Title: Potential for molecular mimicry between the human endogenous retrovirus W family envelope proteins and myelin proteins in multiple sclerosis.

Authors: Ranjan Ramasamy[1*], Blessy Joseph[2], Trevor Whittall[3]

[1] ID-FISH Technology Inc., 797 San Antonio Road, Palo Alto, CA 94303, United States of America; [2]Anglia Ruskin University, East Road, Cambridge, CB1 1PT, United Kingdom; [3]Department of Applied Sciences, University of West of England, Frenchay Campus, Bristol, BS16 1QY, United Kingdom.

[*] Corresponding author (RR)

Email: rjr200911@yahoo.com

Running Title: Molecular mimicry in multiple sclerosis

Abbreviations: BLAST - Basic Local Alignment Search Tool; CNS – Central Nervous System; EAE-Experimental Autoimmune Encephalomyelitis; EBV – Epstein Barr Virus; HERV – Human Endogenous Retrovirus; IEDB – Immune Epitope Data Base; MBP – Myelin Basic Protein; MOG - Myelin Oligodendrocyte Glycoprotein; MS – Multiple Sclerosis; MSRV - Multiple Sclerosis Associated Retrovirus; NCBI – National Centre for Biological Information; PLP - Proteolipid Protein; SMM - Stabilised Matrix Method.

Abstract

Multiple sclerosis is an autoimmune disease caused by the destruction of the myelin sheath in the central nervous system by T cells and antibodies. The major target molecules are the myelin basic protein, the myelin oligodendrocyte glycoprotein and the proteolipid protein but the aetiology of the disease is as yet poorly understood. The HLA Class II allele DRB1*1501 in particular as well as DRB5*0101 and the expression of human endogenous retroviral envelope proteins have been linked to multiple sclerosis but the molecular mechanisms relating these remain to be elucidated. We hypothesised that cross-reactive peptide epitopes in the retroviral envelope proteins and myelin proteins that can be presented by the two Class II DR molecules may play a role in initiating multiple sclerosis. As an initial step to test the hypothesis, sequence homologies between retroviral envelope and myelin proteins and *in silico* predictions of peptides derived from them that are able to bind to the two Class II alleles were examined. The results support the hypothesis that molecular mimicry in peptide epitopes from envelope proteins of the HERV-W family of endogenous retroviruses and myelin proteins is possible and this may potentially trigger multiple sclerosis. Mimicry between syncytin-1, a HERV-W envelope protein that is expressed during placentation, and myelin proteins may also explain the higher prevalence of multiple sclerosis in women. Experimental confirmation of the ability of the identified peptide epitopes to activate T_H cells is a logical extension to the present findings, and can lead to new immunotherapeutic procedures to treat multiple sclerosis.

Key Words: autoimmunity, human endogenous retroviruses, molecular mimicry, multiple sclerosis, myelin proteins.

1. Introduction

Multiple sclerosis (MS) is an inflammatory autoimmune disease of the central nervous system (CNS) that involves the progressive destruction of the myelin sheath and axons resulting in neurodegeneration [1]. Antibodies and T cells are both considered to be responsible effectors [2, 3] and this is consistent with sequences of T and B cell antigen receptors in the CNS of MS patients [2]. Experimental autoimmune encephalomyelitis (EAE) in rodents mimics many aspects of MS and is considered to be a useful animal model of MS. Studies on EAE have indicated the main target molecules in myelin for the autoimmune responses in MS. EAE can be induced in naïve mice by active immunization with the myelin basic protein (MBP), proteolipid protein (PLP) or myelin oligodendrocyte glycoprotein (MOG), and peptides derived from them, or by passive transfer of immune T cells from animals with EAE [4].

Large, multi-population Genome-Wide Association Studies showed that amongst all MHC alleles, the MHC Class II allele DRB1*15:01 had the strongest association with MS [5]. DRB1*15:01 is also associated with early onset of MS [6]. DRB1*1501 is one of the alleles coding for the DR β chain. The DRB1*1501 β chain pairs with the relatively non-polymorphic HLA DR α chain from DRA*0101 to form the HLA DR2b heterodimer. Furthermore, DRB5*0101, which is in linkage disequilibrium with DRB1*1501, is also associated with MS [7, 8] and binds to DRA*0101 to form the HLA DR2a heterodimer. Both HLA DR2a and DR2b, often referred to as representing the HLA DR2(15) haplotype, have been shown present the encephalitogenic MBP peptide 83-99 to T cell clones from MS patients [7, 8]. HLA DR2a and DR2b may also present other peptide epitopes associated with MS to CD4+ T cells. Evidence suggests that CD4+ T cells with a T_H1 phenotype have a central role in the initiation of MS and its progression, the generation of autoantibodies and self-reactive CD8+ T cells, and MS-associated inflammation [9, 10]. HLA DR2a and DR2b restricted T_H1 cells are also able to directly lyse target cells through the perforin or Fas/FasL pathway [11]. Proinflammatory CD4+ T_H17 are also important in the pathogenesis of MS [12, 13].

