.9818	0.7758	0.1918	0.9464	0.6354	0.8938	0.201
AIRE	**0,0342**	0,2412	0,1810	0,1276	0.711	0.274	0.658	0.7804	0.9188	**0.0072**	0.5934	0.99
SIRT7	0,5146	**0,0400**	0,3138	0,3144	0.1894	0.981	0.4086	0.936	0.1748	0.4836	0.1136	0.9566
HDACs	0,9610	0,2060	0,9562	0,9126	0.886	0.6372	0.5472	0.772	0.9158	0.3306	0.368	0.9486
huIFN	0,2176	0,3104	0,3452	0,5364	0.616	0.968	0.673	0.1132	0.7168	0.177	0.2146	0.6716
Human-miRNA targets	0,1514	0,5590	0,4754	0,3818	0.0506	0.3044	**0.0036**	0.7706	0.8784	0.9244	0.9534	0.2258
Inflammasome	0,1470	0,3410	0,0722	0,0510	0.7568	0.093	0.187	0.3338	0.1026	0.3308	0.5632	0.7648
SIRT1	0,3032	0,0724	0,1374	0,1542	0.1684	0.7826	0.373	0.7812	0.727	0.9478	0.6784	0.0538
VDR	0,3020	0,5102	0,2302	0,3212	0.9458	0.9886	0.1324	0.953	0.8164	0.4766	0.7634	0.8654

p value 0,05 0,04 0,03 0,02 0,01 0

B

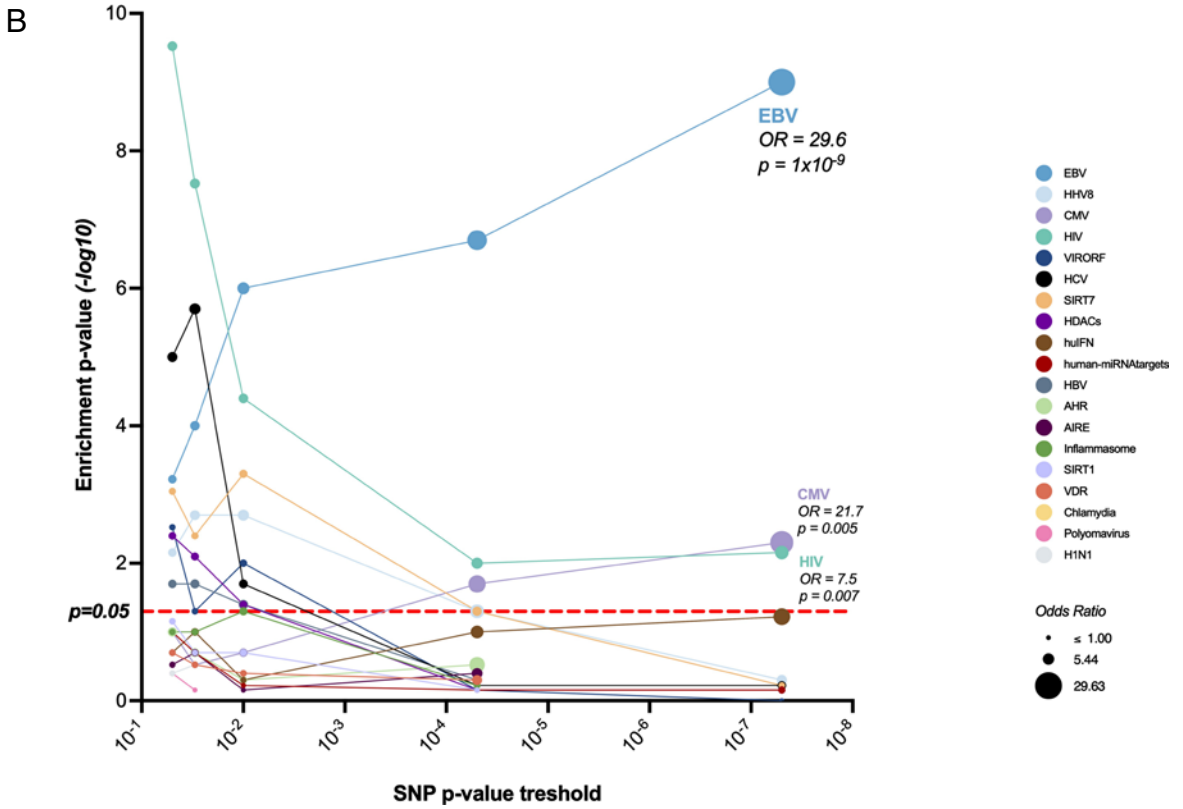


Fig. 1. Enrichment analysis of interactomes in GWAS. (*A*) Heatmap represents the interactomes and related strength of associations with diseases. The statistically relevant associations are indicated in blue, with a color gradient from white ($P > 0.05$) to blue ($P < 0.01$). (*B*) Enrichment of interactomes in MS GWAS 2019, considering different thresholds of SNP *P*-value of association with MS (*x* axis). ChD = Crohn disease; RA = rheumatoid arthritis; CD = celiac disease; T1D = type 1 diabetes; T2D = type 2 diabetes; HT = hypertension; BD = bipolar disorder; CAD = coronary artery disease; AHR = Aryl hydrocarbon receptor; AIRE = autoimmune regulator; BioGRID = Biological General Repository for Interaction Datasets; CMV = Cytomegalovirus; EBV = Epstein–Barr virus; HBV = Hepatitis B virus; HCV = Hepatitis C virus; HDACs = Histone deacetylases; HHV8 = Human Herpesvirus 8; HIV; H1N1 = Influenza A virus; hu-IFN = human innate immunity interactome for type I interferon; Human-miRNA targets = gene targets for human miRNA; Inflammasome = multiprotein complex responsible for activation of inflammatory processes and pyroptosis; JCV = JC virus; SIRT1 = Sirtuin 1; SIRT7 = Sirtuin 7; VIRORF = Virus Open Reading Frame; VDR = vitamin D receptor.

