

Enrichment analysis of GWAS data in autoimmunity delineates the multiple sclerosis-Epstein Barr virus association

Rosella Mechelli,^{1,23*} Renato Umeton,^{2,3,4,23} Virginia Rinaldi,⁵ Gianmarco Bellucci,⁵ Rachele Bigi,⁵ Daniela F. Angelini,⁶ Gisella Guerrera,⁶ Sundararajan Srinivasan,⁷ Silvia Romano,⁵ Maria Chiara Buscarinu,⁵ Raffaella Pizzolato Umeton,^{8,9,10} Eleni Anastasiadou,¹¹ Pankaj Trivedi,¹¹ Arianna Fornasiero,⁵ Michela Ferraldeschi,⁵ IMSCG and WTCCC2,¹² Diego Centonze,^{13,14} Antonio Uccelli,¹⁵ Dario Di Silvestre,¹⁶ Pier Luigi Mauri,¹⁶ Paola de Candia,¹⁷ Giuseppe Matarese,^{18,19} Sandra D'Alfonso,²⁰ Luca Battistini,⁶ Cinthia Farina,⁷ Roberta Magliozzi,^{21,22} Richard Reynolds,²² Marco Salvetti,^{5,13, 24,*} Giovanni Ristori^{5,*}.

¹San Raffaele Roma Open University and IRCCS San Raffaele Pisana, Rome, Italy

²Department of Informatics and Analytics, Dana-Farber Cancer Institute, Boston, MA, USA

³Massachusetts Institute of Technology, Cambridge, MA, USA

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

⁵Centre for Experimental Neurological Therapies (CENTERS), Department of Neurosciences, Mental Health and Sensory Organs, Sapienza University of Rome, Italy

⁶Neuroimmunology Unit, IRCCS Santa Lucia Foundation

⁷Institute of Experimental Neurology & Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy

⁸Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA

⁹Department of Neurology, Massachusetts General Hospital Boston, MA, USA

¹⁰Harvard Medical School, Boston, MA, USA

¹¹Department of Experimental Medicine, Sapienza University of Rome, Rome, Italy.

¹²The International Multiple Sclerosis Genetics Consortium & the Wellcome Trust Case Control Consortium 2, Oxford, UK

¹³IRCCS Istituto Neurologico Mediterraneo Neuromed, Pozzilli, Italy

¹⁴Clinica Neurologica Dipartimento di Medicina dei Sistemi, University of Tor Vergata, Rome, Italy

¹⁵Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health and Centre of Excellence for Biomedical Research, University of Genova, Genoa, Italy

¹⁶Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche (ITB-CNR), 20090 Segrate, Milano, Italy.

¹⁷IRCCS MultiMedica, 20138 Milano, Italy

¹⁸Laboratorio di Immunologia, Istituto di Endocrinologia e Oncologia Sperimentale, Consiglio Nazionale delle Ricerche (IEOS- CNR), Naples, Italy.

¹⁹Treg Cell Lab, Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università di Napoli "Federico II" 80131 Napoli, Italy

²⁰Department of Health Sciences, University of Eastern Piedmont, Novara, Italy

²¹Neurology B, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy.

²²Division of Brain Sciences, Department of Medicine, Imperial College London, London, United Kingdom.

²³These authors contributed equally to this work as co-first authors

²⁴**Lead contact**

***Correspondence:**

Rosella Mechelli, San Raffaele Roma Open University and IRCCS San Raffaele Pisana, Rome, Italy

Via di Val Cannuta, 247

00166 Rome (Italy)

Tel: 00390652252257

e-mail: rosella.mechelli@uniroma5.it

Giovanni Ristori, Centre for Experimental Neurological Therapies (CENTERS), Department of Neurosciences, Mental Health and Sensory Organs, Sapienza University, Rome

Via di Grottarossa, 1035-1039

00189 Rome (Italy)

Tel: 00390633776044; Fax: 0633775900

e-mail: giovanni.ristori@uniroma1.it

Marco Salvetti, Centre for Experimental Neurological Therapies (CENTERS), Department of Neurosciences, Mental Health and Sensory Organs, Sapienza University, Rome

Via di Grottarossa, 1035-1039

00189 Rome (Italy)

Tel: 00390633775994; Fax: 0633775900

e-mail: marco.salvetti@uniroma1.it

SUMMARY

We exploited genetic information to assess non-genetic influences in autoimmunity. We isolated gene modules whose products physically interact with environmental exposures related to autoimmunity, and analyzed their nominal statistical evidence of association with autoimmune and non-autoimmune diseases in genome-wide association studies (GWAS) data. Epstein Barr virus (EBV) and other Herpesviruses interactomes emerged as specifically associated with multiple sclerosis (MS), possibly under common regulatory mechanisms. Analyses of MS blood and brain transcriptomes, cytofluorimetric studies of endogenous EBV-infected lymphoblastoid lines, and lesion immunohistochemistry, confirmed a dysregulation of MS-associated EBV interactors, suggesting their contribution to CD40 signaling alterations in MS. These interactors resulted enriched in modules from inherited axonopathies-causing genes, supporting a link between EBV and neurodegeneration in MS, in accord with the observed transcriptomic dysregulations in MS brains. They were also enriched with top-ranked pharmaceutical targets prioritized on a genetic basis. This study delineates a disease-specific influence of herpesviruses on MS biology.

Keywords:

Genes, Epstein-Barr virus, herpesvirus, interactome, multiple sclerosis, GWAS, gene expression, autoimmune disease, lymphoblastoid cell lines, CNS

INTRODUCTION

By definition, multifactorial diseases are thought to be caused by heritable and nonheritable factors. Genome-wide association studies (GWAS) provide information about the former and sero-epidemiological studies do the same for the latter. Both types of studies are subject to interpretation difficulties. One limitation of GWAS derives from the fragmentation of the information they provide, i.e. lists of single variables, mostly of small effect size, whose putative functions are difficult to reunite and reorganize in one or few disease processes (Boyle et al. 2017; Califano et al. 2012; Li et al. 2019). For seroepidemiological studies, the limitation comes from the correlative and associative nature of most of the investigations on environmental exposures (Fischbach 2018). However, GWAS can inform about the causal relevance of associations between environmental exposures and disease. Reciprocally, extracting value from the study of the interaction between relevant exposures and predisposing genes may contribute to the interpretation of GWAS, in a virtuous 'genes-to-environment-to-genes-again' iterative learning process (Tam et al. 2019; Visscher et al. 2017).

Previous work has shown the potential relevance of GWAS data analyses that focus on the combined effects of many loci of small effect-size (Beecham et al. 2013; Purcell et al. 2009; Saez-Atienzar et al. 2021). Among different efforts to study biologically meaningful combinations of genes, we devised a "candidate interactome" approach to investigate which gene-environment interactions may associate with the development of multiple sclerosis (MS), a multifactorial disease of the central nervous system (CNS) that affects young adults, resulting in significant disability and socioeconomic burden for both affected persons and society. In GWAS data, we looked for statistical enrichments of associations among modules of genes whose products interact with environmental factors of plausible, uncertain, or unlikely relevance for MS pathogenesis ("candidate interactomes") (Mechelli et al. 2013). This approach can be considered akin to Mendelian randomization, which exploits genetic data to improve causal inference in epidemiology (Davey Smith and Hemani 2014). The results provided a preliminary suggestion about the possible relevance of the interaction between host genotype and Epstein Barr virus (EBV) in MS. However, this study did not investigate the disease-specificity and the biological implications of the observed associations.

Integrating GWAS information with transcriptional regulatory data and protein interactions provides plausible models of disease pathogenesis (Baird et al. 2021; Consortium 2019a). Elegant studies are resolving the functional implications of individual disease-associated variants (Gregory et al. 2012; Steri et al. 2017). However, given the complex nature of multifactorial diseases (Wray et al. 2018), and particularly for translation into clinical benefit (Zeggini et al. 2019), a general understanding of how gene-environment interactions generate gene expression alterations is also necessary (Gallagher and Chen-Plotkin 2018).

Finally, it is known that drug targets having genetic associations with a disease significantly increase the probability of success in drug development (Nelson et al. 2015). In our approach, since the genetic associations reflect environmental exposures, the therapeutic relevance may be increased by the opportunity of targeting not only the genetic component but also the exposure. Moreover, in many autoimmune diseases, where the therapeutic scenario is rapidly changing, it is important to understand which targets may be of etiologic relevance in order to design an appropriate therapeutic sequencing.

Taking advantage of the increasing coverage and accuracy of protein-protein interaction data (Menche et al. 2015), and of the availability of additional MS-GWAS data, in this study we replicated and refined the previous findings (Mechelli et al. 2013), extending the "candidate interactome" analysis to GWAS in other multifactorial conditions. This also allowed the assessment of the disease-specificity of the results. We also investigated the functional implications of the observed associations and the distribution of the therapeutic validity (Fang et al. 2019) of the genes

highlighted by the “candidate interactome” approach. Overall, this study outlined an influence of EBV and other Herpesviruses on the biology of MS, which differentiates this disease from other immune-mediated disorders. This influence involves immunological aspects but may also relate to the neurodegenerative component of the disease.

RESULTS

Selection and construction of interactome modules

We isolated modules of genes whose products are known to physically interact (“interactomes”) with environmental exposures or in biological processes of plausible, uncertain, or unlikely relevance for the following autoimmune and non-autoimmune complex disorders: MS, type 1 and type 2 diabetes, rheumatoid arthritis, Crohn disease, celiac disease, bipolar disorder, hypertension, coronary artery disease. We considered 20 interactomes (Figure 1A) (10 related to environmental exposures and 10 to biological processes): interactomes of 9 viruses [Epstein-Barr virus (EBV), Hepatitis B virus (HBV), Cytomegalovirus (CMV), Human Herpes virus 8 (HHV8), JC virus (JCV), Human immunodeficiency virus (HIV), Hepatitis C virus (HCV), H1N1-influenza (H1N1), Polyomavirus], 1 bacterium (Chlamydia), 9 protein/process hubs [Autoimmune regulator (AIRE), Vitamin D receptor (VDR), Aryl hydrocarbon receptor (AHR), proteins targeted by 70 innate immune-modulating viral open reading frames from 30 viral species (VIORF), human innate immunity interactome for type I interferon (h-IFN), histone deacetylases (HDAC), Sirtuin 1 (SIRT1), Sirtuin 7 (SIRT7), Inflammasome] and 1 Human-microRNA targets (human miRNA-mRNA) network. The manually curated interactomes (CMV, EBV, HBV, HHV8 and JCV) and those obtained from publicly available databases (AHR, AIRE and VDR), that we had already selected in our previous work (Mechelli et al. 2013), were updated. We constructed the Inflammasome interactome by analyzing the literature and selecting all the physical-direct interactions among inflammasome proteins and other cellular proteins. The HDACs interactome was obtained from a unique source where different methods had been used to highlight protein-protein interactions among the 11 known HDAC enzymes and other cellular proteins (Joshi et al. 2013). We considered also other deacetylase enzymes belonging to the sirtuin family (SIRT1 and SIRT7), whose interactomes were obtained from BioGRID database (Stark et al. 2006). For the human miRNA-mRNA interactome we used miRecords, taking into consideration only experimentally verified miRNA human targets (Xiao et al. 2009). The Chlamydia interactome was obtained from Mirrashidi et al., that used affinity purification and mass spectrometry approaches to highlight interactions among bacterial and cellular proteins (Mirrashidi et al. 2015). Finally, we used VirusMentha, a protein-protein interaction database specific for virus-host interactions (Calderone et al. 2015), to obtain Polyomavirus interactome.

The MS-GWAS data came from the data set published in 2011 (Sawcer et al. 2011) and from ImmunoChip data (Beecham et al. 2013). To globally 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 (Figure S1). The GWAS data in the other diseases were from the Wellcome Trust Case Control Consortium 2 (Consortium 2007).

From GWAS datasets, we selected significant SNPs associated with MS and with the other complex diseases, 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; see workflow in Figure 1B). Spearman’s coefficients were calculated to evaluate linear correlation between the size of single interactomes and their cumulative p-value of association with the diseases, but no correlation was found in any case, indicating that the size of interactomes does not influence discovery (Figure S2-3).

Herpesvirus interacting proteins are enriched in MS risk genes

When analyzing the frequency of disease-associated genes from distinct GWAS studies in the 20 interactomes, viral interactomes were consistently associated with MS (Figure 1 C). 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-related loci shared by autoimmune disorders), our result may suggest that herpesviruses do not associate to MS through genetic traits common to other autoimmune diseases. In fact, the herpes virus association characterized MS with respect to all other conditions, including immune-mediated disorders like rheumatoid arthritis, Crohn disease and celiac disease (Figure 1C). Significant associations were sparse in non-MS conditions, and herpesvirus interactomes reached significance only in coronary artery disease (CMV) and type 2 diabetes (HHV8). Among other, non-herpes, viral interactomes, the one of HIV was MS-associated in the GWAS 2011, METACHIP1 and METACHIP2 but not in GWAS 2013. Other viral interactomes showing associations with MS were HBV (GWAS 2011), polyomavirus (GWAS 2013) and VIRORF (GWAS 2013). Other occasional associations with MS in the non-viral interactomes were AHR, Chlamydia and SIRT7 in the GWAS 2013 and AIRE in the GWAS 2011. Taken together, these results point to a disease-specific influence of EBV and, possibly, other herpesviruses on the development of MS.

