LUIZ HENRIQUE DA SILVA NALI

**Perfil de expressão dos Retrovírus Endógenos Humanos da família W
em pacientes com Esclerose Múltipla**

Tese apresentada ao Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo para obtenção do título de Doutor em Ciências.

Área de concentração: Doenças Tropicais e Saúde Internacional

Orientadora: Dra. Camila Malta Romano

São Paulo

2018

Ficha catalográfica elaborada pela Biblioteca do Instituto de Medicina Tropical de São Paulo da
Universidade de São Paulo – Bibliotecário Carlos José Quinteiro, CRB-8 5538

© Reprodução autorizada pelo autor

Nali, Luiz Henrique da Silva

Perfil de expressão dos retrovírus endógenos humanos da família W em pacientes com esclerose múltipla / Luiz Henrique da Silva Nali. – São Paulo, 2018.

Tese (Doutorado) – Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Doenças Tropicais e Saúde Internacional

Orientadora: Camila Malta Romano

Descritores: 1. ESCLEROSE MÚLTIPLA. 2. RETROVIRIDAE. 3. TRANSCRIÇÃO GÊNICA. 4. SEQUENCIAMENTO GENÉTICO.

USP/IMTSP/BIB-02/2018.

*Dedico essa tese aos meus pais Magali Moreira e Luiz Nali,
que muito me apoiaram no decorrer dessa caminhada.
E aos pacientes com Esclerose Múltipla,
na incansável luta por melhor qualidade de vida.*

Agradecimento especial à minha orientadora, Dra. Camila Malta Romano, pela confiança depositada em mim para realização desse e de todos os trabalhos paralelos realizados. Pela duradoura amizade, excelência na minha orientação e, sem dúvidas, pelo auxílio no meu crescimento pessoal.

Ao Prof. Dr. Claudio Sérgio Pannuti, por abrir as portas do Laboratório de Virologia do IMT-USP, dando-me a oportunidade e depositando as suas fichas de confiança para que fosse possível realizar esse sonho.

À toda a equipe do CATEM, em especial ao Guilherme do Olival no suporte em toda parte clínica do estudo, as enfermeiras Ana Carolina Sena e Mayra no suporte nas coletas das amostras, e também ao Prof. Dr. Charles Peter Tilbery, por abrir as portas do ambulatório do CATEM para mim.

Aos professores, José Eduardo Levi, Enedina Oliveira, e Thelma Okay pelas valiosas sugestões no exame de qualificação.

Agradecimento especial ao Dr Hugo Naya e Lucía Spangenberg, pela acolhida na *Unidad de bioinformática* do Instituto Pasteur de Montevideu, Uruguai. Os ensinamentos que vocês me proporcionaram foram fundamentais no desenvolvimento desse estudo.

Ao Dr Israel Tojal e Dr Horácio Montenegro, pelo suporte e sugestões nas análises de bioinformática.

À Prof. Dra. Ester Sabino, pela oportunidade de lecionar no curso de aprimoramento em técnicas laboratoriais do Instituto de Medicina Tropical.

Ao Prof. Dr. Heitor Andrade e Dr. Andrés Galisteo, por abrir as portas da Protozoologia para que pudesse dar continuidade nos ensaios de PCR em Tempo Real.

Ao grupo de Neurociências do Instituto de Infectologia Emilio Ribas, em especial ao Dr. Augusto Penalva, que muito colaborou para discussões científicas de muitos estudos.

À Dra. Francielle Tramontini Gomes de Souza Cardoso pelo valioso auxílio e sugestões nas análises do IPA.

Ao Prof. Alúcio Segurado pela oportunidade de trabalhar no Programa de Aperfeiçoamento do Ensino (PAE) da FMUSP.

Aos meus pais Luiz Nali e Magali Moreira, por todo suporte, educação e valores morais que tem me passado até hoje. Serei eternamente grato.

À irmã Bianca Nali pelo companheirismo e por torcer por mim, obrigado por tudo Bi.

À minha família: tia e madrinha Monica, tia Valquíria, tia Rita, às minhas primas Thalita, Karina, Aline e Fernanda, e a minha avó Diva.

Agradecimento especial à Giselle Pacífico Sartori, pelo amor, confiança, apoio e cumplicidade. Sua presença ao meu lado foi fundamental nessa reta final do estudo, amo você.

Aos amigos que o laboratório e a pós-graduação me trouxe: Carol Soares, Giovana Caleiro, Bárbara Brito, Felipe Scassi, Toni Martins, Denis Fujita, Liã Arruda, Renato Oliveira, Paulo Urbano e Cristina Fink.

Aos colegas da Virologia: Tina, Lucy, Luciano, Nathalia, Anderson, Silvia, Dra Cássia, Dra Clarrise, Dra Daisy, Clara, Laurinha, Cyri, Carol Mamana, Débora, Georgina, Alexandre, Sonia, Paulo Braz, Marli, Maria, Luciana, Geovana, Midiã, Karina, Tânia, Maria, Vera, Wilton, Noeli e Jéssica.

Aos novos membros do time da Camila: Cibele Leal e Lucas Soares, obrigado pelo suporte em projetos paralelos que tocamos.

À Secretária da pós-graduação do Instituto de Medicina Tropical, Eliane Araújo, por todo suporte prestado.

À CAPES e a FAPESP (projeto 2013/24223-9) pelo auxílio nas bolsas concedidas.

À pós-graduação do Instituto de Medicina Tropical da Universidade de São Paulo

A todos os voluntários que concordaram em participar desse estudo.

Por fim, agradeço a todos que diretamente ou indiretamente auxiliaram no desenvolvimento desse estudo.

*“Alguns homens veem as coisas como são, e dizem 'Por quê?' Eu sonho com as coisas
que nunca foram e digo 'Por que não?' “*

George Bernard Shaw

Sumário

Resumo.....	19
Abstract.....	20
1. Introdução	21
1.1 A Esclerose Múltipla.....	21
1.1.2 Classificações da Esclerose Múltipla.....	23
1.1.3 Epidemiologia.....	24
1.1.4 Patogenia da Esclerose Múltipla.....	26
1.1.5 Diagnóstico e Tratamento	27
1.1.6 Etiologia da Esclerose Múltipla, possível papel virológico?.....	30
1.2 Os Retrovírus Endógenos.....	32
1.2.1 Organização genômica e classificação.....	33
1.2.2. HERV e seu papel no genoma humano.....	34
1.2.2.1. HERVs e doenças em humanos	35
1.2.2.2 HERV-W e a Esclerose Múltipla	36
2. Justificativa.....	40
3. Objetivos	41
4. Materiais e Métodos.....	42
4.1. Pacientes.....	42
4.2 Critério de exclusão	43
4.3 Casuística	43
4.4 Aspectos Éticos	43
4.5 Logística Laboratorial	43
4.6 Obtenção e purificação do RNA e síntese do DNAc	44
4.7 Padronização e PCR em Tempo Real para HERV-W	45
4.7.1 Quantificação relativa de expressão de HERV-W	47
4.8 Sequenciamento de nova geração	48
4.8.1 Sequenciamento de <i>amplicon</i> – Ion Torrent.....	48
4.8.1.1 Construção da biblioteca.....	48
4.8.1.2 Enriquecimento	49
4.8.1.3 Sequenciamento.....	50
4.8.2 RNASeq Illumina HiSeq	50
4.8.2.1 Construção da biblioteca e enriquecimento	50
4.8.3 Sequenciamento.....	51
4.9 Análises dos sequenciamentos.....	52
4.9.1 Análises de sequências da plataforma Ion Torrent	52
4.9.2 Análise do transcriptoma	52
4.10 Análise de vias de interação gênica	53
4.11 Quantificação de citocinas pró inflamatórias.....	54
4.12 Análises estatísticas.....	54
5. Resultados.....	55
5.1. Dados demográficos dos indivíduos incluídos no estudo.....	55
5.2 Análise da quantidade e integridade de RNA após o tratamento com DNase	56

5.3 Ensaio de PCR em Tempo Real para HERV-W	57
5.3.1 Frequência e quantificação relativa da expressão de HERV-W em pacientes com EM	59
5.4 Análises do transcriptoma.....	60
5.4.1 Sequenciamento de <i>amplicon</i> para gene env do HERV-W	60
5.4.2 Transcriptoma de HERV-W em indivíduos com Esclerose Múltipla	62
5.4.2.1 Procedimento de sequenciamento no Illumina HiSeq	62
5.4.2.2 Análise do transcriptoma	65
5.5 Análise dos <i>pathways</i> gênicos dos indivíduos com EM	76
5.6. Quantificação de citocinas inflamatórias e quimiocinas	83
6. Discussão.....	84
6.1 Pacientes e qualidade do RNA a ser estudado	84
6.2 Expressão relativa de HERV-W em indivíduos com EM.....	85
6.3 Transcriptoma dos pacientes com EM com foco no HERV-W	86
6.4 <i>Pathways</i> gênicos e quantificação de citocinas séricas.....	90
7. Conclusão.....	92
8. Referências.....	93
9. Anexos	110

EM	Esclerose Múltipla
SNC	Sistema Nervoso Central
SCI	Síndrome Clínica Isolada
RR	Forma Remitente Recorrente da Esclerose Múltipla
PP	Forma Primariamente Progressiva da Esclerose Múltipla
SP	Forma Secundariamente Progressiva da Esclerose Múltipla
OMS	Organização Mundial da Saúde
HERV	<i>Human Endogenous Retrovirus</i> ou Retrovírus Endógenos Humanos
SNP	<i>Single Nucleotide Polymorphisms</i> ou Polimorfismo de nucleotídeo único
IL2RA, IL7RA	Receptor alfa de interleucina 2, 7
SUS	Sistema Único de Saúde
BHE	Barreira Hemato-encefálica
LCR	Líquido cefalorraquidiano
PBMC	Células Mononucleares de sangue periférico
MOG	<i>Myelin Oligodendrocyte Glicoprotein</i> ou Glicoproteína de mielina de Oligodendrócito
PBM	Proteína Básica da Mielina
IL	Interleucina
DMT	<i>Disease Modifying Therapies</i> ou terapias modificadoras do curso da doença
HLA	<i>Human Leukocyte Antigen</i> ou Antígeno Linfocitário Humano
MHC-II	<i>Major histocompatibility class 2</i> ou complexo principal de histocompatibilidade classe 2
DNA	Ácido desoxirribonucleico
RNA	Ácido Ribonucleico
DNAc	Ácido desoxirribonucleico complementar
RNA_t	Ácido Ribonucleico transportador
RNA_m	Ácido Ribonucleico mensageiro
RNA_r	Ácido Ribonucleico ribossomal
CD	<i>Cluster of differentiation</i> ou agrupamento de diferenciação
IgG	Imunoglobulina G
ERV	<i>Endogenous Retrovirus</i> ou Retrovírus endógenos
Env	Gene/Proteína do Envelope
Pol	Gene/Proteína da Polimerase

LTR	<i>Long Terminal Repeat</i> ou Longas Repetições Terminais
GAG	<i>group-specific antigen</i> ou antígeno grupo-específico
MSRV	<i>Multiple Sclerosis-associated Retrovirus</i> ou Retrovírus associado a Esclerose Múltipla
GS	Grupo composto por pacientes com EM em condição de surto
GA	Grupo composto por pacientes com EM em condição avançada
GC	Grupo Controle
CATEM	Centro de Atendimento e Tratamento de Esclerose Múltipla
IMT-USP	Instituto de Medicina Tropical da Universidade de São Paulo
EDSS	<i>Expanded Disability Status Scale</i> ou Escala Expandida do Estado de Incapacidade
TCLE	Termo de Consentimento Livre e Esclarecido
EDTA	Ácido etilenoamino tetra-acético
RPM	Rotações por minuto
pb	pares de bases
PCR	Reação em Cadeia da Polimerase
CT	<i>Cycle Threshold</i> ou ciclo limiar
pH	Potencial de hidrogeniônico
ESALQ-USP	Escola Superior de Agricultura Luiz de Queiroz da Universidade de São Paulo
RIN	<i>RNA Integrity Number</i> ou Número de integridade do RNA
GTF	<i>Gene Transfer Format</i> ou Formato de transferência de genes
TNF	<i>Tumoral Necrosis Factor</i> ou Fator de Necrose Tumoral
IFN	Interferon
ILK	<i>Integrin Linked Kinase</i> ou quinase ligada a integrina
PEDF	<i>Pigment epithelium derived factor</i> ou Fator derivado da pigmentação de epitélio
NFκβ	<i>Nuclear Factor kappa beta</i> ou Fator Nuclear kappa beta
TLR	<i>Toll like receptor</i> ou receptor do tipo toll
TH	<i>T helper</i>
EIF2,4	<i>Eukaryotic initiation factor 1, 4</i> ou Fator eucariótico de iniciação 1, 4
VEGF	<i>Vascular Endothelial Growth Factor</i> ou Fator de crescimento vascular endotelial
miRNA	Micro RNA

ACP	Análise de componentes principais
RNASeq	Sequenciamento de RNA

%	Porcentagem
μl	Microlitro
°C	Graus Celsius
mL	Mililitro
Δ	Delta
α	Alfa
β	Beta
γ	Gama
g	gravidade
U	Unidades
M	Mol
Ng	Nanograma
μM	Micromolar

Tabela 1. Sinais e sintomas da Esclerose Múltipla de acordo com a região do SNC afetado, adaptado de Compston e Coles ⁵	22
Tabela 2. Volume e concentração dos reagentes utilizados para amplificação e quantificação relativa do gene env do HERV-W.....	46
Tabela 3. Dados demográficos dos indivíduos incluídos no estudo.	55
Tabela 4. Relação da quantidade de reads utilizadas após aplicação dos filtros	65
Tabela 5. Relação de vias alteradas nos pacientes do GS em comparação ao GC.	77
Tabela 6. Relação de vias alteradas nos pacientes do GA em comparação ao GC.....	78

Figura 1. Prevalência da EM no mundo estimada em números de casos/100.000 habitantes. Fonte: ¹⁰	24
Figura 2. Epidemia de EM nas ilhas Faroe durante a segunda guerra mundial. (A) Incidência dos casos de EM/100.000 habitantes no decorrer dos anos. (B) Mapa das Ilhas Faroe, em círculos os vilarejos nativos e linhas cruzadas mostram as bases das ocupações britânica.....	31
Figura 3. Microscopia eletrônica de células das leptomeninges de paciente com EM evidenciando uma partícula viral. Fonte: ¹³⁵	36
Figura 4. Imunohistoquímica de cérebro de paciente com EM para HERV-W. (A) Substância branca sem lesões; (B) Lesão precoce e ativa de EM com diversas células positivas para env do HERV-W em marrom. Fonte: ¹⁴⁰	37
Figura 5. Representação esquemática do alinhamento entre a proteína env do HERV-W (acima) e da proteína MOG (abaixo), linhas centrais demonstram os aminoácidos idênticos entre as duas proteínas e/ou aminoácidos distintos, porém com mesma carga. Fonte: ¹²⁸	38
Figura 6. Representação esquemática da organização das sequências associadas com os adaptadores e barcodes. Legenda: Adp: Adaptador, BC X: Barcode com a sua sequência escolhida.	49
Figura 7. Análise da integridade do RNA nas amostras antes e após o tratamento com DNase. (A) Quantidade de RNA antes e após dois tratamentos de DNase seguindo o protocolo turbo, $p=0,44$; (B) exemplo de gel de agarose corado com Gel Red (Uniscience) após os tratamentos com DNase, observa-se a integridade das bandas 28S e 18S do RNAr.....	56
Figura 8. PCR em Tempo Real para HERV-W em decaplicata. (A) Curva padrão com os gráficos de fluorescências no decorrer dos ciclos de cada replicata; (B) Curva padrão com valores de y-intercept: 37,65, slope: 3,341 e R2: 0,993.....	58
Figura 9. Análise de expressão diferencial de HERV-W em pacientes com EM e indivíduos controles.....	59
Figura 10. Box plot representando a quantidade de loci ativos em pacientes com EM em comparação ao GC.....	60
Figura 11. Box plot representando a quantidade de loci ativos em pacientes com EM do GS e GA.....	61

Figura 12. Eletroforese em Bioanalyzer para RNA total de parte das amostras. Observa-se dois principais picos, o de maior intensidade representando a subunidade do 28S e o de menor intensidade o do 18S. O RIN de cada amostra está descrito no interior de cada quadro azul. Os quadros cinzas representam ausência de amostra (cavidades vazias)	63
Figura 13. Imagem disponível no BaseSpace referente aos dados de uma das amostras submetidas ao sequenciamento (PET30).	64
Figura 14. Heatmap de expressão dos HERVs em pacientes com EM e em indivíduos do GC. Em destaque o HERV-W, anotado como HERV-17-int.	66
Figura 15. Volcano plot representando a quantidade de loci de HERV-W diferencialmente expressos. Em verde, loci diferencialmente expressos significativamente, enquanto loci em vermelho não apresentaram diferença estatística. Cada imagem representa uma comparação de dois grupos distintos, onde em: (A) Todos pacientes com EM contra o GC; (B) GS contra GC; (C) GA contra GC e; e (D) GA contra GS.	67
Figura 16. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes com EM comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3.....	68
Figura 17. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes do GS comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3.....	69
Figura 18. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes do GA comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3.....	70
Figura 19. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes do GS comparado aos indivíduos do GA. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3 e retângulo vermelho demonstra o locus 7q21.2.....	71
Figura 20. Análise de componente principal baseando-se em todas as reads obtidas de cada amostra. S = GS; A = GA, C= GC.	72

Figura 21. Análise de componente principal baseando-se em todas as reads de cada amostra. A) GS e GC, correspondendo a primeira corrida e B) GA e GC, correspondendo a segunda corrida. Em verde os indivíduos com EM, e em vermelho os indivíduos do GC.	73
Figura 22. Heatmap dos HERVs diferencialmente expressos na comparação GA contra GC. HERV-17= HERV-W.	74
Figura 23. Quantidade de reads dos loci de HERV-W mais frequentes nos indivíduos da primeira corrida (GS contra GC).	75
Figura 24. Quantidade de reads dos loci de HERV-W mais frequentes nos indivíduos da segunda corrida (GA contra GC).	76
Figura 25. Heatmap da análise de comparação de perfil gênico das vias alteradas dos indivíduos do GS e GA. Genes AGU = GS, Genes AVA = GA. Figura gerada automaticamente como output do programa IPA.	80
Figura 26. Representação esquemática da via da IL-6 dos indivíduos do GS. Genes em branco não apresentaram diferença significativa na análise do transcriptoma. Genes em rosa estão mais ativos do que nos indivíduos do GC. Ícones em roxo são citocinas em maior quantidade avaliadas na quantificação de citocinas no soro (próximo tópico). ...	81
Figura 27. Representação esquemática da via da IL-6 dos indivíduos do GA. Genes em branco não apresentaram diferença significativa na análise do transcriptoma. Genes em rosa estão mais ativos do que nos indivíduos do GC. Genes em verde estão menos ativos do que nos indivíduos do GC.	82
Figura 28. Quantificação de citocinas pró inflamatórias e quimiocinas. Legendas: EM SP = GA, EM Surto = GS, Controle = GC, *= $p < 0.01$	83

Anexo 1. Aprovação do comitê de ética da Faculdade de medicina da Universidade de São Paulo.	110
Anexo 2. Aprovação do comitê de ética da Faculdade de Medicina da Santa Casa de Misericórdia de São Paulo	112
Anexo 3. Termo de Consentimento Livre e Esclarecido fornecido aos voluntários da pesquisa.....	115
Anexo 4. Produção do aluno.....	121

Resumo

Nali LHS. *Perfil de expressão dos Retrovírus Endógenos humanos da família W em pacientes com Esclerose Múltipla* [Tese]. São Paulo: Instituto de Medicina Tropical da Universidade de São Paulo; 2018.

Introdução: A Esclerose Múltipla (EM) é uma doença autoimune desmielinizante que afeta drasticamente a capacidade motora, cognitiva, e sensitiva dos pacientes. Acredita-se que o Retrovírus Endógeno Humano da família W (HERV-W) possa ter um papel na patogênese da doença. Assim, o objetivo deste trabalho foi analisar o transcriptoma desses indivíduos, e analisar os *loci* do HERV-W diferencialmente ativos.

Materiais e Métodos: PBMC e soro de pacientes com EM em surto (GS), em condições avançadas (GA) e indivíduos saudáveis (GC) foram coletadas. Amplicons de envelope de HERV-W foram sequenciados em Ion Torrent e o RNAm foi sequenciado na plataforma Illumina HiSeq2500. Além da análise do HERV-W, análises de interação gênica foram feitas e citocinas inflamatórias e quimiocinas foram testadas.

Resultados: Foram analisados 23 indivíduos com EM (16 GS e 7 GA) e 36 do GC. Os pacientes com EM apresentam 3x mais expressão de HERV-W do que os indivíduos controle. O sequenciamento de amplicon revelou que os grupos com EM apresentavam mais *loci* ativos do que o GC. Apesar limitações decorrentes de variações entre corridas, o transcriptoma demonstrou que o HERV-K11 era diferencialmente expresso no GS, e no GA, 19 HERVs estavam diferencialmente expressos. *Loci* novos e já descritos como ativos em outros estudos foram encontrados no presente trabalho. O perfil de interação gênica do GS demonstrou um caráter inflamatório, confirmados pela dosagem de citocinas, onde IL-6, IL-1 β , TNF- α , IFN- γ estavam elevadas nos indivíduos do GS. Já os indivíduos do GA apresentavam um perfil não inflamatório com vias de reparo neuronal inativadas.

Conclusões: Os pacientes com EM apresentam maior nível de expressão e maior diversidade de expressão de HERV-W do que o GC. Apesar os perfis semelhantes de expressão, há *loci* diferencialmente expressos dependente do grupo estudado. Os pacientes com EM apresentam perfis distintos de expressão gênica onde os indivíduos GS apresentam um perfil inflamatório e no GA, um perfil neurodegenerativo.

Descritores: Esclerose Múltipla, Retroviridae, Transcrição gênica, Sequenciamento genético

Abstract

Nali, LHS. *Profile of human endogenous retroviruses W family expression in Multiple Sclerosis patients* [Thesis]. Sao Paulo: Instituto de Medicina Tropical da Universidade de São Paulo, 2018.