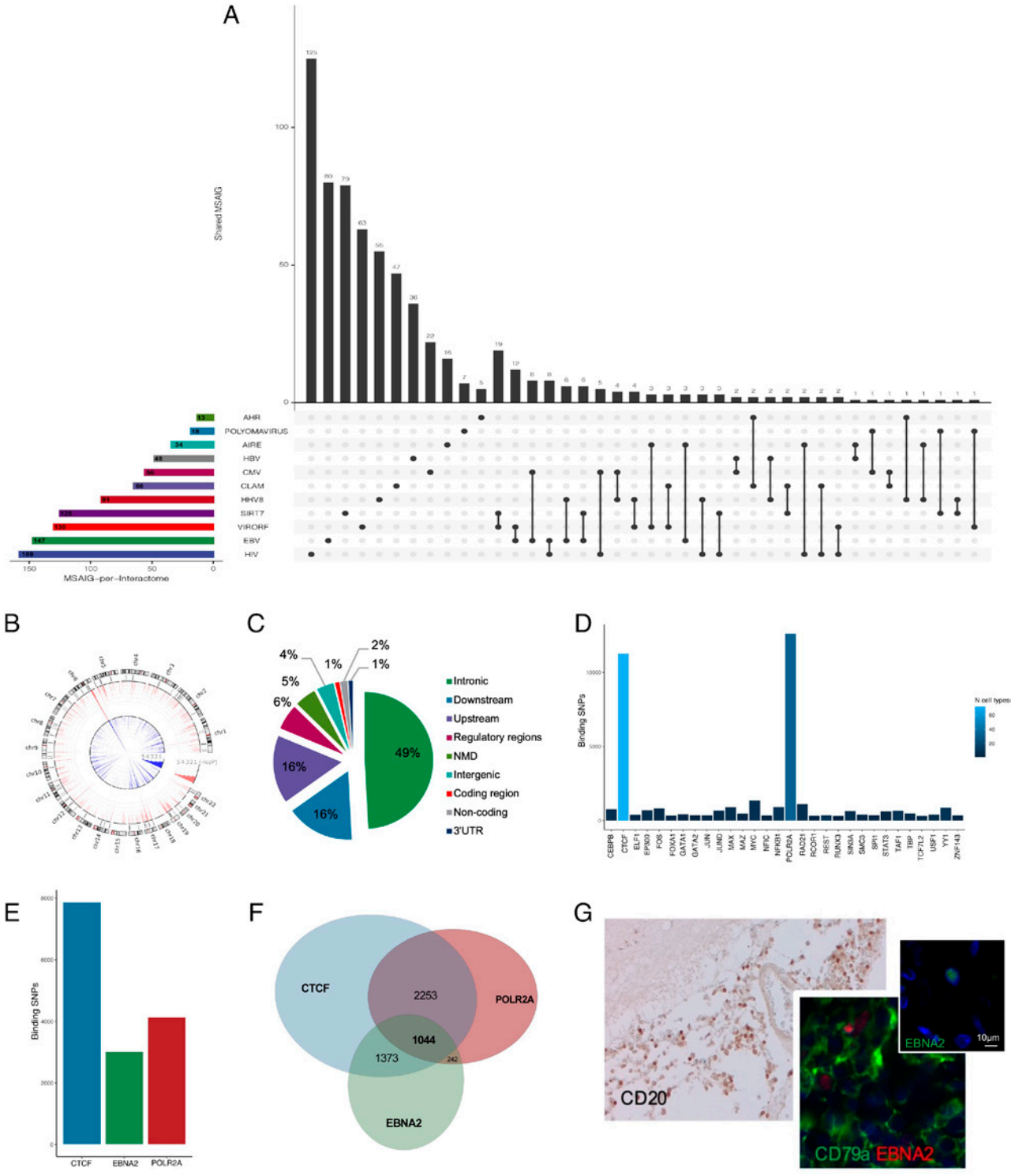


Fig. 2. Genomic distribution of the candidate interactome SNPs and protein binding enrichment near MS-associated SNPs. (*A*) The UpsetPlot displays the number of MS-AIG for each interactome of origin (*x* axis) and the intersections between gene sets with the respective size (*y* axis) (see also *SI Appendix*, Table S2). (*B*) The circos plot shows the genomic distribution of the SNPs obtained from the candidate interactome analysis. The outer circle shows the chromosomes with related cytobands. The underlying gray lines indicate the physical position of the corresponding 741 MS-AIG (each line may represent up to 25 genes). The red and blue lines in the inner circles indicate the −log(*P* value) for each SNP respectively from GWAS 2011 and GWAS 2013 (scaled from 0 to 20, which truncates the signal in several regions). (*C*) The pie chart represents the distribution of SNPs in different regions of the human genome, showing an enrichment in intronic and downstream/upstream regions. (*D*) The bar plot displays the number of SNPs bound by a given protein obtained from different cell types (vertical axis) and the top 30 transcription factors enriched on MS-associated SNPs; color scale is related to the number of cell types. (*E*) The bar plot shows the number of SNPs colocalizing with EBNA2, CTCF, and POLR2A binding sites. (*F*) The Venn graph displays the overlap of SNPs bound by CTCF, EBNA2, and POLR2A. (*G*) The image displays EBNA2 expressed by CD79a+ B cells in the leptomeninges of postmortem progressive MS cases in the presence of a substantial number of meningeal CD20+ B cells. In the *Upper Right Inset* are shown EBNA2+ nuclei. NMD = Nonsense-mediated mRNA decay; 3′UTR = three prime untranslated region.

CTCF-EBNA2 (Fig. 2*F*). These results suggest broad human–viral interactions, with EBNA2 impacting transcriptional regulation through MS-associated loci.

If the pervasive influence of EBNA2 on the regulation of the MS-AIG is directly relevant to CNS inflammation, the expression of this viral gene should be detectable in B cells in MS lesions.

Table 1. MS-AIG differentially expressed in peripheral blood leukocytes and brain areas

MS phenotypes	No. of DEG	*P*-value	OR
Blood			
CIS	166	0.59	0.28
RRMS	170	0.34	0.89
PPMS	249	0.082	3.02
SPMS	226	0.53	0.37
Brain			
NAGM-F +	48	**0.008**	1.53
NAGM-F-	127	**0.001**	1.4
GML-F +	52	**0.003**	1.59
GML-F-	133	**0.001**	1.42

Frequency of MS-AIG in DEG obtained from distinct MS phenotypes and brain areas. The associations reaching significance are highlighted in bold. CIS = clinically isolated syndrome; RRMS = Relapsing-Remitting multiple sclerosis; PPMS = Primary progressive multiple sclerosis; SPMS = Secondary progressive multiple sclerosis; DEG= differentially expressed genes; OR= Odds ratio.

Neuropathological analysis of fixed-frozen brain sections, obtained from postmortem progressive MS cases, revealed the expression of EBNA2 in CD79a+ B cells, in particular in leptomeninges enriched in CD20+ B cells (Fig. 2*G*). Intrathecal inflammatory compartments, such as leptomeninges and perivascular spaces, represent possible niches of EBV persistence, with EBNA2 sustaining a dysregulated B cell function.

Dysregulation of MS-Associated EBV Interactors in Blood and CNS Transcriptomes. To evaluate the functional implications of the above results and to validate them, we matched the 741 MS-AIG with microarray gene expression data obtained from peripheral blood leukocytes of 40 healthy subjects and persons with different MS phenotypes (26). We calculated whether the MS-AIG were enriched in the lists of genes differentially expressed ($P < 0.05$ and fold change > 1.5) in each MS phenotype. In total, 464 MS-AIG were differentially expressed in at least one MS phenotype. We did not observe significant enrichments in any of the four disease subtypes when compared with a random selection of transcripts from the whole transcriptome (Table 1). When considering the MS-AIG from single interactomes separately, we observed an enrichment of differentially expressed MS-AIG from AIRE interactome in CIS patients, from AIRE and SIRT7 interactomes in RRMS, and EBV interactome in PPMS (Table 2).

At the CNS level, we matched the 741 MS-AIG with the microarray gene expression data obtained from postmortem brain tissue of 20 people with SPMS and 10 non-neurological controls. We considered gray matter lesions (GML) and normal appearing gray matter (NAGM) from MS brains that exhibited presence (F+) or absence (F–) of meningeal lymphoid-like structures associated with different degrees of progression (27). We calculated whether the overall frequency of MS-AIG was higher in the dysregulated gene lists ($P < 0.05$ and fold change > 1.5) than its expected frequency in a random selection of transcripts from the whole transcriptome. Overall, 169 MS-AIG were differentially expressed in at least one of the four conditions (GML-F+, GML-F–, NAGM-F+, NAGM-F). The result was significant (Table 1). Considering separately the MS-AIG obtained from each MS-associated interactome, only genes belonging to the EBV interactome were significantly increased in all four conditions (Table 2).

The above results imply a potential role for EBV-related pathogenic responses in the neurodegenerative and inflammatory aspects of MS. This prompted us to consider recent data showing enrichment of EBV interactors in protein modules involved in Hereditary Spastic Paraplegia (HSP) (28), an inherited degenerative axonopathy sharing clinical similarities with progressive MS. Importantly, it has been shown that HSP-related mutations are more frequent in persons with PPMS (29). We, therefore, matched the 741 MS-AIG with the genes of the HSP-protein modules (n = 295) and found 71 genes in common. Of these, 34 were included in EBV interactome (n = 147 from MS-AIG) (*SI Appendix,* Fig. S5). This number was larger than randomly expected (P-value < 0.00001).

MS-Associated EBV and Human Genetic Variants Concur in the Downregulation of CD40 Expression. We used MetaCore© to identify biological functions possibly affected by the dysregulated MS-AIG. The 741 MS-AIG were mapped to the internal MetaCore KnowledgeBase, and a classification in canonical pathways was obtained. The CD40 pathway emerged in both peripheral blood and CNS compartments as the leading biological function (Fig. 3*A*).

Focusing on the CD40 molecule itself, we asked whether host- and virus-specific factors could drive CD40 alterations in MS. By flow cytometry, we evaluated the surface protein expression of CD40 in EBV-infected B cells (LCLs) obtained from 13 persons with MS (untreated at the time of sampling) and 8 healthy donors. Out of 33 LCLs, 21 were spontaneously outgrowing LCLs (spLCLs), i.e., infected by the endogenous EBV, and 12 were LCLs infected with the recombinant EBV strain B95.8 (B95.8LCLs). We also evaluated the expression of CD80 and CD86 costimulatory molecules known to be induced by CD40 and by EBV infection (30). To control for influences from human and EBV genetic variability, we genotyped all donors for the MS-associated SNPs on the CD40 (rs4810485) and CD86 genes (rs9282641), associated with an altered membrane expression of the proteins (31), and for the EBNA2 alleles.

We found a reduction of CD40 protein expression in MS-spLCLs, compared to HD-spLCLs ($P = 0.04$; Fig. 3*B*), a nearly significant reduction of CD80 expression ($P = 0.089$; Fig. 3*D*) and no difference for CD86 (*SI Appendix,* Fig. S6). In line with the notion that estrogens modulate the expression of MS risk genes in EBV-infected B cell lines (32), the reduced CD40 and CD80 expression in MS spLCL was more pronounced in females with MS compared to healthy females ($P = 0.006$ for CD40, Fig. 3*C* and $P = 0.0037$ for CD80, Fig. 3*E*), with a near-significant difference when comparing females and males for CD40 expression within the MS group ($P = 0.049$; Fig. 3*C*). A linear regression analysis revealed an interdependence of CD40 and CD80 protein levels in MS-spLCLs ($r^2 = 0.43$, $P = 0.016$; Fig. 3*F*) but not in HD-spLCLs ($r^2 = 0.32$, $P = 0.16$; Fig. 3*G*).

The differences were not attributable to the host's genetic CD40 and CD86 variability (*SI Appendix,* Fig. S7) and were not observed in B95.8LCL (*SI Appendix,* Fig. S8). A multiple linear regression analysis, controlling for CD80 (since CD80 and CD40 levels were correlated, Fig. 3*F*), highlighted a lower expression of CD40 in MS-spLCLs compared to MS-B95.8LCLs derived from the same subjects at the same timepoint ($P = 0.04$; Fig. 3 *H* and *I*). These results supported a role of the endogenous EBV variants on the downregulation of CD40 expression in MS.

