TESIS DE LA UNIVERSIDAD DE ZARAGOZA
2020
12

Sofía Macías Redondo

Human Endogenous Retrovirus Type W in Multiple Sclerosis

Departamento
Bioquímica y Biología Molecular y Celular

Director/es
Schoorlemmer, Johan Albert

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

HUMAN ENDOGENOUS RETROVIRUS TYPE W IN MULTIPLE SCLEROSIS

Autor

Sofía Macías Redondo

Director/es

Schoorlemmer, Johan Albert

UNIVERSIDAD DE ZARAGOZA

Bioquímica y Biología Molecular y Celular

2018

UNIVERSIDAD DE ZARAGOZA

PROGRAMA DE DOCTORADO EN BIOQUÍMICA Y BIOLOGÍA MOLECULAR

TESIS DOCTORAL

HUMAN ENDOGENOUS RETROVIRUS TYPE W IN MULTIPLE SCLEROSIS

Memoria presentada por SOFÍA MACÍAS REDONDO

para optar al grado de Doctor

ZARAGOZA, 2018

D. JON SCHOORLEMMER, Doctor en Biología,

HACE CONSTAR:

Que Dña. Sofía Macías Redondo ha realizado bajo mi dirección el trabajo de tesis doctoral: **"Human endogenous retrovirus type W in multiple sclerosis"** y corresponde fielmente a los resultados obtenidos.

Una vez redactada la presente memoria, ha sido revisada por mí y posee la calidad científica necesaria para ser presentada y aspirar al grado de Doctor por la Universidad de Zaragoza.

Y para que conste, en cumplimiento de las disposiciones vigentes, expido el presente informe favorable.

Zaragoza, a 28 de septiembre de 2018.

Fdo. Jon Schoorlemmer

El desarrollo y ejecución de esta Tesis Doctoral se han enmarcado dentro del proyecto de investigación en salud (FIS)/Carlos III “La reactivación de los Retrovirus Endógenos Humanos (HERV) en Esclerosis Múltiple: el control epigenético y el potencial diagnóstico” (PI1302518). Sofía Macías Redondo ha disfrutado de una beca/contrato de Personal Investigador Predoctoral en formación de la Diputación General de Aragón, financiada por el Gobierno de Aragón y el Fondo Social y Europeo (C071/2014).

Esta tesis se la dedico a mis padres.

A mi madre, por alimentar mi alma y acallar mis miedos.

A mi padre, por sus consejos e instarme a dar siempre lo mejor de mí misma.

He llegado hasta aquí gracias a vuestro cariño y apoyo incondicional.

Gracias de todo corazón.

"Me he subido a mi mesa para recordar
que hay que mirar las cosas de un modo diferente.
El mundo se ve distinto desde aquí arriba."

El club de los poetas muertos

En primer lugar, quiero agradecer a mi director de tesis el Dr. Jon Schoorlemmer por brindarme la oportunidad de formar parte de su equipo. Gracias por darme autonomía, enseñarme a pensar por mí misma y permitirme cometer errores como parte del proceso de aprendizaje. Esta tesis nos ha dado algún que otro momento de frustración, pero me has enseñado a afrontarlos con tranquilidad, ilusión y nuevos propósitos.

Gracias a todo el equipo del Área de Neurología del Hospital Universitario Miguel Servet por la colaboración prestada. Sin vuestra profesionalidad y generosidad esta tesis no habría sido posible. Y gracias a la ayuda prestada por el personal de los SCT del IACS. En especial a Mark, César y David, por su respaldo e incondicional apoyo durante estos años. Ha sido un placer aprender de grandes profesionales como vosotros.

Gracias a mi familia y a los amigos de siempre, que por mucho tiempo que pase siempre me esperan con una gran sonrisa, enfundándome alegría, aunque no haya podido compartir con ellos todos los momentos que quisiera y el tiempo que merecen. Cada carta recibida, llamada inesperada, reencuentro y escapada a casa, ha hecho que os lleve conmigo y os quiera, si es posible, un poquito más.

Gracias a los nuevos amigos que han compartido conmigo este viaje en primera persona. Yohanna, Marcos, Mati, Luis, nos embarcamos a la vez en la gran aventura del Doctorado y juntos hemos compartido alegrías y penas. Y a todos los amigos, ya casi familia, que me ha brindado la sala de predocs, me quedo con el compañerismo de Pilar, la alegría de Iris y el buen corazón de Natalia. Con la generosidad de Edu, los consejos de Pedro y la sonrisa de Manuel. Con la energía de Samantha, la empatía de Joana y la naturalidad de Ruth. A todos, no cambies nunca porque valéis mucho.

Gracias a mis dos guardianes, Fran y César. Habéis sido mi refugio en los días malos, animándome cada vez que llegaba a casa triste porque la PCR no había salido o las células habían decidido morirse… Me habéis recompuesto siempre con vuestro cariño y comprensión. Habéis crecido personal y profesionalmente conmigo estos años y no puedo estar más orgullosa de vosotros.

Y, por último, gracias a mis ángeles. Hay gente que trae una luz tan especial al mundo que, incluso después de haberse ido, la luz permanece.

LIST OF ABBREVIATIONS

AD	Alzheimer's Disease
APC	Antigen-Presenting Cell
APOBEC	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like
AZFa	Azoospermia Factor A
BBB	Blood-Brain Barrier
BDNF	Brain-Derived Neurotrophic Factor
BER	Base Excision Repair machinery
BMPs	Bone Morphogenetic Proteins
BSA	Bovine Serum Albumin
cAMP	Adenosine cyclic Monophosphate
CIS	Clinically Isolated Syndrome
CFA	Freund's Adjuvant
CLNs	Cervical Lymph Nodes
CNS	Central Nervous System
CSF	Cerebrospinal Fluid
DMEM/F-12	Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12
DNA, cDNA, gDNA	Deoxyribonucleic acid, complementary DNA, genomic DNA
DNMTi, **DNMTs**	DNA Methyltransferases, DNMTs inhibitors
EAE	Experimental Autoimmune Encephalomyelitis
EBs	Embryoid Bodies
EBV	Epstein–Barr Virus
EDSS	Expanded Disability Status Scale
EGF	Epidermal growth factor
ENV	Envelope
FGF2	Basic Fibroblast Growth Factor
GAD	Glutamic Acid Decarboxylase
GFAP	Glial Fibrillary Acidic Protein
GFP	Green Fluorescent Protein

GWASs	Genome-Wide Association Studies
GPCs	Glial Precursors Cells
HERVs, HERV-W/K	Human Endogenous Retroviruses, HERV family W/K
hESCs	Human Embryonic Stem Cells
HGF	Hepatocyte Growth Factor
HHV	Human Herpesvirus
HLA	Human Leukocyte Antigen
HSV-1	Herpes Simplex Virus 1
IFN-α/β/γ	Interferon-alpha/beta/gamma
IGF-1	Insulin-like Growth Factor-1
IgG	Immunoglobulin G
iPSCs	Induced Pluripotent Stem Cells
LINEs	Long Interspersed Nuclear Elements
LTRs	Long Terminal Repeats
MBP	Myelin Basic Protein
MHC (I/II)	Major Histocompatibility Complex (class I/II)
MMR	Mismatch Repair
MOBP	Myelin-associated Oligodendrocyte Basic Protein
MOG	Myelin Oligodendrocyte Glycoprotein
MSRV	Multiple Sclerosis-associated Retrovirus
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
M&M	*Materials and Methods*
NCAM	Neural Cell Adhesion Molecule
NER	Nucleotide Excision Repair
NGS	Next-Generation Sequencing
NO	Nitric Oxide
NPCs	Neural Progenitor Cells
NT3	Neurotrophin 3
OL	Oligodendrocyte
OPCs	Primary Oligodendrocyte Precursor Cells

OPF	Open Reading Frame
PAMP	Pathogen-Associated Molecular Pattern
PBMCs	Peripherical Blood Mononuclear Cells
PDGF-AA	Platelet-Derived Growth Factor AA
PGK	Phosphoglycerate Kinase
PLP	Proteolipid Protein
PPMS	Primary Progressive Multiple Sclerosis
PRMS	Progressive Relapsing Multiple Sclerosis
PT	Pertussis Toxin
qPCR	Real-time Quantitative Polymerase Chain Reaction
RA	Retinoic Acid
RELN	Reelin
RNAi, ds RNA	Ribonucleic Acid interference, double-stranded RNA
RRMS	Relapsing-Remitting Multiple Sclerosis
RT	Reverse Transcriptase
SAG	Smoothened Agonist
SINEs	Short Interspersed Nuclear Elements
SNP	Single Nucleotide Polymorphism
SPMS	Secondary Progressive Multiple Sclerosis
TEs	Transposable Elements
TLRs	Toll-Like Receptors
T3	Thriiodothytonine
VZV	Varicella-Zoster Virus
5-azadC	5-aza-2-deoxycytidine

INDEX

INDEX

I. SUMMARY/ RESUMEN

Summary

Multiple sclerosis (MS) is a chronic inflammatory and degenerative disease of the central nervous system (CNS). Although its etiology is unknown, it is considered an autoimmune disease facilitated by genetic and environmental predispositions (Milo and Kahana, 2010). Already in the 1990's, the presence of retrovirus-like particles was associated with MS pathology (Perron et al., 1997). The RNA of these particles was characterized as the MS-associated endogenous retrovirus (MSRV), which shares good homology with the human endogenous retrovirus family W (HERV-W) (Blond et al., 1999). Both the presence and expression of RNA and proteins of MSRV/HERV-W is elevated in MS patients and levels coincide with progression of the disease (Garcia-Montojo et al., 2013; Mameli et al. al., 2009; Perron et al., 2012). Although MS mostly affects the brain, increased expression of MSRV/HERV-W has also been observed in the peripherical blood mononuclear cells (PBMC). HERV-W is a family composed of more than 600 similar elements integrated throughout the human genome (Pavlicek et al., 2002). While similar, detailed sequence analysis allows for discrimination between elements (Lander et al., 2001). My studies have focussed on elucidating the exact relationship between MSRV and HERV-W on the one hand, and understanding how HERV-W overexpression contributes to MS on the other hand.

I hypothesized that MSRV could be a HERV-W element that is not fixed in the current population, and therefore not described in the genome database. My first objective was to identify copies of HERV-W not represented in the database. A PCR-based method described previously to localize transposable elements (Uren et al., 2009), was adapted to HERV-W type sequences. Using the variant method, several HERV-W elements present in PBMCs from 4 MS patients have been successfully amplified, cloned and localized in the human genome. However, all of the HERV-W sequences identified in this way corresponded to copies represented in the database. NGS has allowed for identification of HERV-W copies not represented in the database. Their implication in MS is presently unknown.

Thanks to a collaboration with the department of Neurology of Miguel Servet University Hospital (Zaragoza, Spain), I have obtained blood samples from MS patients

and healthy controls from which I have extracted PBMCs. The number of copies of HERV-W in the samples was analysed, and I have detected that it is not enhanced in MS patients. HERV-W expression was analysed in PBMCs, and expression was slightly elevated in MS patients. Overexpression among patients could be linked to clinically active patients. I have analysed which HERV-W elements might be overexpressed or overrepresented in the transcriptome of MS patients. Relative expression levels of HERV-W that are most abundantly transcribed in PBMCs do not differ between MS patients and controls. By contrast, several specific HERV-W elements are exclusively transcribed in MS patients, although their relative expression levels are not high (between 1-5%).

Taking other HERVs elements as an example, regulation of HERV-W may be controlled, among other mechanisms, by epigenetic mechanisms such as DNA methylation. As the methylation status of HERV-W has not been analysed in the context of MS, the possibility that the methylation of HERV-W *loci* would control their expression was investigated. Methylation assays for selected HERV-W elements were designed, validated and applied to genomic DNA obtained from MS patients. Results show that the analysed *loci* were highly methylated in PBMCs of both MS patients and controls. No relation between HERV-W expression and methylation levels could be established. Similarly, methylation levels were not altered in MS and did not correlate with the pathology.

Toll-like receptors (TLRs) detect viral proteins or nucleic acids and mediate an antiviral response that includes interferon beta (INFb) induction. While dsRNA transcribed from HERV *loci* is capable of triggering a TLR3-mediated immune response in tumour cells (Chiappinelli et al., 2015), a similar mechanism in MS has not been studied. The possibility that overexpression of HERV-W may induce an inflammatory response within the CNS, specifically in oligodendrocytes and neuronal precursor cells (NPCs), was investigated. Lentivirus-mediated HERV-W overexpression induced expression of INFb in NPCs, possibly through TLR.

The combined results of my studies suggest that the increase in HERV-W expression observed in MS patients does not originate from a single deregulated HERV-W element

that may be associated with the pathology. On the contrary, they suggest several copies of HERV-W that are expressed at low levels, do so exclusively in patients. My results suggest that the deregulation of HERV-W could activate the innate system within the CNS and thus contribute to the neuroinflammation present in MS.

Resumen

La esclerosis múltiple (MS, *multiple sclerosis*) es una enfermedad crónica inflamatoria y degenerativa del sistema nervioso central (CNS, *central nervous system*). Ya en los años 90 se relacionó la presencia de partículas parecidas a retrovirus con la patología de MS (Perron et al., 1997). El ARN (RNA, *ribonucleic acid*) de dichas partículas se caracterizó como el retrovirus endógeno asociado a esclerosis múltiple (MSRV, *multiple sclerosis-associated retrovirus*), el cual comparte mucha homología con la familia W de los retrovirus endógenos humanos (HERV-W, *human endogenous retrovirus family W*) (Blond et al., 1999). Tanto la presencia como la expresión de RNA y proteínas de MSRV/HERV-W se encuentra elevada en pacientes de MS y se asocia al progreso de la enfermedad (Garcia-Montojo et al., 2013; Mameli et al., 2009; Perron et al., 2012). Aunque la MS afecta sobre todo al cerebro, el aumento de expresión de MSRV/HERV-W se ha observado también en las células mononucleares de la sangre periférica (PBMC, *periferal blood mononuclear cells*). HERV-W es una familia compuesta por más de 600 elementos similares integrados a lo largo del genoma humano (Pavlicek et al., 2002). Sin embargo, la relación exacta entre MSRV y HERV-W se desconoce a día de hoy.

Para abordar cuál es la relación entre MSRV y HERV-W y, cuál es la contribución de HERV-W a la patología, el primer objetivo ha consistido en intentar localizar MSRV o copias de HERV-W que no están en la base de datos del genoma en el ADN (DNA, *desoxyribonucleic acid*) genómico de PBMCs de pacientes de MS. Para ello se ha modificado un método que nos permitiera amplificar secuencias de HERV-W. El ensayo ha permitido localizar en el genoma varios elementos de HERV-W presentes en el genoma de pacientes de MS, cuya relación con MS se desconoce a día de hoy.

Gracias a una colaboración con el Servicio de Neurología del Hospital Universitario Miguel Servet (Zaragoza, España), se han obtenido muestras de sangre de pacientes de MS y controles sanos de las cuales se han extraído las PBMCs. Posteriormente, se ha analizado el número de copias de HERV-W en el genoma humano, y se ha detectado que es contante en todos los individuos. También se han analizado los niveles de expresión de HERV-W a nivel de RNA en PBMCs. Se ha detectado que la expresión de

HERV-W se encuentra ligeramente elevada en los pacientes de MS, y que dicha expresión no proviene de una sola copia de HERV-W sino de un grupo de copias de HERV-W menos abundantes pero asociadas a los pacientes de MS.

Como en otros HERV, se supone que la regulación de HERV-W está controlada, entre otros mecanismos, por mecanismos epigenéticos como la metilación del ADN. Como el estado de metilación de HERV-W no ha sido analizado en el contexto de MS, se planteó la posibilidad de que el grado de metilación de HERV-W está relacionado con su expresión. Se han diseñado y aplicado ensayos para medir el grado de metilación de algunas de las copias de HERV-W previamente identificadas en pacientes de MS. Todos los *loci* analizados han resultado estar altamente metilados tanto en PBMCs de pacientes como en controles. Por lo que no parece que los niveles de metilación de HERV-W regulen su expresión.

Los receptores tipo Toll (TLR; *Toll-like receptors*) detectan productos virales en forma de proteína o ácidos nucleicos y median la respuesta antiviral. He iniciado estudios para investigar la posibilidad de que la sobre-expresión de HERV-W indujera una respuesta inflamatoria dentro del CNS. Debido a la dificultad de generar oligodendrocitos, se han utilizado precursores neuronales humanos. Se ha detectado que la sobre-expresión de HERV-W desencadena una respuesta inflamatoria mediada por interferón β. Estos resultados sugieren que la desregulación de HERV-W podría activar el sistema innato residente en el CNS y contribuir así en la neuro-inflamación presente en MS.

II. INTRODUCTION

1. Multiple sclerosis

1.1 Symptoms and epidemiology

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS), characterized by the demyelination of axons and loss of oligodendrocytes (Brück et al., 2002; Lucchinetti et al., 2000; Trapp et al., 1998). Myelin is the protective sheath that surrounds the axons of neurons (Figure 1A), acting as an electric insulator and allowing the fast propagation of electrical impulses along axons. This myelin is produced by oligodendrocytes, a group of specialised cells of the CNS. In addition, oligodendrocytes produce growth factors that provide trophic support to neurons, such as brain-derived neurotrophic factor (BDNF) or insulin-like growth factor-1 (IGF-1) (Du and Dreyfus, 2002). Oligodendrocytes, therefore, support both axon survival and function (Simons and Nave, 2016). Although the exact cause of MS remains unknown, the immune system is triggered to attack the myelin of the patient, causing the disruption of the electrical impulses (Nylander and Hafler, 2012). This interruption of communication leads to MS symptoms such as muscle weakness, fatigue, or paralysis among others (Figure 1B).

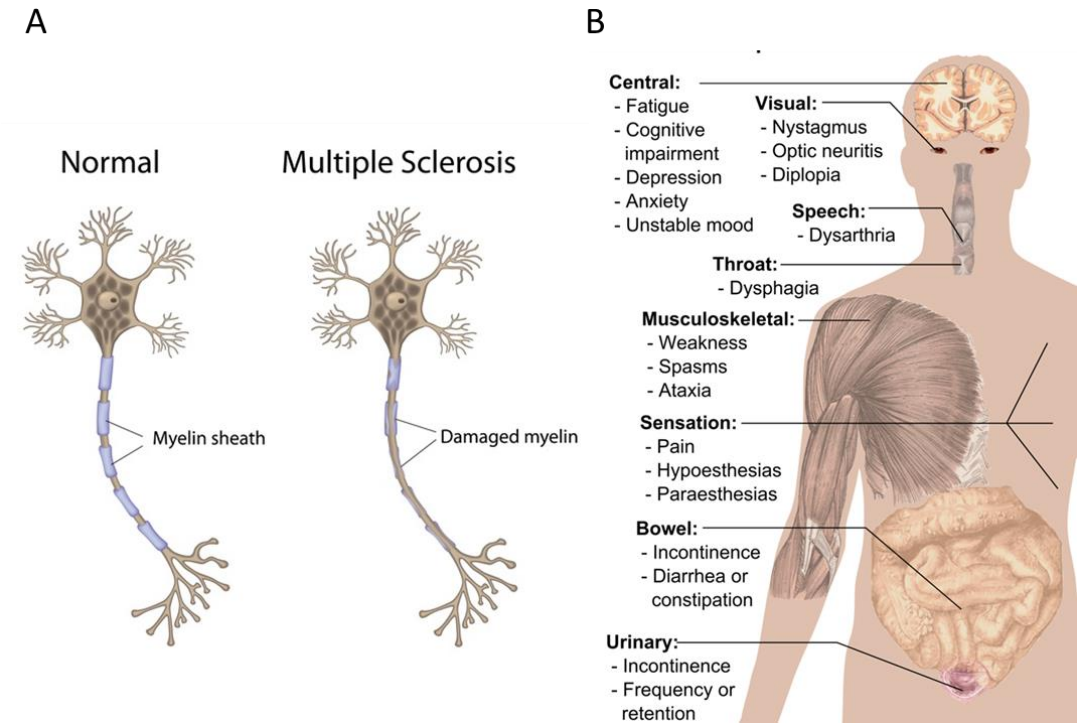


Figure 1. Representation of the structure of a neuron and the main symptoms of multiple sclerosis. A) In multiple sclerosis (MS), the myelin sheath that protects the nerve fibres is damaged and the transmission of nerve impulses impaired. B) The image summarizes the main symptoms of MS, which

range from direct neural pathologies (vision problems, body equilibrium and changes in cognition) to bowel and urinary alterations. Image A is taken from Alila Medical Media/Shutterstock.com. Image B is taken from Häggström and Häggström, 2014.

It is estimated that around 2.5 million people suffer from MS worldwide. MS can affect people of any age or sex. However, it is most commonly diagnosed in people between the ages of 20 and 50 years and is about 2 to 3 times more common in females than males (Goodin, 2016). MS, therefore, stands out as the most common neurological disorder affecting young adults in developed countries.

1.2. Diagnosis and clinical course

MS is difficult to diagnose since there is no single test that can determine that a person has MS. Moreover, the absence of a better explanation for the individual's symptoms remains an essential consideration. Diagnosis is a process that weighs evidence obtained from a clinical examination, medical history, specific laboratory evidence and MRI imaging of the brain and sometimes the spinal cord. It is recommended that MS is best diagnosed by a clinician with MS-related expertise based on the evidence mentioned above. A set of diagnostic criteria has been defined so that it is applied to patients who had suffered a first demyelinating attack, also called clinically isolated syndrome (CIS) patients. Such criteria, named as "McDonald Criteria", was created by the International Panel on diagnosis of MS (McDonald et al., 2001). These guidelines facilitate the diagnosis of MS and are periodically revised (Latest update: Thompson et al., 2018). Briefly, diagnosis of MS requires evidence of the presence of multifocal CNS lesions scattered in time and space. Such CNS lesions and their dissemination in time and space are usually observed using magnetic resonance imaging (MRI) ("Diagnosing MS," n.d.). On a MRI scan, demyelinated areas show up as white spots or darkened areas, depending on the type of scan that is used (Figure 2). Other diagnosis tools are usually included to support MS diagnosis. For instance, cerebrospinal fluid (CSF) is normally analysed for the presence of CNS-specific oligoclonal bands. Oligoclonal bands are antibodies which are detected by separation of CSF proteins. The presence of two or more oligoclonal bands in the CSF but absent from the patient's serum indicates a local B-cell response and inflammation of the CNS, providing powerful evidence for the diagnosis of MS (Link and Huang, 2006). In particular, immunoglobulin

G (IgG) is detectable in CSF of more than 95% of MS patients while undetectable in serum (Link and Huang, 2006).

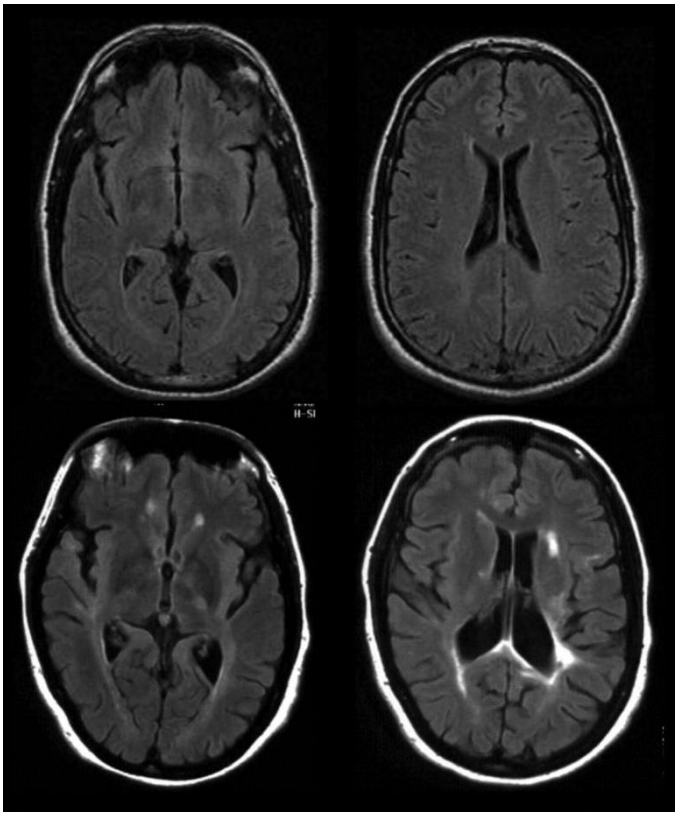


Figure 2. Brain MRI scans. Images are derived from a fluid-attenuated inversion recovery magnetic resonance, a kind of MRI scan. The images show MRI scans of the brain of a MS patient (bottom) and a healthy control (top). White matter MS lesion are shown as white spots against the healthy white matter which looks relatively dark. Images taken from Bakshi et al., 2004.

Once diagnosed, the Kurtzke Expanded Disability Status Scale (EDSS) is used as a measure of physical disability by assigning a severity score (0-10) to the clinical status of the patients (Kurtzke, 1983). The EDSS quantifies disability based on mobility limitations and independence in self-care of MS patients.

According to the clinical course, the disease can manifest itself in different ways (Figure 3) (Goldenberg, 2012; Lublin, 2005). Relapsing-remitting MS (RRMS) is the most common form of the disease (around 85% of MS patients), in which symptoms appear for several days to weeks (relapses), followed by periods of remission when symptoms improve or disappear. As this stage may last for years, tissue damage accumulates over time, changing the pattern of disease from RRMS towards the secondary progressive stage of MS (SPMS). In SPMS, MS-related neurologic deficits gradually worsen over time, with or without periods of remission. About 15% of patients present primary

progressive MS (PPMS), which is characterised by gradually worsening neurologic function from the onset, without early relapses or remissions. Progressive relapsing MS (PRMS) is a rare form, affecting fewer than 5% of patients. It is defined as gradual neurologic worsening from the onset with intermittent relapses along the way with no periods of remission.

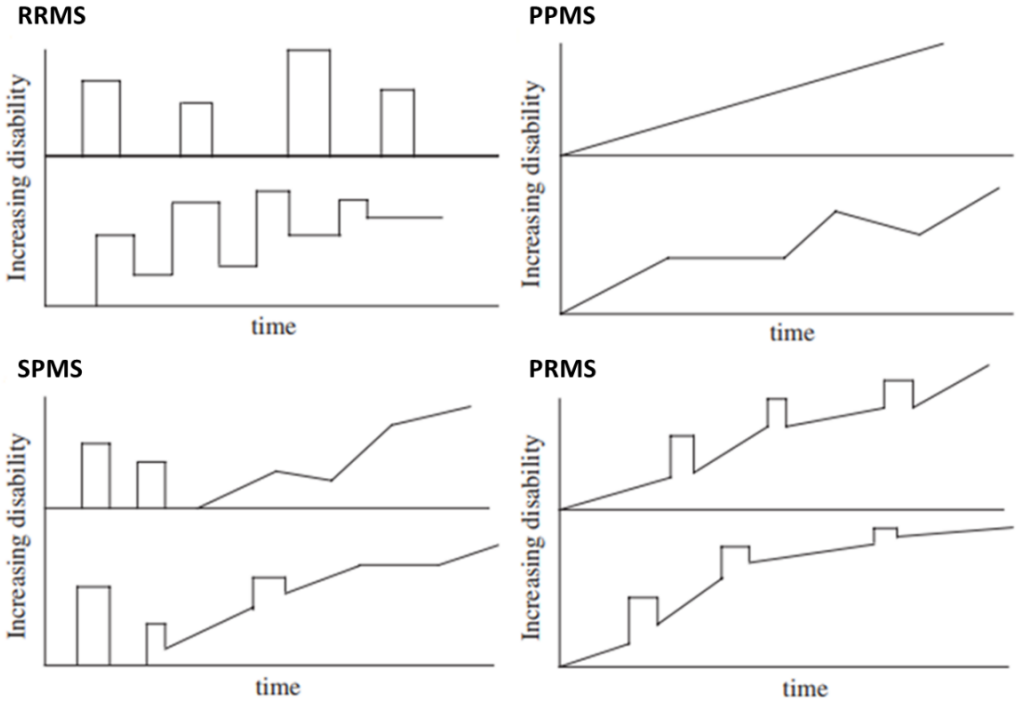


Figure 3. Types of MS. MS patients are grouped into four categories based on the course of disease. RRMS: relapsing-remitting MS; SPMS: secondary progressive MS; PPMS: primary progressive MS; PRMS: progressive relapsing MS. Image adapted from Lublin, 2005.

1.3 MS pathogenesis

The main pathological hallmark of MS is the presence of CNS lesions featuring demyelination, oligodendrocytes loss and infiltration of immune cells. Ample evidence supports the widely held view that pathogenesis of MS involves an autoimmune attack against myelin proteins in the CNS, which subsequently leads to neural degeneration (Garg and Smith, 2015; Sospedra and Martin, 2005). An autoimmune attack consists of a reaction of the immune system against autoantigens in the own patient's body. Although autoimmunity is a hallmark of a variety of diseases including MS and rheumatoid arthritis, its causes are presently unknown. In MS, somehow cells of the immune system (T cells, B cells and macrophages among others) are activated, starting the autoimmune response (Chaplin, 2010; Medina, 2016). Whether this inflammatory

response is directed against myelin or oligodendrocytes is presently debated (Stadelmann et al., 2011). Some studies support that demyelination precedes oligodendrocyte loss in both MS and its animal model, experimental autoimmune encephalomyelitis (EAE) (Romanelli et al., 2016). By contrast, genetically induced loss of oligodendrocytes was sufficient to trigger a CD4 T cells-mediated autoimmune response against myelin and subsequent demyelination in a mouse model (Traka et al., 2016).

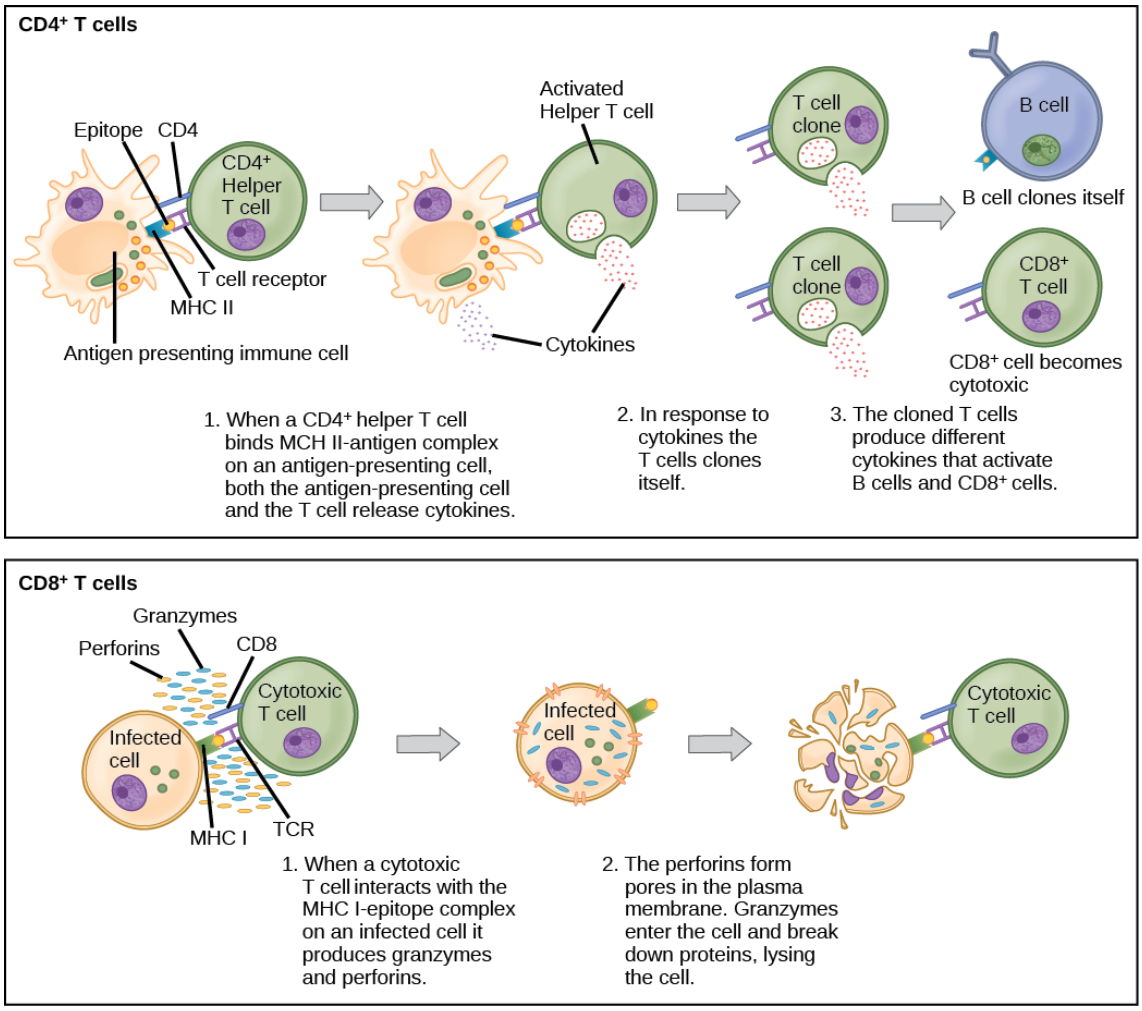


Figure 4. Antigen presentation and T cells responses. Activation of CD4 and CD8 T cells is mediated by the major histocompatibility complex (MHC) present on antigen presenting cells (APC). CD4 and CD8 T cells recognise antigenic peptides associated to MHC II or MHC I molecules respectively. As a result of these interactions, CD4 T cells mature into helper T cells which will collaborate in raising antibody responses and activate CD8 T cells. CD8 T cells, on the contrary, will then differentiate into cytotoxic T cells which will destroy infected cells. Image taken from Wikimedia Commons, the free media repository (Figure 42 02 04).

Immune system mainly protects us against infections. After infection, T cells are activated when they recognise antigens associated with molecules of the major histocompatibility complex (MHC), present on antigen-presenting cells (APCs). The MHC complex is a set of surface proteins that binds antigens, a combination which in turn is recognized by T cells. There are two types of MHC complex proteins: class I proteins which are expressed by most cell types, and class II proteins whose expression is restricted to certain immune cells. In turn, there are two kinds of T cells, CD8 and CD4 T cells, characterized by the surface expression of either CD8 or CD4 molecules. While CD8 T cells only recognise antigens associated with MHC class I in APCs, CD4 T cells exclusively recognise antigens associated with MHC class II in APCs (Figure 4) (Chaplin, 2010; Medina, 2016). Once activated, CD8 and CD4 T cells vary in their mechanism of action. Activated CD8 T cells release different cytotoxic factors and cytokines, which destroy infected APCs. By contrast, stimulated CD4 T cells produce various cytokines which promote activation of MHC class II-containing APCs, such as macrophages and B cells. Subsequently, activated macrophages function destroying pathogens, and activating B cells to produce antigen-specific antibodies (Figure 4) (Chaplin, 2010; Medina, 2016).

In the context of MS, the most widely accepted hypothesis suggests that the early phase of the disease is characterised by the activation of the peripheral immune system (Garg and Smith, 2015; Sospedra and Martin, 2005), although mechanisms that trigger this initial response remain unclear. In this model, it is assumed that leukocytes are activated outside the CNS (Figure 5). Activated leukocytes (specially macrophages, T cells and B cells), subsequently, express specific surface adhesion molecules and release proinflammatory cytokines that affect the endothelium of the blood-brain barrier (BBB) and cause its disruption. BBB is the only barrier separating the CNS from the periphery, preserving local homeostasis and an optimal environment for brain function. Once the BBB is altered, peripherically activated T cells, B cells and macrophages migrate into the brain (Minagar and Alexander, 2003). After entering the CNS, the macrophages, T cells, and antibodies secreted by B cells trigger a local immune response that ends up causing demyelination and the subsequent neuronal/axonal injury (Figure 5).

The most widely accepted hypothesis suggests that autoimmunity in MS is CD4 T cells-mediated, and that autoimmune reactions are directed particularly at myelin proteins (Garg and Smith, 2015; Sospedra and Martin, 2005). Within the CNS, activated CD4 T cells encounter their specific myelin epitopes, presented by MHC class II of local APCs such as dendritic cells, macrophages/microglia, and B cells. Such encounters lead to a massive destructive inflammatory response, mediated by the local release of cytokines and chemokines, proliferation of T and B cells and reactivation of macrophages and microglia. This inflammatory cascade attracts other cells of the immune system and, as the ultimate consequence, the myelin-oligodendrocyte complex is attacked resulting in damage to the myelin sheath, loss of oligodendrocytes and altered neuronal transmission (Figure 5).

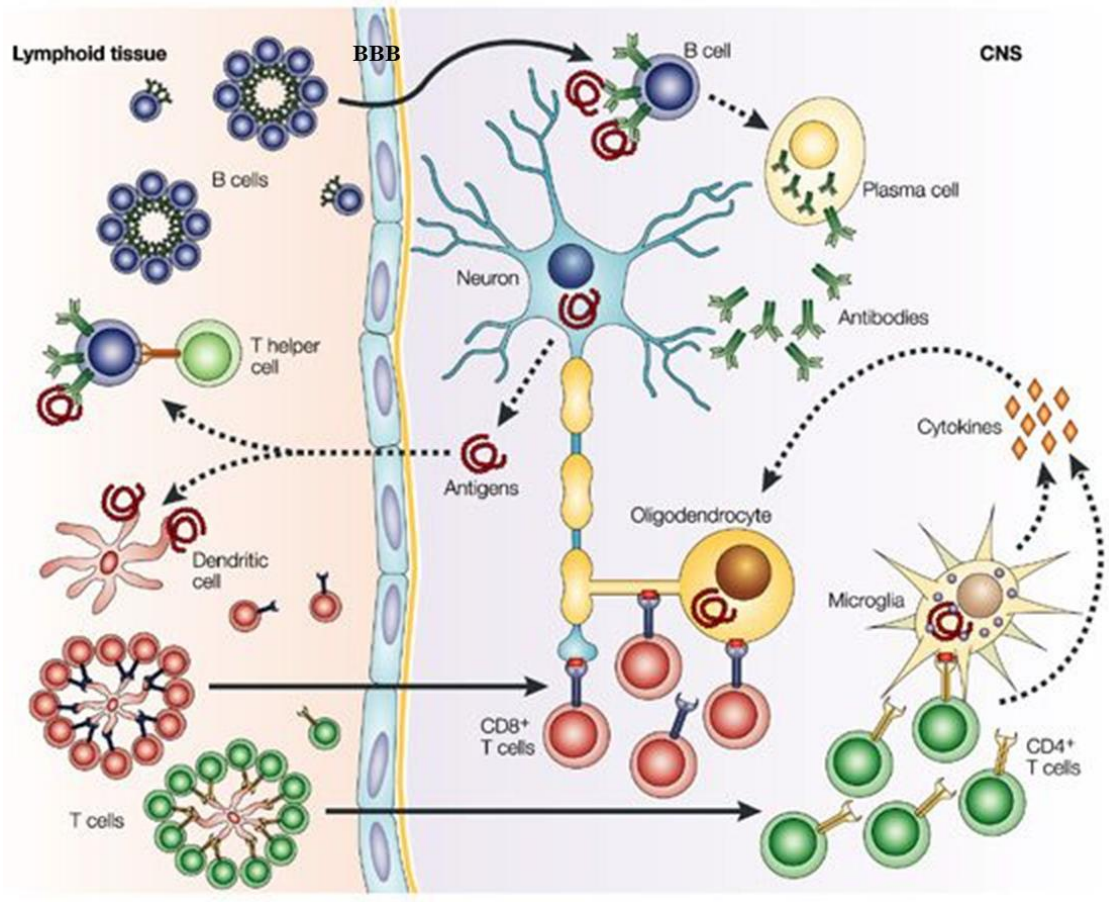


Figure 5. Hypothetical immune responses in multiple sclerosis. After peripherical activation of immune cells (T cells, B cells and macrophages), such cells mediate blood-brain-barrier (BBB) disruption and migrate into the central nervous system (CNS). Activated B cells produce antigen-specific antibodies. Activated CD4 T cells recognise specific antigens presented by microglia, and both CD4 T cells and activated microglia release cytotoxic cytokines that attack oligodendrocytes. Activated CD8 T cells directly attack axons and oligodendrocytes. Image adapted from Hemmer et al., 2002.

This view is based on all major characteristics of CD4 T cells in MS. To begin with, active myelin-specific T cells and their products have been detected in peripherical blood mononuclear cells (PBMCs), cerebrospinal fluid (CSF) and cerebral lesions of MS patients (Allegretta et al., 1990; Elong Ngono et al., 2012; Olsson et al., 1992; Wu and Alvarez, 2011). Interestingly, activation of autologous myelin-specific T cells is enough to CNS inflammation animal models (Genain et al., 1994; MeinL et al., 1997; Stromnes and Goverman, 2006). Medication-based studies have revealed that activation of myelin-specific T cells in the periphery resulted in damage to the CNS of MS patients (Bielekova et al., 2000). Variations in the genes encoding MHC class II molecules represent the strongest human genetic risk factor for MS cells (GAMES and Transatlantic Multiple Sclerosis Genetics Cooperative, 2003). This provides indirect support for the importance of CD4 T cells as MHC class II molecules participate in the antigen presentation and CD4 T cells activation. And in addition, peripherically activated T cells release the cytokines that contribute to the disruption of the BBB and migrate into the CNS more efficiently than other immune cells (Minagar and Alexander, 2003).

Also, although it is widely accepted that myelin-reactive T cells are the main players in the immunopathogenesis of MS, the exact triggers that initially stimulate such cells have not yet been identified. Nevertheless, several candidates have been studied. T cell from MS patients recognise several myelin proteins such as myelin basic protein (MBP), proteolipid protein (PLP), myelin oligodendrocyte glycoprotein (MOG) and myelin-associated oligodendrocytic basic protein (MOBP) among others (Bernard et al., 1997; Kaushansky et al., 2010; Zhang et al., 1994). However, CNS myelin antigens are unlikely to be present in the periphery since they are synthesized by oligodendrocytes that only reside in the CNS. The peripheral activation of myelin-reactive T cells may be triggered by myelin proteins MBP and PLP that reside in cervical lymph nodes (CLNs) (Fabriek et al., 2005). Since CLNs are located in the neck, CLNs-associated myelin proteins could trigger a peripherical T cell activation in MS patients. Another mechanism suggested is based on molecular mimicry, by which peripherical viral or bacterial antigens that have sufficient sequence similarity with autoantigens would be

able to cross-activate T cells specific for myelin epitopes (Figure 6) (Stinissen and Hellings, 2008; Wucherpfennig and Strominger, 1995).

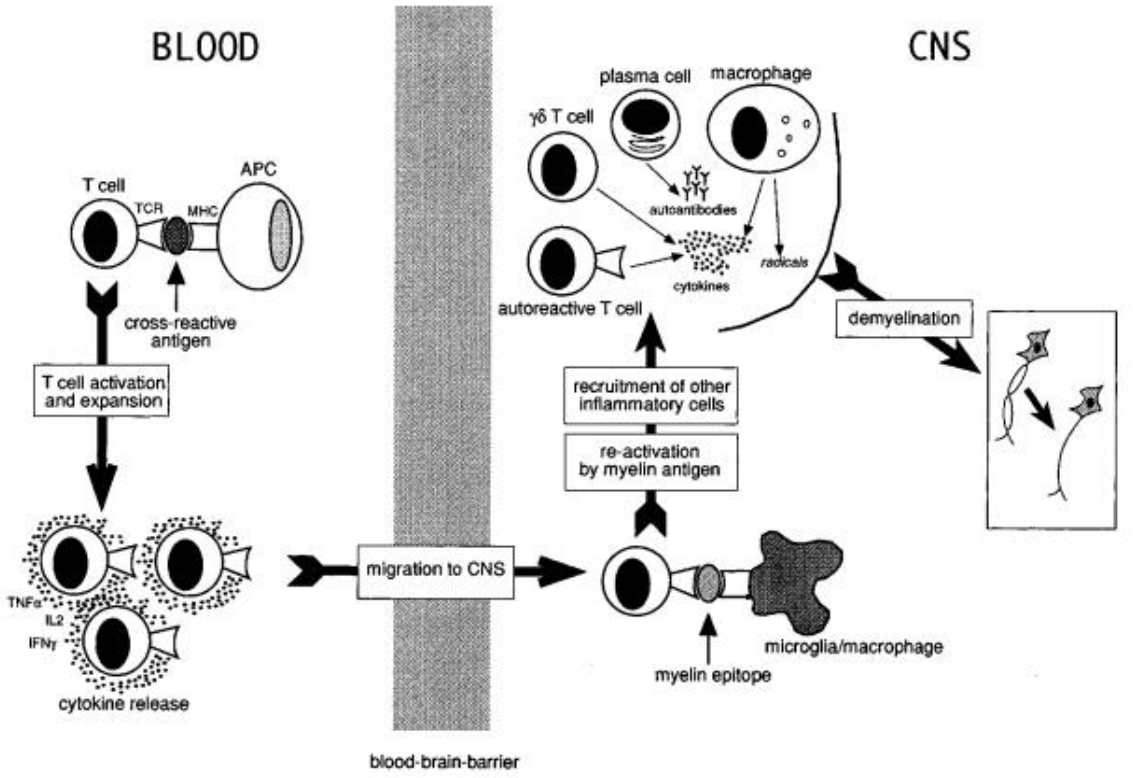


Figure 6. Hypothetical cross-activation of myelin reactive T cells in the periphery. Activation of myelin reactive T cells by molecular mimicry would led to T cells expansion and migration into the central nervous system (CNS) in the absence of myelin proteins. Within the CNS, they would become reactivated when they encounter their specific myelin epitope presented by local antigen presenting cells (APC) such as microglia or macrophages. TCR: T cell receptor. MHC: major histocompatibility complex. Image reproduced from Stinissen and Hellings, 2008.

Beyond the CD4 T cells paradigm, there is also evidence to incriminate CD8 T cells in MS pathology (Friese and Fugger, 2009). Several studies have shown that, in fact, CD8 T cells are the T cell-type more abundantly encountered in CNS lesions, blood and CSF of MS patients, beyond CD4 T cells (Babbe et al., 2000; Booss et al., 1983; Jilek et al., 2007; Traugott et al., 1983). Furthermore, CD8 T cells can induce axonal damage by direct injury to MHC class I APCs such as neurons and oligodendrocytes (Batoulis et al., 2010). In fact, axonal damage within CNS lesions of MS patients, correlates with the number of infiltrating CD8 T cells rather than CD4 T cells (Bitsch et al., 2000; Kuhlmann et al., 2002), although not necessarily in all cases.

B cells also play an important role in MS pathology. B cells can act as APCs for T cells, and produce myelin-specific antibodies (Sospedra and Martin, 2005). The increased presence of specific antibodies in CSF, but not in the serum, of MS patients is one of the main diagnostic tools for MS diagnosis (Thompson et al., 2018). B cells and myelin-specific antibodies have also been detected in lesions of MS patients. Although the antigen specificity of antibodies in MS patients is in some cases not well established, anti-MOG and anti-MBP antibodies have been found with certainty in human MS lesions (Genain et al., 1999; O'Connor et al., 2005; Wucherpfennig et al., 1997).

1.4. Etiology

As of today, the etiology of MS has not been precisely defined. Extensive evidence supports that MS is a multifactorial disease, which results from complex interactions between genetic and environmental factors. Although MS is not considered a hereditary disease, studies support a genetic susceptibility for the disease (Milo and Kahana, 2010). Around 20% of MS patients have one affected relative and MS incidence is increased up to 10 times when both parents are affected (Dyment et al., 1997; Sawcer et al., 1997). Efforts have been made to discover gene variants associated with MS risk. Meta-analysis combining the results of nine whole genome screens for MS-associated genomic regions, revealed that the genetic factor with the strongest association with MS was the human leukocyte antigen gene complex (HLA) class II locus on chromosome 6p21 (GAMES and Transatlantic Multiple Sclerosis Genetics Cooperative, 2003). HLA complex is the human version of the MHC. This was later fine-mapped to HLA-DRB1 gene, or more specifically to some HLA-DRB1 haplotypes. Among them, HLA allele DRB1*1501 accounts for near 50% of the total genetic basis of MS. Subsequent genome-wide association studies (GWASs) yielded the identification of more MS susceptibility *loci*, and many are related to lymphocyte function and/or cytokine pathways (International Multiple Sclerosis Genetics Consortium et al., 2011, 2007; International Multiple Sclerosis Genetics Consortium (IMSGC) et al., 2013). Altogether, several MHC class I and II alleles, and 110 non-MHC variants have been associated with MS risk (Moutsianas et al., 2015). However, no single-gene defect or allele that causes MS has been identified so far and GWASs have

not yielded reliable genetic predictors of the disease. Therefore, other non-genetic causes must contribute to MS risk.

A large body of data supports a contribution of environmental variables to MS. MS frequency increases progressively with geographic latitude and decreasing solar radiation (Kurtzke, 2005). A potential explanation for this association is the involvement of vitamin D, which is produced in the skin by exposure to ultraviolet radiation or supplied in the food. It turns out that both sunlight exposure and vitamin D levels have been inversely associated with the risk of MS (Hayes, 2000). Low levels of vitamin D in serum has been associated with high risk of developing MS, while vitamin D supplementation reduces the risk (Munger et al., 2006, 2004). However, vitamin D levels by itself cannot predict MS risk either. Other environmental factors that have been reported to modulate MS risk, and approximately correlate with geographical distribution, include psychological stress, pollution, dietary fat and smoking (Milo and Kahana, 2010). Smoking, in particular, has emerged as one of the most common environmental factors associated to MS along with latitude and vitamin D levels. Smoking is associated with a higher risk of developing MS, and also correlates with a faster progression of the disease (Hernán et al., 2001).

Proof for additional environmental factors has been obtained in migration studies. Among immigrants that move from an area of higher prevalence of MS to an area of lower prevalence, the prevalence trends towards the lower rate (Dean, 1967). Genetic susceptibility alone, cannot explain such changes in MS incidence. As alternative explanations, the changes have been attributed to sunlight exposure and subsequent vitamin D synthesis. However, similar studies, in which the age of migrants was considered, revealed that immigrants only acquire the low MS risk when migrating before age 15–16, while migration at a later age did not change the risk (Hammond et al., 2000). A plausible explanation for these age-related changes in MS risk is the apparent MS susceptibility conferred by some viral infections. Among the viral agents associated with MS are human endogenous retroviruses (HERVs) (Morandi et al., 2017b), whose association with MS will be further discussed later (page 47). The risk of multiple sclerosis is increased in persons with prior infectious mononucleosis caused by Epstein–Barr virus (EBV) (Ahlgren et al., 2009; Nielsen et al., 2007; Ramagopalan et

al., 2009). Distribution of EBV is widespread, infecting over 95% of the adult population. EBV infection is usually acquired at an early age when the infection is mostly asymptomatic. However, EBV infection in adolescence is commonly known as mononucleosis and triggers a severe inflammatory immune response. After mononucleosis infection is controlled by the immune system, a latent infection remains in the form of infected B cells, that do not produce virus. Although the mechanism is unknown, it is believed that this mononucleosis-mediated activation of immune system increases susceptibility to MS (Ascherio and Munger, 2010). Yet another alternative hypothesis is based upon differential exposure to EBV dependent on latitude. Children born in developed countries (northern latitudes) would tend to experience later exposure to EBV infection than children from undeveloped countries (southern latitudes), and therefore, they face a higher MS risk.

In brief, no single gene or environmental factor has been unambiguously identified as the causative agent of MS. It is likely that the combinatorial effects of several genetic and environmental factors lead to the onset of the disease.

1.5. Strategies to study MS disease etiology, pathology and progression

Epidemiological and genetic studies of patient cohorts have highlighted the influence of environmental factors (solar radiation, vitamin D levels, the incidence and timing of viral infections, smoking habits) and the genetic background, on MS risk. However, the main injury in MS occurs in the CNS, while the peripheral nervous system and other organs are not targeted by the disease process. For obvious ethical reasons, *in vivo* observation of CNS lesions is almost impossible, and the histopathology of CNS lesions has been mainly studied in post-mortem brain samples of patients. Animal models have been useful to better understand etiology and progression of the disease.

In vivo observation of CNS lesions in MS patients has been mostly limited to the use of MRI. MRI scanning allows to diagnose and evaluate MS progression through the observation of demyelinating lesion in the brain (Absinta et al., 2016; Harrison et al., 2015; Louapre et al., 2015). However, *in vivo* data related to the initiation and progression of MS are hard to obtain and changes in brain-specific

immunohistopathology, epigenetics and immune processes cannot be followed in patients.

Understanding of the inflammatory response in MS has been greatly shaped by findings from animal models. Experimental autoimmune encephalomyelitis (EAE) is the most commonly used animal model for disorders characterised by inflammation, myelin damage and neurodegeneration, including MS. To create such a model, already in the first half of the last century, different susceptible animal species were immunised with brain extracts, Freund's adjuvant (CFA) and pertussis toxin (PT) (Freund and McDermott, 1942; Kabat et al., 1947; Munoz et al., 1984). This resulted in alterations of the immune system and neurological disorders. Among all immunised animals generated, mice arose as the best host for an experimental animal model of MS due to the availability of genetically engineered animals, which allows for direct testing of the involvement of MS-specific proteins in EAE pathology or genetic ablation of cell lineages. Susceptible mice, such as SJL/J, C57BL/6 and NOD, were later immunised with human recombinant myelin-associated peptides such as MBP and PLP (Pettinelli and McFarlin, 1981; Zamvil et al., 1985). Either treatment caused EAE pathogenesis. Briefly, immunization triggers the activation of T cells in the periphery. These activated T cells produce interferon-gamma (IFN-gamma) and inflammatory cytokines, which activate other immune cells such as B cells and monocyte/macrophages. All immune cells migrate into the CNS and, along with activated resident glial cells, damage the myelin and axons. The release of cytokines attracts further inflammatory cells to the CNS, thus perpetuating the inflammatory cascade (Constantinescu et al., 2011; Duffy et al., 2014). The symptomatic result is an MS-like progressive ascending paralysis starting from the tail. Immunisation usually gave rise to an acute phase in which animals may recover (remission phase) or not (chronic disease). Models in which animals do recover from the acute phase, or develop subsequent relapses and remissions, have been also developed since they clearly are more useful to study mechanisms behind relapsing-remitting MS (Al-Izki et al., 2012; Baker et al., 1990). In parallel, models in which the neurological deficits accumulate to progressive disability (called secondary progressive EAE), better mimic the clinical and pathological aspects of SPMS (Hampton et al., 2008).

Over the years, EAE models have allowed to study the complex interaction between a variety of immunopathological and neuropathological features of MS: inflammation, demyelination, axonal loss or damage and gliosis. For instance, the involvement of T cell in MS-like pathogenesis has been studied in EAE models, allowing the dissection of the pathogenic role of T cell response to MOG in MS-like pathogenesis. (Mendel et al., 1995; Smith et al., 2005). Staining of EAE brain slices confirmed that CD8-positive T cells are targeted against oligodendrocytes and induce oligodendrocyte apoptosis (Göbel et al., 2009). EAE models also helped to shed light on the role that cytokines play in MS-like pathogenesis, analysing among others, the role of the related molecules IL-12 and IL-23 in the susceptibility to autoimmune demyelination (Becher et al., 2002; Cua et al., 2003).

However, although EAE has been a valuable model for studying MS, not all aspects of the human pathology are represented in the model (Baker et al., 2011; Vesterinen et al., 2010). The main difference between MS and EAE is that animal models require an artificial external immunization step, whereas MS patients suffer a spontaneous initial trigger. The inducing antigens in EAE are known, whereas triggers for MS remain unknown. Several aspects of the disease are not mimicked in EAE models. A major drawback of EAE-like models is that responses to therapies often cannot predict efficacy in humans (Constantinescu et al., 2011). Another relevant question that EAE models cannot address is related to complications restricted to humans, for example the implication of human-specific virus in the pathogenesis of MS. A main example is again EBV (Ascherio and Munger, 2010; Serafini et al., 2007). These combined drawbacks of the mouse model have increased interest in alternatives such as human-derived *in vitro* models.

