EBNA2 Binds to Genomic Intervals Associated with Multiple Sclerosis and Overlaps with Vitamin D Receptor Occupancy

Vito A. G. Ricigliano[1,2], Adam E. Handel[3,4], Geir K. Sandve[5], Viviana Annibali[6], Giovanni Ristori[6], Rosella Mechelli[6]*, M. Zameel Cader[1], Marco Salvetti[6]

1 Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom, 2 Neuroimmunology Unit, Fondazione Santa Lucia (I.R.C.C.S.), Rome, Italy, 3 Medical Research Council Functional Genomics Unit and Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, United Kingdom, 4 Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom, 5 Department of Informatics, University of Oslo, Blindern, Norway, 6 Centre for Experimental Neurological Therapies (CENTERS), Neurology and Department of Neuroscience, Mental Health and Sensory Organs, Faculty of Medicine and Psychology, "Sapienza" University of Rome, Rome, Italy

* rosella.mechelli@uniroma1.it

Abstract

Epstein-Barr virus (EBV) is a non-heritable factor that associates with multiple sclerosis (MS). However its causal relationship with the disease is still unclear. The virus establishes a complex co-existence with the host that includes regulatory influences on gene expression. Hence, if EBV contributes to the pathogenesis of MS it may do so by interacting with disease predisposing genes. To verify this hypothesis we evaluated EBV nuclear antigen 2 (EBNA2, a protein that recent works by our and other groups have implicated in disease development) binding inside MS associated genomic intervals. We found that EBNA2 binding occurs within MS susceptibility sites more than expected by chance (factor of observed *vs* expected overlap [O/E] = 5.392-fold, p < 2.0e-05). This remains significant after controlling for multiple genomic confounders. We then asked whether this observation is significant per se or should also be viewed in the context of other disease relevant gene-environment interactions, such as those attributable to vitamin D. We therefore verified the overlap between EBNA2 genomic occupancy and vitamin D receptor (VDR) binding sites. EBNA2 shows a striking overlap with VDR binding sites (O/E = 96.16-fold, p < 2.0e-05), even after controlling for the chromatin accessibility state of shared regions (p < 0.001). Furthermore, MS susceptibility regions are preferentially targeted by both EBNA2 and VDR than by EBNA2 alone (enrichment difference = 1.722-fold, p = 0.0267). Taken together, these findings demonstrate that EBV participates in the gene-environment interactions that predispose to MS.

Introduction

Epstein-Barr virus (EBV) is associated with the development of multiple sclerosis (MS), as supported by epidemiological surveys, serological evidences and other experimental laboratory

based studies [1–6]. However, its exact contribution to the pathogenesis of MS remains unclear.

Investigating the possible interplay between EBV and genetic susceptibility factors identified by genome-wide association studies (GWAS) may help clarify its possible causative role. Recently, using a "candidate interactome" approach, our group has shown that the interaction between disease-associated genetic variants and EBV proteins is relevant for MS etiology [7]. Viral proteins are able to bind DNA by means of direct [8] or indirect [9] interactions, and localize to promoter and enhancer regions [10] in infected cells, thus affecting transcriptional programs [11].

Among EBV nuclear proteins, Epstein-Barr nuclear antigen 2 (EBNA2) plays a key role in B cell transformation. Furthermore, some evidences suggest its potential implication in MS pathology: EBNA2 expressing cells have been observed in affected brains [12] and specific EBNA2 genotypes associate with disease status [13]. However, so far no analyses have been performed to define whether EBNA2 preferentially controls the expression of genes associated with MS susceptibility.

To fill this gap, using publicly available ChIP-seq data of lymphoblastoid cell lines (LCLs), we characterized the genomic distribution of EBNA2—and of its cognate factor recombination signal binding protein for immunoglobulin kappa J region (RBPJ)[9]—within disease-specific regions. We then investigated its potential interplay with vitamin D, another major environmental risk factor for MS, whose receptor (VDR) preferentially interacts with disease susceptibility regions in LCLs [14] and other immune cell types [15]. To study their relation on a genomic scale, we looked at the overlap between EBNA2 distribution and VDR occupancy both inside and outside of MS associated regions.

Methods

MS without major histocompatibility complex (MHC) single nucleotide polymorphisms (SNPs) were obtained from the latest published list of GWAS variants [16]; MS MHC SNPs included 2 Class II SNPs (rs3135388 and rs3135391) and 20 Class I SNPs contributing to MS risk independently of the HLA-DRB1*15:01 [17]. For the other disorders (obesity, UC, RA and SLE), we collected disease-associated SNPs from the 'Catalog of Published Genome-Wide Association Studies' (http://www.genome.gov/gwastudies/)[18]. Only associations reported from GWAS with replication cohort were selected. Associated regions were defined as the genomic intervals of 25, 50, 100 and 150 kb centered on the MS associated SNPs. Genomic interval of 100 kb were considered for other disease- associated SNPs. GM12878 DNase I hypersensitive sites (DHSs) track and immune response gene track were extracted from ENCODE data [19] and the HyperBrowser track repository [20].