Common regulatory mechanisms exist among all the MS-associated interactomes

To better understand the overall picture that may emerge from the above results we evaluated the possibility that common regulatory mechanisms exist among all the MS-associated interactomes. To verify this hypothesis, we considered all the SNPs and genes that were responsible for the statistical enrichments of association of interactomes with MS. Specifically, we obtained n=10,788 nominally significant SNPs resulting from ALIGATOR analysis and the corresponding 741 MS-associated interactome genes (MS-AIG) (Figure 2A, Table S1). We mapped the SNPs and related MS-AIG in the human genome, showing a wide distribution on autosomal chromosomes (Figure 2B) and a prevalent localization in intronic and upstream/downstream regions (Figure 2C). This result is in line with the concept that genetic variants associated to multifactorial diseases affect the regulation of gene expression (Farh et al. 2015; Freund et al. 2018), reinforcing the plausibility of our observations. We then used RegulomeDB database (Boyle et al. 2012) to look for proteins whose binding close to SNPs extracted from ALIGATOR analysis is enriched. We overlapped the SNPs with the protein binding map generated by Encyclopedia of DNA Elements (ENCODE) and RoadMap Epigenomics (Boyle et al. 2012), obtaining a list of enriched proteins. RNA Polymerase II Subunit A (POLR2A) and CCCTC-Binding Factor (CTCF) displayed the highest regulatory potential (Figure 2D). Since each SNP can be bound by multiple proteins, we carried out an analysis of "co-occurrence" for POLR2A and CTCF evaluated 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 (Tang et al. 2015).

Dysregulations of MS-associated EBV interactors in blood and CNS transcriptomes

To evaluate the functional implications of the above results, we matched the 741 MS-AIG with microarray gene expression data obtained from peripheral blood leucocytes of 40 healthy subjects and persons with different MS phenotypes, including clinically isolated syndrome (CIS, n=46), relapsing-remitting (RR, n=52), secondary progressive (SP, n=21) and primary progressive (PP, n=23) MS (Srinivasan et al. 2017b). We calculated whether the MS-AIG were significantly enriched in the lists of genes differentially expressed in each MS phenotype. We took into consideration

differentially expressed genes with $p < 0.05$ and fold change > 1.5 . Totally, 464 MS-AIG were differentially expressed in at least one MS phenotype but no significant enrichment appeared in any of the 4 disease subtypes when compared with a random selection of transcripts from the whole transcriptome. However, when considering separately the MS-AIG from single interactomes, we observed an enrichment of differentially expressed MS-AIG from AIRE interactome in CIS patients, from AIRE and SIRT7 interactomes in RRMS, and from EBV interactome in PPMS (Table 1).

To understand the functional implications of our findings also at the central nervous system (CNS) level, we matched the 741 MS-AIG with microarray gene expression data obtained from post-mortem brain tissue of 20 people with secondary progressive MS (SPMS) and 10 non-neurological controls. Grey matter lesions (GML) and normal appearing grey matter (NAGM), from MS brains that exhibited the presence (F+) or absence (F-) of lymphoid follicles, were considered (Magliozzi et al. 2019). Again, we calculated whether the overall frequency of MS-AIG was significantly 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 analyzed (GML-F+, GML-F-, NAGM-F+, NAGM-F-). The result was significant (Table 2). Considering separately the MS-AIG obtained from each MS-associated interactome, we observed a significant enrichment only for genes belonging to the EBV interactome in all the four conditions analyzed.

The above results from peripheral blood and CNS transcriptomes imply a potential role for EBV-related neuroimmunological response in the neurodegenerative and neuroinflammatory aspects of the disease. This prompted us to consider recent data showing an enrichment of EBV interactors in protein modules involved in Hereditary Spastic Paraplegia (HSP), an inherited degenerative axonopathy (Bis-Brewer et al. 2019) sharing clinical similarities with progressive MS. Importantly, it has been shown that persons with PPMS are enriched for HSP-related mutations (Jia et al. 2018). 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) (Figure S4). This number resulted larger than randomly expected ($p\text{-value} < 0.00001$). In summary, gene expression data support the results of the candidate interactome analysis and show that the functional implications may be more condensed at the CNS level.

Interactome-driven pathophysiology converges on the CD40 pathway in MS

To highlight the potential biological functions affected by the MS-AIG, we performed a pathway analysis using MetaCore[®]. The 741 MS-AIG were mapped to the internal MetaCore KnowledgeBase and a classification in canonical pathways was obtained. This analysis showed that the most critical biological process involved in MS through gene-environment interactions was CD40 signaling, followed by stress-induced antiviral cell response, oxidative stress and apoptosis (Figure 3A). This result was in agreement with studies showing the role of the CD40 pathway and that of B cells in MS pathophysiology (Afrasiabi et al. 2019; Field et al. 2015; Milo 2019; Smets et al. 2018).

To investigate whether CD40 alterations in MS could be driven by the interaction between host- and virus-specific factors we used flow cytometry to evaluate the surface protein expression of CD40 and co-stimulatory molecules CD80 and CD86 [known to be induced by CD40 and by EBV infection (Morandi et al. 2017)] in spontaneously outgrowing lymphoblastoid cell lines (spLCLs) obtained from 13 persons with MS (MS-spLCLs) and 8 healthy donors (HD-spLCLs). SpLCLs were obtained from B cells infected by the endogenous EBV strain: this condition allowed us to study the immunologic profile in a setting that may be more representative of a putative pathophysiological context where host and virus genetic variants interact in vivo. Results were compared with those obtained in B cells infected with the recombinant EBV strain B95.8 (B95.8LCL).

We found a reduction of CD40 protein level in MS-spLCLs, compared to HD-spLCLs ($p=0.04$; Figure 3B). This was especially evident when we stratified the data according to sex: we observed a more pronounced difference in females with MS compared to healthy females ($p=0.006$), and a near-significant difference ($p=0.06$) when we compared females and males in the MS group (Figure 3C).

Analyzing the CD80 protein level, we found a slight reduction of the surface protein in MS-spLCLs compared to HD-spLCLs ($p=0.089$; Figure 3D); once again a significant difference emerged when we compared the affected and healthy female groups ($p=0.0037$, figure 3E). The CD86 protein levels showed no differences between MS-spLCLs and HD-spLCLs (Figure S5). We then performed a simple linear regression analysis to evaluate a possible interdependence between CD40 and CD80 protein levels. The two protein levels were correlated in MS-spLCLs ($r^2=0.43$ $p=0.016$; Figure 3F) but not in HD-spLCLs ($r^2=0.32$ $p=0.16$; Figure 3G).

The MS associated SNPs rs4810485, which maps within the CD40 gene, and rs9282641, which maps within the CD86 gene, have been reported to be associated with a reduced CD40 expression and an increased expression of CD86 on the B cell surface (Smets et al. 2018). We did not observe this correlation in the spLCLs (Figure S6). We also measured CD40, CD80 and CD86 protein expression in B cells infected with recombinant EBV strain B95.8 (B95.8LCL) in 6 patients and 6 HD but did not find significant differences (Figure S7).

Finally, we performed a multiple linear regression to assess the possible influence of the EBV strain (spLCL vs B95.8LCL) on CD40 expression level controlling for CD80, since CD80 levels were correlated with CD40 protein expression (see Figure 3F). For this analysis, we considered only spLCL and B95.8LCL cell lines obtained from the same subject, at the same time point. This experimental setting allowed us to observe a lower CD40 expression in MS-spLCLs compared to MS-B95.8LCLs ($p=0.04$; Figure 3H-I), suggesting an effect of MS-associated EBV variants on CD40 protein level. With the possible limitation of the reduced sample size of spLCLs under scrutiny, we were not able to link this effect to MS-associated EBV variants previously found in the EBNA2 gene (Mechelli et al. 2015). This issue deserves further investigation, by increasing the number of spLCLs scrutinized and extending the analysis to other EBV genes.

We then used MetaCore[®] to identify the biological functions possibly affected by the MS-AIG dysregulated in the CNS. Once again, we found CD40 signaling as the pathway that more clearly emerged from the disease-predisposing, host-environment interaction in MS, also at the lesion level. By using digital droplet-PCR (ddPCR) we analyzed the CD40 mRNA level in the same RNA samples obtained from cortical lesions of post-mortem progressive MS cases used for the microarray gene expression analysis. CD40 expression was increased in the subpial cortical lesions of MS cases compared to the normal grey matter of non-neurological controls (fold change: 1.9, $p=0.007$; Figure 4A, Table S2). By using immunohistochemistry, the CD40 expression was observed mainly in active lesions, either in the grey matter (Figure 4B, C) or in the white matter (Figure 4E, F). Sporadic CD40⁺ cells were also observed in normal appearing grey matter (Figure 4D) and normal appearing white matter (Figure 4G), while CD40 expression was not detected in inactive lesions and in non-neurological controls. In particular, CD40 expression was detected in cells morphologically resembling activated microglia, as further supported by double immunolocalization with the MHC class II marker (Figure 4 insets in C and in F). Though the experimental settings (spLCLs and brain lesions) do not allow to conduct analyses on large numbers, the results suggest that the interaction between host- and EBV-specific factors may lead to CD40 signaling alterations.

EBV interacting proteins have the highest therapeutic potential in a Priority Index analysis of MS-AIG

A recent genetics-led drug target prioritization approach (Fang et al. 2019) represented an interesting resource to translate our results into therapeutically relevant information. Fang and colleagues developed a Priority Index (Pi) pipeline, based on GWAS variants for several immunopathological traits, including MS, that allowed us to investigate whether, and to what extent, the MS-AIG may contribute to the discovery of new druggable targets. Among the first 150 positions in Pi Ranking for MS, 26 appeared in MS-AIG (Figure 5A). This enrichment was specifically due to 13 genes belonging to the EBV interactome ($p=0.0045$). We also matched the MS-AIG with the list of genes showing the highest druggability among the top 1% genes prioritized in the work by Fang et al. Interestingly, 24 out of the previously identified 26 MS-AIG were included in high-Pi rated genes and were highly-druggable targets. Then we looked at the position of the 741 MS-AIG genes within the actual landscape of drug target genes (including Phase 1 to Phase 4 trials) in immunological diseases. Six of these genes proved to already be the object of clinical investigation for MS, while 10 were not specifically intended to treat MS, but are involved in the treatment of other diseases (Figure 5B and Table S2). Among these were the highest Pi-prioritized MS-AIG, such as IL6, JAK1 (HBV interactome), JAK3 and SYK (EBV interactome). The results therefore suggest that a pharmaceutical targeting of relevant gene-environment interactions is a feasible strategy that may potentially advance the efficacy of MS treatments.

DISCUSSION

The association between MS and EBV is a long recognized one (Ascherio and Munger 2016). Starting from the indications of a preliminary study (Mechelli et al. 2013), we here show that the interaction between genotype and herpesviruses may have causal implications in MS, in a meaningful setting that takes into account new and more accurate data and additional *in silico* and *in vitro* analyses. The results now indicate that this finding is specific for MS and not for other immune-mediated diseases. Peripheral blood and CNS transcriptome data in MS show a dysregulation of MS-associated EBV interactors in both compartments, further validating the result (Gustafsson et al. 2014; Lamparter et al. 2016). Pathway analyses of the MS-AIG and protein expression experiments in spLCLs from persons with MS and controls, show that the interaction between MS-AIG and viral exposures may converge towards the dysregulation of the costimulatory receptor CD40 and its pathway. In accord with the observed transcriptomic dysregulations in MS lesions, matching MS-associated interactome genes with protein interaction modules from inherited axonopathy-causing genes, supported a link between MS-associated EBV interactors and neurodegeneration. Finally, we show that products of the MS-AIG may be prioritized as drug targets.

Compared to our previous results, the new “candidate interactome” analysis confirms the relevance of EBV, shows the emergence of other herpesviruses and appears to disconfirm other viruses (e.g. HBV) that were significant in the previous study. These results are in line with the sero-epidemiological literature, with obvious biological similarities between EBV and other herpesviruses, and with biological plausibility. Furthermore, the results indicate that this kind of analyses become increasingly meaningful and reliable as new genetic and protein-protein interaction data become available (Huttlin et al. 2021; Wang et al. 2021). It is remarkable that, in a recent study considering a multi-omics approach in MS (Badam et al. 2021), the analysis identified a module of 220 genes strongly associated with risk factors in MS. This module shows a significant overlap with our 741 MS-AIG and with the EBV interactome genes whose expression we found dysregulated in the transcriptome analyses (Figure S8). Concerning other viruses, and again in comparison with our previous work, HIV retained a significant association. In spite of the lack of epidemiological links between HIV and MS, the result will deserve further attention as HIV can induce demyelination (Berger et al. 1989; González-Duarte et al. 2011; Johnson 1994). Furthermore, HIV may share some functional kinship with the HERV family of endogenous retroviruses potentially involved in the pathogenesis of MS and of other neurodegenerative diseases (An et al. 2001; Küry et al. 2018). Finally, our finding of CTCF acting as a common regulator of the MS-AIG reinforces the idea that the MS-AIG are disease-relevant genes that may be affected by the regulatory influence of EBV. In fact, it has been shown that CTCF-bound sites show enhancer-blocking activity towards developmental genes associated with human diseases (Martin et al. 2011) and overlap with sites where EBV alters accessible chromatin (Lamontagne et al. 2021), subverting the host’s gene expression program (Ersing et al. 2017; Zhou et al. 2015).

Support to our results comes also from two recent studies on low-frequency and rare-coding variants associated with MS (Consortium 2018; Vilariño-Güell et al. 2019). Altogether, these studies identified eighteen genes harboring MS-associated low-frequency coding variants. Surprisingly, half of these are either EBV interactors or implicated in EBV-related pathophysiologies (for details see Table S1). In a third study (Vidmar et al. 2019), MS-associated rare variants could be linked to molecular mechanisms connecting EBV with interferon-beta signaling and autophagy. Finally, the SNP with the highest association with white matter lesion volume in the recent genome-wide association study of brain magnetic resonance imaging phenotypes of the UK Biobank (Elliott et al. 2018), is in the 5’ UTR of EFEMP1, a gene that codes for another EBV interactor (its product interacts with two EBV proteins, BFLF2 and BRRF1) (Gulbahce et al. 2012).