2. Human stem cells

There are three characteristics that set stem cells apart from other kinds of cells in the body: they are capable of dividing and renewing themselves, they are unspecialized, and they can give rise to specialized cell types (Alison et al., 2002; Gardner, 2002). Unlike specialized, differentiated cells such as muscle cells, blood cells, or nerve cells, which do not normally replicate themselves, stem cells are capable of long-term self-

renewal. This ability allows them to proliferate without losing their differentiation potential and without undergoing senescence (Amit et al., 2000). They can divide symmetrically (in which both daughter cells are either stem cells or differentiated cells), or asymmetrically (yielding one stem cell and one more differentiated cell). The potential to differentiate into other cell types is called developmental potency. Stem cells are endowed with varying potency, the more cell types a stem cell can differentiate into, the greater its potency (De Los Angeles et al., 2015; Singh et al., 2016). Thus, stem cells can be either totipotent (they can differentiate into embryonic and extraembryonic cell types), pluripotent (descendants of totipotent cells, they can differentiate into cells representing any of the three primary germ layers of the early mammalian embryo), multipotent (they can only make cells within a given germ layer), or unipotent (they can make just cells of a single cell type).

Depending on their origin, stems cells can be divided into human embryonic stem cells (hESCs) and tissue-specific or adult stem cells. Adult stem cells are present in most or all human organs and tissues including brain, bone marrow, blood vessels, heart, liver and umbilical cord. Compared to embryonic stem cells, they have less lineage potential and are generally multi- or unipotent. Their main role is as a source of newly generated cells, which are required in the process of regular turn-over in tissues. Therefore, they are essential for maintaining and repairing their tissue of residence (Fuchs, 2008; Montagnani et al., 2016; van der Flier and Clevers, 2009). hESCs are derived from the inner cell mass of the blastocyst-stage embryo, which in human forms three to five days after an egg cell is fertilized by a sperm (Reubinoff et al., 2000; Thomson et al., 1998). Due to the ethical issues of destroying an embryo, hESCs have been mainly derived from surplus frozen embryos that remain after a woman or couple has completed infertility treatment (Lo and Parham, 2009). The cells are pluripotent and provide a renewable resource for studying normal development and disease, and for testing drugs and other therapies.

2.1. Human embryonic stem cells and induced pluripotent stem cells

The pluripotency of hESCs is driven by several molecular mechanism which regulate the expression of genes responsible for the maintenance of pluripotency and

repression of differentiation (De Los Angeles et al., 2015). A defined set of transcription factors (i.e. OCT4, SOX2, and NANOG) has an essential role in maintaining the self-renewal and pluripotency of hESCs, which is regulated by a network of auto-regulatory and feed-forward loops (Boyer et al., 2005; Singh et al., 2016). In fact, some of these pluripotency transcription factors overlap with the ones used to induce somatic cells reprogramming to induced iPSCs (Takahashi et al., 2007). Apart from transcription factors, pluripotency of hESC is also been linked to the presence of other molecular markers such as *REX1, NANOG, OCT4* and *SOX2* (Chen and Lai, 2015).

hESCs can be cultured on either a feeder layer (usually proliferation-deficient mouse embryonic fibroblasts) or under feeder-free conditions (Kent, 2009; Lee and Lee, 2011). When cultured on feeder-free conditions, they must be grown on a matrix including Matrigel, vitronection, laminin, fibronectin, or collagen IV among others (Xu et al., 2001). When culturing hESC, specific growth factors must be added to the culture medium to maintain cell pluripotency. For instance, addition of basic fibroblast growth factor (FGF2) and noggin suppresses bone morphogenetic proteins (BMPs) signalling that would otherwise induce differentiation (Wang et al., 2005, p. 2; Xu et al., 2005).

Over the past 12 years, it has become clear that differentiated, specialized cells can be reprogrammed to resemble cells with expanded developmental potency (reviewed in Shi et al., 2017). Reprogramming was first achieved by nuclear transfer, and later by introduction of transcription factors that induce pluripotency (Takahashi et al., 2007). Reprogrammed cells acquire a pluripotent state and are therefore named induced pluripotent stem cells (iPSCs). This can be achieved in virtually any somatic mammalian cell type, and some of the transcription factors most frequently used are OCT4, SOX2, NANOG, and LIN28 (Stadtfeld and Hochedlinger, 2010; Takahashi et al., 2007). Human iPSCs possess the same differentiation potential as do hESCs.

2.2 Applications of hESCs and hiPSCs

hESCs can differentiate into most if not all cell types of an organism (Figure 7). Therefore, controlled *in vitro* differentiation of hESCs can provide a source of differentiated cells and tissues to replace damaged cells for diseases in which native

cell types are inactivated or destroyed, such as diabetes and heart disease (Kumar et al., 2010; Nadig, 2009). In this context, iPSCs have an important advantage over hESCs. In case of transplantation, tissue derived from patient-specific iPSCs would suffer less rejection than tissue derived from hESCs. Nevertheless, efficient stem cell-based therapy would be restricted to diseases characterised by the loss of a main cell type as opposed to more complex diseases such as cancer (Cherry and Daley, 2013). In parallel, cell types such as cardiomyocytes and hepatocytes generated from hESCs provide ideal populations for predictive toxicology. The use of hESCs, therefore, can can reveal toxicity of new drugs not detected using animal models (Bongso and Richards, 2004; Keller, 2005). Alternatively, hESCs/ iPSCs-based *in vitro* models can be used to study the pathology of diseases which are difficult to access in human patients and/or which are difficult to model in animals (Cherry and Daley, 2013). For instance, introduction of targeted mutations into hESCs allowed the generation of hESC-based models of monogenic cardiac diseases (Moretti et al., 2013). Manipulation of the genome of the hESCs has also provided a model for the Lesh-Nyhan syndrome (Urbach et al., 2010). This model was generated introducing a mutated form of the disease gene HPRT1, and the resulting lines exhibited some of the attributes of the disease.

With the arrival of iPSCs, patient- and disease-specific stem cells are now being generated. Their availability therefore represents a powerful tool for biomedical research and may provide patient-specific *in vitro* models for tissues that are otherwise experimentally not accessible. Human iPSCs from patients with a range of human diseases have already been stablished (Park et al., 2008). For instance, Parkinson's patient-specific iPSCs have been generated and subsequently differentiated into the cell type mostly affected in patients i.e. dopaminergic neurons (Soldner et al., 2009). Similarly, iPSCs generated from patients with amyotrophic lateral sclerosis can be differentiated into motor neurons (Dimos et al., 2008). The generation of iPSCs from a patient enables therefore the production of the cell types affected by that patient’s disease. These cells could in turn be used for disease modelling and drug discovery.

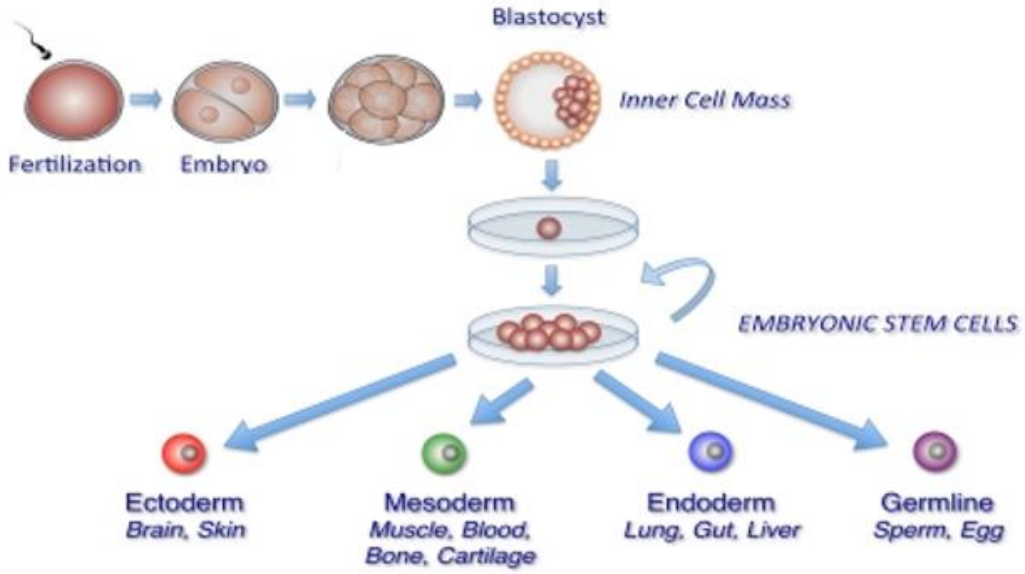


Figure 7. hESC derivation and differentiation. hESC are derived from the inner cell mass of a blastocyst and can self-renew indefinitely *in vitro,* as well as forced to assume more specialized functions cell fates associated with specific differentiated cell types. Adapted from Yabut and Bernstein, 2011.

2.3. In vitro models for multiple sclerosis and others neuroinflammatory disorders

In vitro culture of CNS cells has enormous advantages. And it turns out, a lot of efforts has been invested over the past years to devise and optimise culture systems for all the main cell types of the CNS (Gibbons and Dragunow, 2010; Kipp et al., 2012; van der Star et al., 2012). *In vitro* culture of mature neurons, microglia and oligodendrocytes is particularly challenging since these cell types do not proliferate easily. Cell lines provide unlimited supplies of homogeneous cells, but do not faithfully represent alterations observed in pathologies. While the preparation and culture of primary cells is much more challenging and provides only limited amounts of cells.

Both hESCs and iPSCs can be differentiated into neuronal, microglial and oligodendrocyte precursors. Different authors claim that subsequent differentiation into mature cell types can be achieved using different protocols (Douvaras et al., 2017; Gunhanlar et al., 2018; Hu et al., 2009). Mature differentiated cells are characterised by the expression of typical markers, and functional assays. (Abud et al., 2017; Hu et al., 2009). However, although some of these differentiation protocols seem to be very promising, the truth is that they are difficult to carry out, expensive, time-consuming

and they are not very efficient nor easy to reproduce (from my personal experience, (see Manuscript 4) and personal communications (Anna Duarri Piqué)).

Interestingly, patient iPSCs-derived CNS cells have been successfully generated and used as models of neurological disease such as MS, Alzheimer's disease, Huntington's disease and Parkinson's disease (Ghaffari et al., 2018). The successful generation of iPSCs from MS patients was demonstrated few years ago (Douvaras et al., 2014; Song et al., 2012). MS patient-derived iPSCs have been subsequently differentiated into mature astrocytes, oligodendrocytes and neurons which, by definition, contain the patient's individual genetic susceptibility to MS. For instance, iPSCs-derived neural progenitor cells (NPCs) from MS patients showed inherent defects in myelin injury response *in vivo* when compared to control NPCs (Nicaise et al., 2017).

2.4. Generation of human oligodendrocytes from human pluripotent stem cells

Generally, *in vitro* differentiation is achieved by adding a cocktail of small molecules and growth factors to the stem cell culture in order to induce a specific-cell type differentiation (Singh et al., 2016). It became clear that, applying factors such as FGF2 or the epidermal growth factor (EGF) to hESCs, they could directly differentiate *in vitro* into multipotent neural precursor cells (NPCs), which could subsequently be differentiated into more specified cell types, including astrocytes, oligodendrocytes, and neurons (Reubinoff et al., 2000; Zhang et al., 2001). These NPCs were also able to differentiate into astrocytes, oligodendrocytes, and neurons after transplantation into newborn mouse brains (Reubinoff et al., 2000; Zhang et al., 2001). In all cases, differentiation into the specific cell types were confirmed by the expression of cell type-specific markers, such as expression of the neural cell adhesion molecule (NCAM) in neurons, glial fibrillary acidic protein (GFAP) in astrocytes, and MBP in oligodendrocytes. However, it was not until some years later that a study demonstrated that hESCs-derived oligodendrocytes could resemble functional oligodendrocytes after transplantation in the *shiverer* mouse model of dysmyelination (Nistor et al., 2005). Oligodendrocytes are mature or functional when they have the ability to produce myelin. And it turned out that, while *shiverer* mice did not contain

myelin, MBP immunostaining confirmed myelination by transplanted cells in transplanted *shiverer* mice (Nistor et al., 2005).

In subsequent years, alternative protocols to generate oligodendrocytes from human pluripotent stem cells were developed (Hu et al., 2009; Kang et al., 2007; Kerr et al., 2010; Nistor et al., 2005). Although quite similar, these mentioned protocols differ in the cell line used, culture time, culture method and the growth factors used (reviewed in Alsanie et al., 2013). At any rate, the first step of oligodendrocyte differentiation consists of differentiating hESC/iPSCs into neural precursors. This can be achieved through the generation of embryoid bodies (EBs) in suspension, or in a monolayer of cells by addition of supplements (FGF2, N2, B27…). NPCs are subsequently differentiated into glial precursors cells (GPCs), which are subsequently converted into the different oligodendrocyte stages, consisting of the pre-progenitors, progenitors, pre-oligodendrocytes and mature oligodendrocytes (Figure 8). This conversión is induced by the addition of specific inducers (mostly growth factors) at each stage. Each of these stages can be identified *in vivo* and *in vitro* by the expression of specific markers. These markers can be either transcription factors that regulate neural and oligodendrocyte maturation or growth factors that signal through extracellular receptors. Each marker is expressed at specific and defined stages during the progressive differentiation of human pluripotent stem cells (hESCs) into myelinating oligodendrocytes. Some of the commonly used markers during human oligodendrocyte differentiation are: *SOX1* and *PAX6* in NPCs, *OLIG2* in GPCs, *NG2* in OPCs and *O4* and *Myelin oligodendrocyte glycoprotein (MOG)* in late OPCs and mature oligodendrocytes.

However, these protocols yield low number of mature OPCs. Thus, follow-up studies focused on refining the induction protocols, improve cell growth, reduce the number steps required, in order to accelerate the process and achieve higher proportions of $O4^+$ OPCs and their derived oligodendrocytes (All et al., 2015; Douvaras and Fossati, 2015; Ehrlich et al., 2017; Stacpoole et al., 2013). Recent studies showed that differentiation of hESCs/hiPSCs to oligodendrocytes can be enhanced by overexpression of specific transcription factors, which allows for more efficient generation of O4+ oligodendrocytes. In fact, the overexpression of transcription

factors (such as SOX10, OLIG2, and NKX6.2) in hESCs/iPSCs-derived NPCs resulted in the generation of $O4^+$ oligodendrocytes within a month (Ehrlich et al., 2017; García-León et al., 2018).

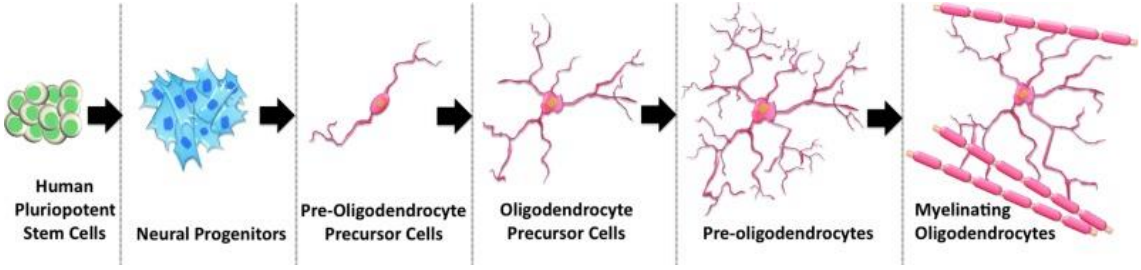


Figure 8. Oligodendrocyte differentiation steps. Schematic of the progressive differentiation of human pluripotent stem cells to myelinating oligodendrocytes. Protocol begins with human pluripotent stem cells which are differentiated into neural progenitor cells (NPCs). Then, the different stages of oligodendrocyte development are induced: pre-progenitor, progenitors, pre-oligodendrocyte and mature oligodendrocytes. Adapted from Alsanie et al., 2013.

3. Epigenetics

Although all cells from a multicellular organism contain the same genetic information, they differ in structure and function because there is a strict regulation of gene expression within each cell type. For instance, in the context of tissue development, some genes are only expressed at a specific developmental timepoint or in a specific cell type. Epigenetic processes allow the regulation of gene expression through changes in the DNA structure and organization, without altering the DNA sequence itself (Dupont et al., 2009; Goldberg et al., 2007). As epigenetics was initially studied in the context of cellular differentiation, the term usually refers to heritable modifications of DNA and chromatin that are passed on, through mitosis, to subsequent cell generations.

DNA is highly compacted and organized within the nucleus (Figure 9). The DNA helix is wrapped around histone proteins giving rise to the DNA-histone complex, which in combination with additional DNA binding proteins is known as chromatin. Chromatin fibres, in turn, are further compacted into chromosomes (Lodish et al., 2000a). Importantly, accessibility of DNA is a major determinant of activity status. While transcriptional activation usually coincides with accessible chromatin, lack of accessibility of positively acting regulators correlates with transcriptional silencing. Therefore, variations in DNA structure or in the structure or activity of associated

proteins, either by chemically altering the DNA itself or modifying the proteins responsible for unpacking the DNA, can influence gene expression. Among the better known mechanisms which are implicated in epigenetic regulation, and able to influence DNA accessibility and the activity of DNA polymerase are DNA methylation and histone modifications (Qiu, 2006).

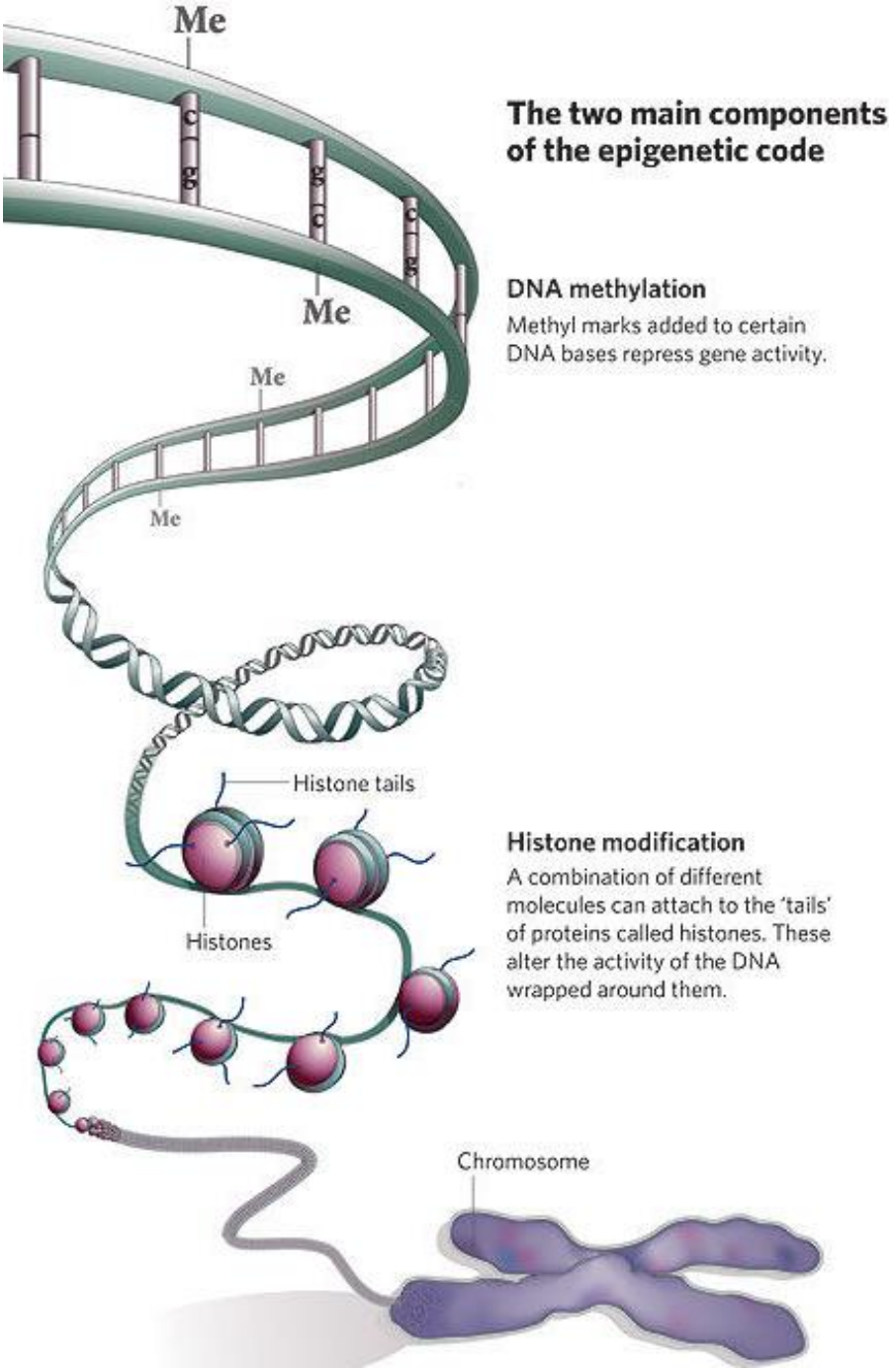


Figure 9. Overview of the organization of DNA into chromosomes. There are two main components of the epigenetic processes: DNA methylation and histone modifications. Image taken from Qiu, 2006.

3.1. DNA methylation

In vertebrates, DNA methylation refers to the addition of a methyl group (-CH3) covalently to a cytosine residue. In the vast majority of methylated cytosines, this modification occurs in at CpG dinucleotides (A Jones, 2012) (Figure 10). The distribution of CpGs across the genome is not random; genomic regions enriched in CpGs, called CpG islands, are usually found at gene promoter sequences. Indeed, CpG islands characterize the promoter regions of more than half of all genes in the human genome (Antequera and Bird, 1993). Although most of the CpG islands are methylated (Bird, 2002), CpG islands in promoters of important genes are principally unmethylated, suggesting the influence of DNA methylation in gene regulation by regulating both chromatin structure and accessibility to regulatory factors.

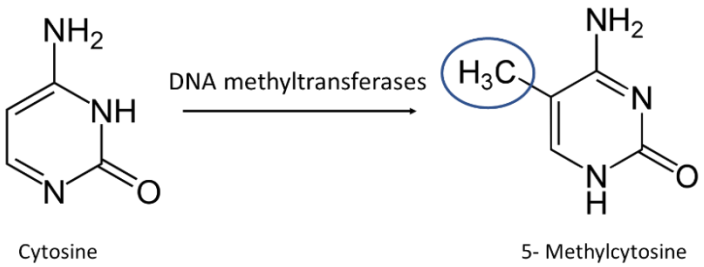


Figure 10. DNA methylation process. DNMTs catalyse the addition of a methyl group to the 5 position of a cytosine in a CpG dinucleotide.

Methylation of DNA is an essential epigenetic control mechanism in mammals. Historically, DNA methylation has been associated with gene repression. DNA methylation of specific promoters often blocks the binding of activating transcription factors. Alternatively, methylated CpGs are recognized by proteins that alter the chromatin structure, turning it more compacted and less accessible to the transcription machinery (Lande-Diner and Cedar, 2005). However, the actual transcriptional effect of DNA methylation at promoter regions is related to the local CpG-density. In general, methylation only has a significant transcriptional effect at high-CpG density promoters, which are usually hypomethylated, and low CpG-density promoters, which are usually methylated. In contrast, intermediate CpG-density promoters can be either methylated or unmethylated and remain transcriptionally active regardless of their methylation state (Hackett and Surani, 2012; Messerschmidt et al., 2014). Besides the role of DNA methylation in transcriptional control at

promoters, DNA methylation also control specific regions. DNA methylation is essential for important processes in mammalian development such as X-chromosome inactivation, transcriptional silencing of endogenous transposable elements, genomic imprinting, long-term gene silencing and regulation of chromatin structure (Bird, 1997; Hore et al., 2007; Lande-Diner et al., 2007).

DNA methylation is catalysed by DNA methyltransferases (DNMTs), particularly DNMT1, DNMT3A and DNMT3B. These enzymes are responsible for regulating normal DNA methylation and, although the mechanism is similar for different DNMTs, they have different responsibilities in various developmental stages within organisms (A Jones, 2012). DNMT1 oversees the maintenance of correct DNA methylation, methylating hemimethylated DNA. On the other hand, DNMT3A and DNMT3B generate new DNA methylation marks (*de novo* DNA methylation). They both are essential for DNA methylation patterns in early development, and their activity contributes to the repression of genes that are no longer required after cell differentiation (A Jones, 2012).

DNA-methylation can be cell type-specific and patterns are often transmitted through cell division. However, DNA methylation is not a static mark and apart from methylating enzymes, demethylating processes are operative. DNA demethylation can occur through passive or active mechanisms (Chen and Riggs, 2011). DNA demethylation can be achieved passively in the absence of methylation of the newly synthesized DNA strand after replication. One passive mechanism, for instance, is the inhibition of DNMT1 by the action of DNMT inhibitors (DNMTi). Active DNA demethylation, on the other hand, is a replication-independent process, mediated by multiple demethylating enzymes. Enzymatic mechanisms result in the direct removal of the 5-methyl group, the 5-methylcytosine base or the 5-methylcytosine nucleotide. Removal of methylated bases is followed by insertion of an unmethylated one using base excision repair machinery (BER), while removal of the entire nucleotide is followed by insertion of a new one by nucleotide excision repair (NER) or mismatch repair (MMR). Both DNMT and DNMTi are essential for maintaining the balance of the DNA methylation level throughout the genome. The gene- or locus-specific alterations that occur by either *de novo* methylation or directed demethylation influence the

acquiescence of cellular identity during differentiation, and its subsequent maintenance across cell divisions.

3.2. DNA methylation in MS and other diseases

Defects and alterations in epigenetic marks have also been associated with a variety of human diseases, such as cancer, neurological disorders and autoimmune diseases (Baylin and A Jones, 2011; Portela and Esteller, 2010). Epigenetic alterations may predispose to subsequent pathological insult or represent the secondary result of disease processes. When it comes to cancer, although research has been focused for many years on genetic mutations, genome-wide methylation studies have revealed the possible association between aberrant DNA methylation and cancer. Overall genomic hypomethylation has been historically reported in several cancers, promoting chromosomal instability and oncogenes upregulation (Ehrlich, 2009; Sharma et al., 2010). Additionally, hypermethylation has been observed to contributes to oncogenesis by silencing tumour suppressor genes (Ehrlich, 2009; Sharma et al., 2010).

Epigenetic modifications have also been related to neurological diseases such as schizophrenia, Alzheimer's Disease and multiple sclerosis among others. For instance, genomic DNA of monozygotic twins discordant for schizophrenia has been analysed, showing differences in DNA methylation patterns of specific genes (Tsujita et al., 1998). This finding suggested that DNA methylation patterns could be associated with disease-specific gene expression changes in the *glutamic acid decarboxylase (GAD67)* and *reelin* (*RELN)* genes. And it turned out that, compared to normal neurons, the decrease of *GAD67 and RELN* expression in neurons of patients with schizophrenia was directly correlated with the DNMT1-mediated hypermethylation on CpG sites within these genes' promoters (Noh et al., 2005). In AD disease, initial genome-wide analysis of DNA methylation levels in post-mortem brain tissues or in lymphocytes did not show differences between patients with AD and controls. However, DNA methylation patterns in specific AD-associated genes differed when comparing patients with controls (Wang et al., 2008). Specific examples are the differential methylation in the CRTC1 gene in the hippocampus of AD patients and methylation-dependent (hydroxymethylcytosine) TREM2 upregulation (Celarain et al., 2016; Mendioroz et al.,

2016) The exact contribution of such differences to AD risk or pathology is presently unknown.

Similar to cancer and AD, searches have been conducted to identify epigenetic alterations relevant for the etiology or progress of MS. In a search of MS-biomarkers, genomic DNA of CD4+ lymphocytes from monozygotic twins discordant for MS was analysed in order to test if any discordance could be attributed to methylation changes in MS susceptibility genes . However, no evidence was found for genetic, epigenetic or transcriptome differences that explained disease discordance. These results contrasted with similar analysis performed for other neurological diseases as mentioned above . In another contemporary study, the methylation patterns of 56 genes, previously shown to be differentially methylated in ovarian cancer tissue (Melnikov et al., 2009), were compared using a microarray-based technique in plasma from MS patients and controls (Liggett et al., 2010). 15 out of 56 genes showed significant differences in methylation patterns between MS patients and controls. Additional studies have been conducted on post-mortem brain tissue. Such studies have revealed epigenomic changes in pathology-free areas of MS–affected brains(Huynh et al., 2014). Specifically, genes affecting oligodendrocyte susceptibility to damage are affected. As an example, the *BCL2L2* and *NDRG1 g*enes were hypermethylated and expressed at lower levels in pathology-free brain tissues from MS patients than in controls. As the products of these genes are involved in oligodendrocyte survival, these results suggest that oligodendrocytes in MS patients might be short lived in MS. Additionally, these results suggested that pathology-free areas in a MS brain might harbour epigenomic changes that might modulate the ability to respond to the pathological process. Other studies have been conducted in peripheral cells of MS patients. The promoter of the protein tyrosine phosphatase *SHP1*, a negative regulator of pro-inflammatory signalling and autoimmune disease, is inappropriately hypermethylated in peripheral blood leukocytes of MS subjects (Kumagai et al., 2012). As a result, SHP1 expression levels and activity are reduced, which is associated with increased leukocyte-mediated inflammation in MS (Kumagai et al., 2012). A genome-wide DNA methylation analysis of CD4+ T cells from MS patients and controls revealed evidence for association of DNA methylation at the HLA-DRB1 locus with risk for MS (Graves et al., 2014). The selected

examples above, along with many others, highlight the presence of DNA methylation alterations in multiple sclerosis (Miyazaki and Niino, 2015).

4. Transposable elements

Mammalian genomes have been usually thought to be stable and unchanging. However, it has become increasingly clear that the genome is not static but dynamic, and that its DNA content is subject to rearrangements, insertions and deletions (Prak and Kazazian, 2000; Smit, 1999). In fact, only half of the human genome is “unique”, while around 42% of the human genome consists of repetitive DNA, mostly in the form of transposable elements (TEs) (Lander et al., 2001). TEs are mobile DNA sequences that are dispersed throughout the genome. Barbara McClintock's discovery of these "jumping genes" earned her a Nobel Prize in 1983 (McClintock, 1950). These sequences can change its position within the genome, sometimes creating or reversing mutations, or resulting in duplication of DNA sequences (Prak and Kazazian, 2000; Smit, 1999). TEs therefore have the potential to drive genome evolution by causing rearrangement of genomes (Ayarpadikannan and Kim, 2014). TEs are increasingly recognized as a potent source of regulatory sequences in eukaryotic genomes (Chuong et al., 2017). Expression of a gene might be disrupted by insertion of a TE by providing novel or stronger promoter sequences. Alternatively, enhancer functions carried within a TE might activate the expression of more distant genes upon transposition (Girard and Freeling, 1999).

Transposable elements can be divided into two broad classes: DNA transposons and retrotransposons, according to their mechanism of action (Figure 11) (Kazazian, 2004; Prak and Kazazian, 2000). DNA transposons encode the protein transposase, which they require for excision and insertion, cutting and pasting themselves into new genomic locations. In contrast, the transposition of retrotransposons occurs by a replicative mechanism. DNA is first transcribed into RNA, which is subsequently reverse transcribed to DNA by a retro-element-encoded reverse transcriptase. This copied DNA is then inserted back into the genome at a new position while the original retrotransposon is maintained *in situ*. Retrotransposon are divided into two classes based on whether they contain a long terminal repeats (LTRs) at both ends (LTR

retrotransposons) or lack LTRs (non-LTR retrotransposons). The main function of such LTRs is to mediate the integration of the retroviral DNA via an LTR-specific integrase. In addition, they function as transcriptional promoters for both retroviral proteins and surrounding genes after genomic insertion. Non-LTR retrotransposons, in turn, consist of two sub-types: long interspersed elements (LINEs) and short interspersed elements (SINEs). The main different in their mechanism of action is that while LINEs encode for a reverse transcriptase and are transcribed by RNA polymerase II, SINEs do not encode reverse transcriptase and are transcribed by RNA polymerase III.

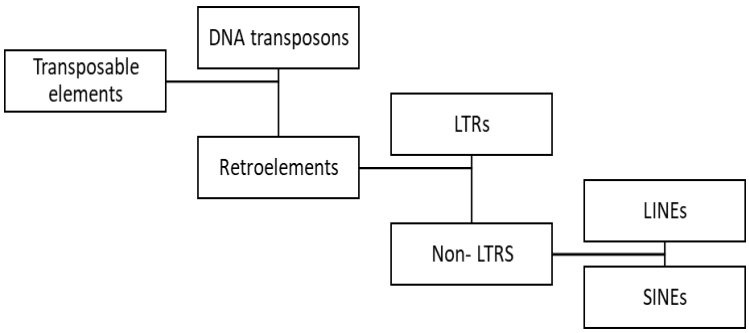


Figure 11. Classification of transposable elements. LTR: long terminal repeats; LINEs: long interspersed elements; SINEs: short interspersed elements.

4.1. Viruses: retroviruses

A virus is the smallest type of infectious particles. A virus cannot reproduce by itself, but once it infects a susceptible cell (a host), however, it can direct the cellular machinery to produce more viral particles. The entire infectious virus particle, called a virion, consists of the nucleic acid (RNA or DNA) enclosed in a protein coat (Lodish et al., 2000b). Animal viruses are classified in 6 groups based on the composition of their genomes (DNA or RNA) and pathway of mRNA formation (Figure 12).

Class VI corresponds to retroviruses. A retroviral particle contains two copies of single-stranded positive-sense molecules of RNA within an internal protein core named capsid (Figure 13). Within this compartment, there are also several viral proteins such as a reserve transcriptase, capsid proteins, an integrase, and a protease. The inner core is composed of capsid protein, is surrounded by an outer layer composed of matrix protein(s), which is in turn surrounded by the envelope glycoprotein.

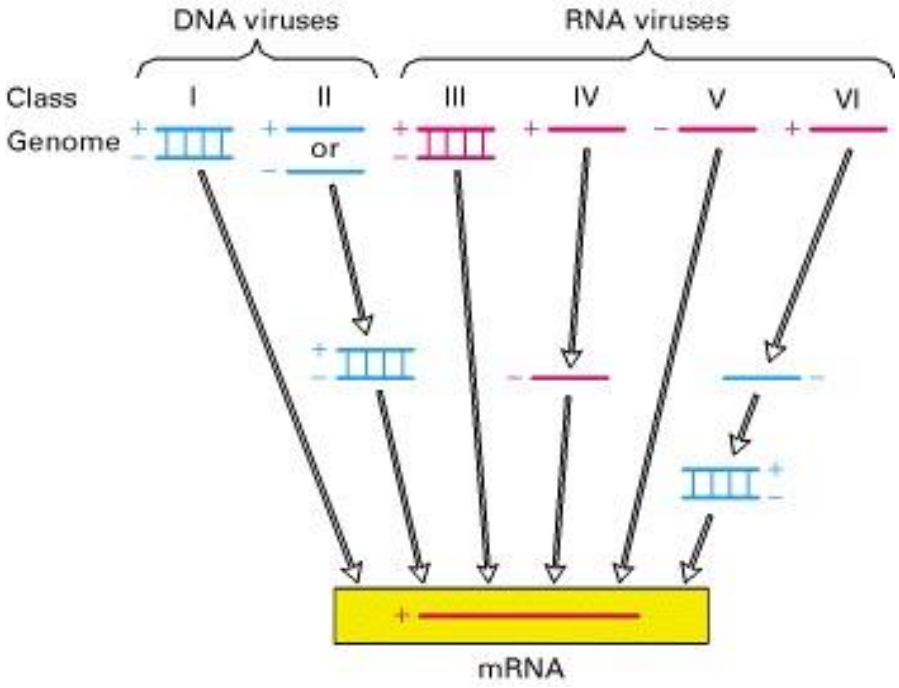


Figure 12. Classification of animal viruses. Viruses are divided into 6 classes (I-VI) depending on the composition of their genomes and pathway of mRNA formation. Genome can consist of a single- or double-stranded nucleic acid of DNA (shown in blue) or RNA (showed in red). The viral mRNA (plus strand) can, in turn, can derived from a minus strand of DNA or RNA. Image taken from Lodish et al., 2000b.

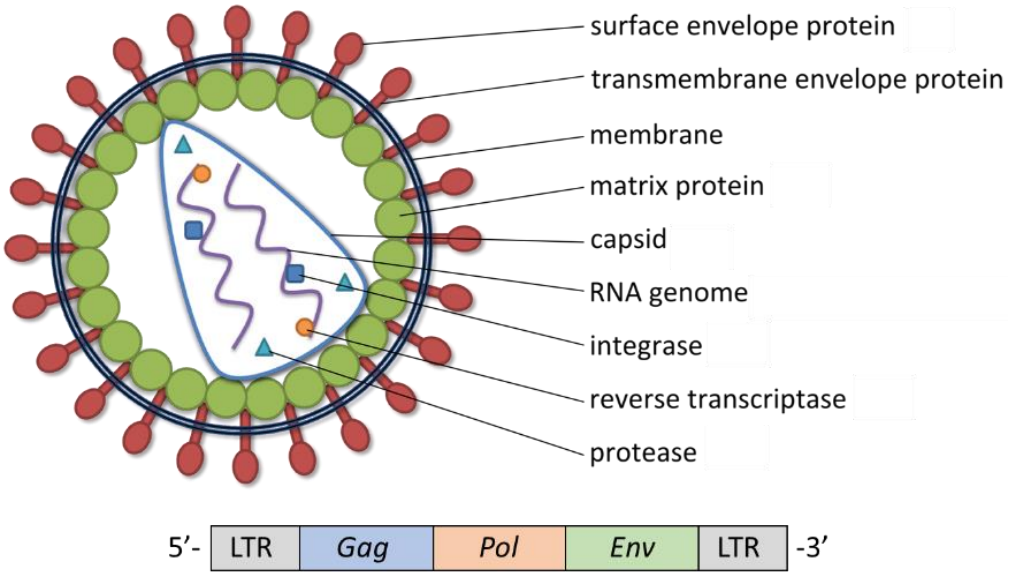


Figure 13. Retrovirus virion and provirus structure. The viral particle contains two copies of positive-stranded RNA within an internal core named capsid. Also located within the capsid are several viral enzymes such as integrase, protease and reverse transcriptase enzymes. The inner core is surrounded by an outer matrix protein core, which is in turn encompassed by a glycoprotein envelope. The viral genome, termed provirus once integrated into the host genome, consists of an internal region containing *GAG, POL*, and *ENV* genes, flanked by LTR. Picture adapted from O'Keefe,2018.

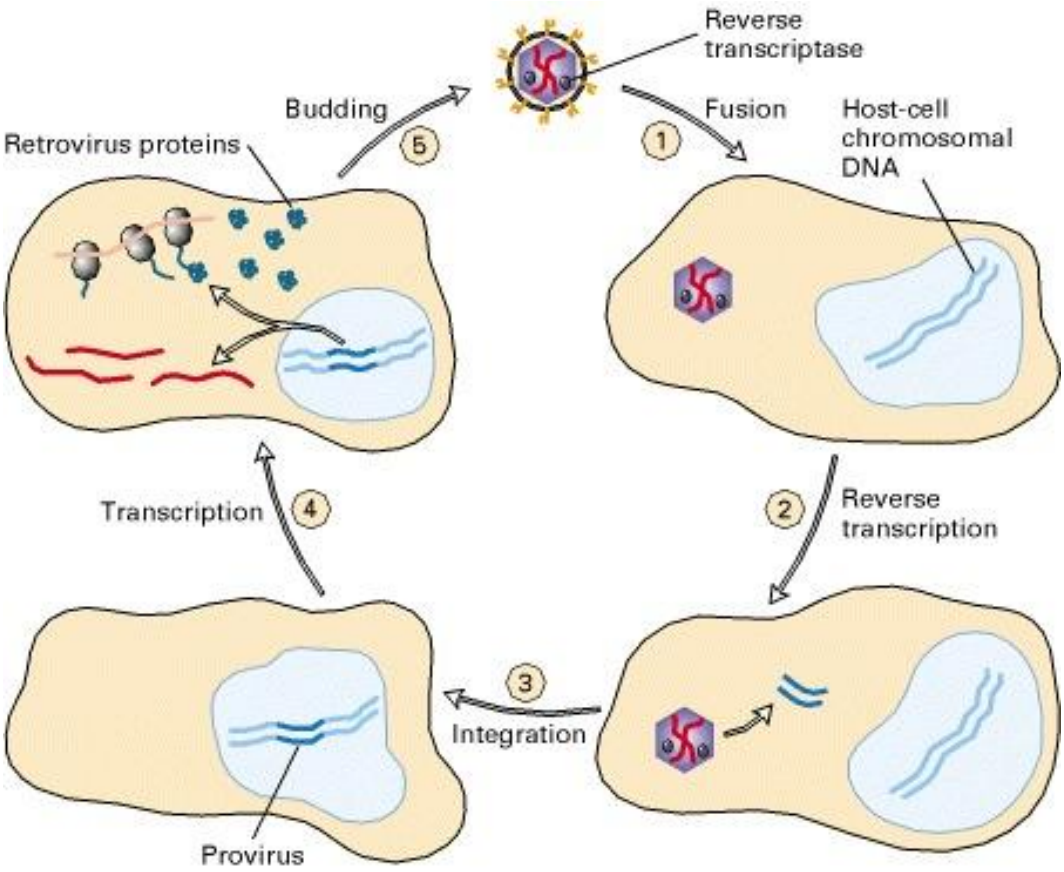


Figure 14. Retroviral life cycle. After a retrovirus enters a host cell by fusion (step 1), reverse transcriptase converts the retroviral RNA genome into double-stranded DNA (step 2). This viral DNA then migrates to the nucleus and becomes integrated into the host genome (step 3). Viral genes are transcribed and translated (step 4). New virus particles assemble, exit the cell, and can infect another cell (step 5). Image taken from Lodish et al., 2000b.

Retroviral infection is initiated when viral glycoproteins embedded in the lipid envelope recognize receptors displayed on the host cell plasma membrane and mediate viral attachment (Figure 14). Subsequent membrane fusion between the viral and host cell membrane follows and allows viral entry. Once inside the host cell cytoplasm, the virus is uncoated and replicates through the process of reverse transcription. First, the virus encodes for reverse transcriptase enzyme, which catalyzes the conversion of its RNA genome into DNA. This new retroviral double-stranded DNA is then integrated into the host cell genome in a process catalyzed by the virus-encoded integrase. Once integrated, the retroviral genome is named named provirus. The provirus consists of an internal region containing *GAG, POL,* and *ENV* genes, flanked by LTRs (Figure 13). *GAG* encodes for the matrix and capsid proteins, *POL* encodes for proteins such as reverse transcriptase, protease and integrase

enzymes, and *ENV* encodes for the envelope proteins including surface envelope glycoproteins and transmembrane glycoprotein. Within the proviral 5'-LTR there is a promoter, which is recognized by the cellular RNA polymerase, and enhancer regions, which are important in the initiation of transcription. As any part of its genome, the host cell starts transcribing the viral genes, producing viral RNAs and translating these into the proteins required to assemble new copies of the virus. Progeny virions are subsequently released by budding from the host-cell membrane. Most retroviruses do not kill their host cells. Infected cells can, therefore, replicate and produce daughter cells with integrated proviral DNA and new progeny virions.

4.2. Human endogenous retrovirus

Human endogenous retroviruses (HERVs) are the result of ancestral retroviral infections that entered the human germline. The resulting permanent retroviral integrations termed as provirus are transmitted to the offspring in a Mendelian fashion (Figure 15). HERVs are part of the LTR retrotransposon group. In fact, the typical structure of a proviral HERV contains an internal region containing *GAG*, *POL*, and *ENV* genes, flanked by two LTRs (Figure 16). The structure and nature of the HERVs indicates that following the initial integration, HERV elements have been subjected to amplification and transposition events, leading to a widespread distribution of complete or partial retroviral sequences throughout the human genome. As a result, nowadays HERVs comprise approximately 8% of our genome (Cho et al., 2008; Löwer et al., 1996). However, since their conversion to Mendelian genetic elements, HERVs have also accumulated mutations and have been subjected to recombination events. Therefore, most (if not all) HERVs are unable to replicate due to deletions and mutations acquired during evolution.

Apart from HERVs with the typical provirus structure, the human genome is littered with HERV-related sequences: pseudoelements, solo LTR and truncated elements (Figure 16). Pseudoelements are the result of LINE-mediated transpositions of the HERVs mRNA. They are characterised by the presence of a poly(A) tail and the partial truncation at the 5' end (Pavlicek et al., 2002). The presence of solo LTRs is the result of homologous recombination events occurring between the two LTRs of an element,

which results in the excision of the internal open reading frame (ORF) (Pavlicek et al., 2002). Truncated elements refer to HERVs with a truncated internal sequence or truncated LTRs. Since the U3 region of the LTR typically contains the binding motifs necessary for transcriptional initiation, only proviral elements were initially considered transcriptionally active. However, all kinds of HERV sequences have been found to be transcribed in human cells and tissues (Li et al., 2011), suggesting that transcription may be driven by control elements outside the HERV sequence. HERV elements are divided in around 30 multicopy families, each derived from an independent infection by an exogenous virus (Belshaw et al., 2005). Each family, in turn, is composed of closely related but distinguishable elements dispersed in the human genome.

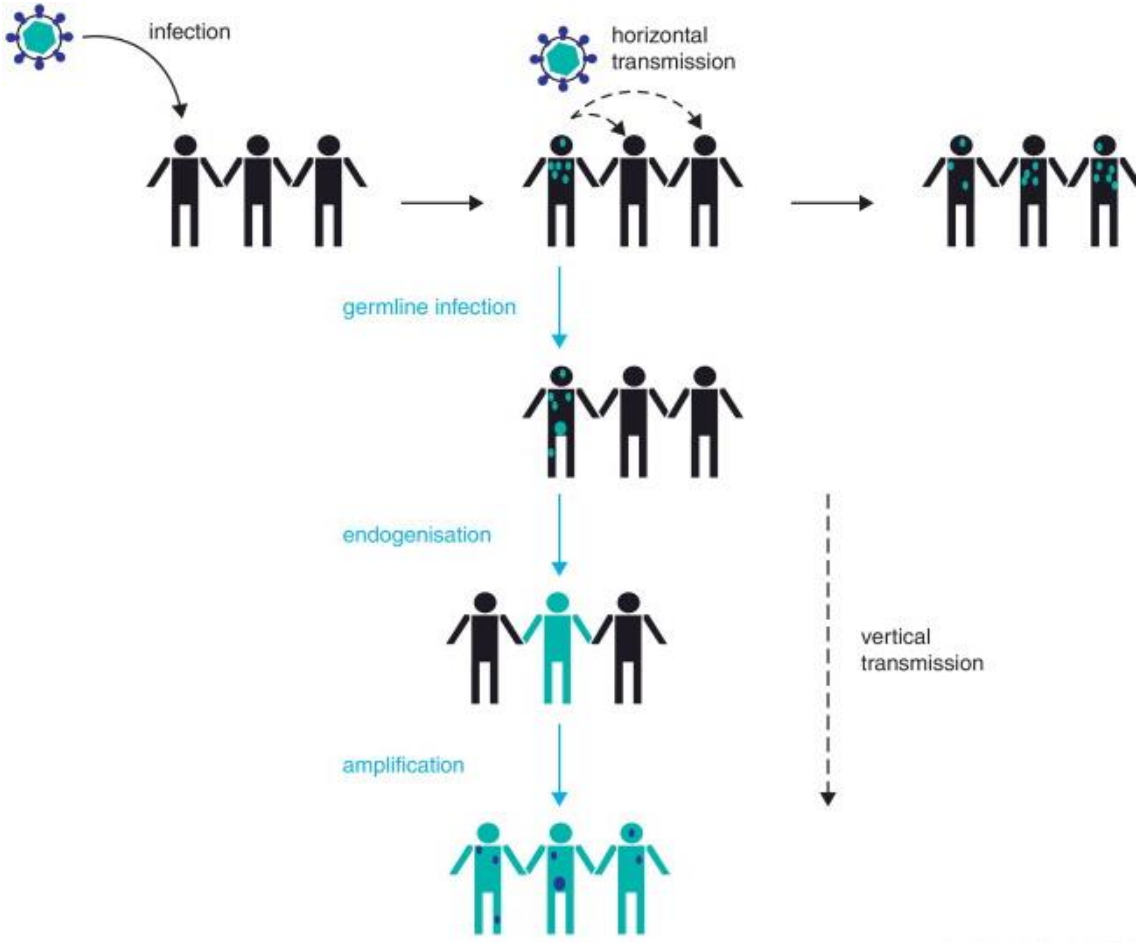


Figure 15. The endogenization process of retroviruses. Transmission of retroviruses can occur through two pathways: horizontal and vertical transmission. In horizontal transmission, retroviruses are transmitted among individuals of the same generation. However, when an exogenous retrovirus infects a germ line cell, the integrated viral DNA become part of the host's genome and is transmitted vertically to the offspring like a human endogenous retrovirus. Imagen taken from Dewannieux and Heidmann, 2013.

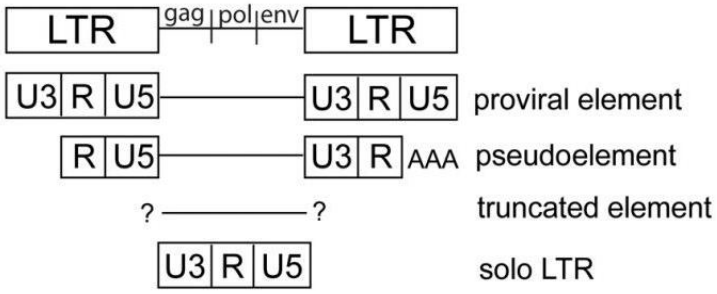


Figure 16. Schematic structure of HERV elements. HERV elements can show a complete proviral structure (an internal open reading frame flanked by two long terminal repeats (LTRs), or other incomplete structures such as pseudoelement, truncated element or solo LTR. A complete LTR is segmented into two unique regions (U3 and U5), and a repeated (R) region in between them. LTR contains a promoter in the U3 region, and a polyadenylation sequence in U5. Imagen taken from Li et al., 2011.

Most HERVs are highly defective, unable to replicate and unable to produce infectious exogenous viruses due to deletions and mutations acquired during evolution. In fact, until recently, many scientists considered HERV a type of "junk" DNA since they apparently served no function but rather were remnants of past infections. However, while most repetitive DNA and HERV sequences may be devoid of biological relevance, particular HERVs have been observed to play biological roles. Although very few of the identified HERV *loci* contain proviruses with preserved open reading frames (ORFs) for the three major structural proteins GAG, POL, and ENV, some HERVs have retained ORFs putatively encoding functional proteins (Bénit et al., 2001; Villesen et al., 2004). HERVs have been principally reported to be implicated in stem cell identity and embryonic development (Grow et al., 2015; Schoorlemmer et al., 2014). The best illustration of the biological role of HERVs is the contribution of *Syncytin* to placental physiology. It turns out that a HERV copy of the family W located on chromosome 7q encodes for a competent HERV-W envelope protein named SYNCITIN-1, which is expressed in placenta and involved in placental development (Blond et al., 2000; Mi et al., 2000). However, although a small number of HERV-encoded proteins can be found under normal physiological conditions, expression of HERVs is usually silenced in adult tissues under normal physiological conditions (Dewannieux and Heidmann, 2013).

4.3. Regulation and dysregulation of retrovirus

Several steps in the viral life cycle are blocked by a variety of silencing mechanisms to inhibit replication of retro-elements including HERV. Anti-retroviral defence mechanisms have been described that interfere with receptor binding and uncoating, trafficking of retroviral particles or components and viral DNA replication (Goff, 2004a, 2004b; Wolf and Goff, 2008). For instance, APOBEC (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like) proteins is a family of cytidine deaminases which deaminates cytosines to uracils during viral reverse transcription, contributing to hypermutation of viral synthesised DNA (Harris and Liddament, 2004). Retroviral particles and ribonucleotides are also detected by the innate immune system (Barbalat et al., 2011; Pertel et al., 2011). Silencing mechanisms are active and inhibit transcription and/or expression, and include DNA methylation, RNA interference (RNAi), repressive histone modifications, and ATP-dependent chromatin remodelling (Maksakova et al., 2008). HERVs expression is mainly repressed by epigenetic mechanism as histone modifications and DNA methylation (Maksakova et al., 2008). In fact, HERVs are heavily methylated in normal tissues(Hurst and Magiorkinis, 2017; Rowe and Trono, 2011). And for those HERV elements that are not fully methylated, as exemplified by the family of HERV-K(HML-2) elements, 5'LTR methylation levels inversely correlated with transcriptional activity of these elements (Lavie et al., 2005). Altogether, the combined activity of restriction mechanisms and especially DNA methylation restricts HERV expression to almost undetectable levels in physiological circumstances.

HERVs have been suspected to be associated with several diseases such as cancer and neurodegenerative diseases (Colmegna and Garry, 2006; Dolei, 2006; Kurth and Bannert, 2010). However, it is not very well-known which mechanism induce upregulation of HERV transcripts and proteins in patients. Although HERVs are usually silenced by epigenetic mechanisms, silent HERVs can be activated by environmental triggers. In fact, it has long been known that CpG DNA demethylation can reverse the silencing of HERV (Stoye, 2012). Epigenetic changes observed in pathological conditions could explain HERV activation, and their direct or indirect contribution to diseases. For instance, genomic hypomethylation of HERV elements has been observed

in ovarian tissue of patients with ovarian clear cell carcinoma, and in T cells from systemic lupus erythematosus patients (Iramaneerat et al., 2011; Nakkuntod et al., 2013). HERVs are induced in some tissues and at different development stages, which means that they are capable of responding to different external and internal signals (Figure 17).

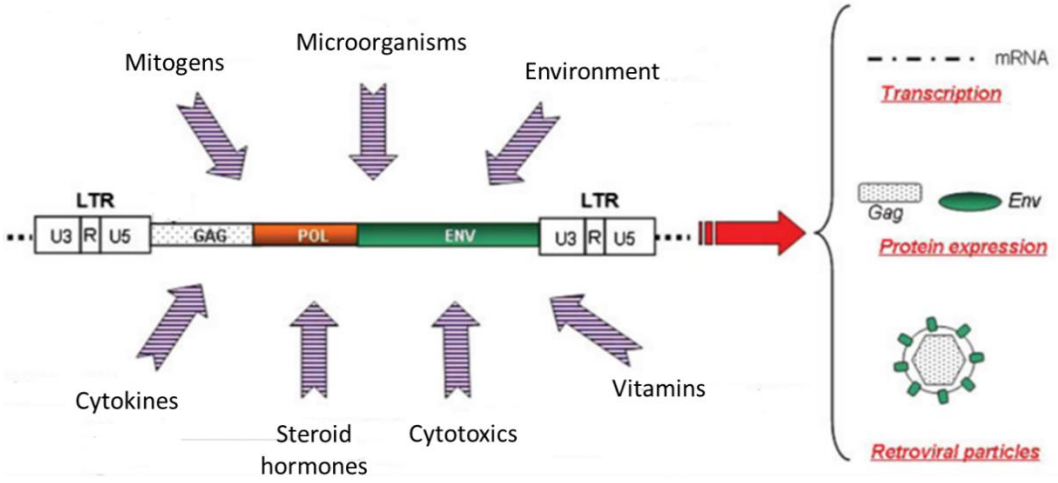


Figure 17. Schematic view of the different inducing factor that may act on HERVs. Several inducing factor may induce HERV mRNA synthesis, protein expression and even the formation of retroviral particles. Image adapted from Balada et al., 2009.

Environmental factors and different molecules have been suggested to induce HERV expression (Balada et al., 2009). Environmental stimuli such as radiation, chemicals or infectious agents have been reported to modulate HERV expression. For instance, UVB irradiation of human epidermal keratinocytes and infection of the neuroepithelial cell line SK-N-MC with *Toxoplasma gondii* resulted in an increased transcriptional activity of HERV elements (Frank et al., 2006; Hohenadl et al., 1999). Similarly, infection of the SK-N-MC cell line with of herpes simplex virus 1 (HSV-1) or influenza viruses led to increased levels of transcripts from HERV-W elements (Nellåker et al., 2006). And HIV-1 infected cell lines presented an up-regulated HERV expression compared to the corresponding uninfected cells (Vincendeau et al., 2015). All the above-mentioned findings

4.4. Association of human endogenous retrovirus with diseases

Although expression of HERVs is usually silenced in adult tissues under normal physiological conditions, contributions of HERV to pathologic processes have been

described, especially with respect to autoimmune diseases and cancer (Young et al., 2013). There are several possible mechanisms that could explain the pathogenicity of HERVs. The most obvious pathogenic mechanism that can be considered for a HERV is the production of infectious virions. In theory, replication-competent retroviruses could arise from a series of recombination events between two or more replication-defective HERVs, which would restore their respective defects. Although *in vitro* alignment of several defective members of the family HERV-K resulted in a replication-competent recombinant retrovirus (Lee and Bieniasz, 2007), the frequency of this *in vivo* event was very low and no infectious HERV has been yet identified. We therefore assume, that the association of HERV with disease does not depend on retroviral infection, although particles may be present.