EBNA2 and RBPJ ChIP-seq raw data were obtained from a previous publication [9] and processed on the Lifeportal platform (https://lifeportal.uio.no). We used BWA [21] for Illumina to map sequences to the reference genome hg19 and MACS (Model-based analysis of ChIP-Seq) [22] for peak-calling. MACS parameters were defined as follows: default genome size, 26 bp tag size, 300 bp band width and a p-value cutoff for peak detection of 1e-05. Peaks with a false discovery rate (FDR) of <1% were selected. VDR genomic distribution was retrieved from the literature [23]. All the analyses on processed ChIP-seq data and disease regions were performed using the Genomic HyperBrowser (http://hyperbrowser.uio.no/hb/) [20]. Enrichment between two different tracks was determined accordingly to what has been described elsewhere [14] and was expressed as the factor of observed *vs* expected overlap (O/E), which calculates the degree of coverage by track A inside track B segments *vs* outside them.

In order to assess statistical significance, we defined a null model where the location of track A intervals varied randomly, while preserving inter-segment gaps distribution and segment lengths. Track B intervals were kept fixed. Overlap analysis between the tracks was measured on real data, as well as on 50.000 Monte Carlo samples from the null model. Statistical significance was taken at p<0.05.

To estimate the overlap of ChIP-seq sites (EBNA2 or RBPJ) and MS associated regions relative to that measured for other disorder-associated regions or for immune response gene regions, we generated case–control tracks for each disease pair (MS track-non MS track) and for the MS-immune response gene pair by discarding the overlapping segments between tracks and marking the remaining intervals as case (MS specific intervals) or control (other disorder specific intervals, immune response gene specific intervals). We then calculated the fold enrichment difference for the overlap with ChIP-seq regions between MS associated intervals *vs* control intervals.

To control for any confounding effects of chromatin state ('open' *vs* 'closed') on EBNA2-MS, EBNA2-VDR and RBPJ-MS relations, we adjusted for confounding track GM12878 DHSs, using the strategy for confounder handling previously described [24]. In the statistical hypothesis testing, MS regions (or VDR segments, for the EBNA2-VDR relations) were kept fixed.

Plots of VDR, RBPJ and DHSs peak distribution around EBNA2 peak position were drawn using Homer (http://homer.salk.edu/homer/ngs/quantification.html).

Prediction of function of *cis*-regulatory regions for EBNA2, VDR and the joint overlap track EBNA2-VDR was performed using GREAT [25].

For binding motif analysis, we first enlarged the SNP position track (including both confirmed MS SNPs and those in LD [$r^2 \geq 0.8$] with them, obtained from SNAP [26] on CEU individuals) of 100 bp on each side, in order to intersect it with RBPJ ChIP-seq peaks. We then used MEME and FIMO tools (http://meme.nbcr.net/meme/intro.html) to screen for potential disruptive effects of MS variants on the RBPJ motif.

Results

We performed an overlap analysis of EBNA2 and RBPJ ChIP-seq peaks and found a 85.6% EBNA2-RBPJ joint occupancy at the base pair (bp) level, calculated on the overall number of bp covered by EBNA2 peaks (S1A Table). This result confirmed literature data [9], thus indicating the reliability of our method. We also measured the enrichment (i.e., degree of coverage by a track within the segments of a second track *vs* outside them) of EBNA2 peaks inside RBPJ binding sites and demonstrated a high co-localization of EBNA2 and RBPJ (factor of observed *vs* expected overlap [O/E] = 431.4, p < 2.0e-05, S1B Table).

Specifically looking at the distribution of EBNA2 and RBPJ in relation to MS associated regions, defined as described in Methods, we performed an enrichment analysis of their binding sites inside disease associated intervals, using two different MS tracks, one including and the other one excluding the MHC class I and class II regions. This distinction was made in order to account for any predominant effects of MHC single nucleotide polymorphisms (SNPs), which show the strongest association with the disease. We found that MS susceptibility loci are substantially enriched for both EBNA2 (O/E calculated for MS regions, MHC removed = 5.392-fold, p < 2.0e-05; MHC included = 5.047-fold, p < 2.0e-05) and RBPJ (O/E calculated for MS regions, MHC removed = 6.841-fold, p < 2.0e-05; MHC included = 6.402-fold, p < 2.0e-05) (Table 1).

To verify the sensitivity of our results with respect to a choice (SNP interval 100 kb) that is not obvious based on the literature published so far, we evaluated different SNP intervals (25,

Table 1. EBNA2/RBPJ overlap with MS regions.