Besides genetic factors, viral infections have been associated with MS and three mechanisms that are not mutually exclusive have been proposed to explain this [3]. Firstly, viral infection of the CNS can potentially cause inflammation and damage oligodendrocytes that produce myelin, thereby releasing myelin fragments that activate autoreactive T cells in an inflammatory milieu. Subsequent epitope spreading can produce more demyelination and axon death. Secondly, persistent viral infection of the CNS can produce inflammatory demyelinating disease caused by the immune response attempting to eliminate infected cells within the CNS. Thirdly, viral infection outside the CNS can activate cross-reactive T cells that then enter the CNS and cause inflammatory demyelinating lesions leading to MS. Accumulating evidence suggests that human endogenous retroviruses (HERVs), generally inactive remnants of exogenous retroviruses that became integrated into primate genomes, are important in the aetiology and pathogenesis of MS [14]. An MS associated retrovirus (MSRV), a member of the HERV-W family, has been particularly implicated in MS because virus particles and reverse transcriptase activity are detected in MS patients [14 - 16]. A role in MS pathogenesis has been ascribed to the MSRV *env* gene product which is nearly identical to syncytin-1 of the HERV-W family [15]. HERV-W encoded syncytin-1, itself a viral envelope protein remnant and a membrane glycoprotein, has evolved to perform an essential fusion function in forming the placental syncytiotrophoblast layer in humans [17, 18]. Syncytin-1 is highly conserved between members of the HERV-W family including MSRV [19]. Syncytin-1 is homologous to syncytin-2, another fusogenic envelope glycoprotein encoded in a different HERV family *viz*. HERV FRD, which has also evolved to play an important role in forming the syncytiotrophoblast [18]. The expression of the MSRV *env* gene product is significantly higher in brain lesions in MS plaques and correlates with the extent of active demyelination and inflammation [15, 16]. Furthermore, the observed temporal relationship between Epstein Barr Virus (EBV) infection and MS has been ascribed to activation of the expression of MSRV envelope protein/HERV-W syncytin-1 by EBV infection [16].

We hypothesised, based on the existing data, that the presentation of peptide epitopes derived from syncytin-1 or the MSRV envelope protein that cross-react with epitopes from myelin

proteins to CD4+ T cells in the context of the MHC Class II DR2b and DR2a molecules may be a molecular mimicry trigger for MS. We therefore determined *in silico* the potential for cross-reactive epitopes between syncytin-1, syncytin-2 and the MSRV envelope protein on one hand and human MBP, MOG and PLP on the other, that can be presented to CD4+ T cells by HLA DR2a and DR2b molecules on antigen presenting cells.

2. Materials and Methods

2.1 Sequence homologies between HERV-W and HERV FRD envelope proteins and the myelin proteins MBP, MOG and PLP

The predicted protein coding sequences of the 538 amino acid (aa) HERV-W syncytin-1 (Uniprot accession number Q9UQF0), the 538 aa HERV FRD syncytin-2 (NP_997465), the 542 aa MSRV envelope protein (AAK18189.1), the 304 aa human MBP (P02686.3), the 247 aa human MOG (Q16653.2), and the 277 aa human PLP (P60201.2) were obtained from the NCBI data base. Amino acid sequences of the proteins were compared by pairwise Basic Local Alignment Search Tool (BLAST) analysis performed online using default parameters (https://www.ncbi.nlm.nih.gov/blast).

2.2 Prediction of peptides in HERV envelope proteins and myelin proteins potentially binding to HLA DR2a and HLA DR2b molecules

Syncytin-1, Syncytin-2 and the MSRV envelope protein and the three myelin proteins were analysed for peptides potentially binding to HLA DR2a and HLA DR2b molecules using the Immune Epitope Data Base or IEDB (www. iedb.org) procedures [20 - 23]. The default peptide length of 15 aa was used in the analysis but the results also show the core nonamer peptides that are expected to bind to the HLA DR molecule and constitute the major portion of the T cell epitope [20]. The analysis method selected was the Stabilised Matrix Method (SMM) where the peptides are ranked according to their predicted binding affinities or IC_{50} which indicates the concentration of

peptide in nM expected to achieve 50% saturation of the HLA molecule. Therefore a lower IC_{50} indicates higher affinity. As a guide, peptides with IC_{50} values <50 nM are considered to bind with high affinity, between 50nM to 500 nM with intermediate affinity and between 500nM to 5000 nM with low affinity [21 - 23]. For each peptide, a percentile rank is generated by comparing the peptide's score against the scores of five million random 15 mers selected from the SWISSPROT protein database. Therefore smaller percentile rank values, typically <10, also indicate higher affinity and specificity of binding to the HLA molecule [21 - 23].