Introduction: Multiple Sclerosis (MS) is an autoimmune disease which drastically affects motor, cognitive and sensitive capability of the patients. It seems that Human Endogenous Retrovirus W family (HERV-W) may play a role in MS pathogenesis. Therefore, the aim of this study was to analyze the transcriptome of these individuals and to analyze the HERV-W *loci* differentially expressed.

Materials and Methods: PBMC and serum samples were collected from MS patients in relapsing conditions (GS), MS patients in advanced conditions (GA) and healthy individuals (GC). HERV-W Env amplicon was sequenced in Ion Torrent and mRNA was sequenced in illumina HiSeq2500 platform. Besides HERV-W analysis, genic pathways analysis was performed and inflammatory cytokines and chemokines were tested.

Results: A total of 23 MS patients (16 from GS and 7 from GA) and 36 from GC were enrolled in the study. MS patients presented 3-fold higher expression than healthy individuals. Amplicon sequencing revealed that MS groups presented more active *loci* than GC. Despite the limitations due to variations between sequencing runs, the transcriptome revealed that HERV-K11 was differentially expressed in GS, and 19 HERVs were differentially expressed in GA. New HERV-W *loci* and other *loci* that were reported as active *loci* previously are described here. The genic pathway analysis revealed that individuals from GS presented an inflammatory profile, also confirmed by cytokines dosage, where IL-6, IL-1 β , TNF- α , IFN- γ were significantly higher in GS. In the other hand, individuals from GA presented a non inflammatory profile with neuronal repair pathways inactivated

Conclusions: MS patients present higher level and diversity of HERV-W expression than GC. Regardless the similar profile of HERV-W expression in MS groups, there are differentially expressed HERV-W *loci* depending of each MS group. MS patients present distinct genic expression profile, where GS presented an inflammatory pathway with vascular permeability, whereas GA presented a neurodegenerative profile

Descriptors: Multiple Sclerosis, Retroviridae, Genic Transcription, Genetic Sequencing

1. Introdução

1.1 A Esclerose Múltipla

As primeiras descrições da Esclerose Múltipla (EM) começaram no século XIV¹. Achados clínicos e histológicos foram descritos posteriormente ² e, por fim, Jean-Martin Charcot uniu esses achados e intitulou a doença como Esclerose em Placas. Seu nome, por sua vez, vem das diversas “cicatrices” ou placas que se formam em virtude do dano ao tecido lesionado em diferentes regiões do Sistema Nervoso Central (SNC) ³. A primeira descrição da doença no Brasil foi publicada apenas em 1923 por Aluizio Marques ⁴.

A EM é definida como uma doença autoimune do SNC, caracterizada por lesões desmielinizantes que ocorrem de forma não padronizada, disseminadas no tempo e espaço. Manifesta-se, principalmente, por sinais neurológicos e possui um caráter evolutivo e posteriormente degenerativo, com sinais clínicos sucessivos causados por danos contínuos à mielina e aos axônios ⁵. Os sinais e sintomas da doença são extremamente variados, no entanto, estão relacionados a região do SNC a qual sofreu lesão desmielinizante. Alguns dos principais sinais e sintomas relacionados a região da lesão estão descritos na Tabela 1.

Tabela 1. Sinais e sintomas da Esclerose Múltipla de acordo com a região do SNC afetado, adaptado de Compston e Coles ⁵.

Área do SNC	Sintomas	Sinais
Cérebro	Deficiência cognitiva	Dificuldades de atenção, raciocínio e de execução
	Alterações de sensibilidade	
	Epilepsia	
	Déficit focal cortical	
	Depressão	
Nervo Óptico	Perda de visão unilateral e visão periférica	Redução da acuidade visual e reconhecimento de cores
	Visão borrada	Defeito pupilar relativo
	Dores oculares	aferente, Escotoma
Cerebelo	Tremor	Tremor postural e de ação, disartria
	Perda de equilíbrio	Descoordenação de membros e ataxia de marcha
Tronco encefálico	Diplopia, oscilopsia	Nistagmo, oftalmoplegia
	Vertigem	
	Dificuldades para deglutir	Disartria
	Dificuldades na fala e oscilações de humor	Paralisia Pseudobulbar

	Fraqueza	Sinais de neurônio motor superior
	Rigidez muscular e espasmos dolorosos	
Medula Espinhal	Disfunção de bexiga	
	Disfunção erétil	
	Constipação	
Outros	Dor, Fadiga, Intolerância a calor e a exercícios	

1.1.2 Classificações da Esclerose Múltipla

Apesar da maioria dos casos iniciar com um surto com associação de diversos sintomas neurológicos, o curso da doença é variável e está baseado em como a progressão clínica da doença se apresenta, sendo assim alguns tipos de critérios foram elaborados para determinar os diferentes fenótipos da EM ^{4,6}.

Os principais fenótipos de EM descritos foram: Síndrome Clínica Isolada (SCI); Forma Remitente Recorrente (RR), Forma Primariamente Progressiva (PP) e Forma Secundariamente Progressiva (SP). A SCI é definida como a primeira manifestação neurológica relatada que pode estar associada a EM e, basicamente, essa manifestação possui caráter inflamatório e desmielinizante ⁷. Quanto a forma RR, pode ser caracterizada pelos surtos e remissões no decorrer do tempo. Ainda, essa pode ser caracterizada de forma ativa e não ativa. Basicamente, a forma RR ativa caracteriza-se pelo surgimento de novas lesões e sintomas clínicos no período de 1 ano ⁶. As formas progressivas, por sua vez, possuem como principal característica o caráter degenerativo ao invés de inflamatório. A forma SP é caracterizada de forma retrospectiva baseada na piora gradativa comparada ao quadro inicial da forma RR ⁶. E por fim, a forma PP apresenta o mesmo perfil que a forma SP ^{6,8}, no entanto, precede a forma SP, característica de pacientes que evoluíram da forma RR em um tempo mais longo.

Apesar da excelente caracterização clínica dos tipos de EM, ainda não há marcadores para definir quais pacientes irão desenvolver os diferentes tipos da doença.

1.1.3 Epidemiologia

Apesar da EM representar uma das principais doenças neurológicas, dados epidemiológicos são escassos. Em 2008, a Organização Mundial da Saúde estimou que, no mundo, existiam cerca de 1.3 milhões de pessoas vivendo com EM (OMS, 2008). No entanto, em 2013, último período com dados mundiais publicados, estima-se que havia 2.3 milhões de pessoas vivendo com a doença¹⁰. Essa diferença pode ser explicada por um maior acesso a equipamentos de ressonância magnética utilizados para realizar o diagnóstico da doença bem como maior adesão de países participantes das pesquisas de prevalência¹⁰. A prevalência da EM parece apresentar um padrão geográfico, com maior prevalência no hemisfério norte^{9,10}.

Figura 1. Prevalência da EM no mundo estimada em números de casos/100.000 habitantes. Fonte: ¹⁰

Mais recentemente, um trabalho mais detalhado revisou a prevalência da EM na América Latina, e observou que o número de casos aparentemente aumenta nos países mais distantes da linha do Equador. Por exemplo, Argentina e Uruguai apresentavam prevalências de 25.6 e 21 casos/100.000 habitantes, respectivamente, enquanto Peru e Colômbia, apresentavam prevalências bem menores, com 3.8 e 5 casos/100.000 habitantes, respectivamente ¹¹.

No Brasil, os dados epidemiológicos são escassos, não há um estudo nacional que busque entender a fundo a prevalência de EM no país. No entanto, os dados brasileiros disponíveis demonstram um cenário onde a prevalência da EM é menor do que nos países do hemisfério norte, bem como alguns países da América do Sul, como Argentina e Uruguai ^{10,11}. Além disso, a prevalência da EM no Brasil parece não concordar com a latitude, uma vez que a prevalência é baixa na cidade de São Paulo, com 15,1 casos/100.000 habitantes ¹², corroborando com esse dado antigo, a cidade de Santos apresenta prevalência de 15,5 casos/100.000 habitantes ¹³, podendo ser menor ainda, como em João Pessoa, com 12 casos/100.000 habitantes ¹⁴. Em Belo Horizonte, foi descrito uma prevalência de 18,1 casos/100.000 habitantes ¹⁵ e em Uberaba 12,5 casos/100.000 habitantes ¹⁶. Por outro lado, temos algumas regiões com prevalências mais elevadas, como em Santa Maria, no Rio Grande do Sul, com 27,2 casos/100.000 habitantes ¹⁷ e em Volta Redonda, no Estado do Rio de Janeiro, com 30,7 casos/100.000 habitantes ¹⁸, sendo a maior descrita para o país. Tais dados sugerem que não há um padrão de distribuição de casos de acordo com a latitude do país.

A doença afeta em sua maioria indivíduos caucasianos ¹⁹ e é mais prevalente em mulheres ^{10,20}. Estima-se que existam pelo menos 2 vezes mais mulheres com EM do que homens ²¹, no entanto essa escala pode ser mais elevada, atingindo valores como 3,5:1 nos EUA ²², e, até valores 4,5:1 em São Paulo²³. Muitas teorias sobre quais poderiam ser os fatores que levaram a essa diferença foram levantadas. Tais teorias basearam-se em alguns fatores de risco previamente descritos como: maior número de mulheres fumantes ²⁴, respostas diferentes a fatores ambientais, como a vitamina D ²⁵ e uma possível associação do cromossomo X com a autoimunidade ²⁶. No entanto, nenhuma dessas possibilidades de fato explicou quais seriam as reais razões para uma maior prevalência entre as mulheres, muito menos em caucasianos.

1.1.4 Patogenia da Esclerose Múltipla

A patogenia da EM é basicamente causada por um processo autoimune inflamatório que leva a desmielinização e dano aos axônios. O processo pelo qual desenvolve-se a autoimunidade ainda não está totalmente estabelecido, porém, existem duas formas plausíveis para o desenvolvimento do processo: uma delas sugere a quebra da tolerância imunológica da mielina^{27,28}. Essa quebra de tolerância, conhecida amplamente como *bystander activation*, é originada pela presença de algum fator ambiental, ainda não conhecido, que levaria a uma resposta inflamatória inespecífica e exacerbada, onde por consequência os auto-antígenos poderiam ativar células dendríticas resultando em expansão clonal de linfócitos T auto-reativos^{27,29}. O segundo mecanismo que visa explicar a imunopatogenicidade da EM é pautado no processo de mimetismo molecular. Tal evento trata-se de um processo evolutivo onde organismos adquiriram similaridades genéticas a outros organismos, o que por sua vez levaria a aquisição de vantagens de sobrevivência. No entanto, o mimetismo molecular na EM pode gerar consequências aos indivíduos com a doença. No caso, antígenos de um agente externo que são similares a antígenos da mielina ou demais antígenos do SNC, poderiam induzir uma resposta imune inespecífica, tanto para agentes externos quanto para antígenos de algum tecido do hospedeiro, resultando em autoimunidade^{29,30}. Esse segundo fator será mais bem explorado no tópico HERV e a Esclerose Múltipla.

Ainda, há fatores genéticos que podem estar associados a EM. Por exemplo, um grande estudo procurou por possíveis *single nucleotide polymorphisms* (SNPs) associados a EM em famílias. Ao todo foram relatados 49 SNPs, desses, dois polimorfismos dentro do gene receptor alfa de interleucina 2 (IL2RA) e um no gene receptor alfa de interleucina 7 (IL7RA) foram fortemente associados a Esclerose Múltipla quando comparados a indivíduos controle³¹. Além disso, polimorfismos no gene do Antígeno Linfocitário Humano (HLA – do inglês *Human Leukocyte Antigen*) tem sido fortemente associado a uma maior susceptibilidade para o desenvolvimento da doença. O DRB1*15:01 tem o maior destaque, apresentando uma frequência cerca de 3 vezes maior em indivíduos com EM³².

Muito tem-se discutido sobre quais fatores ambientais que representam fatores de risco associados a patogênese da EM, dentre estes: poluição do ar ³³, deficiência de vitamina D ^{25,34,35}, pouca exposição aos raios ultravioleta ³⁶, fatores hormonais ³⁶ e infecções virais ³⁷⁻⁴⁰.

Independentemente do mecanismo no qual a autoimunidade é desencadeada, sabe-se que há um infiltrado de linfócitos T CD8+ e macrófagos nas regiões das lesões. Ainda, esses linfócitos podem ser ativados fora do SNC e depois passar através da barreira hemato-encefálica (BHE) ^{41,42}. Com o tempo, essa infiltração linfocitária se torna mais difusa, o que resulta em danos mais disseminados pelo SNC, levando a atrofia de substância branca e cinzenta ⁴³. Por sua vez, os linfócitos B podem participar da patogênese da EM, especialmente quando se diz respeito a produção de anticorpos oligoclonais no líquido cefalorraquidiano (LCR). Estes anticorpos, presentes na grande maioria dos pacientes com EM, são reativos a diversos vírus ^{44,45} e antígenos da mielina ^{42,46}. Ainda, os linfócitos B presentes nas células mononucleares de sangue periférico (PBMC) também apresentam reatividade a glicoproteína de mielina de oligodendrócito (do inglês - *Myelin Oligodendrocyte Glicoprotein* - MOG) e a proteína básica da mielina (PBM), podendo contribuir como células apresentadoras de antígenos ⁴⁷. A progressão da doença está relacionada com degeneração, consequência da constante desmielinização em lesões previamente remielinizadas, estresse oxidativo causado pelos diversos efeitos gerados pela inflamação e alterações nos canais de íon dos axônios ^{43,48}.

Considerando todos os fatores conhecidos até hoje, a EM trata-se de uma doença autoimune multifatorial progressiva, resultado da interação de um conjunto de fatores, envolvendo fatores epigenéticos e susceptibilidade genética ³⁵.

1.1.5 Diagnóstico e Tratamento

O diagnóstico da EM passou por diversas revisões no decorrer da história, onde foram adotados critérios clínicos, laboratoriais e de imagem para determinar parâmetros necessários para caracterizar a doença. Ainda, muitos desses parâmetros foram revisados e aprimorados a fim de melhorar a sensibilidade e especificidade do diagnósticos ⁴⁹⁻⁵². Por fim, o diagnóstico da doença é baseado na determinação da

disseminação das lesões no tempo e espaço, ou, quando as lesões se originaram e quais regiões foram afetadas, respectivamente ⁵¹.

A EM é tratada de acordo com as fases em que a doença se encontra, bem como a capacidade do paciente em responder aos diferentes tratamentos. Sendo assim, os surtos da doença, ou a forma aguda da doença, são tratados com altas doses de corticoesteróides^{53,54}. Esse tipo de tratamento pode reduzir a duração do surto, no entanto, não possui capacidade de retardar a progressão da doença, tampouco a probabilidade de novos surtos ⁵⁵. O tratamento a longo prazo pode ser conduzido com medicamentos imunossupressores ou com outros capazes de modificar o curso natural da doença. Dentre os imunossupressores, os principais exemplos são a Ciclofosfamida e a Azatioprina.

A Ciclofosfamida é um agente alquilante relacionado a mostarda nitrogenada, que se liga ao DNA impedindo a replicação celular⁵⁶. O medicamento é capaz de suprimir respostas inflamatórias potencializadas por IL-12 e resposta Th1 ⁵⁷, enquanto promove a ação de interleucina anti-inflamatórias, como IL-4 e IL-10 ⁵⁸. A Azatioprina por sua vez, é um agente não esteroide análogo de purina que atua em células de crescimento rápido, como linfócitos T e B, impedindo a sua proliferação⁵⁶.

Outra classe de medicamentos utilizados para o tratamento da EM são os chamados modificadores do curso da doença ou, do inglês *Disease Modifying Therapies* (DMT). Basicamente, o objetivo do tratamento com tais drogas é reduzir a quantidade de surtos e a progressão da doença. Essa é uma classe de medicamentos que tem crescido consideravelmente nos últimos anos, sendo que atualmente existem 13 medicamentos dessa classe disponíveis para tratamento ⁵⁹. Dentre eles, os de maior destaque são, acetato de glatirâmer, Interferon β -1^a, Natalizumabe e o Fingolimode.

O acetato de glatirâmer é um polímero sintético com estruturas do auto antígeno da Proteína Básica da Mielina (PBM). Basicamente, o mecanismo desse medicamento é mimetizar os antígenos da PBM e indiretamente competir com os antígenos da PBM, evitando assim que auto-antígenos se liguem ao complexo principal de histocompatibilidade de classe 2 (MHC II, do inglês, *major histocompatibility class II*) ⁶⁰, e em teoria, reduzir os efeitos inflamatórios da autoimunidade.

O Interferon 1 é produzido de forma endógena por células eucarióticas em resposta a infecções virais, no entanto, o Interferon β -1a é uma molécula sintética

produzida através de recombinação de DNA em células de mamíferos ⁶¹. O seu mecanismo para o tratamento contra a EM ainda não é totalmente compreendido, porém sabe-se que o Interferon β -1a é capaz de interferir nas sinalizações celulares, sejam elas parácrinas ou autócrinas ⁵⁹. Ainda, o medicamento foi capaz de elevar os níveis de citocinas anti-inflamatórias como IL-10 e IL-14, enquanto a produção de algumas citocinas pró-inflamatórias não apresentaram alterações ⁶¹. Por fim, acredita-se que esse medicamento possa também inibir a capacidade secretora de células B ⁶².

O Fingolimode é um dos novos medicamentos utilizados por via oral, que tem apresentado resultados promissores ⁶³. Trata-se de um composto derivado da miriocina, um metabólito produzido pelo fungo *Isaria sinclairii* ⁶⁴. O medicamento possui algumas ações, dentre elas: impedir o egresso de linfócitos ativados ou *naïve* dos linfonodos de volta para a circulação, e consequentemente, para o SNC. Isso é realizado através da internalização de receptores de esfingosina 1-fosfato em linfócitos, que são dependentes dessa sinalização para migrar entre tecidos ⁶⁵⁻⁶⁷.

Ainda, dentro da classe dos DMTs, existe o tratamento com anticorpos monoclonais. Dentre eles, destaca-se o Natalizumabe, que tem como alvo a integrina $\alpha 4\beta 1$ expressa na superfície de linfócitos T ⁶⁸. Esse anticorpo bloqueia a migração de linfócitos T através da BHE. Atualmente, esse medicamento representa uma das principais ferramentas para o tratamento da EM em sua forma RR, sendo capaz de reduzir consideravelmente a quantidade de surtos e a progressão da doença ⁶⁹. O Natalizumabe está disponível gratuitamente pelo Sistema Único de Saúde (SUS) para tratamento da forma RR da EM ⁷⁰. Outro representante do grupo dos monoclonais é o Ocrelizumab ^{71,72}. Esse monoclonal tem como alvo a molécula CD20, expressa em linfócitos B ⁷². Esse medicamento também demonstrou bons resultados na redução de surtos bem como na progressão da doença ⁷¹. Por fim, a escolha do tratamento a ser utilizado a longo prazo é baseada no histórico de tratamento do paciente e na progressão da doença.

1.1.6 Etiologia da Esclerose Múltipla, possível papel virológico?

Muito foi pesquisado sobre possíveis fatores de risco para a autoimunidade. Mas infelizmente, ainda não existe um fator que seja patognomônico da EM, como discutido no tópico “*Patogenia da Esclerose Múltipla*”. A EM é descrita como uma doença multifatorial onde há interações genéticas e de fatores ambientais. A susceptibilidade genética sozinha, aparentemente não é determinante para desencadear a autoimunidade, uma vez que podemos observar diferenças nas incidências entre gêmeos monozigóticos⁷³. Logo, os fatores ambientais são decisivos para desencadear a autoimunidade. Dentre os fatores ambientais há uma teoria sobre um possível papel infeccioso para a doença. Historicamente, essa hipótese surgiu nos anos 1900, com a ideia de que a doença poderia ser oriunda de uma zoonose⁷⁴. Cerca de 30 anos após, pesquisadores britânicos clamaram que haviam isolado um vírus, denominado *Spherula Insularis*, de LCR de pacientes com EM, no entanto os achados foram rapidamente contestados⁷⁴.

Apesar disso, a lista de vírus que foram pesquisados como possíveis agentes de uma etiologia autoimune da EM é vasta, dentre eles podemos incluir: Vírus da Raiva, Coronavírus, Sarampo, Vírus Torque Teno, praticamente todos da família herpesviridae (com exceção do 7 e do 8) e parainfluenza⁷⁵.

A hipótese de que há um fator ambiental infeccioso na patogênese da EM tem suporte epidemiológico, dado alguns exemplos de agrupamentos de casos descritos na literatura. O primeiro, mais conhecido, é o caso das Ilhas Faroe. Lá foi registrada uma epidemia de EM após a ocupação das tropas britânicas nas ilhas durante a segunda guerra mundial (Figura 2). Basicamente, as tropas britânicas levantaram acampamento em regiões muito próximas dos vilarejos dessas ilhas, e por fim, os indivíduos locais passaram a desenvolver sintomas característicos de EM após a ocupação⁷⁶.

Figura 2. Epidemia de EM nas ilhas Faroe durante a segunda guerra mundial. (A) Incidência dos casos de EM/100.000 habitantes no decorrer dos anos. (B) Mapa das Ilhas Faroe, em círculos os vilarejos nativos e linhas cruzadas mostram as bases das ocupações britânica.

Muito foi questionado a respeito das técnicas utilizadas para diagnosticar a EM durante aquele período, e com total razão, pois os avanços das técnicas de imagem foram determinantes para aprimorar o diagnóstico da doença na atualidade. No entanto, um estudo mais recente também demonstrou a presença de agrupamentos de casos de EM em um bairro na cidade de Ottawa no Canadá ⁷⁷. Ainda, essa não é a única cidade do Canadá a demonstrar esse aspecto de agrupamento de casos, em Winnipeg também foi observada essa característica ⁷⁸.

Apesar de não ter sido descrito um único agente etiológico da EM, os vírus representam fortes candidatos em virtude de consequências de interações entre eles e o hospedeiro. A princípio, existem vírus capazes de causar desmielinização por efeito diretamente citopático ⁷⁵, como é o caso do vírus JC, um poliomavírus ubíquo na população humana, causador da Leucoencefalopatia Multifocal Progressiva, uma doença desmielinizante ⁷⁹. Outro exemplo de doença desmielinizante é a pan-encefalite esclerosante subaguda, causada por uma infecção crônica pelo vírus do Sarampo ⁸⁰.