ANALYSIS DESCRIPTION	enrichment (O/E)	p-value	correction for confounder track GM12878 DHSs
'RBPJ peaks' inside 'MS non-MHC regions'	6.841	2.00e-05	9.99e-04
'EBNA2 peaks' inside 'MS non-MHC regions'	5.392	2.00e-05	9.99e-04
'RBPJ peaks' inside 'MS regions including MHC'	6.402	2.00e-05	9.99e-04
'EBNA2 peaks' inside 'MS regions including MHC'	5.047	2.00e-05	9.99e-04

Results of EBNA2/RBPJ peak overlap with MS associated regions, defined as SNP position ± 50kb

O/E: factor of observed *vs* expected overlap

EBNA2 = Epstein-Barr nuclear antigen 2; RBPJ = Recombination signal-binding protein for immunoglobulin Kappa J region; DHS = DNase I hypersensitive sites; MHC = Major histocompatibility complex; MS = multiple sclerosis.

doi:10.1371/journal.pone.0119605.t001

50 and 150 kb). These analyses supported the consistency of the results (S2 Table), showing that SNP intervals do not affect the enrichment of EBNA2 and RPBJ within MS susceptibility loci.

Specific regions of intersection between EBNA2 or RBPJ peaks and MS regions/MS SNPs—both original GWAS SNPs or SNPs in linkage disequilibrium (LD)≥0.8 with them—are listed in detail in S3A–F Table.

We then assessed whether any of the SNPs intersecting RBPJ peaks induced a disruption of its binding motif, but we did not find any significant alterations (none of the SNPs showed a *q*-value <0.05, S4 Table).

A potential confounder for the EBNA2/RBPJ overlap with MS intervals is that transcription factors normally tend to bind to regions of open chromatin. To address this issue, we controlled for regions of open chromatin in LCLs, using data on DNase hypersensitivity peaks (GM12878 DHSs) from the ENCODE project [19]. This represents a reliable strategy to rule out gene proximity effects. Even when adjusting for DHSs track, the enrichment for both EBNA2 and RBPJ binding within MS susceptibility regions remained significant (p = 9.99e-04, Table 1). Another possible confounder for the results of EBNA2 overlap with MS associated sites could be the co-localization of viral proteins with immune-related genes. We indeed found an enrichment of EBNA2 in immune regions (O/E = 3.261-fold, *p* = 0.00012) (Table 2). To control for this, we normalized the EBNA2-enriched MS intervals for a general track of immune response gene regions derived from the Genomic HyperBrowser track repository [20]. We found an excess enrichment of EBNA2 binding to MS regions relatively to the immune gene track: enrichment difference (EED) = 1.603-fold p = 6.0e-05 (Table 2). This result shows that EBNA2 enrichment inside MS loci is not simplistically dependent on its tendency to interact with regions mapping to immune-related genes.

Table 2. EBNA2 peak overlap with immune response gene regions and excess enrichment of case-control analysis.

ANALYSIS DESCRIPTION	enrichment (O/E)	p-value
'EBNA2 peaks' inside 'immune response gene regions'	3.261	0.00012
Preferential overlap of 'EBNA2 peaks' on 'MS regions including MHC' (case) relatively to 'immune response gene regions' (control)	1.603 [1]	6.00e-05

O/E: factor of observed *vs* expected overlap; [1]: enrichment difference (EED).

EBNA2 = Epstein-Barr nuclear antigen 2; RBPJ = Recombination signal-binding protein for immunoglobulin Kappa J region; MHC = Major histocompatibility complex; MS = multiple sclerosis.

doi:10.1371/journal.pone.0119605.t002

Table 3. EBNA2 overlap with genomic region associated to other disorders.

ANALYSIS DESCRIPTION	enrichment (O/E)	p-value	case/control analysis [1]	p-value
'EBNA2 peaks' inside 'obesity egions'	0.457	0.8378	10.98	0.00028
'EBNA2 peaks' inside 'UC regions including MHC'	1.887	0.06522	2.888	0.03684
'EBNA2 peaks' inside 'RA regions including MHC'	4.776	9.99e-05	1.044	0.2174
'EBNA2 peaks' inside 'SLE regions including MHC'	6.4	9.99e-05	0.8339	0.3436

Results of the overlap analysis between EBNA2 peaks and other disorder-associated regions.

O/E: factor of observed *vs* expected overlap [1]: excess enrichment inside MS regions (case) relatively to control regions.

EBNA2 = Epstein-Barr nuclear antigen 2; UC = ulcerative colitis; MHC = Major histocompatibility complex; RA = Rheumatoid arthritis; SLE = Systemic lupus erythematosus.

doi:10.1371/journal.pone.0119605.t003

To verify whether our findings could be MS-specific or common to other diseases sharing with MS a multifactorial etiology, we took into account four complex disorders, characterized by a different contribution of genetic, environmental and immune factors: obesity, ulcerative colitis (UC), rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE). Specifically, RA and SLE share with MS part of the genetic background (i.e., MHC region) and the association with EBV as major environmental risk factor [27, 28]. On selected disorders, we measured EBNA2 overlap and found no significant enrichment inside obesity and UC susceptibility loci (O/E for obesity = 0.4565-fold, p = 0.8378; O/E for UC = 1.892-fold, p = 0.06522); on the contrary, RA and SLE regions showed results comparable to those seen in MS (O/E for RA, MHC included = 4.776-fold, p = 9.99e-05; O/E for SLE, MHC included = 6.4-fold, p = 9.99e-05) (Table 3); analyses performed without considering MHC regions yielded similar results (data not shown). We then performed case-control enrichment analyses using MS regions with MHC as cases and obesity, UC, RA and SLE as controls. We observed an excess enrichment of EBNA2 binding within MS associated regions relatively to control tracks only for obesity and UC (EED relative to obesity track = 10.98-fold p = 0,00028; EED relative to UC track = 2.888-fold, p = 0,03684; EED relative to RA track = 1.044-fold, p = 0.2174; EED relative to SLE track = 0.8339-fold, p = 0,3436) (Table 3). Our results show that EBNA2 preferentially binds to regions genetically associated with diseases sharing EBV as major environmental risk factor.