Even if not transcribed or translated, HERV integrations may impact genome regulation and stability (Jern and Coffin, 2008; Young et al., 2013). Depending on their position and transcriptional orientation in relation to a host gene, HERV LTRs can act as transcriptional promoters or repressors of neighbouring genes. They could also disrupt genes at their integration site or alter ORFs to generate truncated proteins. HERVs could also promote genomic rearrangements and genome instability by providing regions of homology for DNA recombination. Non-allelic homologous recombination may cause deletions and duplications, resulting in genomic disorders and diseases (Campbell et al., 2014). For instance, apparent homologous recombination between HERV15 elements results in deletion of azoospermia factor a (AZFa) within the human chromosome Y, and leads to male infertility (Sun et al., 2000).

As HERVs exist in the genome of all individuals, how HERVs may cause disease in some individuals and not in others is not obvious. A plausible explanation for this could lie in the existence of genetic polymorphisms in HERVs. For instance, a specific allelic variant or single nucleotide polymorphism (SNP) could be related to disease, an example of which is the suggested association of HERV-K18 polymorphisms with type 1 diabetes (Marguerat et al., 2004), and the association of HERV-W polymorphism in chromosome X is associated with MS risk (García-Montojo et al., 2014). On the other hand, insertional polymorphisms may also exist and be related to specific diseases. HERV-K113 and HERV-K115 are insertional polymorphic proviruses present in the genomes of

29 and 16% of individuals respectively and, HERV-K113 prevalence has been associated with MS and Sjögren's syndrome (Moyes et al., 2007).

HERVs may cause disease by inducing or altering immune responses. Although most HERVs are replication-defective, many still retain the potential to generate transcripts and express retroviral proteins. Upregulated HERV transcripts or proteins synthesis would therefore be detected as foreign by the immune system and trigger an immune response (Fierabracci, 2009). In fact, elevated expression of several HERV transcripts and proteins has been reported to be associated with disease. In response to viral infection, antiviral innate immunity is activated, mediated by intracellular toll-like receptors (TLRs) that recognize viral DNA, RNA or protein as pathogen-associated molecular patterns (PAMP). TLR activation in turn triggers signalling pathways that induce transcriptional activation of genes encoding proinflammatory cytokines and type I interferons (IFN-α/β). This response could subsequently trigger anti-viral innate immune responses and modulate adaptive immunity (Xagorari and Chlichlia, 2008). For example, HERV-encoded ENV protein has been reported to interact with TLR4 in human monocyte, resulting in the production of proinflammatory cytokines (Rolland et al., 2006). And dsRNA overexpressed from HERV *loci* has been shown to trigger a TLR3 mediated-immune response in cancer cells (Chiappinelli et al., 2015).

4.5. Association of human endogenous retrovirus family W with multiple sclerosis

An association between HERVs and multiple sclerosis (MS) was suggested when in the late 1980's extracellular retroviral-like particles with reverse transcriptase (RT) activity were discovered and isolated from MS patient-derived cell cultures (Perron et al., 1991, 1989). Molecular characterization of the RNA encapsidated in these particles, allowed for the identification of a novel retrovirus named multiple sclerosis-associated retrovirus (MSRV) (Perron et al., 1997). MSRV presented a strong homology with the family W of human endogenous retroviruses (Blond et al., 1999; Komurian-Pradel et al., 1999; Perron et al., 2000). Both MSRV and HERV-W have the structure typical of retrovirus (Figure 16).

The HERV-W family is a multicopy family consisting of approximately 650 closely related but distinguishable elements dispersed throughout the human genome. The

HERV-W family is composed by complete and truncated HERV-W elements (Figure 16). Whether MSRV represents an exogenous member of the HERV-W family or whether it originates from modified endogenous HERV-W copies is presently unknown (Komurian-Pradel et al., 1999; Laufer et al., 2009; Perron et al., 2000). Continuing uncertainty about the origin of MSRV and its relation to HERV-W led to the MS-associated virus being relabelled as MSRV/HERV-W in several studies.

The linkage between the presence or expression of MSRV/HERV-W and MS has been repeatedly underlined (Hon et al., 2013; Morandi et al., 2017b, 2017a). On one hand, the number of MSRV virions detected in serum, blood and cerebrospinal fluid (CSF) have been reported to be higher in MS patients (Dolei et al., 2002; Garson et al., 1998). On the other hand, in a large study of over a 200 samples, MSRV-like DNA copy number was also increased in peripherical blood mononuclear cells (PBMCs) of MS patients (Garcia-Montojo et al., 2013). In other studies, MSRV/HERV-W mRNA was detected at higher levels in plasma, CSF, PBMCs and autopsied brain from MS patients (Mameli et al., 2007; Perron et al., 2012). MSRV*ENV* and *SYNCYTIN-1* transcripts are closely related and difficult to distinguish from each other. In fact, they have 94% identity and only differ by an insertion of 12 nucleotides in the MSRV *ENV* RNA that is lacking in *SYNCYTIN-1*, in a region corresponding to the trans-membrane domain of MSRV ENV protein. The lack of discriminating tools has generated confusing reports in the literature for years, during which *SYNCYTIN-1* was thought to be involved in MS pathology as well. Confusion was exacerbated by the fact that antiSYNCYTIN antibodies cross-react with all HERV-W-encoded ENV protein. As a result, the ENV protein detected in autopsied normal and MS human brain tissue (Antony et al., 2006, 2004) was described as SYNCYTIN. However, in 2009, the 12 bp insert that distinguishes MSRV *ENV* and *SYNCYTIN-1* transcripts was selected for the development of a discriminatory real-time PCR assay (Mameli et al., 2009). Both transcripts were expressed in autopsied human brain and PBMCs, but only MSRV *ENV* sequences were found more frequently in MS patients compared to controls (Mameli et al., 2009). Based on these results, the hypothesis that associated *SYNCYTIN-1* and MS was finally rejected.

Many searches aimed at the identification of transcribed HERV-W *loci* in MS patients have been performed, but few clear results have been achieved accomplished. Seven HERV-W *loci* are reportedly transcribed in human PBMCs of MS patients and controls (Laufer et al., 2009). However, transcriptional activity of individual identified HERV-W elements did not significantly differ between patients and controls. A similar analysis was performed in autopsied brain tissue (Schmitt et al., 2013). Likewise, among more than 100 transcribed HERV-W *loci* identified, none of them was found to be upregulated in brains of MS patients. To date, neither MSRV as a unique HERV, nor novel copies of HERV-W whose is associated with MS have been identified. Therefore, the origin of MSRV and its location within the human genome still remain unknown.

At the protein level, HERV-W ENV and GAG proteins have been detected in normal and MS autopsied human brain tissue, but reports varied. Some studies reported HERV-W proteins in brain tissue from MS patients but also in controls (Perron et al., 2005), while others only detected HERV-W proteins in MS brain and serum (Mameli et al., 2007; Perron et al., 2012). As no antibody specific for a unique HERV-W had been identified, the choice of antibody along with the technique applied and the patient cohort may influence in the results. HERV-W ENV proteins have also been reported to be expressed in higher quantities on the surface of B cells and monocytes in patients with active MS (Brudek et al., 2009).

Once overexpressed, the is no doubt regarding the pathogenic potential of MSRV/HERV-W proteins. *In vitro* infection of human peripheral blood lymphocytes from healthy donors with MSRV particles induced a T-lymphocyte-mediated immune response (Perron et al., 2001). On the other hand, MSRV-type envelope protein has the potential to trigger a TLR4-mediated immune response in human monocytes, with the consecutive cytokine production (Rolland et al., 2006). Consistent with the latter finding, immunohistochemical analysis of brain and spinal cord from both control and multiple sclerosis patients, revealed enhanced expression of TLR3 and TLR4 in inflamed CNS tissues (Bsibsi et al., 2002). More recently, stimulation of human cultured primary oligodendrocyte precursor cells (OPCs) with MSRV-related ENV protein provoked a TLR4-mediated induction of proinflammatory cytokines and an enhanced nitric oxide (NO) production (Kremer et al., 2013). Altogether, these results indicate the ability of

MSRV/HERV-W proteins may have to trigger a TLR-mediate immune response and contribute to diseases.

Following the repeated association of MSRV ENV protein with MS, a selective humanized antibody targeting MSRV ENV protein was developed and named GNbAC1 (Curtin et al., 2014, 2012). GNbAC1 was obtained after immunizing mice with recombinant MSRV ENV protein, expressed from a cloned RT-PCR amplicon, which was obtained using purified extracellular MSRV virions. GNbAC1 selectively binds with high affinity to the SU domain of the MSRV ENV protein and neutralizes its TLR4 binding. GNbAC1, therefore, could be used as a treatment for MS and is currently being tested in clinical trials with promising results (Curtin et al., 2016; Derfuss et al., 2015; Zimmermann et al., 2015). Altogether, it remains entirely possible that dysregulation of the immune system triggered by MSRV-type transcripts and proteins contributes to MS pathogenesis.

III. OBJECTIVES

Since the first description and characterization of multiple sclerosis-associated retrovirus (MSRV) (Perron et al., 1997), the connection between MSRV and multiple sclerosis (MS) has been addressed in a variety of studies. Despite the strong homology between MSRV and the family W of human endogenous retroviruses (HERV-W), the exact relationship between MSRV and HERV-W remains unknown (Blond et al., 1999). Both the presence and expression of MSRV/HERV-W RNA and proteins is elevated in peripherical blood mononuclear cells (PBMCs) and cerebral lesions of MS patients (Garcia-Montojo et al., 2013; Mameli et al., 2009; Perron et al., 2012). However, neither the location of MSRV within the human genome, nor MS-associated HERV-W elements have been identified so far (Laufer et al., 2009; Perron et al., 2000; Schmitt et al., 2013). Regulation of retroviruses involves several mechanisms including epigenetic regulation. Nevertheless, while HERV hypomethylation is associated with a range of diseases such cancer (Gimenez et al., 2010; Iramaneerat et al., 2011), HERV-W methylation status in the context of MS has not yet been evaluated. Therefore, the possibility that hypomethylation-induced overexpression of HERV-W may serve as a diagnostic marker, or even contributes to disease, remains viable. In response to viral infection, toll-like receptors (TLRs) play an important role in the activation of antiviral innate immunity pathways. As retroviral and HERV-derived RNAs have the potential to trigger a TLR3-mediated immune response in cancer cells (Chiappinelli et al., 2015), I was interested in the possibility that TLR recognition of upregulated HERV-W RNA expression and subsequent immune responses (including interferon-beta production) play a role in MS.

Objectives:

1. Identify the location of MSRV and/or MS-associated HERV-W elements within the human genome.

2. Confirm the increased expression of MSRV/HERV-W in MS patients.

3. Identify which among the 600 described HERV-W elements are overexpressed in PBMCs of MS patients.

4. To determine whether DNA methylation may be involved in the regulation of HERV-W expression.

5. To evaluate the TLR-mediated activation of the immune system by HERV-W overexpression in oligodendrocytes or cell types relevant for MS.

IV. RESULTS/ MANUSCRIPTS

Manuscript I. Development of a splinkerette-PCR method for the genome-wide amplification of multiple sclerosis-associated HERV-W copies in humans.

Abstract

While the presence and expression of Human Endogenous Retroviruses family or type W (HERV-W) has been repeatedly associated with Multiple Sclerosis (MS), disease-specific copies have not been identified in the human genome. To date, no systematic searches for disease-specific proviruses of HERV-W have been reported. Therefore, the possibility that unfixed copies or rare polymorphic *loci* contribute to the disease has not been directly addressed. Here we describe a variant of so-called splinkerette amplification of proviral sequences, in which DNA fragments containing genomic sequences flanking the termini of HERV-W proviral sequences are selectively amplified by PCR. Subsequent sequencing of the products obtained allows genomic positioning. A total of 7 alternative primer sets were successfully applied in an attempt to identify potential unfixed HERV-W proviruses specific to MS patients. We surveyed the HERV-W proviral content of 4 unrelated humans, including several that had been previously shown to carry a high genomic load of HERV-W-like elements. In an initial analysis of about 100 cloned products, we failed to identify unfixed HERV-W proviruses specific to MS patients. The assay also revealed the presence of a non-disease related polymorphism in HERV-W Xq21 *loci*. Splinkerete-PCR products obtained from 2 MS patients were analysed by next-generation sequencing (NGS). The NGS data obtained were analysed using the software RetroSeq, which can accurately discover non-reference transposable element insertions. While several non-reference insertions were identified in this analysis, their relation to MS has not been established yet. The method we describe, in combination with NGS, allows for systematic searches for uncharacterized HERV-W proviral sequences and thereby provides a direct means to assess inter-individual genetic variation associated with HERV-W proviruses and their potential relation to MS.

Introduction

Human endogenous retroviruses (HERVs) are the result of ancestral retroviral infections that entered the human germ line, leading to a permanent retroviral integration transmitted to the offspring in a Mendelian fashion. After initial integration, HERVs continued to generate self-copies and as a result they comprise approximately 8% of our genome (Cho et al., 2008; Löwer et al., 1996). Typical structure of a HERV contains an internal region containing *GAG, POL,* and *ENV* genes, flanked by two long terminal repeats (LTR). Although most HERVs are highly defective and unable to replicate due to deletions and mutations acquired during evolution, some HERVs have retained open reading frames (ORFs) putatively encoding functional proteins (Bénit et al., 2001; Villesen et al., 2004). As an example, a HERV-W copy located on chromosome 7q is been reported to encode for a competent HERV-W envelope protein named SYNCYTIN-1, which is expressed in placenta and involved in placental development (Blond et al., 2000; Mi et al., 2000). HERVs have been stated as well to be associated to several diseases such as cancer and neurodegenerative diseases (Colmegna and Garry, 2006; Dolei, 2006; Kurth and Bannert, 2010).

In the late 1980's an association between HERVs and multiple sclerosis (MS) was suggested, as extracellular retroviral-like particles with reverse transcriptase (RT) activity were discovered and isolated from MS patient-derived cell cultures (Perron et al., 1989, 1991). Follow-up studies using PCR-based approaches yielded the molecular characterization of the RNA encapsidated in these particles, and a novel retrovirus named multiple sclerosis-associated retrovirus (MSRV) was reported (Perron et al., 1997). However, no replication-competent MSRV provirus encoding such particles could be identified (Perron et al., 2000). MSRV presents a strong sequence homology with the HERV-W family, although the exact relationship has not been clarified yet (Blond et al., 1999; Komurian-Pradel et al., 1999; Perron et al., 2000). The HERV-W family is a multicopy family consisting of approximately 650 closely related but distinguishable elements dispersed in the human genome. Whether MSRV represents an exogenous member of the HERV-W family or whether it originates from modified endogenous HERV-W copies remains unknown (Komurian-Pradel et al., 1999; Laufer et al., 2009; Perron et al., 2000).

The linkage between the presence or expression of MSRV/ HERV-W and MS has been repeatedly underlined (Hon et al., 2013; Morandi et al., 2017a, 2017b). Not only the number of MSRV virions detected in serum, blood and cerebrospinal fluid (CSF) have been reported to be higher in MS patients over controls (Dolei et al., 2002; Garson et al., 1998), but also MSRV/ HERV-W expression has been established to be higher in plasma, CSF, peripheral blood mononuclear cells (PBMCs) and autopsied brain from MS patients versus controls (Mameli et al., 2007; Perron et al., 2012). In addition, MSRV/ HERV-W proteins express in human brain and can trigger an immune response (Perron et al., 2001, 2005; Rolland et al., 2006), thus suggesting their potential pathogenicity. Nevertheless, the mechanism of HERV action in MS has not been conclusively demonstrated to date (Antony et al., 2011; Grandi and Tramontano, 2017).

The genomic distribution of HERV-W and the number of copies has been described (Grandi et al., 2016). Moreover, searches for transcribed HERV-W *loci* have clarified to which extent HERV-W *loci* described in the human genome database are active in MS patients (Laufer et al., 2009; Schmitt et al., 2013). Most HERVs copies are present in the human population and described in the human reference genome sequence because they integrated a long time ago (fixed copies). However, if some HERVs have continued replicating in the human population, it is possible that some copies have integrated in some individuals' genome but are absent from the reference genome, since they have not had time to drift to fixation (unfixed copies) (Marchi et al., 2014). This hypothesis is reinforced by the fact that novel retroviral insertions of HERV-K elements have been reported (Kahyo et al., 2013; Wildschutte et al., 2016). In addition, other studies have supported that retroelements preserve their mobile properties in spite of their evolutionary age (Kremer et al., 2013), as well as revealing that some HERVs actually occur polymorphically in the human genome (Moyes et al., 2007; Turner et al., 2001). Therefore, there is a possibility that MS might be associated only with specific unfixed HERV-W/MSRV *loci*. This issue has not been amply addressed, as few searches for additional genomic copies of HERV in MS patients have been reported (Nissen et al., 2012). In order to address this issue, further searches for the genomic location of MSRV or novel HERV-W copies are required. To the best of our knowledge,

we present the first systematic approach to discover potential MS-associated copies of HERV-W. We have successfully adapted a previously reported splinkerette-PCR method (Uren et al., 2009), initially designed for the isolation of retroviral insertion sites, to identify HERV-W-like sequences. This new method was validated and easily allowed for the identification of referenced HERV-W copies present in gDNA samples from MS patients. To increase the scope of the search, we carried out a next-generation sequencing (NGS) analysis that enlarged the number of HERV-W copies identified in gDNA samples from MS patients. With the aim of identifying unfixed HERV-W *loci*, we analysed the NGS data using the software RetroSeq, which can accurately discover non-reference transposable element (TE) insertions (Keane et al., 2013; Marchi et al., 2013). Several non-reference HERV-W copies could be identified using this approach. Further research is required to clarify the importance of these unfixed HERV-W copies to MS.

Materials and methods

PBMCs isolation and gDNA extraction

Blood samples were collected from MS patients and healthy controls from the Neurology Department of Miguel Servet University Hospital (Zaragoza, Spain). Clinical samples were obtained with "Informed Consent", in the context of the study "Reactivation of Human Endogenous Retrovirus in Multiple Sclerosis as a diagnostic tool: Epigenetic control and inducing factors". This study has been approved by the local Ethics committee (CEICA, PI14/0021) in 2014. Whole fresh blood was drawn into vacutainer tubes (Becton Dickinson Vacutainer) containing EDTA. Within the day, peripheral blood mononuclear cells (PBMC) were isolated using Histopaque-1077 (Sigma) as previously described (Böyum, 1968) (Figure s1). Cells were incubated with Genomic Lysis Buffer (10 mM Tris-Cl pH8, 25 mM EDTA, 0.5% SDS, 100 mM NaCl) and 1 U/ml Proteinase Q (QIAGEN) at 65 °C overnight. Genomic DNA was subsequently extracted using phenol-chloroform, precipitated using sodium acetate-ethanol precipitation, re-suspended in 100 µl of TE buffer (10 mM Tris-Cl pH 8, 1 mM EDTA pH 8) and quantified using Nanodrop (Thermo Scientific). Two additional gDNA samples from MS patients, obtained from the Multiple Sclerosis Unit of San Carlos Clinic Hospital (Madrid, Spain), were generous provided by Roberto Alvarez Lafuente. They

had been included in a previous study (Garcia-Montojo et al., 2013), and represent samples with high HERV-W copy numbers (Roberto Alvarez Lafuente, personal communication).

Library construction and amplification

A splinkerette-PCR method previously described (Uren et al., 2009) was adapted for HERV-W sequences and carried out as follows. Genomic DNA (2 µg) was digested with 40 U of a restriction enzyme (Table s1) in a 30-µl reaction containing 1 mM DTT. Digestion was incubated overnight at 37 °C, followed by a heat inactivation of the enzyme at 75 °C for 10 minutes. The digestion product was cleaned up using phenol-chloroform and precipitated with sodium acetate-ethanol according to standard protocols. The final product was re-suspended in 20 µl of TE buffer and 300 ng of digested DNA was ligated to the splinkerette adaptors (Table s2). The splinkerette adapter mix was prepared containing according to the instructions given (Uren et al., 2009). The ligation reaction was carried out in a 40-µl volume using 25 pmol splinkerette adaptor mix, 10 U of T4 DNA Ligase and 1X T4 DNA Ligase Buffer. Ligation was incubated at 4 °C overnight, followed by heat inactivation of the T4 DNA Ligase at 75 °C for 10 minutes. The 40-µl ligation was then digested at 37 °C overnight with 10U of a second restriction enzyme (Table s1), in the presence of 1 mM DTT. Following digestion, the T4 DNA ligase was heat inactivated at 75 °C for 10 minutes and the digested product was cleaned up using the Geneclean Turbo kit (MP Biomedicals), re-suspended in 20 µl of TE buffer and stored at -20 °C until used. For the following nested-PCR, primers required have been listed (Table s1) and cycling conditions were as described (Uren et al., 2009). Primary PCR amplification was carried out in a 50-µl volume using 5 µl of purified splinkerette adaptor ligated DNA, 0.1 µg splink1 primer, 0.1 µg LTR1 primer, 250 µM dNTPs, 1.5 mM MgCl2, 1.25 U Platinum® Taq DNA Polymerase and 1x PCR Buffer. For the nested PCR, we used 1 µl of the first PCR product as template and it was performed in a 50-µl volume containing 0.1 µg splink2 primer, 0.1 µg LTR2 primer, 250 µM dNTPs, 2 U Taq DNA Polymerase from Thermus aquaticus (Sigma-Aldrich) and 1x PCR reaction buffer containing MgCl2 (Sigma-Aldrich). The presence of amplified products was confirmed on a 2% agarose gel (Figure s2A). All enzymes were obtained from Thermo Scientific™ unless otherwise indicated.

Splinkerette-PCR product sequencing

Splinkerette-PCR product was purified using the GeneClean Turbo Kit (MP Biomedicals), eluted in 20 µl of TE buffer and cloned into the pGEM®-T Easy Vector System I (Promega) according to manufacturer's instructions. The vector was transformed into E. coli XL1 cells and cultured in Luria–Bertani (LB)-agar plates supplemented with 50 µg/mL ampicillin at 37 °C. Colonies were screened for the presence of insert by colony PCR and further analysed by plasmid miniprep and sequencing. DNA sequencing analyses were performed by Secugen S.L (Madrid, Spain) and the genomic locations of HERV-W insertions were mapped to the human genome (GRCh38) using the blastn option of BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) (Altschul et al., 1990) (Figure s2B). Species-specific filters for retroviral elements were not activated (or de-activated). To avoid mismapping as a result of the high degree of homology among HERV-W elements, a similarity of > 97% between the database sequence and the total length of a particular sequence was used as criterium to assign the sequences to a particular locus.

Polymorphism analysis

For the PCR we used as template 100 ng of gDNA, previously fragmented by sonication for 5 minutes at high power with *cycles* of *30 seconds* ON, and 30 seconds OFF. PCR primers were designed using Integrated DNA Technologies (IDT) PrimerQuest tool (https://eu.idtdna.com/PrimerQuest/Home/Index) (Forward primer: 5'-CACCATCTTGGGAGCTCTG-3', reverse primer: 5'-[Btn]CTAAACATCAACCTGAGAACTAAGC-3'). The PCR reaction was carried out in a 50-µl volume using 200 µM dNTPs, forward and reverse primers (400 nM each), 1 mM $MgCl_2$, 0.5 U Taq DNA Polymerase from *Thermus aquaticus* (Sigma-Aldrich) and 1x PCR reaction buffer containing $MgCl_2$ (Sigma-Aldrich). gDNA was boiled for 3 minutes and cycling conditions were as follows: 94 °C for 2 minutes and then 35 cycles of 94 °C for 1 minute, 60 °C for 1minute and 72 °C for 1 minute, with a final extension of 72 °C for 10 minutes. The presence of amplified product was confirmed on a 2% agarose gel. PCR product was sequenced using PSQ 96MA System (Pyrosequencing) according to manufacturer's instructions. The resulting data was analysed and quantified with the

PSQ 96MA 2.1 software (Biotage) applying the default settings. The forward PCR primer was used as sequencing primer, and the nucleotide dispensation order for the assay was as follows 5'- CTGCAGCAGTACTGCGCT-3'. Several negative controls were included in the sequence to assess the absence of undesired *PCR products.*

Next-generation sequencing analysis

NGS and subsequent analysis including RetroSeq was carried out by the Sequencing and Functional Genomics Service (Servicio Científico Técnico de Secuenciación y Genómica Funcional) of the Aragon Health Sciences Institute (IACS, Zaragoza). Methods for sequencing were recommended or approved for IonTorrent technology by the relevant manufacturers. PCR products were end repaired and purified with Agencourt AMPure XP paramagnetic beads (Beckman Coulter) and amplicons were quantified using the Qubit 3 fluorometer (Qubit BR dsDNA assay, Thermofisher). Adapters and barcodes (Ion Xpress Barcode Adapters, Thermofisher) were added by ligation, followed by product purification with AMPure XP beads. Libraries were prepared separately for each simple using the Kapa Lib Prep Kit for Ion Torrent (Roche). Libraries were quantified by qPCR (Ion Library TaqMan™ Quantitation Kit, Thermofisher) and equimolar amounts of sample were pooled before template preparation. Emulsion PCR, template enrichment and chip loading were carried out manually using the Ion 520&530 Kit-OT2 (Thermofisher) and samples were sequenced on the Ion Torrent S5XL platform using a Ion 530 chip.

The Fastq files produced by the sequencer (barcodes removed) were further trimmed to remove PCR primer sequences and sequence reads were mapped to the human reference genome (version hg19) using the Torrent MApping Program (TMAP v5.2.22). Mapping criteria were evaluated based on the error rate, and adjusted accordingly. Strict criteria were applied to maximize mapping differences between different copies (soft clipping off, mismatch penalty 100, open gap penalty 100, gap extension penalty 10, minimum alignment length 135, "unique best hit and stage1 map4"). For each sample, the number of reads corresponding to individual known HERVW copies was extracted, as well as the cumulative error rate.

Discovery of HERV-W elements not represented in the reference genome

Fastq files produced by IonTorrent sequencing (Figure 3A) were trimmed to remove PCR primer sequences. Reads were converted "in silico" to first double stranded DNA sequences, and subsequently into 50 bp long paired-end reads. The resulting read pairs were mapped to the human reference genome (version hg19) using BWA (version 0.7.16a-r1181). The resulting bed file was loaded into RetroSeq (Keane et al., 2013), a software that specifically designed screen NGS data to discover proviral insertions not present in the reference genome. HERV-W sequences present in this version of the genome database were downloaded from RepBase (Bao et al., 2015). RetroSeq analysis can be separated into 3 steps: reads pairs that map to different genomic locations are first identified, compared to locations of known HERVW elements (one of the paired reads is required to overlap), and scored based on the number of overlapping reads and the mapping scores. The result is a list of breakpoints, which indicates the positions in the reference genome where a not-listed HERVW insertion is present in the sample analysed.

Results

Strategy for identification of HERV-W-like sequences in human genome based on a variation of a splinkerette-PCR method

Although MSRV has been repeatedly associated with MS, systematic searches for its genomic location have not been reported. To develop assays allowing such systematic analysis, a previously reported splinkerette-PCR method to isolate retroviral insertion sites (Uren et al., 2009) *has been effectively adapted (Figure 1A). This highly efficient technique is a variant of ligation-mediated PCR, using a splinkerette 'hairpin loop' to avoid artefacts caused by 'end-repair priming' and improve specificity by requiring synthesis from a primer that anneals within the proviral sequences. Briefly, genomic DNA is first digested with a high-frequency-cleavage restriction enzyme, linked to compatible splinkerette adaptors and selectively digested with a low-frequency-cleavage restriction enzyme to remove one of the splinkerette adaptors. The final product is amplified in two rounds of PCR, and the resulting fragments are cloned and sequenced. Therefore, this technique allows the amplification, sequencing and location*

of the genome sequence that lies outside of a known retroviral sequence. Due to the strong homology shared between MSRV and HERV-W, we adapted this method to selectively amplify HERV-W-like sequences with the aim of identifying MSRV-like sequences or HERV-W copies not described in the human genome database*. Since MSRV is not describe in any human genome database to date, we consequently gathered the human genome sequences that did match 100% with a set of external primers that specifically amplify MSRV-type ENV* (Mameli et al., 2009)*, and we used the consensus sequence of these MSRV/HERV-W sequences as a template to adapt the method. Furthermore, as HERV-W elements contain different elements (complete provirus, pseudoelement, truncated copies and solo LTR)* (Costas, 2002; Li et al., 2011)*, different splinkerette-PCR assays were designed by selecting the appropriate restriction enzymes and primers in order to cover these different structures (Figure 1B). Design A aims to amplify the area downstream of the U5 region of the 3' LTR of a complete HERV-W element. Design B intends to amplify the area downstream of the U3 region of the 3' LTR of a HERV-W pseudoelement. Design C focuses on amplifying the area upstream of the U3 region of the 5' LTR of a complete HERV-W element. Design D amplifies from the ENV gene region towards 3' LTR. This last assay is quite relevant because it takes advantage of a set of primers used to detect higher MSRV-like sequences expression in PBMCs of MS patients over controls* (Mameli et al., 2009)*, pretending this way to enclose the pool of sequences where to look for MSRV-like sequences.* Assay designs A to D were carried out using genomic DNA from PBMCs of two MS patients. Among a total of 73 clones that were sequenced after cloning, 70 corresponded to sequences closely related to HERV-W (>97% identity between the cloned sequence and the database sequence was used as criterium) (Table 1A). The 3 remaining clones corresponded to PCR artefact (data not shown). Out of these 70 clones, 41 did not contain suficient sequence outside of the HERV-W-part to allow for the identification of their genomic position. In contrast, the 29 remaining clones could be unequivocally localized and corresponded to copies represented in the UCSC Genome database.

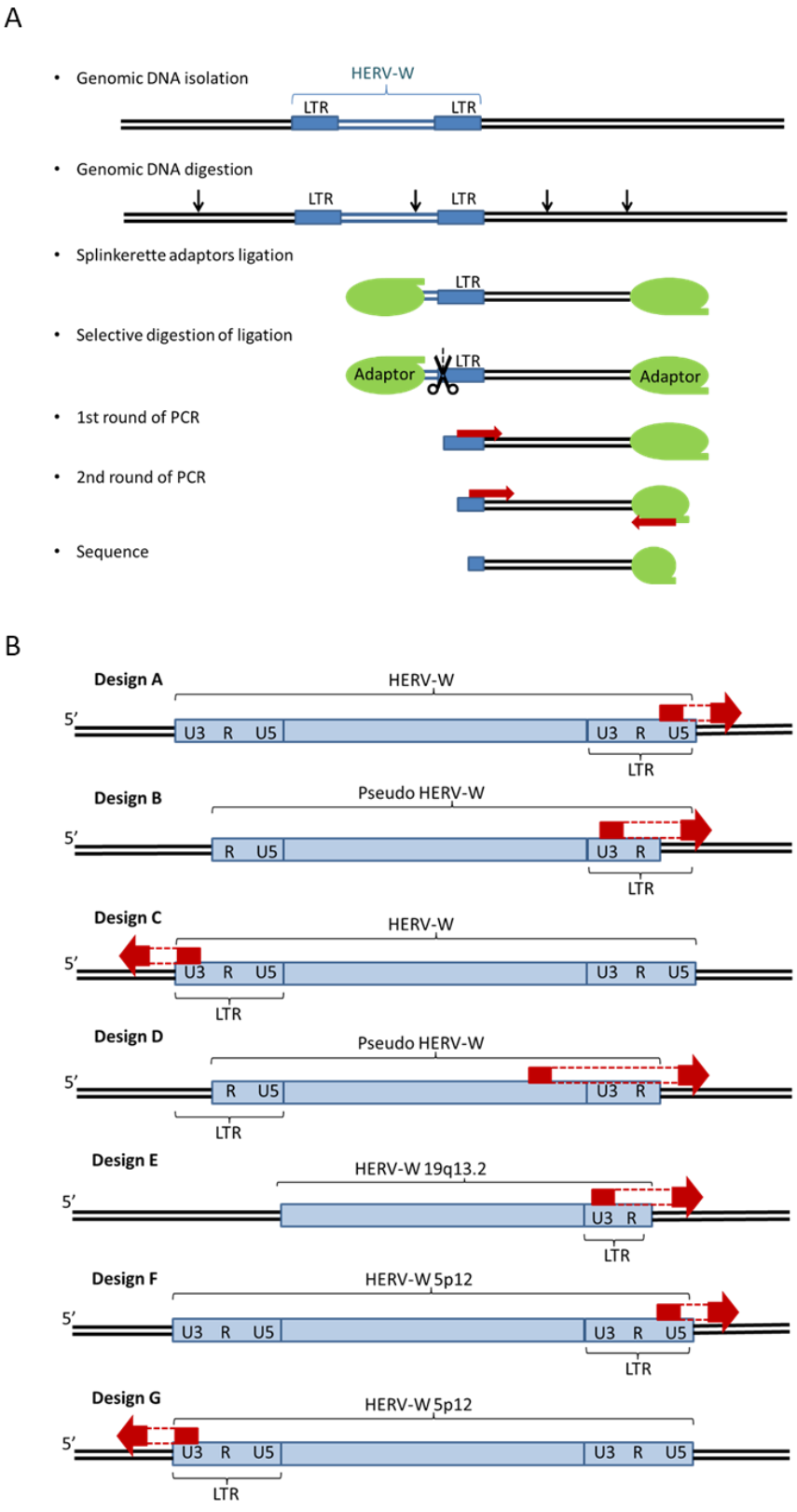


Figure 1. Overview of the splinkerette-PCR protocol used to determine the location of HERV-W/ MSRV copies. A) Steps performed to generate splinkerette-PCR products for sequencing. B) Different splinkerette-PCR designs carried out in order to cover most of the HERV-W copies described by Mameli G et al., 2009. The red arrows indicate the location of the primer in the 1st round of PCR. It is not drawn to scale.

A

Design	Clones analyzed	Sequences with high homology to HERV-W LTR	Located	New location identified
A	35	35	22	0
B	4	4	0	0
C	23	23	4	0
D	11	8	3	0
Total	73	70	29	0

B

Design	Clones analyzed	Sequences with high homology to HERV-W LTR	Located (copy expected)	Located (other copies)	New location identified
E	8	8	1	0	0
F	17	17	6	10	0
G	16	16	0	6	0
Total	41	41	7	16	0

C

Design F	Clones analyzed	Sequences with high homology to HERV-W LTR	Located (copy expected)	Located (other copies)	New location identified
Sample 1	11	11	2	6	0
Sample 2	11	11	0	8	1
Total	22	22	2	14	1

Table 1. Numbers represented above are derived from the sequencing of splinkerette-PCR products. Designs A-G are explained in Figure 1B. Clones were generated by cloning splinkerette-PCR products. Located: non-LTR sequence that allows localization in human genome (>97% homology with GRCh38). Copy expected: located in copies expected based on primer design. Other copies: located with >97% homology to HERV-W copies in the reference genome different from the copies expected. New location: LTR location not present in the reference genome (GRCh38).

In a separate study, we cloned and sequenced MSRV/HERV-W-*ENV*-specific PCR products, generated after reverse transcription of RNA of MS patients to identify *loci* transcribed in PBMCs (Manuscript II). 21 transcribed HERV-W *loci* were found within more than 100 sequenced clones, of which HERV-W copies on chromosome 19q13.2 or 5p12 were among the most abundantly transcribed. However, if unfixed MSRV/ HERV-W copies existed, this high level of transcription could mask the expression of non-reference HERV-W copies with high similarity to HERV-W 19q13.2 or 5p12, if such copies existed. To discriminate weather the high level of transcription really originated from the HERV-W 19q13.2 or 5p12 copies as opposed to unfixed copies with high similarity, we developed more splinkerette assays. These novels designs (E-G) aimed to

specifically amplify HERV-W *loci* 19q13.2 and 5p12 or close relatives (Figure 1B). Once validated, assays were carried out using genomic DNA from PBMCs of one MS patients, allowing us to efficiently locate both HERV-W copies and others HERV-W *loci*. However, all clones analysed could be unequivocally localized, indicating they are not unfixed HERV-W copies (Table 1B).

Additionally, we applied the splinkerette-PCR design F in two gDNA samples presenting a high copy number of HERV-W elements (Garcia-Montojo et al., 2013) *(M&M). All 22 sequenced clones corresponded to HERV-W-like sequences, and 17 out of 22 clones could be unequivocally localized. Except for one particular clone (see below), the sequence of all of them corresponded to known HERV-W copies represented in the UCSC Genome database (Table 1C).*

Validation of a HERV-W polymorphism

When the sequence of one clone was mapped to the human genome (GRCh38) using BLAST, the best match found shared only a 93% identity and corresponded to the location X: 147021204 to 147021360 (Figure 2A). Although the search aimed to undoubtedly locate the clone sequence, there were mismatches between the retroviral sequence of the clone and its best match in the database. This lack of homology meant that the clone sequence was not described in the database and it could be an unfixed HERV copy. To study if the polymorphism was linked to multiple sclerosis, a SNP genotyping analysis was performed by pyrosequencing. First, a set of primers was designed to specifically amplify both the cloned sequence and its best genomic match. PCR amplification was carried out using these primers and using the gDNA sample in which the polymorphism had been found as a template. The PCR-amplified product was analysed by pyrosequencing (Figure 2B). It turned out that 98.5% of the PCR product corresponded to the sequence described in the database, while just a 1.5% of the PCR product matched the sequence of the clone. The low percentage of polymorphism detected in the sample indicate that it is not a frequent allele in PBMCs. Given this confirmation, it could not be confirmed if the clone sequence corresponded to a newly discovered unfixed copy or a PCR-generated

artefact. However, these data ratified the usefulness of the adapted method allowing the identification of possible *loci* of interest.

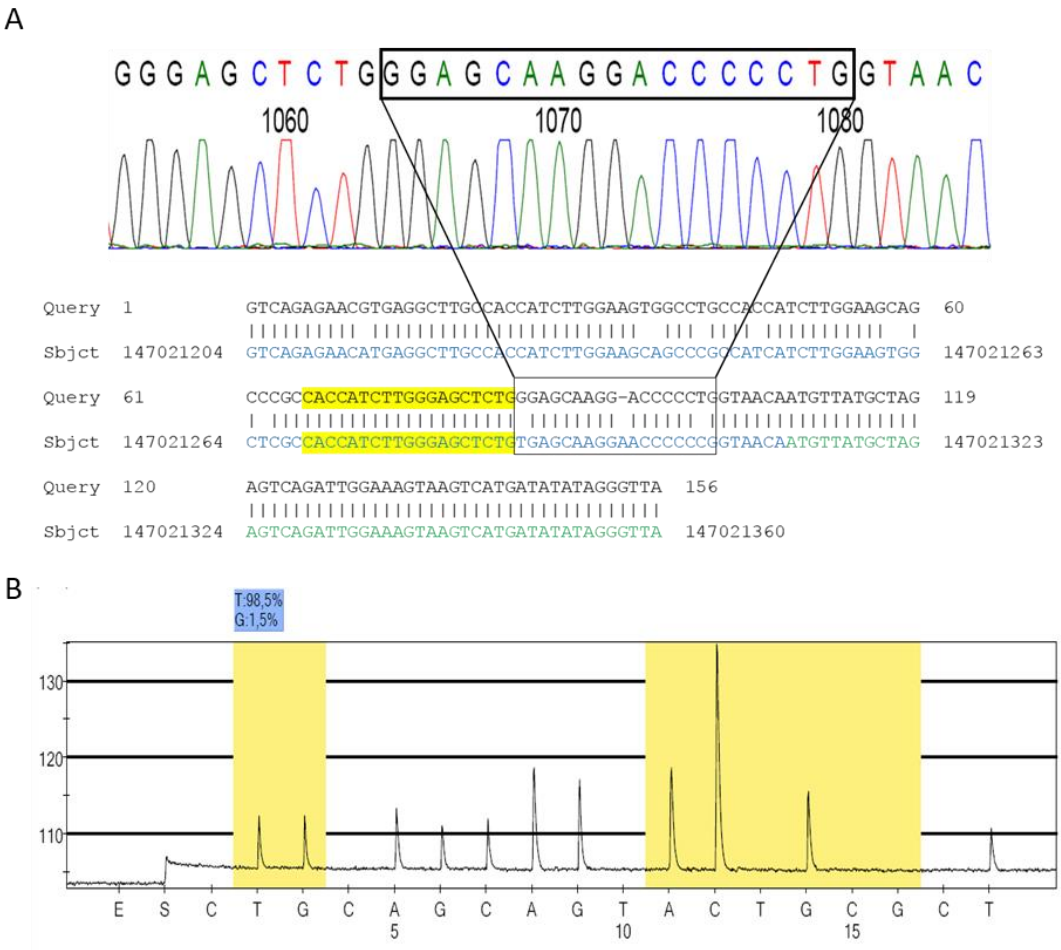


Figure 2. Analysis of a suspected genetic polymorphism. A) Alignment of the sequenced polymorphic clone that carries the (Query) with its best match in the human genome database (Sbjct). Location: Homo sapiens chromosome X, GRCh38.p7 Primary Assembly 147021204 to 147021360, identities: 146/157(93%), gaps: 1/157(0%), strand: Plus/Plus. LTR sequence is marked in blue, the sequence that lies outside the retroviral element is marked in green. The sequencing primer used in the pyrosequencing is highlighted in yellow. Partial representation of its chromatogram is represented above and the rectangle indicates the region analysed by pyrosequencing. B) Pyrogram resulting from the SNP analysis. Vertical axis: signal intensity; horizontal axis: order of residues analysed. The signal intensity is proportional to the number of incorporated nucleotides, and the DNA sequence can be read based on the appearance of a peak and the height of the signal in the pyrogram. Position of SNPs is highlighted in yellow. The quality of the result at each position is estimated by the software and represented through a colour-based score: blue ("passed"), yellow ("to check") and red ("failed").

Next-generation sequencing analysis

To extend the genomic search for HERV-W *loci* not represented in the genome database, splinkerette-PCR products were analysed with NGS technology. Splinkerette-

PCR design A was applied to DNA samples from 2 MS patients with verified high HERV-W copy numbers (Roberto Alvarez Lafuente, personal communication), splinkerette-PCR products were generated, and subjected to IonTorrent based NGS (Figure 4A). NGS data was examined using the software RetroSeq (Keane et al., 2013), which detects non-reference TE insertions from paired-ends NGS data. In the context of splinkerette-PCR products, the paired reads of the DNA sequence correspond to the sequence of a HERV-W element and its bordering genomic sequence, respectively (Figure 3). Briefly, RetroSeq software selects pairs with homology to HERV-W and compares the location of mapped paired-end reads in the human reference genome. If the different reads of a pair are mapped close to each other, the DNA sequence is described in the human reference genome and corresponds to a copy present in the reference genome. While if the couple of reads are not mapped in nearby positions, they are a "discordant read pair" and it could be a variant HERV-W integration site. Additional filters incorporated in the software require a minimum number of reads to overlap in the same location. RetroSeq software did identify non-reference HERV-W insertions within our NGS data (Figure 4B). The relationship between these insertions and MS is unknown at present, and their presence should be compared between MS patients and a control cohort.

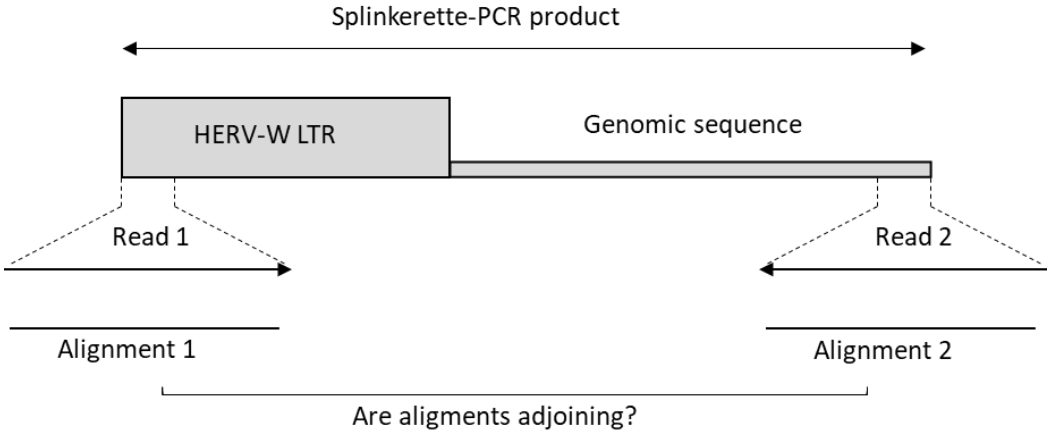


Figure 3. Schematical representation of RetroSeq analysis. A typical splinkerete-PCR product is composed of a partial HERV-W element and its bordering genomic sequence. DNA fragments generated by splinkerette-PCR were sequenced and data obtained was converted to short paired-end reads (indicated as Read 1 and Read 2 in the figure). The obtained pairs were mapped to the human reference genome (version hg19). RetroSeq software analysis determines whether the two alignments are adjoining or not and detects whether the sequence corresponds to the human reference genome or represents a variant HERV-W integration site.

A

	Sample 1	Sample 2
Total reads	2443604	2580868
Average read length	265 bp	270 bp
Mapped reads	1380598	1450291

B

Sample 1	Sample 2
chr1: 199458028	chr1: 98199153
	chr1: 214490727
chr2: 30492384	chr2: 30492384
	chr2: 31508426
chr2: 76106617	chr2: 76106595
chr2: 157502124	chr2: 157502029
chr2: 175529994	
chr3: 83564952	chr3: 83564902
chr4: 68173705	chr4: 8422654
	chr4: 95840422
	chr4: 99411128
	chr4: 134883158
	chr4: 169979087
	chr5: 12947364
	chr5: 59955059
chr5: 143678022	chr5: 143678027
chr6:4192111	chr6: 11649536
chr6: 11649694	
chr6: 41506963	chr7: 35736724
chr6: 71050562	chr7: 96704546
chr8: 124423466	chr8: 124423468
chr9: 114099731	chr11: 23463305
chr11: 36346450	chr11: 36346446
chr11: 77573928	chr11: 124716691
chr12: 78853910	chr12: 78853903
	chr15: 51553635
chr16: 24245427	chr16: 24245427
chr18: 63891213	chr18: 63891151
chrX: 71433744	chrX: 71433713
N= 20	N= 27

Figure 4. A) The table shows the Fastq files produced by IonTorrent sequencing for samples 1 and 2 (total reads), the average read lengths for each sample (bp, base pair), along with the number of total mapped reads. B) The table indicates the non-reference HERV-W insertions that RetroSeq software identified within the NGS data (total reads). The genomic location within the human genome (version hg19) is indicated for each HERV-W insertion and underlined for those non-reference HERV-W insertions that overlaps (within hundred base pairs) between the two samples.

Discussion

Since MSRV particles were detected in MS patient-derived cell cultures, the enlarged presence of MSRV/ HERV-W elements has been repeatedly observed in MS patients. However, limited searches with the aim of identifying the genomic location of MSRV or MS-associated HERV-W copies have been conducted. The distribution and genetic composition of several HERV-W elements in the human reference genome has been approached (Grandi et al., 2016). And other studies have yielded the identification of the HERV-W element which are transcribed in MS patients (Laufer et al., 2009; Schmitt et al., 2013). However, to date, neither replication-competent MSRV provirus encoding such particles nor MS-specific HERV-W *locus* has been identified within the human genome.

To the best of our knowledge, our study represents the first method to discover potential MS-associated copies of MSRV/HERV-W. Adapted from a previously reported splinkerette-PCR method (Uren et al., 2009), this new assay was used successfully to amplify MSRV/HERV-W-like sequences (Table 1). A variety of different HERV-W copies present in gDNA samples from MS patients were detected, showing that the method is not restricted to single element. Sequence reads were mapped to unique HERVW copies with more than 97% sequence identity. Copies identified using this mapped, all represented known HERVW copies. Around 100 PCR products were analysed, of which 97% corresponded to HERV-W elements. We observed that among the 7 alternative splinkerette-PCR designs that were applied, design F (Figure 1B) was the most successful in generating traceable PCR products (Table 1). Such design allowed not only to locate 94% of amplified HERV-W sequences (Table 1) but also revealed a possible non-disease related polymorphism. Nevertheless, all designs applied only identified reference HERV-W elements, failing to identify MSRV or novel HERV-W elements (Figure 2).

MSRV has been defined by several overlapping cDNA sequences derived from MSRV particle-associated RNA isolated from MS patients. However, to date, no complete infectious MSRV sequence has been found within the human genome. Likewise, the origin and nature of MSRV remains unknown. Although at first MSRV was proposed to be an exogenous replication-competent member of the HERV-W family (Perron et al.,

1997), posterior studies are increasingly remarking or suggest that MSRV-like cDNA sequences could be derived from recombination events between HERV-W elements (Laufer et al., 2009). In this study, we intended to locate MSRV sequences using a set of primers published and used to detect higher MSRV-like sequences expression in PBMCs of MS patients over controls (Mameli et al., 2009). Nevertheless, our method (Figure 1, design D) only yielded HERV-W elements present in the reference genome. Therefore, we failed so far to amplify novel HERVW sequences that may represent MSRV, or sequences identical to MSRV-sequences published (Blond et al., 1999; Perron et al., 1997). We therefore did not find evidence supporting the existence of a unique HERV-W related copy that could represent MSRV.

On the other hand, we have hypothesized that unfixed HERV-W *loci* and/or unmapped polymorphisms could play a relevant role in MS disease. To date, several retrotransposition events have been associated with human disease (Hancks and Kazazian, 2012, 2016; O'Donnell and Burns, 2010). Most of the causative agents corresponded to the known TEs active subfamilies such as Alu and LINE-1. Within HERV subfamily, in theory, novel retroviral insertions have only been reported for elements of the HERV-K family (Belshaw et al., 2005, 2005; Kahyo et al., 2013; Wildschutte et al., 2016), but cannot be totally rejected that does not occur for HERV-W family members. In relation to polymorphisms, on the contrary, there are already some described polymorphisms associated with MS susceptibility, such as one near HERV-W *locus* in chromosome Xq22.3 and another one near HERV-Fc1 *locus* (an HERV-H family member) on chromosome X (García-Montojo et al., 2014; Hansen et al., 2011). However, all except one of the sequences amplified by the technique were described in the human reference genome. The sequence missing in the databases resulted to be a low frequently allele (Figure 2) and, therefore, it seems unlikely that is MS-related.

In recent years different informatics tools have been developed to identify novel transposable elements (TEs) insertions in the human genome (Rishishwar et al., 2017). Such methods could help in identifying disease-causing insertions, and better interpreting the consequences that genetic variation may have on human phenotype. All these informatics tools work using NGS sequence data as input. Here we took advance of one of such informatic tools, named RetroSeq (Keane et al., 2013), which

can accurately discover non-reference TEs insertions. Splinkerette-PCR was carried out on DNA obtained from 2 MS patients, NGS was applied to the resulting PCR products, and data was analysed using the RetroSeq software. Several non-reference insertions were identified in this analysis (Figure 3B), but their relation to MS has not been established yet. In conclusion, the method we describe here along with NGS, allows for systematic searches for uncharacterized HERV-W proviral sequences. However, further research is required to clear up the origin and nature of MSRV and identify possible MS-associated unfixed HERV-W copies.

Supplementary materials

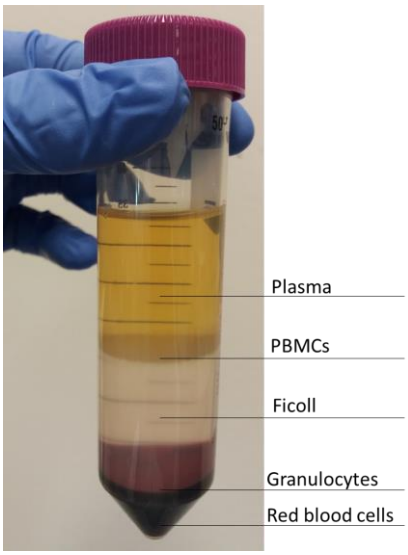


Figure s1. Isolation of mononuclear cells from human peripheral blood by density gradient centrifugation. Differential migration of blood cells during centrifugation results in the formation of layers containing different cell types. From top to bottom: plasma and other constituents, a layer of mononuclear cells (PBMC), Ficoll, and erythrocytes and granulocytes which should be present in pellet form.

Design	1st enzyme	2nd enzyme	LTR primer sequence (5'-3')
A	MseI	BlpI	LTR1:TCCTGCATGGCTAAGTGCCCAG LTR2:TCTTCCGTGACCCATGGCTTCT Degenerate primers: LTR1:TCCTGCAYGGCTAAGTGCCYRGG LTR2:TYTTCCRTGACCCAYGGCTTCT
B	TaqI or Sau3AI	(AvrII+ NheI) + BlpI	LTR1:GAATYCCTAAGCCTAGCTGGGAA LTR2:GGGCTTGYAACTYAGCTCAC
C	TaqI	AhaIII or (AhaIII + NheI)	LTR1:TCACCTTCCCAGCTAGGCTTAG LTR2:AGGAAATCCAGCTAGTCCTGTCTCT
D	TaqI or Sau3AI	EarI	LTR1:CTTCCAGAATTGAAGCTGTAAAGC LTR2:TTCTTCAAATGGARCCYCAGAKG
E	Sau3AI	NaeI	LTR1:CCGGCAACGACTACCCTCTT LTR2:CCTTTGTATGGGAGCTCTGT
F	MseI	BlpI+ SpeI	LTR1:ATAACACTCACCGCATGGCCCAA LTR2:GTCAGAGAACGTGAGGCTTG
G	Sau3AI	AhaIII+ ScaI	LTR1:TCACCTTCCCAGCTAGGCTTAG LTR2:AGGAAATCCAGCTAGTCCTGTCTCT

Table s1. List of all splinkerette designs. The table shows the first enzyme (1st enzyme) and the second enzyme (2nd enzyme) used in the library construction of each splinkerette design. The table also includes a list of primer sequences used for the first (LTR1) and second (LTR2) round of PCR in each design. Y= C or T.

Primer name	Primer sequence (5'-3')
Long strand adapter	CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTGGTGTCGACACTAGTGG
Short strand adapter MSEI	TACCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
Short strand adapter TAQI	CGCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
Short strand adapter SAU3AI	GATCCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
Splink1	CGAAGAGTAACCGTTGCTAGGAGAGACC
Splink2	GTGGCTGAATGAGACTGGTGTCGAC

Table s2. List of additional primers used in the library construction and amplification.