Além disso, efeitos imunopatológicos podem ser induzidos pela infecção viral. A capacidade dos vírus de exercer algum papel na regulação gênica do hospedeiro abre o leque de possibilidades de como estes poderiam desencadear alterações na resposta

imunológica. Ainda, destacam-se dois eventos autoimunes em que os vírus atuam, como o mimetismo molecular e *bystander activation*, já discutidos no tópico *Patogenia da Esclerose Múltipla*.

Uma das principais características clínico-laboratoriais de indivíduos com EM é a presença de bandas oligoclonais no LCR. As bandas oligoclonais são fruto de uma resposta imune humoral, IgG principalmente, presente no LCR⁸¹. A presença dessas bandas pode auxiliar no diagnóstico da EM, uma vez que essas são encontradas em cerca de 90% dos indivíduos com a doença^{51,81}. Um dado interessante é que boa parte dessas bandas apresentam anticorpos contra uma série de vírus simultaneamente, como por exemplo, Sarampo, Vírus da Varicela, Rubéola, Epstein-Barr e herpes simplex 1 e 2^{45,82-84}.

Por fim, há um vírus em particular, que tem demonstrado forte associação com a doença, o Retrovírus Endógeno Humano da família W (HERV-W).

1.2 Os Retrovírus Endógenos

Os retrovírus podem infectar diversas células somáticas de muitas espécies de vertebrados. No entanto, no decorrer da história, alguns desses retrovírus exógenos, infectaram células germinativas de grande parte das espécies de vertebrados⁸⁵. Como uma etapa do ciclo retroviral é integrar-se no genoma do hospedeiro, diversos eventos de integração viral no genoma de células germinativas ocorreram, sendo que algumas fixaram-se, gerando um indivíduo com esse provírus no genoma de todas as suas células⁸⁵. Esses retrovírus foram transmitidos no decorrer de gerações de forma Mendeliana, e foram então denominados Retrovírus Endógenos (ERV)⁸⁶⁻⁸⁸.

Diversos foram os eventos de integração de retrovírus exógenos distintos em nossos ancestrais a milhões de anos atrás⁸⁹⁻⁹². O HERV-W, por exemplo, endogenizou e fixou-se no genoma de um ancestral em comum há 40 milhões de anos, após a divergência entre os primatas do Novo Mundo dos primatas do Velho Mundo⁹³⁻⁹⁵.

No entanto, a identificação desses ERVs em humanos ocorreu apenas em 1970, revisado por Patience *et al.*⁸⁸. Posteriormente, após o sequenciamento do genoma humano, identificou-se que aproximadamente 8% deste apresentava sequências de

retrovírus endógenos, os quais foram então classificados como retrovírus endógenos humanos (HERV) ^{86,96,97}.

1.2.1 Organização genômica e classificação

Os Retrovírus endógenos podem ser classificados como retrovírus simples, pois apresentam apenas a estrutura principal genômica e característica dos retrovírus, três genes principais: o *gag*, *env* e *pol*, flanqueados por duas regiões promotoras idênticas denominadas *Long Terminal Repeats* (LTRs) ^{86,98}.

O gene *gag* recebe esse nome devido a expressão em inglês *group-specific antigen*, e as proteínas codificadas por ele são geralmente conservadas e menos imunogênicas ⁹⁸. Os peptídeos derivados dessa proteína são responsáveis pela produção do capsídeo viral ⁹⁸. O gene *env* é responsável por sintetizar as glicoproteínas do envelope viral. Essas são mais imunogênicas do que as do capsídeo, e por estarem expostas ao sistema imune do hospedeiro, comumente são alvo de grande pressão seletiva. O gene *pol* é responsável por codificar as enzimas virais, incluindo a transcriptase reversa e integrase ⁹⁸. Por fim, os ERVs possuem as regiões promotoras de transcrição, idênticas entre si, divididas em U3, R e U5 ⁹⁹. Essas possuem elementos regulatórios, promotores e intensificadores, sítios de ligação de fatores de transcrição e sinal de poliadenilação ¹⁰⁰.

Os ERVs foram então classificados em três grupos, baseados em sua similaridade com retrovírus exógenos. Portanto, os ERVs da classe I são relacionados aos gamaretrovírus e compreendem inúmeras famílias de HERVs, incluindo os ativos HERV-W e HERV-H. Os retrovírus da classe II, betaretrovírus, compreendem o retrovírus endógeno humano da família K (HERV-K), sendo esta a mais ativa família de HERVs. Finalmente, dentro da classe III (spumaretrovírus) estão o HERV-L e o HERV-S ^{101,102}.

A classificação dos ERVs está relacionada ao aminoácido correspondente ao sítio onde o RNAt se liga para iniciar a transcrição reversa, como por exemplo, o HERV que possui como sítio de início o triptofano, é classificado como HERV-W ⁸⁸.

1.2.2. HERV e seu papel no genoma humano

Os HERVs, no decorrer da evolução humana, sofreram diversos eventos mutacionais, onde houve um acúmulo de substituições, deleções, inserções, além de outros eventos como a presença de um *stop codon* dentro de uma região codificante ou até de perda de genes como um todo ¹⁰³. Todos esses eventos fizeram com que os HERVs fossem fragmentados e dispersos no genoma humano, portanto a formação de partícula viral, aparentemente não seria viável. No entanto, partículas virais podem ser formadas combinando as proteínas virais produzidas a partir de genes situados em *loci* distintos. Estas então interagem para realizar o empacotamento formando a partícula viral, evento comum para o HERV-K e HERV-W ¹⁰⁴⁻¹⁰⁷. Ainda, as proteínas virais isoladamente podem atuar em diferentes atividades fisiológicas humanas, bem como as regiões promotoras do vírus, as LTRs, que exercem papel regulador em genes adjacentes. Alguns desses exemplos serão descritos nesse tópico.

Como exemplo de genes regulados por LTR de HERVs, destacam-se o gene da apolipoproteína-C1, no fígado, o gene receptor de endotelina-B e o gene da pleiotrofina, ambos na placenta ¹⁰⁸. Adicionalmente, o HERV-E contribui com atividade intensificadora de transcrição (*enhancer*) para o gene da amilase humana, o que consequentemente permite o controle da produção de amilase pelas glândulas salivares ¹⁰⁹.

Outro exemplo, onde destaca-se o HERV-W, em especial a sua proteína do envelope. Um dos sítios do gene env do HERV-W está localizado no cromossomo 7q21.2 e é responsável pela produção da sincitina, proteína fundamental em uma etapa da embriogênese humana. Esta é responsável pela formação do sinciciotrofoblasto, auxiliando nas etapas iniciais da formação da placenta ^{110,111}. Ainda, existe outra proteína de HERV, dessa vez o HERV-FRD, que é responsável por sintetizar a sincitina 2, que também atua nessa etapa embriológica ¹¹². Aparentemente, a sincitina não atua exclusivamente na embriogênese. O potencial fusogênico da proteína também exerce papel na miogênese. Um estudo descreveu que animais machos *knockout* para sincitinas demonstraram uma perda significativa na massa muscular. Em posterior análise *ex vivo* foi observado que a proteína atuava na fusão dos mioblastos para a fusão dos miotubos, tanto na formação ou no reparo dos mesmos, demonstrando que a proteína exerce um papel sistêmico em tecido muscular e potencialmente em diferentes momentos da vida do indivíduo ¹¹³.

Ainda, os HERVs podem atuar indiretamente auxiliando no controle de algumas doenças. Acredita-se que a expressão de HERV-K possa interferir positivamente no controle da infecção pelo HIV em virtude de sua expressão mais elevada em indivíduos infectados por esse vírus. Basicamente, linfócitos T específicos para antígenos de HERV-K tem a capacidade de eliminar seletivamente células infectadas pelo HIV, que também expressem antígenos de HERV-K ¹¹⁴⁻¹¹⁶. Além disso, a expressão dos HERVs também pode auxiliar no controle do carcinoma renal. Alguns estudos demonstraram a expressão completa de env HERV-E exclusivamente em tecido tumoral de indivíduo com carcinoma renal ¹¹⁷. Os peptídeos virais expressos na superfície das células tumorais estimulam linfócitos TCD8+, que poderiam induzir uma resposta seletiva contra células cancerosas. Ainda, acredita-se que essa expressão possa ser responsável pela regressão de alguns tipos de tumores ¹¹⁸.

Por outro lado, os HERVs também podem ter participação no aparecimento/desenvolvimento de certas doenças, como será discutido a seguir.

1.2.2.1. HERVs e doenças em humanos

Apesar das inúmeras interações fisiológicas benéficas ao organismo, existem evidências que mostram que o aumento da expressão dos retrovírus endógenos possa estar associado a certas doenças autoimunes e diversos tipos de neoplasias. Como exemplo, foram descritas altas taxas de expressão dos HERVs em indivíduos com câncer de mama ^{119,120}, próstata ¹²¹, tumor cerebral ¹²², melanoma ¹²³, câncer de cólon ¹²⁴, linfoma ¹²⁰ e câncer gastrointestinal ¹²⁵. Entretanto, não é conhecido o real motivo que leva as altas taxas de expressão nesses indivíduos, nem mesmo o papel que essa expressão pode exercer na patogênese dessas neoplasias. Apesar de não haver evidências suficientes sugerindo um papel etiológico do HERV em câncer, há o potencial de utilização dos mesmos como marcador diagnóstico ou prognóstico em alguns tipos de câncer, como o de mama ^{126,127}.

Acredita-se que os HERVs possam estimular processos que levam a autoimunidade, possivelmente através de reações inflamatórias ou mimetismo molecular ^{75,128,129}, podendo portanto, exercer um papel na patogênese de algumas doenças autoimunes. Dentre as principais candidatas estão: o lúpus eritematoso

sistêmico ¹³⁰⁻¹³², artrite reumatoide ¹³³, síndrome de Sjögren ¹³⁴ e a esclerose múltipla, que será discutida a seguir em maiores detalhes.

1.2.2.2 HERV-W e a Esclerose Múltipla

A história do envolvimento do HERV-W com EM existe há 30 anos. Pesquisadores franceses isolaram e cultivaram células das leptomeninges circulantes no LCR e identificaram partículas virais com atividade de transcriptase reversa (Figura 3). Com a ascensão da retrovirologia na época, hipotetizou-se que essas partículas fossem oriundas de HIV ou HTLV, no entanto as células foram testadas com anticorpos monoclonais anti-p24 do HIV e anti-p19 do HTLV e foram negativas para ambos os vírus ^{135,136}.

Figura 3. Microscopia eletrônica de células das leptomeninges de paciente com EM evidenciando uma partícula viral. Fonte: ¹³⁶

Mais tarde o vírus foi clonado e caracterizado, diferentes genes foram identificados através dos clones e em então deu-se o nome ao vírus de MSRV (do inglês, *Multiple Sclerosis-associated Retrovirus*). Posteriormente, identificou-se que o MSRV, na verdade, fazia parte da família HERV-W ¹³⁷.

Uma vez com o vírus devidamente identificado, diversos trabalhos buscaram estudar os níveis de expressão em indivíduos com EM. Por análise de expressão gênica, identificaram expressão de HERV-W mais elevada em indivíduos com EM em comparação a indivíduos saudáveis. Tal achado foi identificado em diversos materiais biológicos, como nas placas escleróticas ^{138,139}, em PBMC ^{129,140-144} e um estudo

descreve a presença de RNA de HERV-W em LCR ¹⁴⁵. Boa parte desses trabalhos relataram expressão de HERV-W de cerca de 1.5 a 3 vezes maior do que indivíduos controles. Aparentemente, o nível de expressão de HERV-W é elevado em todas as formas clínicas da EM ¹⁴¹. Uma das possíveis causas dessa maior expressão de HERV-W em indivíduos com EM pode ser explicada pelo fato dos pacientes apresentarem maior quantidade de cópias de HERV-W no DNA genômico quando comparado a indivíduos controles ¹⁴⁶. No entanto, o inverso também pode ocorrer, por algum motivo, esse HERV está mais ativo e consequentemente pode se retro-inserir em outras posições do genoma, contribuindo para uma maior quantidade de sequências de HERV-W. Os HERVs, assim como os demais retroelementos possuem essa capacidade ¹⁴⁷.

A proteína do envelope do HERV-W é imunogênica *in vitro*, pois é capaz de estimular a produção de citocinas inflamatórias através dos efeitos agonistas iniciais de *toll like receptor 4* (TLR4). Esse processo pode desencadear uma complexa cascata de resposta inflamatória ^{148,149}. Além das alterações imunológicas e da expressão elevada de RNAm de HERV-W, não é incomum pacientes com EM apresentarem a expressão da própria proteína env do HERV-W em lesões desmielinizantes características da doença (Figura 4) ^{37,138,141,150}.

Figura 4. Imunohistoquímica para HERV-W em cérebro de paciente com EM. (A) Substância branca sem lesões; (B) Lesão precoce e ativa de EM com diversas células positivas para env do HERV-W em marrom. Fonte: ¹⁴¹

Perron e colaboradores descreveram o potencial imunopatogênico da proteína env do HERV-W. Os autores imunizaram camundongos com a MOG e em seguida expuseram grupos distintos a doses variadas de proteína env do HERV-W. Os autores descreveram então que os camundongos expostos a proteína env vieram a desenvolver encefalomielite alérgica experimental. Interessantemente, os graus da doença foram dependentes da dose administrada, ou seja, quanto maior a quantidade de proteína env do HERV-W, maior o grau da doença nos animais ¹⁵¹.

De fato, existem evidências *in silico* de que o HERV-W teria capacidade de exercer um papel autoimune através do mimetismo molecular. Um estudo realizado pelo nosso grupo encontrou 5 regiões de similaridade entre a MOG e a proteína env do HERV-W Figura 5 ¹²⁹.

```

HERV-W 449 LPFLGPLAAII 459
          +P LGPL A+I
MOG 215 VPVLGPLVALI 225

252 IRWVTPP-TRIVCL 264
      + W  PP +R+V L
66  VGWYRPPFSRVVHL 79

308 HVVPKPHNKRVP 319
      H  PH  RVP
194 HRTFDPHFLRVP 205

49  APSYRSLSK 57
      AP YR  ++
91  APEYRGRTE 99

219 LVGPLVS 225
      ++GPLV+
217 VLGPLVA 223

```

Figura 5. Representação esquemática do alinhamento entre a proteína env do HERV-W (acima) e da proteína MOG (abaixo), linhas centrais demonstram os aminoácidos idênticos entre as duas proteínas e/ou aminoácidos distintos, porém com mesma carga. Fonte: ¹²⁹

Curiosamente, uma das melhores alternativas para o tratamento da EM, o Natalizumabe é capaz de reduzir os níveis de expressão de HERV-W após um certo período de tratamento ¹⁵², bem como diminuir a resposta humoral anti-env do HERV-W ¹⁵³. Esses achados levaram então a produção de um anticorpo monoclonal, intitulado

GNbAC1¹⁵⁴. Esse anticorpo é direcionado ao domínio extracelular da proteína env do HERV-W e hoje representa um dos candidatos ao tratamento da EM. O anticorpo é seguro e atualmente encontra-se em estudo fase IIb¹⁵⁵⁻¹⁵⁷.

Apesar dos inúmeros dados de expressão gênica do HERV-W em indivíduos com EM, poucos foram os trabalhos que buscaram se aprofundar mais no transcriptoma desse e dos demais HERVs em indivíduos com EM. Esses trabalhos são contraditórios, muito provavelmente por utilizarem metodologias diferentes^{158,159}. Um deles utilizou-se de uma combinação de *primers* e sequenciamento de nova geração para caracterizar o perfil de HERV-W em amostras de cérebro de indivíduos com EM. No entanto, não foram encontradas diferença nos níveis e tampouco no perfil de expressão entre os indivíduos com a doença e indivíduos controles¹⁵⁹.

Sendo assim, torna-se de suma importância analisar o nível de expressão de HERV-W de forma mais ampla.

2. Justificativa

A EM é uma doença autoimune que acomete o SNC, caracterizada por seu perfil progressivo e degenerativo que muitas vezes leva a sequelas importantes, afetando drasticamente a capacidade motora, cognitiva e sensitiva dos pacientes portadores. Estudos de prevalência demonstram que a doença apresenta distribuição global sendo mais frequente em algumas regiões geográficas. Ainda, os tratamentos atuais não garante a não progressão da doença e o aparecimento de novos surtos.

Infelizmente, a EM não tem etiologia definida, porém no decorrer das incansáveis buscas para a compreensão da sua patogênese, o HERV-W demonstrou uma relação particularmente forte com a doença. Essa relação surgiu em virtude de inúmeras evidências que foram desde o isolamento do vírus e maiores níveis de expressão de HERV-W em indivíduos com EM até detecção da proteína viral em lesões ativas da doença e indução do modelo experimental da EM em camundongos expostos a proteína viral. Contudo, ainda estamos distantes de compreender o real papel dos HERVs, em especial o W, na patogênese da EM. Considerando a distribuição dos HERVs no genoma humano, bem como o impacto que esses *loci* possam representar na transcriptômica do hospedeiro, estudar esses vírus de forma pontual pode mascarar uma série de outros *loci* que também podem estar ativos, atuando na regulação da expressão gênica, interagindo com fatores ambientais, e talvez estimulando uma resposta imune direta. Sendo assim, é de suma importância que esses transcritos sejam identificados e em qual grau de expressão se encontram, para que assim possamos dar o primeiro passo para compreender o panorama de expressão de HERVs em indivíduos com EM e possivelmente avançar na compreensão da patogênese da doença.

3. Objetivos

Este trabalho teve como objetivo central comparar o transcriptoma, com ênfase no HERV-W, de indivíduos com EM e indivíduos saudáveis.

Objetivos específicos:

- Realizar o sequenciamento em larga escala dos RNA mensageiros obtidos de células mononucleares do sangue periférico dos indivíduos do estudo;
- Identificar e analisar, através de ferramentas de bioinformática, a presença, distribuição, nível de expressão e frequência de transcritos de HERVs, com ênfase na família W;
- Identificar os provírus que possuem nível de transcrição diferenciado entre os grupos analisados;
- Realizar a quantificação relativa do HERV-W em pacientes com Esclerose Múltipla;
- Investigar citocinas pró-inflamatórias no soro dos indivíduos do estudo;
- Determinar as vias de interação gênica ou de sinalização imunológicas alteradas em indivíduos com EM.

4. Materiais e Métodos

4.1. Pacientes

Os indivíduos incluídos no estudo foram pacientes diagnosticados com EM de acordo com os critérios de McDonald, adequados por Polman *et al.*,^{50,51}. Além disso, foram incluídos indivíduos saudáveis, adultos de ambos os gêneros sem histórico de EM na família ou imunossupressão no momento da coleta. Os indivíduos foram divididos em três grupos:

- Grupo Esclerose Múltipla em Surto (GS) – composto por indivíduos com EM em condição de surto antes do início do tratamento com corticoides. O diagnóstico do surto foi estabelecido por um Neurologista colaborador do estudo observando sintomas clínicos característicos de surto.

- Grupo Esclerose Múltipla Avançado (GA) – composto por indivíduos com EM com quadro mais avançado. Esses pacientes foram selecionados de acordo com alguns critérios: tempo de diagnóstico da doença maior que 10 anos, alta carga lesional, com pelo menos uma lesão em cada uma das regiões do SNC e *Expanded Disability Status Scale* (EDSS) maior que 6. Essa escala é uma ferramenta que avalia o grau de dano neurológico baseando em diversos sistemas funcionais neurológicos, essa escala varia de 0 a 10, onde 0 significa que o paciente não expressa nenhum dano neurológico, enquanto 10 seria morte causada pela EM¹⁶⁰. Portanto, pacientes com EDSS acima de 6 apresentam importantes lesões neurológicas.

- Grupo Controle (GC) – composto por indivíduos adultos, saudáveis e que não apresentassem histórico de doenças autoimunes na família.

Os pacientes dos grupos GS e GA são pacientes oriundos da Santa Casa de Misericórdia de São Paulo e são regularmente atendidos no ambulatório de Esclerose Múltipla do Centro de Atendimento e Tratamento de Esclerose Múltipla (CATEM). O GC foi composto por pesquisadores, funcionários e alunos de pós-graduação do Instituto de Medicina Tropical da Universidade de São Paulo (IMT-USP).

Os dados disponíveis no prontuário dos pacientes tais como o tipo da EM, EDSS, tratamento atual, data e duração do último surto, foram recuperados.

4.2 Critério de exclusão

Pacientes pertencentes ao GS que tivessem algum histórico de doença infecciosa ou câncer no momento da coleta, que não apresentassem o diagnóstico confirmado de surto, ou que tivesse iniciado o tratamento com imunossupressor antes da coleta seriam excluídos do estudo.

Pacientes do grupo GA que não apresentassem uma das características definidas como critério de inclusão, bem como terem câncer ou HIV, seriam excluídos do estudo.

Indivíduos do GC que tivessem qualquer histórico de EM ou demais doenças autoimunes na família seriam excluídos do estudo. Além disso, mulheres grávidas não foram incluídas no estudo.

4.3 Casuística

Para realizar as análises propostas foram coletadas 2 amostras de sangue de cada um dos voluntários em um único momento do estudo. Sendo, um tubo com ácido etilenodiamino tetra-acético (EDTA) para obtenção de PBMC para posterior análises moleculares e uma amostra em tubo seco com gel separador para obtenção de soro para a quantificação de citocinas.

4.4 Aspectos Éticos

O presente projeto foi submetido e aprovado pelos comitês de ética e pesquisa da Faculdade de Medicina da Universidade de São Paulo e da Irmandade Santa Casa de Misericórdia de São Paulo sob os números de protocolo 620.741 e 896.341, respectivamente (Anexo 1 e 2). Os participantes da pesquisa que concordaram em participar do projeto de estudo assinaram um Termo de Consentimento Livre e Esclarecido (TCLE) (Anexo 3).

4.5 Logística Laboratorial

As amostras coletadas foram enviadas ao Laboratório de Virologia do IMTUSP. Lá, as amostras foram identificadas com código alfanumérico e processadas.

As amostras de sangue em tubo EDTA foram centrifugadas a 1.500g por 10 minutos para separação do plasma, leucócito e hemácias, o *buffy coat* foi coletado e tratado com *Buffer EL (Erythrocyte Lysis Buffer – Qiagen)* para completa remoção de eritrócitos. Esse *buffer* rompe os eritrócitos rapidamente por choque hipotônico e térmico, os leucócitos intactos foram recuperados após a centrifugação e remoção do sobrenadante e assim as amostras foram armazenadas em freezer -80°C até o momento da etapa de extração de RNA.