The predicted peptides binding with higher-affinity in myelin proteins were then examined to determine whether they were located in regions that were homologous to syncytin-1, syncytin-2 and the MSRV envelope protein. Homologies between the predicted syncytin-1, syncytin-2 and the MSRV envelope protein and myelin protein peptides were additionally tested by pairwise BLAST analysis of the peptides.

3. Results

3.1 Sequence homologies between HERV envelope proteins and the three myelin proteins

A comparison of the syncytin-1 and MSRV envelope protein by BLAST revealed that the two proteins were 87% identical with 90% positives and 4 gaps (Supplementary Figure S1). A comparison of syncytin-1 with syncytin-2 showed that these two proteins were more distantly related with significant homology present only between parts of the two proteins (Figure S1b). The results obtained with pair-wise BLAST analysis of syncytin-1 and MBP, MOG, and PLP are presented in Supplementary Figures S2, S3 and S4 respectively. The results obtained with pair-wise BLAST analysis of MSRV envelope protein with MBP, MOG, and PLP are presented in Supplementary Figures S5, S6 and S7 respectively

There are three regions of homology between syncytin-1 and MBP, with greatest homology of 29% with an E value of 0.07 being found between aa 223-274 of MBP and aa 8-58 of

syncytin-1 allowing for a total of nine gaps in both proteins (Figure S2). The region between aa 223-274 of MBP was also homologous to MSRV envelope protein aa 8-58, with an E value of 1.8 and nine gaps (Figure S5). Two other regions of homology identified between syncytin-1 and MBP possessed weaker homologies with E values of 0.7 and 6 (Figure S2). Similarly two other regions of weaker homology were detected between MBP and MSRV envelope protein with E values of 2.7 and 9.4, with only the latter region being identical to a region of homology detected between MBP and syncytin-1 (Figure S5).

A short region of significant homology of 62% identity without gaps with an E value of 0.03 was seen between MOG aa 214-226 and syncytin-1 aa 448-460 (Figure S3). However this region was not detected in the pairwise BLAST comparison of MOG and MSRV envelope protein but four other regions of weak homologies with E values >2.3 were detected (Figure S6).

Between syncytin-1 and PLP, only one region of weak homology was identified *viz.* between aa 114-154 of PLP and aa 411-453 of syncytin-1 with 28% identity and an E value of 5.9 allowing for two gaps in the PLP sequence (Figure S4). The corresponding region of MSRV envelope protein aa 411-464 was also homologous to PLP aa 114-167 with 29% identity, E value of 0.9 and four gaps (Figure S7).

No regions of significant homology were observed between syncytin-2 and the three myelin proteins (data not shown).

3.2 Peptides in HERV envelope proteins and the three myelin proteins predicted to bind to HLA DR2b molecules

The results of IEDB analysis of syncytin-1, MSRV envelope protein, MBP, MOG and PLP binding HLA DR2b are shown in Supplementary Tables S8, S9, S10, S11 and S12 respectively.

3.2.1 Syncytin-1

There were twelve 15mer peptides in syncytin-1 with an IC_{50} <50 nM and percentile rank ≤ 0.4 predicted to be able to bind to the HLA DR2b molecule (Table S8). Four of these twelve high affinity predictions contained the core nonamer sequence ILPFLGPLA corresponding to aa 448 to 456 in the syncytin-1 sequence that is homologous to the sequence TLPFLGPLA (aa 448-456) in the MSRV envelope protein (Figure S1). This sequence in syncytin-1 is within a short sequence significantly homologous to MOG that was identified by BLAST analysis (Figure S3). In the corresponding homologous region of MOG are two overlapping peptides with core nonamer sequences of IVPVLGPLV (aa 214-222) and ITLFVIVPV (aa 209-217) predicted to have a high affinity of binding to HLA DR2b (Table S11 and Section 3.2.4 below). The syncytin-1 nonamer ILPFLGPLA shows 6/9 positional identities with the MOG nonamer IVPVLGPLV.