Autoimmune diseases have common and disease-specific signatures, recently highlighted by gene expression studies of target tissues (Szymczak et al. 2021). In the present study, we applied the interactome analysis to immune-mediated diseases other than MS, where epidemiological studies have found associations with EBV, though probably less robust compared to MS. In these conditions, our data do not support the idea that the interaction between EBV and the genetic background of the host may contribute to disease etiology. The difference between MS and the other immune-mediated conditions appears plausible in the light of differences in pathophysiology and response to treatments. Distinctions have been proposed in genetic (Gregory et al. 2012) and immunological studies on the tumor necrosis factor (TNF) system (Gao et al. 2017; Veroni et al. 2010), one of the most ancient antiviral mechanisms (tenOever 2016). In addition to suggestions coming from mechanistic studies, clinical evidence is clear cut: it involves not only the well-known, opposite effects of anti-TNF therapies (highly beneficial in RA and detrimental in MS) but also immunotherapies that exploit cytokines with antiviral actions, such as interferon-beta. These are effective in MS but not in RA (Genovese et al. 2004; van Holten et al. 2005) or in other immune-mediated diseases where they may concur in the development of the diseases (Axtell et al. 2011). Furthermore, the response to anti-B lymphocyte therapies in MS, including in the progressive form of the disease, may be indicative of a role of EBV in disease pathophysiology as this virus remains for the lifetime of the host in memory B cells. Of note, in RA, anti-B lymphocyte therapies may not be as effective as anti-TNF treatments (González-Vacarezza et al. 2014). Despite these differences, other mechanisms of EBV involvement in autoimmunity cannot be excluded. For example, the binding of EBNA2 to genomic intervals associated with various autoimmune diseases, including MS, suggests another mechanism that operates across diseases (Harley et al. 2018; Ricigliano et al. 2015). Furthermore, also in view of results about HHV-6 and HHV-7 in Alzheimer's disease and in a non-human primate model of MS (Leibovitch et al. 2018; Readhead et al. 2018), it will be necessary to study more immune-mediated diseases and various other "candidate interactomes" of viral origin to draw a complete picture of the long-suspected role of viruses in autoimmunity and in neurodegeneration. Incidentally, given the pleiotropic effects of EBV on the immune response, the negative result in immune-mediated diseases other than MS to some extent controls for the possibility that the association between EBV interactome and MS derives from effects of the genetic variants on the outcome (MS) that are not exclusively mediated by the exposure (EBV) (horizontal pleiotropy) (Davey Smith and Hemani 2014; Davies et al. 2018).

An elegant study comparing gene expression in unselected peripheral blood B cells and B95.8 infected B cells in healthy individuals recently showed that EBV infection affects MS risk genes (Afrasiabi et al. 2019). We now extend this result to the disease state: the analysis of peripheral blood and brain cortex transcriptome data confirmed that, among the MS-AIG, those that are part of the EBV interactome are more frequently dysregulated than expected. Of the eight sets of gene expression data (four from the peripheral blood and four from brain cortical areas), MS-AIG belonging to the EBV interactome were significantly dysregulated in five. The dysregulation was more consistent in transcriptomes from the four brain tissue types (4/4) than in those from the peripheral blood (1/4). This result, particularly in the datasets from cortical tissue devoid of lymphoid follicles, may be to some extent unexpected given the propensity of EBV to infect B lymphocytes and influence the immune response. However, astrocytes and microglia activation, which have a key role in MS inflammatory response (James et al. 2020) in MS, may express the EBV receptor CD21 (Lindblom et al. 2015) and may be infected by the virus in this disorder (Hassani et al. 2018), suggesting that EBV may entertain pathogenic interactions also outside the B cell compartment. Analyses of the most recent MS-GWAS (Consortium 2019b), which shows an enrichment for MS genes in microglia, may clarify this issue, though it is already intriguing that the MS variant with the strongest eQTL in microglia, CLECL1, is an EBNA3C-regulated gene in lymphoblastoid cell lines (Zhao et al. 2011) and is upregulated in tonsillar B cells after CD40 ligation (Ryan et al.

2002). Here we show that EBV interactors within the MS-AIG are enriched in protein interaction modules derived from genes that cause hereditary spastic paraplegia (Bis-Brewer et al. 2019). In this disease, that has no clear evidence of an immune-mediated component but shares neurodegenerative features and clinical similarities with MS, a potential link with EBV infection (Bis-Brewer et al. 2019) and with MS (Jia et al. 2018) had been already suggested. Overall, these data imply that the contribution of EBV to CNS damage may, to some extent, be independent of a canonical inflammatory context. Whether or not the hypothesis of EBV acting at the CNS level be supported by its presence in latently infected B cells at the lesion level or in meningeal lymphoid follicles remains a controversial issue (Aloisi et al. 2010; Lassmann et al. 2011; Ramesh et al. 2020; Sargsyan et al. 2010; Serafini et al. 2007; Willis et al. 2009).

The CD40 pathway (a key factor in antigen presentation, immune regulation, B lymphocyte development, germinal center formation, somatic hypermutation and class-switch recombination) emerges among the biological functions affected by the MS-AIG, again in accord with data comparing peripheral blood and B95.8 infected B cells (Afrasiabi et al. 2019). This result is coherent with what is known about the role of the CD40 molecule in MS, where the risk allele affects the membrane-bound molecule and splice forms of decoy receptors (Field et al. 2015; Smets et al. 2018), once again in an opposite way between MS and other immune-mediated diseases such as rheumatoid arthritis (Raychaudhuri et al. 2008), systemic lupus erythematosus (Vazgiourakis et al. 2011), autoimmune thyroiditis (Tomer et al. 2002) and Kawasaki disease (Onouchi et al. 2012). With respect to the CD40 expression in LCLs, the data suggest an effect of MS-associated EBV variants on the downregulation of the membrane-bound molecule. However, due to the small numbers of spLCLs available, we were unable to understand which EBV variants are responsible for this effect. In more general terms the results are in line with the idea that viruses and disease risk alleles may induce or enhance diseases through similar mechanisms (Chen and Xia 2019; Gulbahce et al. 2012) and with the observation that the EBV-induced regulatory perturbations are often in the direction of a reduced gene expression (Zhou et al. 2015). Here, we suggest also a broader involvement of the CD40 signaling pathway and outline the gene-environment interactions that may generate its involvement in MS. This information is relevant as the CD40 pathway may participate in the creation of memory B cells (Baker et al. 2018) and in the activation of a repertoire of autoreactive, brain-homing CD4+ T cells in a way that does not seem to directly depend on the CD40 molecule itself (Jelcic et al. 2018). Interestingly, EBV-specific CD4+ T cells that cross-react with myelin and neuronal epitopes are part of this repertoire (Wang et al. 2020), suggesting the possibility that EBV participates in MS pathogenesis through multiple and synergic mechanisms.

The fact that we found increased CD40 expression by microglia/macrophages in active grey matter and white matter lesions of progressive MS cases may support the hypothesis that this molecule contributes to the active phase of lesion pathology and may therefore represent a potential candidate target for future therapeutic strategies (Jäckle et al. 2020). Previous experimental autoimmune encephalomyelitis (EAE) studies demonstrated that CD40 is necessary for the activation of microglia cells since the initial disease stages (Gerritse et al. 1996; Ponomarev et al. 2006). Our findings support and expand these data in the context of the human disease: the increased CD40 expression in NAGM and NAWM (even in absence of demyelination) in cell resembling activated microglia points to a pivotal role of brain resident innate immunity in disease initiation. CD40 may thus contribute to bridge the innate and adaptive immune response in neuroinflammation, deserving further and more in-depth analysis, given its clinical implications. In this context, our approach offers the opportunity to pinpoint therapeutic targets, prioritized on a genetic basis (Fang et al. 2019), that are related to “environmental” exposures. As such, these targets are in principle more actionable than others based on an ongoing host pathophysiology, particularly for preventive purposes. The Pi analysis of the MS-AIG, once more, highlighted the EBV interactome as a relevant reservoir of therapeutic targets and exposed MS-specific target pathways (e.g. interferon signaling) among which CD40 ranked highest.

In accord with recent suggestions about the importance of understanding genetic variability at the systems level rather than in individual genes (Li et al. 2019) this study indicates that the interaction between host genotype and EBV (including other Herpesviruses) is relevant for MS etiology, reinforcing the rationale for future studies that attempt to decipher disease mechanisms by looking at virus-mediated pathology (Müller-Schiffmann et al. 2021). It also strengthens the rationale for experimental medicine approaches aimed at advancing therapies while proving the causality of viral agents in MS. The results of a recent study on EBV-specific adoptive T cell therapy in patients with progressive MS support this perspective (Pender et al. 2018).

Acknowledgments

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

CENTERS is a Special Project of the Italian Multiple Sclerosis Foundation (FISM).

G.M. is funded by Fondazione Italiana Sclerosi Multipla (FISM; grants 2016/R/18 and 2018/S/5), Progetti di Rilevante Interesse Nazionale (PRIN; grant 2017K55HLC 001), and Ministero della Salute (grant RF-2019-12371111). P.d.C. is supported by FISM (grant number 2018/R/4).

Author contributions

R.M., R.U., G.R., M.S. conceived and designed the study. R.M. selected and constructed the candidate interactomes and performed digital droplet-PCR analysis. R.U., V.R., G.B. and R.U.P. performed the bioinformatics works. R.M., R.U., D.C., L.B., G.R., M.S. analyzed the data. S.S. and C.F. provided blood transcriptome data and were involved in data analysis. R.M.A. and R.R. provided brain transcriptome data and were involved in data analysis. R.M.A. performed neuropathological brain tissue analysis. D.F.A. and G.G. performed cytofluorimetric analysis. S.R. and G.B. conceived and drawn the graphical abstract. R.B., E.A. and P.T. provided lymphoblastoid cell lines and were involved in data analysis. A.F., M.F., MCB and S.R. provided reagents, analysis methods, and instrumentation. S.D.A. carried out sample genotyping and was involved in data analysis. R.M., G.R. and M.S. provided data interpretation and wrote the manuscript. C.F., L.B., S.D.A., P.D.C., G.M., A.U., D.D.S., P.M. and D.C. provided critical reading of the manuscript. IMSGC and WTTTC2 provided GWAS data.

Competing interests:

Mechelli R., Renato U., Rinaldi V., Bellucci G., Angelini D.F., Guerrero G., Bigi R., Fornasiero A., Ferraldeschi M., Romano S., Buscarinu MC., Srinivasan S., Anastasiadou E., Trivedi P., Di Silvestre D., Mauri P., De Candia P., Battistini L., D'Alfonso S., Magliozzi R. and Ristori G. declare no conflict of interest.

Centonze D. 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. His preclinical and clinical research was supported by grants from Bayer Schering, Biogen Idec, Celgene, Merck Serono, Novartis, Roche, Sanofi- Genzyme e Teva

Pizzolato Umeton R. is supported by the National MS Society.

Farina C. received research support from Merck-Serono, Teva, Novartis.

Uccelli A. has received consulting honoraria including as part of advisory board and/or speaker fees from Sanofi Genzyme, Roche, Biogen, Novartis, TEVA and Merck.

Matarese G. reports receiving research grant support from Merck, Biogen, and Novartis and advisory board fees from Merck, Biogen, Novartis, and Roche

Reynolds R. received research support from the UK MS Society and MedImmune and has received speaker honoraria from Roche and Novartis.

Salveti M. received research support and consulting fees from Biogen, Merck-Serono, Novartis, Roche, Sanofi, Teva

Figure legends

Figure 1: Disease-associated interactomes

(A) Schematic representation of candidate-interactome analysis.

(B) List of interactomes and related sizes and sources.

(C) 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 ($p < 0.01$).

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= Human Immunodeficiency virus; 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;

Figure 2: Genomic distribution of the candidate interactome SNPs and protein binding enrichment near MS-associated SNPs

(A) The UpsetPlot displays the number of MSAIG for each interactome of origin (x axis) and the intersections between gene sets with the respective size (y axis) (see also Table S1).

(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 grey 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 $-\log(P\text{value})$ for each SNP respectively from GWAS 2011 and GWAS 2013 (scaled from 0-20, 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 bubble plot displays the number of SNPs bound by a given protein obtained from different cells types (vertical axis) and the top 50 proteins enriched on MS-associated SNPs (horizontal bar).

NMD = Nonsense-mediated mRNA decay; 3'UTR = three prime untranslated region.

Figure 3: CD40 signaling analysis.

(A) The top 10 pathways obtained from the analysis of the MS-AIG. The pathway analysis was performed using MetaCore and its canonical pathway maps. The lower p-value ($-\log P\text{value}$) means higher relevance of the genes within the pathway datasets. The CD40 signaling reached the highest relevance.

(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 under-expressed in females of MS group (F-MS) compared to the females of 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 under-expressed in F-MS compared to the F-HD (**p=0.0037).

(F-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). Dashed curves represent the confidence interval, the straight lines represent the correlation slope.

The protein levels are represented as median fluorescence intensity. Data analysis was performed using Unpaired t-test F= female; M= male.

(H-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.

Figure 4: CD40 analysis in post-mortem progressive MS brain tissues

(A) ddPCR analysis of CD40 gene expression, referred to the housekeeping gene GAPDH, in 6 GML and in 3 NGM from controls. The CD40 mRNA level is up-regulated in diseased tissue at Unpaired t-test. Data are shown as mean \pm SD. **p= 0.007.

(B-G) Neuropathological assessment of the CD40 protein expression in post-mortem MS cases. CD40 protein expression was observed mainly in active lesions, either in the grey matter (B, C) or in the white matter (D, F), in cells morphologically resembling activated microglia, as further supported by double immunolocalization with the MHC class-II marker of activated microglia/macrophages (insets in C and in F). Sporadic CD40+ cells were also observed in normal appearing grey matter, NAGM (D) and normal appearing white matter, NAWM (G). Original magnifications: 10x (B, D); 100x (C, D, F, G); 400x (insets in C and in F).