As amostras de sangue em tubo seco foram também centrifugadas a 1.500g por 10 minutos para obtenção do soro através da separação pelo gel separador e armazenadas em freezer -20°C até o momento da sua utilização.

4.6 Obtenção e purificação do RNA e síntese do DNAc

Para obtenção de RNA para as análises de expressão gênica e sequenciamento de RNA, realizou-se a extração de RNA pelo método de Trizol-clorofórmio. Tal método baseia-se na quebra da membrana celular com 1mL de Trizol e posterior adição de clorofórmio na proporção 1:6. As amostras foram homogeneizadas em vórtex e centrifugadas a 15.000 rotações por minuto (RPM) por 15 minutos a 4°C. Após a centrifugação forma-se então a fase aquosa, que contém o RNA. Retirou-se toda a fase aquosa e o RNA foi então precipitado com Isopropanol 100% e lavado com Etanol 75%. Em ambas as etapas o material foi centrifugado a 15.000 RPM por 10 minutos a 4°C para precipitação do RNA e remoção do sobrenadante. Em seguida, os tubos contendo o RNA precipitado foram secos em temperatura ambiente, e o RNA foi eluído em 40µl de H₂O Nuclease-Free.

Naturalmente, após a extração de RNA ainda existem quantidades consideráveis de DNA genômico contaminante. Para eliminar esse DNA utilizou-se o kit *DNase-Free Turbo* (Ambion). Optou-se pelo método denominado “Turbo” de remoção de DNA genômico que consiste na incubação com DNase por um tempo mais prolongado que o protocolo convencional e com adição de DNase em dois intervalos distintos. Brevemente, esse protocolo consiste na adição de 0.1X volume de *DNase Buffer* e 2U de DNase e 40U *RNase-out Recombinant RNase inhibitor* (Thermo). Em seguida, as amostras foram incubadas a 37°C por 30 minutos. Posteriormente, foi adicionado mais 2U de DNase às amostras, e essas foram novamente incubadas a 37°C por 30 minutos. A inativação da enzima foi realizada com auxílio do *DNase Inactivation reagent* que

inativa a enzima e auxilia na remoção de cátions presentes nas amostras. Esse protocolo foi repetido para a completa remoção do DNA contaminante.

As amostras foram quantificadas em aparelho Nanodrop (Thermo) e para certificação da integridade do RNA, submetemos as amostras a eletroforese em gel de agarose 1,5%. Para corar o RNA utilizamos o *Gel Red* (Uniscience) 10000X, diluído na proporção de 1:3 em H₂O, com 0,1M de NaCl. Esse corante se liga ao RNA de forma eletrostática, tornando possível a visualização de material de fita simples como RNAm e algumas regiões do RNA ribossomal (RNAr).

A construção do DNA complementar (DNAc) foi realizada após rigorosa confirmação de ausência de DNA genômico nas amostras por PCR em Tempo Real com *primers* complementares a região do gene env do HERV-W (protocolo descrito a seguir). Portanto, utilizou-se o kit *High capacity cDNA Reverse Transcription* (Thermo Fisher) para construção do DNAc, seguindo as instruções do fabricante. A síntese do DNAc por esse kit está baseada na utilização de *primers* randômicos que se ligam ao RNA e com auxílio da enzima transcriptase reversa o DNAc é sintetizado. Nessa etapa também foram adicionadas 40U de *RNase-out Recombinant RNase inhibitor* (Thermo) para evitar a degradação do RNA durante o procedimento. As amostras, agora em DNAc, foram armazenadas em *freezer* -20 °C até o momento da utilização.

4.7 Padronização e PCR em Tempo Real para HERV-W

Para a análise de expressão gênica de HERV-W em pacientes com EM utilizamos *primers* disponíveis na literatura, os parâmetros foram readequados e a técnica foi padronizada. O método para o ensaio de PCR em Tempo Real adotado foi o *Sybr Green*. Esse método consiste na detecção do material nucleico de acordo com sua amplificação no decorrer dos ciclos. Para a detecção da amplificação do material nucleico utiliza-se o *Sybr Green*, que é um intercalante de DNA e se liga às pontes de hidrogênio conforme as fitas duplas vão sendo reassociadas ¹⁶¹.

Sendo assim, utilizamos o kit *Power Sybr Green PCR Master Mix* e *primers* complementares ao gene env do HERV-W, que amplifica um fragmento de 80pb ¹⁶² (ENV-F: CCAATGCATCAGGTGGGTAAC, ENV-R: GAGGTACCACAGACAAAAAATATTCCT).

As condições de ciclagem foram: 50°C por 2 minutos, 95°C por 10 minutos, seguidos de 40 ciclos de 95°C por 10 segundos, 50°C por 45 segundos, 60°C por 1

minuto. Após o término dos ciclos, uma etapa de dissociação foi realizada para correta discriminação dos produtos gerados através da curva de *melting*.

A descrição do volume e das concentrações dos *primers* utilizados estão descritos na Tabela 2.

Tabela 2. Volume e concentração dos reagentes utilizados para amplificação e quantificação relativa do gene env do HERV-W.

Reagente	Volume	Concentração final
H ₂ O	7,5µl	
2X Power Sybr Green Master Mix	12,5µl	1X
Primer Forward e Reverse 10µM	1µl cada	0,2µM
DNAc	3µl	500ng
Volume Total	25µl	

Para a construção da curva padrão, foi adquirido comercialmente um plasmídeo (GenScript), contendo uma sequência de 468 pares de bases (pb) compreendendo as regiões de interesse dos HERV-W e K. O material foi entregue em solução contendo 4µg de DNA do plasmídeo/inserto. Para a conversão de µg em cópias utilizou-se a seguinte fórmula, que considera o peso molecular bem como o tamanho do plasmídeo contendo o inserto (2710+468 nucleotídeos).

Fórmula: Peso em Daltons (g/mol) = (tamanho do plasmídeo com inserto)

Onde: (3178pb)(330Da x 2nucleotídeos/pb) = 2.097.480 g/mol

Logo: (g/mol)/número de Avogadro $6,02214199 \times 10^{23}$ = g / molécula = n de cópias

$2.097.480 \text{ g/mol} / 6,02214199 \times 10^{23} = 3,48 \times 10^{-18} \text{ g / molécula}$.

Assim, a concentração de plasmídeo (g/µl)/número de cópias = $(4\mu\text{g/ml}) / (3,48 \times 10^{-18} \text{ gramas/moléculas})$, o que resulta em um total de $1,17 \times 10^{12}$ cópias/ml.

Após obter a conversão de µg em cópias/mL, realizou-se diluições seriadas na ordem de 1:10 até a concentração de $1,17 \times 10^{-1}$ cópias/ml. A seguir, as amostras foram submetidas a Reação em Cadeia da Polimerase (PCR) em Tempo Real em decaplicata para análise de sensibilidade analítica. Ainda, três concentrações iniciais distintas de primers foram utilizadas (2µM, 5µM, 10µM) e verificadas quanto a reprodutibilidade dos *cycle threshold* (Ct) e sensibilidade. Por fim, controles negativos da reação, da

mistura dos reagentes utilizados para a construção do DNAc foram adicionadas em todas as reações a fim de averiguar a possibilidade de DNA contaminante.

4.7.1 Quantificação relativa de expressão de HERV-W

Em virtude do HERV-W ser considerado um gene endógeno não é possível calcular a sua expressão em quantificação absoluta, mas sim, mensurar a sua expressão com base em algum gene endógeno e com isso, calcular o *fold change* de expressão do HERV-W. Para isso utilizamos a fórmula $2^{-\Delta\Delta Ct}$, onde por exemplo:

$$2^{-(CT_{HERV-W \text{ Gi}} - CT_{\text{gene endógeno Gi}}) - (CT_{HERV-W \text{ GC}} - CT_{\text{gene endógeno GC}})}, \text{ onde Gi = GS ou GA.}$$

Como gene endógeno foi utilizado o gene hMLH1, que sintetiza uma proteína responsável pelo reparo de *mismatches* das bases durante a replicação do DNA¹⁶³.

Os *primers*, hMLH1-F: TGGTGTCTCTAGTTCTGG, hMLH1-R: CATTGTTGTAGTAGCTCTGC, amplificam um fragmento de aproximadamente 300 pb sob os seguintes parâmetros: 50°C por 2 minutos para inativação da Uracila n-glicosilase, 95°C por 5 minutos para desnaturação inicial, seguido de 40 ciclos de 95°C por 15 segundos e 60°C por 1 minuto para anelamento do primers e extensão da fita e mais um ciclo foi inserido para determinar a temperatura de *melting* de cada uma das amostras. As amostras foram testadas em duplicatas em cada uma das reações e o Ct foi obtido para a análise da expressão diferencial. Para o GC, foi calculado o valor médio do ΔCt , que foi adotado para comparar cada um dos pacientes dos demais grupos. As amostras foram consideradas positivas se: as amostras fossem positivas para o gene endógeno, o nível de fluorescência ultrapassasse o *cycle threshold* antes da diluição determinada como menor nível de sensibilidade para o HERV-W; os picos da temperatura de *melting* fossem os mesmos dos controles positivos de cDNA (76,4°C $\pm$ ou - 0,3).

4.8 Sequenciamento de nova geração

Como parte inicial do estudo, foi realizado um sequenciamento de *amplicon* a fim de identificar a quantidade de *loci* de HERV-W ativos em indivíduos com EM. Para isso foi desenhado um novo set de *primers* destinados a amplificar o env de HERV-W (WF: CCCATACCTCAAACCTCACC e WR:ACAGGGTCCAAAGAGGAGTA). Esses *primers*, por sua vez, amplificam um fragmento de aproximadamente 730bp sob os seguintes parâmetros: 95°C por 7 minutos para desnaturação inicial, seguido de 35 ciclos de 95°C por 1 minuto, 55°C para anelamento dos *primers* e 72°C para extensão da fita, por fim, 72°C para extensão final. Os pacientes que apresentavam expressão de HERV-W foram submetidos a esse segundo PCR para serem sequenciados na plataforma Ion Torrent. Essa etapa do trabalho teve caráter de estudo piloto para averiguar de forma pontual alguns *loci* ativos.

As amostras de RNA, por sua vez, foram encaminhadas ao RNASeq para obtenção de sequências para análise de transcriptoma.

4.8.1 Sequenciamento de *amplicon* – Ion Torrent

Essa plataforma de sequenciamento de próxima geração possui algumas etapas fundamentais de execução, que são: construção da biblioteca, enriquecimento da biblioteca e sequenciamento.

4.8.1.1 Construção da biblioteca

Os produtos de PCR das amostras positivas foram purificados com o *GFX PCR DNA and Gel Band purification kit* (GE) para a remoção de excesso de oligonucleotídeos ao redor de 20 bases e excesso de grupo fosfato. Para a construção da biblioteca precisava-se que as amostras estivessem em uma concentração entre 100ng e 1µg de DNA. Para isso as amostras foram quantificadas em aparelho Qubit com o kit Qubit dsDNA BR (*Broad Range*) Assay Kit.

As amostras foram submetidas a fragmentação mecânica em aparelho *Biorruptor* UCD200 (*Thermofisher*). Brevemente, as amostras permanecem em meio aquoso e são submetidas a três sessões de ondas ultrassônicas de 15 minutos cada. Esse

procedimento pode deixar bases extras/grupo fosfatos extras em uma das fitas, que devem ser então reparadas para garantir a ação da ligase. Para reparo das extremidades das fitas utilizamos o *Ion Plus Fragment Kit*.

A próxima etapa foi de ligação dos adaptadores e dos *barcodes* às extremidades das fitas com auxílio de enzima ligase. Os adaptadores são sequências que auxiliam no enriquecimento da biblioteca por PCR, pois são complementares a *primers* utilizados na etapa do PCR de emulsão, enquanto os *barcodes* são sequências distintas para cada amostra que são reconhecidas pelo software do sequenciador que irá identificar a qual amostra cada leitura pertence. Após a conclusão dessa etapa os fragmentos ficaram com uma estrutura semelhante a descrita na Figura 6.

Figura 6. Representação esquemática da organização das sequências associadas com os adaptadores e barcodes. Legenda: Adp: Adaptador, BC X: Barcode com a sua sequência escolhida.

Em seguida as amostras foram purificadas e submetidas a eletroforese em gel de agarose próprio do kit de construção (*E-gel*), com o objetivo de selecionar o tamanho de fragmento desejado de acordo com a biblioteca construída. Brevemente, todo o DNA contendo os adaptadores e *barcodes* atravessam o gel juntamente com um marcador de peso molecular. No caso desejava-se uma biblioteca com fragmento médio de aproximadamente 330pb, quando o tamanho de interesse atinge a cavidade inferior coletamos a solução contendo fragmentos de 330pb.

4.8.1.2 Enriquecimento

A etapa subsequente é o enriquecimento da biblioteca por PCR de emulsão. Essa metodologia é semelhante a PCR tradicional, exceto por ser realizado em uma emulsão e as amplificações ocorrerem na superfície de *beads*. O mix de PCR contém *beads* que irão servir de base sólida para a amplificação clonal dos fragmentos, e o óleo, que após o turbilhão promovido pelo equipamento *One Touch* (Thermo Fisher) gera a formação

de micelas onde as *beads* ficarão alocadas. Em seguida as *beads* contendo fragmentos amplificados são capturadas com auxílio de sondas de biotina.

4.8.1.3 Sequenciamento

A etapa de sequenciamento é realizada em um chip condutor. As *beads* contendo o DNA são inseridas em um chip que contém milhões de cavidades, sendo que o sequenciamento propriamente dito ocorre em ciclos. Nesses ciclos ocorre a introdução de um nucleotídeo por vez, e no caso de incorporação do nucleotídeo na fita nascente, é liberado um íon de hidrogênio que altera o potencial hidrogeniônico (pH) do meio na cavidade que a *bead* está localizada. Assim, o sequenciador identifica essa alteração e registra a incorporação daquele nucleotídeo. Logo, ao final dos 400 ciclos, teremos diversas sequências sintetizadas paralelamente.

4.8.2 RNASeq Illumina HiSeq

Para essa etapa, foram selecionadas algumas amostras de cada um dos grupos para realizar o transcriptoma. As amostras foram enviadas ao Laboratório de Genética da Escola Superior de Agricultura Luiz de Queiroz da Universidade de São Paulo (ESALQ USP) onde acompanhou-se todo o procedimento. Assim como o sequenciamento na plataforma Ion Torrent, o RNASeq no Illumina HiSeq conta com três principais etapas: construção da biblioteca, enriquecimento e sequenciamento.

4.8.2.1 Construção da biblioteca e enriquecimento

As amostras de RNA foram submetidas inicialmente a uma eletroforese no equipamento Bioanalyzer com o *kit Agilent RNA nano* (Agilent). Ao final da corrida é possível quantificar as amostras, bem como determinar a integridade do RNA, expressas em uma unidade denominada *RNA Integrity Number* (RIN), que avalia as áreas de fluorescência das duas subunidades do RNA ribossomal (RNAr) 28S e 18S. Posteriormente, as amostras foram quantificadas em aparelho Qubit com o *Qubit RNA HS Assay kit*.

Cerca de 100 ng de RNA foram utilizados para o preparo das bibliotecas. Utilizamos o *kit TruSeq Stranded Total RNA sample* (Illumina). O método adotado por esse kit é baseado na fragmentação iônica, que por possuir elevadas quantidades de Mg^{+2} garante a fragmentação controlada do RNA em altas temperaturas por um determinado tempo. Além disso, o kit, através de sondas com poli dT, captura somente sequências de RNA mensageiro (RNAm) deixando para trás sequências de RNAr. Por fim, foi utilizado ainda um kit de remoção de sequências de globina.

As sequências de RNA são então retrotranscritas, e após a síntese do DNAc, realiza-se o processo de adenilação, onde é adicionada uma Adenina na extremidade 3' das fitas de RNA. Esse processo é crucial para a ligação dos adaptadores que possuem uma Timina complementar em sua extremidade 5'. Esses adaptadores servem para identificar a origem de cada leitura, bem como possui regiões complementares para fixar-se na *flowcell* para a realização do sequenciamento e regiões complementares para o enriquecimento da biblioteca. Após a ligação dos adaptadores, as bibliotecas são então enriquecidas através de PCR com *primers* complementares aos adaptadores. Por fim, as bibliotecas são então validadas em equipamento *Bioanalyzer* e quantificadas por PCR em Tempo real.

4.8.3 Sequenciamento

As bibliotecas são normalizadas a 10nM de DNAc para a correta construção e posterior dispersão dos *clusters* pela *flowcell*. Os *clusters* de sequenciamento são pequenos grupos de DNA que são dispersos pela *flowcell*. Constituem aproximadamente em 1000 cópias de DNA e são construídos em aparelho *cbot* com o kit *HiSeq PE Cluster kit v4 cbot* (Illumina). Nessa primeira etapa são adicionados nucleotídeos não marcados para formação das pontes de amplificação clonal que serão posteriormente utilizadas para o sequenciamento de ambas fitas.

O conceito de sequenciamento pelo sistema Illumina é baseado na metodologia de Sanger, porém de forma reversiva. Basicamente, todos os nucleotídeos marcados são inseridos na *flowcell* que se alinham com a fita molde, ocorrendo a captura da fluorescência, e ao final de cada ciclo, essa base marcada com o fluorocromo é removida, permitindo a ligação de um novo nucleotídeo. Ao final da síntese da fita e captura de todas as bases o processo é repetido para a síntese da fita reversa. O kit utilizado para o sequenciamento foi *HiSeq SBS v4 kit* (Illumina).

Ao final do sequenciamento pode-se observar os primeiros controles de qualidade das amostras. Um arquivo fastq, que possibilita a análise de qualidade de cada uma das *reads*, é gerado para cada amostra.

4.9 Análises dos sequenciamentos

4.9.1 Análises de sequencias da plataforma Ion Torrent

As *reads* geradas pelo sequenciador foram exportadas em um arquivo fastq. A partir de então, toda a etapa de análise foi realizada no programa *CLC Genomics Workbench 7.0* (www.clcbio.com).

As extremidades das *reads* foram removidas para desconsiderar as sequências oriundas de adaptadores. As sequências de baixa qualidade também foram descartadas pelo software. Para a análise dos transcritos presentes nos pacientes, um banco contendo 138 sequências do gene *env* correspondendo aos *loci* de HERV-W foi construído. Para encontrar tais sequências, utilizou-se como *query* a sequência da Sincitina na ferramenta de busca de sequências humanas BLAT (<https://genome.ucsc.edu/FAQ/FAQblat.html>). Por fim, separamos sequências de *env* de HERV de diferentes *loci* do genoma. As sequências obtidas de cada paciente foram mapeadas contra esse banco seguindo padrões rigorosos, sendo eles: 99% de similaridade e custo máximo de inserções, deleções e *mismatches*.

O total de *reads* de HERV-W oriundas de cada cromossomo foi utilizado para determinar os transcritos mais frequentes bem como quantidade de *loci* mais ativos nos diferentes grupos de pacientes estudados. As análises estatísticas e elaboração de gráficos foram realizadas no programa R ¹⁶⁴ e para as comparações das variáveis utilizamos o teste de wilcox.

4.9.2 Análise do transcriptoma

Os arquivos fastq foram alinhados com o genoma humano (GRCh38/hg38) como referência, no programa *STAR Aligner* ¹⁶⁵. Esse programa possui algoritmos específicos para alinhamentos de sequências oriundas de RNA. Alguns parâmetros de qualidade foram adotados para o alinhamento dessas sequências: *reads* pequenos (<20 bases) e de baixa qualidade foram excluídos, além disso para garantir a estridência

elevada, utilizamos *reads* que apresentassem *single hit*, ou seja, mapeados apenas uma vez, e que apresentassem no máximo dois *mismatches* em relação a referência. Ainda para garantir a quantidade elevada de *reads* aumentou-se o limite de mapeamentos múltiplos, isso se dá pelo fato de estarmos analisando dados oriundos de elementos repetitivos, e assim seria possível obter quantidade de *reads* mapeadas mais relevantes dos HERVs.

Em seguida, as amostras foram analisadas utilizando um arquivo *Gene Transfer Format* (GTF) desenvolvido por Jin *et al.*,¹⁶⁶. Esse arquivo é ideal para análise de elementos repetitivos pois incluem esses elementos ao final da análise, sendo que a maioria dos outros GTFs disponíveis desconsideram esses elementos, o que inviabiliza as análises. As contagens das *reads* utilizadas bem como o mapeamento dos mesmos foram baseados no programa Salmon¹⁶⁷.

Por fim, utilizou-se o DeSeq2¹⁶⁸ para normalizar os dados e determinar a expressão diferencial dos HERVs bem como de outros genes diferencialmente expressos em relação aos grupos a serem comparados. Os dados foram então exportados e os gráficos montados no programa R (<https://www.r-project.org>).

4.10 Análise de vias de interação gênica

Para compreender o perfil do transcriptoma dos indivíduos com EM bem como o papel que algumas vias alteradas teriam na fisiologia desses pacientes, o programa *IPA Ingenuity* (Qiagen) foi utilizado. Com esse programa é possível analisar o perfil de expressão dos genes e a sua relação com vias canônicas, bem como interação entre vias distintas. Logo espera-se ao final das análises compreender o panorama de vias possivelmente alteradas de indivíduos com EM em duas condições distintas.

Sendo assim, todo o painel de genes analisados pelo programa DESeq2 foi exportado em uma planilha contendo informações importantes como Gene ID, *log2foldchange* e o valor de p. Após carregar essa planilha no IPA, a *core analysis* foi realizada. Essa determina quais são as principais vias alteradas, possíveis relações com doenças, bem como *pathways* de interações gênicas dessas vias. Ainda, para refinar as análises, estabeleceu-se como ponto de corte apenas os genes diferencialmente expressos ($p < 0.05$), e apenas as vias determinadas experimentalmente foram selecionadas.

As vias são expressas em imagens com quadros conectados e cores que diferenciam os genes com maior e menor expressão, bem como um sumário das funções de cada uma das vias alteradas.