3.2.2 MSRV envelope protein

The MSRV envelope protein was predicted to have 10 peptides with high affinity to HLA DR2b (Table S9). Four, with nonamer sequence of FRPYISIPV (as 194-202) and four with the nonamer sequence LVKFVSSRI (aa 472-480) were homologous to the predicted high affinity sequences FRPYVSIPV (aa 194-202) and LVNFYSSRI (aa 472-480) respectively in syncytin-1 (the different amino acids in the nonamers are underlined). Two other predicted nonamers of high affinity to the MSRV envelope protein *viz.* LPLHFRPYI (aa 190-198) and FNFLVKFVS (aa 469-477) show two differences from the syncytin-1 high affinity nonamers LPLNFRPYV (aa 190-198) and FNLLVNFVS (aa 469-477) respectively.

The nonamer with sequence TLPFLGPLA (aa 448-456) in the MSRV envelope protein that is homologous to high affinity nonamer sequences in syncytin-1 and MOG (section 3.2.1) lies within a MSRV envelope protein peptide predicted to have an intermediate affinity to HLA DR2b (Table S9).

3.2.3 MBP

IEDB analysis of MBP showed that there were five 15mer peptides that were predicted to bind with high affinity to HLA DR2b and with a percentile rank ≤0.14 (Table S10). The core nonamer peptides for these were all derived from residues 220/221 to 228/229 with the sequences (V)VHFFKNIV(T) that did not show detectable homology with syncytin-1, syncytin-2 or the MSRV envelope protein.

3.2.4 MOG

IEDB analysis of MOG predicted five peptides with high affinity of for HLA DR2b and percentile rank ≤0.34 (Table S11). The homology of the MOG overlapping core nonamer sequence IVPVLGPLV (residues 214-222) from these with the predicted syncytin-1 core nonamer ILPFLGPLA and the MSRV nonamer TLPFLGPLA has been described in section 3.2.1. Syncytin-2 did not possess a homologous nonamer to the MOG nonamer IVPVLGPLV.

3.2.5 PLP

IEDB analysis of PLP showed that there were five 15mer peptides predicted to have a high affinity of binding to HLA DR2b with percentile rank ≤0.43 (Table S12). These had core nonamer sequences of which one FFFLYGALL (aa 78-86) showed homology with a syncytin-1/MSRV envelope protein sequence FLFTVLL (aa 8-14) with 4/7 identities and an E value of 0.19 (Figure S12b). The sequence FLFTVLL appeared in many syncytin-1 and MSRV envelope protein peptides that were predicted to have an intermediate affinity for HLA DR2b (Tables S8 and S9) but syncytin-2 did not possess a homologous sequence.

3.2.6 Syncytin-2

The four nonamers in syncytin-2 that were predicted to have a high affinity of binding to HLA DR2b were located within the signal peptide and not homologous to the three myelin proteins (data not shown).

3.3 Peptides in HERV envelope proteins and the three myelin proteins predicted to bind to HLA DR2a molecules

The results of IEDB analysis of syncytin-1, MSRV envelope protein, MBP, MOG and PLP peptides binding to HLA DR2a are shown in Supplementary Tables S13, S14, S15, S16 and S17 respectively.

3.3.1 Syncytin-1

There were 14 peptides (15mers) in syncytin-1 with a predicted high affinity of binding to HLA DR2a. Nine of these peptides had an affinity of <10nM IC_{50} with a percentile rank = 0.01. However none of the core nonamers derived from these sequences showed a significant homology with the three myelin proteins.

3.3.2 MSRV envelope protein

One peptide with a predicted high affinity to HLA DR2a containing the core nonamer FLVKFVSSR (aa 471-479) was identified in the MSRV envelope protein. However this was not significantly homologous to sequences within the three myelin proteins in BLAST analysis.

3.3.3 MBP, MOG and PLP

Seven 15mer peptides from MBP, five 15mers from MOG and seventeen 15mers from PLP were predicted to bind to HLA DR2a with high affinity and percentile ranks ≤ 0.13, 0.22 and 0.23 respectively. However none of the core nonamers derived from these peptides were significantly homologous to sequences within syncytin-1, syncytin-2 or the MSRV envelope protein.

3.3.4 Syncytin-2

In syncytin-2, the ten nonamers with predicted high affinity for HLA DR2a did not show significant homology to the three myelin proteins (data not shown).

3.4 MBP peptide that binds to both HLA DR2a and HLA DR2b

The encephalitogenic MBP peptide 83-99 that had been shown to be presented in the context of HLA DR2a and HLA DR2b to T cell clones from MS patients [7, 8], was shown in the IEDB analysis to lie in a region predicted to have an intermediate affinity of binding to both HLA DR2b and DR2a and with percentile ranks of ≤2.36 and ≤2.84 respectively (Tables S10 and S15), consistent with the validity of the IEDB analytical procedure used here. However MBP 83-99 is in a region that was not homologous to syncytin-1 and the MSRV envelope protein by the BLAST analysis (Figures S2 and S5 respectively) or to syncytin-2 (data not shown).