In the attempt to identify a more MS-specific signature in EBNA2 distribution, we then investigated a possible overlap with VDR genomic occupancy in the same cell types [23]. We first intersected EBNA2 and VDR ChIP-seq peaks and then measured their binding inside MS regions (performing the evaluation with or without the MHC region, as above): a significant enrichment (O/E without MHC = 8.667-fold, p = 8.88e-03; O/E with MHC = 8.112-fold, p = 8.64e-03) was found, even after controlling for DHSs track (Table 4). When performing a

Table 4. EBNA2-VDR joint binding inside MS and other disease-associated regions.

ANALYSIS DESCRIPTION	enrichment (O/E)	p-value	correction for confounder track GM12878 DHSs	case/control analysis [1]	p-value
'EBNA2 and VDR joint occupancy track' inside 'MS non-MHC regions'	8.667	8.88e-03	9.99e-04	1.722	0.0267
'EBNA2 and VDR joint occupancy track ' inside 'MS regions including MHC'	8.112	8.64e-03	9.99e-04	1.722	0.0253

Overlap results for EBNA2-VDR joint occupancy track within MS regions and excess enrichment relatively to EBNA2 exclusive peaks.

O/E: factor of observed *vs* expected overlap; [1]: EED for EBNA2-VDR peaks (case) relatively to control (EBNA2 exclusive peaks) inside MS regions.

EBNA2 = Epstein-Barr nuclear antigen 2; VDR = Vitamin D receptor; MS = multiple sclerosis; DHS = DNase I hypersensitive sites

doi:10.1371/journal.pone.0119605.t004

Table 5. EBNA2-VDR joint binding inside RA and SLE associated regions.

ANALYSIS DESCRIPTION	enrichment (O/E)	p-value	case/control analysis [1]	p-value
'EBNA2 and VDR joint occupancy track' inside 'RA non MHC regions'	1.29	0.3443	6.841	0.0193
'EBNA2 and VDR joint occupancy track' inside 'RA regions including MHC'	1.232	0.3426	6.679	0.0213
'EBNA2 and VDR joint occupancy track' inside 'SLE non MHC regions'	3.329	0.3849	2.366	0.1805
'EBNA2 and VDR joint occupancy track' inside 'SLE regions including MHC'	3.195	0.3776	2.321	0.1772

Overlap results for EBNA2-VDR joint occupancy track within RA and SLE regions, and excess enrichment of case-control analysis

O/E: factor of observed *vs* expected overlap; [1]: EED for EBNA2-VDR joint occupancy track on MS regions (case) relatively to control regions (RA, SLE).

EBNA2 = Epstein-Barr nuclear antigen 2; VDR = Vitamin D receptor; RA = Rheumatoid arthritis; SLE = Systemic lupus erythematosus; MHC = Major histocompatibility complex.

doi:10.1371/journal.pone.0119605.t005

case-control analysis using EBNA2-VDR joint track as case and EBNA2 as control, we found an excess enrichment of the first track inside MS intervals relatively to EBNA2 single binding (EED on MS regions with MHC = 1.722-fold, p = 0.0267) (Table 4). We then evaluated EBNA2-VDR joint overlap with RA and SLE regions, given the evidence of a role for vitamin D in these disorders. Analyzed intervals showed no significant enrichment: O/E for RA without MHC = 1.29-fold, p = 0.3443; O/E for RA with MHC = 1.232-fold, p = 0.3426; O/E for SLE without MHC = 3.329-fold, p = 0.3849; O/E for SLE with MHC = 3.195-fold, p = 0.3776. We compared these results to those observed for MS, using the same case-control strategy described above, setting MS regions as case and RA or SLE as control. We found an excess enrichment of EBNA2-VDR joint overlap inside MS regions relatively to each control track: EED relative to RA track with MHC = 6.679-fold p = 0,0213; EED relative to SLE track with MHC = 2.321-fold, p = 0,1772 (not significant) (Table 5).

These findings suggest that EBNA2-VDR joint occupancy sites are more selectively located inside MS associated regions when compared to other EBV and vitamin D-related autoimmune diseases. Furthermore, they indicate that EBNA2 and VDR partially share the occupancy sites on the DNA.

To confirm this in a broad perspective, we intersected their distribution on the genomic scale.

We observed that VDR binding sites are strikingly enriched for EBNA2 peaks (O/E = 96.16-fold, $p < 2.0e-05$;) even after controlling for DHSs track (p = 9.99e-04). EBNA2-VDR propensity to localize in similar genomic regions was confirmed by the analysis of VDR peak distribution around EBNA2 peak position (Fig. 1).