4.11 Quantificação de citocinas pró inflamatórias.

Selecionou-se, por conveniência, um painel de algumas citocinas pró inflamatórias (IL-6, IL-1 β , TNF- α , IFN- γ), quimiocinas IP/CXCL10, MCP-1 e CD14 solúvel (CD14s). Neste ensaio a tecnologia Luminex xMAP foi utilizada. Essa plataforma consiste em dosar diversos analitos simultaneamente. Ainda, a técnica é semelhante ao ELISA, no entanto envolve o uso de microesferas de poliestireno marcadas com dois fluoróforos, que são utilizadas para a identificação e diferenciação de cada uma das esferas, as quais são posteriormente identificadas pelo equipamento. Adicionalmente, as esferas são marcadas com anticorpos específicos (contra cada um dos analitos descritos). Após o analito se ligar aos anticorpos de captura localizados na superfície das esferas, a detecção é realizada através de um terceiro marcador fluorescente, conjugado ao anticorpo de detecção. Por fim, o equipamento movimenta as esferas em fila única através de dois *lasers* diferentes em um citômetro de fluxo, onde o primeiro identifica a esfera e o segundo quantifica o sinal de reporte em cada esfera.

Para a análise do CD14s utilizou-se o *kit Quantikine ELISA* para CD14 solúvel (RnD Systems).

4.12 Análises estatísticas

Para analisar a diferença estatística de expressão de HERV-W, da quantificação de citocinas, e quantidade de *loci* ativos pela metodologia de sequenciamento de amplicon, foi utilizado o teste de Wilcox.

5. Resultados

5.1. Dados demográficos dos indivíduos incluídos no estudo

Ao todo, foram incluídos no estudo 23 pacientes com Esclerose Múltipla e 36 indivíduos saudáveis. A descrição detalhada dos pacientes encontra-se na Tabela 3.

Tabela 3. Dados demográficos dos indivíduos incluídos no estudo.

Dados demográficos	Grupos Esclerose Múltipla(n)		Indivíduos saudáveis(n)
Grupos	GS (16)	GA (7)	GC (36)
Gênero	14 Mulheres	4 Mulheres	26 Mulheres
	2 Homens	3 Homens	10 Homens
Idade(média)	18-61 (33,5)	24-60 (40)	20-57 (30,5)
Tempo de Surto* (media)	2-14 dias (5 dias)	NA	NA
EDSS(média)	1-4,5 (2)	6-8 (6,5)	NA
Tipo de EM•	15 RR	6 SP	NA
	1 SIC	1RR	
Tempo de doença(média)	1-20 anos (4,5 anos)	10-20 anos (15 anos)	NA
Tratamento atual	5 S/ Tratamento		
	5 Avonex		
	1 Natalizumabe		
	1 Copaxone	6 Natalizumabe♦	NA
	1 Rebiff	1 Copaxone	
	1 Fingolimode		
	2 β Interferon		

*Aplicável apenas para GS. • CIS= Síndrome Clínica Isolada, RR= Remitente Recorrente, SP = Secundariamente progressiva, aplicável apenas para GS e GA. NA= não se aplica. ♦ 5 desses receberam menos de 5 infusões do medicamento.

5.2 Análise da quantidade e integridade de RNA após o tratamento com DNase

Após a extração de RNA, todas as amostras foram tratadas com o protocolo turbo de DNase por duas vezes para completa remoção do DNA genômico. Devido a agressividade do protocolo (tempo e temperatura de exposição das amostras), adicionamos 2U de *RNAse out inhibitor* (Thermo). Após os tratamentos, uma eletroforese foi feita para analisar as bandas do RNA ribossomal e o RNA presente nas amostras antes e após os tratamentos com DNase foi quantificado (Figura 7 A e B).

Figura 7. Análise da integridade do RNA nas amostras antes e após o tratamento com DNase. (A) Quantidade de RNA antes e após dois tratamentos de DNase seguindo o protocolo turbo, $p=0,44$; (B) exemplo de gel de agarose corado com Gel Red (Uniscience) após os tratamentos com DNase, observa-se a integridade das bandas 28S e 18S do RNAr

Não foi observada perda significativa de RNA. Em média, as amostras precisaram ser tratadas com DNase turbo por pelo menos duas vezes para que houvesse completa remoção do DNA contaminante. Ainda, o RNA presente nas amostras demonstrava-se estar íntegro conforme a figura 7B e portanto, em condições de ser utilizado.

5.3 Ensaio de PCR em Tempo Real para HERV-W

Uma análise paralela com poucas replicatas da nossa diluição seriada do plasmídeo demonstrou que a concentração final de 0,2 μ M dos *primers* foi a mais reprodutível. Ainda, a reação de decaplicata demonstrou uma positividade em pelo menos 90% das réplicas na diluição contendo 117 cópias/mL, portanto o teste apresentou uma sensibilidade de 117 cópias/mL. Ainda, a análise de regressão linear da curva padrão revelou valores bem satisfatórios com y-intercept: 37,65, slope: 3,341 e R^2 : 0,993 (Figura 8 A e B). Vale ressaltar que esse teste foi utilizado para verificar a ausência de DNA genômico em todas as amostras após os tratamentos com DNase.

Figura 8. PCR em Tempo Real para HERV-W em decaplicata. (A) Curva padrão com os gráficos de fluorescências no decorrer dos ciclos de cada replicata; (B) Curva padrão com valores de y-intercept: 37,65, slope: 3,341 e R2: 0,993.

5.3.1 Frequência e quantificação relativa da expressão de HERV-W em pacientes com EM

Todos os pacientes com EM avaliados, tanto do GS quanto do GA, apresentaram alguma expressão do gene env do HERV-W. Com relação ao GC, 31 dos 36 indivíduos apresentaram expressão de HERV-W.

Apesar da frequência elevada de expressão do gene env de HERV-W nos indivíduos controles, notamos que esses apresentavam baixos níveis de expressão, enquanto os pacientes do grupo GS e GA apresentavam cerca de 3 x mais expressão do vírus do que os indivíduos do GC ($p < 0,01$). A análise dos níveis de expressão entre os grupos dos pacientes com EM revelou que os indivíduos do GA apresentam uma expressão ligeiramente maior do que as dos indivíduos GS, no entanto sem diferença estatística significativa ($p = 0,6$) (Figura 9).

Figura 9. Análise de expressão diferencial de HERV-W em pacientes com EM e indivíduos controles.

5.4 Análises do transcriptoma

5.4.1 Sequenciamento de *amplicon* para gene *env* do HERV-W

Essa etapa consistiu em um estudo piloto para compreender a diversidade de expressão do HERV-W em indivíduos com EM. Sendo assim, foram selecionados aleatoriamente 10 pacientes com EM, sendo 4 do GS e 6 do GA além de 10 indivíduos do GC que expressavam HERV-W. Todos os pacientes do GA estavam sob tratamento com Natalizumabe, enquanto do GS todos estavam sob tratamento com Avonex. Para o grupo GS, foram selecionados 3 mulheres e 1 homem, para o GA, foram selecionados 5 mulheres e 1 homem e para o GC foram selecionados 6 mulheres e 4 homens.

Após o sequenciamento, as *reads* obtidas foram mapeadas contra o banco de *loci* de *env* HERV-W. Foram considerados *loci* ativos quaisquer *loci* de HERV-W que apresentasse *reads* que contemplassem ao menos alguma região do fragmento de referência e com uma cobertura de pelo menos 3 vezes.

Então, foi possível observar que os pacientes com EM, de uma forma geral, apresentavam cerca de 15 *loci* ativos, enquanto os indivíduos do GC que expressavam HERV-W apresentaram em média 10 *loci* ativos ($p < 0.01$) (Figura 10).

Figura 10. Box plot representando a quantidade de *loci* ativos em pacientes com EM em comparação ao GC.

Ainda, observou-se que os transcritos mais expressos em pacientes com EM foram: 1p34.2, 6q27, 9q31.3, 14q21.2 e Xq22.3, enquanto nos indivíduos saudáveis os *loci* mais ativos foram o 1p34.2, 6q27, 7q21.2 (Sincitina) e 14q21.2. Buscou-se entender a diferença na diversidade de transcritos de env de HERV-W presente entre os grupos dos indivíduos com EM (Figura 11). Os indivíduos com GA aparentemente apresentam uma diversidade ligeiramente maior do que as dos indivíduos do GS. Ainda, os *loci* que apresentaram maior número de *reads* no grupo GS foram 1p34.2, 6q27 e 14q21.2, enquanto os *loci* mais ativos no grupo GA foram Xq22.3, 1p34.2 e 6q27 e 9q31.2.

Figura 11. Box plot representando a quantidade de *loci* ativos em pacientes com EM do GS e GA.

Nesse caso pudemos observar perfis completamente distintos: os pacientes do grupo GA apresentavam uma maior diversidade de expressão de HERVs em comparação aos pacientes do GS ($p < 0.01$).

Uma análise detalhada do alinhamento do banco contra os *primers* utilizados revelou que havia *mismatches* na região de ligação dos *primers* das sequências de

alguns cromossomos, sugerindo que alguns transcritos possam ser amplificados parcialmente, portanto com menor eficiência, ou ainda que outros possam nem estar sendo detectados. Mesmo assim, vale ressaltar que foram observadas diferenças de *loci* ativos entre os grupos, e podemos assumir que, se determinados *loci* são amplificados com certo viés, isso deve ocorrer em ambos os grupos. Portanto, as diferenças observadas não devem refletir problemas experimentais, mas sim, diferenças reais entre os objetos. De qualquer forma, essa metodologia claramente não é adequada para se estudar a fundo a diversidade de transcritos de HERV.

5.4.2 Transcriptoma de HERV-W em indivíduos com Esclerose Múltipla

5.4.2.1 Procedimento de sequenciamento no Illumina HiSeq

Para dar prosseguimento ao sequenciamento das amostras, uma eletroforese em *Bioanalyzer* foi realizada para a obtenção do RIN. Brevemente, esse valor representa uma escala obtida através de análise de alguns fatores que vão desde o pico de absorbância da região 28S e 18S do RNAr até oscilações de absorbância antes do primeiro pico e áreas de cada pico das bandas do RNAr. O valor obtido varia de 0-10, onde o valor do algoritmo obtido representa proporcionalmente a integridade de RNA¹⁶⁹. Sendo assim, estabeleceu-se que um valor ideal para o sequenciamento e posterior análise dos dados seria de 7,0¹⁷⁰. Todas as amostras apresentaram valores de RIN satisfatórios variando de 7,3 a 9,3, vide Figura 12.

Figura 12. Eletroforese em Bioanalyzer para RNA total de parte das amostras. Observa-se dois principais picos, o de maior intensidade representando a subunidade do 28S e o de menor intensidade o do 18S. O RIN de cada amostra está descrito no interior de cada quadro azul. Os quadros cinzas representam ausência de amostra (cavidades vazias)

As amostras foram então sequenciadas. Ao final do sequenciamento as *reads* ficam disponíveis na nuvem em uma plataforma fornecida pela empresa desenvolvedora, o *BaseSpace*. Lá podemos observar as quantidades e tamanhos das *reads* obtidas em cada uma das amostras. Todas elas apresentaram acima de 30 milhões

de *reads* na forma *pair-ended*. Um exemplo de uma das corridas está descrito na Figura 13.

PET30 Launch App	
Sample name	PET30
Date created	06/28/2015 3:16:03
Genome	
Paired end	Paired
Number of Clusters	21,813,737
Number of Reads	43,627,474
Read 1 length	126
Read 2 length	126
Size	2.88 GB

Figura 13. Imagem disponível no BaseSpace referente aos dados de uma das amostras submetidas ao sequenciamento (PET30).

Após o alinhamento e o mapeamento das *reads* com o genoma humano utilizando os filtros de estringência, verificamos que mais de 87% das *reads* foram mapeadas corretamente, alcançando uma cobertura de até 46x do transcriptoma humano, Tabela 4.

Tabela 4. Relação da quantidade de reads utilizadas após aplicação dos filtros

Amostra	<i>Paired-Reads</i> mapeadas	<i>Reads</i> Mapeadas após filtros de estringência	% utilizada para análises	Cobertura em vezes
PET10	22.961.092	21.591.157	94,03366791	43,93
PET12	19.797.357	18.124.313	91,54915477	36,80
PET15	21.886.468	20.664.788	94,41810346	42,05
PET16	17.953.569	16.025.894	89,26299835	32,61
PET27	22.970.517	21.539.649	93,77084982	43,83
PET28	16.736.064	14.662.326	87,60916545	29,83
PET29	24.375.119	22.870.542	93,82740655	46,54
PET30	21.086.819	19.858.084	94,17297128	40,41
PET8	22.323.082	20.654.878	92,52699963	42,03
PET9	18.654.953	16.826.476	90,19843684	34,24

5.4.2.2 Análise do transcriptoma

A análise dos dados teve início com uma parceria com a equipe de bioinformática do Hospital AC Camargo. Para alinhar as *reads* utilizou-se o programa STAR *aligner*¹⁶⁵, utilizando o genoma hg19 como referência; para mapear os elementos transponíveis foi empregado o programa TETranscript¹⁶⁶. E finalmente para análise de expressão diferencial utilizou-se o programa DESeq2¹⁶⁸.

Buscou-se compreender quais *loci* de HERV-W estavam mais ativos em cada um dos grupos, ou seja, qual o perfil transcricional desses elementos em indivíduos com EM e nos controles. Sendo assim, foi analisado inicialmente o panorama geral de expressão dos HERVs. Para isso, as quantidades de *reads* de cada um dos *loci* foram transformadas em uma escala de log2, basicamente a função `rlog()` normaliza os dados e calcula a expressão diferencial dos genes a serem analisados. Esse dado foi utilizado para a construção dos *heatmaps*.

Com essa análise inicial pode-se observar que os HERVs das famílias 1, H, 9, K9, L, L74, K22, IP, 3 e K3 são os mais expressos nos indivíduos com EM. Ainda, a

análise revelou que o HERV-W, anotado como HERV-17-int também faz parte dos HERVs mais ativos nos indivíduos do nosso estudo, Figura 14.

Figura 14. Heatmap de expressão dos HERVs em pacientes com EM e em indivíduos do GC. Em destaque o HERV-W, anotado como HERV-17-int.

Posteriormente, os *loci* diferencialmente expressos do HERV-17 foram estratificados em *heatmaps* separados para verificar quais eram de fato ativos dentro da família.

Uma forma de identificar quais *loci* estavam ativos seria plotar os dados de log de *fold change* contra o valor de p de cada *loci* em um gráfico do tipo *volcano* (*volcano plot*), assim identificar quais *loci* estavam mais e quais estavam menos ativos em cada grupo. Sendo assim, realizou-se diferentes comparações: EM total contra GC, GA contra GC, GS contra GC e GA contra GS. Os dados podem ser visualizados na Figura 15.

Figura 15. *Volcano plot* representando a quantidade de *loci* de HERV-W diferencialmente expressos. Em verde, *loci* diferencialmente expressos significativamente, enquanto *loci* em vermelho não apresentaram diferença estatística. Cada imagem representa uma comparação de dois grupos distintos, onde em: (A) Todos pacientes com EM contra o GC; (B) GS contra GC; (C) GA contra GC e; e (D) GA contra GS.

Identificou-se, na Figura 15, uma série de *loci* de HERV-W diferencialmente expressos em todas as comparações. Ainda, verificou-se posteriormente a localização de cada um desses *loci* diferencialmente expressos, que serão apresentados a seguir.

Inicialmente, os dados comparativos de EM contra GC foram utilizados para construir um *heatmap*, (Figura 16). Nessa análise verificou-se que existiam 16 *loci* diferencialmente expressos, sendo 6 mais ativos e 10 menos ativos. Ainda, verificou-se que os *loci* de HERV-W das famílias 1 e 2 (ERVW-1 e ERVW-2) estavam expressos em indivíduos com EM, no entanto o ERVW-1 era o mais frequente. Uma análise detalhada demonstrou que 75% dos *loci* mais ativos eram oriundos da LTR de HERV-W. Ainda, dentre os *loci* mais ativos, destacam-se os Xq22.3, 2q11.2, 12p13.31 (Figura16).

Figura 16. Heatmap dos *loci* de HERV-W diferencialmente expressos dos pacientes com EM comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o *locus* Xq22.3.

Em seguida foram comparados os indivíduos do GS e indivíduos do GC, Figura 17. Assim como na análise inicial, foram vistos mais *loci* ativos nos indivíduos GS do que nos indivíduos GC. Identificamos 24 *loci* diferencialmente expressos significativamente, sendo 6 menos ativos e 18 mais ativos. Ainda, detectou-se a expressão de ambas as famílias de HERV-W, sendo que a família ERVW-1 foi a mais frequente. Além disso, boa parte dos *loci* diferencialmente expressos eram originários de Gag (50%) e LTR (50%). Os *loci* mais expressos no GS em comparação ao grupo GC foram: 1p32.2, 2q35, 11q14.1, (Figura 17).

Figura 17. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes do GS comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3.

A próxima comparação foi entre indivíduos do GA e indivíduos do GC (Figura 18). Assim como nas análises anteriores, os indivíduos GA apresentavam mais loci ativos significativamente do que os indivíduos do GC. Foram identificados um total de 45 loci diferencialmente expressos, sendo 11 menos ativos e 34 mais expressos. Desta vez, o gene mais expresso foi o env (70%). Ainda, a família de HERV-W mais frequentemente expressa foi a ERVW-1 e dentre os loci mais expressos, destacaram-se os loci 6q21, Xp11.2 e Xq22.3, (Figura 18).

Figura 18. *Heatmap* dos *loci* de HERV-W diferencialmente expressos dos pacientes do GA comparado aos indivíduos do GC. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o *locus* Xq22.3.

Por fim, a comparação entre os indivíduos GS e GA mostrou que os indivíduos do GA apresentavam mais *loci* ativos em comparação aos indivíduos GS (Figura 19). Ao todo foram 100 *loci* diferencialmente expressos, sendo 38 mais expressos em GS e 62 mais expressos no GA. O gene *env* foi o mais prevalente, expresso em 62% e 67% em indivíduos com GS e GA respectivamente. Ainda, a família de HERV-W mais expressa foi a ERVW-1, e dentre os *loci* mais expressos destacam-se 1p34.2, 5p14.3, Xq21.2 do grupo GS e os *loci* 1q22, 3q25.1, 9p13.3, Xq22.3 no grupo GA, (Figura 19).

Figura 19. Heatmap dos loci de HERV-W diferencialmente expressos dos pacientes do GS comparado aos indivíduos do GA. Pontos em vermelho representam ERVW-1, enquanto pontos em azul demonstram ERVW-2. Retângulo azul demonstra o locus Xq22.3 e retângulo vermelho demonstra o locus 7q21.2

Interessantemente, foi observado um perfil de expressão distinto dependente do grau de progressão da doença em que o paciente se encontra. E, o gene env foi mais frequente quando comparamos os grupos de EM entre si. Adicionalmente, foi encontrado o locus 7q21.2(Sincitina), um envelope do HERV-W, mais ativo nos indivíduos do GS do que nos indivíduos GA. Por outro lado, os indivíduos do GA apresentaram outros loci mais ativos em relação ao GS, como por exemplo, o Xq22.3.

Na segunda metade do meu doutoramento, realizei um estágio na *Unidad de Bioinformática* do Instituto Pasteur de Montevideu, Uruguai. Lá tive como objetivo adquirir *expertise* em análise de bioinformática, com ênfase em transcriptômica. Sendo assim, algumas análises foram refeitas e realizamos uma Análise de Componentes Principais (ACP) para verificar a distribuição do perfil transcricional dos indivíduos como um todo. Observamos que a maioria dos indivíduos dos grupos com EM

agrupavam-se separadamente dos indivíduos do GC. No entanto os indivíduos do grupo GC agrupavam-se dependente dos grupos relacionados a cada corrida de sequenciamento. Brevemente, foram realizadas duas corridas de sequenciamento, a primeira contendo os pacientes do GS e alguns indivíduos do GC, e a corrida posterior incluiu pacientes do GA e novos indivíduos do GC (Figura 20).

Figura 20. Análise de componente principal baseando-se em todas as reads obtidas de cada amostra. S = GS; A = GA, C= GC.

Ao retornar, conversamos com o bioinformata responsável da ESALQ, local onde foram realizadas as corridas do sequenciamento, para tentar entender o que poderia estar acontecendo. Identificou-se que as amostras dos indivíduos da primeira corrida apresentavam uma contaminação com RNAr. Mesmo após tentativas de remoção de transcritos ribossomais das amostras da primeira corrida, algumas

apresentaram quantidade de *reads* muito baixas, o que poderia desviar o eixo do perfil transcricional destas amostras. Assim, decidimos remover algumas amostras da primeira corrida para as análises posteriores (8, 9 do GS; 12, 16 e 28 do GC), (Figura 21). Após a remoção destas, bem como dos *reads* de genes ribossomais das demais, observou-se que em ambos os ACPS os indivíduos com EM agruparam de forma independente dos indivíduos do GC.

Figura 21. Análise de componente principal baseando-se em todas das *reads* de cada amostra. A) GS e GC, correspondendo a primeira corrida e B) GA e GC, correspondendo a segunda corrida. Em verde os indivíduos com EM, e em vermelho os indivíduos do GC.

Devido a essa inesperada intercorrência, decidimos reanalisar o perfil transcricional dos HERVs de cada corrida separadamente. A primeira análise, GS contra GC, demonstrou apenas 1 HERV diferencialmente expresso, o HERV-K11. Enquanto na análise do GA contra o GC, encontrou-se 19 HERVs diferencialmente expressos (incluindo os HERVs 17, 9 e algumas subfamílias do K), Figura 22.

Figura 22. Heatmap dos HERVs diferencialmente expressos na comparação GA contra GC. HERV-17= HERV-W.

Apesar do único HERV diferencialmente expresso na comparação do GS contra o GC, ter sido o HERV-K11, verificou-se quais eram os *loci* com maior número de *reads* do HERV-17 (HERV-W). Esses *loci* bem como a quantidade de *reads* para cada um dos pacientes pode ser observada na Figura 23. Ainda, a maioria dos *loci* expressos foram do ERVW-1 (90%) e dentre os genes expressos, destaca-se o env em 80% dos casos.

Figura 23. Quantidade de *reads* dos *loci* de HERV-W mais frequentes nos indivíduos da primeira corrida (GS contra GC).