Having shown, in the previous section, a possible impact of EBNA2 on the transcription of MS-AIG, we evaluated the effect of the 1.2 EBNA2 allele, possibly associated with MS (33), on CD40 protein levels. We compared the 1.2 EBNA2-infected spLCLs with the B95-8-infected LCLs, irrespective of the MS and HD condition. Cell lines expressing the 1.2 allele showed a lower level of CD40 compared to those expressing the B95.8 allele ($P = 0.013$, Fig. 3*J*). To confirm that this phenotype is related to the EBNA2 alleles and not to other viral proteins, we quantified

Table 2. Frequency of interactome-specific MS-AIG in peripheral blood leukocytes and brain areas

MS phenotypes	Frequency of MS-AIG DEG (%)											
	AHR	AIRE	CLA	CMV	EBV	HBV	HHV8	HIV	PV	SIRT7	VIRORF	GT
Blood												
CIS	25.0	**50.0**	29.0	34.6	24.3	16.0	22.1	21.4	45.4	28.0	30.5	25.6
RRMS	25.0	**40.6**	25.8	19.2	22.9	24.0	27.3	21.9	45.4	**34.2**	2.0	23.7
PPMS	25.0	40.6	38.7	38.5	**44.6**	46.0	42.9	32.9	27.3	31.5	34.7	33.9
SPMS	31.2	37.5	27.4	32.7	33.1	28.0	37.7	32.9	9.1	30.6	34.7	32.4
Brain												
NAGM-F +	7.7	5.9	4.5	5.4	**10.1**	4.1	7.7	5.0	11.1	7.2	7.7	6.6
NAGM-F-	7.7	14.7	18.2	17.8	**21.5**	16.7	15.4	16.6	22.2	16.8	18.5	17.1
GML-F +	0.0	11.8	3.0	5.4	**12.7**	6.2	6.6	7.5	5.6	5.6	6.1	7.3
GML-F-	15.4	17.6	16.7	14.3	**22.1**	18.7	16.5	16.6	16.7	17.6	17.0	18.1

The enrichment of MS-AIG was calculated considering separately each interactome in distinct MS phenotypes and brain areas. The associations reaching significance are highlighted in bold. CIS = clinically isolated syndrome; RRMS = relapsing–remitting multiple sclerosis; PPMS = primary progressive multiple sclerosis; SPMS = secondary progressive multiple sclerosis; DEG = differentially expressed genes; NAGM-F+ = normal-appearing gray matter follicle positive; NAGM-F− = normal-appearing gray matter follicle negative; GML-F+ = gray matter lesions follicle positive; GML-F− = gray matter lesions follicle negative; CLA = *Chlamydia*; PV = polyomavirus; GT = global transcriptome.

CD40 and EBNA2 protein levels in a U2OS cell line transfected with plasmid expressing the 1.2 or the B95-8 EBNA2 alleles (the empty plasmid was used as control-mock) (Fig. 3 *K* and *L*). Using western blot analysis, we observed that at the same EBNA2 level, the cell lines expressing the 1.2 allele showed a lower level of CD40 protein compared to those expressing the B95.8 allele ($P = 0.025$) (Fig. 3*L*).

Additionally, we observed an inverse correlation between CD40 and HLA-DR protein levels in MS-spLCLs ($r^2 = 0.41$, $P = 0.018$). No clear correlation was identified with the LCL EBNA2 allele. The analysis of higher numbers of LCLs will be needed to determine whether this correlation is under the influence of EBNA2 genetic variability.

EBV Genetic Variability Affects MS Risk. To ascertain that the above EBNA2-related mechanisms are relevant for the disease, we attempted to replicate previous preliminary findings from our group showing that the EBNA2 1.2 allele was associated with the risk of MS (33).

In an independent cohort of 380 therapy-free persons with RRMS (mean age 39.4 ± 11.7; female/male ratio 1.9) and 250 HD (mean age 36.4 ± 11.1; female/male ratio 2.3), using a droplet digital PCR approach, we found that the EBNA2 1.2 allele increases the risk of MS by 2.5-fold (95% CI = 1.527 to 4.169; $P = 0.0002$).

EBV-Interacting Proteins Are Prioritized Therapeutic Targets to Address MS Gene–Environment Interactions. Using the Priority Index (Pi) pipeline (33) we found that MS-AIG were enriched in prioritized therapeutic targets for MS, with 34 MS-AIG appearing in the first 150 positions of the PI ranking ($P < 0.0001$, Fig. 4*A*). The EBV interactome was the most enriched with highly druggable targets (16 MS-AIG, $P < 0.0001$, Fig. 4 *B* and *C*), followed by HBV, HHV8, AHR, and CMV interactomes. We constructed an integrative gene–environment therapeutic module for MS, including the 34 prioritized MS-AIG with 35 MS-specific crosstalk genes, representing central hubs for multiple disease-relevant molecular pathways (34). The resultant therapeutic network included 61 highly interconnected nodes (PPI enrichment *P*-value: <1.0e-16), in which EBV interactors and the CD40 pathway MS-AIG appeared as a supporting structure of the therapeutic module, consistently with the above 'omics results (Fig. 4*D*).

We prioritized candidate therapies addressing the MS gene–environment module. Biological therapies were ranked based on their target's PI rating (*SI Appendix,* Table S3): The top-scoring hits were anti-CD40 monoclonal antibodies. Nonbiological compounds were prioritized through a network-based algorithm. The top-ranked drugs included multitarget tyrosine-kinase inhibitors, Bruton kinase inhibitors that have already entered the MS pipeline (35), leukotriene receptor antagonists, and a proteasome inhibitor (*SI Appendix,* Table S4).

Discussion

We show that the interaction between MS-predisposing genes and EBV (possibly involving other herpesviruses) has causal and disease-specific implications in MS and harbors a high therapeutic potential. The dysregulation of MS-associated EBV interactors validates the result, especially at the CNS level, where we could confirm the EBNA2 expression. This viral element emerges as a possible, pervasive regulator of the disease-predisposing, human–viral gene interaction, affecting CD40 costimulatory receptor biology. Similarly to the human CD40 risk allele, the 1.2 EBNA2 EBV variant associates with an accentuated downregulation of the CD40 molecule on EBV-infected B cells.

Besides MS, we observed an enrichment in genes coding for EBV interactors in SLE. Based on the literature (24, 36), this was an expected result. As such, it supports the methodological soundness of our approach. We also observed an enrichment of CMV interactors in MS but not in SLE risk loci, suggesting a disease specificity with no obvious explanation. Though CMV exposure may be protective in MS (37) and predisposing in SLE (38), the direction of the observed associations in different studies may vary, with evidence of regional differences, at least in MS (36). Studies comparing the complex interplay between EBV, CMV, and human predisposing genes in both diseases will be needed to clarify this issue. At this stage, we can only remark that among the three genes showing the strongest association in the CMV interactome (Dataset S1), two (TAP1 and TAPBP) are essential components of the MHC class I antigen-processing pathway (39) while the third (MICA) is involved in immune activation via NKG2D signaling and MHC class I and class II upregulation (40).

Though supported by the occurrence of autoimmunity in CD40 or CD40L primary immune deficiencies (41), the association between a reduced expression of a costimulatory molecule (CD40)