Lastly, we defined the biological meaning of genomic regions bound by EBNA2, VDR or by both using GREAT [25]. Not surprisingly, we found a strong autoimmunity signature for EBNA2-bound sites (with a q-value = 6.59e-07 for the 'altered susceptibility to autoimmune disorder' category) and several phenotypes related to immune response (e.g. B cell function, interferon gamma signaling, thymus-specific gene expression) or deregulation of pluripotency genes for the genomic intervals bound by both factors (S5 Table).

Discussion

Looking at the distribution of EBNA2 and its co-factor RBPJ in genomic intervals associated with MS, we found that both proteins tend to fall inside MS intervals more than expected by chance. This overlap is not due to a shared preference for genomic regions of accessible chromatin, as demonstrated by the significance of the analysis (p = 9.99e-04) even after taking into account the effect of confounder track DHSs. The consistency and implications of this observation are strengthened by data showing that disease-related distributions of EBNA2 binding
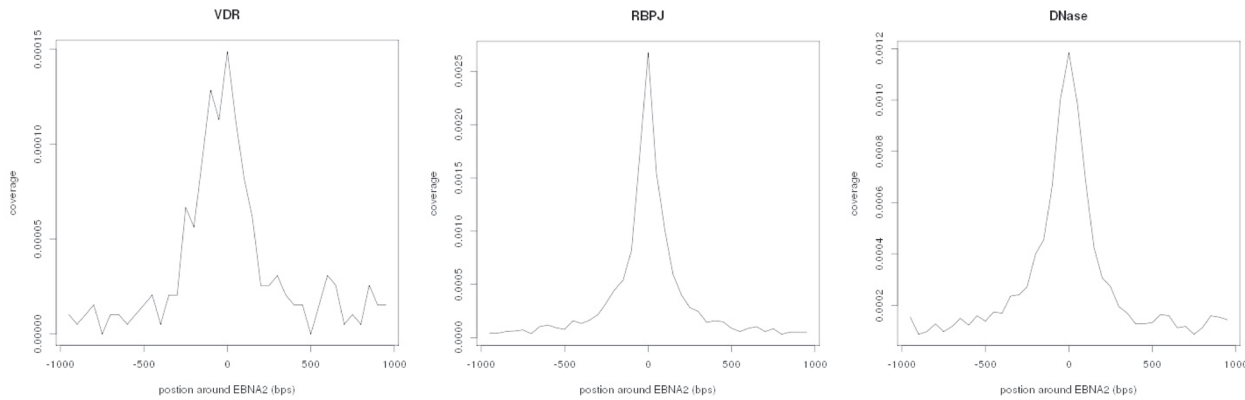
Fig 1. Histogram plots of VDR, RBPJ and DNase peak distribution centered on EBNA2 peak position. In similar analyses, when the position of two factors (e.g., co-factors) across the genome tends to coincide, the correspondent graph shows a clear middle-point peak, as seen for the relation between EBNA2 and RBPJ (plot in the middle). Refer to the text for interpretation.

doi:10.1371/journal.pone.0119605.g001

occur in EBV-associated autoimmune conditions (MS, SLE, RA), it is less evident in other immunological diseases not closely related to EBV (UC), and it disappears in non-immune disorders (obesity).

The biological significance of EBNA2 and RBPJ interaction with MS associated loci appears even more important if we consider that RBPJ may be related to MS pathology [29] and is a major downstream effector in NOTCH signaling pathway [30]. NOTCH-1 activity, which is also altered in MS, is up-regulated by Bacille Calmette-Guérin (BCG), whose protective effects have been reported in individuals with clinically isolated syndromes (CIS) and early MS [31, 32]. It has been proposed that BCG positive outcomes may act through the restoration of a normal NOTCH signaling [33], which can be perturbed by EBV infection via a RBPJ-mediated mechanism [34, 35]. These and other evidences implicate a role for EBNA2 and RBPJ in MS pathogenesis.

When looking at EBNA2 localization with respect to VDR genomic occupancy, we observed a striking overlap between EBNA2 and VDR binding sites. This finding was confirmed after controlling for DHSs track, thus suggesting that the tendency of EBNA2 and VDR to occupy similar genomic positions in B cell lines is not determined by a shared preference for accessible chromatin regions. Furthermore, intervals of joint EBNA2-VDR distribution significantly overlapped with MS regions more than observed for EBNA2 or VDR [14] separately. This result appears to be MS-specific when compared to results of the same analysis performed for RA and SLE.