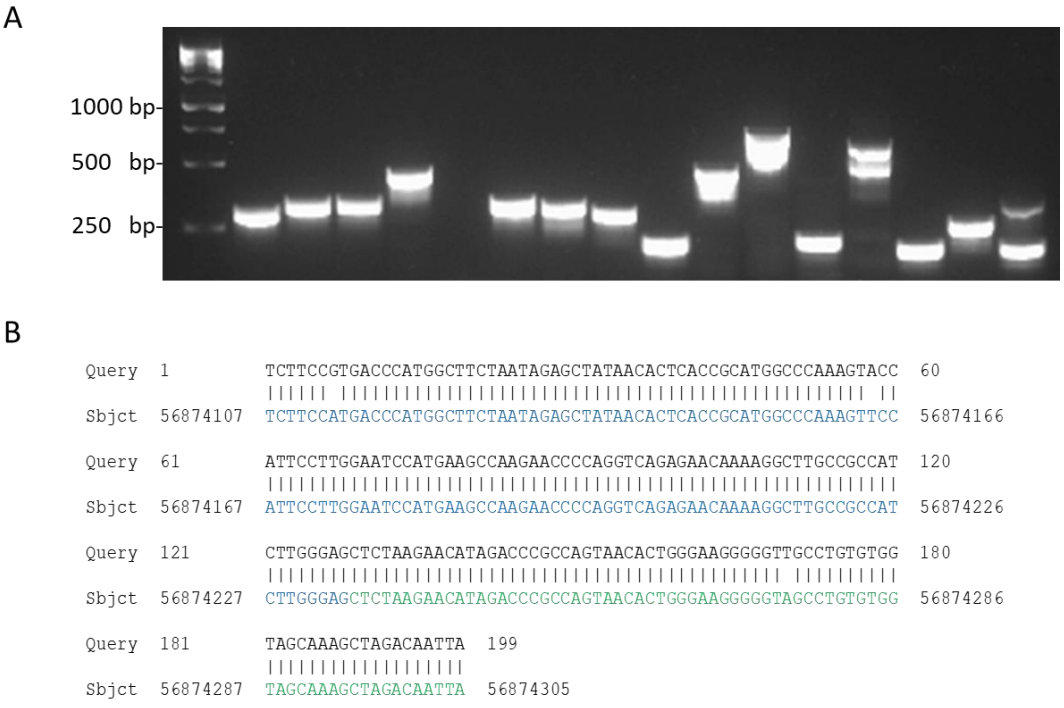


Figure s2. Example of a gel electrophoresis related to splinkerette-PCR products that have been cloned and an alignment between a sequence clone (Query) with its best match in the human genome database (Sbjct). A) Colony PCR amplicons indicate plasmids containing insert. Note the difference in size of the amplicons representing the different sequences amplified by the splinkerette-PCR design. Bp= base pairs. B) Location: Homo sapiens chromosome 11, GRCh38.p7 Primary Assembly 56874107 to 56874305, identities: 196/199(98%), gaps: 0/199(0%), strand: Plus/Plus. LTR sequence is marked in blue and the sequence that lies outside the retroviral element is marked in green.

References

Antony, J.M., Deslauriers, A.M., Bhat, R.K., Ellestad, K.K., and Power, C. (2011). Human endogenous retroviruses and multiple sclerosis: innocent bystanders or disease determinants? Biochim. Biophys. Acta *1812*, 162–176.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA *6*, 11.

Belshaw, R., Dawson, A.L.A., Woolven-Allen, J., Redding, J., Burt, A., and Tristem, M. (2005). Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J. Virol. *79*, 12507–12514.

Bénit, L., Dessen, P., and Heidmann, T. (2001). Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. J. Virol. *75*, 11709–11719.

Blond, J.L., Besème, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet, F. (1999). Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. J. Virol. *73*, 1175–1185.

Blond, J.L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F.L. (2000). An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J. Virol. *74*, 3321–3329.

Cho, K., Lee, Y.-K., and Greenhalgh, D.G. (2008). Endogenous retroviruses in systemic response to stress signals. Shock Augusta Ga *30*, 105–116.

Colmegna, I., and Garry, R.F. (2006). Role of endogenous retroviruses in autoimmune diseases. Infect. Dis. Clin. North Am. *20*, 913–929.

Costas, J. (2002). Characterization of the intragenomic spread of the human endogenous retrovirus family HERV-W. Mol. Biol. Evol. *19*, 526–533.

Dolei, A. (2006). Endogenous retroviruses and human disease. Expert Rev. Clin. Immunol. *2*, 149–167.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., Cirotto, M.C., Rosati, G., and Sotgiu, S. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology *58*, 471–473.

Garcia-Montojo, M., Dominguez-Mozo, M., Arias-Leal, A., Garcia-Martinez, Á., De las Heras, V., Casanova, I., Faucard, R., Gehin, N., Madeira, A., Arroyo, R., et al. (2013). 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*, e53623.

García-Montojo, M., de la Hera, B., Varadé, J., de la Encarnación, A., Camacho, I., Domínguez-Mozo, M., Arias-Leal, A., García-Martínez, Á., Casanova, I., Izquierdo, G., et al. (2014). HERV-W polymorphism in chromosome X is associated with multiple sclerosis risk and with differential expression of MSRV. Retrovirology *11*, 2.

Garson, J.A., Tuke, P.W., Giraud, P., Paranhos-Baccala, G., and Perron, H. (1998). Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet Lond. Engl. *351*, 33.

Grandi, N., and Tramontano, E. (2017). Type W Human Endogenous Retrovirus (HERV-W) Integrations and Their Mobilization by L1 Machinery: Contribution to the Human Transcriptome and Impact on the Host Physiopathology. Viruses *9*.

Grandi, N., Cadeddu, M., Blomberg, J., and Tramontano, E. (2016). Contribution of type W human endogenous retroviruses to the human genome: characterization of HERV-W proviral insertions and processed pseudogenes. Retrovirology *13*.

Hancks, D.C., and Kazazian, H.H. (2012). Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. *22*, 191–203.

Hancks, D.C., and Kazazian, H.H. (2016). Roles for retrotransposon insertions in human disease. Mob. DNA *7*.

Hansen, B., Oturai, A.B., Harbo, H.F., Celius, E.G., Nissen, K.K., Laska, M.J., Søndergaard, H.B., Petersen, T., and Nexø, B.A. (2011). Genetic association of multiple sclerosis with the marker rs391745 near the endogenous retroviral locus HERV-Fc1: analysis of disease subtypes. PloS One *6*, e26438.

Hon, G.M., Erasmus, R.T., and Matsha, T. (2013). Multiple sclerosis-associated retrovirus and related human endogenous retrovirus-W in patients with multiple sclerosis: a literature review. J. Neuroimmunol. *263*, 8–12.

Kahyo, T., Tao, H., Shinmura, K., Yamada, H., Mori, H., Funai, K., Kurabe, N., Suzuki, M., Tanahashi, M., Niwa, H., et al. (2013). Identification and association study with lung cancer for novel insertion polymorphisms of human endogenous retrovirus. Carcinogenesis *34*, 2531–2538.

Keane, T.M., Wong, K., and Adams, D.J. (2013). RetroSeq: transposable element discovery from next-generation sequencing data. Bioinformatics *29*, 389–390.

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E., Mallet, F., Mandrand, B., et al. (1999). Molecular Cloning and Characterization of MSRV-Related Sequences Associated with Retrovirus-like Particles. Virology *260*, 1–9.

Kremer, D., Schichel, T., Förster, M., Tzekova, N., Bernard, C., van der Valk, P., van Horssen, J., Hartung, H.-P., Perron, H., and Küry, P. (2013). Human endogenous retrovirus type W envelope protein inhibits oligodendroglial precursor cell differentiation. Ann. Neurol. *74*, 721–732.

Kurth, R., and Bannert, N. (2010). Beneficial and detrimental effects of human endogenous retroviruses. Int. J. Cancer *126*, 306–314.

Laufer, G., Mayer, J., Mueller, B.F., Mueller-Lantzsch, N., and Ruprecht, K. (2009). Analysis of transcribed human endogenous retrovirus W env loci clarifies the origin of multiple sclerosis-associated retrovirus env sequences. Retrovirology *6*, 37.

Li, F., Nellåker, C., Yolken, R.H., and Karlsson, H. (2011). A systematic evaluation of expression of HERV-W elements; influence of genomic context, viral structure and orientation. BMC Genomics *12*, 22.

Löwer, R., Löwer, J., and Kurth, R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. U. S. A. *93*, 5177–5184.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., Serra, C., Sotgiu, S., Bonetti, B., and Dolei, A. (2007). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J. Gen. Virol. *88*, 264–274.

Mameli, G., Poddighe, L., Astone, V., Delogu, G., Arru, G., Sotgiu, S., Serra, C., and Dolei, A. (2009). 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*, 98–106.

Marchi, E., Kanapin, A., Byott, M., Magiorkinis, G., and Belshaw, R. (2013). Neanderthal and Denisovan retroviruses in modern humans. Curr. Biol. *23*, R994–R995.

Marchi, E., Kanapin, A., Magiorkinis, G., and Belshaw, R. (2014). Unfixed Endogenous Retroviral Insertions in the Human Population. J. Virol. *88*, 9529–9537.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785–789.

Morandi, E., Tarlinton, R.E., Tanasescu, R., and Gran, B. (2017a). Human endogenous retroviruses and multiple sclerosis: Causation, association, or after-effect? Mult. Scler. J. *23*, 1050–1055.

Morandi, E., Tanasescu, R., Tarlinton, R.E., Constantinescu, C.S., Zhang, W., Tench, C., and Gran, B. (2017b). The association between human endogenous retroviruses and multiple sclerosis: A systematic review and meta-analysis. PLoS ONE *12*.

Moyes, D., Griffiths, D.J., and Venables, P.J. (2007). Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. Trends Genet. TIG *23*, 326–333.

Nissen, K.K., Laska, M.J., Hansen, B., Pedersen, F.S., and Nexø, B.A. (2012). No additional copies of HERV-Fc1 in the germ line of multiple sclerosis patients. Virol. J. *9*, 188.

O'Donnell, K.A., and Burns, K.H. (2010). Mobilizing diversity: transposable element insertions in genetic variation and disease. Mob. DNA *1*, 21.

Perron, H., Geny, C., Laurent, A., Mouriquand, C., Pellat, J., Perret, J., and Seigneurin, J.M. (1989). Leptomeningeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. Res. Virol. *140*, 551–561.

Perron, H., Lalande, B., Gratacap, B., Laurent, A., Genoulaz, O., Geny, C., Mallaret, M., Schuller, E., Stoebner, P., and Seigneurin, J.M. (1991). Isolation of retrovirus from patients with multiple sclerosis. Lancet Lond. Engl. *337*, 862–863.

Perron, H., Garson, J.A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P.W., Voisset, C., Blond, J.L., et al. (1997). Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc. Natl. Acad. Sci. U. S. A. *94*, 7583–7588.

Perron, H., Perin, J.P., Rieger, F., and Alliel, P.M. (2000). Particle-associated retroviral RNA and tandem RGH/HERV-W copies on human chromosome 7q: possible components of a "chain-reaction" triggered by infectious agents in multiple sclerosis? J. Neurovirol. *6 Suppl 2*, S67-75.

Perron, H., Jouvin-Marche, E., Michel, M., Ounanian-Paraz, A., Camelo, S., Dumon, A., Jolivet-Reynaud, C., Marcel, F., Souillet, Y., Borel, E., et al. (2001). Multiple sclerosis retrovirus particles and recombinant envelope trigger an abnormal immune response in vitro, by inducing polyclonal Vbeta16 T-lymphocyte activation. Virology *287*, 321–332.

Perron, H., Lazarini, F., Ruprecht, K., Péchoux-Longin, C., Seilhean, D., Sazdovitch, V., Créange, A., Battail-Poirot, N., Sibaï, G., Santoro, L., et al. (2005). Human endogenous retrovirus (HERV)-W ENV and GAG proteins: physiological expression in human brain and pathophysiological modulation in multiple sclerosis lesions. J. Neurovirol. *11*, 23–33.

Perron, H., Germi, R., Bernard, C., Garcia-Montojo, M., Deluen, C., Farinelli, L., Faucard, R., Veas, F., Stefas, I., Fabriek, B.O., et al. (2012). Human endogenous retrovirus type W envelope expression in blood and brain cells provides new insights into multiple sclerosis disease. Mult. Scler. J. *18*, 1721–1736.

Rishishwar, L., Mariño-Ramírez, L., and Jordan, I.K. (2017). Benchmarking computational tools for polymorphic transposable element detection. Brief. Bioinform. *18*, 908–918.

Rolland, A., Jouvin-Marche, E., Viret, C., Faure, M., Perron, H., and Marche, P.N. (2006). The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. J. Immunol. Baltim. Md 1950 *176*, 7636–7644.

Schmitt, K., Richter, C., Backes, C., Meese, E., Ruprecht, K., and Mayer, J. (2013). Comprehensive Analysis of Human Endogenous Retrovirus Group HERV-W Locus Transcription in Multiple Sclerosis Brain Lesions by High-Throughput Amplicon Sequencing. J. Virol. *87*, 13837–13852.

Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M.I., Kidd, K.K., and Lenz, J. (2001). Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr. Biol. CB *11*, 1531–1535.

Uren, A.G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A.H., Wilson, C.H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A., et al. (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat. Protoc. *4*, 789–798.

Villesen, P., Aagaard, L., Wiuf, C., and Pedersen, F.S. (2004). Identification of endogenous retroviral reading frames in the human genome. Retrovirology *1*, 32.

Wildschutte, J.H., Williams, Z.H., Montesion, M., Subramanian, R.P., Kidd, J.M., and Coffin, J.M. (2016). Discovery of unfixed endogenous retrovirus insertions in diverse human populations. Proc. Natl. Acad. Sci. U. S. A. *113*, E2326–E2334.

Manuscript II. HERV-W expression in multiple sclerosis

Abstract

The upregulated presence and expression of multiple sclerosis (MS)-associated retrovirus (MSRV) and its relative human endogenous retroviruses family W (HERV-W) has been repeatedly detected in autopsied brain tissues and peripheral blood mononuclear cells (PBMCs) of MS patients. However, little is known about which HERV-W elements are overexpressed in MS patients and none MS-associated HERV-W *loci* have been identified so far. We therefore intended to assess HERV-W *ENV* expression in PBMCs of MS patients and search for MS-associated HERV-W *loci.*

We analysed HERV-W *ENV* transcription levels by real-time quantitative PCR (qPCR) in PBMCs of MS patients and controls. An initial assay revealed that while MS patients did not show differences in HERV-W *ENV* expression over controls, CIS patients presented a 1.37-fold increase in HERV-W *ENV* expression compared to controls (p=0.006). We analysed HERV-W *ENV*-specific PCR products generated after reverse transcription of RNA of 6 CIS patients and 6 controls to identify *loci* transcribed in PBMCs. After cloning, sequencing identified 21 transcribed HERV-W *ENV loci* within 141 sequenced clones. Moreover, one *locus* on chromosome 5p12 showed relative transcript levels 4-time higher in MS patients over controls (p=0.016). In addition, a next-generation sequencing (NSG) analysis yielded the identification and genomic location of HERV-W *ENV loci* transcribed in PBMCS of 4 CIS patients and 4 controls. While the most abundantly transcribed HERV-W *ENV loci* did not differ between patients and controls, several MS patient-associated HERV-W *ENV loci*, including the HERV-W 5p12 element, were identified. To extend these finding to larger study cohorts, HERV-W *ENV* expression was analysed in a larger collection (n=116). Results showed that HERV-W *ENV* expression was just slightly induced in MS patients, which presented a 1.2-fold increase in HERV-W *ENV* expression compared to controls. A second cohort of PBMCs samples from MS patients confirmed the overexpression of HERV-W *ENV* in MS patients, especially in those being clinically active, which presented a 2-fold increase in HERV-W *ENV* expression compared to controls (p=0). We therefore identified specific copies that may contribute to HERV-W overexpression in PBMCs MS patients

Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS), characterized by demyelination of axons of nerve cells and loss of oligodendrocytes (Compston and Coles, 2002, 2008). When neural damage occurs, the communication among different parts of the nervous system is blocked and MS symptoms appear, including muscle weakness, weak reflexes, impaired body movement, and miss-coordination among others. There are several types of MS which are mainly defined by how often a person experiences symptoms (Lublin et al., 1996). Relapsing-remitting MS (RRMS), the most common type of MS, is characterized by clearly defined attacks (relapses) of new or increasing neurologic symptoms. These attacks are followed by periods of partial or complete recovery (remissions). Initial RRMS usually leads to a secondary progressive MS (SPMS), in which there is a progressive worsening of neurologic function over time. Primary progressive MS (PPMS) is characterized by worsening neurologic function from the onset of symptoms, without early relapses or remissions. Later, new disease courses were determined, such as clinically isolated syndrome (CIS) (Miller et al., 2005). CIS refers to a first neurological episode, caused by inflammation or demyelination in the CNS, which may be either the onset of MS, or represent an isolated episode not related to MS.

Human endogenous retroviruses (HERVs) are the result of ancestral retroviral infections that entered the human germ line, leading to a permanent retroviral integration transmitted to the offspring in a Mendelian fashion. After initial integration, HERVs continued to generate self-copies and as a result they now comprise approximately 8% of our genome (Cho et al., 2008; Löwer et al., 1996).

The association between human endogenous retroviruses (HERVs) and MS was first studied in the late 1980's, when free virus-like particles were purified from cells and plasma of MS patients (Perron et al., 1989, 1991). Molecular characterisation of the viral RNA within these particles revealed partial retroviral-like sequences, upon which the retrovirus was named MS-associated or MSRV (Perron et al., 1997). MSRV turned out to be closely related to human endogenous retroviruses family W (HERV-W), a multicopy family consisting of approximately 650 closely related but distinguishable elements dispersed in the human genome (Pavlícek et al., 2002). However, the exact

relationship between MSRV and HERV-W has not been clarified yet (Blond et al., 1999; Komurian-Pradel et al., 1999; Perron et al., 2000) and therefore, the term MSRV/HERV-W is often used. The linkage between expression of MSRV/HERV-W and MS has been repeatedly underlined since MSRV discovery. Not only was DNA load of MSRV/ HERV-W found to be increased in peripherical blood mononuclear cells (PBMC) from MS patients (Garcia-Montojo et al., 2013; Perron et al., 2012), but also MSRV/HERV-W expression was found to be increased in cerebrospinal fluid (CSF), PBMCs and post-mortem brain tissue of MS patients (Dolei et al., 2002; Mameli et al., 2007; Serra et al., 2001). In line with these findings, the levels or amounts of MSRV/HERV-W ENV protein detected in post-mortem brain tissue or serum of MS patients were higher compared to controls (Mameli et al., 2007; Perron et al., 2012). Although most HERVs are highly defective and unable to replicate due to the acquisition of deletions and mutations during evolution, within the HERV-W family there is a copy located on chromosome 7q that encodes for a competent HERV-W envelope protein named SYNCYTIN-1, which is expressed in placenta and involved in placental development (Blond et al., 2000; Mi et al., 2000). Nevertheless, a discriminatory PCR study showed that RNA encoding MSRV-type envelope protein but not *SYNCYTIN-1* is expressed at increased levels in PBMCs of MS patients (Mameli et al., 2009).

Further pursuing this HERV-W association with MS, several studies have tried to identify which HERV-W *ENV loci* are transcribed and the relative transcript levels of those *loci* in PBMC and brain of MS patients and controls (Laufer et al., 2009; Schmitt et al., 2013). However, transcriptional activity of individual HERV-W *ENV* elements did not significantly differ between patients with MS and controls in these studies.

In the present study, we hypothesise that specific copies of HERV-W are overexpressed in MS, and therefore represent MSRV. We tried to identify MS-associated HERV-W *loci.* We collected PBMCs samples from MS patients and healthy controls and both HERV-W *ENV* DNA load and RNA levels were assessed. HERV-W *ENV*-specific cDNA derived-PCR products were sequenced to identify transcribed HERV-W *ENV loci* in PBMCs of MS patients and controls. To the best of our knowledge, we present the first evidence that HERV-W *ENV* expression might be limited to clinically active stages of the disease. We also identify candidates for MS-associated HERV-W *loci.*

Materials and methods

PBMCs isolation and gDNA extraction

PBMCs isolation and gDNA extraction was performed as described before (M&M, Manuscript I).

HERV-W ENV copy number quantification: qPCR

For the PCR we used as template 100 ng of gDNA, previously fragmented by sonication for 5 minutes at high power with *cycles* of *30 seconds* ON, and 30 seconds OFF, using a bioruptor sonicator (Diagenode). Real-Time PCR was performed under standard thermal cycling conditions on a ViiA™ 7 Real-Time PCR System (Applied Biosystems) using the Premix Ex Taq™ (Probe qPCR) Master Mix (RR390A, Takara) according to manufacturer's protocol. The assay was carried out in a total volume of 20 µl using 1x TaqMan® Copy Number Reference Assay RNase P (4403326, Life Technologies), a specific pair of primers and probe (used at 200 nM) for MSRV *ENV* (Mameli et al., 2009), and 1 µl of sonicated genomic DNA as template. Amplification efficiency of all Taqman assays between 90% and 110% was assessed in standard curves (Appendix I). All reactions were carried out in triplicate and only measurements with a standard deviation < 0.2 were considered. Copy number quantification, normalized to the house-keeping control *RNAseP*, was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

RNA isolation and cDNA synthesis

PBMCs were homogenized with TRIzol® reagent (Invitrogen). Total RNA was subsequently isolated from the resulting homogenate using an in-house method (Pérez-Palacios et al., 2016). Isolated RNA was quantified using a Nanodrop spectrophotometer (Thermo Sci., Nanodrop 2000) and kept at -80ºC until further use. cDNA was synthesized from 500 ng of RNA with either random hexamer primer or oligo(dT) using the ThermoScript™ RT-PCR System (11146-024, Invitrogen) according to manufacturer's protocol.

HERV-W ENV expression analysis: qPCR

cDNA was analyzed by Real-Time PCR on a ViiA™ 7 Real-Time PCR System (Applied Biosystems), under standard thermal cycling conditions (ViiA™ 7 System). HERV-W *ENV* expression analysis was performed using a specific pair of primers and probe for MSRV *ENV* (Mameli et al., 2009) in a TaqMan qPCR, using the Premix Ex Taq™ (Probe qPCR) Master Mix (RR390A, Takara) according to manufacturer's protocol. HERV-W *locus*-specific expression analyses were performed in a SYBR Green qPCR, using the SYBR Green Master Mix (SYBR Premix Ex Taq II (Tli RNase H Plus), RR820A, Takara) according to manufacturer's protocol. In all cases, *hGAPDH* expression analysis was performed in a SYBR Green qPCR. Sequences of all primers used are described in Table s1. Amplification efficiency of all set of primers between 90% and 110% was assessed in standard curves (Appendix I) and, sequences of all primers used (used at 200 nM) are listed in Table s1. All reactions were carried out in triplicate and only measurements with a standard deviation < 0.2 were considered. Target gene expression levels were recalculated as $2^{-\Delta\Delta Ct}$ with respect to the median ΔCt value of the samples in the control group, using *hGAPDH* as a reference gene (Livak and Schmittgen, 2001).

Identification and localization of transcribed HERV-W loci

cDNA from PBMCs obtained from MS patients and healthy controls, and primed with either oligo dT or random primer, was amplified employing the specific MSRV *ENV* pair of primers described by Mameli et al., 2009 (Table s1). PCR amplification was carried out in a 50-µl volume using 2 µl of cDNA, the pair of primers previously mentioned (used at 400 nM), 200 µM dNTPs and 0.5 U Taq DNA Polymerase (D1806, Sigma-Aldrich) along with its corresponding 1X PCR reaction buffer containing 1.5 mM $MgCl_2$. PCR cycling conditions were as follows: initial activation at 98 ºC for 2 min, followed by 35 cycles of 95 ºC for 30 seconds, 60 ºC for 30 seconds and 72 ºC for 30 seconds. The final extension step was at 72 ºC for 10 minutes. The PCR product was purified using the Diffinity RapidTip®2 (D2947, Sigma) and precipitated with sodium acetate-ethanol according to standard protocols. The final product was re-suspended in 3 µl of TE buffer and cloned into the pGEM®-T Easy Vector System I (A1360, Promega) according

to manufacturer's instructions. The vector was transformed into E. coli XL1 cells and cultured in Luria–Bertani (LB)-agar plates supplemented with 50 µg/mL ampicillin at 37 ºC. Colonies were screened for the presence of insert by colony PCR and further analysed by plasmid miniprep and sequencing. DNA sequencing analyses were performed by Secugen S.L (Madrid, Spain) and chromosomal location of cloned HERV-W/MSRV cDNA sequences was mapped to the human genome (GRCh38) using the blastn option of BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) (Altschul et al., 1990). Species-specific filters for retroviral elements were not activated (or de-activated). More than 97% identity between the cloned sequence and the database sequence was used as criterium to map the sequenced fragments.

Next Generation Sequencing analysis

Next Generation Sequencing analysis was performed as described before (M&M, Manuscript I).

Statistical analysis

SPSS software was used for all analyses (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 15.0). Normality was assessed with Shapiro-Wilk test (when n< 50) or Kolmogorov-smith test (when n> 50). For the assessment of the statistical significance of differences, a specific test was performed depending on normality and number of groups of samples compared. Normally-distributed data: t-student test (2 groups) or ANOVA test (>2 groups). Non-normally distributed data: U de Mann-Whitney test (2 groups) or Kruskal-Wallis test (>2 groups). The graphs were generated using SPPS software as well. Within a box plot, the bottom line of the box represents the first quartile (Q1) of the data, the middle of the box (shown as a line) represents the median (or second quartile Q2), and the top line of the box represents the third quartile (Q3). The interquartile range (IR) is a measure of dispersion and it is defined as Q3-Q1. The vertical lines extend from the ends of the box to the minimum and maximum values of the data. Outliers are indicated as individual points and defined as being more than 1.5*IR outside the box.

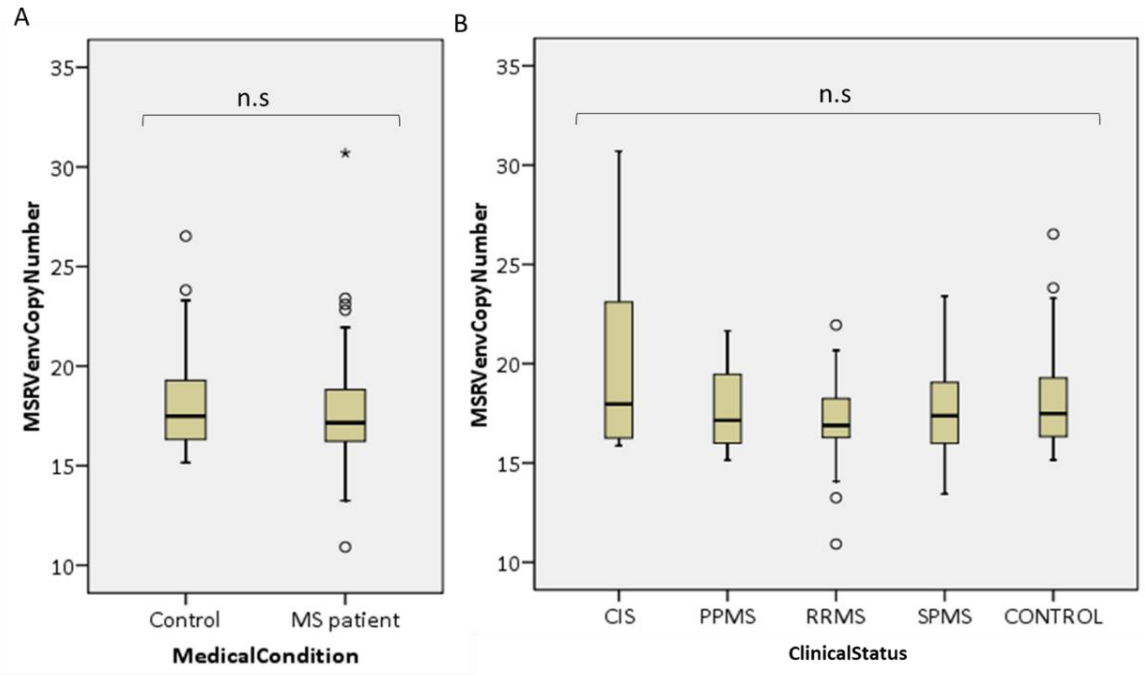
Results

DNA copy number of HERV-W ENV in PBMCs

Aiming at analysing HERV-W *ENV* DNA load in genomic DNA of PBMC from MS patients, HERV-W *ENV* copy number was analysed by a specific quantitative PCR (qPCR). 79 MS patients and 37 healthy controls were included in the analysis (Table 1). Although CIS not necessarily indicates the onset of MS, in the absence of additional clinical data, throughout this manuscript they are considered patients. HERV-W *ENV* DNA load (copies/haploid genome) was measured by a duplex real-time PCR with specific primers and probe for MSRV *ENV*-like sequences (Mameli et al., 2009), and for the endogenous reference gene *RNAseP*. *RNAseP*, a known single-copy gene, was included to normalize the results. MSRV *ENV*-like DNA load was mostly the same in MS patients (17.16 ± 2.66) and controls (17.48 ± 3.44) (U. de Mann Whitney; p= 0.46) (Figure 1A). The stratification of MS patients by clinical status (CIS, PPMS, RRMS, SPMS), showed that there was not difference either in the HERV-W *ENV* DNA load among clinically defined groups (Krustal-Wallis; p= 0.82) (Figure 1B). Therefore, in our cohort HERV-W *ENV* DNA load in PBMCs was not associated with MS.

COHORT 1		
	MS patients (n=79)	**Controls (n=37)**
Disease form		
CIS	n=6	
PPMS	n=10	
RRMS	n=36	
SPMS	n=27	

Table 1. Clinical data of cohort 1. For this study, a cohort of 116 participants, including patients with multiple sclerosis (MS) and controls, were recruited. MS patients presented different diasease forms: relapsing-remitting MS (RRMS), secondary progressive MS (SPMS), primary progressive MS (PPMS) and CIS (clinically isolated syndrome).



	N	Median (IR)	P- value U. De Mann Whitney
MS patients	79	17.16 (2.66)	P= 0.46
Controls	37	17.48 (3.44)	

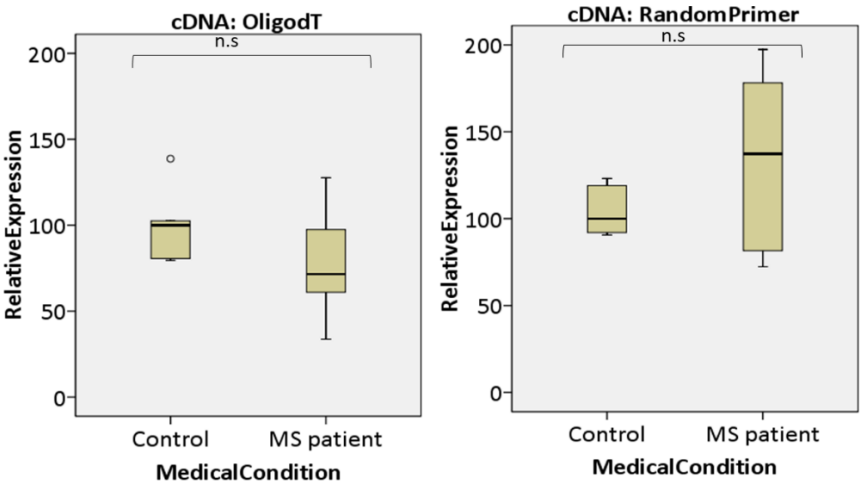
	N	Median (IR)	P- value Kruskal-Wallis
CIS	6	17.96 (8.85)	P= 0.82
PPMS	10	17.15 (3.81)	
RRMS	36	16.88 (2.01)	
SPMS	27	17.37 (3.26)	
CONTROL	37	17.48 (3.44)	

Figure 1. DNA copy number of human endogenous retrovirus-W quantification in PBMCs. HERV-W *ENV* DNA load represents HERV-W *ENV* copies per haploid genome in PBMC samples from MS patients and controls, normalized to the RNAse P reference gene. Analysis was performed comparing A) MS patients versus controls and B) different clinical status in MS and controls. Tables show numerical values taken from the graph.

Human endogenous retrovirus-W ENV expression in PBMCs

Previous studies had indicated a substantial increase in HERV-W *ENV* expression in a small cohort of patients (Mameli et al., 2009), and a small but statistically significant increase in a much larger cohort (García-Montojo et al., 2014). As we were interested in identifying specific copies of HERV-W overexpressed in MS patients, we first confirmed overexpression within a small test cohort of MS patients (n= 7) and healthy controls (n= 5). HERV-W *ENV* expression was measured in PBMCs by quantitative PCR (qPCR), using the same set of primers and specific probe for MSRV *ENV*-like sequences used in the above-mentioned publications. To compare different assays to determine expression levels, cDNA primed with either random primer and oligo dT was obtained

and analysed (Figure 2). While oligo dT-primed cDNA analysis showed that HERV-W *ENV* expression was decreased in MS patients (71.52% ± 59.75) compared to controls (100% ± 40.53) (t-student; p= 0.247), the study of random-primed cDNA showed an increase of HERV-W *ENV* expression in MS patients (137.3% ± 119.39) compared to controls (100% ± 29.8) (U. de Mann Whitney; p= 0.685) (Figure 2). Therefore, the choice of primer used to initiate reverse transcription was found to alter or influence HERV-W *ENV* detection in PBMCs. Since HERV-W *ENV* overexpression in patients was uniquely detected in random primer-primed cDNA samples, we decided on the use of random primer for all future reverse transcriptions. Interestingly, we observed that among the 7 MS samples analysed, those with higher levels of HERV-W *ENV* expression corresponded to the clinical status CIS (data not shown). We therefore decided to include all CIS samples available in the analysis.

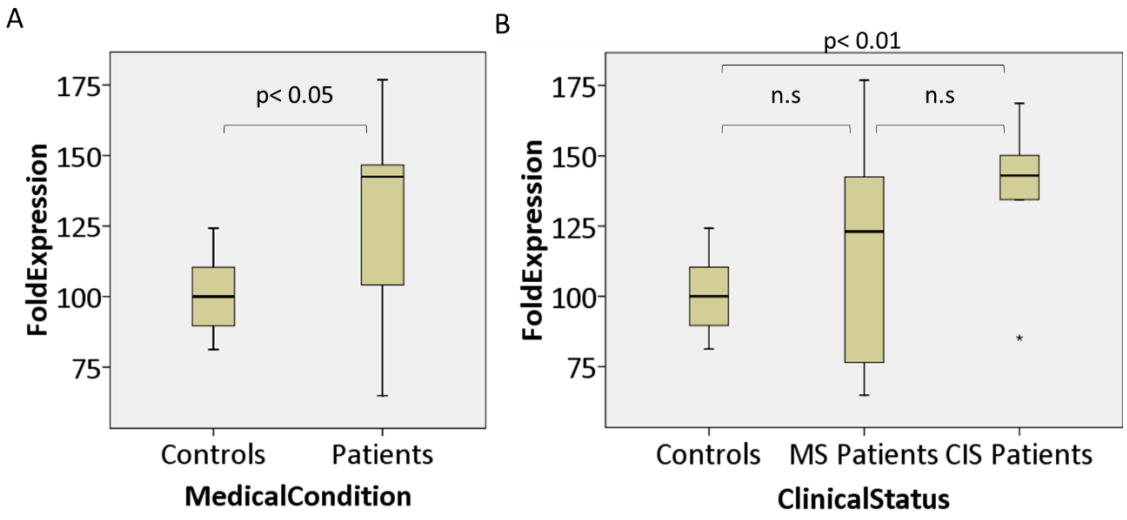


	Median (IR)		P-value
	Controls (N=5)	Patients (N=7)	
Oligo dT	100 (40.53)	71.52 (59.75)	p= 0.247 (t-student)
Random primer	100 (29.8)	137.3 (119.39)	p= 0.685 (U de Mann-Whitney)

Figure 2. Human endogenous retrovirus-W *ENV* expression in PBMCs. HERV-W *ENV* expression was measured by quantitative PCR (qPCR), using a specific set of primers for MSRV *ENV* (Mameli et al., 2009). **A**nalysis was performed using cDNA, primed with either oligo dT or random primer, from 7 MS patients and 5 controls. Results were normalized using *GAPDH* as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. Tables show numerical values taken from the graph.

The optimized qPCR assay described above was repeated to discriminate whether HERV-W *ENV* overexpression was biased by the clinical status of MS patients. For

comparison, the overall sample size was slightly increased to MS patients (n= 11) and healthy controls (n= 9). Of those 11 MS patients, 6 samples corresponded to CIS patients (not properly classified as MS patients yet) and 5 to MS patients (RRMS and SPMS). Results were calculated and represented in the same way as above. In this cohort, HERV-W *ENV* expression was significantly increased in MS patients (142.48% ± 64.92) compared to controls (100% ± 28.51) (t-student; p= 0.045) (Figure 3A). Interestingly, the stratification by clinical status of patients (CIS patients, other MS patients) showed that although HERV-W *ENV* overexpression was detected in both MS and CIS patients, it was higher in CIS patients (142.97% ± 32.66) (t-student; p= 0.006) than in MS patients (123.06% ± 89.01) (t-student; p= 0.449), compared to controls (100% ± 28.51) (Figure 3B). Therefore, while a trend towards HERV-W *ENV* overexpression was detected in MS patients, overexpression was more prominent in CIS patients.



	N	Median (IR)	P- value (t-student)
Controls	9	100 (28.51)	p=0.045
Patients	11	142.48 (64.92)	

	N	Median (IR)	P- value (t-student)
Controls	9	100 (28.51)	
MS patients	5	123.06 (89.01)	p=0.449 (MS-Controls)
CIS patients	6	142.97 (32.66)	p=0.006 (CIS-Controls)

Figure 3. Human endogenous retrovirus-W *ENV* expression in CIS patients. Analysis was performed using cDNA primed with random primer from 11 MS patients and 9 controls. Results were represented comparing A) control group versus MS patients group and, B) control group versus MS patients group divided into CIS patients (n= 6) and the rest of MS patients (n=5). Results were normalized using *GAPDH* as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. Tables show numerical values taken from the graph.

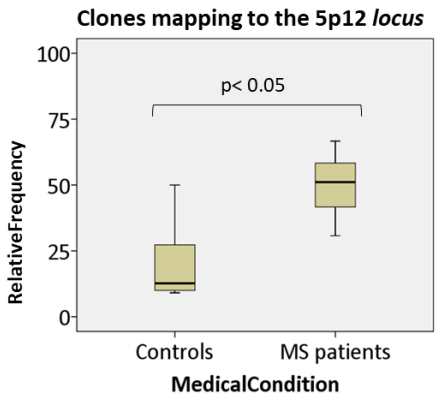
Identification of transcribed HERV-W ENV loci in PBMCs

Aiming at finding MS-associated HERV-W *ENV loci,* we intended to identify which HERV-W *ENV* copies were more abundantly expressed in PBMC from MS patients relative to controls. Oligo dT-primed cDNA was obtained from PBMCs of 6 MS patients and 6 controls, and amplified using the specific PCR for MSRV *ENV* described by Mameli et al., 2009. After molecular cloning, individual clones were sequenced, the sequences were mapped to the human genome and relative cloning frequencies were calculated (the number of cDNA clones mapping to an individual HERV-W *ENV* element relative to the total number of cDNA clones analysed). A total of 76 and 65 cDNA clones, generated from MS patients and controls, respectively, mapped to 21 different HERV-W *ENV* elements. Among those, clones originating from HERV-W *ENV* copies located on chromosomes 5p12, 19q13.2 and Xq22.3 were most frequently encountered (Figure 4A). The relative cloning frequency of individual HERV-W *ENV* copies did not significantly differ between MS patients and controls, except for one HERV-W element. It turned out that the HERV-W copy located on chromosome 5p11 (HERV-W *ENV* 5p12) was 4 times more frequently cloned in samples from MS patients (51.11% ±21.47) compared to controls (12.69% ± 23.18) (U. de Mann Whitney; p= 0.016) (Figure 4B).

A

HERV-W *loci*	MS patients (%) (N=6)	Controls (%) (N=6)
5p12	50	20
19q13.2	20	22
Xq22.3	14	12
12q23.3	5	6
3q11.2	3	3
4q21.22	3	3
15q21.3	3	2
Others	<1	<5
No. of clones	76	65

B

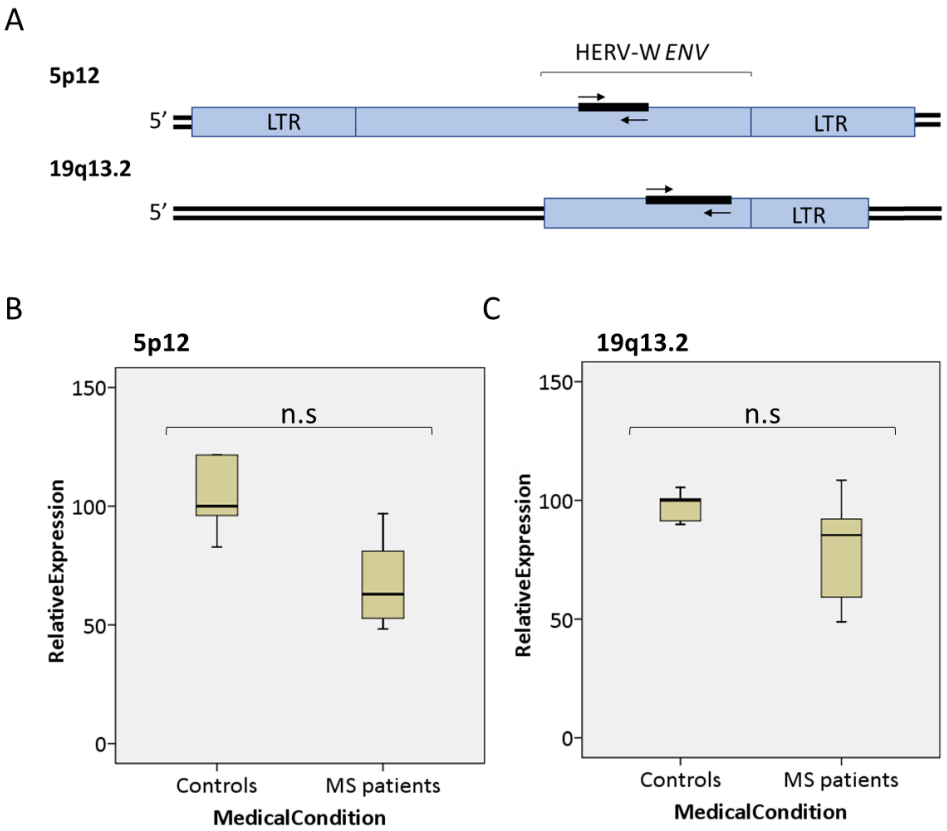


	N	Median (IR)	P- value (U. De Mann Whitney)
Controls	6	12.69 (23.18)	p=0.016
MS patients	6	51.11 (21.47)	

Figure 4. Analysis of transcribed HERV-W *ENV loci* in PBMCs. A specific PCR (Mameli G et al., 2009) was applied to a subset of the samples analysed previously (Figure 2). After molecular cloning, individual clones were sequenced, and their location was determined by comparing the resulting sequences the human genome. A) The chromosomal location of HERV-W *ENV* clones is indicated in the first column. Numbers in the second and third columns represent the percentage of clones mapping to the *loci* indicated in 6 MS patients and 6 control, respectively. B) The relative frequency of clones derived from the 5p12 *locus* is indicated. Tables show numerical values taken from the graph.

Therefore, results suggested that HERV-W *ENV* 5p12 overexpression might be associated with MS.

A similar search using random primer-primed cDNA yielded data suggesting that HERV-W *ENV* 19q13.2 overexpression was potentially MS-associated (data not shown). To verify the possible association between both HERV-W *ENV* 5p12 and 19q13.2 elements and MS prevalence, we performed locus-specific expression analyses. HERV-W *ENV* 5p12 and 19q13.2-specific qPCR assays were designed (Figure 5A). Primer pairs were designed so that they do not match the sequences most similar to elements HERV-W *ENV* 5p12 and 19q13.2. Melting curve analyses of these qPCR assays showed a single peak, verifying the presence of a single PCR product and the specificity of the qPCR assays (data not shown). Expression levels were assessed in PBMCs from 7 MS patients and 5 controls. MS patients showed less HERV-W *ENV* 5p12 expression levels (62.91% ± 44.2) than controls did (100% ± 124.04) (U. de Mann Whitney; p= 0.062) (Figure 5B). HERV-W *ENV* 19q13.2 expression levels were also found to be lower in MS patients (85.38% ± 40.01) than in controls (100% ± 12.45) (U. de Mann Whitney; p= 0.063) (Figure 5C). Therefore, although initial analyses pointed out HERV-W copies 5p12 and 19q13.2 as potential MS-associated elements, we were unable to verify those data in copy-specific expression analyses.



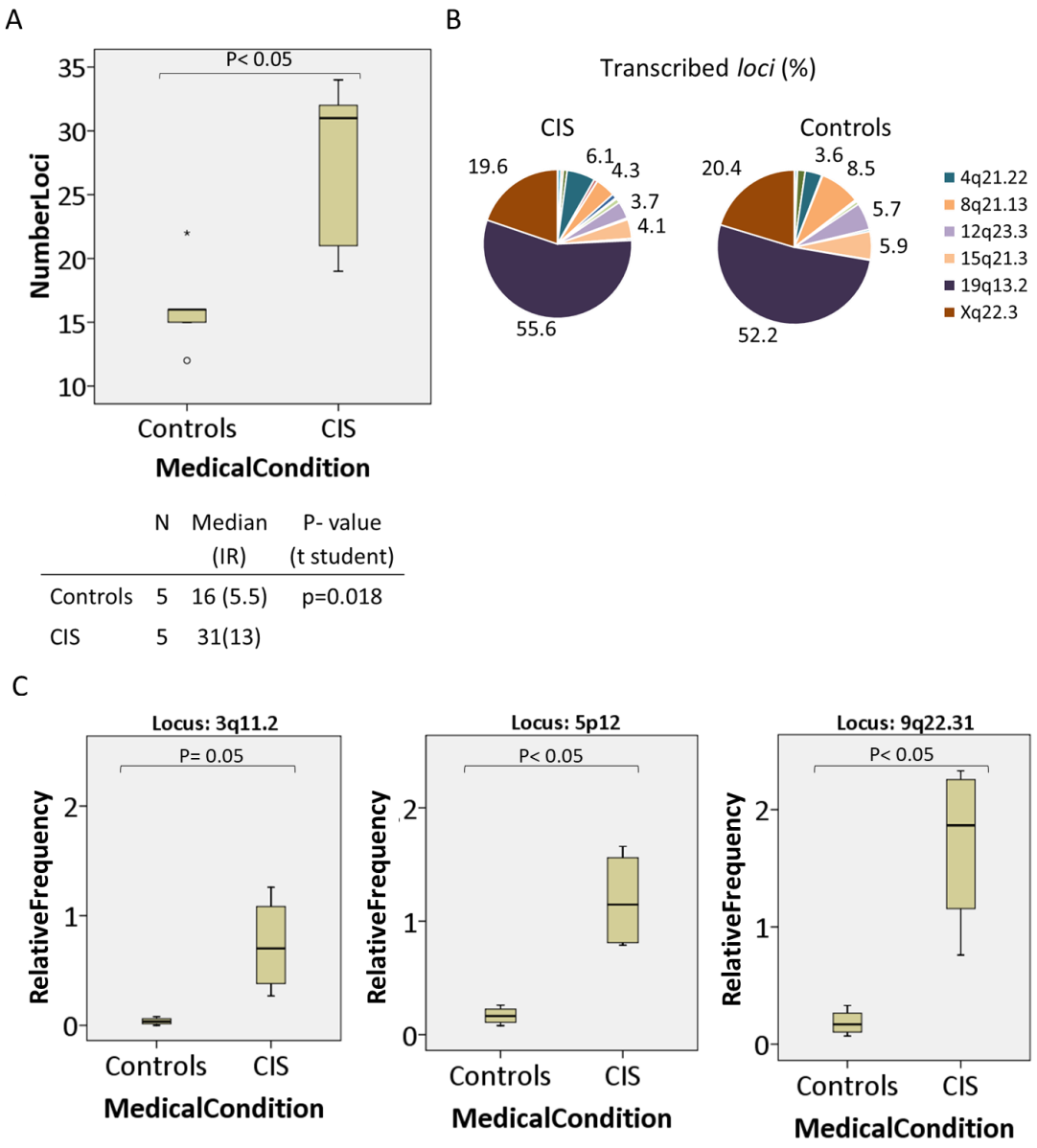
	Median (IR)		p- value
	Controls (N=5)	**Patients (N=7)**	
HERV-W 5p12	100 (124.04)	62.91 (44.2)	p= 0.062 (U de Mann-Whitney)
HERV-W 19q13.2	100(12.45)	85.38 (40.01)	p=0.063 (t-student)

Figure 5. qPCR specific expression analysis for HERV-W *ENV* 5p12 and 19q13.2 in PBMCs. A) Schematical representation of the copy-specific expression assays designed. Arrows represent the pair of primers used in the qPCR to specifically amplify the elements HERV-W 5p12 and 19q13.2 within the *ENV* region. LTR: long terminal repeats. B) Copy-specific expression analysis of HERV-W 5p12 in oligo dT-primed cDNA. C) Copy-specific expression analysis of HERV-W 19q13.2 in random primer-primed cDNA. B, C) Samples, 7 MS patients and 5 controls, were taken randomly from the samples analyzed previously (Figure 2). Results were normalized using *GAPDH* as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. Table shows numerical values taken from the graph.

Next-generation sequencing analysis of transcribed HERV-W ENV loci in PBMCs

We next performed a NGS analysis to identify new candidate MS-associated HERV-W copies. NGS analysis was performed on amplification products obtained using the specific PCR described by Mameli et al., 2009 and random-primed cDNA from PBMCs of CIS patients (n=5) and control (n=5) as templates. Reads (around 60,000 per sample) were assigned to a HERV-W *ENV locus* (the sequences were mapped to the human genome according to criteria described in M&M) and relative frequencies were calculated (the number of reads mapping to an individual HERV-W *ENV* element relative to the total number of reads). The resulting NGS data showed that sequences obtained from CIS patients mapped to a significant higher number of different HERV-W *ENV loci* (31 ± 13), compared to those obtained from controls (16 ± 5.5) (t-student; p= 0.018) (Figure 6A). Reads mapped with high frequency to a limited number of *loci*, especially to HERV-W *ENV* copies located on chromosomes 19q13.2, Xq22.3, 8q21.13, 15q21.3, 12q23.3 and 4q21.22 (Figure 6B). We found no significant differences between CIS patients and controls in the relative frequency of reads mapping to these *loci* (Figure 6B). However, reads mapped with intermediate frequency to another subset of *loci*, located on chromosome 3q11.2, 5p12 and 9q22.31. Reads mapping to these *loci* were mainly identified in CIS patients. The relative frequency of reads mapping to these showed an increase of 17.5, 6.8 and 11 times respectively in CIS patients compared to controls (Figure 6C). It stands to reason, that the relative frequency after mapping is the equivalent of relative overexpression. As most reads obtained from both CIS patients and controls mapped to a limited set of apparently highly transcribed copies, our data indicate that the observed HERV-W *ENV* overexpression in CIS patients (Figure 3B) did not originate from a single HERV-W *ENV* copy. Such overexpression appears to be derived from less abundantly transcribed but CIS-related HERV-W *ENV loci.*

To verify the overexpression of CIS-related HERV-W copies in CIS patients, specific qPCR assays were designed and applied. A copy-specific qPCR assay to selectively amplify HERV-W 9q22.31 was designed. Due to the high identity shared among HERV-W elements, one primer was designed to match outside of the HERV-W-part in order to increase PCR specificity (Figure 7A). This assay, along with the already optimised



	N	Median (IR)	P- value (t student)
Controls	5	16 (5.5)	p=0.018
CIS	5	31(13)	

Locus	Relative frequency (%) Median (IR)		P- value (t-student)
	Controls	CIS	
3q11.2	0.04 (0.06)	0.7 (0.85)	p= 0.05
5p12	0.17 (0.15)	1.15 (0.81)	p= 0.017
9q22.31	0.17 (0.21)	1.87 (1.34)	p= 0.022

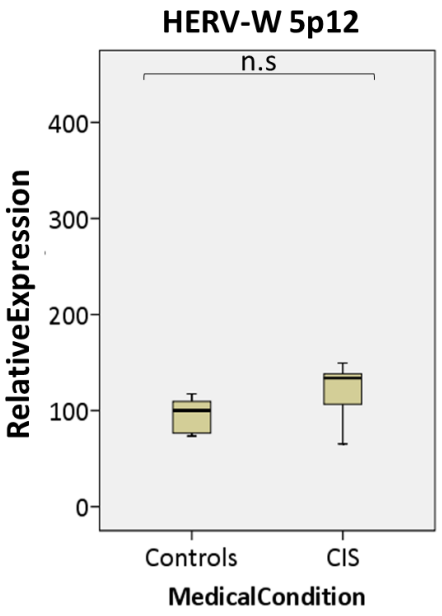
Figure 6. NGS analysis of transcribed HERV-W *ENV* copies in PBMCs. A specific MSRV *ENV* PCR described by Mameli G et al., 2009 was applied to samples taken randomly from the samples analyzed previously (Figure 3) and sequenced by NGS. Relative frequency (%) of individual HERV-W *ENV* copies was calculated as the number of reads from individual HERV-W *ENV loci* relative to the total number of reads analysed). A) HERV-W *ENV* copies were identified in PBMC of 5 CIS patients and 5 controls. Numerical values taken from the graph are indicated. B) Representation of chromosomal location and relative frequency (%) for the most abundant HERV-W copies in CIS patients and controls. C) Relative

frequency (%) in CIS patients or controls, of reads mapped to the individual HERV-W *ENV* copies indicated. Tables show numerical values taken from the graph.

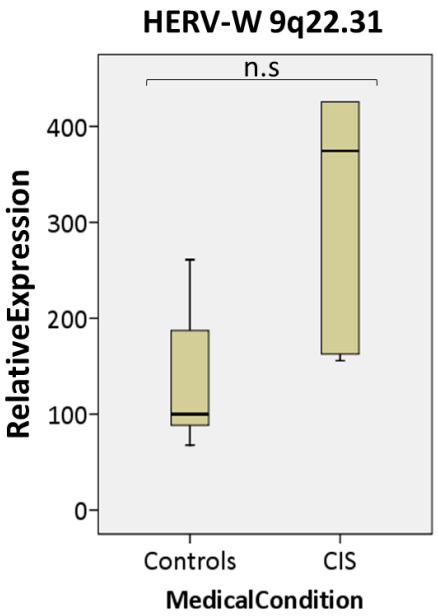
To verify the overexpression of CIS-related HERV-W copies in CIS patients, specific qPCR assays were designed and applied. A copy-specific qPCR assay to selectively amplify HERV-W 9q22.31 was designed. Due to the high identity shared among HERV-W elements, one primer was designed to match outside of the HERV-W-part in order to increase PCR specificity (Figure 7A). This assay, along with the already optimised

A

B



C



Locus	Relative expression (%) Median (IR)		p- value (t-student)
	Controls (N=6)	CIS patients (N=6)	
5p12	100 (35.74)	133.94 (44.72)	p= 0.117
9q22.31	100 (146.12)	374.51 (663.3)	p= 0.139

Figure 7. qPCR specific expression analysis for HERV-W 5p12 and 9q22.31 in PBMCs. Schematical representation of the new copy-specific expression assay designed is illustrated (A). Arrows represent the pair of primers used in the qPCR to specifically amplify the element HERV-W 9q22.31. LTR: long terminal repeat. Random primer-primed cDNA from 6 CIS patients and 6 controls was analyzed with the copy-specific expression analysis for HERV-W *ENV* 5p12 (B) and HERV-W 9q22.31 (C). Samples were taken randomly from the samples analyzed previously (Figure 6). Results were normalized using *GAPDH*

as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. Table shows numerical values taken from the graph.

copy-specific qPCR assay for HERV-W *ENV* 5p12 (Figure 5A), were applied in PBMCs from 6 CIS patients and 6 controls. HERV-W *ENV* 5p12 relative expression was higher in CIS patients (133.94% ± 44.72) compared to controls (100% ± 35.74) (t-student; p= 0.117) (Figure 7B). Likewise, HERV-W 9q22.31 relative expression levels were increased in CIS patients (374.51% ± 663.3) compared to controls (100% ± 146.12) (t-student; p= 0.139) (Figure 7C). This time, both copy-specific qPCR analyses detected and quantified HERV-W overexpression in CIS patients. Although not statistically significant, expression levels for both HERV-W *loci* were increased in CIS patients. Data, therefore, are evidence that CIS patients present elevated expression of certain specific MS-related HERV-W copies such as HERV-W *loci* 5p12 and 9q22.31.