Figure 5: MS-AIG Pi-Prioritization and distribution in actual drug targets for immunological diseases.

(A) The figure displays MS-AIG present in the top 150 Pi-prioritized genes for MS (see text, and Fang et al. 2019). *Y axis* shows the number of genes and their identifier, *x axis* shows the corresponding interactomes.

(B) The diagram shows the overlap between the 741 MS-AIG and the actual clinical proof-of-concept drug targets for immunological diseases, including MS, as resulting from Fang et al, 2019. The 16 overlapping genes are split between MS and other immunopathological conditions and reported in the charts with the corresponding interactome.

Table 1: MS-AIG differentially expressed in peripheral blood leucocytes. The enrichment of MS-AIG was calculated considering separately each interactome in distinct MS phenotypes. The associations reaching significance are highlighted in bold.

MS phenotypes	N° of DE MS-AIG	P-value	OR	Frequency of MS-AIG in blood DEG (%)											
				AHR	AIRE	CHLAMYDIA	CMV	EBV	HBV	HHV8	HIV	POLYOMAVIRUS	SIRT7	VIRORF	GLOBAL TRANSCRIPTOME
CIS	166	0.59	0.28	25.00	50	29.03	34.62	24.32	16.00	22.09	21.43	45.45	27.93	30.51	25.56
RRMS	170	0.34	0.89	25.00	40.63	25.81	19.23	22.97	24.00	27.27	21.90	45.45	34.23	27.97	23.66
PPMS	249	0.082	3.02	25.00	40.63	38.71	38.46	44.59	46.00	42.86	32.86	27.27	31.53	34.75	33.83
SPMS	226	0.53	0.37	31.25	37.50	27.42	32.69	33.11	28.00	37.66	32.86	9.09	30.63	34.75	32.38

CIS = clinically isolated syndrome; RR-MS = Relapsing-Remitting multiple sclerosis; PP-MS = Primary progressive multiple sclerosis; SP-MS = Secondary progressive multiple sclerosis; DEG= differentially expressed genes.

Table 2: MS-AIG differentially expressed in brain.

The frequency of MS-AIG was calculated in DEG of 4 post-mortem brain areas from secondary progressive MS. The enrichment of MS-AIG was also calculated considering separately each interactome in distinct brain areas. The associations reaching significance are highlighted in bold.

Brain areas	N° of DEG	P-value	OR	Frequency of MS-AIG in brain DEG (%)											
				AHR	AIRE	CHLAMYDIA	CMV	EBV	HBV	HHV8	HIV	POLYOMAVIRUS	SIRT7	VIORF	GLOBAL TRANSCRIPTOME
NAGM-F+	48	0.008	1.53	7.69	5.88	4.55	5.36	10.07	4.17	7.69	4.98	11.11	7.20	7.69	6.61
NAGM-F-	127	0.001	1.4	7.69	14.71	18.18	17.86	21.48	16.67	15.38	16.60	22.22	16.80	18.46	17.14
GML-F+	52	0.003	1.59	0.00	11.76	3.03	5.36	12.75	6.25	6.59	7.47	5.56	5.60	6.15	7.29
GML-F-	133	0.001	1.42	15.38	17.65	16.67	14.29	22.15	18.75	16.48	16.60	16.67	17.60	16.92	18.08

DEG= differentially expressed genes; FC= fold change; OR= Odd ratio; NAGM-F+= normal appearing grey matter follicle positive; NAGM-F-= normal appearing grey matter follicle negative; GML-F+= grey matter lesions follicle positive; GML-F-= grey matter lesions follicle negative

Supplementary figure legends:

Figure S1: Schematic representation of METACHIP datasets construction, related to Figure 1.

Metachip 1 and 2 were built as the position-wise union of GWAS and Immunochip. Each dataset is represented by a line in which sample SNPs are indicated with different colors depending on their source dataset (black round for SNPs coming from GWAS and empty round for SNPs coming from Immunochip). When a given SNP existed in both GWAS and Immunochip, we gave preference to GWAS in Metachip1 and to Immunochip in Metachip2.

Figure S2: Spearman correlation in MS GWAS, related to Figure 1.

Linear correlations between the size of single interactomes and their cumulative p-value of association with MS calculated for each GWAS dataset (GWAS 2011, GWAS 2013, METACHIP 1 and METACHIP 2). The 95% confidence intervals of the Spearman correlation are greyed out in all the analyses.

Figure S3: Spearman correlation in non-MS GWAS, related to Figure 1.

Linear correlations between the size of single interactomes and their cumulative p-value of association with complex diseases calculated for each GWAS dataset (Bipolar disorder, celiac disease, Chron disease, coronary artery disease, hypertension, rheumatoid arthritis, T1 diabetes, T2 diabetes). The 95% confidence intervals of the Spearman correlation are greyed out in all the analyses.

Figure S4: Overlap between MS-AIG and genes of HSP-dominant expanded modules.

(A) Overlap between MS-AIG and HSP-protein modules genes. Out of 71 shared genes, 34 resulted EBV interactors. (B) Null distribution of EBV interactors included in 100.000 random samples of the same size of the 71 shared genes, extracted from MS-AIG. The observed number of EBV interactors (N=34 out of 71 shared genes, red line) was larger than randomly expected (p value<0.00001).

Figure S5: The protein expression of CD86 on spLCLs is comparable between MS (MS-spLCLs, n=10) and controls (HD-spLCLs, n=6; A), related to Figure 3. The protein levels are represented as median fluorescence intensity. The data analysis was performed using Unpaired t-test.

Ns= not statistically significant.

Figure S6: Genotype independent CD40 and CD86 protein expression on spLCLs, obtained from MS patients and controls (HD), related to Figure 3. (A) The rs4810485*T is the CD40 MS risk allele. The protein expression levels (represented as median fluorescence intensity) are independent from the genotype when analyzed overall (GG, n=12; TG, n=7; TT, n=2) and when MS (GG, n=7; TG, n= 4; TT, n=2) and HD (GG, n=5; TG, n=3; TT =not detected) were analyzed separately.

(B) The rs9282641*G is the CD86 MS risk allele. The protein expression levels are independent from the genotype when analyzed overall (GG, n= 15; GA, n=2) and when MS (GG, n=10) and HD (GG, n= 5; GA, n=2) were analyzed separately.

The data analysis was performed using 1-way ANOVA, Bonferroni post-test. Ns= not statistically significant

Figure S7: Protein expression of CD40, CD80 and CD86 on B95-8LCLs surface, related to Figure 3.

The protein levels are comparable between MS (n=6) and HD (n=6)-derived B95-8LCLs.

On *Y-axis* are represented the median fluorescence intensity of each protein. The data analysis was performed using Unpaired t-test. Ns= not statistically significant.

Figure S8: Overlap between MS-AIG and the MS disease module derived from multi-omics analysis by Badam et al, 2021.

(A) Overlap between all 741 MS-AIG and the 220 genes forming the MS module defined by Badam et al, 2021 through comparative multi-omics analysis.

(B) Overlap between the MS Multi-Omic Module and the MS-AIG differentially expressed in the brain, blood and both. The overlap is considered significant if $p < 0.05$ (Fisher's exact test, calculated through geneOverlap package in R software). ns= not significant.

STAR METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Marco Salvetti (marco.salvetti@uniroma1.it). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Lymphoblastoid cells lines generation and DNA extraction

Sp-LCLs, carrying endogenous EBV, were generated from 13 subjects with relapsing remitting MS (Polman et al. 2011) and 8 healthy donors matched for age, sex and geographic origins (Chiara et al. 2016). Briefly, peripheral blood mononuclear cells (PBMCs) obtained by density centrifugation over Ficoll–Hypaque were seeded in 96-flat well plates and cultured in RPMI1640 medium, supplemented with 20% FBS (HyClone), 1% L-Glutamine, 100 IU penicillin, and 100 µg/ml streptomycin. Cyclosporin A (1 µg/ml, Calbiochem) was added to the medium to inhibit T-cell activation and cultures were fed twice a week.

PBMC infected with EBV B95.8 strain (B95.8-LCL) were generated from 5 relapsing remitting MS subjects and 4 healthy donors incubating for 2 hours at 37°C the primary cells in a mix containing: virus (10x concentrated), Cyclosporin A (1 µg/ml) and RPMI1640 medium supplemented with 20%FBS, 1% L-Glutamine, 100 IU penicillin and 100 µg/ml streptomycin. Cells were seeded in 24-well plates and fed twice a week.

Genomic DNA was extracted from five million cells using the QIAamp DNA mini kit (Qiagen, Venlo, the Netherlands).

All subjects with MS were sampled during the stable phase of the disease and were free from disease modifying therapies. The local institutional review board approved the study and all participating subjects gave written informed consent.

Flow cytometry analysis

The phenotypic characterization of LCLs obtained from MS patients and HD was performed through multiparametric flow cytometric (MPFC) analysis. The following monoclonal antibodies were used: anti-CD40 PE (Pharmingen), B7-1 (anti-CD80 PE-Cy7) (Beckman Coulter), B7-2 (anti-CD86 APC) (Miltenyi), anti-CD19 APC-eFluor780 (eBioscience) and LIVE/DEAD Fixable Aqua Dead Cell Stain (ThermoFisher). The staining was performed with 0.5×10^6 cells with antibodies at previously defined optimal concentrations and incubated for 30 minutes at 4°C in the dark. The samples were acquired on a CyAn flow cytometer (Beckman Coulter), equipped with three lasers and able to measure up to 15 parameters simultaneously on each cell. For each sample, approximately 300,000 cells were selected based on scatter parameters, physical size (FSC) and grain size (SSC), and the analysis was conducted after the exclusion of dead cells and coincident events. Data were compensated and analyzed using FlowJo v10.5 (TreeStar, Ashland, OR, USA).

Statistical analyses were performed with GraphPad Prism 5.0 program.

CD40 and CD86 SNPs genotyping

We genotyped rs4810485 SNP in CD40 gene and rs9282641 SNP in CD86 gene in 26 spLCLs, using PCR amplification followed by Sanger sequencing.

The following PCR primers were used:

CD40 rs4810485: primers: forward: 5'- TGGTGATCAGAGGGCTGGAGAAA -3', reverse: 5'-

TCTCCACTCCTACCACAAGGGC-3'

CD86 rs9282641: primers: forward 5'- AGGGCTTTACACTCATGCTCCGAG-3', reverse: 5'-

TTTTACTCATCTCTGCGCCGCCA-3'

PCR amplicons were analyzed for SNP genotypes by Sanger direct sequencing using the Big-Dye Terminator v3.1 sequencing kit (Applied Biosystems, Foster City, California, USA) on an automated 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

PCR and sequencing conditions are available on request.

Candidate interactome analysis

The “candidate interactome” approach was applied as previously described (Mechelli et al. 2013). In brief, we obtained 20 candidate interactomes from the literature. Six were manually curated (EBV, HBV, CMV, HHV8, JCV, Inflammasome); 7 were obtained from databases of molecular interactions: AIRE, VDR, AHR, SIRT1, SIRT7 from BIOGRID (<http://thebiogrid.org>) (Chatr-Aryamontri et al. 2015), Polyomavirus from VirusMentha database (Calderone et al. 2015), Human-microRNA targets from miRecords database (Xiao et al. 2009); 7 were obtained from published high-throughput experimental approaches: VIORF (Pichlmair et al. 2012), HIV (Jäger et al. 2011), HCV (de Chassey et al. 2008), h-IFN (Li et al. 2011), H1N1 (Shapira et al. 2009), HDAC (Joshi et al. 2013), Chlamydia (Mirrashidi et al. 2015) (Figure 1A).

The manually curated interactomes were obtained by selecting only those interactions that were reported by two independent sources or were confirmed by the same source with distinct experimental approaches. In all cases, we considered only physical-direct interactions (Mechelli et al. 2013). All SNPs which did not pass quality checks in the GWAS studies were filtered out from the original data. As reference to gather gene and single nucleotide polymorphism (SNP) details from their HUGO Gene Nomenclature Committee (HGNC) Ids and rsids, a local copy of the Ensembl Human databases (version 75, databases core and variation, including SNPs coming from the 1000 Genome project) was employed; the annotation adopted for the whole analysis was GRCh37-p13, that includes the release 6 patches (Genome Reference Consortium: human assembly data - GRCh37.p13 - Genome Assembly (www.ncbi.nlm.nih.gov/assembly/GCF_000001405.25)). We used ALIGATOR (Holmans et al. 2009; Houlston et al. 2010) program to evaluate how SNPs and single genes get summed to provide total contribution of candidate interactomes (Mechelli et al. 2013). The idea behind ALIGATOR’s strategy is to evaluate gene category significance by means of a bootstrapping approach: compare each interactome with the null hypothesis, that was built using random permutations of the data based on a non-parametric bootstrap analysis that uses the Gene Ontology database. Filtering criteria for the SNP selection included (i) a p-value significance taken at p-value cut-off (P-CUT) in the summary statistics; (ii) linkage disequilibrium (LD) filtering was applied. The parameters used to configure ALIGATOR are those reported in its reference paper (Holmans et al. 2009): P-CUT was taken at 0.05, the SNP-gene association parameter was maintained at its default value of 20kbp, and the LD filter was kept at $r^2 < 0.2$. For all disease-associated interactomes, SNPs were extracted after the ALIGATOR filtering and bootstrapping analysis steps. The idea behind this extraction was to pinpoint which SNPs were considered relevant by the ALIGATOR algorithm. Genes (MS-AIG) were mapped to SNPs by the ALIGATOR internal mapping step, for bp distance, p-value, and LD, using the default settings optimized in Holmans paper (Holmans et al. 2009).