The propensity of EBNA2 and VDR to have similar genomic localization in LCLs may be particularly relevant in the context of MS etiology. Based on the observed findings, it is possible to provide different functional interpretations for the relationship between the two proteins:

1. EBNA2 and VDR may compete for binding to MS-related sites and other genomic regions in immune cell subsets. This hypothesis is supported by some literature data showing that EBV and vitamin D tend to have antagonistic effects on B cell function: EBNA2 has selected its targets to drive B-lymphocyte proliferation and survival in primary infection [9], whereas vitamin D (and its analogs) preferentially down-regulate B cell function and play a prominent role in the maintenance of their homeostasis [36–38]. A major regulator at the crossing point between pro- and anti-apoptotic fate may be represented by *MYC*: its transactivation by EBNA2 causes continuous B-cell proliferation [9], while its suppression by 1,25(OH)2 D3 modulates cell growth [39, 40]. As far as MS is concerned, the potential antagonistic

activity of these two factors in disease-associated loci may reflect the opposite tendency of EBV infection and high vitamin D levels to predispose to or protect against MS, respectively.

2. EBNA2 and VDR may interact with each other inside DNA binding sites. In this scenario, it has been shown that RBPJ is the preferential—but not exclusive—co-factor used by EBNA2 to bind to DNA, therefore a potential interaction with VDR, similarly to what has been documented for another EBV-encoded protein (EBNA3, which binds to VDR and affects its normal function) [41, 42], cannot be excluded. A further element in support of the above hypothesis is that the plot of VDR peak distribution centered on EBNA2 peak position resembles the one obtained for EBNA2 relation with its well-known co-factor RBPJ (Fig. 1). These observations may lead to speculate on two alternative outcomes for the EBNA2-VDR cooperative binding: this could either result in deactivation, as described for the EBNA3-VDR interplay [41], or in a relation analogous to that of EBNA2 with RBPJ. Once again, the latter interpretation is not supported by epidemiological evidences in MS, which suggest opposing effects of EBV and vitamin D.

One limitation of the present study is that EBNA2-RBPJ or VDR genomic distributions have been characterized by separate research groups and in different cell samples, although in both cases the analyzed lines were LCLs. Therefore, a simultaneous evaluation of the EBNA2 and VDR distribution in the same samples will be required to experimentally demonstrate whether the ability to bind to similar genomic intervals occurs in a mutually exclusive or joint (co-occupancy) fashion.

Taken together, our data suggest that EBNA2 and RBPJ tend to bind to MS-related intervals more than expected by chance; EBNA2 distribution across the genome overlaps with VDR occupancy in B cell lines, suggesting that non-heritable risk factors influence each other and interact with MS susceptibility loci in pathogenically relevant immune cell subsets [43, 44].

Supporting Information

S1 Table. EBNA2 specific binding to DNA at the base pair (bp) level (A) and analysis results for 'EBNA2 peaks' inside 'RBPJ peaks' (B).
(XLS)

S2 Table. EBNA2/RBPJ peak overlap with MS associated regions, defined as SNP position interval of 25, 50kb and 150 kb.
(XLS)

S3 Table. Shared genomic intervals between EBNA2/RPBJ peaks and MS regions (defined as SNP position +/−50kb)/MS SNPs. A. Shared genomic intervals between EBNA2 peaks and MS excluding MHC regions; **B**. Shared genomic intervals between EBNA2 peaks and MS regions including MHC; **C**. Shared genomic intervals between EBNA peaks and MS SNPs (original GWAS SNPs, indicated in bold, and SNPs in LD≥0.8); **D**. Shared genomic intervals between RBPJ peaks and MS excluding MHC regions; **E**. Shared genomic intervals between RBPJ peaks and MS regions including MHC; **F**. Shared genomic intervals between RBPJ peaks and MS SNPs (original GWAS SNPs, indicated in bold, and SNPs in LD≥0.8).
(XLS)

S4 Table. RBPJ motif screening results for MS SNP (original GWAS SNPs + SNPs in LD≥0.8) sequences intersecting RBPJ peaks. MS SNP sequences were defined as follows: SNP position +/−100 bp. Original GWAS SNPs are indicated in bold.
(XLS)

S5 Table. Prediction of function for genomic regions bound by EBNA2, VDR or by both factors.
(XLS)

Author Contributions

Conceived and designed the experiments: VAGR AEH RM. Performed the experiments: VAGR AEH GKS. Analyzed the data: VAGR AEH GKS RM. Wrote the paper: VAGR AEH VA GR RM MZC MS.

References

1. Banwell B, Krupp L, Kennedy J, Tellier R, Tenembaum S, Ness J, et al. Clinical features and viral serologies in children with multiple sclerosis: a multinational observational study. Lancet Neurol. 2007; 6: 773–781. PMID: 17689148

2. Kakalacheva K, Lünemann JD. Environmental triggers of multiple sclerosis. FEBS Lett. 2011; 585: 3724–3729. doi: 10.1016/j.febslet.2011.04.006 PMID: 21486562

3. Pender MP. The essential role of Epstein-Barr virus in the pathogenesis of multiple sclerosis. Neuroscientist 2011; 17: 351–367. doi: 10.1177/1073858410381531 PMID: 21075971

4. Almohmeed YH, Avenell A, Aucott L, Vickers MA. Systematic review and meta-analysis of the sero-epidemiological association between Epstein Barr virus and multiple sclerosis. PLoS One 2013; 8: e61110. doi: 10.1371/journal.pone.0061110 PMID: 23585874