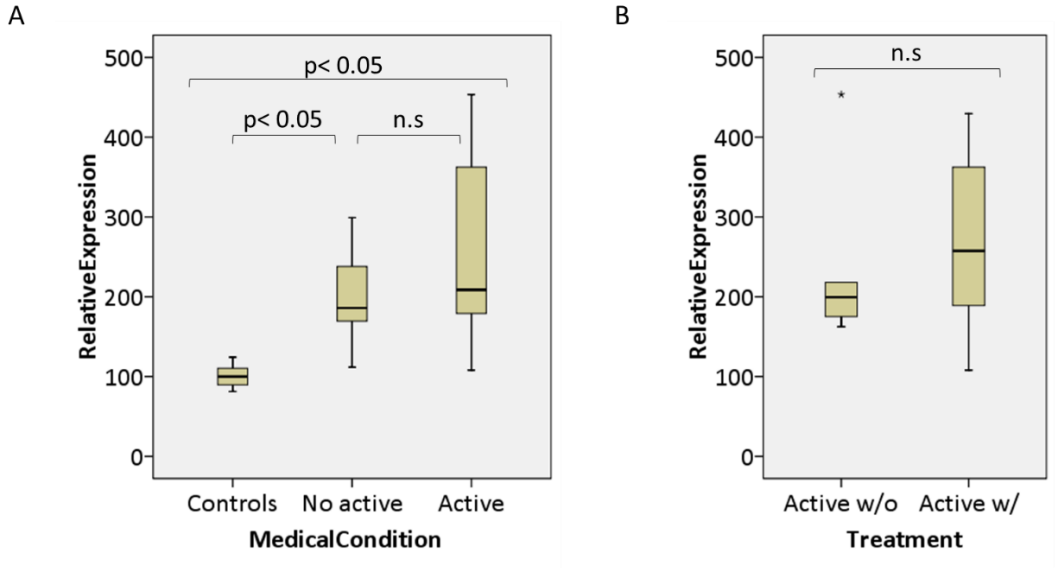
Is HERV-W overexpression associated with active disease?

Initial results pointed out that HERV-W *ENV* overexpression was specifically found in CIS patients. Due to the close association between CIS patients and the presence of clinical activity, we intended to distinguish whether such retroviral overexpression could be linked to active disease rather than being exclusively CIS-associated. We gathered a new collection of PBMCs samples, taken from MS patients either presenting (n=12) or not presenting (n=12) clinical activity (Table 2). Within the group of MS patients, some were under treatment (n=7) and others were not undergoing treatment (n=5). HERV-W *ENV* expression was analysed using the same expression assay mentioned above, and results were compared to the median expression of the control group defined within the initial cohort (Figure 3). Both, MS patients with and without clinical activity showed a significant increase of HERV-W *ENV* expression compared to controls (t-student; p= 0 and p= 0.019 respectively) (Figure 8A). And although not statistically significant, HERV-W *ENV* expression was indeed slightly higher in active MS patients (208.75% ± 199.5) than in non-active MS patients (185.95% ± 77.03), compared to controls (100% ± 28.51) (t-student; p= 0.158) (Figure 8A). On the other hand, HERV-W *ENV* expression was not significantly influenced in active MS patients receiving treatment compared to those without treatment (U de Mann-Whitney; p= 0.685) (Figure 8B). In this new cohort, HERV-W *ENV* expression was

therefore not only associated with MS, but specially associated with the presence of clinical activity. These results strengthen the plausible linkage between HERV-W *ENV* expression and MS and indicate a possible correlation between such overexpression and active stages of the disease.

COHORT 2	
	MS patients (n=24)
With clinical activity	n=7
Treated	n= 5
Non-treated	
Without clinical activity	n=12

Table 2. Clinical and treatment data of cohort 2. A new cohort of 24 MS patients were recruited either presenting or not presenting clinical activity. Within the group of MS patients with clinical activity, some were under treatment and others were not undergoing treatment.



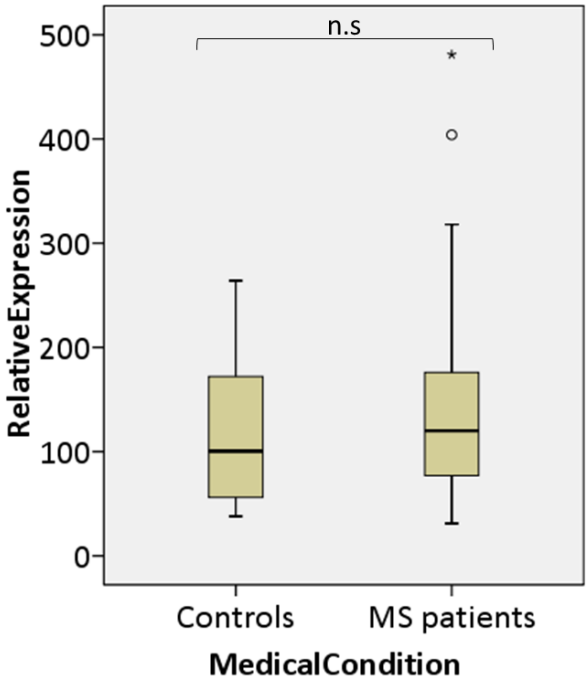
	N	Median (IR)	P- value (t-student)
Controls	9	100 (28.51)	p=0.158 (No active-Active)
No active	12	185.95 (77.03)	p= 0.019 (No active- Controls)
Active	12	208.75 (199.5)	p=0.000 (Active-Controls)

	N	Median (IR)	P- value (U de Mann-Whitney)
Active w/o	5	199.4 (166.95)	p=0.685
Active w/	7	257.5 (207.8)	

Figure 8. Human endogenous retrovirus-W *ENV* expression. A) Analysis was performed using cDNA primed with random primer from 24 MS patients, of which 12 were non-active MS patients and 12 were active MS patients. Control group was the same used in figure 3. Results were normalized using *GAPDH* as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. B) Results were also stratified according to whether MS active patients were undergoing treatment (w/) or not (w/o). Table below shows numerical values taken from the graph.

Human endogenous retrovirus-W ENV expression in PBMCs of the larger cohort

After confirming HERV-W *ENV* overexpression in two different small cohorts (Figures 3A and 8A), we analysed HERV-W *ENV* expression in the larger collection of samples from which the initial small group of samples was taken (Figure 3). The collection was composed of samples of MS patients with different clinical status (CIS, PPMS, RRMS and SPMS) (n=79) and controls (n=34). HERV-W *ENV* expression was analysed using random-primed cDNA and the specific PCR assay described by Mameli et al., 2009. Results were calculated and normalized as already described. HERV-W *ENV* expression turned out to be quite similar in MS patients (120% ± 101) and control (100% ± 117.75) treatment (U de Mann-Whitney; p= 0.272) (Figure 9). Although these results did not



	N	Median (IR)	P- value (U de Mann-Whitney)
Controls	34	100 (117.75)	P= 0.272
MS patients	79	120 (101)	

Figure 9. Human endogenous retrovirus-W *ENV* expression in PBMCs. qPCR analysis was performed using cDNA primed with random primer from 79 MS patients and 34 controls. Results were normalized using *GAPDH* as a reference gene, calculated using the $2^{-\Delta\Delta Ct}$ method, and represented as a percentage of the fold expression compared to the median expression level in controls. Tables below show numerical values taken from the graph.

completely correlate with the previous ones (Figure 3), they did show the general tendency for HERV-W *ENV* expression to be increased in MS patients (120% versus 100% in controls). Since our findings indicated that MS-associated HERV-W elements are not very highly expressed (Figure 6C), it is not surprising the difference in HERV-W *ENV* expression between patients and control is not very big.

Discussion

All assays discussed here were done on PBMC unless indicated otherwise. An discriminatory real-time PCR assay was developed to selectively amplify MSRV (Mameli et al., 2009). As HERV-W *ENV* sequences are related to MSRV, independent of the exact relationship between HERV-W and MSRV, I have used the term HERV-W to refer to these sequences. And all assays discussed here were done on PBMC unless indicated otherwise. Using this assay, higher HERV-W copy numbers have been detected in MS patients (Garcia-Montojo et al., 2013; Mameli et al., 2009; Perron et al., 2012). However, here we did not find differences in the DNA load of HERV-W elements when applying this same assay in a cohort of 116 samples (Figure 1A). We suggest that the difference between individuals, populations and cohorts may reside in genetic variability. In fact, copy number may vary due to the presence of unfixed HERV-W copies derived from novel LINE-1-mediated retrotransposition events in modern populations (Costas, 2002; Grandi and Tramontano, 2017; Pavlícek et al., 2002). Therefore, the copy number of HERV-W elements may differ between cohorts. As opposed to published studies (Perron et al., 2012), we did not observe any differences in the DNA copy number of HERV-W elements when MS samples were stratified according to their clinical status (Figure 1B). Gender-specific differences in DNA copy number of HERV-W elements have also been reported (Garcia-Montojo et al., 2013). Although we would like to analyse our data likewise, this data is unavailable as of yet for our cohort.

The assay described by Mameli et al. that discriminates between MSRV and *SYNCYTIN-1*, has been used to detect upregulated HERV-W *ENV* expression in MS (Mameli et al., 2009; Perron et al., 2012). We re-evaluated this assay, comparing cDNA priming procedures and securing calibration curves for low expression samples (Appendix I). After these verifications, we present a robust and reproducible assay to measure

expression of MS-associated copies of HERV-W (Figure 2). We found in a small cohort (n=12) that non-active MS patients did show increased levels of HERV-W *ENV* expression (185.95% versus 100% in controls; p= 0.019) (Figure 8A). By contrast, analysis of a different, larger cohort showed a weak tendency for HERV-W *ENV* expression to be increased in MS patients (120% versus 100% in controls; p= 0.272) (Figure 9). This supports the hypothesis that HERV-W *ENV* overexpression somehow contributes to MS. The high variability in the larger cohort suggests that stratification based on clinical activity or disease pattern, may yield better information. Unfortunately, we do not have access to the information on the clinical activity of MS patients recorded at the time of sample taking, precluding analysis. To better evaluate this, cohorts should be enlarged, and data stratified according to clinical history, clinical status and type of treatment.

Expression analyses using the optimized assay described by Mameli et al. (Mameli et al., 2009) in a first cohort of MS patients demonstrated upregulated HERV-W *ENV* expression in CIS patients (p= 0.006) (Figure 3B). Several considerations suggest that this overexpression may be associated with active disease in general, and not to the clinical status CIS itself. Of course, CIS patients are defined as active patients showing a first disease outbreak. When copy-specific expression assays were designed to detect HERV-W 5p12, increased expression was evident in CIS patients (Figure 7B), as opposed to a mixed group of patients (Figure 5B). To further investigate the hypothesis that correlation exists between the presence of clinical activity in MS patients and higher HERV-W *ENV* expression levels, a new set of samples was collected. Samples were separated in groups according the presence or absence of clinical activity. Both subgroups showed upregulated HERV-W *ENV* expression (Figure 8A). Stratification by clinical activity revealed that active MS patients displayed a slightly higher HERV-W *ENV* expression levels compared to non-active MS patients (208.75% versus 185.85%; p= 0.158). Pending confirmation, this finding may suggest that, independent of the MS subtype, HERV-W *ENV* expression is linked to the presence of clinical activity.

Previous studies failed to identify MS-specific *loci* or expression (Laufer et al., 2009; Schmitt et al., 2013). We succeeded in identifying which overexpressed HERV-W sequences are detected by the MSRV-specific qPCR assay of Mameli et al. Although

initial analysis suggested HERV-W *loci* 5p12 and 19q13.2 as possible MS-associated *loci* (Figure 4), copy-specific qPCR assays developed for these *loci* failed to detect increased expression levels in MS patients. To increase the scope of the search, cDNA sequences were sequenced by NGS. Posterior analysis revealed that most reads map to a selected set of 6 HERV-W *loci* with a similar relative frequency in MS patients and controls (Figure 6B). However, the relative frequency of reads mapping to HERV-W *loci* 5p12, 3q11.2 and 9q22.31 was much higher in CIS patients ($p< 0.05$) (Figure 6C). Moreover, copy-specific qPCR assays for HERV-W *loci* 5p12 and 9q22.3 confirmed increased expression in CIS patients (Figure 7). As most reads obtained from both CIS patients and controls mapped to a limited set of apparently highly transcribed copies, our data indicate that the observed HERV-W *ENV* overexpression in CIS patients (Figure 3B) did not originate from a single HERV-W *ENV* copy. Such overexpression appears to be derived from less abundantly transcribed but CIS-related HERV-W *ENV loci.* In the light of these finding, the hypothesis that a given HERV-W element could uniquely provoke MS seems highly unlikely since the most abundant HERV-W elements are apparently the same in all individuals. However, MS appears associated with the upregulation of specific HERV-W *loci*. This finding now awaits confirmation in a larger cohort.

Finally, it deserves to be mentioned that since the specific assay used throughout our studies (Mameli et al., 2009) only amplifies a limited number of MSRV-like HERV-W elements, our study may miss other HERV-W elements. Moreover, transcripts of other HERV families (such as HERV-K and HERV-H) have been occasionally associated with MS (Morandi et al., 2017). Overexpression of MSRV/HERV-W has been associated with MS in a variety of studies (Dolei et al., 2002; Mameli et al., 2007, 2009; Perron et al., 2012). We studied an additional cohort and found a weak overexpression (Figure 9), which did not reach statistical significance. It should be noted, that other study also reported a lack of overexpression in MS brain (Schmitt et al., 2013). Our detailed studies aimed at which HERV-W may be overexpressed in CIS patients, have yielded two separate findings (Figure 6). First of all, after sequencing of cDNA prepared from overexpressed HERV-W, reads from CIS patients mapped to almost double the number of different HERV-W *loci* compared to reads obtained from controls ($p=0.018$) (Figure 6A). We conclude that more HERV-W *loci* are expressed in CIS patients than in control

subjects. This finding may be more relevant than overexpression per se. Similar findings have been reported previously (Schmitt et al., 2013), who analysed MS brain. In addition, we show that the HERV-W copies relatively overexpressed in CIS patients do not produce high levels of transcripts and only represent 1-3% of total transcripts (Figure 6C). The combined findings on low levels of overexpression, activation of more *loci,* and activation of low-expressing HERV-W elements in MS suggest that their potential contribution to the pathology may not reside simply in high expression levels. CIS-associated copies may produce proteins that are especially active in TLR4 activation (Rolland et al., 2006). Alternatively, HERV-W products in the form of RNA may activate the native immune system through TLR3 (Hurst and Magiorkinis, 2015; Stetson, 2012). It remains perfectly possible, that additional HERV-W copies or the CIS-associated HERV-W copies (Figure 6), produce or overproduce LTR3-activating agents functioning as Pathogen Associated Molecular Pattern. We therefore highlight the potential contribution to MS of TLR3 activation through RNAs that originate from HERV-W, and stress the importance of searching for such substances in MS patients.

Supplementary materials

Target	Sequence (5'-3')	1 mM $MgCl_2$ addition
MSRV *ENV* (Mameli et al., 2009)	For: CTTCCAGAATTGAAGCTGTAAAGC Rev: GGGTTGTGCAGTTGAGATTTCC TaqMan® probe: FAM-5'-TTCTTCAAATGGAGCCCCAGATGCAG-3'-TAMRA	No
hGAPDH	For: ATCAGCAATGCCTCCTGCAC Rev: TGGCATGGACTGTGGTCATG	No
HERV-W 19q13.2	For: GCAGGATTTCCTAGGCCGATT Rev: CTTACCTGATGGGTCGGGTG	Yes
HERV-W 5p12	For: CCTGCTAGACTATGCTCTGATG Rev: CTGCTCTAACTGCTTCCTACTG	Yes
HERV-W 9q22.31	For: GCTTGCAACTTAGCCCACAC Rev: AAACAAAACCTTACCATCTTAACCA	Yes

Table s1. List of set of primers used for the qPCR expression analysis. The first column shows the target, the second column indicates the oligo sequence (For= forward; Rev= reverse) and, the third column points out whether the PCR reaction needs an extra addition of 1 mM $MgCl_2$ or not.

References

Blond, J.L., Besème, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet, F. (1999). Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. J. Virol. *73*, 1175–1185.

Blond, J.L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F.L. (2000). An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J. Virol. *74*, 3321–3329.

Cho, K., Lee, Y.-K., and Greenhalgh, D.G. (2008). Endogenous retroviruses in systemic response to stress signals. Shock Augusta Ga *30*, 105–116.

Compston, A., and Coles, A. (2002). Multiple sclerosis. The Lancet *359*, 1221–1231.

Compston, A., and Coles, A. (2008). Multiple sclerosis. The Lancet *372*, 1502–1517.

Costas, J. (2002). Characterization of the intragenomic spread of the human endogenous retrovirus family HERV-W. Mol. Biol. Evol. *19*, 526–533.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., Cirotto, M.C., Rosati, G., and Sotgiu, S. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology *58*, 471–473.

Garcia-Montojo, M., Dominguez-Mozo, M., Arias-Leal, A., Garcia-Martinez, Á., De las Heras, V., Casanova, I., Faucard, R., Gehin, N., Madeira, A., Arroyo, R., et al. (2013). 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*, e53623.

García-Montojo, M., de la Hera, B., Varadé, J., de la Encarnación, A., Camacho, I., Domínguez-Mozo, M., Arias-Leal, A., García-Martínez, Á., Casanova, I., Izquierdo, G., et al. (2014). HERV-W polymorphism in chromosome X is associated with multiple sclerosis risk and with differential expression of MSRV. Retrovirology *11*, 2.

Grandi, N., and Tramontano, E. (2017). Type W Human Endogenous Retrovirus (HERV-W) Integrations and Their Mobilization by L1 Machinery: Contribution to the Human Transcriptome and Impact on the Host Physiopathology. Viruses *9*.

Hurst, T.P., and Magiorkinis, G. (2015). Activation of the innate immune response by endogenous retroviruses. J. Gen. Virol. *96*, 1207–1218.

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E., Mallet, F., Mandrand, B., et al. (1999). Molecular Cloning and Characterization of MSRV-Related Sequences Associated with Retrovirus-like Particles. Virology *260*, 1–9.

Laufer, G., Mayer, J., Mueller, B.F., Mueller-Lantzsch, N., and Ruprecht, K. (2009). Analysis of transcribed human endogenous retrovirus W env loci clarifies the origin of multiple sclerosis-associated retrovirus env sequences. Retrovirology *6*, 37.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif *25*, 402–408.

Löwer, R., Löwer, J., and Kurth, R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. U. S. A. *93*, 5177–5184.

Lublin, F.D., Reingold, S.C., and National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis* (1996). Defining the clinical course of multiple sclerosis: Results of an international survey. Neurology *46*, 907–911.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., Serra, C., Sotgiu, S., Bonetti, B., and Dolei, A. (2007). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J. Gen. Virol. *88*, 264–274.

Mameli, G., Poddighe, L., Astone, V., Delogu, G., Arru, G., Sotgiu, S., Serra, C., and Dolei, A. (2009). 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*, 98–106.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785–789.

Miller, D., Barkhof, F., Montalban, X., Thompson, A., and Filippi, M. (2005). Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. Lancet Neurol. *4*, 281–288.

Morandi, E., Tanasescu, R., Tarlinton, R.E., Constantinescu, C.S., Zhang, W., Tench, C., and Gran, B. (2017). The association between human endogenous retroviruses and multiple sclerosis: A systematic review and meta-analysis. PLoS ONE *12*.

Pavlícek, A., Paces, J., Elleder, D., and Hejnar, J. (2002). Processed pseudogenes of human endogenous retroviruses generated by LINEs: their integration, stability, and distribution. Genome Res. *12*, 391–399.

Pérez-Palacios, R., Macías-Redondo, S., Climent, M., Contreras-Moreira, B., Muniesa, P., and Schoorlemmer, J. (2016). In Vivo Chromatin Targets of the Transcription Factor Yin Yang 2 in Trophoblast Stem Cells. PLOS ONE *11*, e0154268.

Perron, H., Geny, C., Laurent, A., Mouriquand, C., Pellat, J., Perret, J., and Seigneurin, J.M. (1989). Leptomeningeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. Res. Virol. *140*, 551–561.

Perron, H., Lalande, B., Gratacap, B., Laurent, A., Genoulaz, O., Geny, C., Mallaret, M., Schuller, E., Stoebner, P., and Seigneurin, J.M. (1991). Isolation of retrovirus from patients with multiple sclerosis. Lancet Lond. Engl. *337*, 862–863.

Perron, H., Garson, J.A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P.W., Voisset, C., Blond, J.L., et al. (1997). Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc. Natl. Acad. Sci. U. S. A. *94*, 7583–7588.

Perron, H., Perin, J.P., Rieger, F., and Alliel, P.M. (2000). Particle-associated retroviral RNA and tandem RGH/HERV-W copies on human chromosome 7q: possible components of a “chain-reaction” triggered by infectious agents in multiple sclerosis? J. Neurovirol. *6 Suppl 2*, S67-75.

Perron, H., Germi, R., Bernard, C., Garcia-Montojo, M., Deluen, C., Farinelli, L., Faucard, R., Veas, F., Stefas, I., Fabriek, B.O., et al. (2012). Human endogenous retrovirus type W envelope

expression in blood and brain cells provides new insights into multiple sclerosis disease. Mult. Scler. J. *18*, 1721–1736.

Rolland, A., Jouvin-Marche, E., Viret, C., Faure, M., Perron, H., and Marche, P.N. (2006). The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. J. Immunol. Baltim. Md 1950 *176*, 7636–7644.

Schmitt, K., Richter, C., Backes, C., Meese, E., Ruprecht, K., and Mayer, J. (2013). Comprehensive Analysis of Human Endogenous Retrovirus Group HERV-W Locus Transcription in Multiple Sclerosis Brain Lesions by High-Throughput Amplicon Sequencing. J. Virol. *87*, 13837–13852.

Serra, C., Sotgiu, S., Mameli, G., Pugliatti, M., Rosati, G., and Dolei, A. (2001). Multiple sclerosis and multiple sclerosis-associated retrovirus in Sardinia. Neurol. Sci. Off. J. Ital. Neurol. Soc. Ital. Soc. Clin. Neurophysiol. *22*, 171–173.

Stetson, D.B. (2012). Endogenous retroelements and autoimmune disease. Curr. Opin. Immunol. *24*, 692–697.

Manuscript III. HERV-W methylation in multiple sclerosis

Abstract

The presence and expression of human endogenous retroviruses family W (HERV-W) has been repeatedly associated with multiple sclerosis (MS). However, little is known of the regulation of HERV-W expression and the pathological significance of overexpression. A number of studies have reported that DNA hypomethylation correlates with HERV activation in a range of diseases. Nevertheless, the possible role of DNA hypomethylation in the induction of HERV-W expression in MS has not yet been directly addressed. HERV-W *ENV* copy-specific methylation assays were designed for HERV-W *ENV loci* 19q13.2, 15q21.3 and Xq22.3. Assays span 100 bp amplicons and allow assessment of methylation level of individual CpG sites by pyrosequencing. After validation, they were applied to analyse methylation levels in genomic DNA from peripherical blood mononuclear cells (PBMCs) of 5 MS patients and 6 controls. Methylation patterns in the *loci* analysed did not differ between MS patients and controls. In all cases, DNA sequences were highly methylated, and no MS-specific changes were observed. Therefore, 5-aza-2-deoxycytidine (5-azadC) was used as a demethylating agent in an attempt to force HERV-W demethylation and address the relation to expression levels. Neither in oligodendrocyte precursor cells (OPCs), nor in lymphocytes 5-azadC treatment was effective and demethylation could not be demonstrated. Human embryonic stem cells (hESCs) were more susceptible to 5-azadC, and treatment caused demethylation of *POU5F1*, *LEFTY1* and *SERPINA3* genes, However, DNA methylation of the HERV-W elements was not altered. It seems that DNA methylation levels of HERV-W elements are ordinarily high in PBMCs and do not correlate with MS. Efforts to artificially force HERV-W demethylation using 5-azadC, demonstrated that the high methylation levels at HERV-W elements maybe unusually stable and resistant to demethylation.

Introduction

Human endogenous retroviruses (HERVs) are permanent retroviral integration which comprise approximately 8% of our genome (Cho et al., 2008; Löwer et al., 1996). HERVs can be divided into multicopy families, consisting of closely related but distinguishable elements dispersed in the human genome. Typical structure of a HERV provirus contains an internal coding region containing *GAG, POL,* and *ENV* genes, flanked by two long terminal repeats (LTR). HERVs are present within full length or incomplete sequences and, most of them have lost their protein-coding capacity by an accumulation of mutations (Kim, 2012). However, an increasing number of studies have reported HERV expression in several tissues and cells (Balestrieri et al., 2015; Seifarth et al., 2005).

HERVs have been found to be active in human cells in a tissue-specific manner, with a considerable expression in skin, thyroid gland, placenta, and tissues of reproductive organs (Seifarth et al., 2005). Control of HERVs expression depends on regulation at the level of the LTRs, which contain multiple binding sites for transcription factors and act as promoters (Kovalskaya et al., 2006). It widely assumed that transcriptional activity of HERVs is regulated by epigenetic mechanisms, mainly by DNA methylation (Hurst and Magiorkinis, 2017). In fact, tissue-specific epigenetic modifications in the LTR promoter region have been reported to be associated with HERVs expression. For instance, promoter of HERV elements expressed in placenta tissue, have been reported to be unmethylated in placenta but heavily methylated in other cell types, where these HERV elements do not express (Matoušková et al., 2006; Reiss et al., 2007).

Upregulated HERV expression has repeatedly been associated to a range of diseases such as cancer and neurodegenerative diseases (Colmegna and Garry, 2006; Dolei, 2006; Kurth and Bannert, 2010). And such aberrant HERVs expression has also been linked to altered DNA methylation. DNA hypomethylation has been observed to correlate with HERVs activation in cancer cells and in CD4+ T cells from lupus patients (Goering et al., 2011; Wu et al., 2015). In particular, demethylation of HERV-K LTR has been reported to induce HERV-K expression in cancer cells (Stengel et al., 2010). However, although HERV family W (HERV-W) presence and expression has consistently

been associated to multiple sclerosis (MS) (Dolei et al., 2002; Garson et al., 1998; Mameli et al., 2007; Perron et al., 2012), HERV-W methylation status in the context of MS has not yet been addressed. In a previous study (Manuscript II), we identified HERV-W *loci* that were overexpressed in peripherical blood mononuclear cells (PBMCs) of MS patients. We now intended to address the relationship between genomic methylation of HERV-W *loci* and induced expression. Based on that initial study, copy-specific methylation assays for specific HERV-W *loci* were designed, developed, validated and applied to PBMC samples taken from MS patients and controls. As no MS-related changes in DNA methylation were observed, we resorted to an alternative strategy to analyse whether HERV-W hypomethylation correlates with HERV-W activation. Strategies to study epigenetics include the use of drugs which inhibit DNA methyltransferases (DNMTs), resulting in demethylation and reactivation of silenced genes (Chistiakov et al., 2017; Palii et al., 2008; Ramos et al., 2015). Different cell types were treated with the DNA methyltransferase inhibitor 5-aza-2′-deoxycytidine (5-azadC), and the effect on HERV-W methylation status was assessed.

Materials and methods

PBMCs isolation and gDNA extraction

PBMCs isolation and gDNA extraction was performed as described before (M&M, Manuscript I).

Sodium bisulfite modification

Genomic DNA samples (500 ng) were bisulfite modified using the EZ DNA Methylation™ Kit (D5001, Zymo Research) following the manufacturer's instructions. The converted DNA was eluted in 20 µl of TE Buffer and stored at -20 °C until used.

PCR amplification and methylation assays

Bisulfite-converted genomic DNA was amplified by PCR or nested PCR depending on the *locus* analyzed. For the first round of amplification, PCR reaction was carried out in a 50-µl volume using 4 µl of bisulfite-converted DNA (previously boilt at 100 °C for 3 minutes), 200 µM dNTPs, forward and reverse primers (400 nM each), and 0.5 U Taq DNA Polymerase (D1806, Sigma-Aldrich) in PCR reaction buffer containing 2.5 mM

$MgCl_2$. As an exception, the PCR reaction buffer used in the first PCR round of HERV-W *ENV locus* 5p12 only contained 1.5 mM MgCl2). Components of nested PCR were the same as listed above, except for the template which consisted of 1 µl of the first PCR product. Sequences of all primers used and cycling conditions for all assays are described in Table s1 and Table s2. All set of primers were in-house designed, except for those primers used in the pyrosequencing assays for *POU5F1*, *LEFTY1* and *SERPINA3* (Kim et al., 2011). PCR amplifications were performed in duplicate and PCR products from duplicate reactions were mixed before analyzed. The presence of amplified products was confirmed on a 2% agarose gel. Bisulfite-converted PCR products were sequenced using PSQ 96MA System (Pyrosequencing). Sequencing primer, the nucleotide dispensation order and the number of CpGs sites analysed in all pyrosequencing assays are described in Table s1 and Table s2. Some negative controls were included in the sequence to assess the absence of undesired *PCR products (Figure s1).* Technical replicates using different amount of DNA input were performed to exclude variability and validate every pyrosequencing assay. Percentage of methylation for each CpG site was calculated using the Pyro Q-CpG software (version 1.0.9) (Biotage) applying the default settings. Only high-quality quantitative results, determined by the quality control thresholds of the software, were considered. Examples of the pyrograms obtained from assays described in Table s1 and Table s2 can be observed in Figure s1 and Figure s2, respectively.

Oligodendrocyte precursor cells culture and 5-aza-2'-deoxycytidine treatment

Oligodendrocyte precursor cells (OPCs) characterized by expression of *NG2*, were differentiated from hESCs ES[4] as described in Manuscript IV (All et al., 2015). OPCs were treated with 100nM, 250nM, 500nM and 1µM 5-aza-2'-deoxycytidine (5-aza-dC) or untreated (mock) for 96 hours (drug-supplemented culture medium was renovated daily). After 96 hours, cells were harvested and gDNA was isolated as above.

Lymphocytes culture, cell viability assay and 5-aza-2'-deoxycytidine treatment

PBMCs were isolated from buffy coats (obtained by the Aragón Blood and Tissue Bank). PBMC were plated at a density of 1 x 10^6 cells/mL and cultured in RPMI-1640 medium, supplemented with 1X glutamax, 10% heat-inactivated FBS and 10 ng/ml LPS

(Sigma). After 24 hours in culture, non-adherent cells (lymphocytes, both B and T cells) were centrifuged and plated aside the adherent cells (monocytes), in the same conditions used so far. The following day, lymphocytes were treated with 100nM, 1µM, 3µM and 10 µM 5-aza-dC or untreated (mock) for 96 hours (drug-supplemented culture medium was renovated daily). After 96 hours, cells were harvested and gDNA was isolated.

Human embryonic stem cell line ES[4] cell culture, cell viability assay and 5-aza-2'-deoxycytidine treatment

The human embryonic stem cell line ES[4] was provided by the Stem Cell Bank of Barcelona (BLCB) and propagated on Matrigel matrix (BD) in mTeSR1 medium (STEMCELL Technologies) as described by BLCB. ES[4] cells were passaged when they were approximately 80-85% confluent using ReLeSR™ (STEMCELL Technologies) according to manufacturer's' protocol at a 1:8-1:10 split ratio, ensuring this way that cells were passaged every 3-4 days. To test for dose-dependent toxicity of 5-aza-2′-deoxycytidine, ES[4] cells were plated, allowed to grow for 48 hours, and treated with 0.1 nM, 1 nM, 10 nM, 100nM, 1µM and 5µM 5-aza-2′-deoxycytidine (5-aza-dC) or untreated (as control) for 96 hours (fresh culture medium and drug was added every 24 hours). During treatment, cell viability was observed under the microscope and defined based on cell adherence and morphology. After treatment, cells were harvested and gDNA was isolated

Statistical analysis

Statistical analysis was performed as described before (M&M, Manuscript II).

Results

Methylation level of HERV-W-ENV-specific loci in PBMCs

To analyse the potential changes methylation status related to MS, we decided to assess methylation levels of HERV-W *ENV* in PBMCs of MS patients. We took advantage of a previous work in which we defined HERV-W *ENV* copies overexpressed in PBMCs of a set of MS patients (Manuscript II). We now designed copy-specific assays for some of the most representative HERV-W *ENV* copies. After validation, these assays

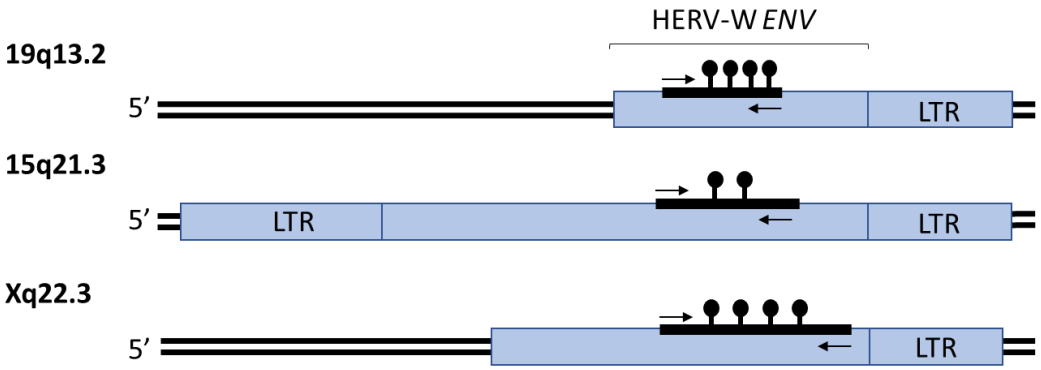
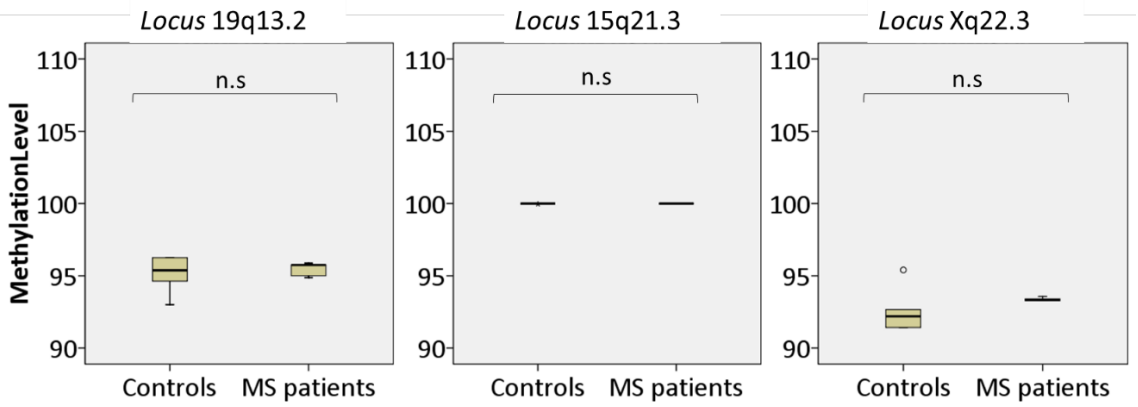


Figure 1. Schematic representation of methylation assays. Methylation assays for *loci* HERV-W *ENV* indicated on the left (19q13.2, 15q21.3 and Xq22.3). Arrows represent the pair of primers used in PCR to selectively amplify each HERV-W element. Biotinylated and sequence primers are indicated in Table s1. The round symbols ( ) indicate individual CpG sites within the amplicons analyzed. It is not drawn to scale. LTR: long terminal repeats.



Locus	Methylation levels (%) Median (IR)		p- value
	Controls (n=6)	MS patients (n=5)	
19q13.2	95.38 (2.03)	95.75 (0.88)	p= 0.639 (t-student)
15q21.3	100 (0.01)	100 (6.24)	p= 0.787 (U de Mann-Whitney)
Xq22.3	92.2 (3.27)	93.33 (2.71)	p= 0.273 (U de Mann-Whitney)

Figure 2. DNA methylation levels of HERV-W *loci* in PBMCs. Pyrosequencing-based methylation analysis was performed on genomic DNA from PBMCs of MS patients (n=5) and control individuals (n=6). Results are represented as de median of the methylation levels across all CpG sites (Figure 1) analysed. The table shows numerical values taken from the graph.

were applied in a set of PBMCs from MS patients (n=5) and controls (N=6). MSRV/HERV-W *ENV* expression had previously been observed to be upregulated in those MS patients (data not shown). Assays were based on the resistance of methylated cytosine to conversion into uracil by bisulfite treatment (Frommer et al., 1992). Bisulfite-converted DNA was used for the amplification of the region of interest followed by pyrosequencing. Different set of primers were designed around certain CpG sites and used to selectively amplify HERV-W *ENV* loci 19q13.2, 15q21.3 and Xq22.3 (Figure 1). The resulting PCR products were analysed by pyrosequencing and, the methylation status of individual CpG sites within each amplicon was determined. Examples of the pyrograms obtained from the analyses are shown in Figure s1. Amplicon of HERV-W *ENV* 19q13.2 was 95.38% and 97.75% methylated in MS patients and controls, respectively (t-student; p= 0.639) (Figure 2). Amplicon of HERV-W *ENV* 15q21.3 was 100% methylated in both MS patients and controls (U. de Mann Whitney; p= 0.787). And amplicon of HERV-W *ENV* Xq22.3 was 92.3% and 93.33% methylated in MS patients and controls, respectively (U. de Mann Whitney; p= 0.273). Therefore, we did not detect DNA methylation changes related to MS in the *loci* studied.

Effect of 5-azadC on DNA methylation in oligodendrocyte precursor cells and lymphocytes

As we failed to detect alterations in genomic HERV-W methylation, we wanted to address whether there is a correlation between changes in HERV-W methylation levels and the expression levels. As immune responses typical of MS attack oligodendrocytes, we analysed 5-azadC-mediated demethylation in oligodendrocyte precursor cells (OPCs). 5-azadC treatment inhibits methylation of the newly formed DNA strand formed upon DNA synthesis, and therefore depends on cell proliferation. First, we examined the effect of 5-azadC on cell viability. OPCs were treated with increasing concentrations of 5-azadC (100 nM to 1µM) for 96 hours. During treatment, cell viability was observed under the microscope and defined based on cell adherence and morphology. We found that cell viability was not affected by any of the treatments used (data not shown). Therefore, the conditions used in following assays was the

maximum tested exposure to the drug, which consisted of an exposure of 1µM 5-azadC for 96 hours.

To test the effect of 5-azadC on DNA demethylation, methylation levels of an amplicon within HERV-W 19q13.2 was assessed. This HERV-W copy was chosen since it was abundantly transcribed in MS patients (Manuscript II) and we had already optimized a pyrosequencing assay for this *locus* (Figure 1). *LRRC27* was included as a control gene since it is expressed in brain tissues and its expression is regulated by DNA methylation (Huynh et al., 2014). First, a pyrosequencing assay was optimized so that methylation levels of several CpGs within the *LRRC27 locus* could be analysed (Figure 3). Examples of the pyrogram obtained from the assay is shown in Figure s2. The resulting assay was applied to 5-azadC treated-OPCs. We found that the average methylation level of the amplicon of HERV-W 19q13.2 and the *LRRC27* gene did not vary between untreated and treated OPCs (Figure 4A). While *LRRC27* methylation amounted to 47% and 49%, respectively, the levels of methylation level in the amplicon of HERV-W 19q13.2 measured at 94 % and 90%, respectively (Figure 4A). Therefore, while we applied a prolonged treatment with a relatively high concentration of 5-azadC, no changes in the methylation of the amplicon of HERV-W 19q13.2 were observed and we failed to prove that demethylation has occurred.

In the hope of improving demethylation, 5-azadC treatment was applied to cells with a higher proliferative capacity. Lymphocytes were treated with increasing concentrations of 5-azadC (100nM to 10 µM) for 96 hours. During treatment, the number of viable cells was determined using the trypan blue dye exclusion test of cell viability, and cell proliferation was maintained by daily adding fresh LPS. We found that cell viability was neither affected by any of the 5-azadC treatments used, nor induced by LPS (data not shown). Lymphocytes were exposed to 10µM 5-azadC for 96 hours, and DNA demethylation was assayed in the amplicon of *LRRC27*. We found that, just like OPCs, the average LRRC27 methylation level did not vary between untreated and 5-azadC treated lymphocytes, the levels measured at 15 % and 18%, respectively (Figure 4B). Therefore, the 5-azadC treatment also failed to induce DNA demethylation in lymphocytes.

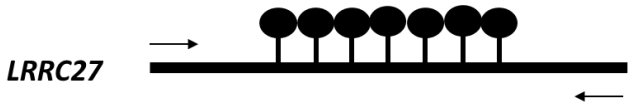
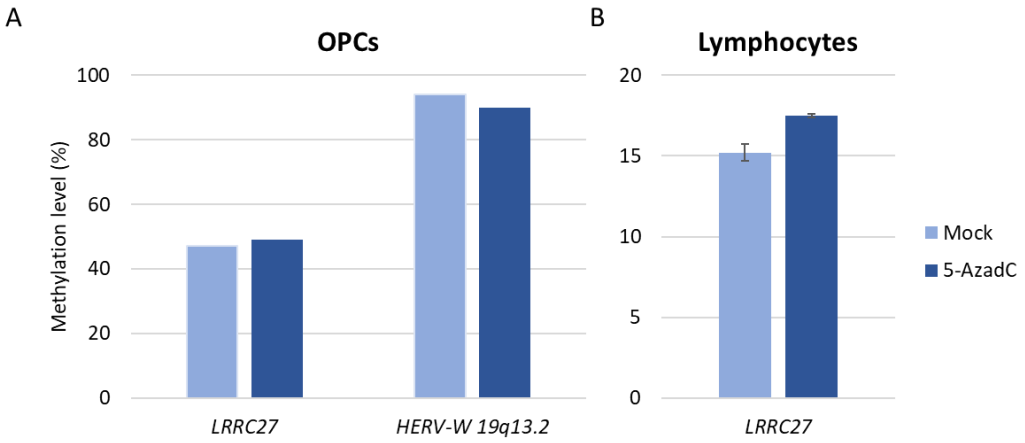


Figure 3. Schematic representation of the methylation assay designed for gene *LRRC27*. Arrows represent the pair of primers used in PCR to selectively amplify LRRC27. Biotinylated and sequence primers are indicated in Table s2. The round symbols ( ) indicate individual CpG sites within the amplicons analyzed.



OPCs	Nº CpGs	N	Methylation level (%)	
Targets			Mock	5-AzadC
LRRC27	7	1	47	49
HERV-W 19q13.2	4	1	94	90

Lymphocytes	Nº CpGs	N	Methylation level (%)	
Target			Mock	5-AzadC
LRRC27	7	2	15 (0.51)	18 (0.1)

Figure 4. Dose-dependent effect of 5'aza2D on methylation levels in oligodendrocyte precursor cells (OPCs) and lymphocytes. A) OPCs were exposed to 1 µM 5-aza-2-deoxycytidine (5-AzadC) for 96 hours or left untreated (mock). B) Lymphocytes were exposed to 10 µM 5-aza-2-deoxycytidine (5-AzadC) for 96 hours or left untreated (mock). In both assays, methylation levels of indicated targets were determined by pyrosequencing. Results are represented as the average of the percentage of methylation of CpG sites across *loci.* For the assay A, none standard deviation could be calculated since there was just one biological replicate (N=1). On the contrary, for assay B, standard deviation was calculated (N=2). The tables below show numerical values taken from the graphs.

Effect of 5-azadC on DNA methylation in human embryonic stem cells

In a last attempt of induce a 5-azadC-mediated DNA demethylation, we applied the drug to human embryonic stem cells (hESCs), which are highly proliferative cells likely to be susceptible to 5-azadC action. We first examined the effect of 5-azadC on cell

viability. ES[4] cells were treated with increasing concentrations of 5-azadC (0.1 nM to 5µM) for 96 hours. After treatment, cell viability was observed under the microscope and defined based on cell adherence and morphology. We found that cell attachment was markedly reduced after 48 hours of treatment with doses of 100 nM 5-azadC and higher (data not shown). In contrast, doses lower than 10 nM 5-azadC were observed not to have any effect on cell adherence and morphology after 96 hours of treatment (data not shown). These results revealed that the most suitable conditions for ES[4] cells treatment were an exposure of 10 nM and 100 nM 5-azadC for 96 and 48 hours, respectively.

To test the effect of 5-azadC on DNA demethylation, methylation levels of the amplicon of HERV-W 19q13.2 and some control genes were assessed. First, pyrosequencing assays for genes expressed in hESCS and whose expression is regulated by DNA methyaltion (Kim et al., 2011) were carried out. Pyrosequencing assays for *POU5F1*, *LEFTY1* and *SERPINA3* (Kim et al., 2011) were optimized and validated, allowing analysis of a single CpG site in the amplicons for *POU5F1*, *LEFTY1* and two CpG sites in the amplicon for *SERPINA3* (Figure 5). Examples of the pyrograms obtained from these assays are shown in Figure s2. Next, ES[4] cells were treated with 10 nM and 100 nM 5-azadC for 96 and 48 hours, respectively or left untreated (mock). We found that *POU5F1*, *LEFTY1* and *SERPINA3* methylation levels in untreated ES[4] cells were 54%, 87%, and 89%, respectively (Figure 6). After treatment with 10nM 5-azadC for 96 hours, *POU5F1* and *LEFTY1* methylation levels were 42%, 79%, and 89%, respectively (Figure 6). And therefore, *POU5F1* and *LEFTY1* methylation levels decreased 22% and 9% respectively, and no variation was detected for *SERPINA3* (Figure 6). After treatment with 100nM 5-azadC for 48 hours, the levels of methylation of *POU5F1*, *LEFTY1* and *SERPINA3* were 37%, 67% and 76%, respectively (Figure 6). Therefore, methylation levels decreased by 31%, 23% and 15%, respectively. Therefore, we observed a 5-azadC dose-dependent DNA demethylation across *loci* in hESCs. When we analysed the methylation levels of the amplicon of HERV-W 19q13.2 (Figure 1), we found that methylation levels in untreated ES[4] cells were 94% (Figure 6). However, we found that methylation levels did not change upon treatment with 10 nM and 100 nM 5-azadC for 96 and 48 hours, and the levels measured at 96 % and

94%, respectively (Figure 6). Unfortunately, the resistance of the amplicon of HERV-W 19q13.2 to 5-azadC-induced demethylation precluded the study of the association between HERV-W methylation and expression levels.

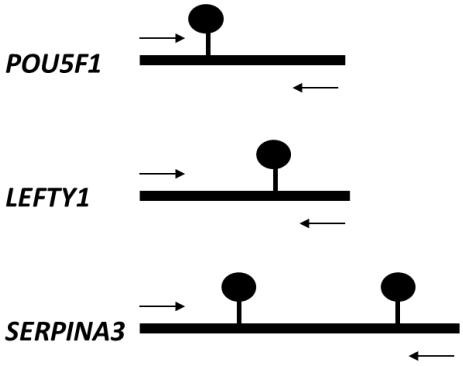
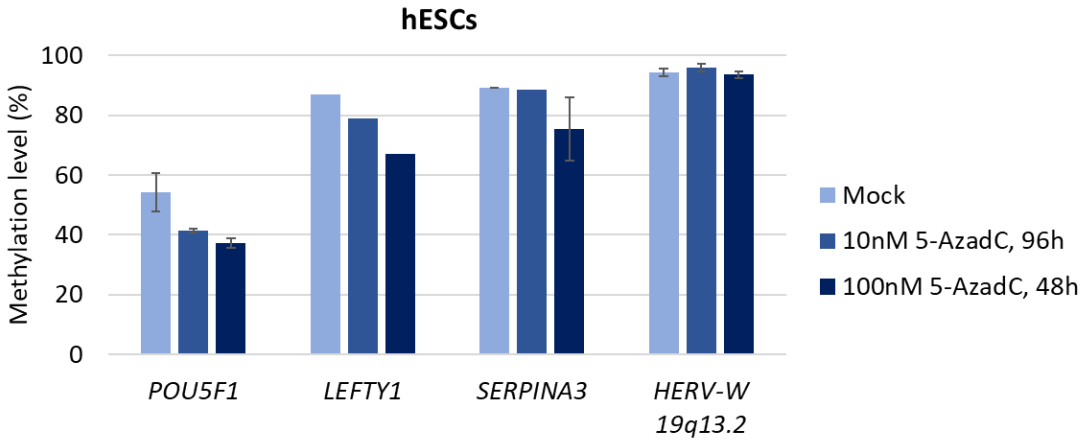


Figure 5. Schematic representation of the methylation assay designed for genes *POU5F1, LEFTY1* and *SERPINA3.* Arrows represent the pair of primers used in PCR to selectively amplify each gene. Biotinylated and sequence primers are indicated in Table s2. The round symbols ( ) indicate individual CpG sites within the amplicons analyzed. It is not drawn to scale.



	Nº CpGs	N	Methylation level (%)		
Targets			**Mock**	**10nM 5-AzadC, 96h**	**100nM 5-AzadC, 48h**
POU5F1	1	2-4	54 (6.43)	42 (0.7)	37 (1.7)
LEFTY1	1	1	87	79	67
SERPINA3	2	1-2	89 (0)	89	76 (10.6)
HERV-W 19q13.2	4	2-3	94 (1.24)	96 (1.2)	94 (1.1)

Figure 6. Dose-dependent effect of 5'aza2D on methylation levels in ES[4] cells. ES[4] cells were exposed to 10 nM and 100 nM 5-aza-2-deoxycytidine (5-AzadC) for 96 and 48 hours respectively or left untreated (mock). Methylation levels of indicated targets were determined by pyrosequencing Results are represented as the average of the percentage of methylation of CpG sites across *loci*. Biological replicates are indicated (N). The corresponding standard deviation was calculated except for those assays performed only once (N=1). The table below shows numerical values taken from the graph.

Discussion

While some reports have shown that changes in DNA hypomethylation correlate with HERVs activation in cancer cells or in CD4+ T cells from lupus patients (Goering et al., 2011; Wu et al., 2015), HERV-W methylation status remains unknown in the context of MS. In a previous study (Manuscript II), we detected that HERV-W expression was slightly increased in PBMCs of MS patients. In the present study, we directly addressed the hypothesis that increased HERV-W expression in MS patients may be promoted by HERV-W hypomethylation. We successfully developed *locus*-specific methylation assays for 3 HERV-W *loci*: 19q13.2, 15q21.3 and Xq22.3. However, our results showed that the methylation status of three HERV-W *loci* did not differ in MS patients and controls, and that HERV-W *loci* were hypermethylated in all cases (Figure 2). The most straightforward conclusion of this data is that methylation levels do not correlate with changes in HERV-W expression in MS patients. However, a potential correlation could be masked by imperfect study design. As HERV-W expression was only slightly increased in MS patients (Manuscript II), a potential small difference in HERV-W methylation levels between MS patients and controls may have fallen below the detection limit of our assays. This may be the case if differences are restricted to a subset of cells. The observed lack of correlation between HERV-W expression and methylation might also be explained by the design of our methylation assays, which targets HERV-W *ENV* regions. We intended to analyse HERV-W *ENV* regions since we had observed a HERV-W *ENV* expression increase in MS patients (Manuscript II). However, the expression of HERVs is thought to be regulated by promoter and enhancer regions within the long terminal repeats (LTRs) located at both ends of HERV sequence (Kovalskaya et al., 2006; Nelson et al., 2003). For the HERV-W *locus ERVWE1* (which encodes Syncytin-1), evidence indicate that the promoter region is located within the 5’ LTR (Cheng et al., 2004), and that hypermethylation of the 5'LTR correlates with decreased *Syncytin*-1 mRNA and protein levels in preeclamptic placentas (Zhuang et al., 2014). Therefore, it is possible that the methylation state of the LTR region of HERV-W is more relevant. While we tried to design and validate copy-specific methylation assays in the LTR, we failed in obtaining such PCR products

(data not shown). But it might be possible that the methylation state of the *ENV* region of HERV-W could be irrelevant for transcription.

As we did not observe differences in DNA methylation within HERV-W *loci* in MS patients, we wanted to study whether DNA methylation regulates gene expression levels. To do so, we artificially induced DNA demethylation in different cell types in culture and analyses the effect on HERV-W expression. We treated OPCs, lymphocytes and hESCs with the demethylating agent 5-azadC. No differences in the DNA methylation level were detected neither in OPCs nor lymphocytes after 5-AzadC treatment (Figure 4). It had to be considered that 5-AzadC does not act demethylating existing DNA but demethylating new DNA. Therefore, effectiveness of 5-AzadC treatment depended on DNA synthesis and, it can be possible that longer treatments or higher 5-AzadC doses are requires to observe changes in DNA methylation. 5-AzadC treatment in hESCs yielded DNA demethylation of some control genes but not HERV-W elements (Figure 6). Both, inefficiency of 5-AzadC treatment and the resistance of HERV-W elements to 5-azadC-induced demethylation, prevented us from analysing whether DNA hypomethylation correlates with HERV-W activation. Taken together, our findings indicate that further studies are required to understand the role that DNA methylation modifications may play in MS.

Supplementary materials

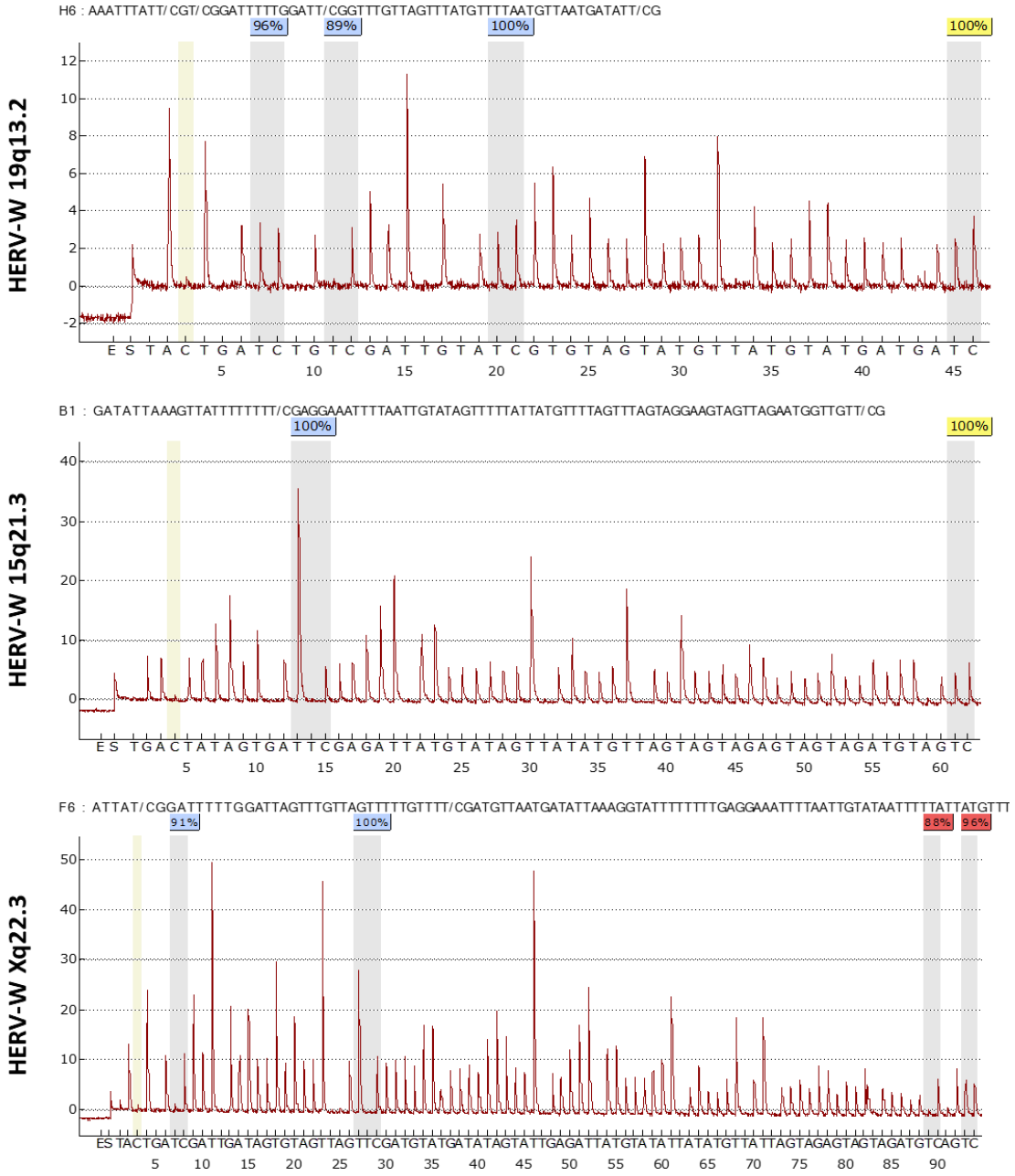


Figure s1. Representative pyrograms of HERV-Wenv *locus*-specific methylation assays. Pyrosequencing analyses for the indicated *loci* are represented in pyrograms. Vertical axis: signal intensity; horizontal axis: order of residues analyzed. The % methylation of each CpG sites analyzed are indicated in colored boxes. The quality of the result at each position is estimated by the software, and represented through a color-based score: blue ("passed"), yellow ("to check") and red ("failed"). The yellow/brown bar shows a C analyzed to confirm complete conversion of non-methylated cytosines.