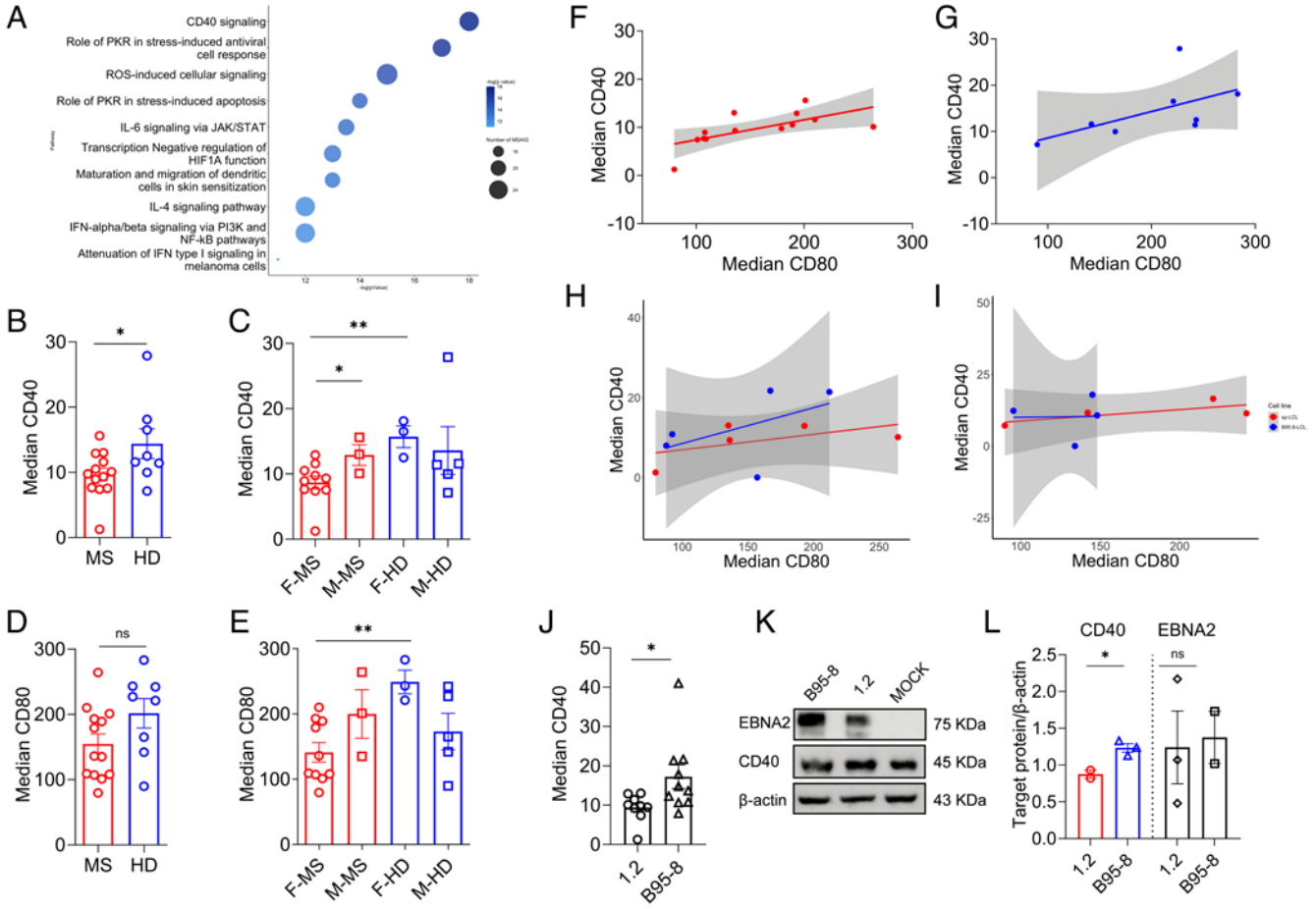
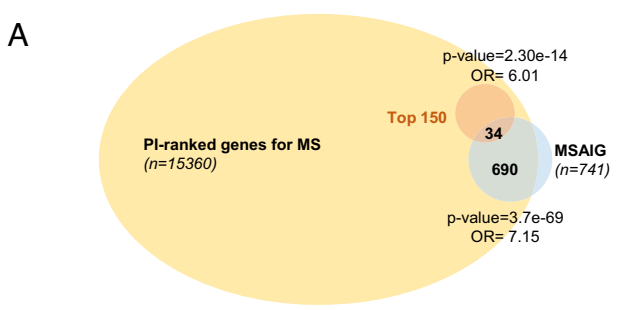


Fig. 3. CD40 signaling analysis. (*A*) The top 10 pathways obtained from the analysis of the MS-AIG. The lower *P*-value (−log Pvalue) means higher relevance of the genes within the pathway datasets. (*B*) The CD40 protein level analyzed in MS-spLCLs (n = 13) and HD-spLCLs (n = 8) is downregulated in MS-spLCLs at unpaired *t*-test (**P* = 0.04). (*C*) The CD40 protein level stratified according to the gender, CD40 protein is underexpressed in females of the MS group (F-MS) compared to males (M-MS, **P* = 0.049) and compared to the females of the control group (F-HD; ***P* = 0.006). (*D*) The CD80 protein level is comparable between MS-spLCLs and HD-spLCLs. (*E*) The CD80 protein level stratified according to the gender, CD80 protein is underexpressed in F-MS compared to the F-HD (***P* = 0.0037). (*F* and *G*) A correlation between CD40 and CD80 protein levels was evaluated by a simple linear regression analysis (Spearman correlation) in MS (r^2= 0.43, *P* = 0.016; *F*) and in HD (r^2= 0.32, P = 0.16; *G*). Gray areas represent the CI, and the straight lines represent the correlation slope. The protein levels are represented as median fluorescence intensity. Data analysis was performed using the unpaired *t*-test F = female; M = male. (*H* and *I*) The effect of the EBV strain (spLCL vs. B95.8LCL) and CD80 levels on CD40 protein expression in MS patients (n = 5; *H*) and controls (n = 4; *I*) were analyzed by a mixed-effect multiple linear regression model, setting subjects as random effect, and cell line context and CD80 levels as fixed effects. (*J*) CD40 protein levels analyzed in spLCL carrying the MS-associated 1.2 EBNA2 allele and in B95.8LCL. CD40 is reduced in spLCL with 1.2 EBNA2 (n = 8) respect to those infected by EBNA2 B95.8 (n = 10; **P* = 0.013). (*K*) Representative images of western blot probed for CD40, EBNA2, and ß-actin (internal controls) on U2OS cell lysate. (*L*) The expression of CD40, normalized to the level of mock condition, and EBNA2 (both calculated over the internal control) in U2OS transfected with 1.2 EBNA2 (CD40, n = 2; EBNA2, n = 3) and with EBNA2 B95.8 (CD40, n = 3; EBNA2, n = 2, **P* = 0.025). The data are expressed as mean ± SEM analyzed by the unpaired *t*-test.

and autoimmunity is counterintuitive. However, CD40 signaling in intrathymic B cells is a central tolerance mechanism (42) and has recently been shown to affect the T cell response to the aquaporin-4 autoantigen in neuromyelitis optica (43). Moreover, being CD40 signaling an essential component of the affinity maturation processes, its defective function may explain the diverse mimicry mechanisms highlighted by recent studies (4–7), in the context of a dysregulation that may be broader than previously thought.

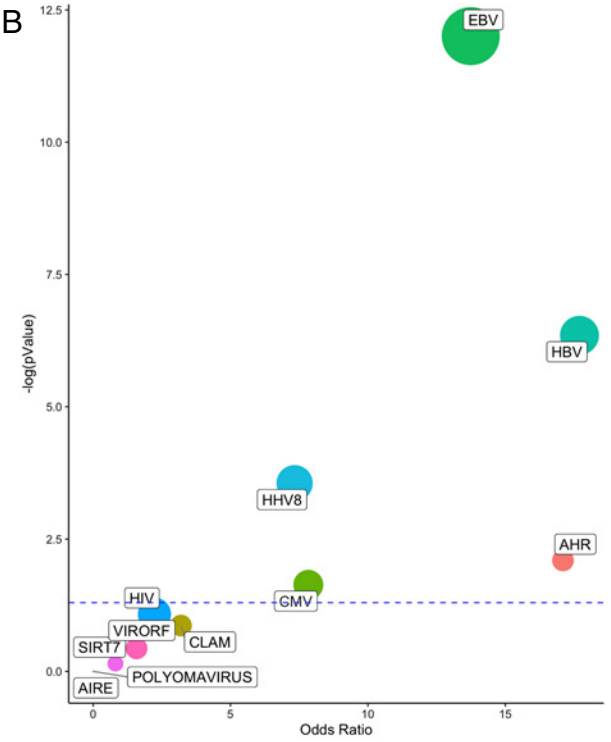
The association between EBNA2 alleles and CD40 expression has not been previously explored. Prior studies examining EBNA2 targets among MS (44) and SLE (45) risk genes demonstrated that, in B95.8 LCLs, EBNA2 expression positively correlated with CD40 levels while inversely correlating with EBV DNA copy number. Additionally, CD40 stimulation was shown to inhibit LCL proliferation. Here, we observed that, at the same EBNA2 expression level, the spLCLs carrying the 1.2 EBNA2 allele displayed lower levels of CD40 protein. This may weaken a mechanism that helps control the viral infection, possibly via the ability of CD40 to compete with its viral homologue, LMP1, for intracellular signaling molecules (44). Furthermore, reduced CD40 expression, particularly if combined with elevated HLA-DR expression, may promote molecular mimicry and autoproliferation, particularly in a *DRB1*15:01* context (46). Notably, LMP1 is also involved in NK cell-mediated protection from molecular mimicry events, in accord with recent results from other groups (8). It will be important to evaluate the effects of second-generation anti-CD40L monoclonal antibodies on the EBNA2–CD40–LMP1 interaction. These therapies are currently in clinical development (47). Their use may enable new preclinical studies to be designed around questions that emerge from the clinical trials in a "bench-to-bedside-to-bench" approach.