5. Simon KC, Munger KL, Ascherio A. Vitamin D and multiple sclerosis: epidemiology, immunology, and genetics. Curr Opin Neurol 2012; 25: 246–251. doi: 10.1097/WCO.0b013e3283533a7e PMID: 22547098

6. Koch MW, Metz LM, Kovalchuk O. Epigenetic changes in patients with multiple sclerosis. Nat Rev Neurol 2013; 9: 35–43. doi: 10.1038/nrneurol.2012.226 PMID: 23165337

7. Mechelli R, Umeton R, Policano C, Annibali V,Coarelli G, Ricigliano VAG, et al. A "candidate-interactome" aggregate analysis of genome-wide association data in multiple sclerosis. PLoS One 2013; 8: e63300. doi: 10.1371/journal.pone.0063300 PMID: 23696811

8. Arvey A, Tempera I, Tsai K, Chen HS, Tikhmyanova N, Klichinsky M, et al. An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host virus regulatory interactions. Cell Host Microbe 2012; 12: 233–245. doi: 10.1016/j.chom.2012.06.008 PMID: 22901543

9. Zhao B, Zou J, Wang H, Johannsen E, Peng CW, Quackenbush J, et al. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. Proc Natl Acad Sci USA 2011; 108: 14902–14907. doi: 10.1073/pnas.1108892108 PMID: 21746931

10. McClellan MJ, Wood CD, Ojeniyi O, Cooper TJ, Kanhere A, Arvey A, et al. Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. PLoS Pathog 2013; 9: e1003636. doi: 10.1371/journal.ppat.1003636 PMID: 24068937

11. Maier S, Staffler G, Hartmann A, Höck J, Henning K, Grabusic K, et al. Cellular target genes of Epstein-Barr virus nuclear antigen 2. J Virol 2006; 80: 9761–9771. PMID: 16973580

12. Serafini B, Rosicarelli B, Franciotta D, Magliozzi R, Reynolds R, Cinque P, et al. Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. J Exp Med 2007; 204: 2899–912. PMID: 17984305

13. Mechelli R, Manzari C, Policano C, Annese A, Picardi E, Umeton R et al. EpsteinBarr virus genetic variants associates with multiple sclerosis. Neurology. In press

14. Disanto G, Sandve GK, Berlanga-Taylor AJ, Ragnedda G, Morahan JM, Watson CT, et al. Vitamin D receptor binding, chromatin states and association with multiple sclerosis. Hum Mol Genet 2012; 21: 3575–3586. doi: 10.1093/hmg/dds189 PMID: 22595971

15. Handel AE, Sandve GK, Disanto G, Berlanga-Taylor AJ, Gallone G, Hanwell H, et al. Vitamin D receptor ChIP-seq in primary CD4+ cells: relationship to serum 25-hydroxyvitamin D levels and autoimmune disease. BMC Med 2013; 11: 163. doi: 10.1186/1741-7015-11-163 PMID: 23849224

16. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013; 45: 1353–1360. doi: 10.1038/ng.2770 PMID: 24076602

17. Cree BA, Rioux JD, McCauley JL, Gourraud PFD, Goyette P, McElroy J, et al. A major histocompatibility Class I locus contributes to multiple sclerosis susceptibility independently from HLA-DRB1*15:01. PLoS One 2010; 5: e11296. doi: 10.1371/journal.pone.0011296 PMID: 20593013

18. Hindorff LA, Sethupathyb P, Junkinsa HA, Ramosa EM, Mehtac JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci USA 2009; 106: 9362–9367. doi: 10.1073/pnas.0903103106 PMID: 19474294

19. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, et al. The UCSC Genome Browser database: update 2011. Nucleic Acids Res 2011; 40 (Database issue): D876–D882. doi: 10.1093/nar/gkr1011 PMID: 22075997

20. Sandve GK, Gundersen S, Johansen M, Glad IK, Gunathasan K, Holden M, et al. The Genomic Hyper-Browser: an analysis web server for genome-scale data. Nucleic Acids Res 2013; 41(Web Server issue): W133–W141. doi: 10.1093/nar/gkt342 PMID: 23632163

21. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25: 1754–1760. doi: 10.1093/bioinformatics/btp324 PMID: 19451168

22. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008; 9: R137. doi: 10.1186/gb-2008-9-9-r137 PMID: 18798982

23. Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrel A, et al. A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. Genome Res 2010; 20: 1352–1360. doi: 10.1101/gr.107920.110 PMID: 20736230

24. Sandve GK, Gundersen S, Rydbeck H, Glad IK, Holden L, Holden M, et al. The Genomic HyperBrowser: inferential genomics at the sequence level. Genome Biol 2010; 11: R121. doi: 10.1186/gb-2010-11-12-r121 PMID: 21182759

25. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010; 28: 495–501. doi: 10.1038/nbt.1630 PMID: 20436461

26. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI, et al. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008; 24: 2938–2939. doi: 10.1093/bioinformatics/btn564 PMID: 18974171