Target	Sequence (5'-3')	Cycling conditions (first PCR round)	Cycling conditions (second PCR round)	Nucleotide dispensation order (5'→3')
HERV-W locus 19q13.2	For: TTTTTTAAGAATGAGGAGATGGA	40X: 94 °C 30 seconds		AAATTTATCGCGGATTTTTGGATCG
	Rev: [Btn]TCTCAACAAAAAAACCCAAA	60 °C 30 seconds		GTTTGTTAGTTTATGTTTTAATGTTA
	Seq: AATGGAGTTTTAGATGTAGTTTATGATTA	72 °C 30 seconds		ATGATATCG
		1X: 72 °C 10 minutes		
HERV-W locus 15q21.3	For-ext: GTAGAGTAGAGGAGTTTTAAAAT	35X: 94 °C 1 minute	1X: 98 °C 2 minutes	GATATTAAAGTTATTTTTTTCGAGG
	Rev-ext: AATAAACATTCATTTACCAACT	48 °C 1 minute	35X: 94 °C 1minute	AAATTTTAATTGTATAGTTTTTATTAT
	For-int: GGGTTTTTTTAGTTAATGGATGT	72 °C 1 minute	53 °C 1 minute	GTTTTAGTTTAGTAGGAAGTAGTTA
	Rev-int: [Btn]TTCTTTTCCTCCTATTTTTACCT	1X: 72 °C 10 minutes	72 °C 1 minute	GAATGGTTGTCG
	Seq: TTTGTTAGTTTATGTTTTAATGTTAAT		1X: 72 °C 10 minutes	
HERV-W locus Xq22.3	For-ext: GGGGTTTTTTTAGTTAATGGAT	35X: 94 °C 1 minute	1X: 98 °C 2 minutes	ATTACGGATTTTTGGATTAGTTTGT
	Rev-ext: AAACATACCATCTAAACTTA	53 °C 1 minute	35X: 94 °C 1minute	TAGTTTTTGTTTCGATGTTAATGATA
	For-int: GGATGTTTTGGGTTTTTTTT	72 °C 1 minute	50 °C 1 minute	TTAAAGGTATTTTTTTTGAGGAAAT
	Rev-int: [Btn]AACTAACTATTTCTTTACCTCCT	1X: 72 °C 10 minutes	72 °C 1 minute	TTTAATTGTATAATTTTTATTATGTTT
	Seq: TGTAGTTTATGATTAAAATTT		1X: 72 °C 10 minutes	TATTTTAGTAGGAAGTAGGTAGAGC GGTCG

Table s1. HERV-Wenv *locus*-specific PCR amplification and DNA methylation assays. The first column shows the HERV-Wenv *locus* target and the second column indicates the oligo sequences used for individual HERV-Wenv *locus*-specific PCR amplification and pyrosequencing analysis. (For= forward; Rev= reverse; ext: first nested PCR round; int: second nested PCR round; seq: sequence primer). The third and fourth columns indicate the cycling conditions for the first and the second PCR round respectively. The last column shows the nucleotide dispensation order for the pyrosequencing analysis. CpGs sites are highlighted in yellow.

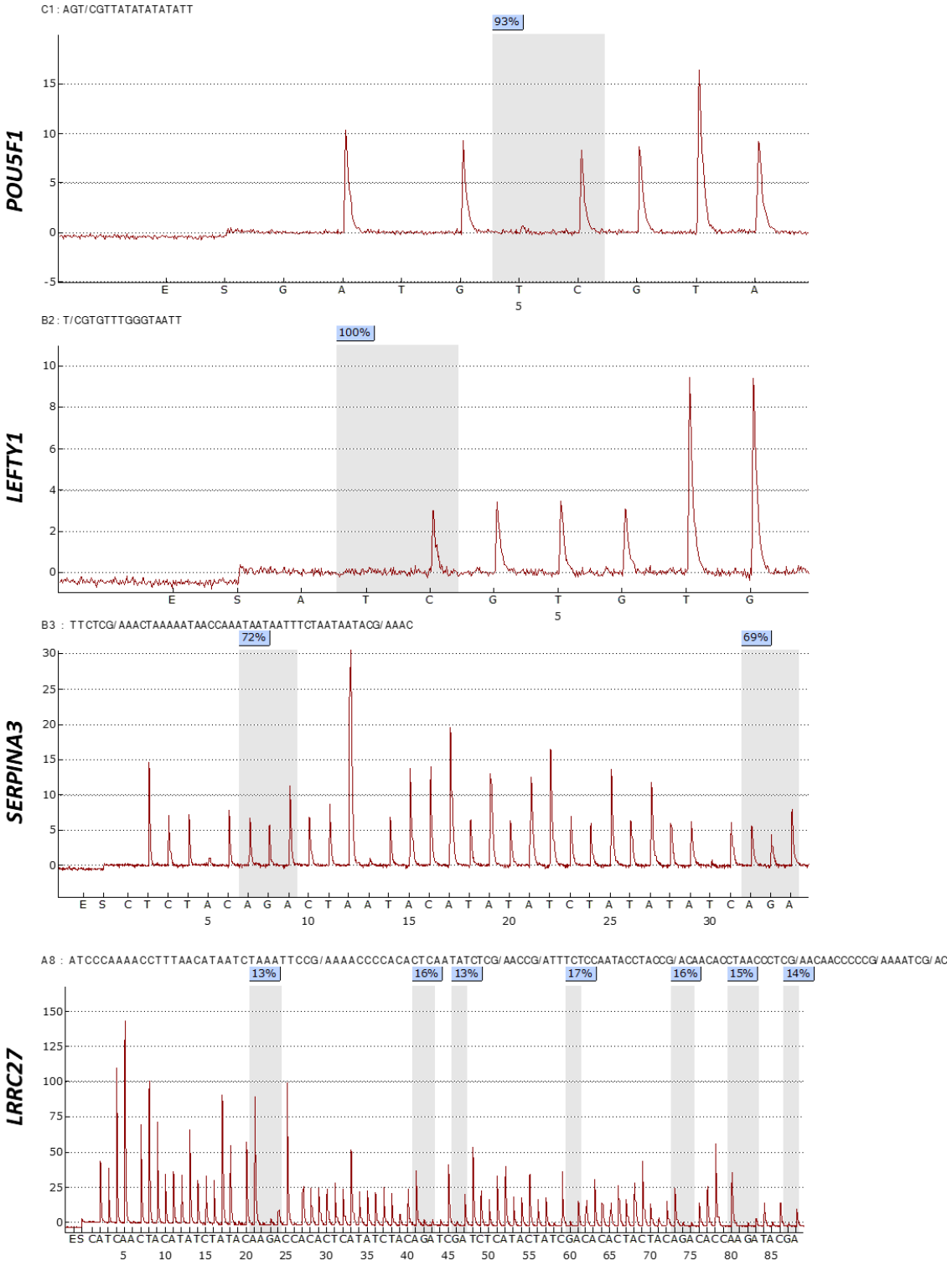


Figure s2. Representative pyrograms of gene-specific methylation assays. Pyrosequencing analyses for the indicated *loci* are represented in pyrograms. Vertical axis: signal intensity; horizontal axis: order of residues analyzed. The % methylation of each CpG sites analyzed are indicated in colored boxes. The quality of the result at each position is estimated by the software, and represented through a color-based score: blue ("passed"), yellow ("to check") and red ("failed").

Target	Sequence (5'-3')	Cycling conditions (first PCR round)	Nucleotide dispensation order (5'→3')
POU5F1	For: GGAAGGTATAAGGGAGTA Rev: [Btn] CACACCTCCATATTCTTC Seq: TGTTGATGTATTGAGGGAT	35X: 94 °C 1 minute 50 °C 1 minute 72 °C 1 minute 1X: 72 °C 10 minutes	AGCGTTATATATATATT
LEFTY1	For: TGGGGTATAATTATGTTTTG Rev: [Btn]CCCCTAACAACCAAATAT Seq: TGAGTTATTGTAAATAGAAG	35X: 94 °C 1 minute 50 °C 1 minute 72 °C 1 minute 1X: 72 °C 10 minutes	CGTGTTTGGGTAATT
SERPINA3	For: [Btn] TATTTGAGTAGGTTTAATAAGTT Rev: CAAAACAACCAAATATTCTA Seq: CAACCAAATATTCTATTCTC	35X: 94 °C 1 minute 50 °C 1 minute 72 °C 1 minute 1X: 72 °C 10 minutes	TTCTCGAACTAAAAATAACCAAA TAATAATTTCTAATAATACGAAC
LRRC27	For: [Btn] TGATGTGTTAATTTGTTTTTGAAATG Rev= seq: ACTAACCCAACACACCACTATTTCT	35X: 94 °C 1 minute 60 °C 1 minute 72 °C 1 minute 1X: 72 °C 10 minutes	ATCCCAAAACCTTTAACATAATC TAAATTCCGAAACCCCACACTCA ATATCTCGACCGTTTCTCCAATAC CTACCGCAACACCTAACCCTCGA CAACCCCCGAAATCGC

Table s2. Gene-specific PCR amplification and DNA methylation assays. The first column shows the gene and the second column indicates the oligo sequences used for PCR amplification and pyrosequencing analysis. (For= forward; Rev= reverse; Btn: biotin; seq: sequence primer). The third and fourth columns indicate the cycling conditions for PCR and the last column shows the nucleotide dispensation order for the pyrosequencing analysis. CpGs sites are highlighted in yellow.

References

Balestrieri, E., Pica, F., Matteucci, C., Zenobi, R., Sorrentino, R., Argaw-Denboba, A., Cipriani, C., Bucci, I., and Sinibaldi-Vallebona, P. (2015). Transcriptional activity of human endogenous retroviruses in human peripheral blood mononuclear cells. BioMed Res. Int. *2015*, 164529.

Cheng, Y.-H., Richardson, B.D., Hubert, M.A., and Handwerger, S. (2004). Isolation and characterization of the human syncytin gene promoter. Biol. Reprod. *70*, 694–701.

Chistiakov, D.A., Myasoedova, V.A., Orekhov, A.N., and Bobryshev, Y.V. (2017). Epigenetically Active Drugs Inhibiting DNA Methylation and Histone Deacetylation. Curr. Pharm. Des. *23*, 1167–1174.

Cho, K., Lee, Y.-K., and Greenhalgh, D.G. (2008). Endogenous retroviruses in systemic response to stress signals. Shock Augusta Ga *30*, 105–116.

Colmegna, I., and Garry, R.F. (2006). Role of endogenous retroviruses in autoimmune diseases. Infect. Dis. Clin. North Am. *20*, 913–929.

Dolei, A. (2006). Endogenous retroviruses and human disease. Expert Rev. Clin. Immunol. *2*, 149–167.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., Cirotto, M.C., Rosati, G., and Sotgiu, S. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology *58*, 471–473.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc. Natl. Acad. Sci. U. S. A. *89*, 1827–1831.

Garson, J.A., Tuke, P.W., Giraud, P., Paranhos-Baccala, G., and Perron, H. (1998). Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet Lond. Engl. *351*, 33.

Goering, W., Ribarska, T., and Schulz, W.A. (2011). Selective changes of retroelement expression in human prostate cancer. Carcinogenesis *32*, 1484–1492.

Hurst, T.P., and Magiorkinis, G. (2017). Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (LTRs). Viruses *9*.

Huynh, J.L., Garg, P., Thin, T.H., Yoo, S., Dutta, R., Trapp, B.D., Haroutunian, V., Zhu, J., Donovan, M.J., Sharp, A.J., et al. (2014). Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. Nat. Neurosci. *17*, 121–130.

Kim, H.-S. (2012). Genomic Impact, Chromosomal Distribution and Transcriptional Regulation of HERV Elements. Mol. Cells *33*, 539–544.

Kim, M., Kang, T.-W., Lee, H.-C., Han, Y.-M., Kim, H., Shin, H.D., Cheong, H.S., Lee, D., Kim, S.-Y., and Kim, Y.S. (2011). Identification of DNA methylation markers for lineage commitment of in vitro hepatogenesis. Hum. Mol. Genet. *20*, 2722–2733.

Kovalskaya, E., Buzdin, A., Gogvadze, E., Vinogradova, T., and Sverdlov, E. (2006). Functional human endogenous retroviral LTR transcription start sites are located between the R and U5 regions. Virology *346*, 373–378.

Kurth, R., and Bannert, N. (2010). Beneficial and detrimental effects of human endogenous retroviruses. Int. J. Cancer *126*, 306–314.

Löwer, R., Löwer, J., and Kurth, R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. U. S. A. *93*, 5177–5184.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., Serra, C., Sotgiu, S., Bonetti, B., and Dolei, A. (2007). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J. Gen. Virol. *88*, 264–274.

Matoušková, M., Blažková, J., Pajer, P., Pavlíček, A., and Hejnar, J. (2006). CpG methylation suppresses transcriptional activity of human syncytin-1 in non-placental tissues. Exp. Cell Res. *312*, 1011–1020.

Nelson, P.N., Carnegie, P.R., Martin, J., Davari Ejtehadi, H., Hooley, P., Roden, D., Rowland-Jones, S., Warren, P., Astley, J., and Murray, P.G. (2003). Demystified. Human endogenous retroviruses. Mol. Pathol. MP *56*, 11–18.

Palii, S.S., Van Emburgh, B.O., Sankpal, U.T., Brown, K.D., and Robertson, K.D. (2008). DNA Methylation Inhibitor 5-Aza-2'-Deoxycytidine Induces Reversible Genome-Wide DNA Damage That Is Distinctly Influenced by DNA Methyltransferases 1 and 3B. Mol. Cell. Biol. *28*, 752–771.

Perron, H., Germi, R., Bernard, C., Garcia-Montojo, M., Deluen, C., Farinelli, L., Faucard, R., Veas, F., Stefas, I., Fabriek, B.O., et al. (2012). Human endogenous retrovirus type W envelope expression in blood and brain cells provides new insights into multiple sclerosis disease. Mult. Scler. J. *18*, 1721–1736.

Ramos, M.-P., Wijetunga, N.A., McLellan, A.S., Suzuki, M., and Greally, J.M. (2015). DNA demethylation by 5-aza-2'-deoxycytidine is imprinted, targeted to euchromatin, and has limited transcriptional consequences. Epigenetics Chromatin *8*.

Reiss, D., Zhang, Y., and Mager, D.L. (2007). Widely variable endogenous retroviral methylation levels in human placenta. Nucleic Acids Res. *35*, 4743–4754.

Seifarth, W., Frank, O., Zeilfelder, U., Spiess, B., Greenwood, A.D., Hehlmann, R., and Leib-Mösch, C. (2005). Comprehensive Analysis of Human Endogenous Retrovirus Transcriptional Activity in Human Tissues with a Retrovirus-Specific Microarray. J. Virol. *79*, 341–352.

Stengel, S., Fiebig, U., Kurth, R., and Denner, J. (2010). Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. Genes. Chromosomes Cancer *49*, 401–411.

Wu, Z., Mei, X., Zhao, D., Sun, Y., Song, J., Pan, W., and Shi, W. (2015). DNA methylation modulates HERV-E expression in CD4+ T cells from systemic lupus erythematosus patients. J. Dermatol. Sci. *77*, 110–116.

Zhuang, X.-W., Li, J., Brost, B.C., Xia, X.-Y., Chen, H.B., Wang, C.-X., and Jiang, S.-W. (2014). Decreased expression and altered methylation of syncytin-1 gene in human placentas associated with preeclampsia. Curr. Pharm. Des. *20*, 1796–1802.

Manuscript IV. Immune responses in human oligodendrocyte precursor cells cultured *in vitro*

Abstract

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS), characterized by the loss of myelin and the damage of oligodendrocytes. Evidence suggests that upregulated human endogenous retrovirus family W (HERV-W) expression in autopsied brain tissues and peripheral blood mononuclear cells (PBMCs) is associated to MS. The first line of defence against viral infection depends on the activation of intracellular receptors called toll-like receptors (TLR). For instance, HERV-W proteins provoke a TLR4-mediated immune response in PBMCs and monocytes. However, the possible relationship between upregulated HERV-W RNA expression and TLR signalling has not yet been addressed in cell types from the CNS. Here we aimed to study the possible TLR3-mediated innate immune system stimulation by RNA transcribed from a HERV-W in human oligodendrocytes. We implemented two published protocols to differentiate human embryonic stem cells (hESCs) into $O4^+$ oligodendrocyte cells. In my hands, none yielded $O4^+$ oligodendrocytes starting from ES[4] cells. A partial HERV-W *ENV* copy (*locus* 5p12) named HERV-W *ENV* 5p12 was cloned, linked to a ubiquitous promoter, and inserted into a lentiviral vector. After transduction into human neural progenitor cells (NPCs), we observed an increased HERV-W *ENV* expression compared to non-transduced NPCs. Such increase was associated with an increase of both *TLR3* and type I *interferon β (IFNβ*) expression. We conclude that HERV-W overexpression triggers a type I interferon response in NPCs, possibly via TLR3-activation.

Introduction

Multiple sclerosis (MS) is an autoimmune inflammatory disease of the central nervous system (CNS) characterized by the loss and damage to oligodendrocytes, the highly specialized cell type of the CNS that produce the myelin (Brück et al., 2002; Lucchinetti et al., 2000; Trapp et al., 1998). As a result, myelin sheaths are affected resulting in neuronal dysfunction (Nylander and Hafler, 2012). Although mechanisms that trigger the autoimmune response remain unclear, genetic, environmental, and viral causes contribute to the development of the disease (Koriem, 2016). Inflammation of the CNS is the result of both a peripherical and local immune response. The massive infiltration of peripherical mediators of the immune system in the brain causes an inflammatory response that causes oligodendrocyte and axonal damage (Brück, 2005). Also, microglia, the resident immune cells of the CNS not only can interact with cells of the adaptive immune system (such as T cells and B cells), but can also directly cause neuroinflammatory tissue damage (Hemmer et al., 2015; Macchi et al., 2015).

As the human brain is inaccessible for experimental approaches for obvious reasons, *in vitro* models are used to study the pathology in human oligodendrocyte associated with MS. In recent years, a number of protocols for generating *in vitro* human pluripotent stem cells-derived oligodendrocytes have been stablished (All et al., 2015; Douvaras and Fossati, 2015; Ehrlich et al., 2017; García-León et al., 2018; Stacpoole et al., 2013). In such protocols pluripotent cells are converted first into neural progenitor cells (NPCs), then into oligodendrocyte precursor cells (OPCs) and finally into $O4^+$ oligodendrocytes. This works for both human embryonic stem cells (hESCs), and induced pluripotent stem cells (iPSCs) derived from MS patients (Douvaras et al., 2014), the latter representing an important approach for the study of MS pathophysiology.

Human endogenous retroviruses family W (HERV-W) have been related to MS pathogenesis since different studies have reported that HERV-W expression is elevated in cerebrospinal fluid (CSF), autopsied brain tissues and peripheral blood mononuclear cells (PBMCs) from MS patients versus controls (Dolei et al., 2002; Garson et al., 1998; Mameli et al., 2007). In response to viral infection, antiviral innate immunity is mostly mediated by toll-like receptors (TLRs). TLRs belong to a family of innate immune

system receptors whose stimulation causes a response of the innate immune system. In particular, TLRs activation leads to the transcriptional activation of genes encoding proinflammatory cytokines, chemokines and type I interferons (IFN-α/β). The latter subsequently trigger innate immune responses and modulate adaptive immunity (Hernández-Pedro et al., 2013; Xagorari and Chlichlia, 2008). TLR receptor are usually expressed on innate immune response cells such as microglia, but also on other cells of the CNS, including oligodendrocytes, astrocytes and neurons (Carpentier et al., 2008; Kielian, 2006; Trudler et al., 2010). TLRs are fundamental for the induction of antiviral innate immune responses (Lester and Li, 2014; Xagorari and Chlichlia, 2008). For instance, HERV-W proteins have the potential to trigger a TLR4-mediated immune response in PBMCs and monocytes (Rolland et al., 2006). Also, recent data showed that overexpression of viral double-stranded (ds) RNA in cancer cells activated a type I interferon-mediated immune response through TLR3 signalling (Chiappinelli et al., 2015). However, the possible relationship between upregulated HERV-W RNA expression in brain of MS patients and TLR recognition has not yet been addressed.

In the present study, we first aimed to establish hESC-derived $O4^+$ oligodendrocytes and oligodendrocyte precursor cells (OPCs), as a simplified model for disease-related processes. We took advantage of two already stablished differentiation protocols (All et al., 2015; García-León et al., 2018). We aimed to monitor immune responses provoked by HERV-W overexpression. To this end, a cloned partial HERV-W copy (*locus* 5p12) was overexpressed in NPCs. After induction of HERV-W expression, the expression level of genes relevant to innate immunity was assessed.

Materials and methods

Human embryonic stem cell line ES[4] cell culture

The human embryonic stem cell line ES[4] was provided by the Stem Cell Bank of Barcelona (BLCB) and propagated on Matrigel matrix (BD) in mTeSR1 medium (STEMCELL Technologies) as described by BLCB. ES[4] cells were passaged when they were approximately 80-85% confluent using ReLeSR™ (STEMCELL Technologies) according to manufacturer's' protocol at a 1:8-1:10 split ratio, ensuring this way that cells were passaged every 3-4 days. For hESCs ES[4] morphology see Figure s.1.

Oligodendrocyte differentiation (All et al., 2015)

Oligodendrocyte precursor cells were differentiated from hESCs ES[4] as described (All et al., 2015)**.** In brief, ES[4] cells were lifted using ACCUTASE™ (STEMCELL Technologies) and seeded in suspension so that embryoid bodies (EBs) were generated. On the second day, medium was changed from mTeSR1 medium to N2B27 medium supplemented with 150 ng/ml LDN-193189 (axonmedchem) and 20 ng/ml basic fibroblast growth factor (FGF2) (Peprotech). N2B27 medium contained Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) (Thermo Fisher Scientific) supplemented with 1 X non-essential amino acids (SIGMA), 1 X Glutamax (Thermo Fisher Scientific), 1 X N2 supplement (Thermo Fisher Scientific), 1 X B27 supplement (Thermo Fisher Scientific), 50 µM 2-Mercaptoethanol**,** 7,5% Bovine Serum Albumin (BSA) and 1 X Penicillin-Streptomycin**.** EBs were daily fed for 15 days with N2B27 medium, then they were plated onto Matrigel-coated 60mm culture plates in N2B27 medium supplemented with 20 ng/ml FGF2 for 5 days, leading to differentiation into neural progenitors (NPCs). NPs were expanded in this media for 7 days and glial progenitors (GPCs) were subsequently induced by adding 10 ng/ml epidermal growth factor (EGF) (STEMCELL Technologies) for 7 days. After 7 days in GP media, cells were split with collagenase IV (Thermo fisher) and transferred into new Matrigel-coated culture plates. Oligodendrocyte precursor cells (OPs) were induced by supplementing N2B27 medium with 20 ng/ml EGF and 20 ng/ml platelet-derived growth factor AA (PDGF-AA) (STEMCELL Technologies) for 25 days.

Oligodendrocyte differentiation (García-León et al., 2018)

Oligodendrocytes were differentiated from hESCs ES[4] using SOX10 overexpression as described (García-León et al., 2018). Briefly, on day 0, NPCs were induced by switching medium from mTeSR™1 medium to NPC differentiation medium supplemented with 25 ug/ml insulin (I9278, Sigma), 10 mM SB431542 (301836-41-9, axonmedchem), 1 mM LDN193189 and 100 nM retinoic acid (RA) (R2625, Sigma) until day 8. NPC differentiation medium consisted of Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) (Thermo Fisher Scientific), 1 X non-essential amino acids (SIGMA), 1 X Glutamax (Thermo Fisher Scientific), 1 X N2 supplement (Thermo Fisher Scientific), 1 X B27 supplement (Thermo Fisher Scientific), 50 µM 2-Mercaptoethanol

and 1 X Penicillin-Streptomycin. On day 8, SB431542 and LDN193189 were removed and 1 mM smoothened agonist (SAG) (73412, STEMCELL Technologies) added until day 12, when NPCs were dissociated, plated, and transduced with lentiviral particles containing FUW-SOX10 vector. FUW-Sox10 was a gift from Catherine M. Verfaillie (García-León et al., 2018). The next day, medium was changed to OL differentiation medium consisting of NPC differentiation medium supplemented with 25 ug/ml insulin (Sigma), 10 ng/ml of PDGF-AA (Peprotech), 10 ng/ml insulin-like growth factor 1 (IGF1) (STEMCELL Technologies), 5 ng/ml hepatocyte growth factor (HGF) (STEMCELL Technologies), 10 ng/ml neurotrophin 3 (NT3) (Peprotech), 100 ng/ml biotin (Sigma), 1 µM adenosine cyclic monophosphate (cAMP) (Sigma), 60 ng/ml thriiodothytonine (T3) (Sigma) and 1 µg/ml doxycycline (Sigma). Cells were maintained for 7–10 days.

RNA isolation and cDNA synthesis

RNA isolation and cDNA synthesis were performed as described before (M&M, Manuscript II).

Semiquantitative analysis by PCR

PCR reaction was carried out in a 50-µl volume using 2 µl of cDNA, 200 µM dNTPs, forward and reverse primers (400 nM each), 0.5 U Taq DNA Polymerase (D1806, Sigma) and its appropriated buffer. For primer sequences see Table s1. PCR was performed in a Veriti Thermal Cycler (Applied Biosystems) and the cycling conditions consisted of 98 °C for 2 minutes and then 35 cycles of 94 °C for 1 minute, 60 °C for 1minute and 72 °C for 1 minute, with a final extension of 72 °C for 5 minutes. The presence of amplified products was confirmed on 1.5 % agarose gels, visualized using Gel Red™ (41003, Biotium) and photographed on a Gel Doc transilluminator (BioRad).

Flow cytometry: extracellular staining protocol

A single-cell suspension was prepared lifting cells with ACCUTASE™ (STEMCELL Technologies). All antibodies were suspended in Blocking buffer (5% fetal bovine serum (FBS) in PBS). Cells were pelleted at 200 g for 3 minutes and incubated with primary antibody for 1 hour at 4 °C after resuspension. Cells were subsequently washed with Blocking buffer and incubated with secondary antibody for 30 minutes at

4 °C. When using a conjugated primary antibody, cells were incubated for 30 minutes at 4 °C. After labelling, cells were washed, resuspended in 200 µl PBS and stained with SYTOX™ Dead Cell Stain Sampler Kit (S34862, Thermo Fisher Scientific) prior to data collection. Antibodies and dilutions used: CD 29 Mouse Anti-Human (CD2901, Thermo Fisher Scientific; 1:10), Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 (A11029, Invitrogen; 1:50), Oligodendrocyte Marker O4 Alexa Fluor 488-conjugated antibody (FAB1326G, R&D Systems; 1.25µg). Quantification was carried out using a FACSAria™ instrument (BD BioSciences) equipped with a 488-nm laser. Data were analysed using the WEASEL Flow Cytometry Software. Gates were established based on the corresponding unstained and isotype controls. To analyse GFP expression, gates were set based on non-transduced cells.

Plasmid constructs

To induce HERV-W overexpression, HERV-W copy located on chromosome 5p12 (HERV-W 5p12) was amplified and fused to the phosphoglycerate kinase (PGK) promoter (HERV-W 5p12-PGK). Both sequences, HERV-W 5p12 and HERV-W 5p12-PGK were then fused to the plasmid pSIN-EF2-Lin28-Pur which contains the EF-1α promoter (Figure s.2). Plasmids were constructed through the following steps. All enzymes mentioned were from Thermo Scientific unless specified otherwise. For primer sequences see Table s.2.

A genomic HERV-W 5p12 copy was amplified by PCR as follows. The PCR reaction was carried out in a 50-µl volume using 100 ng of gDNA, 200 µM dNTPs, forward and reverse primers (400 nM each), 0.5 U Taq DNA Polymerase (D1806, Sigma) and its appropriate buffer. gDNA was boiled at 100 °C for 3 minutes before addition to the reaction. Cycling conditions consisted of 35 cycles of 94 °C for 1 minute, 60°C for 1minute and 72°C for 1 minute, with a final extension of 72 °C for 5 minutes. The presence of amplified products was confirmed on 1% agarose gels, visualized using Gel Red™ (41003, Biotium) and products were vizualized on a Gel Doc transilluminator (BioRad). PCR product was cleaned up using the Geneclean Turbo kit (MP Biomedicals), precipitated with sodium acetate-ethanol and cloned into the pGEM®-T Easy Vector System I (Promega) according to manufacturer's instructions. The vector was

transformed into E. coli XL1 cells and cultured in Luria–Bertani (LB)-agar plates supplemented with 50 µg/mL ampicillin at 37 °C. Colonies were analysed by plasmid miniprep, screened for the presence of insert by digestion with EcoRI and further sequencing (Secugen S.L). Among all clones analysed, one with a 5'-T7 promoter-5'insert-3'insert-SP6 promoter-3' orientation was selected and purified using a midiprep kit (K210004, Invitrogen). Purified plasmid was digested with SpeI and EcoRI. Digested product was cleaned up using phenol-chloroform, precipitated with sodium acetate-ethanol and re-suspended in TE buffer. The resulting product was gel-purified (Geneclean Turbo kit, MP Biomedicals).

The PGK promoter was amplified by PCR as above, using a PGK-containg plasmid as a template and specific primers (Table s2). Resulting PCR product was purified using the Diffinity RapidTip®2 (Sigma), digested two with BamHI and XbaI and cloned after purification into appropriately digested pSIN-EF2-Lin28-Pur plasmid (16580, Addgene).

pSIN-EF2-Lin28-Pur plasmid (16580, Addgene) was digested with EcoRI, and either BamHI or SpeI. Digestion products were treated with Shrimp Alkaline Phosphatase (Fermentas) at 37 °C for 30 minutes and subsequently, gel-purified as above.

Both PGK promoter fragment (BamHI- XbaI) and HERV-W 5p12 (EcoRI- SpeI) fragment were ligated into pSIN-EF2-Pur After transformation, colonies were screened for the presence of insert by digestion and verified further sequencing (Secugen S.L). The resulting clones were named pSIN-EF2-HERV-W 5p12 and pSIN-EF2-HERV-W 5p12-PGK. High-quality plasmid DNAs was isolated using a midiprep kit (K210004, Invitrogen).

Plasmid DNA transfection of 293T cells

Human embryonic kidney 293 T cells were cultured in 293T cells were culture in Dulbecco's modified Eagle's medium. 293T cells were transfected on PolyDLysin-coated dishes using LipofectamineTM reagent (Invitrogen) with either pSIN-EF2, pSIN-EF2-HERV-W 5p12 or pSIN-EF2-PGK-HERV-W 5p12 vector. After 48 hours, cells were homogenized with TRIzol® reagent (Invitrogen) for future RNA isolation.

Production of lentiviral particles and cell transduction

FUW-SOX10, pSIN-EF2-HERV-W 5p12, pSIN-EF2-PGK-HERV-W 5p12 and pLOX-CWgfp (a green fluorescent protein (GFP)-containing vector; addgene #12241) vectors were independently co-transfected with the packaging (psPAX2, addgene #12260) and the envelope (pMDs.G, addgene #12259) plasmids into the 293T cell line using FUGENE 6 Transfection reagent (Promega). The supernatant containing lentiviruses was collected after 48 hours, filtered through 0.45µm filter and stored in aliquots at -80°C for further use. For the $O4^+$ oligodendrocyte cells differentiation (García-León et al., 2018), NPCs at day 12 were lifted using ACCUTASE™ (STEMCELL Technologies) and plated at 50.000 cells/cm^2 on 0.1 mg/ml poly-Lornithine/10 µg/ml laminin (P3655, L2020 Sigma) coated dishes before transduction. The original protocol could not be replicate due to cell adhesion problems. For the induction of HERV-W expression, NPCs at day 12 (García-León et al., 2018) were separately transduced with pSIN-EF2-HERV-W 5p12, pSIN-EF2-PGK-HERV-W 5p12 and pLOX-CWgfp lentiviral vectors for two consecutive days. Each day, cells were cultured with 1 ml of culture medium plus 1 ml of lentiviral supplemented with polybrene. Transduction efficiency was assessed by flow cytometry, and estimated as the percentage of cells expressing GPP, that is, the percentage of cells transduced with pLOX-CWgfp. After transduction, lentiviral supernatant was removed, cells were washed three times with PBS (14190144, Thermo Fisher Scientific), and maintained in NPC differentiation medium with 100 nM RA and 1 mM SAG until collecting data.

Quantitative real-time PCR

cDNA was analyzed by Real-Time PCR on a ViiA™ 7 Real-Time PCR System (Applied Biosystems) using the SYBR Green Master Mix (SYBR Premix Ex Taq II (Tli RNase H Plus), RR820A, Takara) according to manufacturer's protocol. Amplification efficiency of all set of primers between 90% and 110% was assessed in standard curves (Appendix I) and, sequences of all primers used (used at 200 nM) are listed in Figure s3. All reactions were carried out in triplicate and only measurements with a standard deviation < 0.2 were considered. Target gene expression levels were recalculated as 2^-

$^{\Delta\Delta Ct}$ with respect to the mean ΔCt value of the samples in the control group, using *GAPDH* as a reference gene (Livak and Schmittgen, 2001).

Results

Generation and characterization of in vitro human oligodendrocytes

With the aim of generating *in vitro* $O4^+$ oligodendrocyte cells, we followed a publish differentiation protocol (Figure 1) (All et al., 2015). The first step consisted on the formation of embryoid bodies (EBs). For this, human embryonic stem cells (hECs) were seeded in suspension and within 24 hours EBs were formed. Such EBs increased in size and complexity over the course of differentiation, until showing a spheroid shape with a darker centre (Figure 2A). EBs were subsequently plated and, within 24 hours, cells with a neuroectodermal morphology were observed with projections extending outwards from the centre of the EBs (Figure 2B), leading to differentiation into neural progenitor cells (NPCs) (Figure 2C). NPCs were expanded for a week (Figure 2D) and glial progenitor cells (GPCs) were subsequently induced (Figure 2E), and finally converted into oligodendrocyte progenitor cells (OPCs). The obtained cells had the characteristic rounded shape with multiple filopodial-like extensions (Figure 2F) (All et al., 2015). In short, we found that the cells underwent morphologic changes consistent with the transition from the undifferentiated pluripotent stage into oligodendrocyte precursor cells.

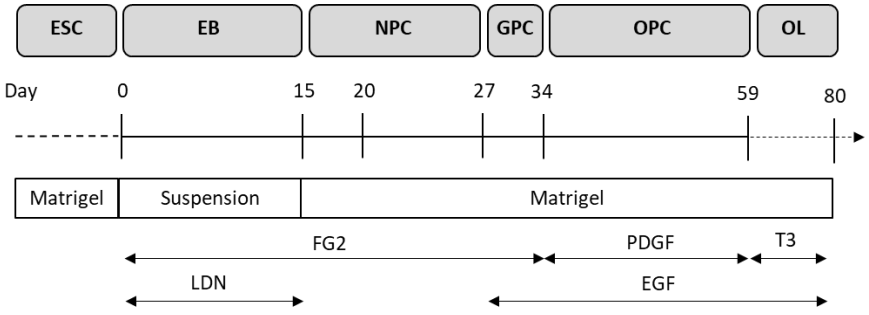


Figure 1. Timeline of oligodendrocyte differentiation based on All et al., 2015. Time is given in days, adherent of suspension culture is indicated as well as coating used, growth factors added to the medium (daily) are also pointed out. Starting with embryoid bodies (EB) formation at day 0, followed by the neural precursor cells (NPC) expansion. Then, glial precursor cell (GPC) are induced, and subsequent differentiated into oligodendrocyte precursor cells (OPC). LDN, LDN193189; FGF 2, fibroblast growth factor 2; EGF, Epidermal growth factor; PDGF, platelet-derived growth factor; T3, thriiodothytonine.

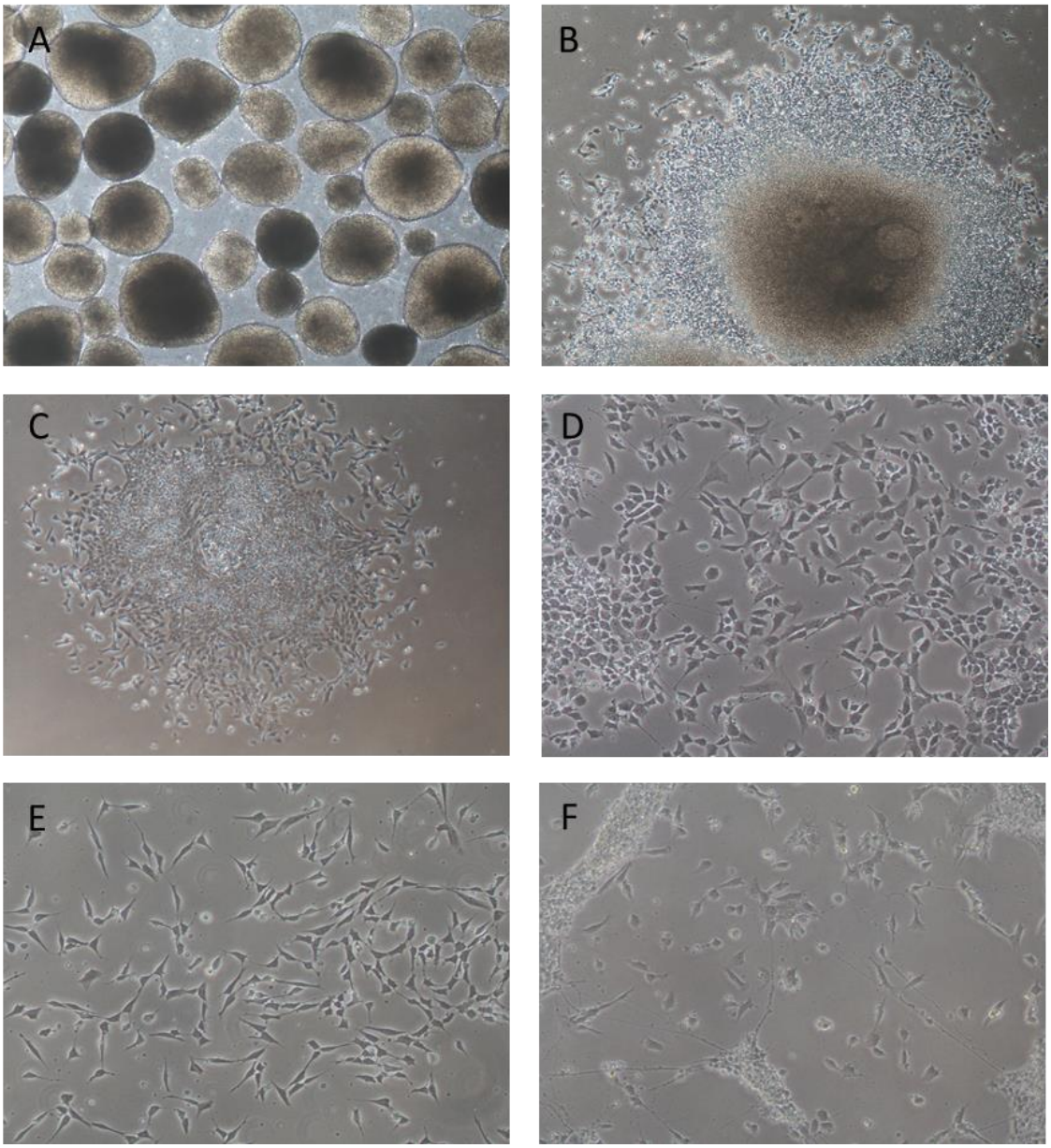


Figure 2. Bright field microscopy images of the different stages of oligodendrocyte differentiation. Representative pictures of the following intermediates are depicted: Embryoid bodies (EBs) (A). Neural precursors cells (NPCs) (B, C, D). Glial precursors cells (GPCs) (E), oligodendrocyte progenitors (OPCs) (F). Images A-C were taken using magnification 5X, while images D-F were taken using magnification 10X.

The proper progression from a pluripotent state into OPCs was also determined by gene expression and flow cytometry analyses (Figure 3 and 4). We found that pluripotency marker *REX1* expression was increased in both ES[4] cells and at the EBs stage, but was undetectable in NPCs, GPCs and OPCs (Figure 3). On the contrary, NPCs and GPCs exhibited upregulated markers of early neural and glial differentiation respectively. For instance, *SOX10* was expressed in NPCs while *OLIG2* expression was increased in GPCs (Figure 3). Final OPCs identity was confirmed by the pronounced

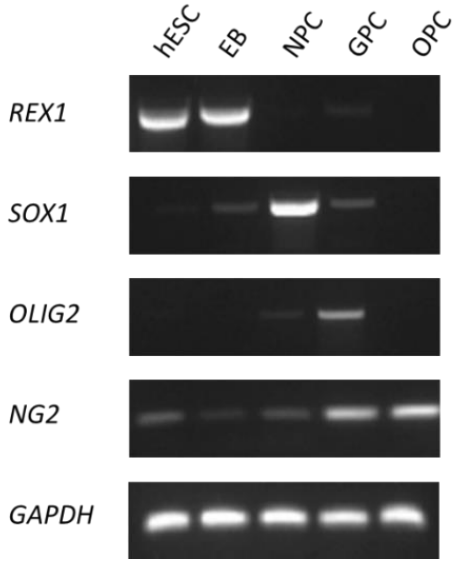


Figure 3. mRNA expression of stage-specific genes during oligodendrocyte differentiation of hESC. Semi-quantitative RT-PCR analysis indicates gene expression of the markers indicated specific for pluripotency (*REX1*), neural precursors (*SOX1*), glial precursors (*OLIG2*) and oligodendrocyte precursors (*NG2*) at various steps during the differentiation process outlined in Figure 2. Stages analyzed: undifferentiated human embryonic stem cells (hESC); embryoid bodies (EB); neural precursors cells (NPC); glial precursors cells (GPC) and oligodendrocyte progenitors (OPC).

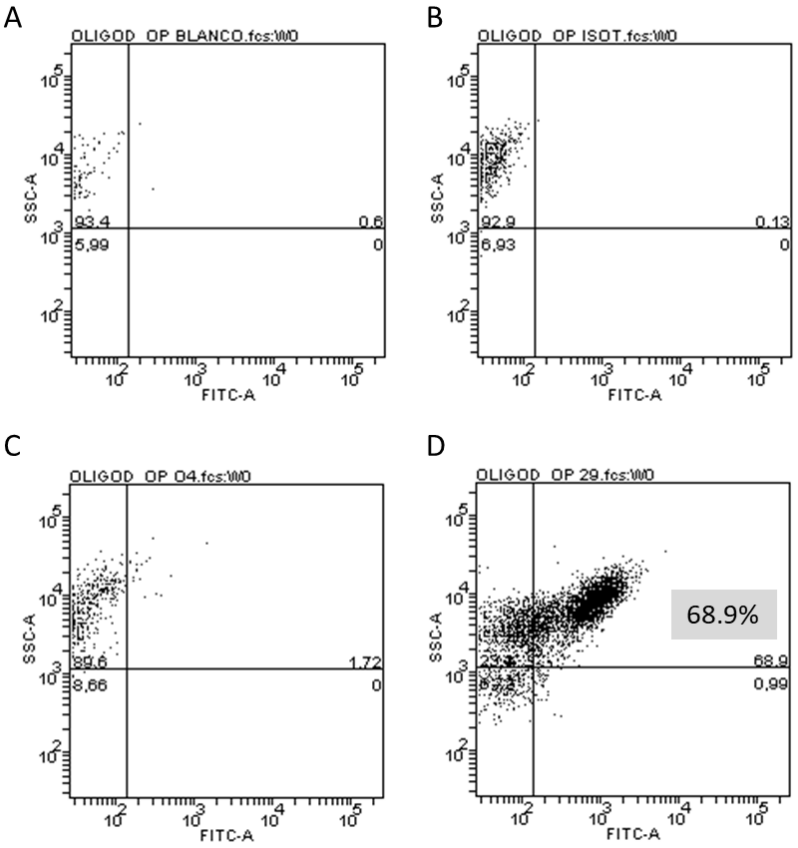


Figure 4. Flow cytometry analysis of oligodendrocyte precursors marker O4 in OPCs. Flow cytometry dot plots showed the amount of oligodendrocyte precursor cells expressing surface antigen marker beta1 integrin (CD29) (D) and oligodendrocyte precursors marker O4 (C). Gating and data analysis were adjusted according to an unstained control (A) and an isotype control (B).

NG2 expression, a marker indicative of oligodendrocyte precursor cells (Figure 3). However, flow cytometry analysis revealed that OPCs failed to be oligodendrocyte precursors marker O4 positive cells (Figure 4C). Therefore, morphological characterization, marker gene expression analysis and flow cytometry analysis confirmed that we obtained OPCs but not mature oligodendrocyte.

While this work was in progress, an improved method to generate more mature O4-positive cells was published (García-León et al., 2018). We aimed to replicate this protocol, which ensured the generation of a $O4^+$ population within a month, thanks to the lentivirus-mediated overexpression of the transcription factor *SOX10* (Figure 5). Firstly, NPCs were induced. NPCs identity was confirmed by the expecting morphology, a homogenous population of closely packed NPC which exhibited a neuroepithelium morphology (Figure 6A). Also, neuronal progenitor markers, such as *PAX6* and *OLIG2,* expression was increased in NPCs, while pluripotency marker *OCT4* expression was decreased in NPC stage compared to ES[4] (Figure 6B).

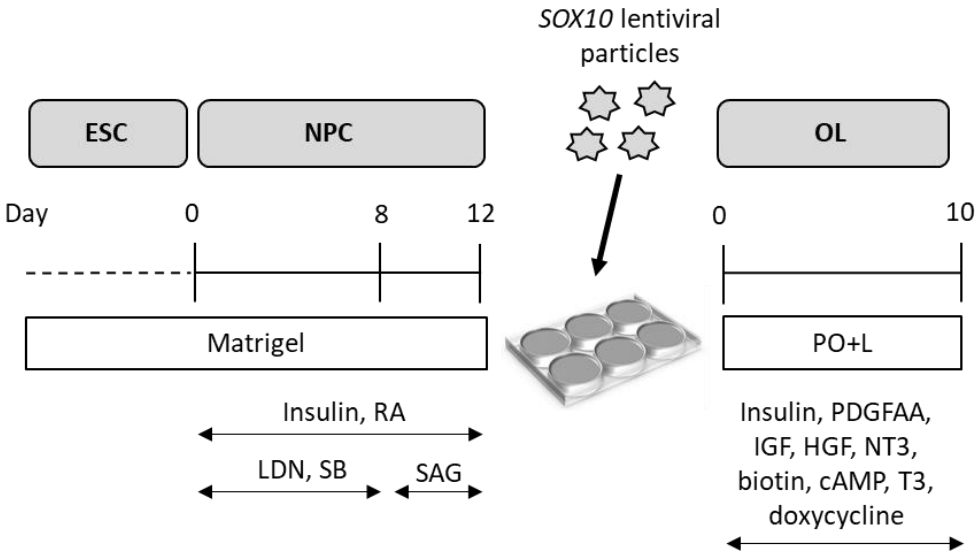


Figure 5. Timeline of oligodendrocyte differentiation based on García-León et al., 2018. Time is given in days, starting with neural precursor cells (NPC) differentiation at day 0. At day 12, cells were transfected with lentiviral particles with *SOX 10* gene cloned. Following the successful transfection, cells were differentiated into oligodendrocyte (OL). RA, retinoic acid; LDN, LDN193189; SB, SB431542, SAG, Smoothened Agonist; PDGF, platelet-derived growth factor; IGF, insulin-like growth factor 1; HGF, hepatocyte growth factor; NT3, neurotrophin-3; cAMP, adenosine cyclic monophosphate; T3, thriiodothytonine.

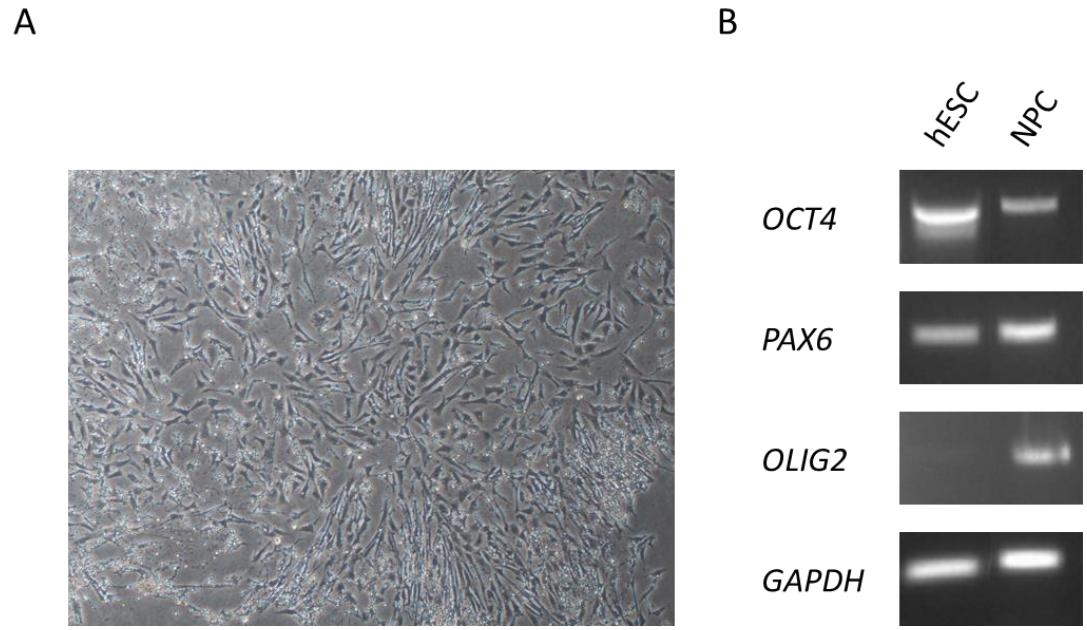


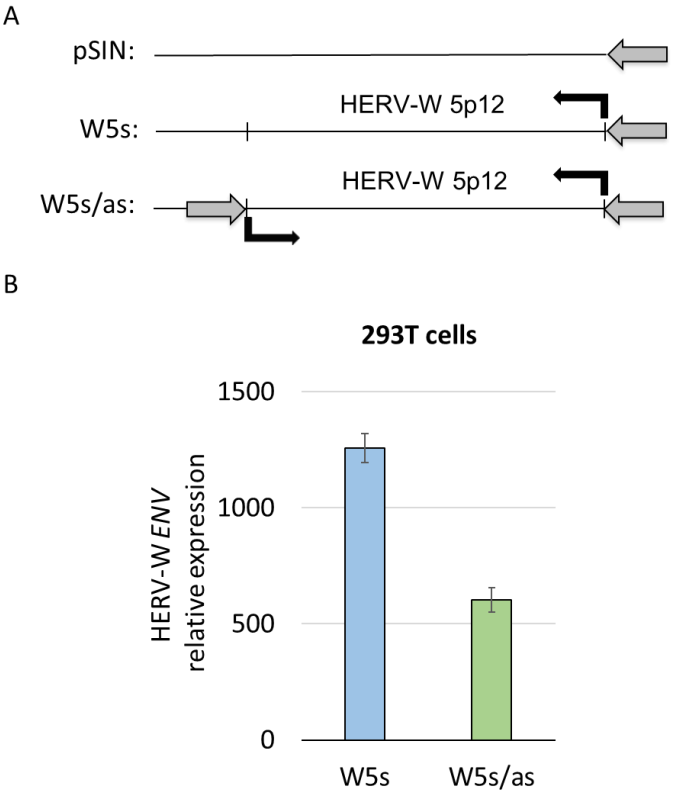
Figure 6. Bright field microscopy image of neural precursor cells obtained using the protocol described by García-León et al., 2018, and characterized by semi-quantitative RT-PCR. A) Human neural precursor cells (NPC) after 8 days of differentiation were closely packed, had a distinctive morphology, and the culture appeared to be homogeneous. Image was taken using magnification 5X. B) Analysis of mRNA expression shows different expression at different time points of markers specific for pluripotency (*OCT4*) and neural precursors (*PAX6, OLIG2*) markers through the steps of differentiation: human embryonic stem cells (hESC) and neural precursors cells (NPC) at day 5 and 11 of differentiation.

However, transduction of NPCs with lentiviral particles containing the FUW-SOX10 vector could not be performed due to ES[4] cell line-specific adhesion problem. ES[4] cell line inability to adhere to matrigel coating after being lifted precluded the successful replication of the differentiation protocol. Hence, $O4^+$ oligodendrocyte cells could not be generated. And at this point, we decided to continue pursuing the aim of this study with cells in the previous differentiation stag: NPCs.

Induction of HERV-W ENV RNA expression triggers a type I interferon response via TLR3 signalling

We next intended to study whether HERV-W overexpression in NPCs could trigger an inflammatory immune response. First, to induce the overexpression of HERV-W *ENV* 5p12, a genome DNA partial sequence of HERV-W *ENV* 5p12 was cloned into the pSIN-EF2 vector and linked to EF1a promoter or PGK promoter (Figure s2). To prove promoters' activity, 293T cells were independently transfected with either pSIN-EF2 (pSIN), pSIN-EF2-HERV-W 5p12 (W5s) or pSIN-EF2-PGK-HERV-W 5p12 (W5s/as) vectors (Figure 7A). After transfection, HERV-W *ENV* expression levels were assessed using a

specific qPCR for MSRV *ENV* (Mameli et al., 2009). We found a 1258-fold induction of HERV-W *ENV* expression in cells transfected with the W5s vector, and a 602-fold induction in cells transfected with the W5s/as vector, compared to HERV-W *ENV* expression in 293T cells transfected with the pSIN vector (Figure 7B). Promoter's activity from both plasmid constructions was therefore validated to induce HERV-W *ENV* overexpression.

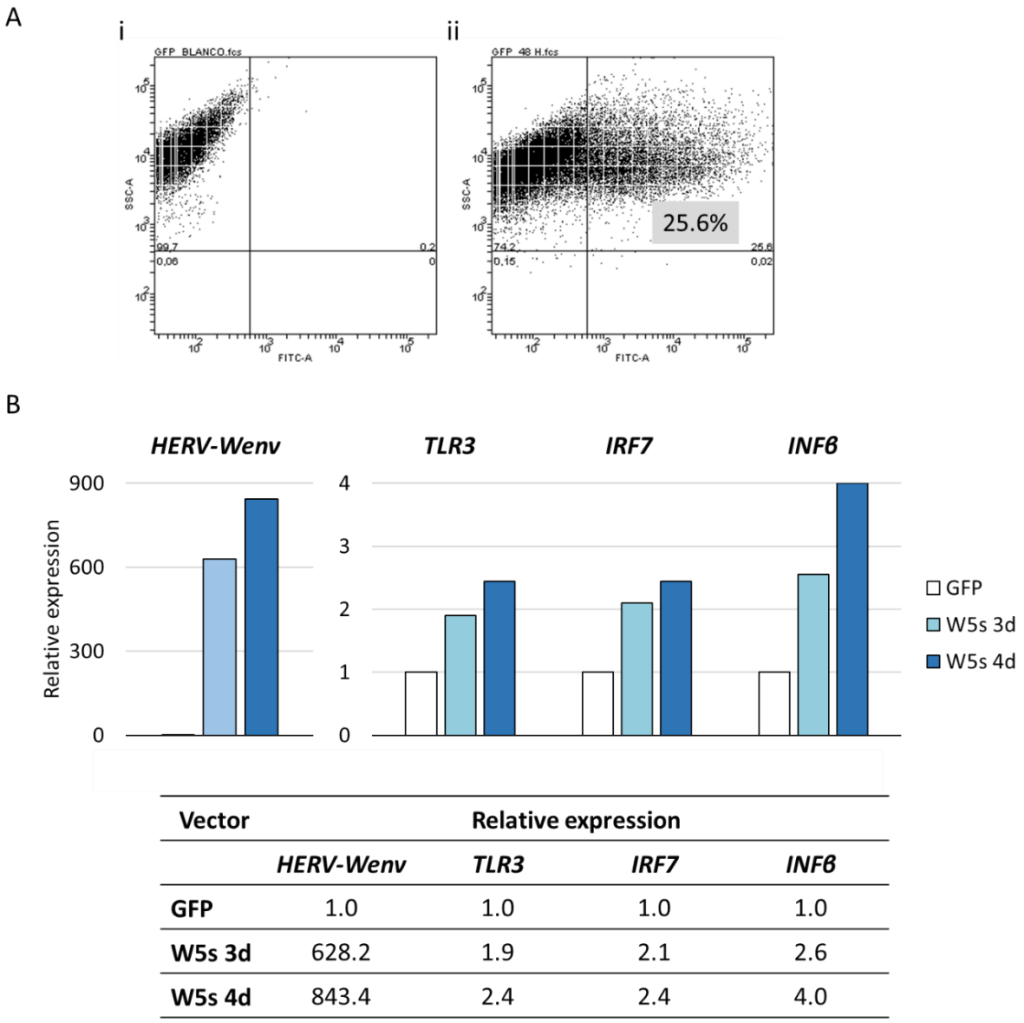


HERV-W *ENV* relative expression

Vector	N	Mean (SD)
pSIN	2	1 (0)
W5s	2	1258 (63)
W5s/as	2	602 (53)

Figure 7. Transfection of 293T cells. A) Graphical representation of the general structure of the vectors used in this experiment: pSIN-EF2 vector (pSIN); pSIN-EF2-HERV-W 5p12 vector (W5s) and pSIN-EF2-HERV-W 5p12-KGPp vector (W5s/as). Green arrows show promoters, while black arrows indicate direction of transcription. B) Quantification of HERV-W *ENV* expression in transfected 293T cells. Vectors were individually transfected in 293T cells in duplicate. After 48 hours, cells were harvested and analyzed for HERV-W *ENV* expression. Results are represented as mean fold changes compared to the mean expression in 293T cells transfected with pSIN vector. Biological replicates are indicated (N). The table shows the numerical values of the graph.