The association between 1.2 EBNA2 allele and MS risk provides an initial indication of how the compelling evidence of an implication of EBNA2 in autoimmunity may depend on EBNA2 genetic variability. It also advances the interpretation of a recent comparison of EBNA 2 type 1 and 2 binding to human genetic risk loci for autoimmune diseases (48), showing



C

INTERACTOME	PRIORITIZED GENES
CMV	*PML, STAT3, SPI1, STAT1*
EBV	*HLA-DRB1, HLA-DQA1, HLA-DRA, HLA-DQB1, HLA-DQA2, MYC, LYN, PML, TYK2, IRF7, SRC, TRAF1, SPI1, SYK, JAK3, TRAF3*
HIV	*DDX6, HLA-B, GRB2*
HHV8	*IRF1, IRF6, IRF7, STAT3, STAT5A, NFKBIA*
HBV	*MAPK1, IL6, JAK1, NFKB1, NFKBIA, CD44, STAT1*
AHR	*RELA, SRC*
CLAM	*RELA, HLA-A*
SIRT7	*OASL*
VIRORF	*STAT1, DDX6*



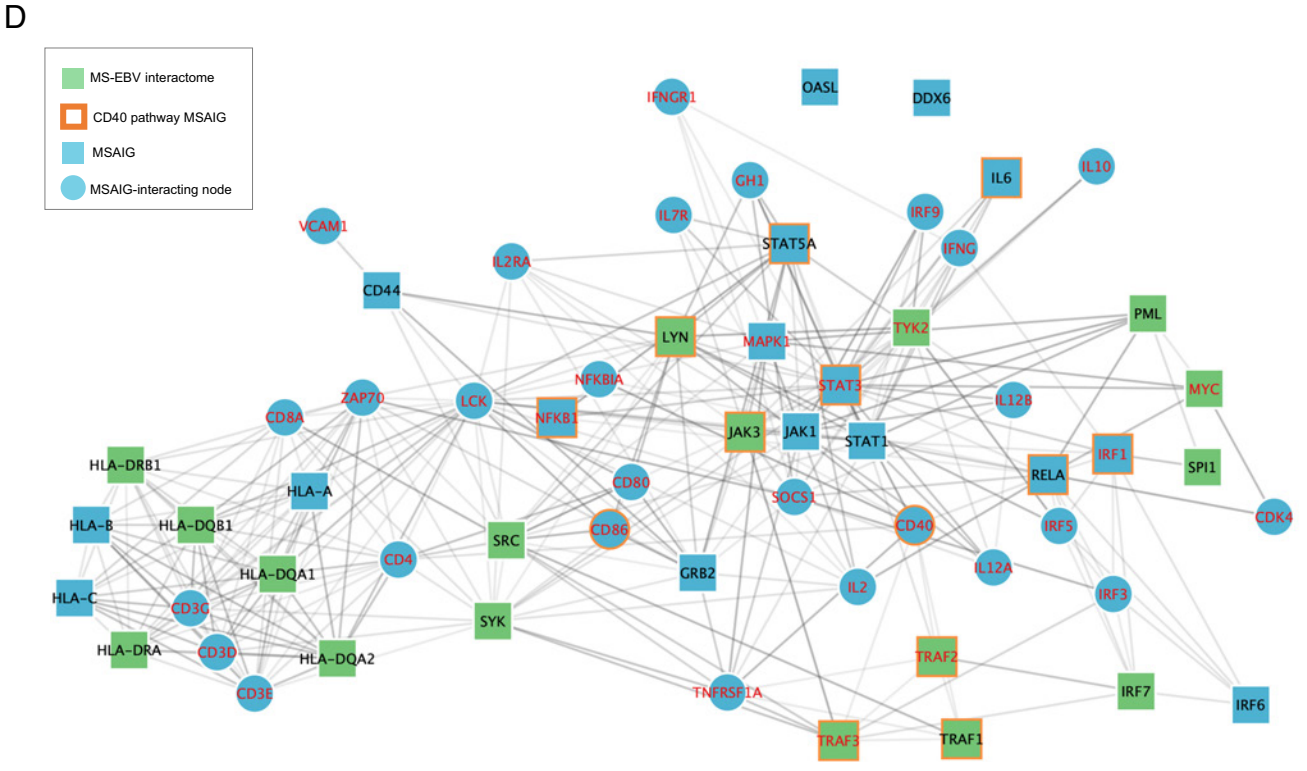


Fig. 4. Enrichment of MS-AIG in prioritized therapeutic targets and therapeutic module construction. (*A*) Overlap between MS-AIG and Priority Index genes for MS (Fisher's exact test, *P*-value after adjusting using the Benjamini–Hochberg method, OR: odds ratio). (*B*) The chart shows the enrichment of MS-AIG in the top 150 Priority index positions for MS by interactome of origin. The *x* axis shows the odds ratio, the y axis shows the −log(*P* value), and dot size is proportional to the number of items. (*C*) The table includes the prioritized genes by interactome of origin. (*D*) Illustration of the therapeutic module, made of prioritized MS-AIG and PI crosstalk genes for MS. As shown in the legend, the EBV interactome genes are displayed in green, while those related to the CD40 pathway are highlighted with an orange contour. Edge intensity is proportional to the PPI score.

that the EBNA2 type 1 occupancy patterns may be of specific relevance in MS.

Humans are the only natural hosts for EBV, so it has been difficult to fully reproduce the sophisticated interaction between this virus and the human immune system. This is particularly true for studies, including the present one, that explored the function of a pervasive transcriptional regulator such as EBNA2. For this reason, we were limited in the in vivo exploration of the

pathophysiology of EBNA2 variability in the context of an animal model of MS. More work will be needed to develop new animal models or generate new knowledge in clinical trials with antivirals. In this context, the strong effects of Tenofovir prodrugs on EBV lytic DNA replication (49), along with recent case reports suggesting some clinical efficacy in MS (50), may represent a concrete opportunity.

It will be essential to verify whether other examples of convergence between viral and human genetic variability hold for other aspects of MS pathogenesis to reach "peak genetic insights" (51) and pinpoint new and etiology-based therapeutic targets. For this purpose, future EBV GWAS in MS may complement the information from human GWAS and large commensal gut microbiome studies (52).

Materials and Methods

The methodology used in this study is described in more detail in *SI Appendix*.

A flowchart outlining the study design is available in supplementary figures (*SI Appendix,* Fig. S9).

Candidate Interactome Analysis. The "candidate interactome" approach was applied as previously described (53). In brief, we obtained 20 candidate interactomes from the literature. Six were manually curated (EBV, HBV, CMV, HHV8, JCV, and Inflammasome); seven were obtained from databases of molecular interactions: AIRE, VDR, AHR, SIRT1, SIRT7 from BIOGRID (http://thebiogrid.org) (54), Polyomavirus from VirusMentha database (55), and Human–microRNA targets from miRecords database (56); seven were obtained from published high-throughput experimental approaches: VIRORF (57), HIV (58), HCV (59), h-IFN (60), H1N1 (61), HDAC (62), and *Chlamydia* (63) (Fig. 1*A*). Further details are provided in *SI Appendix*.

Protein Binding Enrichment. Genomic regions in Fig. 1*D* were identified using Variant Effect Predictor version 98. After the ALIGATOR analysis, we used the RegulomeDB database (20), that is fed with live data from ENCODE and Roadmap Epigenomics, to look for proteins binding in regions containing "postmatch" SNPs that displayed nominally significant associations with MS. When SNPs are bound by two or more proteins a Fisher's exact test (2 × 2 table), controlling q-values for multiple testing (FDR), Benjamini–Hochberg, was computed to evaluate whether this co-occurrence is statistically relevant using Python v3.6.0.

Colocalization Analysis. The 10,788 nominally significant SNPs derived from the ALIGATOR analysis on MS-associated interactomes and were used as region of interest (ROI), by defining genomic intervals centered on the SNPs and extended of 20kb upstream and downstream. EBNA2 binding sites (n = 17,019) were extracted from a ChIP-Seq performed on GM12878 EBV-transformed LCLs by Hong et al. (25). CTCF binding sites (N = 62,580) and POLR2A binding sites (N = 29,824) were extracted from ChIP-Seq data in Unibind (64), both performed on the same cell line (GM12878 EBV-transformed LCLs) to maximize consistency and avoid possible bias deriving from cell-specific TF activities. To provide a "threshold" against which the matches ROI<>EBNA2/CTCF/POLR2A would be benchmarked, we used as a statistical background ("Universe") 13,202,334 regions from ENCODE database of Transcription Factors Binding Sites (65) downloaded on 31/1/2020.

The colocalization analysis was performed through LOLA (66).

Postmortem Study. By using immunohistochemistry/immunofluorescence, the presence, extent, and distribution of potential cells expressing EBNA2 were examined in 4% formaldehyde-fixed frozen brain tissues obtained from postmortem cases with pathologically confirmed multiple sclerosis (N = 10). All postmortem samples were obtained from the UK Multiple Sclerosis Society Tissue Bank (UKMSTB) at Imperial College under ethical approval (08/MRE09/31). The diagnosis of multiple sclerosis was confirmed by neuropathology according to the International Classification of Diseases of the Nervous System criteria (www.ICDNS.org). Further details are provided in *SI Appendix*.

Frequency of Interactome Genes in MS Transcriptomes. The human PBMC microarray datasets analyzed in this study were published in Srinivasan papers (26, 67) and deposited at the EBI Array express database.

The human brain microarray datasets were published (27) and deposited at Gene Expression Omnibus. Further details are provided in *SI Appendix*.