27. Lossius A, Johansen JN, Torkildsen Ø, Vartdal F, Holmøy T. Epstein-Barr virus in systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis—association and causation. Viruses 2012; 4: 3701–3730. PMID: 23342374

28. Pelajo CF, Lopez-Benitez JM, Miller LC. Vitamin D and autoimmune rheumatologic disorders. Autoimmun Rev 2010; 9: 507–510. doi: 10.1016/j.autrev.2010.02.011 PMID: 20146942

29. Querol L, Clark PL, Bailey MA, Cotsapas C, Cross AH, Hafler DA, et al. Protein array-based profiling of CSF identifies RBPJ as an autoantigen in multiple sclerosis. Neurology 2013; 81: 956–963. doi: 10.1212/WNL.0b013e3182a43b48 PMID: 23921886

30. Guruharsha KG, Kankel MW, Artavanis-Tsakonas S. The Notch signalling system: recent insights into the complexity of a conserved pathway. Nat Rev Genet 2012; 13: 654–666. doi: 10.1038/nrg3272 PMID: 22868267

31. Ristori G, Buzzi MG, Sabatini U, Giugni E, Bastianello S, Viselli F, et al. Use of Bacille Calmette Guèrin (BCG) in multiple sclerosis. Neurology 1999; 53: 1588–1589. PMID: 10534275

32. Ristori G, Romano S, Cannoni S, Visconti A, Tinelli E, Mendozzi L, et al. Effects of Bacille Calmette-Guèrin after the first demyelinating event in the CNS. Neurology 2014; 82: 41–48. doi: 10.1212/01.wnl.0000438216.93319.ab PMID: 24306002

33. Brenner SR. BCG vaccination may restore normal Notch signaling, which is disrupted in MS, possibly due to infection by Epstein-Barr virus causing auto-immunity. Re: "Effects of Bacille Calmette-Guèrin after the first demyelinating event in the CNS" Ristori, et al, Neurology 2014; 82: 41–48. doi: 10.1212/01.wnl.0000438216.93319.ab PMID: 24306002

34. Hsieh JJ, Henkel T, Salmon P, Robey E, Peterson MG. Truncated mammalian Notch1 activates CBF1/RBPJk-repressed genes by a mechanism resembling that of Epstein-Barr virus EBNA2. Mol Cell Biol 1996; 16: 952–959. PMID: 8622698

35. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. Nat Rev Cancer 2004; 4(10):757–768. PMID: 15510157

36. Chen S, Sims GP, Chen XX, Gu YY, Chen S, Lipsky PE. Modulatory effects of 1,25-dihydroxyvitamin D3 on human B cell differentiation. J Immunol 2007; 179: 1634–1647. PMID: 17641030

37. Terrier B, Derian N, Schoindre Y, Chaara W, Geri G, Zahr N, et al. Restoration of regulatory and effector T cell balance and B cell homeostasis in systemic lupus erythematosus patients through vitamin D supplementation. Arthritis Res Ther 2012; 14: R221. doi: 10.1186/ar4060 PMID: 23075451

38. Pepper C, Thomas A, Hoy T, Milligan D, Bentley P, Fegan C. The vitamin D3 analog EB1089 induces apoptosis via a p53-independent mechanism involving p38 MAP kinase activation and suppression of ERK activity in B-cell chronic lymphocytic leukemia cells in vitro. Blood 2003; 101: 2454–2460. PMID: 12446453

39. Reitsma PH, Rothberg PG, Astrin SM, Trial J, Bar-Shavit Z, Hall A, et al. Regulation of *myc* gene expression in HL-60 leukaemia cells by a vitamin D metabolite. Nature 1983; 306: 492–494. PMID: 6689055

40. Rohan JN, Weigel NL. 1α,25-dihydroxyvitamin D3 reduces c-Myc expression, inhibiting proliferation and causing G1 accumulation in C4-2 prostate cancer cells. Endocrinology 2009; 150: 2046–2054. doi: 10.1210/en.2008-1395 PMID: 19164469

41. Yenamandra SP, Hellman U, Kempkes B, Darekar SD, Petermann S, Sculley T, et al. Epstein-Barr virus encoded EBNA-3 binds to vitamin D receptor and blocks activation of its target genes. Cell Mol Life Sci 2010; 67: 4249–4256. doi: 10.1007/s00018-010-0441-4 PMID: 20593215

42. Disanto G, Meier U, Giovannoni G, Ramagopalan SV. Vitamin D: a link between Epstein-Barr virus and multiple sclerosis development? Expert Rev Neurother 2011; 11:1221–1224. doi: 10.1586/ern.11.97 PMID: 21864064

43. Franciotta D, Salvetti M, Lolli F, Serafini B, Aloisi F. B cells and multiple sclerosis. Lancet Neurol 2008; 7: 852–858. doi: 10.1016/S1474-4422(08)70192-3 PMID: 18703007

44. Hauser SL, Waubant E, Arnold DL, Vollmer T, Antel J, Fox RJ, et al. B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 2008; 358: 676–688. doi: 10.1056/NEJMoa0706383 PMID: 18272891