Com relação a análise dos *loci* de HERV-W nos indivíduos do GA, foi encontrado um perfil de expressão com alguns *loci* também prevalentes nos indivíduos do GS, como o 4q21.22, 6q27, 7q21.2(sincitina), 11q14 e 15q21.3 (Figura 24). Assim como na comparação anterior o ERVW-1 foi o mais frequente e os genes mais expressos por sua vez, foram env e LTR, em 52% e 31% dos casos, respectivamente. Apesar das semelhanças entre os *loci* mais frequentemente expressos, a atividade transcricional desses *loci* HERV-W aparentemente é maior nos indivíduos do GS, dado o número das *reads* do grupo mencionado.

Figura 24. Quantidade de *reads* dos *loci* de HERV-W mais frequentes nos indivíduos da segunda corrida (GA contra GC).

5.5 Análise dos *pathways* gênicos dos indivíduos com EM

Foram selecionados para essa análise apenas genes diferencialmente expressos de forma significativa ($p < 0.05$). Sabendo do problema de viés de corrida, duas análises distintas foram feitas, uma baseada na corrida 1 (GS em comparação ao GC) e outra baseada na corrida 2 (GA em comparação a GC). Foram identificados 353 genes diferencialmente expressos nos indivíduos do GS e 365 genes diferencialmente expressos nos pacientes do GA.

Realizamos a *core analysis* desses genes de forma individual para cada uma das corridas. A análise dos genes diferencialmente expressos entre GS em comparação ao GC revelou 12 principais vias canônicas alteradas com suporte biológico para a doença. Todas as vias, bem como as funções biológicas de cada gene estão descritas na Tabela 5. Além disso, o valor de Z representa uma ferramenta matemática para determinar os desvios-padrões de uma base de dados, expressão gênica nesse caso, dentro de uma via a ser analisada. Portanto, valores positivos podem estar relacionados de forma empírica

a uma ativação da via, enquanto valores negativos podem estar associados a uma inativação da via estudada.

Tabela 5. Relação de vias alteradas nos pacientes do GS em comparação ao GC.

Via alterada	Função Biológica	Valor de Z
Sinalização de extravasamento linfocitário	Diapedese durante processos inflamatórios	+3,05
IL-12	Resposta inflamatória	+0,06
IL-7	Resposta Inflamatória	+1,342
ILK (<i>Integrin linked kinase</i>)	Conexão de moléculas de integrinas ao citoesqueleto	+1,897
PEDF(<i>Pigment epithelium derived factor</i>)	Fator neurotrópico promotor de sobrevivência neuronal	+1,633
Sinalização de Integrinas	Moléculas de adesão celular, adesão celular de barreira hematoencefálica	+2,121
NFκβ	Atuação em resposta imunológica adaptativa e inata	+2,828
IL-6	Resposta inflamatória	+2,646
IL-2	Proliferação e homeostase de linfócitos	+2,00
TH1	Resposta intracelular contra microrganismos	+1,32

Maturação de células dendríticas	Potencializador de resposta inata e adaptativa	+2,646
Paxilina	Mobilidade celular, adesão em matriz extra-celular	+2,45

A comparação de vias alteradas dos indivíduos do GA em comparação ao GC revelou por sua vez 15 principais vias canônicas alteradas, com suporte biológico. Todas as vias, funções biológicas de cada via e valor de Z estão descritos na Tabela 6.

Tabela 6. Relação de vias alteradas nos pacientes do GA em comparação ao GC.

Via alterada	Função Biológica	Valor de Z
EIF2 (<i>Eukaryotic initiation factor</i>)	Apoptose, dano celular	+2,6
EIF4	Crescimento celular, proliferação e desenvolvimento	-1,13
mTOR	Crescimento celular, proliferação e desenvolvimento	-0,33
VEGF (<i>Vascular endothelial growth factor</i>)	Crescimento e desenvolvimento de células endoteliais, vasculares e linfáticas	-2,6
IL-1	Resposta inflamatória	-3,3
iNOS	Sinalização celular	-2,6
P13k (AKT)	Proliferação e crescimento celular	-2,7
IL-6	Resposta Inflamatória	-3,05

Receptor de células B	Ativação de células B	-3,1
Toll Like Receptor	Resposta humoral e celular	-2,12
Fator de ativação de linfócitos B	Ativação de células B	-2,23
NGF (<i>Neuronal Growth Factor</i>)	Crescimento e proliferação de células neuronais	-3,31
Neuregulina	Manutenção, crescimento e reparo de células neuronais	-1,4
IL8	Resposta inflamatória	-3,3
Neurotrofina	Crescimento neuronal	-2,6

Outras vias, sem suporte biológico que estavam alteradas em ambas as análises estão ligadas ao diabetes, alguns tipos de cânceres, hipertrofia cardíaca e resposta a antígenos por vias aéreas.

Finalmente, realizou-se uma análise de comparação das vias alteradas. Essa análise utiliza-se do valor de Z para cada uma das vias estudadas em ambas as análises as agregando para determinar perfis de expressão gênica de populações distintas. Assim, utilizamos essa comparação para averiguar mais detalhadamente o perfil de expressão gênica dos pacientes GS e GA em comparação aos indivíduos controles. De acordo com a Figura 25 verificou-se de início uma série de vias ativadas no GS em comparação ao GA. Obviamente, muitas dessas vias não apresentam suporte biológico para a doença estudada, porém vale ressaltar o valor de Z score positivo nos indivíduos do GS, e o valor de Z negativo para essas vias nos indivíduos do GA para diversas vias de resposta inflamatória, IL6, IL-8, por exemplo, permeabilidade endotelial (extravasamento linfocitário), moléculas de reconhecimento celular (*reelin*, integrinas) e de fatores de crescimento e manutenção neuronal (*Neuronal Growth Factor*, Neurotrofina).

Figura 25. Heatmap da análise de comparação de perfil gênico das vias alteradas dos indivíduos do GS e GA. Genes AGU = GS, Genes AVA = GA. Figura gerada automaticamente como output do programa IPA.

Interessantemente, foi possível observar que os genes envolvidos nas vias alteradas, fossem eles mais ou menos ativos, em cada uma das análises (GS e GA) não foram os mesmos. Esse exemplo fica mais claro ao observarmos as Figuras 26 e 27.

Figura 26. Representação esquemática da via da IL-6 dos indivíduos do GS. Genes em branco não apresentaram diferença significativa na análise do transcriptoma. Genes em rosa estão mais ativos do que nos indivíduos do GC. Ícones em roxo são citocinas em maior quantidade avaliadas na quantificação de citocinas no soro (próximo tópico).

৯৭

5.6. Quantificação de citocinas inflamatórias e quimiocinas

Ao todo foram quantificadas 4 citocinas pró inflamatórias, as quimiocinas MCP-1, IP/CXCL10 e CD14s em todos os grupos. Verificamos que as 4 citocinas pró inflamatórias analisadas (IL-6, IL-1 β , TNF- α , IFN- γ) estavam elevadas nos indivíduos do GS em comparação aos indivíduos do GC ($p<0,01$). Por outro lado, a quimiocina MCP-1 estava mais elevada nos indivíduos do GA em comparação aos indivíduos do GC, Figura 28.

Figura 28. Quantificação de citocinas pró inflamatórias e quimiocinas. Legendas: EM SP = GA, EM Surto = GS, Controle = GC, *= $p<0,01$.

Com o objetivo de contribuir com a validação de determinadas vias inflamatórias, os dados observados aqui também foram utilizados para a análise nos *pathways* gênicos (tópico 5.5 Análise dos *pathways* gênicos dos indivíduos com EM). Por exemplo, na figura 27 observamos que a via de sinalização da IL-6 estava alterada, com uma série de genes mais expressos em relação aos indivíduos do GC. Ainda, dentro dessa mesma via a IL-6 e o TNF- α , encontravam-se aumentadas no nosso ensaio, sendo que essas moléculas podem interagir com os genes dessa via.

6. Discussão

6.1 Pacientes e qualidade do RNA a ser estudado

Os pacientes do grupo GS foram selecionados de acordo com a sua condição clínica dependente do surto, enquanto os pacientes do GA foram selecionados com base em: progressão clínica da EM e os critérios adotados pelo nosso grupo para caracterizar esses pacientes em condição mais extrema, como tempo de doença e a quantidade de lesões no SNC. Esses critérios podem ter interferido na homogeneidade da idade entre os grupos, onde em média, os indivíduos do GS tinham 33,5 anos, os do GA 40, e o GC 30,5. No entanto, essa diferença é de certa forma esperada. Uma extensa revisão avaliou os dados clínicos de diversos pacientes e mostrou que apesar do início da doença ser praticamente o mesmo para pacientes nas formas RR e SP (29 anos, em média), aqueles que evoluíram para a forma SP tinham, em média, 39,1 anos no momento ¹⁷¹. Sendo assim, a diferença entre as idades dos presentes grupos pode ser explicada, talvez, pelo tempo do curso natural da doença.

Outro fator que vale a pena ser destacado é que grande parte dos indivíduos SP (pertencentes ao grupo GA) falharam em suas terapias anteriores e hoje estão sob a terapia com Natalizumabe. De fato, 6/7 pacientes do GA recebem esse medicamento. Sabe-se que o Natalizumabe pode atuar na modulação gênica dos indivíduos tratados, portanto poderia representar um impacto negativo no estudo. Pode-se citar algumas situações onde o tratamento com Natalizumabe pode alterar a expressão de genes. Por exemplo, a expressão de alguns miRNAs, como no caso do miR-155 e miR-132, let-7c, miR-125a-5p e miR-642, e algumas citocinas pró-inflamatórias são modulados pelo Natalizumabe após um período de 6 meses de tratamento ^{172,173}, outros após 12 meses, como no caso dos miR-320, miR-320b e miR-629 ¹⁷³. Além disso, o Natalizumabe também pode aumentar a expressão de alguns genes como POU2AF1/Spi-B ¹⁷⁴ após 2 anos de tratamento. Um pouco mais preocupante é a já descrita capacidade do Natalizumabe de reduzir a expressão de HERV-W e também a resposta humoral anti-HERV-W em pacientes com EM que tratam por longos períodos, no caso >24 meses ^{152,153}. Felizmente, todos os pacientes do GA incluídos no estudo migraram para o Natalizumabe pouco tempo antes da coleta, sendo que todos receberam menos de 5 infusões, com exceção do indivíduo 41, que recebeu cerca de 45 infusões. Ainda, todos os pacientes do GA apresentavam expressão de HERV-W e de acordo com a análise de

quantificação relativa, os indivíduos do GA expressam HERV-W em níveis semelhantes ao dos indivíduos em surto e superiores ao de indivíduos saudáveis (Discussão no tópico 6.2). Portanto, é plausível que o tratamento tenha representado pouco ou nenhum impacto na expressão diferencial dos indivíduos incluídos no estudo.

6.2 Expressão relativa de HERV-W em indivíduos com EM

Identificamos que todos os indivíduos com EM, independentemente do grupo, estavam expressando HERV-W. Por outro lado, 31/36 (86,1%) dos indivíduos do GC também expressavam HERV-W. A elevada frequência de expressão em indivíduos com EM é esperada, porém a frequência de expressão é considerada alta para o GC quando comparada a estudos anteriores que, não detectaram, ou detectaram em baixa frequência a expressão em indivíduos saudáveis ^{99,135,140,141,150}.

Diversas podem ser as explicações para essa diferença. Por exemplo, alguns *primers* desenhados e utilizados em estudos anteriores ¹⁴⁴ foram específicos para o *locus* 7q21.2 (Sincitina). Além disso, o ensaio de PCR foi padronizado novamente, e ao final das análises observamos que o limite de sensibilidade foi de 117 cópias/ml, podendo portanto ter contribuído para uma maior frequência de detecção de RNAm dos provírus. Por fim, descarta-se a presença de DNA contaminante, pois um rigoroso controle de qualidade foi adotado a fim de evitar a contaminação por DNA genômico, e as etapas de construção de cDNA e do PCR em Tempo Real foram realizadas apenas após a confirmação da ausência do DNA.

Apesar da frequência elevada de expressão do HERV-W em indivíduos do GC, observou-se que o nível de expressão nesses indivíduos era basal. Por outro lado, os indivíduos com EM (GS ou GA) apresentaram um nível de expressão 3x maior do que indivíduos saudáveis. Esse dado corrobora grande parte dos estudos descritos até agora, onde os pacientes com EM apresentam uma expressão mais elevada em comparação a indivíduos saudáveis ^{99,129,139,141,144,149}.

Além disso, foi observado que os indivíduos do GS e do GA apresentavam níveis semelhantes de expressão de HERV-W. Esse dado também corrobora um estudo anterior que visou compreender os níveis de expressão de HERV-W em diferentes tipos de EM. Perron *et al.* também não encontraram diferença nos níveis de expressão entre indivíduos com diferentes tipos de EM ¹⁴¹. Portanto, esse dado sugere que o nível transcricional de HERV-W não se diferencie de acordo com o tipo de EM.

Como já levantado anteriormente, sabe-se que o Natalizumabe pode diminuir a expressão de HERV-W em pacientes com EM tratados por longos períodos ¹⁵². Todos os pacientes que estavam sob o tratamento com esse medicamento apresentaram níveis elevados de expressão de HERV-W, possivelmente pelo curto período de tratamento com o medicamento. Interessantemente, o único paciente que tratava com Natalizumabe por mais de 4 anos também apresentou níveis elevados de expressão de HERV-W (~8x mais que a média dos indivíduos do GC). Aparentemente o medicamento não foi capaz de interferir negativamente na expressão do provírus nos pacientes que estavam sob essa terapia e em especial o paciente 41. Vale ressaltar que apesar dos níveis de expressão serem satisfatórios decidiu-se não incluir o paciente no estudo do transcriptoma a fim de evitar possíveis vieses nos dados gerados.

O real motivo para os pacientes com EM apresentarem um maior nível de expressão pode ser explicado pela diferença na “constituição genômica” desses pacientes. Os indivíduos com EM apresentam um maior número de cópias do HERV-W em comparação a indivíduos saudáveis ^{141,146}. Nesse sentido, a expressão elevada levaria a um maior número de novas integrações, que por sua vez, foram retrotranscritas e detectadas no pool de RNAm de HERV. Logo, trata-se de um *feedback* positivo.

Como mencionado no tópico “HERVs e doenças em humanos” os HERVs têm sido associados a uma série de doenças autoimunes ^{131–134,140,141}. No entanto, o HERV-W especificamente tem sido pouco pesquisado em outras doenças além da EM. Não há trabalhos que pesquisaram a relação do HERV-W com Lúpus e síndrome de Sjögren, por exemplo. Com relação a Artrite Reumatóide, apenas um trabalho encontrou partículas virais e RNAm de HERV-W em amostras de sangue e líquido sinovial em pacientes afetados pela doença, entretanto, sem associação direta ¹⁷⁵.

Essa ausência de dados do HERV-W com outras doenças autoimunes pode aparentar que a expressão desse retrovírus não seja tão ubíqua nessas doenças, mas sim apresenta um caráter específico para a EM.

6.3 Transcriptoma dos pacientes com EM com foco no HERV-W

Os primeiros dados analisados foram originados através de sequenciamento de *amplicon* do gene env do HERV-W, onde o nosso objetivo foi começar a compreender qual era a diversidade dos HERVs nos indivíduos com EM. Obviamente, os dados

observados nessa etapa do estudo foram mais restritos em comparação aos gerados pelo RNASeq. No entanto, pôde-se observar alguns resultados interessantes ao comparar os resultados de ambos os sequenciamentos, como por exemplo uma maior diversidade de *loci* ativos de HERV-W, bem como um perfil transcricional distinto entre os indivíduos dos grupos de EM. Com base nisso, foi possível avaliar as primeiras características do perfil transcricional do HERV-W na EM.

Durante a análise dos resultados de RNASeq observou-se um importante fator que tornou necessário readequar os parâmetros da mesma. O ACP revelou que apesar dos grupos de cada corrida estarem distantes, GA contra o GC, por exemplo, as duas corridas estavam distantes umas das outras (ao se comparar indivíduos comuns a ambas, como o Grupo Controle). Esse fator impossibilitou a comparação direta dos dois grupos de EM, bem como inviabilizou a utilização de todos os indivíduos do GC como um único grupo. Muitas podem ser as explicações para a ocorrência dessa discrepância no perfil transcricional desses indivíduos. No entanto, o principal fator está no preparo da biblioteca. Como nossas amostras eram oriundas de células do PBMC, a possibilidade de haver uma quantidade grande de sequências oriundas de globinas era muito grande, então além da depleção do RNAr, optou-se por depletar as sequências de globinas também. Logo, utilizou-se na primeira corrida o *TruSeq Stranded Total RNA Ribo-Zero Globin* (Illumina) e na segunda corrida, o *GLOBINclear-Human kit* (Ambion). O primeiro kit é baseado na hibridização de oligos com sequências de RNAr e sequências de globinas. Já o segundo kit usa metodologia baseada em microesferas com oligos complementares às sequências de globina. Essas microesferas, após hibridização com o RNA alvo são descartadas e o RNA restante é, dessa vez, purificado com microesferas com oligodT, para a captura do RNAm. Em conjunto, é possível que as diferentes metodologias adotadas nas duas corridas possam ter interferido no perfil transcricional observado dos indivíduos de cada corrida.

Variações indesejadas são eventos comuns em sequenciamentos de larga escala, sendo que protocolos específicos para a normalização dos dados foram desenvolvidos para contornar esse tipo de problema¹⁷⁶. Tais abordagens também foram adotadas no nosso ensaio, removendo sequências indesejadas e repetindo os ACPS. Ainda, existem algoritmos que foram desenvolvidos para tentar integrar os dados gerados por bibliotecas de RNA geradas de formas distintas (cauda poli-T ou depleção de RNAr)¹⁷⁷. As amostras da primeira corrida, mesmo tendo sido submetidas à remoção de sequências não RNAm, aparentemente tinham muitas leituras ainda de RNAr, o que

potencialmente pode ter interferido no perfil transcricional e desviado o eixo do ACP desse grupo de indivíduos. Infelizmente, as metodologias utilizadas parecem ter impedido o agrupamento dos indivíduos do GC das duas corridas, e sendo assim, decidimos analisar as corridas de forma independente.

Apesar dos fatores complicadores observados durante essa etapa do estudo, foi possível identificar que dos 22 dos *loci* mais ativos de HERV-W do GA observados na primeira análise, 15 também apareceram na segunda análise. Para o grupo GS, verificou-se que 10/11 dos *loci* mais ativos na primeira análise também estavam presentes na segunda análise. Ainda, alguns desses *loci* (6q27, 7q21.2 e Xq22.3) também se apresentaram com número significativo de leituras no sequenciamento de *amplicon*. Tomando em conjunto, podemos sugerir que a diferença encontrada no ACP não represente, necessariamente, o que acontece com os *loci* de HERV-W, já que aparentemente esses retroelementos são extremamente ativos em ambas as fases da doença.

O perfil de expressão de HERV-W em cada uma das análises revelou determinados *loci* ativos em ambos os tipos da doença, como 1p22.2, 1p12, 2p16, 4q21.22, 6p22.3, 6q27, 7q21.2(Sincitina), 11q14.2, 15q21.3 e Xq22.3. Por outro lado, foram identificados 8 *loci* distintos para o GS e outros 8 *loci* para o GA. Esse dado sugere duas possibilidades interessantes: a) existem alguns *loci* que são expressos em indivíduos com EM independentemente do tipo de EM em que o paciente se encontra; e b) outros *loci* talvez tragam uma assinatura do perfil transcricional do HERV-W dependente do tipo de EM. Sem dúvida, com esse perfil de transcritos em mãos é possível dedicar esforços em estudos experimentais dirigidos a esses *loci* afim de averiguar o possível papel destes na patogênese da doença.

Vale ressaltar que poucos foram os trabalhos que tiveram como foco determinar quais eram os *loci* mais ativos nos indivíduos com EM. Ainda, nesses trabalhos foram adotadas abordagens metodológicas mais restritas do que a utilizada nesse estudo. O trabalho mais recente investigou a presença de RNA viral em amostras de cérebro de indivíduos com EM através de RNASeq. Nesse trabalho, os autores encontraram além de sequências de outros vírus, transcritos de HERV-K, HERV-K10 e HERV-E¹⁷⁸. Curiosamente, os autores não encontraram sequências de HERV-W.

No entanto, foram dois os trabalhos que buscaram as mesmas respostas que o que foi proposto nesse estudo. O primeiro, realizado por Lauffer *et al.*, utilizou clonagem e sequenciamento para identificar os *loci* mais ativos em amostras de PBMC

de indivíduos com EM. Os autores não encontraram diferença significativa na frequência da maioria dos clones na comparação dos grupos de estudo (EM e Controle), com exceção de clones que representavam dois *loci*, 15q21.3 e 6q21. Ainda, outros *loci* também foram encontrados como o Xq22.3 e 7q21.2 e o 17q12¹⁵⁸. Tais *loci* também foram encontrados no presente estudo como mais frequentemente expressos. No entanto, todos esses foram encontrados com maior número de *reads* nos indivíduos com EM do que no GC. Vale mencionar que dentre estes, o *loci* 17q12 demonstrou níveis semelhantes de expressão entre o GA e o GC. No presente estudo foram avaliados mais pacientes e com metodologia mais adequada para responder esse tipo de pergunta, e com isso foi possível identificar esses e novos *loci* diferencialmente expressos entre os indivíduos com EM.