Enrichment of EBV Interactors in MS-AIG and HSP-Protein Modules. HSP-protein modules were identified by Bis-Brewer (28) through a protein–protein interaction criteria using the Human Integrated Protein-Protein Interaction rEference (HIPPIE) database and the DIseAse Module Detection (DIAMOnD). We performed an overlap between MS-AIG (n = 741) and the Hereditary Spastic Paraplegia dominant-protein modules (n = 295). To verify the apparent enrichment of EBV interactors in HSP modules, we randomly extracted from MS-AIG, 100.000 samples of the same size of shared genes and counted the number of EBV interactors included in each random sample. We considered the shared genes as enriched for EBV interactors, if the *P* value of the observed number under the null distribution was <0.05. The analysis was performed using R version 1.1.456.

Network and Pathway Analysis. MetaCore© (version 6.29 build 68613; GeneGO, Thomson Reuters, New York) was employed to classify genes, carrying SNPs obtained from ALIGATOR analysis, in "canonical pathways" and evaluate their possible enrichment in specific categories. Canonical pathway maps represent a set of signaling and metabolic maps covering human biological processes in a comprehensive way. All maps are created by Thomson Reuters scientists by a high-quality manual curation process based on published peer-reviewed literature and validation experiments (68).

Lymphoblastoid Cell Line Generation and DNA Extraction. Sp-LCLs, carrying endogenous EBV, were generated from 13 subjects with relapsing remitting MS (69) and eight healthy donors matched for age, sex, and geographic origins (70). Further details are provided in *SI Appendix*.

Flow Cytometry Analysis. The phenotypic characterization of LCLs obtained from MS patients and HD was performed through multiparametric flow cytometric (MPFC) analysis. Further details are provided in *SI Appendix*.

EBNA2 Variant Identification. DNA samples were extracted from CD19+ B cells from peripheral blood of RRMS and age-and-sex-matched HD. The EBNA2 variant identification was performed using droplet digital PCR. Further details are provided in *SI Appendix*

CD40 and CD86 SNP Genotyping. We genotyped rs4810485 SNP in the CD40 gene and rs9282641 SNP in the CD86 gene in 26 spLCLs, using PCR amplification followed by Sanger sequencing. Further details are provided in *SI Appendix*.

Drug Target Prioritization of MS-AIG. Looking at the therapeutic potential of MS-AIG, we assessed the overlap of MS-AIG with the list of Priority Index (Pi) genes for Multiple Sclerosis, and their frequency in the top 150 positions of druggable targets ranking. Through the Pi pipeline, GWAS studies are leveraged to prioritize druggable genes based on their genomic proximity, functional relevance (physical interaction, expression modulation through eQTL), and association with rare diseases or immune phenotypes. Statistical analysis of the overlap was performed through the geneOverlap package in R, by using Fisher's exact test with Benjamini–Hochberg correction for multiple comparisons. Selected MS-AIG with high Pi rating were considered to build an integrated druggable network, merged with the "crosstalk genes" (operating across molecular pathways) from the Pi ranking for Multiple Sclerosis. Protein–protein interactions were derived from STRING (71): Interaction score cut-off was set at 0.70 (high confidence). The resultant network was imported in Cytoscape v.3.9.1 (72). Finally, the network-based drug repurposing analysis was performed in Cytoscape through the NedReX implementation of the "Closeness centrality" metrics for molecule ranking (73).

Statistical Analysis. The quantitation of protein levels was performed by FlowJo v10.5 and expressed as median fluorescence intensity. Unpaired two-tailed *t* test and one-way ANOVA with Tukey correction for multiple comparisons were used for the comparisons between groups. The correlations between protein levels were performed by a simple linear regression analysis. The enrichment of MS-AIG in DEG was measured by the chi-square test with Yates's correction. All the above analyses were performed using GraphPad Prism 8.

Fisher's exact test (2 × 2 table), controlling q-values for multiple testing (false discovery rate [FDR], Benjamini–Hochberg), was used to evaluate the statistical relevance of protein co-occurrence by using Python v3.6.0.

Linkage disequilibrium was considered as described in Sheffield and Bock (66). As in previous work (19), resulting −log(*P* value), support, and OR were combined into a single score, where the spacing parameter kp was set to 10.0 and we consider all three contributors equally, setting therefore weights wi to 1.0. Statistical significance was taken at $P < 0.05$.

$$\text{Harmonic Score} = \text{kp} \times \left(\sum \text{i wi} \right) / (\text{w1}/(-\log\text{P}) + \text{w2}/\text{Supp} + \text{w3}/\text{OR}).$$

The mixed-effect multiple linear regression analyses were performed using the R program, Version 1.1.456 (function "lmer" from "lme4" package). All the significances were considered when $P < 0.05$. Differences in proportions of EBNA2 allele distribution were compared using Fisher's exact tests (Graph Pad Prism 8).

Data, Materials, and Software Availability. All other data are included in the manuscript and/or supporting information. Previously published data were used for this work [The interactome previously analyzed in Mechelli et al. (52) (CMV, EBV, HBV, HHV8, JCV, AHR, AIRE, and VDR) and included in the present paper has been updated with new data. Gene expression data obtained from peripheral blood leukocytes and brain are available respectively at EBI Array express database [E-MTAB-4890 (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-4890); E-MTAB-5151 (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-5151), refs. 26 and 70] and Gene Expression Omnibus (GSE135511, ref. 27). GWAS data from MS are available at IMSGC (https://imsgc.net/); GWAS data from other diseases are available at WTCCC2. GWAS variants associated with SLE were pooled from SPOKE (https://spoke.cgl.ucsf.edu/; DOID:9074). EBNA2 binding sites were extracted from a ChIP-Seq performed on GM12878 EBV-transformed LCLs by Hong et al. (25). CTCF binding sites and POLR2A binding sites were extracted from ChIP-Seq data by Puig et al. (67).].

ACKNOWLEDGMENTS. The CRESM Biobank, AU San Luigi Gonzaga, Orbassano, Turin (Italy) provided part of the samples used in this study. This study makes use of data generated by the Wellcome Trust Case-control Consortium (WTCCC2). A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113, 085475, and 090355. This research was funded by Italian Multiple Sclerosis Foundation (FISM) and financed or cofinanced with the "5 per mille" public funding: "CENTERS" as Special Project to M.S., grant number 2014/R/12 and 2017/R/18 to R. Mechelli, grant number 2018/B/5 to M.P., grant number 2018/S/5 to L.B., grant number 2016/R/10 and 2018/R/4 to P.d.C., 2021/Special Multi/002 to M.S.; National Multiple Sclerosis Society (NMSS): RG-1807-32021 to RM; PP-1606-24687 to P.d.C; Italian Ministry of Health (Ricerca corrente) to R. Mechelli; RFA-2305-41332 to R. Magliozzi; RF-2019- 12371111 to G.M. Bando PNRR 2022 PNRR-MAD-2022-12375634 to G. Matarese. and L.B.; Sapienza University of Rome (Grandi Attrezzature) to M.S.; Horizon-HLTH-2023-DISEASE-3 (Project 101136991; EBV-MS) to M.S. and R.M.; Ministry of University and Research (MUR) (Bando PRIN 2022 Prot. 2022LNHZAP to G.M.) and MUR PNRR Extended Partnership (INF-ACT no. PE00000007 and MNESYS no. PE00000006 to G.M.); Juvenile Diabetes Research Foundation (JDRF; grant 1-SRA-2018-477-S-B) to P.d.C. P.T. is supported by a grant from Lega Italiana per la lotta contro Tumori (Lilt) and the Italian Ministry of Education and Research (Mur-PRIN 2022 grant). We are indebted to Prof. Francesco Cucca for his critical review and insightful comments on the manuscript. S.M. received honoraria for speaking and consulting from Biogen, Novartis, Merck and Roche; D.L. received travel funding from Biogen, Merck Serono, Sanofi-Genzyme and Teva, honoraria for speaking from Sanofi-Genzyme and Teva, and consultation fees from Merck Serono and Teva. She is a sub-investigator in clinical trials being conducted for Biogen, Merck Serono, Novartis, Roche and Teva; A.D.S. received honoraria for speaking and consulting by Biogen, Novartis, Roche, Sanofi, Merck, Alexion and Sandoz and has been reimbursed by Merck, Biogen, Sanofi, Novartis and Roche for attending several conferences; G.M. is an Advisory Board member of Biogen Idec, Genzyme, Merck-Serono, Novartis, Roche and received honoraria for speaking or consultation fees from Almirall, Bayer Schering, Biogen Idec, Merck Serono, Novartis, Sanofi-Genzyme, Roche, Mylan, Alexion and BMS. She is the principal investigator in clinical trials for Biogen Idec, Merck Serono, Novartis, Roche, Sanofi-Genzyme, Merck Serono and BMS; L.L. received honoraria for consultancy or speaking and travel grant from Biogen, Novartis, Sanofi Genzyme, Bristol, Merck and Roche; E.C. received honoraria for consultancy or speaking and travel grant from Biogen, Novartis, Sanofi Genzyme, Bristol, Merck and Roche; D.C. is an Advisory Board member of Almirall, Bayer Schering, Biogen, GW Pharmaceuticals, Merck Serono, Novartis, Roche, Sanofi-Genzyme, Teva and received honoraria for speaking or consultation fees from Almirall, Bayer Schering, Biogen, GW Pharmaceuticals, Merck Serono, Novartis, Roche, Sanofi-Genzyme, Teva. He is also the principal investigator in clinical trials for Bayer Schering, Biogen, Merck Serono, Mitsubishi, Novartis, Roche, Sanofi-Genzyme, Teva; R.P.U. is supported by the National MS Society; L.B. received speaking honoraria from Baxter, Merck, Novartis, Roche, Bristol, Janssen, Horizon, GSK and Sanofi; G.M. reports receiving advisory board fees from Merck, Biogen, Novartis, and Roche; R.R. received research support from the UK MS Society and has received speaker honoraria from Roche and Novartis; M.S. received consulting fees from Allergan, Biogen, BMS, Horizon, Merck-Serono, Novartis, Roche, Sanofi, Teva. D.C. preclinical and clinical research was supported by grants from Bayer Schering, Biogen Idec, Celgene, Merck Serono, Novartis, Roche, Sanofi- Genzyme e Teva; C.F. received research support from Merck-Serono, Teva, Novartis; L.B. received research support from Merck, Biogen, Novartis, and Roche; R.R. received research support from Medimmune; M.S. Allergan, Biogen, BMS, Horizon, Merck-Serono, Novartis, Roche, Sanofi, Teva.