3.5 Immunodominant peptides in MOG recognised in the context of HLA DR2(15)

Among synthetic peptides specifically able to stimulate T cells from MS patients, the MOG aa 63-87 peptide was shown to be immunodominant in an independent study [24]. The IEDB analysis showed that this peptide region has a high affinity for HLA DR2a and an intermediate affinity for HLA DR2b (Tables S11 and S16) consistent with the experimental findings. However MOG 63-87 was not detectably homologous to syncytin-1 and the MSRV envelope protein in BLAST analysis (Figures S3 and S6 respectively) or to syncytin-2 (data not shown).

Another human MOG peptide aa 35-55 was shown to be immunogenic in the context of HLA DR2b in transgenic mice and therefore potentially encephalitogenic in humans [25]. This region of MOG is predicted to be have weak affinity for HLA DR2b [IC_{50} 524 nM, percentile rank 10.21, core nonamer 32-40 FRVIGPRHP in Table S11] and intermediate affinity for HLA DR2a [IC_{50} 57 nM, percentile rank 0.3, with the same core nonamer 32-40 FRVIGPRHP in Table S16] in the IEDB analysis. However this region of MOG was not detected as being homologous to syncytin-1 and the MSRV envelope protein by BLAST analysis [Figures S3 and S6 respectively] or to syncytin-2 (data not shown).

3.6 PLP peptides recognised in the context of HLA DR2(15)

The PLP peptide 175-194 has been shown to be presented by transgenic murine T cells in the context of HLA DR2b [26] and this is consistent with the intermediate level of affinity for HLA DR2b and HLA DR2a observed in our present analysis (Tables S12 and S17 respectively). However this region of PLP was not significantly homologous to sequences in the two HERV-W envelope proteins (Figures S4 and S7) or to syncytin-2 (data not shown).

Discussion

The IEDB analyses identify several peptides in the two HERV-W envelope proteins that can be potentially presented to T_H cells with high affinity (<50nM IC_{50}) by HLA DR2b (12 peptides) and HLA DR2a (14 peptides), two Class II molecules that have been significantly associated with MS [5 – 8]. Some syncytin-1 peptides predicted to bind to HLA DR2a were expected to do so with particularly high affinity. Many potential binding peptides in the envelope proteins with intermediate (50-500nM IC_{50}) and weak (500-5000nM IC_{50}) affinities for HLA DR2b and HLA DR2a were also identified in the analyses. The findings are therefore consistent with the hypothesis that HERV-W envelope proteins contain many potential epitopes for CD4+ T cells that can be presented by MS- associated Class II HLA-DR molecules on antigen presenting cells.

The IEDB analyses also identified several peptides in the three myelin proteins that are predicted to bind with high affinity to HLA DR2b and HLA DR2a. As in the case of the HERV-W envelope proteins, many other potential peptides with intermediate (50n-500nM IC_{50}) and weak (500-5000nM IC_{50}) affinities for HLA DR2b and HLA DR2a were also identified in the three myelin proteins. The results show the potential for many peptides from the three myelin proteins to be presented to T_H cells in the context of HLA DR molecules known to be associated with a predisposition to MS. Furthermore, peptides from the three myelin proteins that had been shown experimentally to be recognised by T_H cells in the context of presentation by HLA DR2b and HLA

DR2a molecules, were shown by the *in silico* analysis here to have the potential for presentation by HLA DR2b and HLA DR2a.

The identification of peptides from syncytin-1 and the MSRV envelope protein that bind with a predicted high or intermediate affinity respectively to HLA DR2b that also show a degree of sequence homology to predicted high affinity peptides derived from MOG is consistent with the molecular mimicry hypothesis for MS. The identification of a PLP peptide with a predicted high affinity for HLA DR2b that shows sequence homology with predicted peptides from the two HERV-W envelope proteins with intermediate affinity for HLA DR2b is also in keeping with this hypothesis. The observation that immunisation with the MSRV envelope protein can induce EAE in mice is supportive of the hypothesis [27].