Protein binding enrichment

Genomic regions in Figure 1D were identified using Variant Effect Predictor version 98. After the ALIGATOR analysis, we used RegulomeDB database (Boyle et al. 2012), that is fed with live data from ENCODE and Roadmap Epigenomics, to look for proteins binding in regions containing "post-match" SNPs that displayed nominally significant associations with MS. When SNPs are bound by two or more proteins a Fisher's exact test (2x2 table), controlling q-values for multiple-testing (FDR), Benjamini-Hochberg, was computed to evaluate whether this co-occurrence is statistically relevant using StatsModels v0.10 in Python v3.6.0.

Frequency of interactome genes in MS transcriptomes

The human PBMC microarray datasets analyzed in this study were published in Srinivasan papers (Srinivasan et al. 2017a; Srinivasan et al. 2017b) and deposited at EBI Array express database. The dataset was generated by Illumina Human Ref-8 v2 microarrays and contained PBMC transcriptomes of persons with MS (46 with clinically isolated syndromes, 52 with relapsing-remitting, 23 with primary-progressive, 21 with secondary-progressive disease) and 40 healthy controls.

The human brain microarray datasets were published (Magliozzi et al. 2019) and deposited at Gene Expression Omnibus. Briefly, the datasets were obtained by the Illumina whole genome HumanRef8 v2 BeadChip arrays, analyzing RNA extracted from chronic-active subpial GML (type III) and one nearby area of NAGM from the same tissue block in 10 cases of SPMS F+, 10 cases of SPMS F-, as well as in 10 control cases of non-neurological disease. The post-mortem tissues have been previously characterized for neuronal and glial alterations in cortical pathology, and for the presence of lymphoid-like immune cell aggregates (Magliozzi et al. 2010).

The enrichment of MS-AIG in DEG and in a random selection of transcript from the databases, was measured by chi-square test with Yates's correction using Graph Pad.

Enrichment of EBV interactors in MS-AIG and HSP-protein modules

HSP-protein modules were identified by Bis-Brewer (Bis-Brewer et al. 2019) through a Protein-Protein interaction criteria using 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, N.Y) 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 (Bolser et al. 2012).

Pi-index based analysis for MS-AIG

Looking at the therapeutic potential of MS-AIG, we measured their frequency in the top 150 positions of a recent published repository (Pi-Ranking) based on GWAS variants for several immunopathological traits, including MS (Fang et al. 2019). We verified whether the resulting gene lists was statistically enriched in each candidate interactome using chi-square test (Graph Pad).

Immunohistochemistry on post-mortem MS brain samples

Snap frozen tissue blocks from 10 post-mortem brains of secondary progressive MS cases (Figure 4, Table S2) and of 3 age-matched post-mortem controls (2 female/1 male) without neurological conditions have been analyzed. Brains were obtained at autopsy from the UK MS Society Tissue Bank at Imperial College, under ethical approval by the National Research Ethics Committee (08/MRE09/31). Six (*in Figure 4A) out of the 10 examined MS cases were also included in the previous study of Illumina gene expression analysis (Magliozzi et al. 2019) and in the ddPCR described below. Air dried, 10µm thick cryosections were rehydrated with PBS and immunostained with myelin oligodendrocyte glycoprotein (MOG) or major histocompatibility complex (MHC) class-II antibodies for neuropathological assessment of lesion presence and activity following previously published procedures (Magliozzi et al. 2010). Serial frozen sections, cut from the same blocks, were stained using a specific anti-CD40 rabbit polyclonal antibody (Santa Cruz). Double immunohistochemistry staining was performed combining antibodies specific for CD40 and markers MHC-II for microglia/macrophages. Antibody binding was visualised using peroxidase or alkaline phosphatase systems (Vector Labs, Peterborough, UK). For MOG immunostaining treatment with cold methanol was performed, whereas for CD40 and MHC-II cold ethanol fixation was used. Images were acquired using an Axiophot microscope (Carl Zeiss, Jena), equipped with a digital camera (Axiocam MRC) and analyzed using Axiovision 6 AC software.

Digital droplet PCR assays

RNA sample obtained from subpial cortical lesions from 6 post-mortem MS cases and from NAGM of 3 healthy controls matched and previously examined (Magliozzi et al. 2019), was further analysed to validate the CD40 gene expression. Briefly, five-hundred ng of total RNA was reverse-transcribed using Omniscript Reverse Transcription kit (Qiagen) following the manufacturer's instructions. Each sample was analyzed by a multiplex ddPCR approach according to the manufacturer's instructions, using predesigned gene expression assays for CD40 (Hs01002915_g1, FAM-labelled, Lifetechnologies) and GAPDH housekeeping gene (dHsaCPE5031597, HEX-labelled, Bio-rad). Relative quantity of CD40 versus GAPDH gene expression was evaluated using QuantaSoft software, version 1.7 (Bio-Rad, Hercules, CA). For a given sample copies/microliter were calculated by averaging over the duplicate wells.

QUANTIFICATION AND STATISTICAL ANALYSIS

The quantitation of protein levels (Figure 3) was performed by FlowJo v10.5 and expressed as median fluorescence intensity. Unpaired two-tailed T-test was used for the comparisons between groups showed in Figure 3B-E, in Figure 4A and in Figure 5B. The correlations between protein levels (Figure 3 F and G) were performed by a simple linear regression analysis. The enrichment of MS-AIG in DEG (Table 1 and 2) as well as the enrichment of MS-AIG in Pi-ranking (Figure 5) were measured by chi-square test with Yates's correction. All the above analyses were performed using Graph Pad Prism 8.

Fisher's exact test (2x2 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 StatsModels v0.10 in Python v3.6.0.

The mixed-effect multiple linear regressions analyses showed in Figure 3H and I were performed using R program, Version 1.1.456 (function "lmer" from "lme4" package). All the significances were considered when $p < 0.05$

DATA AND CODE AVAILABILITY

Gene expression data obtained from peripheral blood leukocytes and brain are available respectively at EBI Array express database and Gene Expression Omnibus. GWAS data from MS are available at IMSGC; GWAS data from other diseases are available at WTCCC2.

References

- Afrasiabi A, Parnell GP, Fewings N, Schibeci SD, Basuki MA, Chandramohan R, Zhou Y, Taylor B, Brown DA, Swaminathan S et al. 2019. Evidence from genome wide association studies implicates reduced control of epstein-barr virus infection in multiple sclerosis susceptibility. *Genome Med.* 11(1):26.
- Aloisi F, Serafini B, Magliozzi R, Howell OW, Reynolds R. 2010. Detection of epstein-barr virus and b-cell follicles in the multiple sclerosis brain: What you find depends on how and where you look. *Brain.* 133(Pt 12):e157.
- An DS, Xie Ym, Chen IS. 2001. Envelope gene of the human endogenous retrovirus herv-w encodes a functional retrovirus envelope. *J Virol.* 75(7):3488-3489.
- Ascherio A, Munger KL. 2016. Epidemiology of multiple sclerosis: From risk factors to prevention-an update. *Semin Neurol.* 36(2):103-114.
- Axtell RC, Raman C, Steinman L. 2011. Interferon- β exacerbates th17-mediated inflammatory disease. *Trends Immunol.* 32(6):272-277.
- Badam TV, de Weerd HA, Martínez-Enguita D, Olsson T, Alfredsson L, Kockum I, Jagodic M, Lubovac-Pilav Z, Gustafsson M. 2021. A validated generally applicable approach using the systematic assessment of disease modules by gwas reveals a multi-omic module strongly associated with risk factors in multiple sclerosis. *bioRxiv.2020.2010.2026.351783.*
- Baird DA, Liu JZ, Zheng J, Sieberts SK, Perumal T, Elsworth B, Richardson TG, Chen CY, Carrasquillo MM, Allen M et al. 2021. Identifying drug targets for neurological and psychiatric disease via genetics and the brain transcriptome. *PLoS Genet.* 17(1):e1009224.
- Baker D, Pryce G, Amor S, Giovannoni G, Schmierer K. 2018. Learning from other autoimmunities to understand targeting of b cells to control multiple sclerosis. *Brain.* 141(10):2834-2847.
- Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, Cotsapas C, Shah TS, Spencer C, Booth D, Goris A et al. 2013. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet.* 45(11):1353-1360.
- Berger JR, Sheremata WA, Resnick L, Atherton S, Fletcher MA, Norenberg M. 1989. Multiple sclerosis-like illness occurring with human immunodeficiency virus infection. *Neurology.* 39(3):324-329.
- Bis-Brewer DM, Danzi MC, Wuchty S, Züchner S. 2019. A network biology approach to unraveling inherited axonopathies. *Sci Rep.* 9(1):1692.
- Bolser DM, Chibon PY, Palopoli N, Gong S, Jacob D, Del Angel VD, Swan D, Bassi S, González V, Suravajhala P et al. 2012. Metabase--the wiki-database of biological databases. *Nucleic Acids Res.* 40(Database issue):D1250-1254.
- Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S et al. 2012. Annotation of functional variation in personal genomes using regulomedb. *Genome Res.* 22(9):1790-1797.
- Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: From polygenic to omnigenic. *Cell.* 169(7):1177-1186.
- Calderone A, Licata L, Cesareni G. 2015. Virusmentha: A new resource for virus-host protein interactions. *Nucleic Acids Res.* 43(Database issue):D588-592.
- Califano A, Butte AJ, Friend S, Ideker T, Schadt E. 2012. Leveraging models of cell regulation and gwas data in integrative network-based association studies. *Nat Genet.* 44(8):841-847.
- Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, Boucher L, Heinicke S, Chen D, Stark C, Breitkreutz A, Kolas N, O'Donnell L et al. 2015. The biogrid interaction database: 2015 update. *Nucleic Acids Res.* 43(Database issue):D470-478.

- Chen YF, Xia Y. 2019. Convergent perturbation of the human domain-resolved interactome by viruses and mutations inducing similar disease phenotypes. *PLoS Comput Biol.* 15(2):e1006762.
- Chiara M, Manzari C, Lionetti C, Mechelli R, Anastasiadou E, Chiara Buscarinu M, Ristori G, Salvetti M, Picardi E, D'Erchia AM et al. 2016. Geographic population structure in epstein-barr virus revealed by comparative genomics. *Genome Biol Evol.* 8(11):3284-3291.
- Consortium IMMSG. 2018. Low-frequency and rare-coding variation contributes to multiple sclerosis risk. *Cell.* 175(6):1679-1687.e1677.
- Consortium IMMSG. 2019a. A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. *Nat Commun.* 10(1):2236.
- Consortium IMMSG. 2019b. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science.* 365(6460).
- Consortium WTCC. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 447(7145):661-678.
- Davey Smith G, Hemani G. 2014. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 23(R1):R89-98.
- Davies NM, Holmes MV, Davey Smith G. 2018. Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ.* 362:k601.
- de Chassey B, Navratil V, Tafforeau L, Hiet MS, Aublin-Gex A, Agaugué S, Meiffren G, Pradezynski F, Faria BF, Chantier T et al. 2008. Hepatitis c virus infection protein network. *Mol Syst Biol.* 4:230.
- Elliott LT, Sharp K, Alfaro-Almagro F, Shi S, Miller KL, Douaud G, Marchini J, Smith SM. 2018. Genome-wide association studies of brain imaging phenotypes in uk biobank. *Nature.* 562(7726):210-216.
- Ersing I, Nobre L, Wang LW, Soday L, Ma Y, Paulo JA, Narita Y, Ashbaugh CW, Jiang C, Grayson NE et al. 2017. A temporal proteomic map of epstein-barr virus lytic replication in b cells. *Cell Rep.* 19(7):1479-1493.
- Fang H, De Wolf H, Knezevic B, Burnham KL, Osgood J, Sanniti A, Lledó Lara A, Kasela S, De Cesco S, Wegner JK et al. 2019. A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat Genet.* 51(7):1082-1091.
- Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shores N, Whitton H, Ryan RJ, Shishkin AA et al. 2015. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature.* 518(7539):337-343.
- Field J, Shahjani F, Schibeci S, Johnson L, Gresle M, Laverick L, Parnell G, Stewart G, McKay F, Kilpatrick T et al. 2015. The ms risk allele of cd40 is associated with reduced cell-membrane bound expression in antigen presenting cells: Implications for gene function. *PLoS One.* 10(6):e0127080.
- Fischbach MA. 2018. Microbiome: Focus on causation and mechanism. *Cell.* 174(4):785-790.
- Freund MK, Burch KS, Shi H, Mancuso N, Kichaev G, Garske KM, Pan DZ, Miao Z, Mohlke KL, Laakso M et al. 2018. Phenotype-specific enrichment of mendelian disorder genes near gwas regions across 62 complex traits. *Am J Hum Genet.* 103(4):535-552.
- Gallagher MD, Chen-Plotkin AS. 2018. The post-gwas era: From association to function. *Am J Hum Genet.* 102(5):717-730.
- Gao H, Danzi MC, Choi CS, Taherian M, Dalby-Hansen C, Ellman DG, Madsen PM, Bixby JL, Lemmon VP, Lambertsen KL et al. 2017. Opposing functions of microglial and macrophagic tnfr2 in the pathogenesis of experimental autoimmune encephalomyelitis. *Cell Rep.* 18(1):198-212.