O segundo estudo foi realizado por Schmitt *et al.* Nesse trabalho os autores utilizaram a metodologia de sequenciamento de próxima geração baseada em *amplicon* com um conjunto de *primers* complementares ao gene env e LTRs virais em amostras cerebrais de indivíduos com EM. Nesse estudo, os autores encontraram uma diversidade de expressão extremamente similar entre os indivíduos com EM e indivíduos controles com predominância de expressão de 6 *loci*: 6q21, 7q21.2, 14q21.3, 15q21.3 e Xq22.3¹⁵⁹. Todos estes foram encontrados também no nosso estudo, e estavam entre os que tinham maior número de *reads*. Porém, estavam mais expressos nos indivíduos com EM em comparação aos controles, com exceção do *loci* 15q21.3, que apresentou níveis similares de expressão entre o GA e o GC. Ainda, demais *loci* dados como baixa atividade no estudo de Schmitt, aqui também são mais expressos. Como mencionado, nesse estudo foram utilizadas amostras de sangue, enquanto no estudo de Schmitt *et al.*, foram utilizadas amostras de cérebro. Foi observado ainda uma diversidade de expressão mais heterogênea nas amostras deste estudo, contrastando com as amostras de tecido cerebral. Ademais, os 6 *loci* mais ativos apresentavam níveis semelhantes tanto para os indivíduos com EM quanto para indivíduos controles, também contrastando com os resultados descritos aqui. Isso sugere uma expressão heterogênea de determinados *loci* em diferentes tecidos/fluidos corporais. De fato a expressão de HERVs tecido-dependente existe^{108,110–113}. Possivelmente, no caso do HERV-W há um perfil e nível distinto de expressão de HERV-W dependendo do tecido estudado, onde a diversidade de expressão pode ser maior em células do PBMC e menos diversa e mais homogênea em tecido cerebral. Essa possibilidade poderia representar um novo rumo a investigação da patogênese da doença. Considerando a diversidade de HERV-W

presente no PBMC, e eventos de mimetismo molecular entre proteínas do HERV-W e proteínas da MOG¹²⁹, a ativação de células imunológicas poderia ser realizada fora do SNC, e posteriormente migrar para o SNC. Essa ativação de células B e T, de fato, pode ocorrer na via sanguínea, e essas células, agora autoreativas, atravessam a barreira hemato-encefálica causando dano autoimune^{41,42}.

Acreditamos que os dados gerados e descritos aqui representam nada mais do que o início de outras investigações, servindo de base para futuros estudos de interações gênicas e proteicas, a fim de compreender totalmente o papel desses retrovírus na patogênese da doença.

6.4 *Pathways* gênicos e quantificação de citocinas séricas

Embora o foco desse estudo tenha sido investigar os HERVs na população estudada, estavam disponíveis dados de expressão de todos os outros genes desses indivíduos. Sendo assim, realizou-se a análise de expressão diferencial gênica dos indivíduos com EM. Investigou-se inicialmente quais eram os genes diferencialmente expressos em cada uma das corridas, foram identificados mais de 350 genes diferencialmente expressos em cada uma das corridas analisadas. Esses dados serviram de *input* para averiguar as possíveis interações gênicas, bem como possíveis vias alteradas em cada um dos grupos. A *core analysis*, realizada no IPA, revelou cerca de 12 vias alteradas para o grupo GS, e 15 vias alteradas para o GA.

Basicamente, observou-se que os indivíduos do GS apresentavam diversas vias inflamatórias com valor de Z positivo, ou seja, ativadas. Além disso, outras vias de sinalização de integrinas e diapedese em processos inflamatórios também estavam ativadas. Isso parece indicar que os pacientes do GS apresentam vias ativadas de produção de citocinas inflamatórias, permeabilidade vascular e fatores de reparo neuronal. De fato, alterações em algumas dessas vias parecem atuar na EM, como IFN γ , IL-12, IL-2, IL-6^{57,179-182}, além disso, algumas das integrinas são alvo de tratamento da doença, já que podem impedir a transmigração de linfócitos T através da BHE¹⁸³.

Ainda, temos vias que promovem a migração de linfócitos a determinados sítios. Dentre as vias alteradas, destaca-se a via da Paxilina, que também mostrou-se ativada no GS. Essa molécula é capaz de interagir com as moléculas de integrinas, promovendo uma resposta inflamatória a diferentes regiões do organismo^{184,185}. Ainda, a ação dessa molécula tem aparentemente papel fundamental na inflamação, onde em estudo usando

modelo animal, a inibição da produção dessa proteína foi capaz de reduzir drasticamente a migração de linfócitos para regiões de inflamação¹⁸⁶. Interessantemente, um estudo descreveu que PBMCs estimulados com a proteína env do HERV-W, foram capazes de produzir uma série de citocinas pró-inflamatórias, como IFN- γ , IL-6 e TNF- α ¹⁸⁷. Considerando todos esses fatores pode-se imaginar um cenário onde os pacientes com EM em surto teriam um perfil inflamatório e maior susceptibilidade a migração de células inflamatórias a sítios específicos. Ainda, com a possibilidade do retroelemento representar um papel fundamental na resposta autoimune por mimetismo molecular, os HERVs portanto poderiam contribuir para a promoção de inflamação no SNC.

Vale ressaltar que investigamos os níveis de algumas citocinas pró inflamatórias no soro dos pacientes com EM e verificamos uma correlação positiva nas vias em que essas citocinas atuam, como por exemplo a via da IL-6 (Figura 27), corroborando os dados encontrados através do transcriptoma.

Com relação ao grupo GA, encontramos um perfil com ausência de inflamação, onde algumas citocinas pro-inflamatórias estavam inativadas e com uma série de vias de reparo e crescimento neuronal inativadas. Dentre elas, destaca-se a Neuregulina, que tem como principal papel atuar na oligodendrogênese promovendo a remielinização¹⁸⁸. Níveis desregulados dessa proteína e mutações no gene estão associados as formas progressivas da doença^{188,189}. Além dessas vias, outras responsáveis pelo crescimento neuronal também estavam inativadas no caso da Neurotrofina e NGF. Portanto, para os indivíduos do GA, aparentemente temos um perfil mais degenerativo, onde as principais vias de crescimento e reparo neuronal estão inativadas.

Em resumo, os dados gerados neste estudo sugerem que de fato existam dois cenários distintos para a doença em seus dois estágios, onde para o GS foi descrito um perfil totalmente inflamatório e com vias atuando diretamente para um recrutamento das células responsáveis pela inflamação. Ainda, é possível que o sistema de reparo celular esteja tentando conter os danos causados pela autoimunidade. Os pacientes do GA, por outro lado, não apresentam resposta inflamatória, e determinadas vias responsáveis pelo reparo e crescimento celular estão inativadas, demonstrando exaustão do sistema em reparar a região lesionada, promovendo portanto, a degeneração neurológica.

7. Conclusão

Todos os indivíduos com EM expressam HERV-W, enquanto 86,1% dos indivíduos do GC expressam os provírus. O nível de expressão de HERV-W em indivíduos com EM é maior do que em indivíduos saudáveis. Entre os indivíduos com EM em diferentes fases da doença o nível de expressão é semelhante;

Os HERVs FC1, L32, 1, H48, K11, W, S71, L, K22, 15, 3, 9N E, L18, IP10FH, K3, H, IP10B3 foram os HERVs mais ativos em indivíduos com EM.

Os *loci* mais ativos de HERV-W foram 1p34.2, 6q21, 6q27, 7q21.2(Sincitina), 11q14.2 e 15q21.3 nos pacientes do GS quando comparados ao grupo controle; por outro lado os *loci* HERV-W 6q27, 7q21.2(Sincitina), 11q14.1, 13q13.3, 15q21.3 e Xq22.3 foram os mais prevalentes no GA. Tanto a diversidade quanto o nível de expressão de HERV-W foram mais elevados nos pacientes com EM;

Os dados gerados pelo RNASeq foram mais informativos dos que os dados gerados pela metodologia de sequenciamento de *amplicon*;

Perfis totalmente distintos de expressão gênica estão relacionados ao tipo de EM em que o paciente se encontra: os indivíduos do GS apresentam fundamentalmente um perfil inflamatório, com confirmação dos níveis séricos de algumas citocinas elevadas significativamente, e com sinalização de moléculas responsáveis por adesão e migração celular para sítios de inflamação; e os indivíduos do GA apresentam todas as vias de sinalização inflamatória inativadas e com vias de reparo do sistema neuronal também inativadas;

As citocinas inflamatórias IL-6, IL-1 β , TNF- α , IFN- γ estão mais elevadas em pacientes do GS em comparação aos indivíduos do GC.

8. Referências

1. Murray TJ. The history of multiple sclerosis: the changing frame of the disease over the centuries. *J Neurol Sci.* 2009/02/10. 2009;277 Suppl:S3-8.
2. Murray TJ. Robert Carswell: the first illustrator of MS. *Int MS J.* 2009/11/03. 2009;16(3):98–101.
3. Charcot J-M (1825-1893). *Histologie de la sclérose en plaques, leçon faite à l'hospice de la Salpêtrière par M. Charcot et recueillie par M. Bourneville.* 1869;
4. Tilbery CP. *Esclerose Múltipla no Brasil: Aspectos Clínicos e Terapeuticos . Vol. 1.* São Paulo: Atheneu; 2004.
5. Compston A, Coles A. Multiple sclerosis. *Lancet.* 2008/10/31. 2008;372(9648):1502–17.
6. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology.* 2014 Jul 15;83(3):278–86.
7. Miller D, Barkhof F, Montalban X, Thompson A, Filippi M. Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. *Lancet Neurol.* 2005 May;4(5):281–8.
8. Confavreux C, Vukusic S. Natural history of multiple sclerosis: a unifying concept. *Brain.* 2006 Jan 6;129(3):606–16.
9. WHO | Atlas: Multiple Sclerosis Resources in the World 2008. WHO. 2013;
10. Browne P, Chandraratna D, Angood C, Tremlett H, Baker C, Taylor B V, et al. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology.* 2014 Sep 9;83(11):1022–4.
11. Correa E, Paredes V, Martínez B. Prevalence of multiple sclerosis in Latin America and its relationship with European migration. *Mult Scler J – Exp Transl Clin.* 2016 Jan 2;2:205521731666640.
12. Callegaro D, Goldbaum M, Morais L, Tilbery CP, Moreira MA, Gabbai AA, et al. The prevalence of multiple sclerosis in the city of Sao Paulo, Brazil, 1997. *Acta Neurol Scand.* 2001/10/09. 2001;104(4):208–13.
13. Fragoso YD, Peres M. Prevalence of multiple sclerosis in the city of Santos, SP. *Rev Bras Epidemiol.* 2007 Dec;10(4):479–82.
14. Negreiros AALV de, Sousa-Munõz RL de, Oliveira BES de, Nóbrega PV da, Monteiro LLD. Clinical and epidemiological profile of patients diagnosed with

- multiple sclerosis in João Pessoa, Paraíba, Brazil. *Arq Neuropsiquiatr*. 2015 Sep;73(9):741–5.
15. Lana-Peixoto MA, Frota ERC, Campos GB, Monteiro LP. The prevalence of multiple sclerosis in Belo Horizonte, Brazil. *Arq Neuropsiquiatr*. 2012 Feb;70(2):102–7.
16. Ribeiro SBF, Maia DF, Ribeiro JB, Cardoso FAG, Silva C. Clinical and epidemiological profile of patients with multiple sclerosis in Uberaba, Minas Gerais, Brazil. *Arq Neuropsiquiatr*. 2011 Apr;69(2A):184–7.
17. Finkelsztejn A, Lopes JS, Noal J, Finkelsztejn JM. The prevalence of multiple sclerosis in Santa Maria, Rio Grande do Sul, Brazil. *Arq Neuropsiquiatr*. 2014 Feb;72(2):104–6.
18. Calmon AB, Pereira F, Marin B, Preux P-M, Alvarenga RP. Prevalence of Multiple Sclerosis in the City of Volta Redonda - Rio De Janeiro, Brazil Using the Capture-Recapture Method. *Neuroepidemiology*. 2016 Jan 6;46(2):88–95.
19. Sumelahti ML, Tienari PJ, Wikstrom J, Salminen TM, Hakama M. Survival of multiple sclerosis in Finland between 1964 and 1993. *Mult Scler*. 2002/08/09. 2002;8(4):350–5.
20. Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple sclerosis. *N Engl J Med*. 2000/09/28. 2000;343(13):938–52.
21. Orton S-M, Ramagopalan S V, Brocklebank D, Herrera BM, Dymment DA, Yee IM, et al. Effect of immigration on multiple sclerosis sex ratio in Canada: the Canadian Collaborative Study. *J Neurol Neurosurg Psychiatry*. 2010 Jan 1;81(1):31–6.
22. Wallin MT, Culpepper WJ, Coffman P, Pulaski S, Maloni H, Mahan CM, et al. The Gulf War era multiple sclerosis cohort: age and incidence rates by race, sex and service. *Brain*. 2012 Jun;135(6):1778–85.
23. Nali LH da S. Investigação da reativação dos poliomavírus humanos JC e BK em pacientes com Esclerose Múltipla (EM) sob tratamento com Natalizumab e pacientes com EM sob outros tratamentos. [São Paulo]: Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo; 2013.
24. Koch-Henriksen N, Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet Neurol*. 2010 May;9(5):520–32.
25. Kragt J, van Amerongen B, Killestein J, Dijkstra C, Uitdehaag B, Polman C, et al. Higher levels of 25-hydroxyvitamin D are associated with a lower incidence

- of multiple sclerosis only in women. *Mult Scler J*. 2009 Jan;15(1):9–15.
26. Selmi C. The X in sex: how autoimmune diseases revolve around sex chromosomes. *Best Pract Res Clin Rheumatol*. 2008 Oct;22(5):913–22.
27. Garg N, Smith TW. An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. *Brain Behav*. 2015 Sep;5(9):n/a-n/a.
28. Hemmer B, Selter B. Update on immunopathogenesis and immunotherapy in multiple sclerosis. *ImmunoTargets Ther*. 2013 Apr;2:21.
29. Mentis A-FA, Dardiotis E, Grigoriadis N, Petinaki E, Hadjigeorgiou GM. Viruses and Multiple Sclerosis: From Mechanisms and Pathways to Translational Research Opportunities. *Mol Neurobiol*. 2017 Jul 28;54(5):3911–23.
30. Grigoriadis N, Hadjigeorgiou G. Virus-mediated autoimmunity in Multiple Sclerosis. *J Autoimmune Dis*. 2006 Feb 19;3(1):1.
31. International Multiple Sclerosis Genetics Consortium, Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med*. 2007 Aug 30;357(9):851–62.
32. Hollenbach JA, Oksenberg JR. The immunogenetics of multiple sclerosis: A comprehensive review. *J Autoimmun*. 2015 Nov;64:13–25.
33. Bergamaschi R, Cortese A, Pichiecchio A, Berzolari FG, Borrelli P, Mallucci G, et al. Air pollution is associated to the multiple sclerosis inflammatory activity as measured by brain MRI. *Mult Scler J*. 2017 Aug 14;135245851772686.
34. Simpson S, der Mei I van, Taylor B. The role of vitamin D in multiple sclerosis: biology & biochemistry, epidemiology and potential roles in treatment. *Med Chem*. 2017 Sep 21;
35. Milo R, Kahana E. Multiple sclerosis: geoepidemiology, genetics and the environment. *Autoimmun Rev*. 2009/11/26. 2010;9(5):A387-94.
36. Ascherio A, Munger KL. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. *Ann Neurol*. 2007 Jun;61(6):504–13.
37. Perron H, Lazarini F, Ruprecht K, Péchoux-Longin C, Seilhean D, Sazdovitch V, et al. Human endogenous retrovirus (HERV)-W ENV and GAG proteins: physiological expression in human brain and pathophysiological modulation in multiple sclerosis lesions. *J Neurovirol*. 2005 Feb;11(1):23–33.
38. Mancuso R, Hernis A, Cavarretta R, Caputo D, Calabrese E, Nemni R, et al. Detection of viral DNA sequences in the cerebrospinal fluid of patients with multiple sclerosis. *J Med Virol*. 2010 May;82(6):1051–7.

39. Geeraedts F, Wilczak N, van Binnendijk R, De Keyser J. Search for morbillivirus proteins in multiple sclerosis brain tissue. *Neuroreport*. 2004/04/27. 2004;15(1):27–32.
40. Ben Fredj N, Rotola A, Nefzi F, Chebel S, Rizzo R, Caselli E, et al. Identification of human herpesviruses 1 to 8 in Tunisian multiple sclerosis patients and healthy blood donors. *J Neurovirol*. 2012 Feb;18(1):12–9.
41. Dendrou CA, Fugger L. Immunomodulation in multiple sclerosis: promises and pitfalls. *Curr Opin Immunol*. 2017 Dec 16;49:37–43.
42. Lehmann-Horn K, Kinzel S, Weber M. Deciphering the Role of B Cells in Multiple Sclerosis—Towards Specific Targeting of Pathogenic Function. *Int J Mol Sci*. 2017 Sep 23;18(10):2048.
43. Popescu BFG, Pirko I, Lucchinetti CF. Pathology of Multiple Sclerosis. *Contin Lifelong Learn Neurol*. 2013 Aug;19(4 Multiple Sclerosis):901–21.
44. Reiber H, Ungefehr S, Jacobi C. The intrathecal, polyspecific and oligoclonal immune response in multiple sclerosis. *Mult Scler J*. 1998 Jun 2;4(3):111–7.
45. Huss A, Buttmann M, Brecht I, Weishaupt A, Otto M, Tumani H. Validation of a multiplexing technique to determine the intrathecal, polyspecific antiviral immune response in multiple sclerosis. *Expert Rev Mol Diagn*. 2016 Dec;16(12):1353–6.
46. Brennan KM, Galban-Horcajo F, Rinaldi S, O’Leary CP, Goodyear CS, Kalna G, et al. Lipid arrays identify myelin-derived lipids and lipid complexes as prominent targets for oligoclonal band antibodies in multiple sclerosis. *J Neuroimmunol*. 2011 Sep 15;238(1–2):87–95.
47. Harp CT, Ireland S, Davis LS, Remington G, Cassidy B, Cravens PD, et al. Memory B cells from a subset of treatment-naïve relapsing-remitting multiple sclerosis patients elicit CD4⁺ T-cell proliferation and IFN- γ production in response to myelin basic protein and myelin oligodendrocyte glycoprotein. *Eur J Immunol*. 2010 Oct;40(10):2942–56.
48. Coleman M. Axon degeneration mechanisms: commonality amid diversity. *Nat Rev Neurosci*. 2005 Nov 14;6(11):889–98.
49. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. *Ann Neurol*. 1983 Mar;13(3):227–31.
50. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al.

- Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol.* 2001 Jul;50(1):121–7.
51. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann Neurol.* 2011 Feb;69(2):292–302.
52. Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol.* 2016 Mar;15(3):292–303.
53. (UK) NCGC. Multiple Sclerosis. Multiple Sclerosis: Management of Multiple Sclerosis in Primary and Secondary Care. National Institute for Health and Care Excellence (UK); 2014.
54. Galea I, Ward-Abel N, Heesen C. Relapse in multiple sclerosis. *BMJ.* 2015 Apr 14;350(apr14 8):h1765–h1765.
55. Citterio A, La Mantia L, Ciucci G, Candelise L, Brusaferri F, Midgard R, et al. Corticosteroids or ACTH for acute exacerbations in multiple sclerosis. In: Filippini G, editor. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd; 2000. p. CD001331.
56. Okuda DT. Immunosuppressive treatments in multiple sclerosis. In: *Handbook of clinical neurology*. 2014. p. 503–11.
57. Balashov KE, Smith DR, Khoury SJ, Hafler DA, Weiner HL. Increased interleukin 12 production in progressive multiple sclerosis: induction by activated CD4+ T cells via CD40 ligand. *Proc Natl Acad Sci U S A.* 1997 Jan 21;94(2):599–603.
58. Weiner HL. Immunosuppressive treatment in multiple sclerosis. *J Neurol Sci.* 2004 Aug 15;223(1):1–11.
59. Pardo G, Jones DE. The sequence of disease-modifying therapies in relapsing multiple sclerosis: safety and immunologic considerations. *J Neurol.* 2017 Sep 6;
60. Damal K, Stoker E, Foley JF. Optimizing therapeutics in the management of patients with multiple sclerosis: a review of drug efficacy, dosing, and mechanisms of action. *Biologics.* 2013 Nov;7:247–58.
61. Du Pasquier RA, Pinschewer DD, Merkler D. Immunological Mechanism of Action and Clinical Profile of Disease-Modifying Treatments in Multiple Sclerosis. *CNS Drugs.* 2014 Jun 11;28(6):535–58.