The question arises as to how syncytin-1 which plays a normal role in foetal development and also reported to be involved in maintaining maternal tolerance to the foetus [17-19] may become immunogenic in the context of MS. Syncytin-1 and lsyncytin-2 are membrane glycoprotein found in the plasma membrane of trophoblast cells that are also shed as exosomes into the maternal circulation [19]. The MSRV envelope glycoprotein, which is nearly identical to syncytin-1, is on the outer membrane of complete viral particles [18] that have been detected in MS lesions in the CNS [14-16], including in monocytes and macrophages [16]. It is possible therefore syncytin-1 in circulating exosomes and MSRV envelope protein whose expression may be induced within or outside the CNS, can be taken up by antigen presenting cells, processed and presented to T_H cells on HLA DR2b or HLA DR2a molecules. Antigen presenting cells in the CNS include microglia, astrocytes and perivascular monocytes and macrophages and also inflammation-activated endothelium [28]. MSRV envelope protein has been shown to activate human monocytes via the pattern recognition receptors CD14 and TLR4 to induce the formation of the proinflammatory cytokines TNFα, IL-1β and IL-6, and to activate dendritic cells to promote T_H1 responses [29]. It is therefore possible that HERV-W envelope proteins in the form of syncytin-1 or the MSRV envelope protein are a source of peptide

311 epitopes that can initially activate T_H cells in the context of HLA DR2b or HLA DR2a, particularly in an
312 inflammatory context. One source of inflammation can be the activation of MSRV transcription to
313 produce MSRV envelope protein by infections such as EBV [16]. A proportion of the activated T_H cells
314 may then be expected to react with the cross-reactive MOG and PLP peptide epitopes predicted
315 here that are processed and presented by CNS resident antigen presenting cells on HLA DR2b
316 molecules. The resulting amplified inflammation, recruitment of cytotoxic T cells and formation of
317 anti-MOG and anti-PLP antibodies can further damage oligodendrocytes that make myelin and the
318 myelin sheath on nerve cells. This can result in release of more myelin proteins that can be taken up
319 by antigen presenting cells to cause epitope spreading and activation of additional T_H cells.
320 Additional myelin protein – derived peptides may be presented in the context of HLA DR2b or HLA
321 DR2a as predicted here or in the context of other HLA molecules as has been experimentally shown
322 to be possible [30] to result in full-blown MS. We therefore suggest that the *in silico* analysis
323 provides support for molecular mimicry between peptides derived from HERV-W envelope proteins
324 and myelin proteins presented by HLA DR2b and HLA DR2a helping to initiate MS. However the
325 available data does not support a similar role for HERV FRD encoded syncytin-2 in MS. Experimental
326 investigations on the ability of the peptides identified in this study to be recognised by antigen
327 receptors on T_H cells in the context of HLA DR2b and HLA DR2a, and then become appropriately
328 activated, are needed to confirm the hypothesis.

329 The identification of important peptide epitopes can lead to the development of small
330 molecule inhibitors that impede their presentation by HLA DR2b and HLA DR2a to treat MS in its
331 early stages [31]. Two other relevant immunological approaches to treat MS that have shown
332 promise in the murine EAE model include infusing syngeneic antigen presenting cells containing
333 chemically coupled immunogenic MOG peptide which leads to induction of regulatory T_H cells,
334 downregulation of inflammatory T_H cells and reversal of EAE [32, 33], as well as immunisation with
335 cryptic peptide epitopes from MOG that ameliorates EAE [34].

Women have a higher prevalence of MS than men during and after the child bearing age and this gender difference has been attributed to hormonal effects and differential environmental exposure to as yet undetermined agents [35, 36]. It has been reported that MSRV copy numbers are increased in female MS patients compared to male patients, that this increase correlates with disease severity [36], and that the increase may be caused by HERV-W copies in the X chromosome [37, 38]. Our findings suggest that the expression of syncytin-1 during pregnancy and molecular mimicry between syncytin-1 and myelin proteins may contribute, at least in part, to the higher prevalence of MS among women.

Conflict of interest statement

The authors declare no conflict of interest.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

TW and RR designed the project, BJ performed the initial analysis, and RR carried out detailed analysis and drafted the manuscript. All authors read and approved the final manuscript.

References

[1] Compston A, Coles A. Multiple sclerosis. Lancet. 372 (2008)1502-1517. doi: 10.1016/S0140-6736(08)61620-7.

[2] Lossius A, Johansen JN, Vartdal F, Holmøy T. High-throughput sequencing of immune repertoires in multiple sclerosis. Ann. Clin. Transl. Neurol. 3 (2016) 295-306. doi: 10.1002/acn3.295.

[3] Cusick MF, Libbey JE, Fujinami RS. Multiple sclerosis: autoimmunity and viruses. Curr. Opin. Rheumatol. 25 (2013) 496-501. doi: 10.1097/BOR.0b013e328362004d.

[4] Rangachari M, Kuchroo VK. Using EAE to better understand principles of immune function and autoimmune pathology. J. Autoimmun. 45 (2013) 31-39. doi: 10.1016/j.jaut.2013.06.008.

[5] International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium , Sawcer S, Hellenthal G, Pirinen M, Spencer CC, *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 476 (2012) 214–219. doi:10.1038/nature10251.