- Genovese MC, Chakravarty EF, Krishnan E, Moreland LW. 2004. A randomized, controlled trial of interferon-beta-1a (avonex(r)) in patients with rheumatoid arthritis: A pilot study [isrctn03626626]. *Arthritis Res Ther.* 6(1):R73-R77.
- Gerritse K, Laman JD, Noelle RJ, Aruffo A, Ledbetter JA, Boersma WJ, Claassen E. 1996. Cd40-cd40 ligand interactions in experimental allergic encephalomyelitis and multiple sclerosis. *Proc Natl Acad Sci U S A.* 93(6):2499-2504.
- González-Duarte A, Ramirez C, Pinales R, Sierra-Madero J. 2011. Multiple sclerosis typical clinical and mri findings in a patient with hiv infection. *J Neurovirol.* 17(5):504-508.
- González-Vacarezza N, Alemán A, González G, Pérez A. 2014. Rituximab and tocilizumab for the treatment of rheumatoid arthritis. *Int J Technol Assess Health Care.* 30(3):282-288.
- Gregory AP, Dendrou CA, Attfield KE, Haghikia A, Xifara DK, Butter F, Poschmann G, Kaur G, Lambert L, Leach OA et al. 2012. Tnf receptor 1 genetic risk mirrors outcome of anti-tnf therapy in multiple sclerosis. *Nature.* 488(7412):508-511.
- Gulbahce N, Yan H, Dricot A, Padi M, Byrdsong D, Franchi R, Lee DS, Rozenblatt-Rosen O, Mar JC, Calderwood MA et al. 2012. Viral perturbations of host networks reflect disease etiology. *PLoS Comput Biol.* 8(6):e1002531.
- Gustafsson M, Edström M, Gaweł D, Nestor CE, Wang H, Zhang H, Barrenäs F, Tojo J, Kockum I, Olsson T et al. 2014. Integrated genomic and prospective clinical studies show the importance of modular pleiotropy for disease susceptibility, diagnosis and treatment. *Genome Med.* 6(2):17.
- Harley JB, Chen X, Pujato M, Miller D, Maddox A, Forney C, Magnusen AF, Lynch A, Chetal K, Yukawa M et al. 2018. Transcription factors operate across disease loci, with ebna2 implicated in autoimmunity. *Nat Genet.* 50(5):699-707.
- Hassani A, Corboy JR, Al-Salam S, Khan G. 2018. Epstein-barr virus is present in the brain of most cases of multiple sclerosis and may engage more than just b cells. *PLoS One.* 13(2):e0192109.
- Holmans P, Green EK, Pahwa JS, Ferreira MA, Purcell SM, Sklar P, Owen MJ, O'Donovan MC, Craddock N, Consortium WTC-C. 2009. Gene ontology analysis of gwa study data sets provides insights into the biology of bipolar disorder. *Am J Hum Genet.* 85(1):13-24.
- Houlston RS, Cheadle J, Dobbins SE, Tenesa A, Jones AM, Howarth K, Spain SL, Broderick P, Domingo E, Farrington S et al. 2010. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat Genet.* 42(11):973-977.
- Huttlin EL, Bruckner RJ, Navarrete-Perea J, Cannon JR, Baltier K, Gebreab F, Gygi MP, Thornock A, Zarraga G, Tam S et al. 2021. Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell.*
- James RE, Schalks R, Browne E, Eleftheriadou I, Munoz CP, Mazarakis ND, Reynolds R. 2020. Persistent elevation of intrathecal pro-inflammatory cytokines leads to multiple sclerosis-like cortical demyelination and neurodegeneration. *Acta Neuropathol Commun.* 8(1):66.
- Jelicic I, Al Nimer F, Wang J, Lentsch V, Planas R, Madjovski A, Ruhrmann S, Faigle W, Frauenknecht K, Pinilla C et al. 2018. Memory b cells activate brain-homing, autoreactive cd4. *Cell.* 175(1):85-100.e123.
- Jia X, Madireddy L, Caillier S, Santaniello A, Esposito F, Comi G, Stuve O, Zhou Y, Taylor B, Kilpatrick T et al. 2018. Genome sequencing uncovers phenocopies in primary progressive multiple sclerosis. *Ann Neurol.* 84(1):51-63.
- Johnson RT. 1994. The virology of demyelinating diseases. *Ann Neurol.* 36 Suppl:S54-60.
- Joshi P, Greco TM, Guise AJ, Luo Y, Yu F, Nesvizhskii AI, Cristea IM. 2013. The functional interactome landscape of the human histone deacetylase family. *Mol Syst Biol.* 9:672.
- Jäckle K, Zeis T, Schaeren-Wiemers N, Junker A, van der Meer F, Kramann N, Stadelmann C, Brück W. 2020. Molecular signature of slowly expanding lesions in progressive multiple sclerosis. *Brain.* 143(7):2073-2088.

- Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K et al. 2011. Global landscape of hiv-human protein complexes. *Nature*. 481(7381):365-370.
- Küry P, Nath A, Créange A, Dolei A, Marche P, Gold J, Giovannoni G, Hartung HP, Perron H. 2018. Human endogenous retroviruses in neurological diseases. *Trends Mol Med*. 24(4):379-394.
- Lamontagne RJ, Soldan SS, Su C, Wiedmer A, Won KJ, Lu F, Goldman AR, Wickramasinghe J, Tang HY, Speicher DW et al. 2021. A multi-omics approach to epstein-barr virus immortalization of b-cells reveals ebna1 chromatin pioneering activities targeting nucleotide metabolism. *PLoS Pathog*. 17(1):e1009208.
- Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. 2016. Fast and rigorous computation of gene and pathway scores from snp-based summary statistics. *PLoS Comput Biol*. 12(1):e1004714.
- Lassmann H, Niedobitek G, Aloisi F, Middeldorp JM, Group NEW. 2011. Epstein-barr virus in the multiple sclerosis brain: A controversial issue--report on a focused workshop held in the centre for brain research of the medical university of vienna, austria. *Brain*. 134(Pt 9):2772-2786.
- Leibovitch EC, Caruso B, Ha SK, Schindler MK, Lee NJ, Luciano NJ, Billioux BJ, Guy JR, Yen C, Sati P et al. 2018. Herpesvirus trigger accelerates neuroinflammation in a nonhuman primate model of multiple sclerosis. *Proc Natl Acad Sci U S A*. 115(44):11292-11297.
- Li J, Li X, Zhang S, Snyder M. 2019. Gene-environment interaction in the era of precision medicine. *Cell*. 177(1):38-44.
- Li S, Wang L, Berman M, Kong YY, Dorf ME. 2011. Mapping a dynamic innate immunity protein interaction network regulating type i interferon production. *Immunity*. 35(3):426-440.
- Lindblom RP, Berg A, Ström M, Aeinehband S, Dominguez CA, Al Nimer F, Abdelmagid N, Heinig M, Zelano J, Harnesk K et al. 2015. Complement receptor 2 is up regulated in the spinal cord following nerve root injury and modulates the spinal cord response. *J Neuroinflammation*. 12:192.
- Magliozzi R, Howell OW, Durrenberger P, Aricò E, James R, Cruciani C, Reeves C, Roncaroli F, Nicholas R, Reynolds R. 2019. Meningeal inflammation changes the balance of tn timer signalling in cortical grey matter in multiple sclerosis. *J Neuroinflammation*. 16(1):259.
- Magliozzi R, Howell OW, Reeves C, Roncaroli F, Nicholas R, Serafini B, Aloisi F, Reynolds R. 2010. A gradient of neuronal loss and meningeal inflammation in multiple sclerosis. *Ann Neurol*. 68(4):477-493.
- Martin D, Pantoja C, Fernández Miñán A, Valdes-Quezada C, Moltó E, Matesanz F, Bogdanović O, de la Calle-Mustienes E, Domínguez O, Taher L et al. 2011. Genome-wide ctf distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat Struct Mol Biol*. 18(6):708-714.
- Mechelli R, Manzari C, Policano C, Annese A, Picardi E, Umeton R, Fornasiero A, D'Erchia AM, Buscarinu MC, Agliardi C et al. 2015. Epstein-barr virus genetic variants are associated with multiple sclerosis. *Neurology*. 84(13):1362-1368.
- Mechelli R, Umeton R, Policano C, Annibali V, Coarelli G, Ricigliano VA, Vittori D, Fornasiero A, Buscarinu MC, Romano S et al. 2013. A "Candidate-interactome" Aggregate analysis of genome-wide association data in multiple sclerosis. *PLoS One*. 8(5):e63300.
- Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, Barabási AL. 2015. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science*. 347(6224):1257601.
- Milo R. 2019. Therapies for multiple sclerosis targeting b cells. *Croat Med J*. 60(2):87-98.
- Mirrashidi KM, Elwell CA, Verschueren E, Johnson JR, Frando A, Von Dollen J, Rosenberg O, Gulbahce N, Jang G, Johnson T et al. 2015. Global mapping of the inc-human interactome reveals that retromer restricts chlamydia infection. *Cell Host Microbe*. 18(1):109-121.
- Morandi E, Jagessar SA, 't Hart BA, Gran B. 2017. Ebv infection empowers human b cells for autoimmunity: Role of autophagy and relevance to multiple sclerosis. *J Immunol*. 199(2):435-448.

- Müller-Schiffmann A, Trossbach SV, Lingappa VR, Korth C. 2021. Viruses as 'truffle hounds': Molecular tools for untangling brain cellular pathology. *Trends Neurosci.* 44(5):352-365.
- Nelson AG, Zhang X, Ganapathi U, Szekely Z, Flexner CW, Owen A, Sinko PJ. 2015. Drug delivery strategies and systems for hiv/aids pre-exposure prophylaxis and treatment. *J Control Release.* 219:669-680.
- Onouchi Y, Ozaki K, Burns JC, Shimizu C, Terai M, Hamada H, Honda T, Suzuki H, Suenaga T, Takeuchi T et al. 2012. A genome-wide association study identifies three new risk loci for kawasaki disease. *Nat Genet.* 44(5):517-521.
- Pender MP, Csurhes PA, Smith C, Douglas NL, Neller MA, Matthews KK, Beagley L, Rehan S, Crooks P, Hopkins TJ et al. 2018. Epstein-barr virus-specific t cell therapy for progressive multiple sclerosis. *JCI Insight.* 3(22).
- Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A et al. 2012. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature.* 487(7408):486-490.
- Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, Fujihara K, Havrdova E, Hutchinson M, Kappos L et al. 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the mcdonald criteria. *Ann Neurol.* 69(2):292-302.
- Ponomarev ED, Shriver LP, Dittel BN. 2006. Cd40 expression by microglial cells is required for their completion of a two-step activation process during central nervous system autoimmune inflammation. *J Immunol.* 176(3):1402-1410.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P, Consortium IS. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature.* 460(7256):748-752.
- Ramesh A, Schubert RD, Greenfield AL, Dandekar R, Loudermilk R, Sabatino JJ, Koelzer MT, Tran EB, Koshal K, Kim K et al. 2020. A pathogenic and clonally expanded b cell transcriptome in active multiple sclerosis. *Proc Natl Acad Sci U S A.* 117(37):22932-22943.
- Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, Gianniny L, Korman BD, Padyukov L, Kurreeman FA et al. 2008. Common variants at cd40 and other loci confer risk of rheumatoid arthritis. *Nat Genet.* 40(10):1216-1223.
- Readhead B, Haure-Mirande JV, Funk CC, Richards MA, Shannon P, Haroutunian V, Sano M, Liang WS, Beckmann ND, Price ND et al. 2018. Multiscale analysis of independent alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. *Neuron.* 99(1):64-82.e67.
- Ricigliano VA, Handel AE, Sandve GK, Annibali V, Ristori G, Mechelli R, Cader MZ, Salvetti M. 2015. Ebna2 binds to genomic intervals associated with multiple sclerosis and overlaps with vitamin d receptor occupancy. *PLoS One.* 10(4):e0119605.
- Ryan EJ, Marshall AJ, Magaletti D, Floyd H, Draves KE, Olson NE, Clark EA. 2002. Dendritic cell-associated lectin-1: A novel dendritic cell-associated, c-type lectin-like molecule enhances t cell secretion of il-4. *J Immunol.* 169(10):5638-5648.
- Saez-Atienzar S, Bandres-Ciga S, Langston RG, Kim JJ, Choi SW, Reynolds RH, Abramzon Y, Dewan R, Ahmed S, Landers JE et al. 2021. Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types. *Sci Adv.* 7(3).
- Sargsyan SA, Shearer AJ, Ritchie AM, Burgoon MP, Anderson S, Hemmer B, Stadelmann C, Gattenlöhner S, Owens GP, Gilden D et al. 2010. Absence of epstein-barr virus in the brain and csf of patients with multiple sclerosis. *Neurology.* 74(14):1127-1135.
- Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, Freeman C, Hunt SE et al. 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 476(7359):214-219.