Subsequently, lentiviral particles containing W5s, W5s/as and pLOX-CWgfp (GFP) vectors were generated. NPCs were first independently transduced with lentiviral particles containing W5s and GPF vectors, and HERV-W *ENV* expression levels were subsequently assessed using the same assay described above (Mameli et al., 2009). Although only 25.6% of the NPCs were efficiently transduced (Figure 8A), 3 and 4 days after transduction, W5s-transduced NPCs showed a 628.2-fold and an 843.4-fold increased HERV-W *ENV* expression, respectively, compared to GFP-transduced NPCs (Figure 8B).



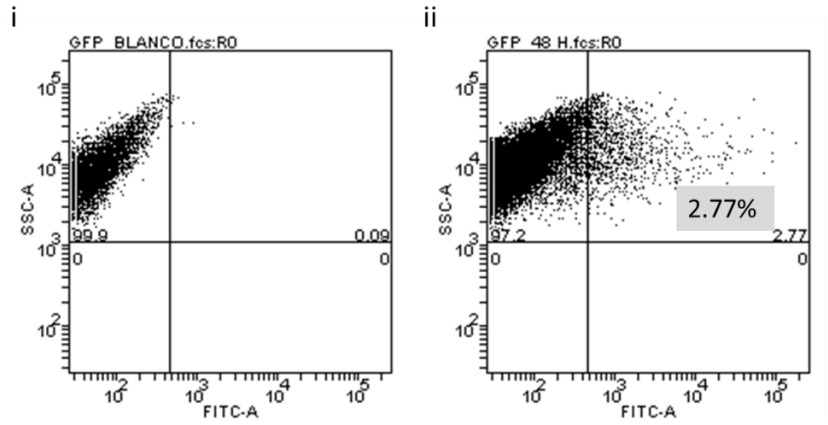
Vector	Relative expression			
	HERV-Wenv	***TLR3***	***IRF7***	***INFß***
GFP	1.0	1.0	1.0	1.0
W5s 3d	628.2	1.9	2.1	2.6
W5s 4d	843.4	2.4	2.4	4.0

Figure 8. Transduction of neural progenitor cells (NPCs). hESCs-derived NPCs were obtained using a published protocol (García-León et al., 2018). In a single experiment, NPCs were then independently transduced with W5s, W5s/as and GFP vector (Figure 7A). A) Flow cytometry dot plot shows the percentage of NPCs expressing GFP (ii). Gating and data analysis were adjusted according to non-transduced cells (i). B) NPCs were harvested 3 (3d) or 4 days (4d) after transfection and expression levels of indicated targets were analysed by qPCR. Results are represented as fold changes compared to the GFP-transfected NPCs. The table shows the numerical values of the graph.

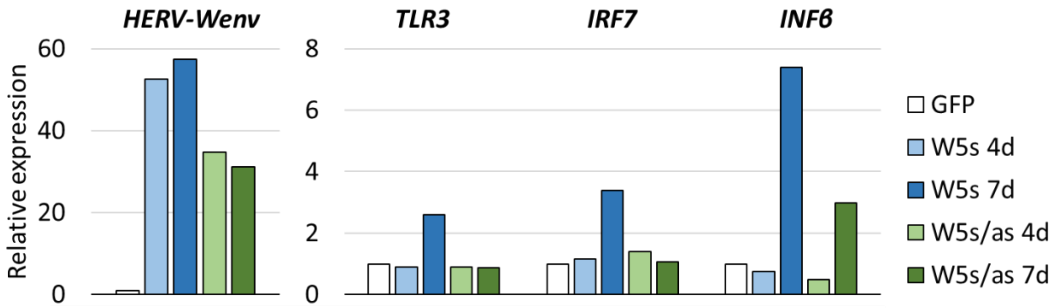
Once the efficient lentivirus-mediated HERV-W *ENV* overexpression in NPCs was confirmed, the possible immune system activation was analysed by assessing *TLR3, IRF7* and *IFN-β* expression levels in the W5s-transduced NPCs. 3 days after transduction, *TLR3, IRF7* and *IFN-β* expression levels were increased 1.9, 2.1 and 2.6 times, respectively, compared to GFP-transduced NPCs (Figure 8B). While 4 days after transduction, expression levels increased up to 2.4, 2.4 and 4 times, respectively (Figure 8B).

Transfection and expression assays were repeated, this time also transducing NPCs with lentiviral particles containing W5s/as. Transduction efficiency was much lower this time at 2.77% (Figure 9A). However, 4 and 7 days after transduction, W5s-transduced NPCs showed a 52.5-fold and a 57.5-fold increased MSRV-like HERV-W expression, respectively, when compared to GFP-transduced NPCs (Figure 9B). And 4 and 7 days after transduction, W5s/as-transduced NPCs showed a 34.8-fold and a 31.2-fold increased MSRV-like HERV-W expression, respectively, when compared to GFP-transduced NPCs (Figure 9B). We observed that HERV-W *ENV* induction was directly proportional to transduction efficiency (%) (Figure 10). 4 days after transduction, *TLR3, IRF7* and *IFN-β* expression was increased 0.9, 1.1 and 0.8 times, respectively, in W5s-transduced NPCs, while 7 days after transduction, *TLR3, IRF7* and *IFN-β* expression increased up to 2.6, 3.4 and 7.4 times, respectively (Figure 9B). In parallel, 4 days after transduction, *TLR3, IRF7* and *IFN-β* expression was increased 0.9, 1.4 and 0.5 times, respectively, in W5s/as-transduced NPCs, while 7 days after transduction, *TLR3, IRF7* and *IFN-β* expression increased up to 0.9, 1.1 and 3 times, respectively (Figure 9B). HERV-W overexpression in NPCs was therefore enough stimulus to trigger a type I interferon immune response possibly via activation of TLR3 and IRF7.

A



B



Vector	Relative expression			
	HERV-Wenv	***TLR3***	***IRF7***	***INFβ***
GFP	1.0	1.0	1.0	1.0
W5s 4d	52.5	0.9	1.1	0.8
W5s 7d	57.5	2.6	3.4	7.4
W5s/as 4d	34.8	0.9	1.4	0.5
W5s/as 7d	31.2	0.9	1.1	3.0

Figure 9. Transduction of neural progenitor cells (NPCs). hESCs-derived NPCs were obtained using a published protocol (García-León et al., 2018). In a single experiment, NPCs were then independently transduced with W5s, W5s/as and GPF vectors (Figure 7A). A) Flow cytometry dot plot shows the percentage of NPCs expressing GFP (ii). Gating and data analysis were adjusted according to non-transduced cells (i). B) NPCs were harvested and expression levels of indicated targets were analysed 4 (4d) and 7 days (7d) after transfection. Results are represented as fold changes compared to the GFP-transfected NPCs. The table shows the numerical values of the graph.

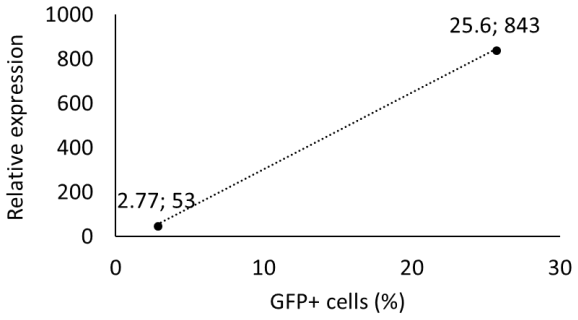


Figure 10. Relationship between the efficacy of transduction of NPCs with pSIN-EF2-HERV-W 5p12 vector (W5s), and the induction of HERV-W *ENV* expression 4 days after transduction. The graph shows a direct relation between the number of efficiently transduced NPCs and the fold-change induction of HERV-W *ENV* expression. Efficacy of transduction was defined as the percentage of GFP^+ NPCs achieved after transduction (Figure 8A and 9A). Relative HERV-W *ENV* expression is expressed as a fold change compared to the expression measured in NPCs transduced with pLOX-CWgfp (GFP) vector (Figure 8B and 9B). Numbers within the graph indicate de x-value and y-value for each point.

Discussion

Although the increased presence and expression of HERV-W elements has been repeatedly associated with MS, little is known about HERV-W contribution to MS pathogenesis. In the present study, we show that HERV-W overexpression activates IRF7 and induces IFNβ upon TLR3 stimulation in NPCs. It is important to highlight the absence of ORFs over 30 aminoacids long in the RNA produced, indicating that it is the RNA itself and not the proteins it encodes that cause immune activation. Overexpression of HERV-W-derived RNAs, therefore, can activate the innate immune system and trigger a local inflammatory response. We therefore hypothesise that abnormal HERV-W expression in MS patients may derive in a similar HERV-W-mediated inflammatory response, which subsequently may contribute to MS pathogenesis. Pending further confirmations showing immune activation in more mature cells and absence of protein products in the RNAs produced, our results support the idea that immune reactions within the CNS, apart from the peripherical immune reactions, may contribute to MS pathogenesis.

Although the most widely accepted hypothesis suggests that MS onset is characterised by the activation of the peripheral immune system (Garg and Smith, 2015; Sospedra and Martin, 2005), local immune response undeniably contributes to MS pathogenesis as well (Hemmer et al., 2015; Macchi et al., 2015). Interestingly, different patterns of demyelination have been found within cerebral lesions of autopsied MS patients

(Lucchinetti et al., 2000). While some studies have mainly observed high amounts of T cells within cerebral MS lesions (Chitnis, 2007; Kivisäkk et al., 2004), other cerebral lesions have been characterised by oligodendrocyte apoptosis and microglial activation with little presence of lymphocytes (Barnett and Prineas, 2004). These latest findings suggest that oligodendrocyte damage and the subsequent local immune response activation may happen before the massive lymphocyte infiltration. In fact, in an animal model, the induced apoptosis of oligodendrocytes resulted in demyelination and microglia activation in the absence of peripherical lymphocytes (Caprariello et al., 2012).

Considering all the above, we hypothesise that one possible mechanism underlying oligodendrocyte damage in the absence of peripherical inflammation could be a HERV-W-mediated local inflammation. Immunohistochemical analysis of autopsied brains of MS patients has revealed enhanced expression of TLR3 and TLR4 in inflamed lesions (Bsibsi et al., 2002). In an mouse model, the induced TLR3 stimulation in glial cells resulted in microglia activation and pre-oligodendrocyte cell death (Steelman and Li, 2011). Likewise, also in a mouse model, induced TLR3 stimulation in neurons resulted in inhibition of axonal growth (Cameron et al., 2007). All these findings showed that TLR3 activation has the potential to cause CNS degeneration. Since TLR3 is expressed in many of the human CNS cell types such as oligodendrocytes, astrocytes and microglia (Bsibsi et al., 2002), HERV-W-mediated TLR3 stimulation may also trigger a detrimental local inflammatory response within the CNS of MS patients.

In vitro differentiation from hESCs and/or iPSCs into relevant cell types can be helpful to generate the cell types affected by diseases. We initially intended to generate *in vitro* hESCs-derived oligodendrocytes since they are one of the main CNS cell types affected in MS. However, two published protocols for differentiation turned out difficult to reproduce, and we finally analysed the effect of HERV-W overexpression in a precursor cell type for neural differentiation. The challenges we found for generating oligodendrocytes using the first protocol (All et al., 2015) were due to the length of culture time required (up to 2 months), and the low amount of OPCs obtained. The second protocol we followed (García-León et al., 2018), was promisingly shorter and more efficient. However, cell attachment problems prevented us from generating

oligodendrocytes. We therefore experienced the arduousness of generating differentiated oligodendrocytes *in vitro*. Although published protocols claim good efficiency of oligodendrocyte production, the reality is that protocols are difficult to replicate and may produce variable results dependent on the stem cell lines used.

Supplementary materials

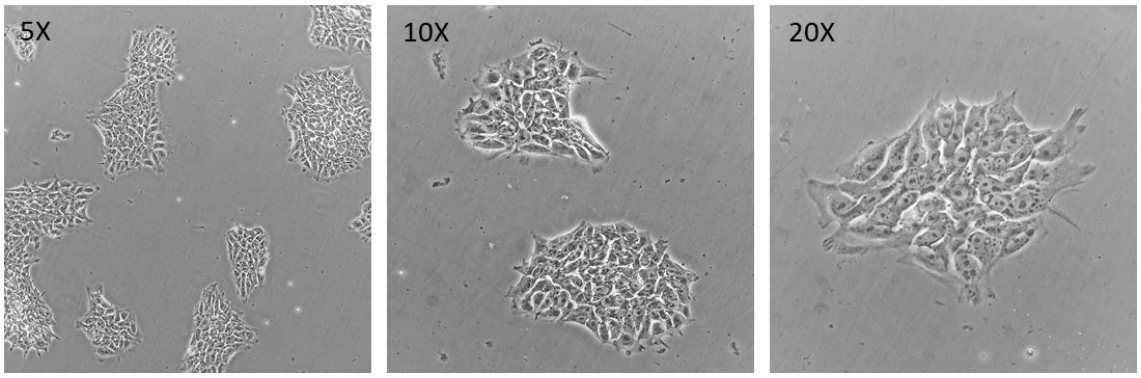


Figure s.1. Bright field microscopy images of [ES]4 cell line cultured on matrigel matrix in mTeSR1 medium. Images illustrate the typical morphology of hESCs. ES[4] cells are rounded, have a high nucleo-cytoplasmic ratio and prominent nucleoli. These cells are tightly packed with each other and form individual colonies. Images were taken using three magnifications: 5X, 10X and 20X.

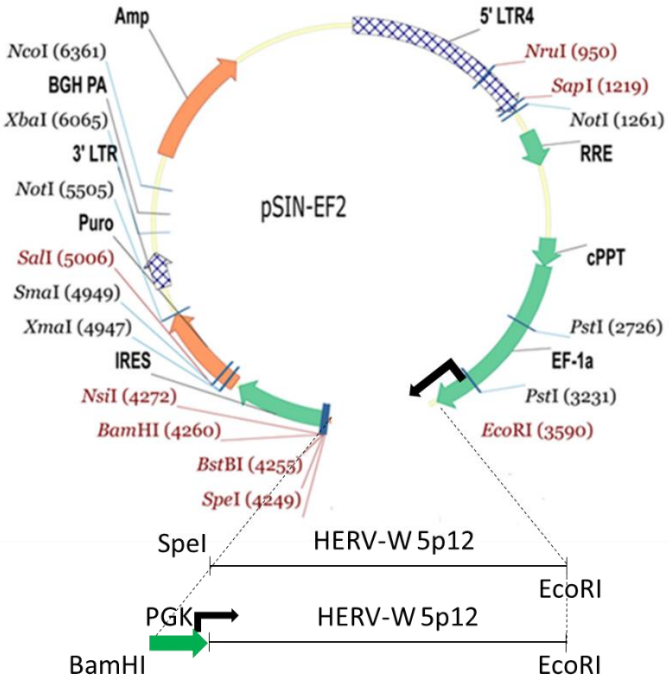


Figure s2. Plasmids construction. To induce unidirectional HERV-W expression, HERV-W copy located on chromosome 5p12 (HERV-W 5p12) was cloned and fused to pSIN-EF2 plasmid, which contains the constitutive EF-1α promoter. Bidirectional HERV-W expression was achieved by also fusing the constitutive phosphoglycerate kinase (PGK) promoter to pSIN-EF2 plasmid, so that PGK and EF-1α promoters control antisense and sense HERV-W 5p12 expression respectively. Green arrows: promoters. Black arrows: direction of transcription.

Gene	Sequence (5'-3')
REX1	For: ACATGACAGGCAAGAAGCTTCCGC Rev: GAGCCCGGATCCTACTTTCCCTCTTGTTCATTCTT
OCT4	For: CGACCATCTGCCGCTTTGAG Rev: CCCCCTGTCCCCCATTCCTA
PAX6	For: AATAACCTGCCTATGCAACCC Rev: AACTTGAACTGGAACTGACACAC
SOX1	For: CAATGCGGGGAGGAGAAGTC Rev: RCTCTGGACCAAACTGTGGCG
OLIG2	For: CCCTGAGGCTTTTCGGAGCG Rev: GCGGCTGTTGATCTTGAGACGC
NG2	For: CTTCACTCAGGCAGAGGTCTACGC Rev: GAGGACAGCTGGAGCTCTAGGGT
GAPDH	For: ATCAGCAATGCCTCCTGCAC Rev: TGGCATGGACTGTGGTCATG

Table s1. List of set of primers used for the analysis of phenotypic markers expression by semi-quantitative PCR. The first column shows the target, the second column indicates the oligo sequence (For= forward; Rev= reverse).

Target	Sequence (5'-3')
HERV-W 5p12	For: TGGAGCAGGAGTGCTAGGTAGA Rev: TGGTCGGGTGTGAGCTAAGT
PGK promoter	For: ATCTGCGGATCCTACCGGGTAGGGGAGGCGC Rev: ATCTGCTCTAGATTCGAAAGGCCCGGAGATGAG

Table s2. List of set of primers used for the plasmids construction and subsequent analysis. The first column shows the target, the second column indicates the oligo sequence (For= forward; Rev= reverse).

Gene	Sequence (5'-3')	1mM $MgCl_2$ addition
MSRV-like HERV-Wenv (Mameli et al., 2009)	For: CTTCCAGAATTGAAGCTGTAAAGC Rev: GGGTTGTGCAGTTGAGATTTCC	No
GAPDH	For: ATCAGCAATGCCTCCTGCAC Rev: TGGCATGGACTGTGGTCATG	No
IFN-β	For: AGTAGGCGACACTGTTCGTG Rev: GCCTCCCATTCAATTGCCAC	Yes
IRF7	For: CCCCCATCTTCGACTTCAGA Rev: CAGGACCAGGCTCTTCTCCTT	No
TLR3	For: GGTGTGTCTTCCCTGGATAG Rev: GCTCCAGCTCCATAAGGAAG	Yes

Table s3. List of set of primers used for the plasmids construction. The first column shows the target, the second column indicates the oligo sequence (For= forward; Rev= reverse).

References

All, A.H., Gharibani, P., Gupta, S., Bazley, F.A., Pashai, N., Chou, B.-K., Shah, S., Resar, L.M., Cheng, L., Gearhart, J.D., et al. (2015). Early Intervention for Spinal Cord Injury with Human Induced Pluripotent Stem Cells Oligodendrocyte Progenitors. PLOS ONE *10*, e0116933.

Barnett, M.H., and Prineas, J.W. (2004). Relapsing and remitting multiple sclerosis: pathology of the newly forming lesion. Ann. Neurol. *55*, 458–468.

Brück, W. (2005). The pathology of multiple sclerosis is the result of focal inflammatory demyelination with axonal damage. J. Neurol. *252*, v3–v9.

Brück, W., Lucchinetti, C., and Lassmann, H. (2002). The pathology of primary progressive multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *8*, 93–97.

Bsibsi, M., Ravid, R., Gveric, D., and van Noort, J.M. (2002). Broad expression of Toll-like receptors in the human central nervous system. J. Neuropathol. Exp. Neurol. *61*, 1013–1021.

Cameron, J.S., Alexopoulou, L., Sloane, J.A., DiBernardo, A.B., Ma, Y., Kosaras, B., Flavell, R., Strittmatter, S.M., Volpe, J., Sidman, R., et al. (2007). Toll-like receptor 3 is a potent negative regulator of axonal growth in mammals. J. Neurosci. Off. J. Soc. Neurosci. *27*, 13033–13041.

Caprariello, A.V., Mangla, S., Miller, R.H., and Selkirk, S.M. (2012). Apoptosis of oligodendrocytes in the central nervous system results in rapid focal demyelination. Ann. Neurol. *72*, 395–405.

Carpentier, P.A., Duncan, D.S., and Miller, S.D. (2008). Glial Toll-like Receptor Signaling in Central Nervous System Infection and Autoimmunity. Brain. Behav. Immun. *22*, 140–147.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., et al. (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974–986.

Chitnis, T. (2007). The role of CD4 T cells in the pathogenesis of multiple sclerosis. Int. Rev. Neurobiol. *79*, 43–72.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., Cirotto, M.C., Rosati, G., and Sotgiu, S. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology *58*, 471–473.

Douvaras, P., and Fossati, V. (2015). Generation and isolation of oligodendrocyte progenitor cells from human pluripotent stem cells. Nat. Protoc. *10*, 1143–1154.

Douvaras, P., Wang, J., Zimmer, M., Hanchuk, S., O'Bara, M.A., Sadiq, S., Sim, F.J., Goldman, J., and Fossati, V. (2014). Efficient Generation of Myelinating Oligodendrocytes from Primary Progressive Multiple Sclerosis Patients by Induced Pluripotent Stem Cells. Stem Cell Rep. *3*, 250–259.

Ehrlich, M., Mozafari, S., Glatza, M., Starost, L., Velychko, S., Hallmann, A.-L., Cui, Q.-L., Schambach, A., Kim, K.-P., Bachelin, C., et al. (2017). Rapid and efficient generation of oligodendrocytes from human induced pluripotent stem cells using transcription factors. Proc. Natl. Acad. Sci. U. S. A. *114*, E2243–E2252.

García-León, J.A., Kumar, M., Boon, R., Chau, D., One, J., Wolfs, E., Eggermont, K., Berckmans, P., Gunhanlar, N., de Vrij, F., et al. (2018). SOX10 Single Transcription Factor-Based Fast and Efficient Generation of Oligodendrocytes from Human Pluripotent Stem Cells. Stem Cell Rep. *10*, 655–672.

Garg, N., and Smith, T.W. (2015). An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. Brain Behav. *5*, n/a-n/a.

Garson, J.A., Tuke, P.W., Giraud, P., Paranhos-Baccala, G., and Perron, H. (1998). Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet Lond. Engl. *351*, 33.

Hemmer, B., Kerschensteiner, M., and Korn, T. (2015). Role of the innate and adaptive immune responses in the course of multiple sclerosis. Lancet Neurol. *14*, 406–419.

Hernández-Pedro, N.Y., Espinosa-Ramirez, G., Cruz, D.L., Pérez, V., Pineda, B., and Sotelo, J. (2013). Initial Immunopathogenesis of Multiple Sclerosis: Innate Immune Response.

Kielian, T. (2006). Toll-Like Receptors in Central Nervous System Glial Inflammation and Homeostasis. J. Neurosci. Res. *83*, 711–730.

Kivisäkk, P., Mahad, D.J., Callahan, M.K., Sikora, K., Trebst, C., Tucky, B., Wujek, J., Ravid, R., Staugaitis, S.M., Lassmann, H., et al. (2004). Expression of CCR7 in multiple sclerosis: implications for CNS immunity. Ann. Neurol. *55*, 627–638.

Koriem, K.M.M. (2016). Multiple sclerosis: New insights and trends. Asian Pac. J. Trop. Biomed. *6*, 429–440.

Lester, S.N., and Li, K. (2014). Toll-like receptors in antiviral innate immunity. J. Mol. Biol. *426*, 1246–1264.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif *25*, 402–408.

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., and Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. Ann. Neurol. *47*, 707–717.

Macchi, B., Marino-Merlo, F., Nocentini, U., Pisani, V., Cuzzocrea, S., Grelli, S., and Mastino, A. (2015). Role of inflammation and apoptosis in multiple sclerosis: Comparative analysis between the periphery and the central nervous system. J. Neuroimmunol. *287*, 80–87.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., Serra, C., Sotgiu, S., Bonetti, B., and Dolei, A. (2007). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J. Gen. Virol. *88*, 264–274.

Mameli, G., Poddighe, L., Astone, V., Delogu, G., Arru, G., Sotgiu, S., Serra, C., and Dolei, A. (2009). 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*, 98–106.

Nylander, A., and Hafler, D.A. (2012). Multiple sclerosis. J. Clin. Invest. *122*, 1180–1188.

Rolland, A., Jouvin-Marche, E., Viret, C., Faure, M., Perron, H., and Marche, P.N. (2006). The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. J. Immunol. Baltim. Md 1950 *176*, 7636–7644.

Sospedra, M., and Martin, R. (2005). Immunology of multiple sclerosis. Annu. Rev. Immunol. *23*, 683–747.

Stacpoole, S.R.L., Spitzer, S., Bilican, B., Compston, A., Karadottir, R., Chandran, S., and Franklin, R.J.M. (2013). High Yields of Oligodendrocyte Lineage Cells from Human Embryonic Stem Cells at Physiological Oxygen Tensions for Evaluation of Translational Biology. Stem Cell Rep. *1*, 437–450.

Steelman, A.J., and Li, J. (2011). Poly(I:C) promotes TNFα/TNFR1-dependent oligodendrocyte death in mixed glial cultures. J. Neuroinflammation *8*, 89.

Trapp, B.D., Peterson, J., Ransohoff, R.M., Rudick, R., Mörk, S., and Bö, L. (1998). Axonal transection in the lesions of multiple sclerosis. N. Engl. J. Med. *338*, 278–285.

Trudler, D., Farfara, D., and Frenkel, D. (2010). Toll-Like Receptors Expression and Signaling in Glia Cells in Neuro-Amyloidogenic Diseases: Towards Future Therapeutic Application.

Xagorari, A., and Chlichlia, K. (2008). Toll-Like Receptors and Viruses: Induction of Innate Antiviral Immune Responses. Open Microbiol. J. *2*, 49–59.

V. DISCUSSION

HERVs and autoimmunity

Innate immune detection of nucleic acids (DNA and RNA) is crucial for the host defence against viral infection (Stetson et al., 2008; Theofilopoulos et al., 2011). In mammals, foreign nucleic acids can be recognised by two kind of intracellular receptors: TLRs and cytoplasmic nucleic acid sensors. TLRs are transmembrane sensors expressed in a cell type specific way in immune cells and other cells types including oligodendrocytes and neurons (Carpentier et al., 2008; Kielian, 2006; Trudler et al., 2010). Within the TLRs family, TLR3, 7, 8 and 9 are nucleic acid-sensing TLRs. TLR3 is activated by dsRNA, TLR7 and 8 detect ssRNA, and TLR9 senses ssDNA that contains unmethylated CpG sequences. All these nucleic acid-sensing TLRs are localized within intracellular endosomal compartments, in which they detect the presence of nucleic acids in compartments where endogenous nucleic acids would not normally reside (Ewald and Barton, 2011; Pelka et al., 2016). In contrast, cytoplasmic nucleic acid sensors are more broadly expressed and trigger a cell-intrinsic antiviral response. These cytosolic sensors are composed of an ample variety of receptors including RIG-I, MDA5 and LGP2 (Theofilopoulos et al., 2011). Activation of either of these two kinds of receptors leads to a type I INF response, a group of cytokines with antiviral activity (Stetson et al., 2008; Theofilopoulos et al., 2011). Although these antiviral responses are usually beneficial, if excessive, they can result in pathogenic inflammatory and autoimmune diseases (Assmann et al., 2015; Barrat et al., 2005; Christensen et al., 2006; Fischer and Ehlers, 2008; Sakata et al., 2018; Theofilopoulos et al., 2011).

One of the main concerns in nucleic acid recognition is the discrimination between viral and self-nucleic acids. Since HERVs are part of the genome, HERV-derived nucleic acids should theoretically be recognised as self-nucleic acids. However, a limited number of situations have been described in which HERV-derived nucleic acids activate the innate immune system (Chiappinelli et al., 2015; Hurst and Magiorkinis, 2015). Expression of HERV elements has been reported in several autoimmune diseases (Gröger and Cynis, 2018; Lättekivi et al., 2018; Morandi et al., 2017). Our data also indicate that HERV-W expression may be slightly induced in MS patients (Manuscript II). In parallel, we demonstrate that the number of transcribed HERV-W *loci* is larger in PBMCs of MS patients than in controls (Manuscript II). Altogether, our findings provide

evidence of the HERV-W RNA accumulation in MS. Combining HERV-induced activation of innate immunity and upregulated expression of HERV in autoimmune disease, a contribution of HERV to autoimmunity has been hypothesised (Hurst and Magiorkinis, 2015; Stetson, 2012).

Independent of the molecular origin of MSRV, the relatedness of MSRV to HERV-W, the copies of HERV-W involved and even the exact levels of accumulation, we aimed to directly test whether HERV-W-derived RNA may activate the innate immune system. Overexpressing HERV-W RNA to high levels in NPCs, we show that HERV-W RNA triggers a type I IFN response in NPCs, potentially via TLR3 signalling (Manuscript IV). We therefore suggest that an IFN-mediated antiviral response mediated by HERV-W transcripts could contribute to MS. This may apply especially to individuals genetically and environmentally predisposed to develop autoimmunity.

In a mouse model for Aicardi-Goutières syndrome (AGS), a HERV-driven INF response was independent of lymphocytes (Gall et al., 2012). *Trex1* is a ssDNA-specific exonuclease whose loss resulted in the accumulation of cytosolic DNA, including HERVs-derived DNA (Stetson et al., 2008). In *Trex1*-deficient mice, this cytosolic DNA accumulation led to a lethal and type I INF-dependent autoimmune response (Stetson et al., 2008). Interestingly, the INF response that drives disease in *Trex1*-deficient mice was found to be independent of lymphocytes (Gall et al., 2012), revealing a non-hematopoietic origin of the antiviral response**.** We provide evidence that overexpressed HERV-W RNA triggers a type I IFN response in NPCs (manuscript IV). As opposed to a DNA-dependent INF response in the mouse autoimmunity model (Gall et al., 2012), an RNA-dependent INF response may contribute to MS. While the event triggering the response in MS may be a local response in the CNS, HERV-W expression could as well activate cell-intrinsic antiviral sensors outside the CNS. Similar to the *Trex1*-deficient mouse model for AGS, the initial type I INF response may subsequently activate lymphocytes in the periphery, which finally would target the CNS.

Of course additional research is required to clarify the potential contribution of HERV expression to autoimmune diseases in general and MS in particular. Our initial experiments forcing HERV-W overexpression in NPCs, provide a framework for future experiments. Immune reactions can be compared between astrocytes, neurons and

oligodendrocytes all derived from stem cells *in vitro*. In subsequent studies patient-derived iPSC can be compared to control cells, and the effect of activated microglia on particular cell types is of interest. The use of cellular models may allow the dissection of immune responses elicited by HERV, and their contribution to MS.

VI. CONCLUSIONS/ CONCLUSIONES

Conclusions

1. PCR-based assays have been developed that allow for the identification of novel HERV-W insertions in the human genome.

2. Application of this technique to a limited set of samples has not produced evidence for neither MSRV nor MS-associated HERV-W elements within the human genome.

3. MSRV/HERV-W copy number has been shown to be similar in PBMCs from MS patients and controls.

4. MSRV/HERV-W expression has been found to be slightly increased in PBMCs from MS patients.

5. It appears unlikely that aberrant expression of a unique HERV-W element causes overexpression of HERV-W in MS, since the most abundantly transcribed HERV-W *loci* are the same in PBMCs of MS patients and controls.

6. The slight difference in MSRV/HERV-W expression between MS patients and controls may arise from the sum of HERV-W elements transcribed, which, although transcribed to a lesser extent, are expressed at higher levels in MS. These HERV-W elements may contribute to MS.

7. The *ENV* region of HERV-W elements are heavily methylated in PBMCs from both MS patients and controls or in human PBMCs, independent of MS. The observed overexpression of HERV-W *ENV* in MS patients, is therefore unlikely to be regulated by DNA methylation.

8. It is been observed that HERV-W overexpression triggers an INF-β response in hESCs-derived NPCs, which resembles a TLR3-mediated antiviral response.

Conclusiones

1. Se ha desarrollado una metodología, basada en ensayos de PCR, que permite la identificación de nuevas inserciones de HERV-W en el genoma humano.

2. El uso de esta técnica en un número limitado de muestras no ha permitido identificar MSRV o copies de HERV-W asociadas a MS en el genoma humano.

3. Se ha detectado que el número de copias de MSRV/HERV-W en el genoma es similar en PBMCs de pacientes de MS y controles.

4. Se ha detectado que los niveles de expresión de MSRV/HERV-W a nivel de RNA se encuentran ligeramente elevados en las PBMCs de pacientes de MS.

5. No parece probable que la sobreexpresión de un único elemento de HERV-W cause la sobreexpresión de HERV-W detectada en pacientes de MS, ya que los *loci* de HERV-W que se transcriben más abundantemente son los mimos en PBMCs de pacientes de MS y controles.

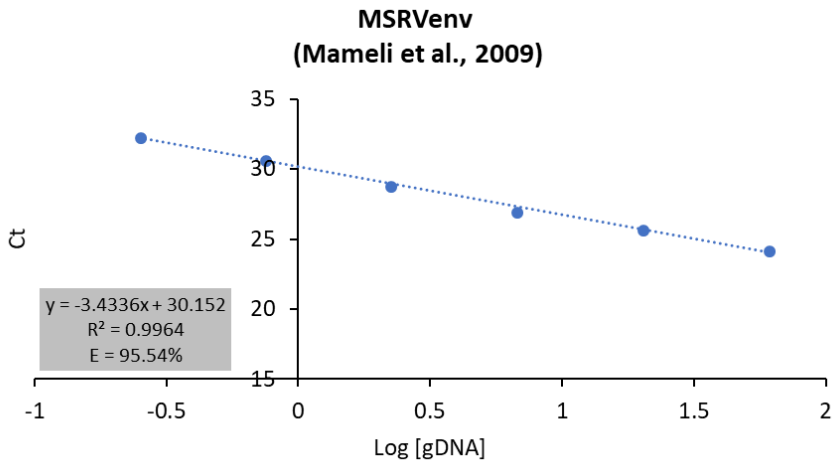
6. El ligero incremento en expresión de MSRV/HERV-W en PBMCs de pacientes de MS podría provenir de la suma de *loci* de HERV-W que, aunque transcritos en menor medida, se expresan más en pacientes de MS. Estos elementos HERV-W podrían contribuir a MS.

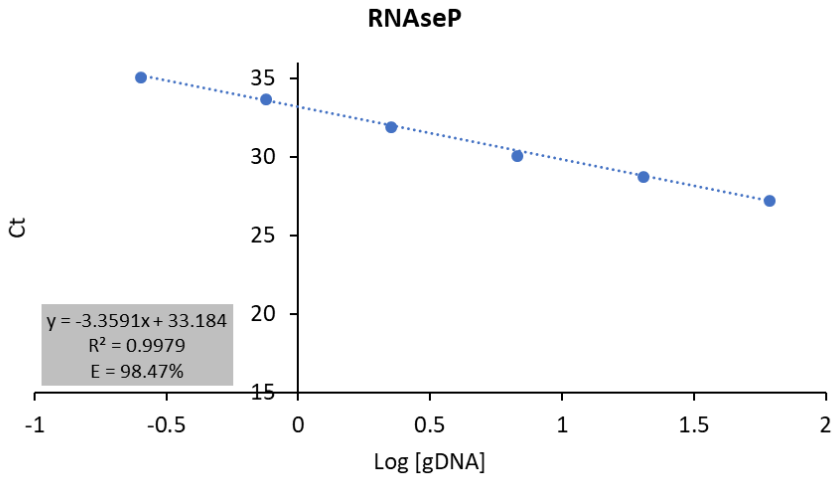
7. La región *ENV* de los elementos HERV-W se encuentra muy metilada en las PBMCs de pacientes de MS y de controles, y en PBMCs humanas no relacionadas con EM. Por lo tanto, la sobreexpresión de HERV-W *ENV* en pacientes de MS no parece estar regulada por la metilación del DNA.

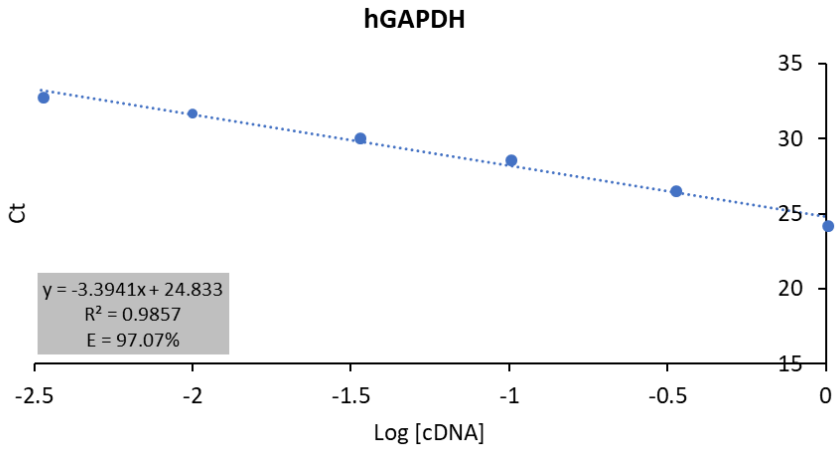
8. Se ha observado que la sobreexpresión de HERV-W desencadena la producción de INF-β en NPCs derivados de hESCs, la cual se asemeja a una respuesta antiviral mediada por TLR3.

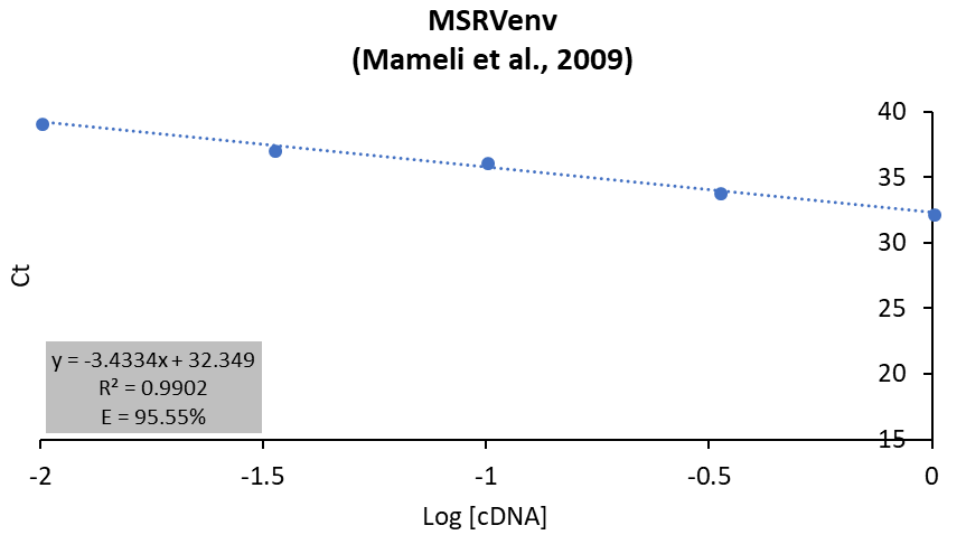
VII. APPENDIX I

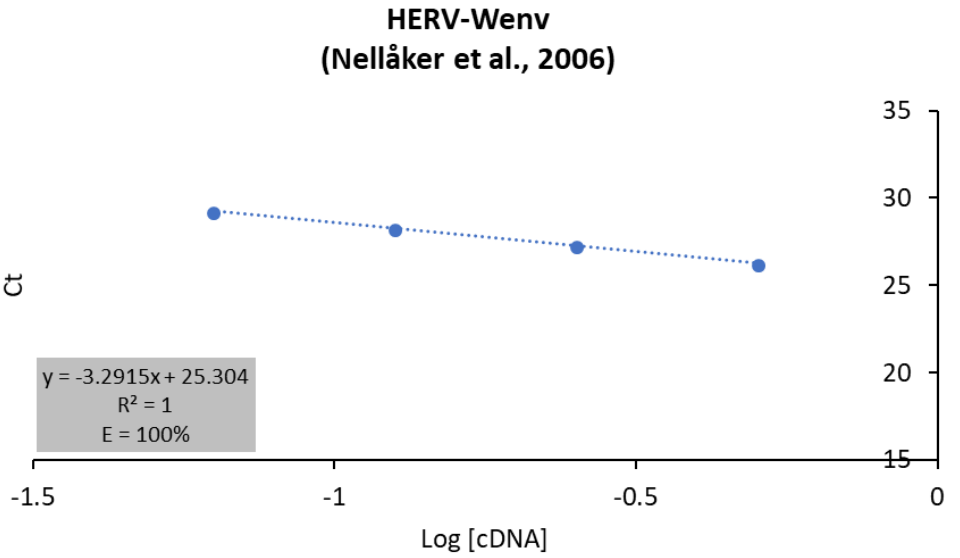
qPCR primer efficiency standard curve analysis. A qPCR standard curve is graphically represented as a semi-log regression line plot of Ct value versus log of input gDNA or cDNA. The efficiency (E) is calculated with the formula $E=10^{(-1/slope)}-1$. Thus, a slope of -3.32 indicates a PCR reaction with 100% efficiency. A range of 80% < E < 110% was considered acceptable.

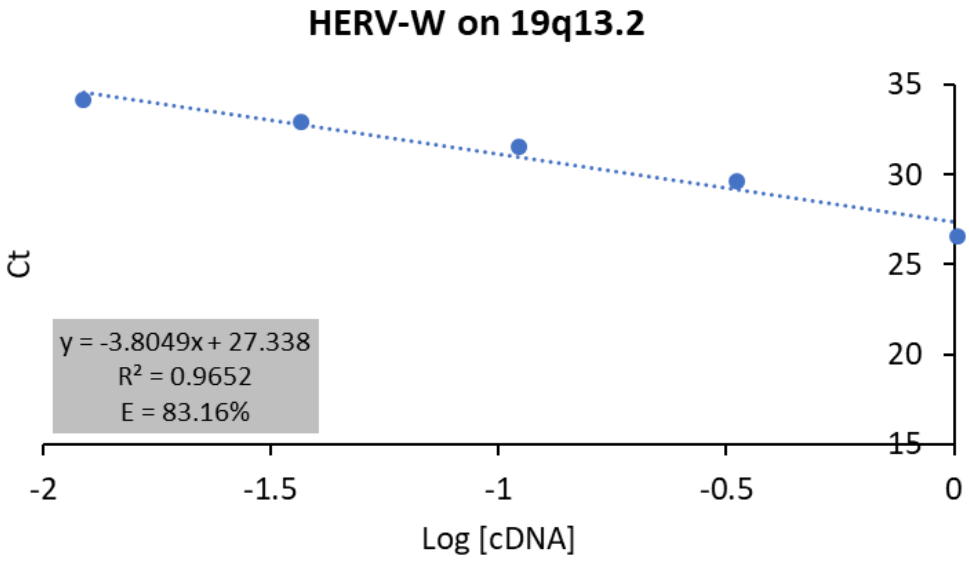


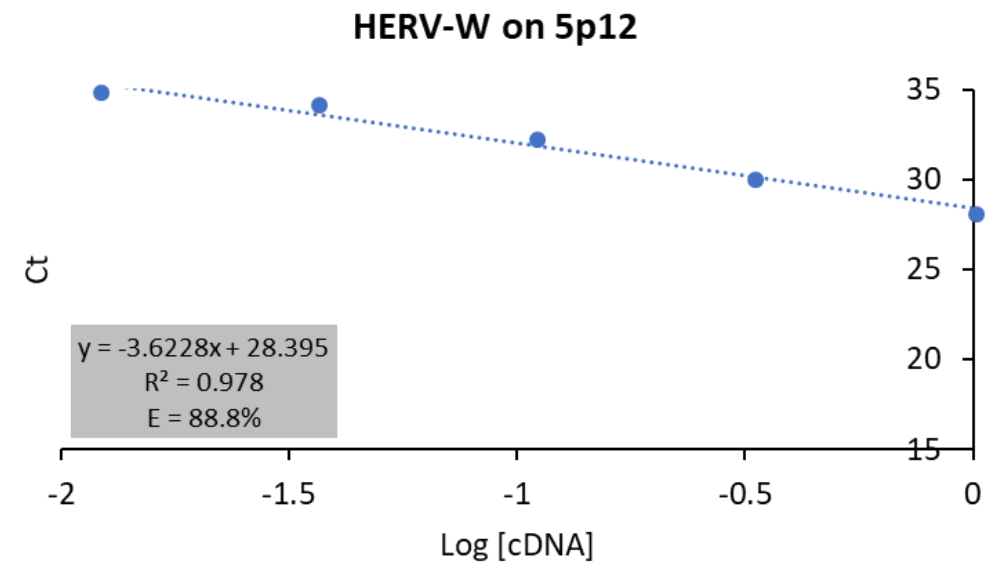


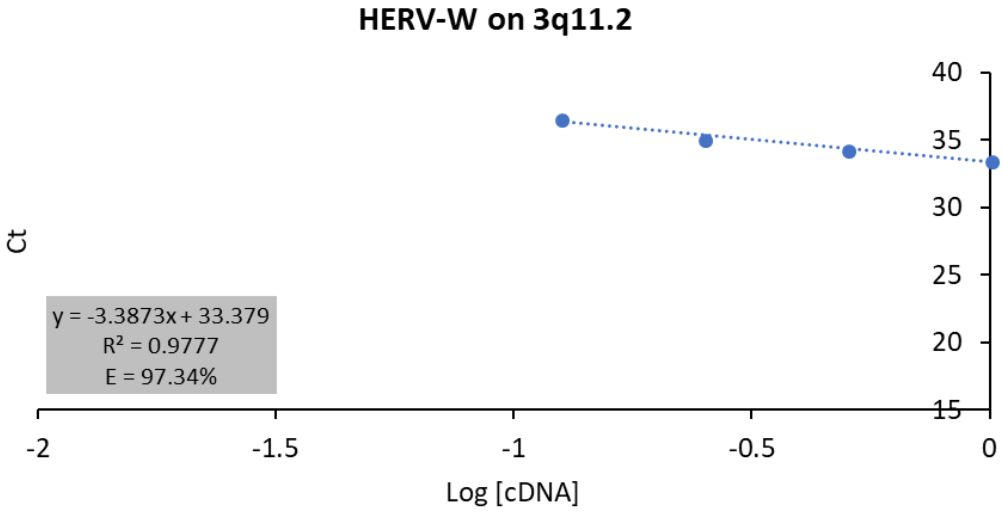


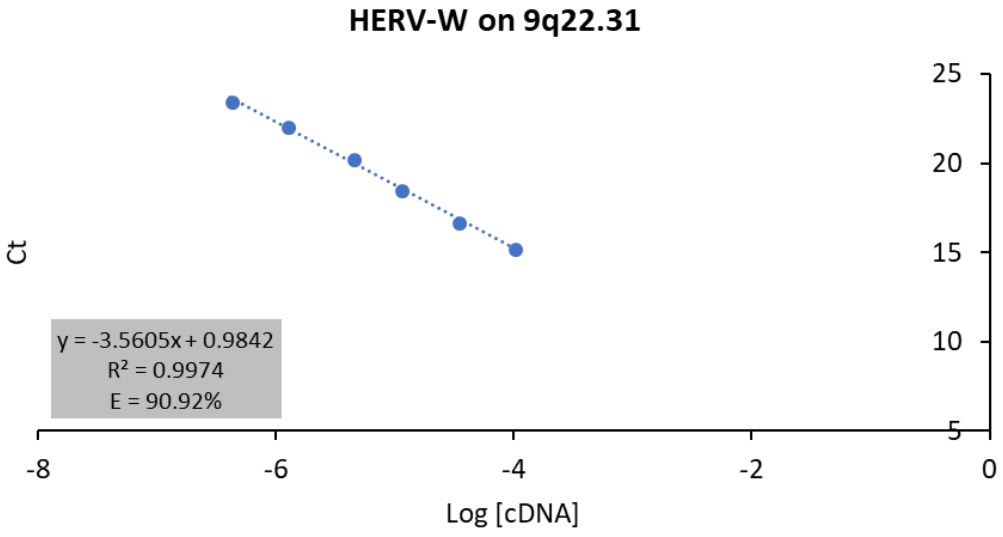

MSRVenv
(Mameli et al., 2009)
Ct
y = -3.4334x + 32.349
R² = 0.9902
E = 95.55%
40
35
30
25
20
15
-2
-1.5
-1
-0.5
0
Log [cDNA]

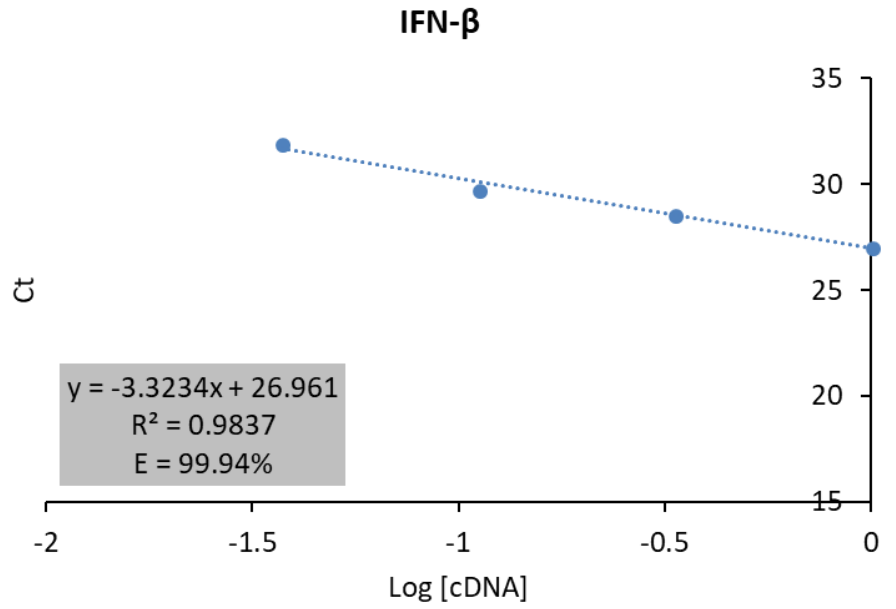

HERV-Wenv
(Nellåker et al., 2006)
Ct
y = -3.2915x + 25.304
R² = 1
E = 100%
35
30
25
20
15
-1.5
-1
-0.5
0
Log [cDNA]

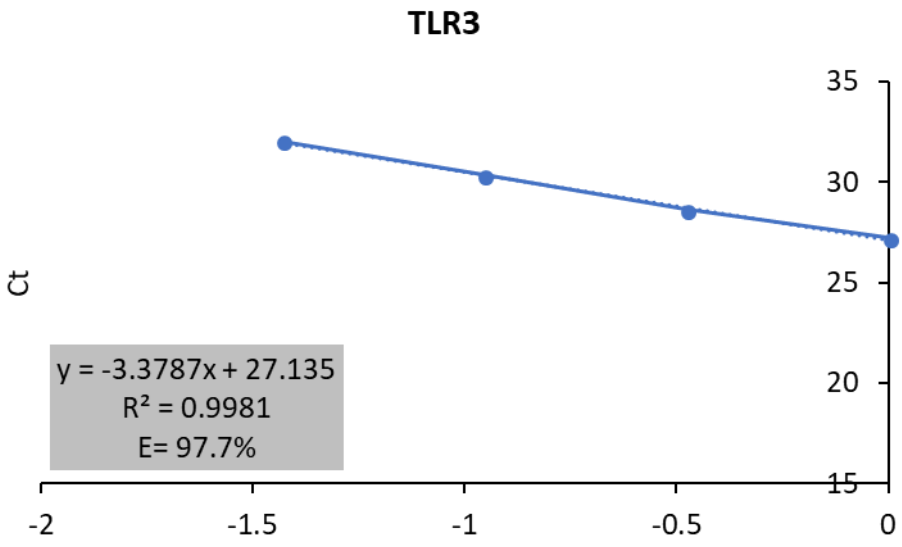

HERV-W on 19q13.2
Ct
y = -3.8049x + 27.338
R² = 0.9652
E = 83.16%
35
30
25
20
15
-2
-1.5
-1
-0.5
0
Log [cDNA]

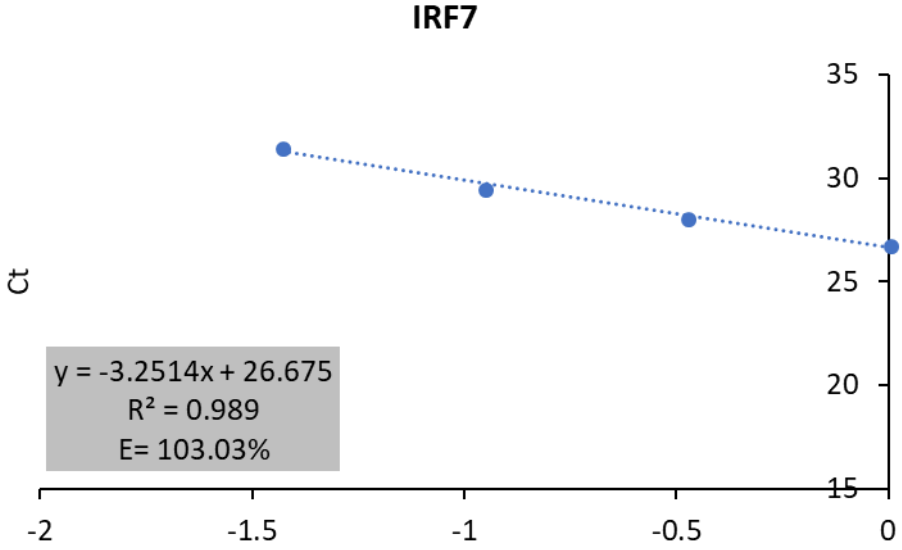

HERV-W on 5p12
Ct
y = -3.6228x + 28.395
R² = 0.978
E = 88.8%
35
30
25
20
15
-2
-1.5
-1
-0.5
0
Log [cDNA]


HERV-W on 3q11.2
Ct
y = -3.3873x + 33.379
R² = 0.9777
E = 97.34%
40
35
30
25
20
15
-2
-1.5
-1
-0.5
0
Log [cDNA]


HERV-W on 9q22.31
Ct
y = -3.5605x + 0.9842
R² = 0.9974
E = 90.92%
25
20
15
10
5
-8
-6
-4
-2
0
Log [cDNA]


IFN-β
Ct
y = -3.3234x + 26.961
R² = 0.9837
E = 99.94%
35
30
25
20
15
-2
-1.5
-1
-0.5
0
Log [cDNA]


TLR3
Ct
y = -3.3787x + 27.135
R² = 0.9981
E= 97.7%
35
30
25
20
15
-2
-1.5
-1
-0.5
0


IRF7
Ct
y = -3.2514x + 26.675
R² = 0.989
E= 103.03%
35
30
25
20
15
-2
-1.5
-1
-0.5
0

VIII. REFERENCES

A Jones, P. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond.

Absinta, M., Reich, D.S., and Filippi, M. (2016). Spring cleaning: time to rethink imaging research lines in MS? J. Neurol. *263*, 1893–1902.

Abud, E.M., Ramirez, R.N., Martinez, E.S., Healy, L.M., Nguyen, C.H.H., Newman, S.A., Yeromin, A.V., Scarfone, V.M., Marsh, S.E., Fimbres, C., et al. (2017). iPSC-Derived Human Microglia-like Cells to Study Neurological Diseases. Neuron *94*, 278-293.e9.