[6] Masterman T, Ligers A, Olsson T, Andersson M, Olerup O, Hillert J. HLA-DR15 is associated with lower age at onset in multiple sclerosis. Ann. Neurol. 48 (2008) 211-219.

[7] Vergelli M, Kalbus M, Rojo SC, Hemmer B, Kalbacher H, Tranquill L, *et al*. T cell response to myelin basic protein in the context of the multiple sclerosis-associated HLA-DR15 haplotype: peptide binding, immunodominance and effector functions of T cells. J. Neuroimmunol. 177 (1997) 195-203.

[8] Quandt JA, Huh J, Baig M, Yao K, Ito N, Bryant M, *et al*. Myelin basic protein-specific TCR/HLA-DRB5*01:01 transgenic mice support the etiologic role of DRB5*01:01 in multiple sclerosis. J. Immunol. 189 (2012):2897-2908. doi: 10.4049/jimmunol.1103087

[9] Sospedra M, Martin R. Immunology of multiple sclerosis. Annu. Rev. Immunol. 23 (2005) 683-747.

[10] Fugger L, Friese MA, Bell J. From genes to function: the next challenge to understanding multiple sclerosis. Nat. Rev. Immunol. 9 (2009) 408-417. doi: 10.1038/nri2554.

[11] Vergelli M, Hemmer B, Muraro PA, Tranquill L, Biddison WE, Sarin A, *et al.* Human autoreactive CD4+ T cell clones use perforin- or Fas/Fas ligand-mediated pathways for target cell lysis. J. Immunol. 158 (1997) 2756–2761.

[12] Johnson MC, Pierson ER, Spieker AJ, Nielsen AS, Posso S, Kita M, *et al*. Distinct T cell signatures define subsets of patients with multiple sclerosis. Neurol. Neuroimmunol. Neuroinflamm. 3 (2016) e278. doi: 10.1212/NXI.0000000000000278.

[13]Gross CC, Schulte-Mecklenbeck A, Hanning U, Posevitz-Fejfár A, Korsukewitz C, Schwab N, *et al*. Distinct pattern of lesion distribution in multiple sclerosis is associated with different circulating T-helper and helper-like innate lymphoid cell subsets. Mult. Scler. (2016) pii: 1352458516662726.

[14]Nissen KK, Laska MJ, Hansen B, Terkelsen T, Villesen P, Bahrami S, *et al*. Endogenous retroviruses and multiple sclerosis-new pieces to the puzzle. BMC. Neurol. 13 (2013) 111. doi: 10.1186/1471-2377-13-111.

[15]Dolei A, Uleri E, Ibba G, Caocci M, Piu C, Serra C. The aliens inside human DNA: HERV-W/MSRV/syncytin-1 endogenous retroviruses and neurodegeneration. J. Infect. Dev. Ctries. 9 (2015) 577-587. doi: 10.3855/jidc.6916.

[16]Morandi E, Tarlinton RE, Gran B. Multiple sclerosis between genetics and infections: Human endogenous retroviruses in monocytes and macrophages. Front. Immunol. 6 (2015) 647. doi: 10.3389/fimmu.2015.00647

[17]Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, *et al*. Syncytin-1 is a captive retroviral envelope protein involved in human placental morphogenesis. Nature. 403 (2000) 785-789.

[18]Lokossou AG, Toudic C, Barbeau B. Implication of human endogenous retrovirus envelope proteins in placental functions. Viruses. 6 (2014) 4609-4627. doi: 10.3390/v6114609.

[19]Mameli G, Poddighe L, Astone V, Delogu G, Arru G, Sotgiu S, *et al*. Novel reliable real-time PCR for differential detection of MSRVenv and syncytin-1 in RNA and DNA from patients with multiple sclerosis. J. Virol. Methods. 161 (2009) 98-106. doi: 10.1016/j.jviromet.2009.05.024.

[20]Nielsen M, Lund O, Buus S, Lundegaard C. MHC class II epitope predictive algorithms. Immunology. 130 (2010) 319-328. doi: 10.1111/j.1365-2567.2010.03268.

[21]Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS. Comput. Biol. 4 (2008) e1000048.