Author affiliations: [a]Department for the Promotion of Human Sciences and Quality of Life, San Raffaele Roma University, Rome, Italy; [b]Istituto Ricovero e Cura a Carattere Scientifico San Raffaele, Rome 00166, Italy; [c]Weill Cornell Medicine, New York, NY 10021; [d]Massachusetts Institute of Technology, Cambridge, MA 02139; [e]Harvard T. H. Chan School of Public Health, Boston, MA 02115; [f]Centre for Experimental Neurological Therapies, Department of Neurosciences, Mental Health and Sensory Organs, Sapienza University, Rome 00189, Italy; [g]Neuroimmunology Unit, Istituto Ricovero e Cura a Carattere Scientifico Fondazione Santa Lucia, Rome 00179, Italy; [h]Biocrystallography Unit, Division of Immunology, Transplantation, and Infectious Diseases, Istituto Ricovero e Cura a Carattere Scientifico San Raffaele Scientific Institute, Milan 20132, Italy; [i]Institute of Experimental Neurology & Division of Neuroscience, Istituto Ricovero e Cura a Carattere Scientifico San Raffaele Scientific Institute, Milan 20132, Italy; [j]Department of Neurology, University of California San Francisco Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA 94158; [k]Clinical Neurobiology Unit, Neuroscience Institute Cavalieri Ottolenghi, Orbassano 10043, Italy; [l]Centri di Riferimento Regionale Sclerosi Multipla Biobank, University Hospital San Luigi Gonzaga, Orbassano 10043, Italy; [m]Department of Neuroscience "Rita Levi Montalcini", University of Turin, Turin 10043, Italy; [n]Department of Neurology and Centri di Riferimento Regionale Sclerosi Multipla, University Hospital San Luigi Gonzaga, Orbassano 10043, Italy; [o]Multiple Sclerosis Clinical and Research Unit, Department of Systems Medicine, Tor Vergata University, Rome 00133, Italy; [p]Multiple Sclerosis Center, Binaghi Hospital, Azienda Sanitaria Locale Cagliari, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari 09126, Italy; [q]Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01655; [r]Department of Neurology, Massachusetts General Hospital, Boston, MA 02114; [s]Harvard Medical School, Boston, MA 02115; [t]Department of Clinical and Molecular Medicine, Sapienza University, Rome 00189, Italy; [u]Department of Experimental Medicine, Sapienza University, Rome 00161, Italy; [v]The International Multiple Sclerosis Genetics Consortium and the Wellcome Trust Case Control Consortium 2, Oxford OX1 2JD, United Kingdom; [w]Department of Systems Medicine, Tor Vergata University, Rome 00133, Italy; [x]Istituto Ricovero e Cura a Carattere Scientifico Istituto Neurologico Mediterraneo Neuromed, Pozzilli-86077, Italy; [y]Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health and Centre of Excellence for Biomedical Research, University of Genova, Genoa 16132, Italy; [z]Ospedale Policlinico San Martino, Istituto Ricovero e Cura a Carattere Scientifico, Genoa 16132, Italy; [aa]Elixir Proteomics Laboratory, Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Segrate, Milan 20054, Italy; [bb]Treg Cell Lab, Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università di Napoli "Federico II", Napoli 80131, Italy; [cc]Department of Health Sciences, University of Eastern Piedmont, Interdisciplinary Research Center of Autoimmune Diseases, Novara 28100, Italy; [dd]Neurology B, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona 37134, Italy; [ee]Department of Brain Sciences, Faculty of Medicine, Imperial College London, London W12 ONN, United Kingdom; and [ff]Laboratorio di Immunologia, Istituto di Endocrinologia e Oncologia Sperimentale, Consiglio Nazionale delle Ricerche, Naples 80131, Italy

Author contributions: R. Mechelli, R.U., G.B., S.E.B., G. Matarese., M.S., and G.R. designed research; R. Mechelli, R.U., G.B., R.B., V.R., D.F.A., G.G., F.C.P., S.I., M.P., S.D., R. Magliozzi, and M.S. performed research; R.B., D.F.A., G.G., F.C.P., M.P., S.R., M.C.B., S. Martire, S. Malucchi, D.L., L.L., E.A., P.T., A.F., M.F., I.W., A.D.S., G. Marfia, E.C., D.C., D.D.S., P.M., L.B., C.F., and R.R. contributed new reagents/analytic tools; R. Mechelli, R.U., G.B., V.R., S.S., G.C., R.P.U., A.U., P.d.C, S.D., C.F., R. Magliozzi, S.E.B., G. Matarese, and M.S. analyzed data; and R. Mechelli, R.U., G.B., R. Magliozzi, G. Matarese, M.S., and G.R. wrote the paper.

Competing interest statement: M.S. and G.R. hold a patent on EBNA2 alleles in multiple sclerosis (IT1417523 EP2981625).

1. B. M. Jacobs *et al.*, Towards a global view of multiple sclerosis genetics. *Nat. Rev. Neurol.* **18**, 613–623 (2022).
2. K. Bjornevik *et al.*, Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* **375**, 296–301 (2022).
3. F. Aloisi, G. Giovannoni, M. Salvetti, Epstein-Barr virus as a cause of multiple sclerosis: Opportunities for prevention and therapy. *Lancet Neurol.* **22**, 338–349 (2023).
4. M. Cortese *et al.*, Serologic response to the Epstein Barr virus peptidome and the risk of multiple sclerosis. *JAMA Neurol.* **81**, 515–524 (2024).