Ahlgren, C., Torén, K., Odén, A., and Andersen, O. (2009). A population-based case-control study on viral infections and vaccinations and subsequent multiple sclerosis risk. Eur. J. Epidemiol. *24*, 541–552.

Alison, M.R., Poulsom, R., Forbes, S., and Wright, N.A. (2002). An introduction to stem cells. J. Pathol. *197*, 419–423.

Al-Izki, S., Pryce, G., O'Neill, J.K., Butter, C., Giovannoni, G., Amor, S., and Baker, D. (2012). Practical guide to the induction of relapsing progressive experimental autoimmune encephalomyelitis in the Biozzi ABH mouse. Mult. Scler. Relat. Disord. *1*, 29–38.

All, A.H., Gharibani, P., Gupta, S., Bazley, F.A., Pashai, N., Chou, B.-K., Shah, S., Resar, L.M., Cheng, L., Gearhart, J.D., et al. (2015). Early Intervention for Spinal Cord Injury with Human Induced Pluripotent Stem Cells Oligodendrocyte Progenitors. PLOS ONE *10*, e0116933.

Allegretta, M., Nicklas, J.A., Sriram, S., and Albertini, R.J. (1990). T cells responsive to myelin basic protein in patients with multiple sclerosis. Science *247*, 718–721.

Alsanie, W.F., Niclis, J.C., and Petratos, S. (2013). Human Embryonic Stem Cell-Derived Oligodendrocytes: Protocols and Perspectives. Stem Cells Dev. *22*, 2459–2476.

Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev. Biol. *227*, 271–278.

Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. Proc. Natl. Acad. Sci. U. S. A. *90*, 11995–11999.

Antony, J.M., Marle, G. van, Opii, W., Butterfield, D.A., Mallet, F., Yong, V.W., Wallace, J.L., Deacon, R.M., Warren, K., and Power, C. (2004). Human endogenous retrovirus glycoprotein–mediated induction of redox reactants causes oligodendrocyte death and demyelination. Nat. Neurosci. *7*, 1088–1095.

Antony, J.M., Izad, M., Bar-Or, A., Warren, K.G., Vodjgani, M., Mallet, F., and Power, C. (2006). Quantitative analysis of human endogenous retrovirus-W env in neuroinflammatory diseases. AIDS Res. Hum. Retroviruses *22*, 1253–1259.

Ascherio, A., and Munger, K.L. (2010). Epstein-barr virus infection and multiple sclerosis: a review. J. Neuroimmune Pharmacol. Off. J. Soc. NeuroImmune Pharmacol. *5*, 271–277.

Assmann, T.S., Brondani, L. de A., Bouças, A.P., Canani, L.H., and Crispim, D. (2015). Toll-like receptor 3 (TLR3) and the development of type 1 diabetes mellitus. Arch. Endocrinol. Metab. *59*, 4–12.

Ayarpadikannan, S., and Kim, H.-S. (2014). The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. Genomics Inform. *12*, 98–104.

Babbe, H., Roers, A., Waisman, A., Lassmann, H., Goebels, N., Hohlfeld, R., Friese, M., Schröder, R., Deckert, M., Schmidt, S., et al. (2000). Clonal expansions of CD8(+) T cells dominate the T cell infiltrate in active multiple sclerosis lesions as shown by micromanipulation and single cell polymerase chain reaction. J. Exp. Med. *192*, 393–404.

Baker, D., O'Neill, J.K., Gschmeissner, S.E., Wilcox, C.E., Butter, C., and Turk, J.L. (1990). Induction of chronic relapsing experimental allergic encephalomyelitis in Biozzi mice. J. Neuroimmunol. *28*, 261–270.

Baker, D., Gerritsen, W., Rundle, J., and Amor, S. (2011). Critical appraisal of animal models of multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *17*, 647–657.

Bakshi, R., Hutton, G.J., Miller, J.R., and Radue, E.-W. (2004). The use of magnetic resonance imaging in the diagnosis and long-term management of multiple sclerosis. Neurology *63*, S3–S11.

Balada, E., Ordi-Ros, J., and Vilardell-Tarrés, M. (2009). Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity. Rev. Med. Virol. *19*, 273–286.

Baranzini, S.E., Mudge, J., van Velkinburgh, J.C., Khankhanian, P., Khrebtukova, I., Miller, N.A., Zhang, L., Farmer, A.D., Bell, C.J., Kim, R.W., et al. (2010). Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature *464*, 1351–1356.

Barbalat, R., Ewald, S.E., Mouchess, M.L., and Barton, G.M. (2011). Nucleic acid recognition by the innate immune system. Annu. Rev. Immunol. *29*, 185–214.

Barrat, F.J., Meeker, T., Gregorio, J., Chan, J.H., Uematsu, S., Akira, S., Chang, B., Duramad, O., and Coffman, R.L. (2005). Nucleic acids of mammalian origin can act as endogenous ligands for Toll-like receptors and may promote systemic lupus erythematosus. J. Exp. Med. *202*, 1131–1139.

Batoulis, H., Addicks, K., and Kuerten, S. (2010). Emerging concepts in autoimmune encephalomyelitis beyond the CD4/T(H)1 paradigm. Ann. Anat. Anat. Anz. Off. Organ Anat. Ges. *192*, 179–193.

Baylin, S., and A Jones, P. (2011). A decade of exploring the cancer epigenome—Biological and translational implications.

Becher, B., Durell, B.G., and Noelle, R.J. (2002). Experimental autoimmune encephalitis and inflammation in the absence of interleukin-12. J. Clin. Invest. *110*, 493–497.

Belshaw, R., Katzourakis, A., Paces, J., Burt, A., and Tristem, M. (2005). High copy number in human endogenous retrovirus families is associated with copying mechanisms in addition to reinfection. Mol. Biol. Evol. *22*, 814–817.

Bénit, L., Dessen, P., and Heidmann, T. (2001). Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. J. Virol. *75*, 11709–11719.

Bernard, C.C., Johns, T.G., Slavin, A., Ichikawa, M., Ewing, C., Liu, J., and Bettadapura, J. (1997). Myelin oligodendrocyte glycoprotein: a novel candidate autoantigen in multiple sclerosis. J. Mol. Med. Berl. Ger. *75*, 77–88.

Bhattacharya, B., Miura, T., Brandenberger, R., Mejido, J., Luo, Y., Yang, A.X., Joshi, B.H., Ginis, I., Thies, R.S., Amit, M., et al. (2004). Gene expression in human embryonic stem cell lines: unique molecular signature. Blood *103*, 2956–2964.

Bielekova, B., Goodwin, B., Richert, N., Cortese, I., Kondo, T., Afshar, G., Gran, B., Eaton, J., Antel, J., Frank, J.A., et al. (2000). Encephalitogenic potential of the myelin basic protein

peptide (amino acids 83-99) in multiple sclerosis: results of a phase II clinical trial with an altered peptide ligand. Nat. Med. *6*, 1167–1175.

Bird, A. (1997). Does DNA methylation control transposition of selfish elements in the germline? Trends Genet. TIG *13*, 469–472.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Bitsch, A., Schuchardt, J., Bunkowski, S., Kuhlmann, T., and Brück, W. (2000). Acute axonal injury in multiple sclerosis. Correlation with demyelination and inflammation. Brain J. Neurol. *123 (Pt 6)*, 1174–1183.

Blond, J.L., Besème, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet, F. (1999). Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. J. Virol. *73*, 1175–1185.

Blond, J.L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F.L. (2000). An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J. Virol. *74*, 3321–3329.

Bongso, A., and Richards, M. (2004). History and perspective of stem cell research. Best Pract. Res. Clin. Obstet. Gynaecol. *18*, 827–842.

Booss, J., Esiri, M.M., Tourtellotte, W.W., and Mason, D.Y. (1983). Immunohistological analysis of T lymphocyte subsets in the central nervous system in chronic progressive multiple sclerosis. J. Neurol. Sci. *62*, 219–232.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell *122*, 947–956.

Brück, W., Lucchinetti, C., and Lassmann, H. (2002). The pathology of primary progressive multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *8*, 93–97.

Brudek, T., Christensen, T., Aagaard, L., Petersen, T., Hansen, H.J., and Møller-Larsen, A. (2009). B cells and monocytes from patients with active multiple sclerosis exhibit increased surface expression of both HERV-H Env and HERV-W Env, accompanied by increased seroreactivity. Retrovirology *6*, 104.

Bsibsi, M., Ravid, R., Gveric, D., and van Noort, J.M. (2002). Broad expression of Toll-like receptors in the human central nervous system. J. Neuropathol. Exp. Neurol. *61*, 1013–1021.

Campbell, I.M., Gambin, T., Dittwald, P., Beck, C.R., Shuvarikov, A., Hixson, P., Patel, A., Gambin, A., Shaw, C.A., Rosenfeld, J.A., et al. (2014). Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination. BMC Biol. *12*, 74.

Carpentier, P.A., Duncan, D.S., and Miller, S.D. (2008). Glial Toll-like Receptor Signaling in Central Nervous System Infection and Autoimmunity. Brain. Behav. Immun. *22*, 140–147.

Celarain, N., Sánchez-Ruiz de Gordoa, J., Zelaya, M.V., Roldán, M., Larumbe, R., Pulido, L., Echavarri, C., and Mendioroz, M. (2016). TREM2 upregulation correlates with 5-hydroxymethycytosine enrichment in Alzheimer's disease hippocampus. Clin. Epigenetics *8*, 37.

Chaplin, D.D. (2010). Overview of the Immune Response. J. Allergy Clin. Immunol. *125*, S3-23.

Chen, Y., and Lai, D. (2015). Pluripotent States of Human Embryonic Stem Cells. Cell. Reprogramming *17*, 1–6.

Chen, Z., and Riggs, A.D. (2011). DNA Methylation and Demethylation in Mammals. J. Biol. Chem. *286*, 18347–18353.

Cherry, A.B.C., and Daley, G.Q. (2013). Reprogrammed cells for disease modeling and regenerative medicine. Annu. Rev. Med. *64*, 277–290.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., et al. (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974–986.

Cho, K., Lee, Y.-K., and Greenhalgh, D.G. (2008). Endogenous retroviruses in systemic response to stress signals. Shock Augusta Ga *30*, 105–116.

Christensen, S.R., Shupe, J., Nickerson, K., Kashgarian, M., Flavell, R.A., and Shlomchik, M.J. (2006). Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus. Immunity *25*, 417–428.

Chuong, E.B., Elde, N.C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. Genet. *18*, 71–86.

Colmegna, I., and Garry, R.F. (2006). Role of endogenous retroviruses in autoimmune diseases. Infect. Dis. Clin. North Am. *20*, 913–929.

Constantinescu, C.S., Farooqi, N., O'Brien, K., and Gran, B. (2011). Experimental autoimmune encephalomyelitis (EAE) as a model for multiple sclerosis (MS). Br. J. Pharmacol. *164*, 1079–1106.

Cua, D.J., Sherlock, J., Chen, Y., Murphy, C.A., Joyce, B., Seymour, B., Lucian, L., To, W., Kwan, S., Churakova, T., et al. (2003). Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature *421*, 744–748.

Curtin, F., Lang, A.B., Perron, H., Laumonier, M., Vidal, V., Porchet, H.C., and Hartung, H.-P. (2012). GNbAC1, a humanized monoclonal antibody against the envelope protein of multiple sclerosis-associated endogenous retrovirus: a first-in-humans randomized clinical study. Clin. Ther. *34*, 2268–2278.

Curtin, F., Perron, H., Kromminga, A., Porchet, H., and Lang, A.B. (2014). Preclinical and early clinical development of GNbAC1, a humanized IgG4 monoclonal antibody targeting endogenous retroviral MSRV-Env protein. MAbs *7*, 265–275.

Curtin, F., Vidal, V., Bernard, C., Kromminga, A., Lang, A.B., and Porchet, H. (2016). Serum pharmacokinetics and cerebrospinal fluid concentration analysis of the new IgG4 monoclonal antibody GNbAC1 to treat multiple sclerosis: A Phase 1 study. MAbs *8*, 854–860.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469–478.

Dean, G. (1967). Annual incidence, prevalence, and mortality of multiple sclerosis in white South-African-born and in white immigrants to South Africa. Br. Med. J. *2*, 724–730.

Derfuss, T., Curtin, F., Guebelin, C., Bridel, C., Rasenack, M., Matthey, A., Du Pasquier, R., Schluep, M., Desmeules, J., Lang, A.B., et al. (2015). A phase IIa randomised clinical study of GNbAC1, a humanised monoclonal antibody against the envelope protein of multiple sclerosis-associated endogenous retrovirus in multiple sclerosis patients. Mult. Scler. Houndmills Basingstoke Engl. *21*, 885–893.

Dewannieux, M., and Heidmann, T. (2013). Endogenous retroviruses: acquisition, amplification and taming of genome invaders. Curr. Opin. Virol. *3*, 646–656.

Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., et al. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. Science *321*, 1218–1221.

Dolei, A. (2006). Endogenous retroviruses and human disease. Expert Rev. Clin. Immunol. *2*, 149–167.

Dolei, A., Serra, C., Mameli, G., Pugliatti, M., Sechi, G., Cirotto, M.C., Rosati, G., and Sotgiu, S. (2002). Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. Neurology *58*, 471–473.

Douvaras, P., and Fossati, V. (2015). Generation and isolation of oligodendrocyte progenitor cells from human pluripotent stem cells. Nat. Protoc. *10*, 1143–1154.

Douvaras, P., Wang, J., Zimmer, M., Hanchuk, S., O'Bara, M.A., Sadiq, S., Sim, F.J., Goldman, J., and Fossati, V. (2014). Efficient Generation of Myelinating Oligodendrocytes from Primary Progressive Multiple Sclerosis Patients by Induced Pluripotent Stem Cells. Stem Cell Rep. *3*, 250–259.

Douvaras, P., Sun, B., Wang, M., Kruglikov, I., Lallos, G., Zimmer, M., Terrenoire, C., Zhang, B., Gandy, S., Schadt, E., et al. (2017). Directed Differentiation of Human Pluripotent Stem Cells to Microglia. Stem Cell Rep. *8*, 1516–1524.

Du, Y., and Dreyfus, C.F. (2002). Oligodendrocytes as providers of growth factors. J. Neurosci. Res. *68*, 647–654.

Duffy, S.S., Lees, J.G., and Moalem-Taylor, G. (2014). The contribution of immune and glial cell types in experimental autoimmune encephalomyelitis and multiple sclerosis. Mult. Scler. Int. *2014*, 285245.

Dupont, C., Armant, D.R., and Brenner, C.A. (2009). Epigenetics: Definition, Mechanisms and Clinical Perspective. Semin. Reprod. Med. *27*, 351–357.

Dyment, D.A., Sadovnick, A.D., Ebers, G.C., and Sadnovich, A.D. (1997). Genetics of multiple sclerosis. Hum. Mol. Genet. *6*, 1693–1698.

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. Epigenomics *1*, 239–259.

Ehrlich, M., Mozafari, S., Glatza, M., Starost, L., Velychko, S., Hallmann, A.-L., Cui, Q.-L., Schambach, A., Kim, K.-P., Bachelin, C., et al. (2017). Rapid and efficient generation of oligodendrocytes from human induced pluripotent stem cells using transcription factors. Proc. Natl. Acad. Sci. U. S. A. *114*, E2243–E2252.

Elong Ngono, A., Pettré, S., Salou, M., Bahbouhi, B., Soulillou, J.-P., Brouard, S., and Laplaud, D.-A. (2012). Frequency of circulating autoreactive T cells committed to myelin determinants in relapsing-remitting multiple sclerosis patients. Clin. Immunol. Orlando Fla *144*, 117–126.

Ewald, S.E., and Barton, G.M. (2011). Nucleic acid sensing Toll-like receptors in autoimmunity. Curr. Opin. Immunol. *23*, 3–9.

Fabriek, B.O., Zwemmer, J.N.P., Teunissen, C.E., Dijkstra, C.D., Polman, C.H., Laman, J.D., and Castelijns, J.A. (2005). In vivo detection of myelin proteins in cervical lymph nodes of MS patients using ultrasound-guided fine-needle aspiration cytology. J. Neuroimmunol. *161*, 190–194.

Fierabracci, A. (2009). Unravelling the role of infectious agents in the pathogenesis of human autoimmunity: the hypothesis of the retroviral involvement revisited. Curr. Mol. Med. *9*, 1024–1033.

Fischer, M., and Ehlers, M. (2008). Toll-like receptors in autoimmunity. Ann. N. Y. Acad. Sci. *1143*, 21–34.

van der Flier, L.G., and Clevers, H. (2009). Stem cells, self-renewal, and differentiation in the intestinal epithelium. Annu. Rev. Physiol. *71*, 241–260.

Frank, O., Jones-Brando, L., Leib-Mösch, C., Yolken, R., and Seifarth, W. (2006). Altered Transcriptional Activity of Human Endogenous Retroviruses in Neuroepithelial Cells after Infection with Toxoplasma gondii. J. Infect. Dis. *194*, 1447–1449.

Freund, J., and McDermott, K. (1942). Sensitization to Horse Serum by Means of Adjuvants. Proc. Soc. Exp. Biol. Med. *49*, 548–553.

Friese, M.A., and Fugger, L. (2009). Pathogenic CD8(+) T cells in multiple sclerosis. Ann. Neurol. *66*, 132–141.

Fuchs, E. (2008). Skin stem cells: rising to the surface. J. Cell Biol. *180*, 273–284.

Gall, A., Treuting, P., Elkon, K.B., Loo, Y.-M., Gale, M., Barber, G.N., and Stetson, D.B. (2012). Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. Immunity *36*, 120–131.

GAMES, and Transatlantic Multiple Sclerosis Genetics Cooperative (2003). A meta-analysis of whole genome linkage screens in multiple sclerosis. J. Neuroimmunol. *143*, 39–46.

García-León, J.A., Kumar, M., Boon, R., Chau, D., One, J., Wolfs, E., Eggermont, K., Berckmans, P., Gunhanlar, N., de Vrij, F., et al. (2018). SOX10 Single Transcription Factor-Based Fast and Efficient Generation of Oligodendrocytes from Human Pluripotent Stem Cells. Stem Cell Rep. *10*, 655–672.

Garcia-Montojo, M., Dominguez-Mozo, M., Arias-Leal, A., Garcia-Martinez, Á., De las Heras, V., Casanova, I., Faucard, R., Gehin, N., Madeira, A., Arroyo, R., et al. (2013). 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*, e53623.

García-Montojo, M., de la Hera, B., Varadé, J., de la Encarnación, A., Camacho, I., Domínguez-Mozo, M., Arias-Leal, A., García-Martínez, Á., Casanova, I., Izquierdo, G., et al. (2014). HERV-W polymorphism in chromosome X is associated with multiple sclerosis risk and with differential expression of MSRV. Retrovirology *11*, 2.

Gardner, R. (2002). Stem cells: potency, plasticity and public perception. J. Anat. *200*, 277–282.

Garg, N., and Smith, T.W. (2015). An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. Brain Behav. *5*, n/a-n/a.

Garson, J.A., Tuke, P.W., Giraud, P., Paranhos-Baccala, G., and Perron, H. (1998). Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. Lancet Lond. Engl. *351*, 33.

Genain, C.P., Lee-Parritz, D., Nguyen, M.H., Massacesi, L., Joshi, N., Ferrante, R., Hoffman, K., Moseley, M., Letvin, N.L., and Hauser, S.L. (1994). In healthy primates, circulating autoreactive T cells mediate autoimmune disease. J. Clin. Invest. *94*, 1339–1345.

Genain, C.P., Cannella, B., Hauser, S.L., and Raine, C.S. (1999). Identification of autoantibodies associated with myelin damage in multiple sclerosis. Nat. Med. *5*, 170–175.

Ghaffari, L.T., Starr, A., Nelson, A.T., and Sattler, R. (2018). Representing Diversity in the Dish: Using Patient-Derived in Vitro Models to Recreate the Heterogeneity of Neurological Disease. Front. Neurosci. *12*, 56.

Gibbons, H.M., and Dragunow, M. (2010). Adult human brain cell culture for neuroscience research. Int. J. Biochem. Cell Biol. *42*, 844–856.

Gimenez, J., Montgiraud, C., Pichon, J.-P., Bonnaud, B., Arsac, M., Ruel, K., Bouton, O., and Mallet, F. (2010). Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. Nucleic Acids Res. *38*, 2229–2246.

Girard, L., and Freeling, M. (1999). Regulatory changes as a consequence of transposon insertion. Dev. Genet. *25*, 291–296.

Göbel, K., Melzer, N., Herrmann, A.M., Schuhmann, M.K., Bittner, S., Ip, C.W., Hünig, T., Meuth, S.G., and Wiendl, H. (2009). Collateral neuronal apoptosis in CNS gray matter during an oligodendrocyte-directed CD8+ T cell attack. Glia *58*, 469–480.

Goff, S.P. (2004a). Genetic control of retrovirus susceptibility in mammalian cells. Annu. Rev. Genet. *38*, 61–85.

Goff, S.P. (2004b). Retrovirus restriction factors. Mol. Cell *16*, 849–859.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. Cell *128*, 635–638.

Goldenberg, M.M. (2012). Multiple Sclerosis Review. Pharm. Ther. *37*, 175–184.

Goodin, D.S. (2016). Chapter 11 - The epidemiology of multiple sclerosis: insights to a causal cascade††This chapter has been partially adapted from Goodin DS (2014). The epidemiology of multiple sclerosis: insights to disease pathogenesis. In Goodin DS (ed.) Multiple Sclerosis and Related Disorders. Handbook of Clinical Neurology, vol. 122. Amsterdam: Elsevier. In Handbook of Clinical Neurology, M.J. Aminoff, F. Boller, and D.F. Swaab, eds. (Elsevier), pp. 173–206.

Graves, M.C., Benton, M., Lea, R.A., Boyle, M., Tajouri, L., Macartney-Coxson, D., Scott, R.J., and Lechner-Scott, J. (2014). Methylation differences at the HLA-DRB1 locus in CD4+ T-Cells are associated with multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *20*, 1033–1041.

Gröger, V., and Cynis, H. (2018). Human Endogenous Retroviruses and Their Putative Role in the Development of Autoimmune Disorders Such as Multiple Sclerosis. Front. Microbiol. *9*.

Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221–225.

Gunhanlar, N., Shpak, G., van der Kroeg, M., Gouty-Colomer, L.A., Munshi, S.T., Lendemeijer, B., Ghazvini, M., Dupont, C., Hoogendijk, W.J.G., Gribnau, J., et al. (2018). A simplified protocol for differentiation of electrophysiologically mature neuronal networks from human induced pluripotent stem cells. Mol. Psychiatry *23*, 1336–1344.

Hackett, J.A., and Surani, M.A. (2012). DNA methylation dynamics during the mammalian life cycle. Philos. Trans. R. Soc. B Biol. Sci. *368*, 20110328–20110328.

Häggström, M., and Häggström, M (2014). Medical gallery of Mikael Häggström 2014. WikiJournal Med. *1*, 8.

Hammond, S.R., English, D.R., and McLeod, J.G. (2000). The age-range of risk of developing multiple sclerosis: evidence from a migrant population in Australia. Brain J. Neurol. *123 (Pt 5)*, 968–974.

Hampton, D.W., Anderson, J., Pryce, G., Irvine, K.-A., Giovannoni, G., Fawcett, J.W., Compston, A., Franklin, R.J.M., Baker, D., and Chandran, S. (2008). An experimental model of secondary progressive multiple sclerosis that shows regional variation in gliosis, remyelination, axonal and neuronal loss. J. Neuroimmunol. *201–202*, 200–211.

Harris, R.S., and Liddament, M.T. (2004). Retroviral restriction by APOBEC proteins. Nat. Rev. Immunol. *4*, 868–877.

Harrison, D.M., Roy, S., Oh, J., Izbudak, I., Pham, D., Courtney, S., Caffo, B., Jones, C.K., van Zijl, P., and Calabresi, P.A. (2015). Association of Cortical Lesion Burden on 7-T Magnetic Resonance Imaging With Cognition and Disability in Multiple Sclerosis. JAMA Neurol. *72*, 1004–1012.

Hayes, C.E. (2000). Vitamin D: a natural inhibitor of multiple sclerosis. Proc. Nutr. Soc. *59*, 531–535.

Hemmer, B., Archelos, J.J., and Hartung, H.-P. (2002). New concepts in the immunopathogenesis of multiple sclerosis. Nat. Rev. Neurosci. *3*, 291–301.

Hernán, M.A., Olek, M.J., and Ascherio, A. (2001). Cigarette smoking and incidence of multiple sclerosis. Am. J. Epidemiol. *154*, 69–74.

Hohenadl, C., Germaier, H., Walchner, M., Hagenhofer, M., Herrmann, M., Stürzl, M., Kind, P., Hehlmann, R., Erfle, V., and Leib-Mösch, C. (1999). Transcriptional activation of endogenous retroviral sequences in human epidermal keratinocytes by UVB irradiation. J. Invest. Dermatol. *113*, 587–594.

Hon, G.M., Erasmus, R.T., and Matsha, T. (2013). Multiple sclerosis-associated retrovirus and related human endogenous retrovirus-W in patients with multiple sclerosis: a literature review. J. Neuroimmunol. *263*, 8–12.

Hore, T.A., Rapkins, R.W., and Graves, J.A.M. (2007). Construction and evolution of imprinted loci in mammals. Trends Genet. TIG *23*, 440–448.

Hu, B.-Y., Du, Z.-W., and Zhang, S.-C. (2009). Differentiation of human oligodendrocytes from pluripotent stem cells. Nat. Protoc. *4*, 1614–1622.

Hurst, T.P., and Magiorkinis, G. (2015). Activation of the innate immune response by endogenous retroviruses. J. Gen. Virol. *96*, 1207–1218.

Hurst, T.P., and Magiorkinis, G. (2017). Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (LTRs). Viruses *9*.

Huynh, J.L., Garg, P., Thin, T.H., Yoo, S., Dutta, R., Trapp, B.D., Haroutunian, V., Zhu, J., Donovan, M.J., Sharp, A.J., et al. (2014). Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. Nat. Neurosci. *17*, 121–130.

International Multiple Sclerosis Genetics Consortium, Hafler, D.A., Compston, A., Sawcer, S., Lander, E.S., Daly, M.J., De Jager, P.L., de Bakker, P.I.W., Gabriel, S.B., Mirel, D.B., et al. (2007). Risk alleles for multiple sclerosis identified by a genomewide study. N. Engl. J. Med. *357*, 851–862.

International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C.A., Patsopoulos, N.A., Moutsianas, L.,

Dilthey, A., Su, Z., et al. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature *476*, 214–219.

International Multiple Sclerosis Genetics Consortium (IMSGC), Beecham, A.H., Patsopoulos, N.A., Xifara, D.K., Davis, M.F., Kemppinen, A., Cotsapas, C., Shah, T.S., Spencer, C., Booth, D., et al. (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat. Genet. *45*, 1353–1360.

Iramaneerat, K., Rattanatunyong, P., Khemapech, N., Triratanachat, S., and Mutirangura, A. (2011). HERV-K hypomethylation in ovarian clear cell carcinoma is associated with a poor prognosis and platinum resistance. Int. J. Gynecol. Cancer Off. J. Int. Gynecol. Cancer Soc. *21*, 51–57.

Jern, P., and Coffin, J.M. (2008). Effects of retroviruses on host genome function. Annu. Rev. Genet. *42*, 709–732.

Jilek, S., Schluep, M., Rossetti, A.O., Guignard, L., Le Goff, G., Pantaleo, G., and Du Pasquier, R.A. (2007). CSF enrichment of highly differentiated CD8+ T cells in early multiple sclerosis. Clin. Immunol. Orlando Fla *123*, 105–113.

Kabat, E.A., Wolf, A., and Bezer, A.E. (1947). THE RAPID PRODUCTION OF ACUTE DISSEMINATED ENCEPHALOMYELITIS IN RHESUS MONKEYS BY INJECTION OF HETEROLOGOUS AND HOMOLOGOUS BRAIN TISSUE WITH ADJUVANTS. J. Exp. Med. *85*, 117–130.

Kang, S.-M., Cho, M.S., Seo, H., Yoon, C.J., Oh, S.K., Choi, Y.M., and Kim, D.-W. (2007). Efficient induction of oligodendrocytes from human embryonic stem cells. Stem Cells Dayt. Ohio *25*, 419–424.

Kaushansky, N., Eisenstein, M., Zilkha-Falb, R., and Ben-Nun, A. (2010). The myelin-associated oligodendrocytic basic protein (MOBP) as a relevant primary target autoantigen in multiple sclerosis. Autoimmun. Rev. *9*, 233–236.

Kazazian, H.H. (2004). Mobile elements: drivers of genome evolution. Science *303*, 1626–1632.

Keller, G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev. *19*, 1129–1155.

Kent, L. (2009). Culture and Maintenance of Human Embryonic Stem Cells. J. Vis. Exp. JoVE.

Kerr, C.L., Letzen, B.S., Hill, C.M., Agrawal, G., Thakor, N.V., Sterneckert, J.L., Gearhart, J.D., and All, A.H. (2010). Efficient differentiation of human embryonic stem cells into oligodendrocyte progenitors for application in a rat contusion model of spinal cord injury. Int. J. Neurosci. *120*, 305–313.

Kielian, T. (2006). Toll-Like Receptors in Central Nervous System Glial Inflammation and Homeostasis. J. Neurosci. Res. *83*, 711–730.

Kipp, M., van der Star, B., Vogel, D.Y.S., Puentes, F., van der Valk, P., Baker, D., and Amor, S. (2012). Experimental in vivo and in vitro models of multiple sclerosis: EAE and beyond. Mult. Scler. Relat. Disord. *1*, 15–28.

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E., Mallet, F., Mandrand, B., et al. (1999). Molecular Cloning and Characterization of MSRV-Related Sequences Associated with Retrovirus-like Particles. Virology *260*, 1–9.

Kremer, D., Schichel, T., Förster, M., Tzekova, N., Bernard, C., van der Valk, P., van Horssen, J., Hartung, H.-P., Perron, H., and Küry, P. (2013). Human endogenous retrovirus type W envelope protein inhibits oligodendroglial precursor cell differentiation. Ann. Neurol. *74*, 721–732.

Kuhlmann, T., Lingfeld, G., Bitsch, A., Schuchardt, J., and Brück, W. (2002). Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. Brain J. Neurol. *125*, 2202–2212.

Kumagai, C., Kalman, B., Middleton, F.A., Vyshkina, T., and Massa, P.T. (2012). Increased promoter methylation of the immune regulatory gene SHP-1 in leukocytes of multiple sclerosis subjects. J. Neuroimmunol. *246*, 51–57.

Kumar, R., Sharma, A., Pattnaik, A.K., and Varadwaj, P.K. (2010). Stem cells: An overview with respect to cardiovascular and renal disease. J. Nat. Sci. Biol. Med. *1*, 43–52.

Kurth, R., and Bannert, N. (2010). Beneficial and detrimental effects of human endogenous retroviruses. Int. J. Cancer *126*, 306–314.

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

Kurtzke, J.F. (2005). Epidemiology and etiology of multiple sclerosis. Phys. Med. Rehabil. Clin. N. Am. *16*, 327–349.

Lande-Diner, L., and Cedar, H. (2005). Silence of the genes--mechanisms of long-term repression. Nat. Rev. Genet. *6*, 648–654.

Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A.G., Rechavi, G., and Cedar, H. (2007). Role of DNA methylation in stable gene repression. J. Biol. Chem. *282*, 12194–12200.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lättekivi, F., Kõks, S., Keermann, M., Reimann, E., Prans, E., Abram, K., Silm, H., Kõks, G., and Kingo, K. (2018). Transcriptional landscape of human endogenous retroviruses (HERVs) and other repetitive elements in psoriatic skin. Sci. Rep. *8*, 4358.

Laufer, G., Mayer, J., Mueller, B.F., Mueller-Lantzsch, N., and Ruprecht, K. (2009). Analysis of transcribed human endogenous retrovirus W env loci clarifies the origin of multiple sclerosis-associated retrovirus env sequences. Retrovirology *6*, 37.

Lavie, L., Kitova, M., Maldener, E., Meese, E., and Mayer, J. (2005). CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K(HML-2). J. Virol. *79*, 876–883.

Lee, J.E., and Lee, D.R. (2011). Human Embryonic Stem Cells: Derivation, Maintenance and Cryopreservation. Int. J. Stem Cells *4*, 9–17.

Lee, Y.N., and Bieniasz, P.D. (2007). Reconstitution of an Infectious Human Endogenous Retrovirus. PLOS Pathog. *3*, e10.

Li, F., Nellåker, C., Yolken, R.H., and Karlsson, H. (2011). A systematic evaluation of expression of HERV-W elements; influence of genomic context, viral structure and orientation. BMC Genomics *12*, 22.

Liggett, T., Melnikov, A., Tilwalli, S., Yi, Q., Chen, H., Replogle, C., Feng, X., Reder, A., Stefoski, D., Balabanov, R., et al. (2010). Methylation patterns of cell-free plasma DNA in relapsing-remitting multiple sclerosis. J. Neurol. Sci. *290*, 16–21.

Link, H., and Huang, Y.-M. (2006). Oligoclonal bands in multiple sclerosis cerebrospinal fluid: an update on methodology and clinical usefulness. J. Neuroimmunol. *180*, 17–28.

Lo, B., and Parham, L. (2009). Ethical Issues in Stem Cell Research. Endocr. Rev. *30*, 204–213.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000a). Organizing Cellular DNA into Chromosomes.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000b). Viruses: Structure, Function, and Uses. Mol. Cell Biol. 4th Ed.

Louapre, C., Govindarajan, S.T., Giannì, C., Langkammer, C., Sloane, J.A., Kinkel, R.P., and Mainero, C. (2015). Beyond focal cortical lesions in MS: An in vivo quantitative and spatial imaging study at 7T. Neurology *85*, 1702–1709.

Löwer, R., Löwer, J., and Kurth, R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. U. S. A. *93*, 5177–5184.

Lublin, F.D. (2005). Clinical features and diagnosis of multiple sclerosis. Neurol. Clin. *23*, 1–15, v.

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., and Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. Ann. Neurol. *47*, 707–717.

Maksakova, I.A., Mager, D.L., and Reiss, D. (2008). Endogenous retroviruses: Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. Cell. Mol. Life Sci. *65*, 3329–3347.

Mameli, G., Astone, V., Arru, G., Marconi, S., Lovato, L., Serra, C., Sotgiu, S., Bonetti, B., and Dolei, A. (2007). Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. J. Gen. Virol. *88*, 264–274.

Mameli, G., Poddighe, L., Astone, V., Delogu, G., Arru, G., Sotgiu, S., Serra, C., and Dolei, A. (2009). 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*, 98–106.

Marguerat, S., Wang, W.Y.S., Todd, J.A., and Conrad, B. (2004). Association of human endogenous retrovirus K-18 polymorphisms with type 1 diabetes. Diabetes *53*, 852–854.

McClintock, B. (1950). The Origin and Behavior of Mutable Loci in Maize. Proc. Natl. Acad. Sci. U. S. A. *36*, 344–355.

McDonald, W.I., Compston, A., Edan, G., Goodkin, D., Hartung, H.P., Lublin, F.D., McFarland, H.F., Paty, D.W., Polman, C.H., Reingold, S.C., et al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann. Neurol. *50*, 121–127.

Medina, K.L. (2016). Overview of the immune system. Handb. Clin. Neurol. *133*, 61–76.

MeinL, E., Hoch, R.M., Dornmair, K., de Waal Malefyt, R., Bontrop, R.E., Jonker, M., Lassmann, H., Hohlfeld, R., Wekerle, H., and ’t Hart, B.A. (1997). Encephalitogenic potential of myelin

basic protein-specific T cells isolated from normal rhesus macaques. Am. J. Pathol. *150*, 445–453.

Melnikov, A., Scholtens, D., Godwin, A., and Levenson, V. (2009). Differential Methylation Profile of Ovarian Cancer in Tissues and Plasma. J. Mol. Diagn. JMD *11*, 60–65.

Mendel, I., Kerlero de Rosbo, N., and Ben-Nun, A. (1995). A myelin oligodendrocyte glycoprotein peptide induces typical chronic experimental autoimmune encephalomyelitis in H-2b mice: fine specificity and T cell receptor V beta expression of encephalitogenic T cells. Eur. J. Immunol. *25*, 1951–1959.

Mendioroz, M., Celarain, N., Altuna, M., Sánchez-Ruiz de Gordoa, J., Zelaya, M.V., Roldán, M., Rubio, I., Larumbe, R., Erro, M.E., Méndez, I., et al. (2016). CRTC1 gene is differentially methylated in the human hippocampus in Alzheimer's disease. Alzheimers Res. Ther. *8*, 15.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. *28*, 812–828.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785–789.

Milo, R., and Kahana, E. (2010). Multiple sclerosis: geoepidemiology, genetics and the environment. Autoimmun. Rev. *9*, A387-394.

Minagar, A., and Alexander, J.S. (2003). Blood-brain barrier disruption in multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *9*, 540–549.

Miyazaki, Y., and Niino, M. (2015). Epigenetics in multiple sclerosis. Clin. Exp. Neuroimmunol. *6*, 49–58.

Montagnani, S., Rueger, M.A., Hosoda, T., and Nurzynska, D. (2016). Adult Stem Cells in Tissue Maintenance and Regeneration.

Morandi, E., Tarlinton, R.E., Tanasescu, R., and Gran, B. (2017a). Human endogenous retroviruses and multiple sclerosis: Causation, association, or after-effect? Mult. Scler. J. *23*, 1050–1055.

Morandi, E., Tanasescu, R., Tarlinton, R.E., Constantinescu, C.S., Zhang, W., Tench, C., and Gran, B. (2017b). The association between human endogenous retroviruses and multiple sclerosis: A systematic review and meta-analysis. PLoS ONE *12*.

Moretti, A., Laugwitz, K.-L., Dorn, T., Sinnecker, D., and Mummery, C. (2013). Pluripotent Stem Cell Models of Human Heart Disease. Cold Spring Harb. Perspect. Med. *3*.

Moutsianas, L., Jostins, L., Beecham, A.H., Dilthey, A.T., Xifara, D.K., Ban, M., Shah, T.S., Patsopoulos, N.A., Alfredsson, L., Anderson, C.A., et al. (2015). Class II HLA interactions modulate genetic risk for multiple sclerosis. Nat. Genet. *47*, 1107–1113.

Moyes, D., Griffiths, D.J., and Venables, P.J. (2007). Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. Trends Genet. TIG *23*, 326–333.

Munger, K.L., Zhang, S.M., O'Reilly, E., Hernán, M.A., Olek, M.J., Willett, W.C., and Ascherio, A. (2004). Vitamin D intake and incidence of multiple sclerosis. Neurology *62*, 60–65.

Munger, K.L., Levin, L.I., Hollis, B.W., Howard, N.S., and Ascherio, A. (2006). Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA *296*, 2832–2838.

Munoz, J.J., Bernard, C.C., and Mackay, I.R. (1984). Elicitation of experimental allergic encephalomyelitis (EAE) in mice with the aid of pertussigen. Cell. Immunol. *83*, 92–100.

Nadig, R.R. (2009). Stem cell therapy – Hype or hope? A review. J. Conserv. Dent. JCD *12*, 131–138.

Nakkuntod, J., Sukkapan, P., Avihingsanon, Y., Mutirangura, A., and Hirankarn, N. (2013). DNA methylation of human endogenous retrovirus in systemic lupus erythematosus. J. Hum. Genet. *58*, 241–249.

Nellåker, C., Yao, Y., Jones-Brando, L., Mallet, F., Yolken, R.H., and Karlsson, H. (2006). Transactivation of elements in the human endogenous retrovirus W family by viral infection. Retrovirology *3*, 44.

Nicaise, A.M., Banda, E., Guzzo, R.M., Russomanno, K., Castro-Borrero, W., Willis, C.M., Johnson, K.M., Lo, A.C., and Crocker, S.J. (2017). iPS-derived neural progenitor cells from PPMS patients reveal defect in myelin injury response. Exp. Neurol. *288*, 114–121.

Nielsen, T.R., Rostgaard, K., Nielsen, N.M., Koch-Henriksen, N., Haahr, S., Sørensen, P.S., and Hjalgrim, H. (2007). Multiple sclerosis after infectious mononucleosis. Arch. Neurol. *64*, 72–75.

Nistor, G.I., Totoiu, M.O., Haque, N., Carpenter, M.K., and Keirstead, H.S. (2005). Human embryonic stem cells differentiate into oligodendrocytes in high purity and myelinate after spinal cord transplantation. Glia *49*, 385–396.

Noh, J.S., Sharma, R.P., Veldic, M., Salvacion, A.A., Jia, X., Chen, Y., Costa, E., Guidotti, A., and Grayson, D.R. (2005). DNA methyltransferase 1 regulates reelin mRNA expression in mouse primary cortical cultures. Proc. Natl. Acad. Sci. U. S. A. *102*, 1749–1754.

Nylander, A., and Hafler, D.A. (2012). Multiple sclerosis. J. Clin. Invest. *122*, 1180–1188.

O'Connor, K.C., Appel, H., Bregoli, L., Call, M.E., Catz, I., Chan, J.A., Moore, N.H., Warren, K.G., Wong, S.J., Hafler, D.A., et al. (2005). Antibodies from Inflamed Central Nervous System Tissue Recognize Myelin Oligodendrocyte Glycoprotein. J. Immunol. Baltim. Md 1950 *175*, 1974–1982.

Olsson, T., Sun, J., Hillert, J., Höjeberg, B., Ekre, H.P., Andersson, G., Olerup, O., and Link, H. (1992). Increased numbers of T cells recognizing multiple myelin basic protein epitopes in multiple sclerosis. Eur. J. Immunol. *22*, 1083–1087.

Park, I.-H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008). Disease-Specific Induced Pluripotent Stem Cells. Cell *134*, 877–886.

Pavlicek, A., Paces, J., Elleder, D., and Hejnar, J. (2002). Processed Pseudogenes of Human Endogenous Retroviruses Generated by LINEs: Their Integration, Stability, and Distribution. Genome Res. *12*, 391–399.

Pelka, K., Shibata, T., Miyake, K., and Latz, E. (2016). Nucleic acid-sensing TLRs and autoimmunity: novel insights from structural and cell biology. Immunol. Rev. *269*, 60–75.

Perron, H., Geny, C., Laurent, A., Mouriquand, C., Pellat, J., Perret, J., and Seigneurin, J.M. (1989). Leptomeningeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. Res. Virol. *140*, 551–561.

Perron, H., Lalande, B., Gratacap, B., Laurent, A., Genoulaz, O., Geny, C., Mallaret, M., Schuller, E., Stoebner, P., and Seigneurin, J.M. (1991). Isolation of retrovirus from patients with multiple sclerosis. Lancet Lond. Engl. *337*, 862–863.

Perron, H., Garson, J.A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P.W., Voisset, C., Blond, J.L., et al. (1997). Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc. Natl. Acad. Sci. U. S. A. *94*, 7583–7588.

Perron, H., Perin, J.P., Rieger, F., and Alliel, P.M. (2000). Particle-associated retroviral RNA and tandem RGH/HERV-W copies on human chromosome 7q: possible components of a "chain-reaction" triggered by infectious agents in multiple sclerosis? J. Neurovirol. *6 Suppl 2*, S67-75.

Perron, H., Jouvin-Marche, E., Michel, M., Ounanian-Paraz, A., Camelo, S., Dumon, A., Jolivet-Reynaud, C., Marcel, F., Souillet, Y., Borel, E., et al. (2001). Multiple sclerosis retrovirus particles and recombinant envelope trigger an abnormal immune response in vitro, by inducing polyclonal Vbeta16 T-lymphocyte activation. Virology *287*, 321–332.

Perron, H., Lazarini, F., Ruprecht, K., Péchoux-Longin, C., Seilhean, D., Sazdovitch, V., Créange, A., Battail-Poirot, N., Sibaï, G., Santoro, L., et al. (2005). Human endogenous retrovirus (HERV)-W ENV and GAG proteins: physiological expression in human brain and pathophysiological modulation in multiple sclerosis lesions. J. Neurovirol. *11*, 23–33.

Perron, H., Germi, R., Bernard, C., Garcia-Montojo, M., Deluen, C., Farinelli, L., Faucard, R., Veas, F., Stefas, I., Fabriek, B.O., et al. (2012). Human endogenous retrovirus type W envelope expression in blood and brain cells provides new insights into multiple sclerosis disease. Mult. Scler. J. *18*, 1721–1736.

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F.A., Uchil, P.D., Chatel, L., et al. (2011). TRIM5 is an innate immune sensor for the retrovirus capsid lattice. Nature *472*, 361–365.

Pettinelli, C.B., and McFarlin, D.E. (1981). Adoptive transfer of experimental allergic encephalomyelitis in SJL/J mice after in vitro activation of lymph node cells by myelin basic protein: requirement for Lyt 1+ 2- T lymphocytes. J. Immunol. Baltim. Md 1950 *127*, 1420–1423.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat. Biotechnol. *28*, 1057–1068.

Prak, E.T., and Kazazian, H.H. (2000). Mobile elements and the human genome. Nat. Rev. Genet. *1*, 134–144.

Qiu, J. (2006). Epigenetics: unfinished symphony. Nature *441*, 143–145.

Ramagopalan, S.V., Valdar, W., Dyment, D.A., DeLuca, G.C., Yee, I.M., Giovannoni, G., Ebers, G.C., Sadovnick, A.D., and Canadian Collaborative Study Group (2009). Association of infectious mononucleosis with multiple sclerosis. A population-based study. Neuroepidemiology *32*, 257–262.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. *18*, 399–404.

Rolland, A., Jouvin-Marche, E., Viret, C., Faure, M., Perron, H., and Marche, P.N. (2006). The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. J. Immunol. Baltim. Md 1950 *176*, 7636–7644.

Romanelli, E., Merkler, D., Mezydlo, A., Weil, M.-T., Weber, M.S., Nikić, I., Potz, S., Meinl, E., Matznick, F.E.H., Kreutzfeldt, M., et al. (2016). Myelinosome formation represents an early

stage of oligodendrocyte damage in multiple sclerosis and its animal model. Nat. Commun. *7*, 13275.

Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. Virology *411*, 273–287.

Sakata, K., Nakayamada, S., Miyazaki, Y., Kubo, S., Ishii, A., Nakano, K., and Tanaka, Y. (2018). Up-Regulation of TLR7-Mediated IFN-α Production by Plasmacytoid Dendritic Cells in Patients With Systemic Lupus Erythematosus. Front. Immunol. *9*, 1957.

Sawcer, S., Goodfellow, P.N., and Compston, A. (1997). The genetic analysis of multiple sclerosis. Trends Genet. *13*, 234–239.

Sawcer, S., Franklin, R.J.M., and Ban, M. (2014). Multiple sclerosis genetics. Lancet Neurol. *13*, 700–709.

Schmitt, K., Richter, C., Backes, C., Meese, E., Ruprecht, K., and Mayer, J. (2013). Comprehensive Analysis of Human Endogenous Retrovirus Group HERV-W Locus Transcription in Multiple Sclerosis Brain Lesions by High-Throughput Amplicon Sequencing. J. Virol. *87*, 13837–13852.

Schoorlemmer, J., Pérez-Palacios, R., Climent, M., Guallar, D., and Muniesa, P. (2014). Regulation of Mouse Retroelement MuERV-L/MERVL Expression by REX1 and Epigenetic Control of Stem Cell Potency. Front. Oncol. *4*.

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*, 2899–2912.

Sharma, S., Kelly, T.K., and Jones, P.A. (2010). Epigenetics in cancer. Carcinogenesis *31*, 27–36.

Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. Nat. Rev. Drug Discov. *16*, 115–130.

Simons, M., and Nave, K.-A. (2016). Oligodendrocytes: Myelination and Axonal Support. Cold Spring Harb. Perspect. Biol. *8*, a020479.

Singh, V.K., Saini, A., Kalsan, M., Kumar, N., and Chandra, R. (2016). Describing the Stem Cell Potency: The Various Methods of Functional Assessment and In silico Diagnostics. Front. Cell Dev. Biol. *4*.

Smit, A.F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet. Dev. *9*, 657–663.

Smith, P.A., Heijmans, N., Ouwerling, B., Breij, E.C., Evans, N., van Noort, J.M., Plomp, A.C., Delarasse, C., 't Hart, B., Pham-Dinh, D., et al. (2005). Native myelin oligodendrocyte glycoprotein promotes severe chronic neurological disease and demyelination in Biozzi ABH mice. Eur. J. Immunol. *35*, 1311–1319.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., et al. (2009). Parkinson's Disease Patient-Derived Induced Pluripotent Stem Cells Free of Viral Reprogramming Factors. Cell *136*, 964–977.

Song, B., Sun, G., Herszfeld, D., Sylvain, A., Campanale, N.V., Hirst, C.E., Caine, S., Parkington, H.C., Tonta, M.A., Coleman, H.A., et al. (2012). Neural differentiation of patient specific iPS cells as a novel approach to study the pathophysiology of multiple sclerosis. Stem Cell Res. *8*, 259–273.

Sospedra, M., and Martin, R. (2005). Immunology of multiple sclerosis. Annu. Rev. Immunol. *23*, 683–747.

Stacpoole, S.R.L., Spitzer, S., Bilican, B., Compston, A., Karadottir, R., Chandran, S., and Franklin, R.J.M. (2013). High Yields of Oligodendrocyte Lineage Cells from Human Embryonic Stem Cells at Physiological Oxygen Tensions for Evaluation of Translational Biology. Stem Cell Rep. *1*, 437–450.

Stadelmann, C., Wegner, C., and Brück, W. (2011). Inflammation, demyelination, and degeneration — Recent insights from MS pathology. Biochim. Biophys. Acta BBA - Mol. Basis Dis. *1812*, 275–282.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev. *24*, 2239–2263.

van der Star, B.J., Vogel, D.Y.S., Kipp, M., Puentes, F., Baker, D., and Amor, S. (2012). In vitro and in vivo models of multiple sclerosis. CNS Neurol. Disord. Drug Targets *11*, 570–588.

Stetson, D.B. (2012). Endogenous retroelements and autoimmune disease. Curr. Opin. Immunol. *24*, 692–697.

Stetson, D.B., Ko, J.S., Heidmann, T., and Medzhitov, R. (2008). Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell *134*, 587–598.

Stinissen, P., and Hellings, N. (2008). Activation of myelin reactive T cells in multiple sclerosis: a possible role for T cell degeneracy? Eur. J. Immunol. *38*, 1190–1193.

Stoye, J.P. (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nat. Rev. Microbiol. *10*, 395–406.

Stromnes, I.M., and Goverman, J.M. (2006). Passive induction of experimental allergic encephalomyelitis. Nat. Protoc. *1*, 1952–1960.

Sun, C., Skaletsky, H., Rozen, S., Gromoll, J., Nieschlag, E., Oates, R., and Page, D.C. (2000). Deletion of azoospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses. Hum. Mol. Genet. *9*, 2291–2296.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872.

Theofilopoulos, A.N., Kono, D.H., Beutler, B., and Baccala, R. (2011). Intracellular Nucleic Acid Sensors and Autoimmunity. J. Interferon Cytokine Res. *31*, 867–886.

Thompson, A.J., Banwell, B.L., Barkhof, F., Carroll, W.M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M.S., et al. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. *17*, 162–173.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Traka, M., Podojil, J.R., McCarthy, D.P., Miller, S.D., and Popko, B. (2016). Oligodendrocyte death results in immune-mediated CNS demyelination. Nat. Neurosci. *19*, 65–74.

Trapp, B.D., Peterson, J., Ransohoff, R.M., Rudick, R., Mörk, S., and Bö, L. (1998). Axonal transection in the lesions of multiple sclerosis. N. Engl. J. Med. *338*, 278–285.

Traugott, U., Reinherz, E.L., and Raine, C.S. (1983). Multiple sclerosis. Distribution of T cells, T cell subsets and Ia-positive macrophages in lesions of different ages. J. Neuroimmunol. *4*, 201–221.

Trudler, D., Farfara, D., and Frenkel, D. (2010). Toll-Like Receptors Expression and Signaling in Glia Cells in Neuro-Amyloidogenic Diseases: Towards Future Therapeutic Application.

Tsujita, T., Niikawa, N., Yamashita, H., Imamura, A., Hamada, A., Nakane, Y., and Okazaki, Y. (1998). Genomic discordance between monozygotic twins discordant for schizophrenia. Am. J. Psychiatry *155*, 422–424.

Urbach, A., Bar-Nur, O., Daley, G.Q., and Benvenisty, N. (2010). Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. Cell Stem Cell *6*, 407–411.

Vesterinen, H.M., Sena, E.S., ffrench-Constant, C., Williams, A., Chandran, S., and Macleod, M.R. (2010). Improving the translational hit of experimental treatments in multiple sclerosis. Mult. Scler. Houndmills Basingstoke Engl. *16*, 1044–1055.

Villesen, P., Aagaard, L., Wiuf, C., and Pedersen, F.S. (2004). Identification of endogenous retroviral reading frames in the human genome. Retrovirology *1*, 32.

Vincendeau, M., Göttesdorfer, I., Schreml, J.M.H., Wetie, A.G.N., Mayer, J., Greenwood, A.D., Helfer, M., Kramer, S., Seifarth, W., Hadian, K., et al. (2015). Modulation of human endogenous retrovirus (HERV) transcription during persistent and de novo HIV-1 infection. Retrovirology *12*.

Wang, G., Zhang, H., Zhao, Y., Li, J., Cai, J., Wang, P., Meng, S., Feng, J., Miao, C., Ding, M., et al. (2005). Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. Biochem. Biophys. Res. Commun. *330*, 934–942.

Wang, S.-C., Oelze, B., and Schumacher, A. (2008). Age-Specific Epigenetic Drift in Late-Onset Alzheimer's Disease. PLOS ONE *3*, e2698.

Wolf, D., and Goff, S.P. (2008). Host Restriction Factors Blocking Retroviral Replication. Annu. Rev. Genet. *42*, 143–163.

Wu, G.F., and Alvarez, E. (2011). The immuno-pathophysiology of multiple sclerosis. Neurol. Clin. *29*, 257–278.

Wucherpfennig, K.W., and Strominger, J.L. (1995). Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell *80*, 695–705.

Wucherpfennig, K.W., Catz, I., Hausmann, S., Strominger, J.L., Steinman, L., and Warren, K.G. (1997). Recognition of the immunodominant myelin basic protein peptide by autoantibodies and HLA-DR2-restricted T cell clones from multiple sclerosis patients. Identity of key contact residues in the B-cell and T-cell epitopes. J. Clin. Invest. *100*, 1114–1122.

Xagorari, A., and Chlichlia, K. (2008). Toll-Like Receptors and Viruses: Induction of Innate Antiviral Immune Responses. Open Microbiol. J. *2*, 49–59.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. Nat. Biotechnol. *19*, 971–974.

Xu, R.-H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nat. Methods *2*, 185–190.

Yabut, O., and Bernstein, H.S. (2011). The promise of human embryonic stem cells in aging-associated diseases. Aging *3*, 494–508.

Young, G.R., Stoye, J.P., and Kassiotis, G. (2013). Are human endogenous retroviruses pathogenic? An approach to testing the hypothesis. Bioessays *35*, 794–803.

Zamvil, S.S., Nelson, P.A., Mitchell, D.J., Knobler, R.L., Fritz, R.B., and Steinman, L. (1985). Encephalitogenic T cell clones specific for myelin basic protein. An unusual bias in antigen recognition. J. Exp. Med. *162*, 2107–2124.

Zhang, J., Markovic-Plese, S., Lacet, B., Raus, J., Weiner, H.L., and Hafler, D.A. (1994). Increased frequency of interleukin 2-responsive T cells specific for myelin basic protein and proteolipid protein in peripheral blood and cerebrospinal fluid of patients with multiple sclerosis. J. Exp. Med. *179*, 973–984.

Zhang, S.C., Wernig, M., Duncan, I.D., Brüstle, O., and Thomson, J.A. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat. Biotechnol. *19*, 1129–1133.

Zimmermann, M., Sanderson, N.S.R., Rasenack, M., Lalive, P.H., Lang, A.B., Curtin, F., Lindberg, R.L.P., Kappos, L., and Derfuss, T. (2015). Immunologic monitoring during a phase 2a trial of the GNbAC1 antibody in patients with MS. Neurol. Neuroimmunol. Neuroinflammation *2*, e144.