411 [22]Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B. Peptide binding predictions
412 for HLA DR, DP and DQ molecules. BMC. Bioinformatics. 11 (2010) 568.
413 [23]Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, *et al*. Immune epitope
414 database analysis resource. Nucleic Acids Res. 40 (2012) W525-W530. doi:
415 10.1093/nar/gks438
416 [24]Wallström E, Khademi M, Andersson M, Weissert R, Linington C, Olsson T. Increased
417 reactivity to myelin oligodendrocyte glycoprotein peptides and epitope mapping in HLA
418 DR2(15)+ multiple sclerosis. Eur. J. Immunol. 28 (1998) 3329-3335.
419 [25]Rich C, Link JM, Zamora A, Jacobsen H, Meza-Romero R, Offner H, *et al*. Myelin
420 oligodendrocyte glycoprotein-35-55 peptide induces severe chronic experimental
421 autoimmune encephalomyelitis in HLA-DR2-transgenic mice. Eur. J. Immunol. 34 (2004)
422 1251-1261.
423 [26]Kaushansky N, Altmann DM, David CS, Lassmann H, Ben-Nun A. DQB1*0602 rather than
424 DRB1*1501 confers susceptibility to multiple sclerosis-like disease induced by proteolipid
425 protein (PLP). J. Neuroinflammation. 9 (2012) 29. doi: 10.1186/1742-2094-9-29.
426 [27]Perron H, Dougier-Reynaud HL, Lomparski C, Popa I, Firouzi R, Bertrand JB, *et al*. Human
427 endogenous retrovirus protein activates innate immunity and promotes experimental
428 allergic encephalomyelitis in mice. PLoS. One. 8(2013) e80128. doi:
429 10.1371/journal.pone.0080128.
430 [28]Chastain EM, Duncan DS, Rodgers JM, Miller SD. The role of antigen presenting cells in
431 multiple sclerosis. Biochim. Biophys. Acta. 1812 (2011) 265-274. doi:
432 10.1016/j.bbadis.2010.07.008.
433 [29]Rolland A, Jouvin-Marche E, Viret C, Faure M, Perron H, Marche PN. The envelope protein of
434 a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and
435 promotes Th1-like responses. J. Immunol. 176 (2006) 7636–7644. doi:
436 10.4049/jimmunol.176.12.7636.

[30]Raddassi K, Kent SC, Yang J, Bourcier K, Bradshaw EM, Seyfert-Margolis V, *et al*. Increased frequencies of myelin oligodendrocyte glycoprotein/MHC class II-binding CD4 cells in patients with multiple sclerosis. J. Immunol. 187 (2011) 1039-1046. doi: 10.4049/jimmunol.1001543.

[31]Ji N, Somanaboeina A, Dixit A, Kawamura K, Hayward NJ, Self C, *et al*. Small molecule inhibitor of antigen binding and presentation by HLA-DR2b as a therapeutic strategy for the treatment of multiple sclerosis. J. Immunol. 191 (2013) 5074-84. doi: 10.4049/jimmunol.1300407.

[32]Turley DM, Miller SD. Peripheral tolerance induction using ethylenecarbodiimide-fixed APCs uses both direct and indirect mechanisms of antigen presentation for prevention of experimental autoimmune encephalomyelitis. J. Immunol. 178 (2007) 2212-2220.

[33]Zhang L, Guo Y, Xia CQ. Infusion of Sulfosuccinimidyl-4-[N-maleimidomethyl]cyclohexane-1-carboxylate-Conjugated MOG35-55-Coupled Spleen Cells Effectively Prevents and Reverses Experimental Autoimmune Encephalomyelitis in Mice. J. Immunol. Res. 2015 (2015)129682. doi: 10.1155/2015/129682.

[34]Lyons JA, Riter MM, Almatrook AM, Ramsbottom MJ, Cross AH. Amelioration of EAE by a cryptic epitope of myelin oligodendrocyte glycoprotein. J. Neuroimmunol. 300 (2016) 66-73. doi: 10.1016/j.jneuroim.2016.06.006.

[35]Leray E, Moreau T , Fromont A , Edan G. Epidemiology of multiple sclerosis. Rev. Neurol (Paris). 172 (2016) 3-13. doi: 10.1016/j.neurol.2015.10.006.

[36]Alonso A, Hernán MA. Temporal trends in the incidence of multiple sclerosis - A systematic review. Neurology. 71 (2008) 129-135.

[37]Garcia-Montojo M, Dominguez-Mozo M, Arias-Leal A, Garcia-Martinez Á, De las Heras V, Casanova I, *et al*. The DNA copy number of human endogenous retrovirus-W (MSRV-type) is increased in multiple sclerosis patients and is influenced by gender and disease severity. PLoS. One. 8 (2013) e53623. doi: 10.1371/journal.pone.0053623.

[38]García-Montojo M, de la Hera B, Varadé J, de la Encarnación A, Camacho I, Domínguez-Mozo M, *et al*. HERV-W polymorphism in chromosome X is associated with multiple sclerosis risk and with differential expression of MSRV. Retrovirology. 11 (2014) 2. doi: 10.1186/1742-4690-11-2.

.