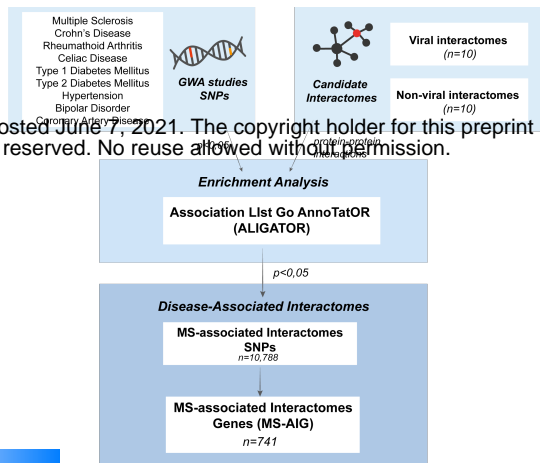
- Serafini B, Rosicarelli B, Franciotta D, Magliozzi R, Reynolds R, Cinque P, Andreoni L, Trivedi P, Salvetti M, Faggioni A et al. 2007. Dysregulated epstein-barr virus infection in the multiple sclerosis brain. *J Exp Med*. 204(12):2899-2912.
- Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, Gupta PB, Hao T, Silver SJ, Root DE et al. 2009. A physical and regulatory map of host-influenza interactions reveals pathways in h1n1 infection. *Cell*. 139(7):1255-1267.
- Smets I, Fiddes B, Garcia-Perez JE, He D, Mallants K, Liao W, Dooley J, Wang G, Humblet-Baron S, Dubois B et al. 2018. Multiple sclerosis risk variants alter expression of co-stimulatory genes in b cells. *Brain*. 141(3):786-796.
- Srinivasan S, Di Dario M, Russo A, Menon R, Brini E, Romeo M, Sangalli F, Costa GD, Rodegher M, Radaelli M et al. 2017a. Dysregulation of ms risk genes and pathways at distinct stages of disease. *Neurol Neuroimmunol Neuroinflamm*. 4(3):e337.
- Srinivasan S, Severa M, Rizzo F, Menon R, Brini E, Mechelli R, Martinelli V, Hertzog P, Salvetti M, Furlan R et al. 2017b. Transcriptional dysregulation of interferome in experimental and human multiple sclerosis. *Sci Rep*. 7(1):8981.
- Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. 2006. Biogrid: A general repository for interaction datasets. *Nucleic Acids Res*. 34(Database issue):D535-539.
- Steri M, Orrù V, Idda ML, Pitzalis M, Pala M, Zara I, Sidore C, Faà V, Floris M, Deiana M et al. 2017. Overexpression of the cytokine baf and autoimmunity risk. *N Engl J Med*. 376(17):1615-1626.
- Szymczak F, Colli ML, Mamula MJ, Evans-Molina C, Eizirik DL. 2021. Gene expression signatures of target tissues in type 1 diabetes, lupus erythematosus, multiple sclerosis, and rheumatoid arthritis. *Sci Adv*. 7(2).
- Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. 2019. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 20(8):467-484.
- Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Włodarczyk J, Ruszczycki B et al. 2015. Ctf-mediated human 3d genome architecture reveals chromatin topology for transcription. *Cell*. 163(7):1611-1627.
- Tomer Y, Concepcion E, Greenberg DA. 2002. A c/t single-nucleotide polymorphism in the region of the cd40 gene is associated with graves' disease. *Thyroid*. 12(12):1129-1135.
- van Holten J, Pavelka K, Vencovsky J, Stahl H, Rozman B, Genovese M, Kivitz AJ, Alvaro J, Nuki G, Furst DE et al. 2005. A multicentre, randomised, double blind, placebo controlled phase ii study of subcutaneous interferon beta-1a in the treatment of patients with active rheumatoid arthritis. *Ann Rheum Dis*. 64(1):64-69.
- Vazgiourakis VM, Zervou MI, Choulaki C, Bertsias G, Melissourgaki M, Yilmaz N, Sidiropoulos P, Plant D, Trouw LA, Toes RE et al. 2011. A common snp in the cd40 region is associated with systemic lupus erythematosus and correlates with altered cd40 expression: Implications for the pathogenesis. *Ann Rheum Dis*. 70(12):2184-2190.
- Veroni C, Gabriele L, Canini I, Castiello L, Coccia E, Remoli ME, Columba-Cabezas S, Aricò E, Aloisi F, Agresti C. 2010. Activation of tnfr receptor 2 in microglia promotes induction of anti-inflammatory pathways. *Mol Cell Neurosci*. 45(3):234-244.
- Vidmar L, Maver A, Drulović J, Sepčić J, Novaković I, Ristić S, Šega S, Peterlin B. 2019. Multiple sclerosis patients carry an increased burden of exceedingly rare genetic variants in the inflammasome regulatory genes. *Sci Rep*. 9(1):9171.
- Vilariño-Güell C, Zimprich A, Martinelli-Boneschi F, Herculano B, Wang Z, Matesanz F, Urcelay E, Vandenbroeck K, Leyva L, Gris D et al. 2019. Exome sequencing in multiple sclerosis families identifies 12 candidate genes and nominates biological pathways for the genesis of disease. *PLoS Genet*. 15(6):e1008180.

- Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J. 2017. 10 years of gwas discovery: Biology, function, and translation. *Am J Hum Genet.* 101(1):5-22.
- Wang H, Ye M, Fu Y, Dong A, Zhang M, Feng L, Zhu X, Bo W, Jiang L, Griffin CH et al. 2021. Modeling genome-wide by environment interactions through omnigenic interactome networks. *Cell Rep.* 35(6):109114.
- Wang J, Jelcic I, Mühlenbruch L, Haunerding V, Toussaint NC, Zhao Y, Cruciani C, Faigle W, Naghavian R, Foege M et al. 2020. Hla-dr15 molecules jointly shape an autoreactive t cell repertoire in multiple sclerosis. *Cell.* 183(5):1264-1281.e1220.
- Willis SN, Stadelmann C, Rodig SJ, Caron T, Gattenloehner S, Mallozzi SS, Roughan JE, Almendinger SE, Blewett MM, Brück W et al. 2009. Epstein-barr virus infection is not a characteristic feature of multiple sclerosis brain. *Brain.* 132(Pt 12):3318-3328.
- Wray NR, Wijmenga C, Sullivan PF, Yang J, Visscher PM. 2018. Common disease is more complex than implied by the core gene omnigenic model. *Cell.* 173(7):1573-1580.
- Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. 2009. Mirecords: An integrated resource for microrna-target interactions. *Nucleic Acids Res.* 37(Database issue):D105-110.
- Zeggini E, Gloyn AL, Barton AC, Wain LV. 2019. Translational genomics and precision medicine: Moving from the lab to the clinic. *Science.* 365(6460):1409-1413.
- Zhao B, Mar JC, Maruo S, Lee S, Gewurz BE, Johannsen E, Holton K, Rubio R, Takada K, Quackenbush J et al. 2011. Epstein-barr virus nuclear antigen 3c regulated genes in lymphoblastoid cell lines. *Proc Natl Acad Sci U S A.* 108(1):337-342.
- Zhou H, Schmidt SC, Jiang S, Willox B, Bernhardt K, Liang J, Johannsen EC, Kharchenko P, Gewurz BE, Kieff E et al. 2015. Epstein-barr virus oncoprotein super-enhancers control b cell growth. *Cell Host Microbe.* 17(2):205-216.

A

	VIRAL INTERACTOMES			NON-VIRAL INTERACTOMES		
	INTERACTOME	SIZE	SOURCE	INTERACTOME	SIZE	SOURCE
HERPESVIRUS	HHV8	123	Manually curated	AIRE	80	BioGRID
	CMV	96	Manually curated	SIRT7	658	BioGRID
	HIV	446	Experimental data	HDACs	287	Experimental data
OTHER VIRUSES	VIRORF	579	Experimental data	hulFN	113	Experimental data
	HBV	123	Manually curated	Human-miRNA targets	294	miRecord
	Polyomavirus	56	VirusMentha	Inflammasome	194	Manually curated
	H1N1	87	Experimental data	SIRT1	285	BioGRID
	HCV	202	Experimental data	VDR	92	BioGRID
	JCV	10	Manually curated			

B

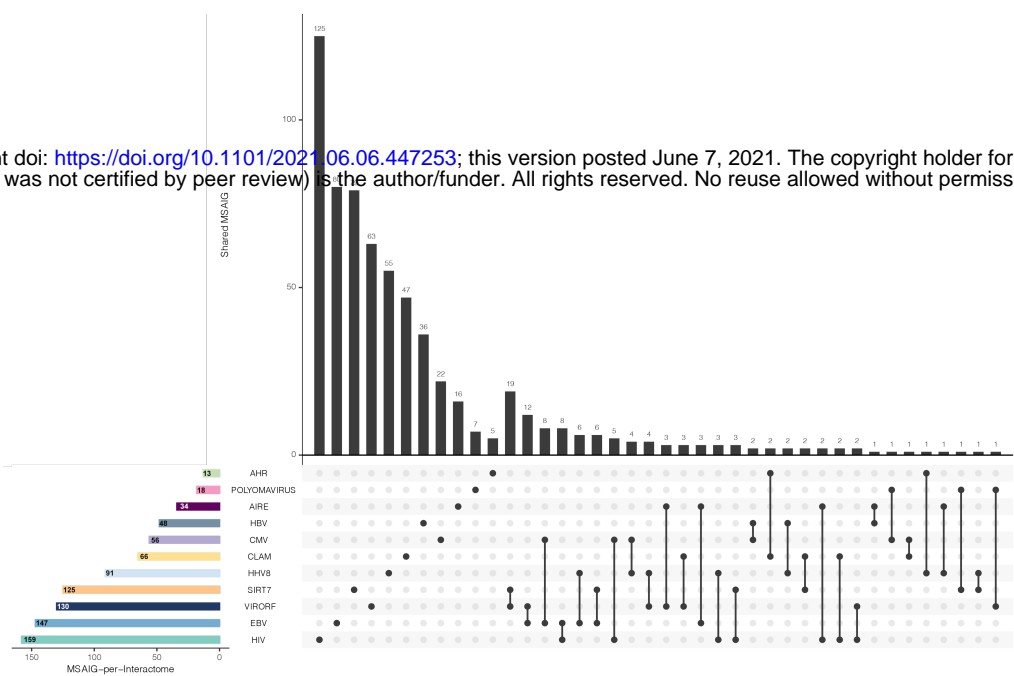


C

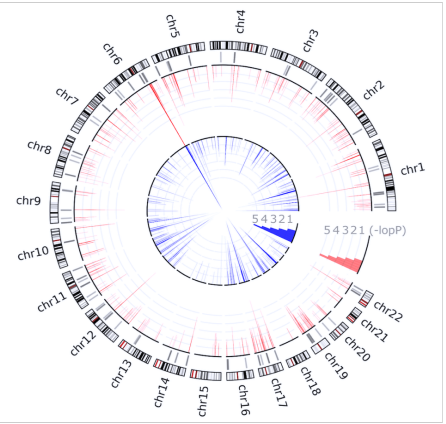
	MULTIPLE SCLEROSIS				OTHER COMPLEX DISEASES							
	GWAS 2011	GWAS 2013	METACHIP1	METACHIP2	ChD	RA	CD	T1D	T2D	HT	BD	CAD
EBV	0.0012	0.0828	0.0016	0.0004	0.4578	0.523	0.4716	0.9968	0.5444	0.1674	0.539	0.4784
HHV8	0.0046	0.2094	0.0052	0.0022	0.8732	0.6424	0.5574	0.3588	0.0214	0.7842	0.6578	0.3154
CMV	0.0444	0.6720	0.0482	0.0298	0.2198	0.5126	0.0704	0.9056	0.3438	0.6956	0.5378	0.0288
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
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
hulFN	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

A

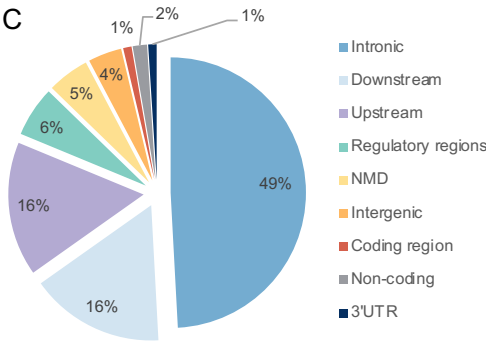
bioRxiv preprint doi: <https://doi.org/10.1101/2021.06.06.447253>; this version posted June 7, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