5. K. Tengvall *et al.*, Molecular mimicry between Anoctamin 2 and Epstein–Barr virus nuclear antigen 1 associates with multiple sclerosis risk. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16955–16960 (2019).
6. T. V. Lanz *et al.*, Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature* **603**, 321–327 (2022).
7. O. G. Thomas *et al.*, Cross-reactive EBNA1 immunity targets alpha-crystallin B and is associated with multiple sclerosis. *Sci. Adv.* **9**, eadg3032 (2023).
8. H. Vietzen *et al.*, Ineffective control of Epstein–Barr-virus-induced autoimmunity increases the risk for multiple sclerosis. *Cell* **186**, 5705–5718.e13 (2023).
9. B. Damania, S. C. Kenney, N. Raab-Traub, Epstein–Barr virus: Biology and clinical disease. *Cell* **185**, 3652–3670 (2022).
10. A. Goris, M. Vandebergh, J. L. McCauley, J. Saarela, C. Cotsapas, Genetics of multiple sclerosis: Lessons from polygenicity. *Lancet Neurol.* **21**, 830–842 (2022).
11. J. Huang *et al.*, Genetics of immune response to Epstein–Barr virus: Prospects for multiple sclerosis pathogenesis. *Brain* **147**, 3573–3582 (2024).
12. S. Sawcer *et al.*, Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219 (2011).
13. A. H. Beecham *et al.*, Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **45**, 1353–1360 (2013).
14. Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
15. International Multiple Sclerosis Genetics Consortium, Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, eaav7188 (2019).
16. E. Mountjoy *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).
17. K. K. Farh *et al.*, Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
18. M. K. Freund *et al.*, Phenotype-specific enrichment of mendelian disorder genes near GWAS regions across 62 complex traits. *Am. J. Hum. Genet.* **103**, 535–552 (2018).
19. R. Umeton *et al.*, Multiple sclerosis genetic and non-genetic factors interact through the transient transcriptome. *Sci. Rep.* **12**, 7536 (2022).
20. A. P. Boyle *et al.*, Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).
21. Z. Tang *et al.*, CTCF-mediated human 3d genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627 (2015).
22. J. T. Keane *et al.*, The interaction of Epstein–Barr virus encoded transcription factor EBNA2 with multiple sclerosis risk loci is dependent on the risk genotype. *EBioMedicine* **71**, 103572 (2021).
23. V. A. Ricigliano *et al.*, EBNA2 binds to genomic intervals associated with multiple sclerosis and overlaps with vitamin D receptor occupancy. *PLoS One* **10**, e0119605 (2015).
24. J. B. Harley *et al.*, Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat. Genet.* **50**, 699–707 (2018).
25. T. Hong *et al.*, Epstein–Barr virus nuclear antigen 2 extensively rewires the human chromatin landscape at autoimmune risk loci. *Genome Res.* **31**, 2185–2198 (2021).
26. S. Srinivasan *et al.*, Transcriptional dysregulation of interferome in experimental and human multiple sclerosis. *Sci. Rep.* **7**, 8981 (2017).
27. R. Magliozzi *et al.*, Meningeal inflammation changes the balance of TNF signalling in cortical grey matter in multiple sclerosis. *J Neuroinflammation* **16**, 259 (2019).
28. D. M. Bis-Brewer, M. C. Danzi, S. Wuchty, S. Züchner, A network biology approach to unraveling inherited axonopathies. *Sci. Rep.* **9**, 1692 (2019).
29. X. Jia *et al.*, Genome sequencing uncovers phenocopies in primary progressive multiple sclerosis. *Ann. Neurol.* **84**, 51–63 (2018).
30. E. Morandi, S. A. Jagessar, B. A. 't Hart, B. Gran, EBV infection empowers human B cells for autoimmunity: Role of autophagy and relevance to multiple sclerosis. *J. Immunol.* **199**, 435–448 (2017).
31. I. Smets *et al.*, Multiple sclerosis risk variants alter expression of co-stimulatory genes in B cells. *Brain.* **141**, 786–796 (2018).
32. J. T. Keane *et al.*, Gender and the sex hormone estradiol affect multiple sclerosis risk gene expression in Epstein–Barr virus-infected B cells. *Front. Immunol.* **12**, 732694 (2021).
33. R. Mechelli *et al.*, Epstein-Barr virus genetic variants are associated with multiple sclerosis. *Neurology* **84**, 1362–1368 (2015).
34. H. Fang, J. C. Knight, Priority index: Database of genetic targets in immune-mediated disease. *Nucleic Acids Res.* **50**, D1358–D1367 (2022).
35. J. Correale, BTK inhibitors as potential therapies for multiple sclerosis. *Lancet Neurol.* **20**, 689–691 (2021).
36. W. H. Robinson, S. Younis, Z. Z. Love, L. Steinman, T. V. Lanz, Epstein Barr virus as a potentiator of autoimmune diseases. *Nat. Rev. Rheumatol.* **20**, 729–740 (2024).
37. S. Thakolwiboon *et al.*, Regional differences in the association of cytomegalovirus seropositivity and multiple sclerosis: A systematic review and meta-analysis. *Mult. Scler. Relat. Disord.* **45**, 102393 (2020).
38. N. S. Rasmussen, A. H. Draborg, C. T. Nielsen, S. Jacobsen, G. Houen, Antibodies to early EBV, CMV, and HHV6 antigens in systemic lupus erythematosus patients. *Scand. J. Rheumatol.* **44**, 143–149 (2015).
39. D. H. Margulies, J. Jiang, J. Ahmad, L. F. Boyd, K. Natarajan, Chaperone function in antigen presentation by MHC class I molecules-tapasin in the PLC and TAPBPR beyond. *Front. Immunol.* **14**, 1179846 (2023), 10.3389/fimmu.2023.1179846.
40. M. Kloss *et al.*, Interaction of monocytes with NK cells upon Toll-like receptor-induced expression of the NKG2D ligand MICA. *J. Immunol.* **181**, 6711–6719 (2008), 10.4049/jimmunol.181.10.6711.
41. R. Yazdani *et al.*, The hyper IgM syndromes: Epidemiology, pathogenesis, clinical manifestations, diagnosis and management. *Clin. Immunol.* **198**, 19–30 (2019).
42. T. Yamano *et al.*, Thymic B cells are licensed to present self antigens for central T cell tolerance induction. *Immunity.* **42**, 1048–1061 (2015).
43. A. M. Afzali *et al.*, B cells orchestrate tolerance to the neuromyelitis optica autoantigen AQP4. *Nature.* **627**, 407–415 (2024).
44. A. Afrasiabi *et al.*, Evidence from genome wide association studies implicates reduced control of Epstein–Barr virus infection in multiple sclerosis susceptibility. *Genome Med.* **11**, 26 (2019).
45. A. Afrasiabi *et al.*, Genetic and transcriptomic analyses support a switch to lytic phase in Epstein Barr virus infection as an important driver in developing Systemic Lupus Erythematosus. *J. Autoimmun.* **127**, 102781 (2022).
46. I. Jelcic *et al.*, Memory B cells activate brain-homing, autoreactive $CD4^+$T cells in multiple sclerosis. *Cell* **175**, 85–100.e23 (2018), 10.1016/j.cell.2018.08.011.
47. P. Vermersch *et al.*; Frexalimab Phase 2 Trial Group, Inhibition of CD40L with Frexalimab in multiple sclerosis. *N. Engl. J. Med.* **390**, 589–600 (2024).
48. K. C. M. F. Viel *et al.*, Shared and distinct interactions of type 1 and type 2 Epstein–Barr Nuclear Antigen 2 with the human genome. *BMC Genomics.* **65**, 273 (2024).
49. N. C. Drosu, E. R. Edelman, D. E. Housman, Tenofovir prodrugs potently inhibit Epstein–Barr virus lytic DNA replication by targeting the viral DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 12368–12374 (2020).
50. Ø. Torkildsen, K. M. Myhr, P. Brugger-Synnes, K. Bjørnevik, Antiviral therapy with tenofovir in MS. *Mult. Scler. Relat. Disord.* **83**, 105436 (2024), 10.1016/j.msard.2024.105436.
51. E. V. Minikel, J. L. Painter, C. C. Dong, M. R. Nelson, Refining the impact of genetic evidence on clinical success. *Nature* **629**, 624–629 (2024).
52. IMSMS Consortium, Gut microbiome of multiple sclerosis patients and paired household healthy controls reveal associations with disease risk and course. *Cell* **185**, 3467–3486.e16 (2022).
53. R. Mechelli *et al.*, A "Candidate-interactome" Aggregate analysis of genome-wide association data in multiple sclerosis. *PLoS One* **8**, e63300 (2013).
54. A. Chatr-Aryamontri *et al.*, The Biogrid interaction database: 2015 update. *Nucleic Acids Res.* **43**, D470–D478 (2015).
55. A. Calderone, L. Licata, G. Cesareni, Virusmentha: A new resource for virus–host protein interactions. *Nucleic Acids Res.* **43**, D588–D592 (2015).
56. F. Xiao *et al.*, Mirecords: An integrated resource for microRNA–target interactions. *Nucleic Acids Res.* **37**, D105–D110 (2009).
57. A. Pichlmair *et al.*, Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* **487**, 486–490 (2012).
58. S. Jäger *et al.*, Global landscape of HIV–human protein complexes. *Nature* **481**, 365–370 (2011).
59. B. de Chassey *et al.*, Hepatitis C virus infection protein network. *Mol. Syst. Biol.* **4**, 230 (2008).
60. S. Li, L. Wang, M. Berman, Y. Y. Kong, M. E. Dorf, Mapping a dynamic innate immunity protein interaction network regulating type I interferon production. *Immunity* **35**, 426–440 (2011).
61. S. D. Shapira *et al.*, A physical and regulatory map of host–influenza interactions reveals pathways in H1N1 infection. *Cell* **139**, 1255–1267 (2009).
62. P. Joshi *et al.*, The functional interactome landscape of the human histone deacetylase family. *Mol. Syst. Biol.* **9**, 672 (2013).
63. K. M. Mirrashidi *et al.*, Global mapping of the inc-human interactome reveals that retromer restricts Chlamydia infection. *Cell Host Microbe* **18**, 109–121 (2015).
64. R. R. Puig, P. Boddie, A. Khan, J. A. Castro-Mondragon, A. Mathelier, Unibind: Maps of high-confidence direct tf-dna interactions across nine species. *BMC Genomics* **22**, 482 (2021).
65. C. A. Sloan *et al.*, Encode data at the encode portal. *Nucleic Acids Res.* **44**, D726–D732 (2016).
66. N. C. Sheffield, C. Bock, Lola: Enrichment analysis for genomic region sets and regulatory elements in r and bioconductor. *Bioinformatics* **32**, 587–589 (2016).
67. S. Srinivasan *et al.*, Dysregulation of MS risk genes and pathways at distinct stages of disease. *Neurol. Neuroimmunol. Neuroinflamm.* **4**, e337 (2017).
68. D. M. Bolser *et al.*, Metabase–The wiki-database of biological databases. *Nucleic Acids Res.* **40**, D1250–D1254 (2012).
69. A. J. Thompson *et al.*, Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2018).
70. M. Chiara *et al.*, Geographic population structure in Epstein–Barr virus revealed by comparative genomics. *Genome Biol. Evol.* **8**, 3284–3291 (2016).
71. D. Szklarczyk *et al.*, String v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
72. P. Shannon *et al.*, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
73. S. Sadegh *et al.*, Network medicine for disease module identification and drug repurposing with the NeDRex platform. *Nat. Commun.* **12**, 6848 (2021).