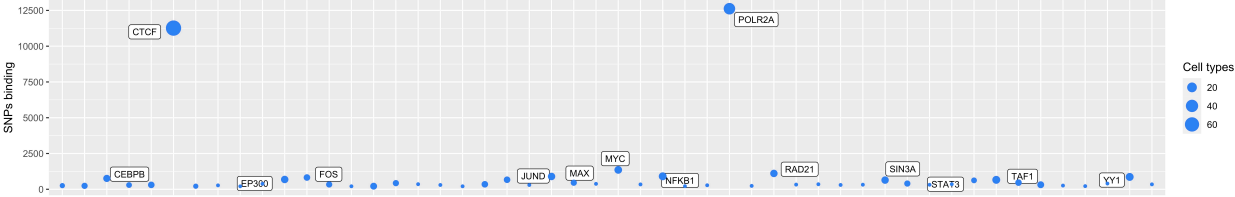
B

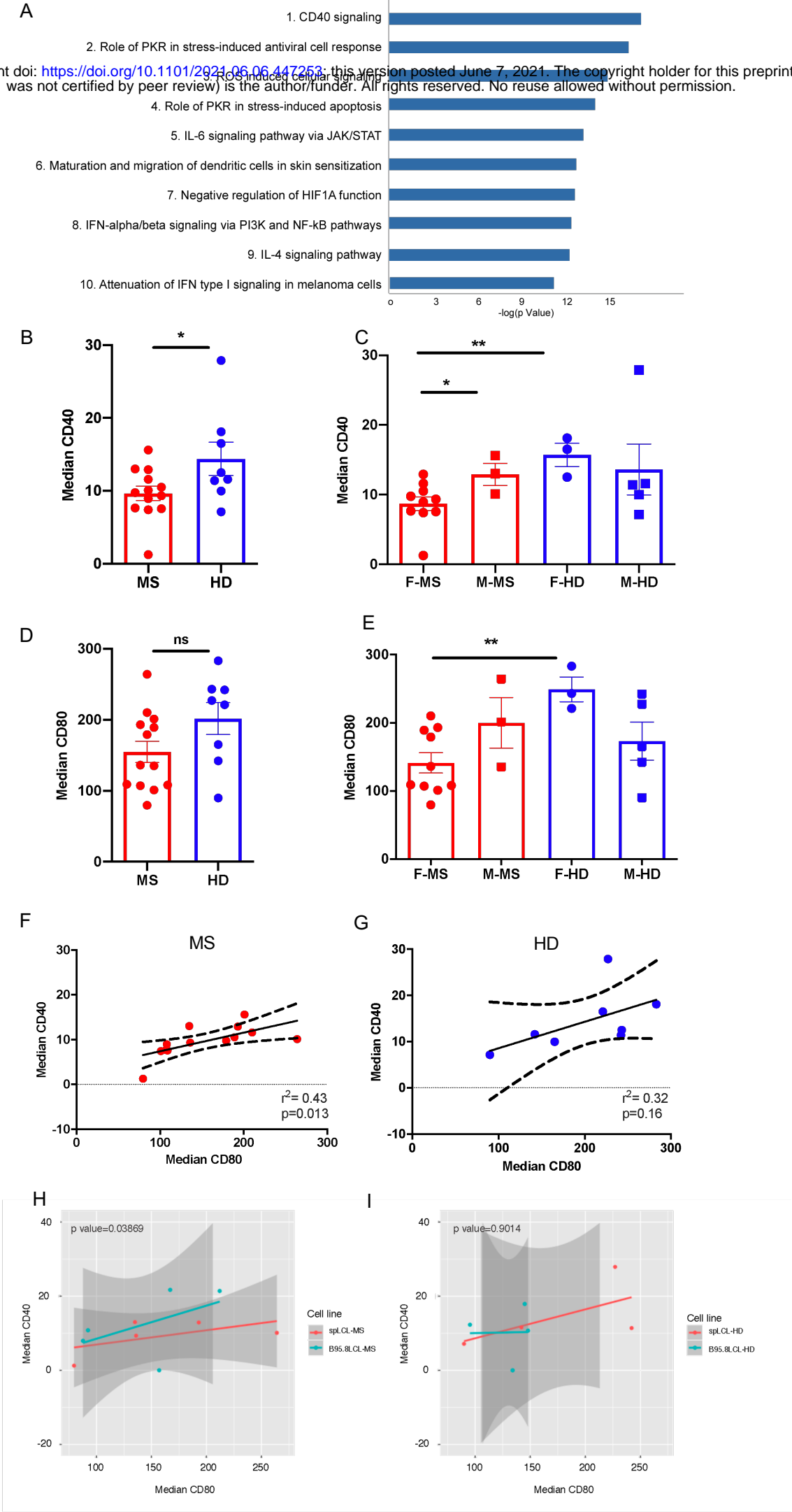


C

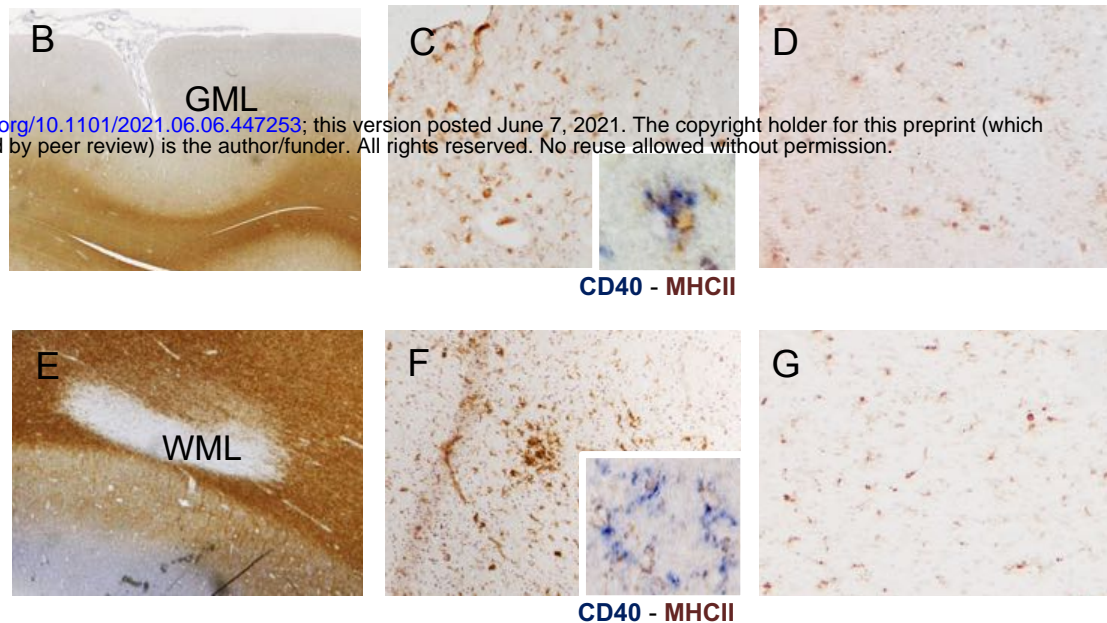
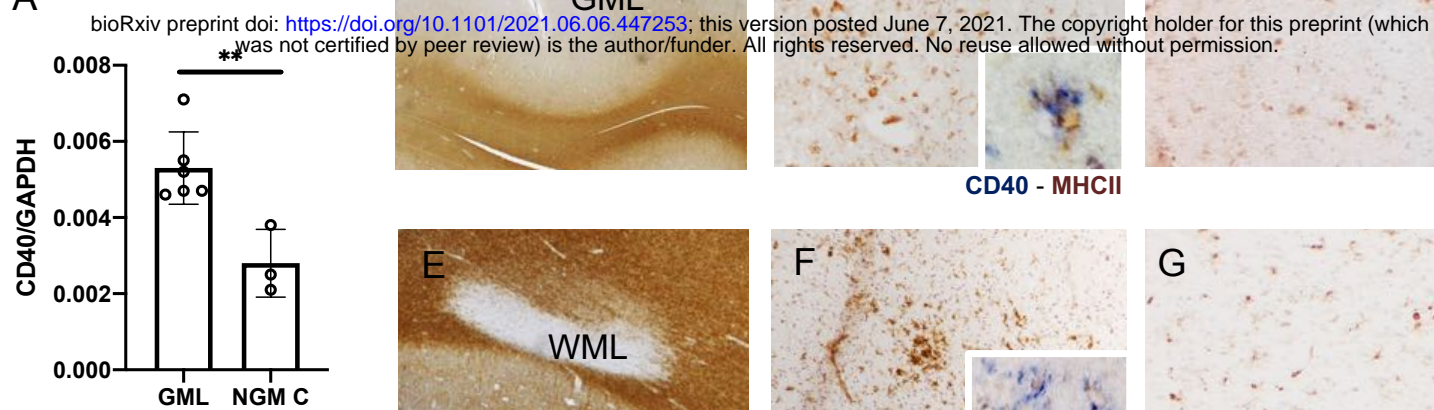


D



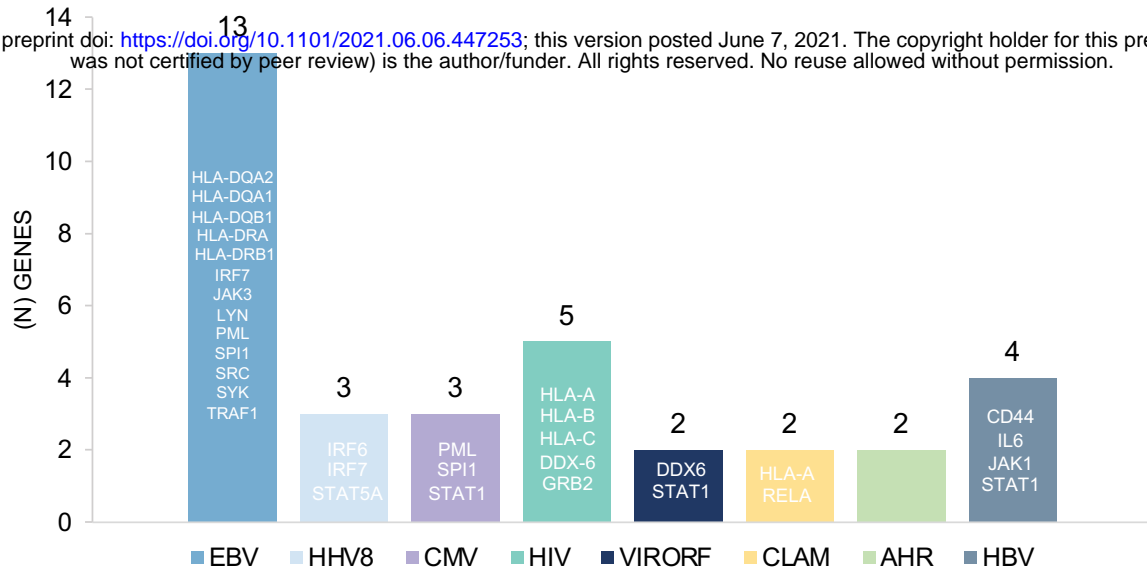


A

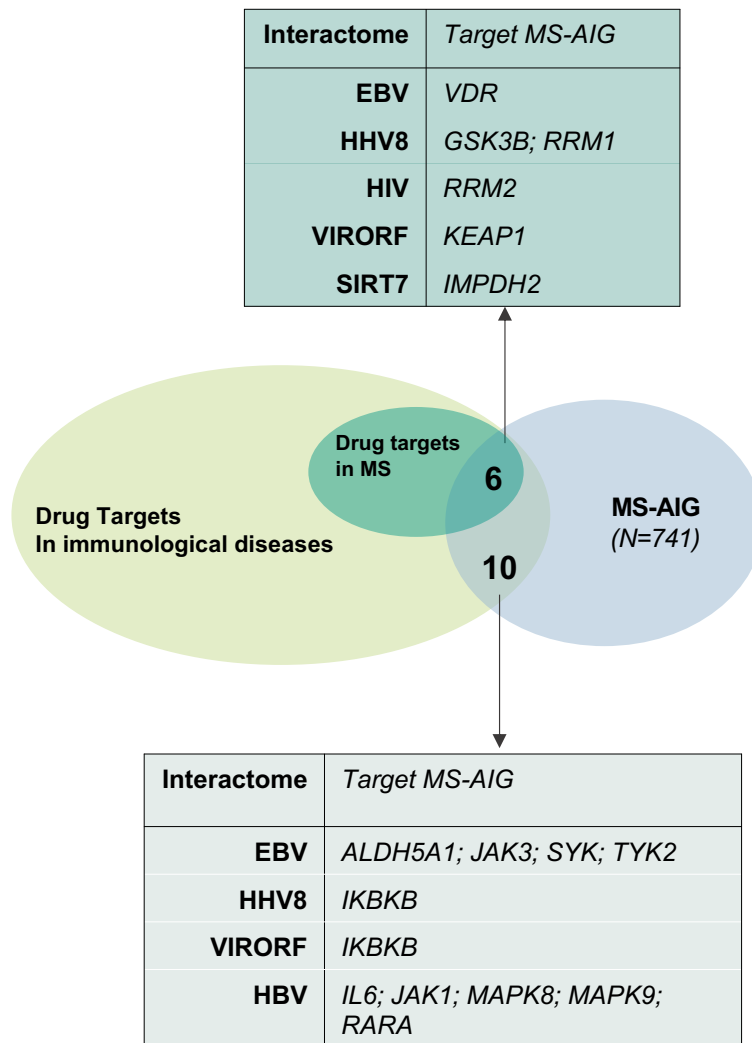


A

bioRxiv preprint doi: <https://doi.org/10.1101/2021.06.06.447253>; this version posted June 7, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



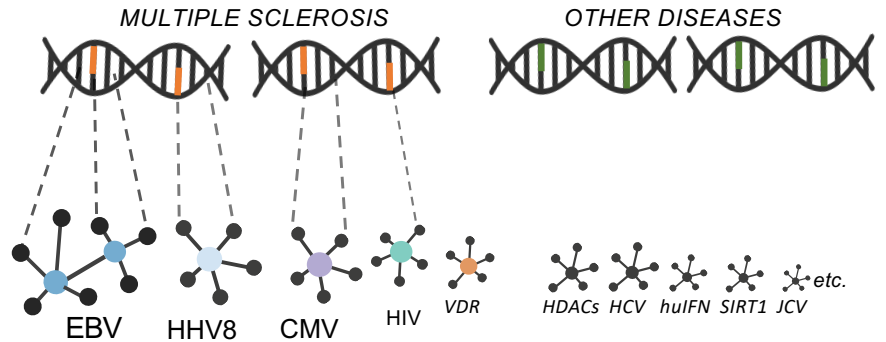
B



INTERACTOME ENRICHMENT OF ASSOCIATION

GENETIC
PREDISPOSITION

INTERACTOMES



BLOOD TRANSCRIPTOME



CLINICAL PHENOTYPES

CIS	RR-MS	SP-MS	PP-MS
			 EBV

CNS TRANSCRIPTOME



PATHOLOGY PHENOTYPES

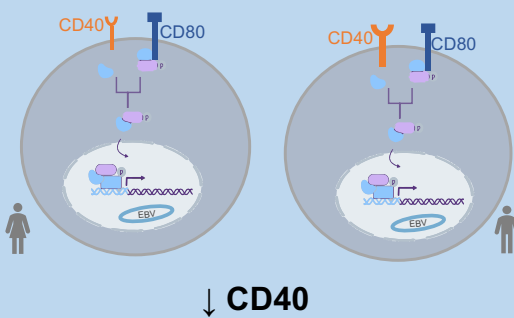
NAGM-F+	NAGM-F-	GML-F+	GML-F-
 EBV	 EBV	 EBV	 EBV

DIFFERENTIALLY
EXPRESSED
GENES

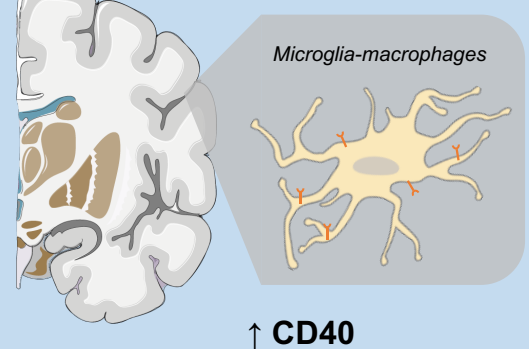
CD40 PATHWAY

IN VITRO
VALIDATION
OF PATHWAY
ANALYSIS

MS PATIENTS' SP-LCLs



MS BRAIN SAMPLES



INTERACTOME ENRICHMENT OF HIGH PRIORITY DRUG TARGETS

Genetics-led list of drug targets