62. Buzzard K, Broadley S, Butzkueven H. What Do Effective Treatments for Multiple Sclerosis Tell Us about the Molecular Mechanisms Involved in Pathogenesis? *Int J Mol Sci*. 2012 Oct 4;13(12):12665–709.
63. Cohen JA, Barkhof F, Comi G, Hartung H-P, Khatri BO, Montalban X, et al. Oral Fingolimod or Intramuscular Interferon for Relapsing Multiple Sclerosis. *N Engl J Med*. 2010 Feb 4;362(5):402–15.
64. FUJITA T, INOUE K, YAMAMOTO S, IKUMOTO T, SASAKI S, TOYAMA R, et al. Fungal metabolites. Part 11. A potent immunosuppressive activity found in *Isaria sinclairii* metabolite. *J Antibiot (Tokyo)*. 1994 Feb 25;47(2):208–15.
65. Chun J, Hartung H-P. Mechanism of Action of Oral Fingolimod (FTY720) in Multiple Sclerosis. *Clin Neuropharmacol*. 2010 Mar;33(2):91–101.
66. Pinschewer DD, Brinkmann V, Merkler D. Impact of sphingosine 1-phosphate modulation on immune outcomes. *Neurology*. 2011 Feb 22;76(Issue 8, Supplement 3):S15–9.
67. Mehling M, Brinkmann V, Antel J, Bar-Or A, Goebels N, Vedrine C, et al. FTY720 therapy exerts differential effects on T cell subsets in multiple sclerosis. *Neurology*. 2008 Oct 14;71(16):1261–7.
68. Yednock TA, Cannon C, Fritz LC, Sanchez-Madrid F, Steinman L, Karin N. Prevention of experimental autoimmune encephalomyelitis by antibodies against alpha 4 beta 1 integrin. *Nature*. 1992/03/05. 1992;356(6364):63–6.
69. Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, et al. A Randomized, Placebo-Controlled Trial of Natalizumab for Relapsing Multiple Sclerosis. *N Engl J Med*. 2006 Mar 2;354(9):899–910.
70. CONITEC. Natalizumabe 300mg (Tysabri®) para Esclerose Múltipla Remitente Recorrente em segunda linha de tratamento. Ministério da Saúde. 2013;
71. Hauser SL, Bar-Or A, Comi G, Giovannoni G, Hartung H-P, Hemmer B, et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. *N Engl J Med*. 2017 Jan 19;376(3):221–34.
72. Hohlfeld R, Meinl E. Ocrelizumab in multiple sclerosis: markers and mechanisms. *Lancet Neurol*. 2017 Apr;16(4):259–61.
73. Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature*. 2010 Apr 29;464(7293):1351–6.
74. Moreira MA, Tilbery CP, Lana-Peixoto MA, Mendes MF, Kaimen-Maciél DR,

- Callegaro D. [Historical aspects of multiple sclerosis]. *Rev Neurol*. 2002/05/22. 2002;34(4):379–83.
75. Virtanen JO, Jacobson S. Viruses and multiple sclerosis. *CNS Neurol Disord Drug Targets*. 2012 Aug;11(5):528–44.
76. Kurtzke JF, Heltberg A. Multiple sclerosis in the Faroe Islands: an epitome. *J Clin Epidemiol*. 2001/02/13. 2001;54(1):1–22.
77. Brown C. Aetiology: Neighbourhood watch. *Nature*. 2016 Nov 30;540(7631):S4–6.
78. Torabi M, Green C, Yu N, Marrie RA. Application of three focused cluster detection methods to study geographic variation in the incidence of multiple sclerosis in Manitoba, Canada. *Neuroepidemiology*. 2014;43(1):38–48.
79. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet*. 1971/06/19. 1971;1(7712):1257–60.
80. Isaacson SH, Asher DM, Godec MS, Gibbs CJ, Gajdusek DC. Widespread, restricted low-level measles virus infection of brain in a case of subacute sclerosing panencephalitis. *Acta Neuropathol*. 1996;91(2):135–9.
81. Dobson R, Ramagopalan S, Davis A, Giovannoni G. Cerebrospinal fluid oligoclonal bands in multiple sclerosis and clinically isolated syndromes: a meta-analysis of prevalence, prognosis and effect of latitude. *J Neurol Neurosurg Psychiatry*. 2013 Aug 1;84(8):909–14.
82. Dias-Carneiro RPC, von Glehn F, Moraes AS, Boldrini VO, Damasceno A, Andrade MD, et al. MRZH reaction increases sensitivity for intrathecal IgG synthesis in IgG Oligoclonal band negative Multiple Sclerosis patients. *J Neuroimmunol*. 2016 Nov 15;300:30–5.
83. Stüve O, Cravens PD, Frohman EM, Phillips JT, Remington GM, von Geldern G, et al. Immunologic, clinical, and radiologic status 14 months after cessation of natalizumab therapy. *Neurology*. 2009 Feb 3;72(5):396–401.
84. Pfuhl C, Oechtering J, Rasche L, Gieß RM, Behrens JR, Wakonig K, et al. Association of serum Epstein–Barr nuclear antigen-1 antibodies and intrathecal immunoglobulin synthesis in early multiple sclerosis. *J Neuroimmunol*. 2015 Aug 15;285:156–60.
85. Vogt P. Historical Introduction to the General Properties of Retroviruses. *Retroviruses*. 1997.

86. Griffiths DJ. Endogenous retroviruses in the human genome sequence. *Genome Biol.* 2001 Jan;2(6):REVIEWS1017.
87. Löwer R, Löwer J, Kurth R. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci U S A.* 1996 May 28;93(11):5177–84.
88. Patience C, Wilkinson DA, Weiss RA. Our retroviral heritage. *Trends Genet.* 1997 Mar;13(3):116–20.
89. Shih A, Coutavas EE, Rush MG. Evolutionary implications of primate endogenous retroviruses. *Virology.* 1991 Jun;182(2):495–502.
90. Kim H-S, Kim D-S, Huh J-W, Ahn K, Yi J-M, Lee J-R, et al. Molecular characterization of the HERV-W env gene in humans and primates: expression, FISH, phylogeny, and evolution. *Mol Cells.* 2008 Jul 31;26(1):53–60.
91. Kim H-S, Lee W-H. Human Endogenous Retrovirus HERV-W Family: Chromosomal Localization, Identification, and Phylogeny. *AIDS Res Hum Retroviruses.* 2001 May 1;17(7):643–8.
92. Mayer J, Meese EU. The human endogenous retrovirus family HERV-K(HML-3). *Genomics.* 2002 Sep;80(3):331–43.
93. Voisset C, Blancher A, Perron H, Mandrand B, Mallet F, Paranhos-Baccala G. Phylogeny of a Novel Family of Human Endogenous Retrovirus Sequences, HERV-W, in Humans and Other Primates. *AIDS Res Hum Retroviruses.* 1999 Nov 20;15(17):1529–33.
94. Kim HS, Takenaka O, Crow TJ. Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates. *J Gen Virol.* 1999 Oct 1;80 (Pt 10)(10):2613–9.
95. Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, et al. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Mol Phylogenet Evol.* 1998 Jun;9(3):585–98.
96. Weiss RA. The discovery of endogenous retroviruses. *Retrovirology.* 2006 Jan;3:67.
97. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. *Science (80-).* 2001 Feb 16;291(5507):1304–51.
98. Goff S. *Fields Virology.* 1st ed. New York: Lippincott Williams & Wilkins; 2005. 165 p.

99. Li F, Nellåker C, Yolken RH, Karlsson H. A systematic evaluation of expression of HERV-W elements; influence of genomic context, viral structure and orientation. *BMC Genomics*. 2011 Jan;12:22.
100. Kovalskaya E, Buzdin A, Gogvadze E, Vinogradova T, Sverdlov E. Functional human endogenous retroviral LTR transcription start sites are located between the R and U5 regions. *Virology*. 2006 Mar 15;346(2):373–8.
101. Tristem M. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol*. 2000 Apr;74(8):3715–30.
102. Weiss RA. The discovery of endogenous retroviruses. *Retrovirology*. 2006 Oct 3;3(1):67.
103. Mager DL, Goodchild NL. Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. *Am J Hum Genet*. 1989 Dec;45(6):848–54.
104. Perron H, Perin JP, Rieger F, Alliel PM. 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*. 2000 May;6 Suppl 2:S67–75.
105. Contreras-Galindo R, Kaplan MH, Markovitz DM, Lorenzo E, Yamamura Y. Detection of HERV-K(HML-2) viral RNA in plasma of HIV type 1-infected individuals. *AIDS Res Hum Retroviruses*. 2006 Oct;22(10):979–84.
106. Ruprecht K, Ferreira H, Flockerzi A, Wahl S, Sauter M, Mayer J, et al. Human Endogenous Retrovirus Family HERV-K(HML-2) RNA Transcripts Are Selectively Packaged into Retroviral Particles Produced by the Human Germ Cell Tumor Line Tera-1 and Originate Mainly from a Provirus on Chromosome 22q11.21. *J Virol*. 2008 Oct 15;82(20):10008–16.
107. Belshaw R, Pereira V, Katzourakis A, Talbot G, Paces J, Burt A, et al. Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci U S A*. 2004 Apr 6;101(14):4894–9.
108. Medstrand P, Landry JR, Mager DL. Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. *J Biol Chem*. 2001 Jan 19;276(3):1896–903.
109. Ting CN, Rosenberg MP, Snow CM, Samuelson LC, Meisler MH. Endogenous retroviral sequences are required for tissue-specific expression of a human

- salivary amylase gene. *Genes Dev.* 1992 Aug;6(8):1457–65.
110. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature.* 2000 Feb 17;403(6771):785–9.
111. Blond JL, Lavillette D, Cheynet V, Bouton O, Oriol G, Chapel-Fernandes S, et al. 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.* 2000 Apr;74(7):3321–9.
112. Blaise S, de Parseval N, Heidmann T. Functional characterization of two newly identified Human Endogenous Retrovirus coding envelope genes. *Retrovirology.* 2005 Mar 14;2(1):19.
113. Redelsperger F, Raddi N, Bacquin A, Vernochet C, Mariot V, Gache V, et al. Genetic Evidence That Captured Retroviral Envelope syncytins Contribute to Myoblast Fusion and Muscle Sexual Dimorphism in Mice. Cox GA, editor. *PLOS Genet.* 2016 Sep 2;12(9):e1006289.
114. SenGupta D, Tandon R, Vieira RGS, Ndhlovu LC, Lown-Hecht R, Ormsby CE, et al. Strong human endogenous retrovirus-specific T cell responses are associated with control of HIV-1 in chronic infection. *J Virol.* 2011 Jul;85(14):6977–85.
115. Jones RB, Garrison KE, Mujib S, Mihajlovic V, Aidarus N, Hunter D V, et al. HERV-K-specific T cells eliminate diverse HIV-1/2 and SIV primary isolates. *J Clin Invest.* 2012 Dec;122(12):4473–89.
116. Garrison KE, Jones RB, Meiklejohn DA, Anwar N, Ndhlovu LC, Chapman JM, et al. T cell responses to human endogenous retroviruses in HIV-1 infection. *PLoS Pathog.* 2007 Nov;3(11):e165.
117. Cherkasova E, Scrivani C, Doh S, Weisman Q, Takahashi Y, Harashima N, et al. Detection of an Immunogenic HERV-E Envelope with Selective Expression in Clear Cell Kidney Cancer. *Cancer Res.* 2016 Apr 15;76(8):2177–85.
118. Takahashi Y, Harashima N, Kajigaya S, Yokoyama H, Cherkasova E, McCoy JP, et al. Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells. *J Clin Invest.* 2008 Feb 1;118(3):1099–109.
119. Wang-Johanning F, Frost AR, Johanning GL, Khazaeli MB, LoBuglio AF, Shaw DR, et al. Expression of human endogenous retrovirus k envelope transcripts in

- human breast cancer. *Clin Cancer Res.* 2001 Jun;7(6):1553–60.
120. Contreras-Galindo R, Kaplan MH, Leissner P, Verjat T, Ferlenghi I, Bagnoli F, et al. Human Endogenous Retrovirus K (HML-2) Elements in the Plasma of People with Lymphoma and Breast Cancer. *J Virol.* 2008 Oct 1;82(19):9329–36.
121. Wang-Johanning F, Frost AR, Jian B, Azerou R, Lu DW, Chen D-T, et al. Detecting the expression of human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma. *Cancer.* 2003 Jul 1;98(1):187–97.
122. Kim H-S, Ahn K, Kim D-S. Quantitative expression of the HERV-W env gene in human tissues. *Arch Virol.* 2008 Aug 5;153(8):1587–91.
123. Katoh I, Mírová A, Kurata S, Murakami Y, Horikawa K, Nakakuki N, et al. Activation of the long terminal repeat of human endogenous retrovirus K by melanoma-specific transcription factor MITF-M. *Neoplasia.* 2011 Nov;13(11):1081–92.
124. Liang Q, Xu Z, Xu R, Wu L, Zheng S. Expression Patterns of Non-Coding Spliced Transcripts from Human Endogenous Retrovirus HERV-H Elements in Colon Cancer. Roemer K, editor. *PLoS One.* 2012 Jan 6;7(1):e29950.
125. Wentzensen N, Coy JF, Knaebel H-P, Linnebacher M, Wilz B, Gebert J, et al. Expression of an endogenous retroviral sequence from the HERV-H group in gastrointestinal cancers. *Int J Cancer.* 2007 Oct 1;121(7):1417–23.
126. Golan M, Hizi A, Resau JH, Yaal-Hahoshen N, Reichman H, Keydar I, et al. Human endogenous retrovirus (HERV-K) reverse transcriptase as a breast cancer prognostic marker. *Neoplasia.* 2008 Jun;10(6):521–33.
127. Zhao J, Rycaj K, Geng S, Li M, Plummer JB, Yin B, et al. Expression of Human Endogenous Retrovirus Type K Envelope Protein is a Novel Candidate Prognostic Marker for Human Breast Cancer. *Genes Cancer.* 2011 Sep 1;2(9):914–22.
128. Mentis A-FA, Dardiotis E, Grigoriadis N, Petinaki E, Hadjigeorgiou GM. Viruses and endogenous retroviruses in multiple sclerosis: From correlation to causation. *Acta Neurol Scand.* 2017 May 23;
129. do Olival GS, Faria TS, Nali LHS, de Oliveira ACP, Casseb J, Vidal JE, et al. Genomic analysis of ERVWE2 locus in patients with multiple sclerosis: absence of genetic association but potential role of human endogenous retrovirus type W elements in molecular mimicry with myelin antigen. *Front Microbiol.* 2013 Jan;4:172.

130. Baudino L, Yoshinobu K, Morito N, Santiago-Raber M-L, Izui S. Role of endogenous retroviruses in murine SLE. *Autoimmun Rev.* 2010 Nov;10(1):27–34.
131. Brodziak A, Ziółko E, Muc-Wierzgoń M, Nowakowska-Zajdel E, Kokot T, Klakla K. The role of human endogenous retroviruses in the pathogenesis of autoimmune diseases. *Med Sci Monit.* 2012 Jun;18(6):RA80-8.
132. Piotrowski PC, Duriagin S, Jagodzinski PP. Expression of human endogenous retrovirus clone 4-1 may correlate with blood plasma concentration of anti-U1 RNP and anti-Sm nuclear antibodies. *Clin Rheumatol.* 2005 Dec 13;24(6):620–4.
133. Freimanis G, Hooley P, Ejtehadi HD, Ali HA, Veitch A, Rylance PB, et al. A role for human endogenous retrovirus-K (HML-2) in rheumatoid arthritis: investigating mechanisms of pathogenesis. *Clin Exp Immunol.* 2010 Mar 16;160(3):340–7.
134. Moyes DL, Martin A, Sawcer S, Temperton N, Worthington J, Griffiths DJ, et al. The distribution of the endogenous retroviruses HERV-K113 and HERV-K115 in health and disease. *Genomics.* 2005 Sep;86(3):337–41.
135. Perron H, Geny C, Laurent A, Mouriquand C, Pellat J, Perret J, et al. Leptomenigeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. *Res Virol.* 1989/11/01. 1989;140(6):551–61.
136. Perron H, Lalande B, Gratacap B, Laurent A, Genoulaz O, Geny C, et al. Isolation of retrovirus from patients with multiple sclerosis. *Lancet (London, England).* 1991 Apr 6;337(8745):862–3.
137. Komurian-Pradel F, Paranhos-Baccala G, Bedin F, Ounanian-Paraz A, Sodoyer M, Ott C, et al. Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. *Virology.* 1999 Jul 20;260(1):1–9.
138. Mameli G, Astone V, Arru G, Marconi S, Lovato L, Serra C, et al. 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.* 2007 Jan;88(Pt 1):264–74.
139. Antony JM, Zhu Y, Izad M, Warren KG, Vodjigani M, Mallet F, et al. Comparative expression of human endogenous retrovirus-W genes in multiple sclerosis. *AIDS Res Hum Retroviruses.* 2007 Oct;23(10):1251–6.
140. Rasmussen HB, Geny C, Deforges L, Perron H, Tourtelotte W, Heltberg A, et al.

- Expression of endogenous retroviruses in blood mononuclear cells and brain tissue from multiple sclerosis patients. *Mult Scler*. 1995 Jun;1(2):82–7.
141. Perron H, Germe R, Bernard C, Garcia-Montojo M, Deluen C, Farinelli L, et al. Human endogenous retrovirus type W envelope expression in blood and brain cells provides new insights into multiple sclerosis disease. *Mult Scler*. 2012 Dec;18(12):1721–36.
142. Sotgiu S, Serra C, Mameli G, Pugliatti M, Rosati G, Arru G, et al. Multiple sclerosis-associated retrovirus and MS prognosis: an observational study. *Neurology*. 2002 Oct 8;59(7):1071–3.
143. Mameli G, Poddighe L, Mei A, Uleri E, Sotgiu S, Serra C, et al. Expression and activation by Epstein Barr virus of human endogenous retroviruses-W in blood cells and astrocytes: inference for multiple sclerosis. *PLoS One*. 2012 Jan;7(9):e44991.
144. Mameli G, Poddighe L, Astone V, Delogu G, Arru G, Sotgiu S, et al. Novel reliable real-time PCR for differential detection of MSR_Venv and syncytin-1 in RNA and DNA from patients with multiple sclerosis. *J Virol Methods*. 2009 Oct;161(1):98–106.
145. Perron H, Garson JA, Bedin F, Beseme F, Paranhos-Baccala G, Komurian-Pradel F, et al. 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*. 1997 Jul 8;94(14):7583–8.
146. Garcia-Montojo M, Dominguez-Mozo M, Arias-Leal A, Garcia-Martinez Á, De las Heras V, Casanova I, et al. The DNA copy number of human endogenous retrovirus-W (MSRV-type) is increased in multiple sclerosis patients and is influenced by gender and disease severity. *PLoS One*. 2013 Jan;8(1):e53623.
147. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet*. 2003 Sep 3;35(1):41–8.
148. Perron H, Jouvin-Marche E, Michel M, Ounanian-Paraz A, Camelo S, Dumon A, et al. Multiple sclerosis retrovirus particles and recombinant envelope trigger an abnormal immune response in vitro, by inducing polyclonal V β 16 T-lymphocyte activation. *Virology*. 2001 Sep 1;287(2):321–32.
149. Rolland A, Jouvin-Marche E, Viret C, Faure M, Perron H, Marche PN. The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. *J Immunol*.

- 2006 Jun 15;176(12):7636–44.
150. Antony JM, van Marle G, Opii W, Butterfield DA, Mallet F, Yong VW, et al. Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. *Nat Neurosci.* 2004 Oct 26;7(10):1088–95.
151. Perron H, Dougier-Reynaud H-L, Lomparski C, Popa I, Firouzi R, Bertrand J-B, et al. Human Endogenous Retrovirus Protein Activates Innate Immunity and Promotes Experimental Allergic Encephalomyelitis in Mice. Villoslada P, editor. *PLoS One.* 2013 Dec 6;8(12):e80128.
152. Arru G, Leoni S, Pugliatti M, Mei A, Serra C, Delogu LG, et al. Natalizumab inhibits the expression of human endogenous retroviruses of the W family in multiple sclerosis patients: a longitudinal cohort study. *Mult Scler.* 2014 Feb;20(2):174–82.
153. Arru G, Caggiu E, Leoni S, Mameli G, Pugliatti M, Sechi G Pietro, et al. Natalizumab modulates the humoral response against HERV-Wenv73-88 in a follow-up study of Multiple Sclerosis patients. *J Neurol Sci.* 2015 Jul 8;
154. Curtin F, Lang AB, Perron H, Laumonier M, Vidal V, Porchet HC, et al. GNBAC1, a humanized monoclonal antibody against the envelope protein of multiple sclerosis-associated endogenous retrovirus: a first-in-humans randomized clinical study. *Clin Ther.* 2012 Dec;34(12):2268–78.
155. Curtin F, Perron H, Kromminga A, Porchet H, Lang AB. Preclinical and early clinical development of GNBAC1, a humanized IgG4 monoclonal antibody targeting endogenous retroviral MSRV-Env protein. *MAbs.* 2015 Jan 2;7(1):265–75.
156. Derfuss T, Curtin F, Guebelin C, Bridel C, Rasenack M, Matthey A, et al. 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.* 2015 Jun;21(7):885–93.
157. Zimmermann M, Sanderson NSR, Rasenack M, Lalive PH, Lang AB, Curtin F, et al. Immunologic monitoring during a phase 2a trial of the GNBAC1 antibody in patients with MS. *Neurol - Neuroimmunol Neuroinflammation.* 2015 Oct 20;2(5):e144.
158. Laufer G, Mayer J, Mueller BF, Mueller-Lantzsch N, Ruprecht K. Analysis of

- transcribed human endogenous retrovirus W env loci clarifies the origin of multiple sclerosis-associated retrovirus env sequences. *Retrovirology*. 2009 Jan;6:37.
159. Schmitt K, Richter C, Backes C, Meese E, Ruprecht K, Mayer J. Comprehensive analysis of human endogenous retrovirus group HERV-W locus transcription in multiple sclerosis brain lesions by high-throughput amplicon sequencing. *J Virol*. 2013 Dec;87(24):13837–52.
160. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology*. 1983 Nov 1;33(11):1444–1444.
161. Zipper H, Brunner H, Bernhagen J, Vitzthum F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res*. 2004 Jul 7;32(12):e103–e103.
162. Nellåker C, Yao Y, Jones-Brando L, Mallet F, Yolken RH, Karlsson H. Transactivation of elements in the human endogenous retrovirus W family by viral infection. *Retrovirology*. 2006 Jan;3:44.
163. Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, et al. Mutation in the DNA mismatch repair gene homologue hMLH 1 is associated with hereditary non-polyposis colon cancer. *Nature*. 1994 Mar 17;368(6468):258–61.
164. Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. *J Comput Graph Stat*. 2012 Feb 21;
165. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan 1;29(1):15–21.
166. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*. 2015 Nov 15;31(22):3593–9.
167. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017 Mar 6;14(4):417–9.
168. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
169. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA

- measurements. *BMC Mol Biol.* 2006 Jan;7:3.
170. REMC Standards and Guidelines for RNA-sequencing DRAFT v3.0 I. Introduction. 2007;0–4.
 171. Confavreux C, Vukusic S. The clinical course of multiple sclerosis. In: *Handbook of clinical neurology.* 2014. p. 343–69.
 172. Mameli G, Arru G, Caggiu E, Niegowska M, Leoni S, Madeddu G, et al. Natalizumab Therapy Modulates miR-155, miR-26a and Proinflammatory Cytokine Expression in MS Patients. *PLoS One.* 2016;11(6):e0157153.
 173. Muñoz-Culla M, Irizar H, Castillo-Triviño T, Sáenz-Cuesta M, Sepúlveda L, Lopetegi I, et al. Blood miRNA expression pattern is a possible risk marker for natalizumab-associated progressive multifocal leukoencephalopathy in multiple sclerosis patients. *Mult Scler.* 2014 Dec;20(14):1851–9.
 174. Meira M, Sievers C, Hoffmann F, Haghikia A, Rasenack M, Décard BF, et al. Natalizumab-induced POU2AF1/Spi-B upregulation: A possible route for PML development. *Neurol Neuroimmunol neuroinflammation.* 2016 Jun;3(3):e223.
 175. Gaudin P, Ijaz S, Tuke PW, Marcel F, Paraz A, Seigneurin JM, et al. Infrequency of detection of particle-associated MSRV/HERV-W RNA in the synovial fluid of patients with rheumatoid arthritis. *Rheumatology (Oxford).* 2000 Sep;39(9):950–4.
 176. Peixoto L, Risso D, Poplawski SG, Wimmer ME, Speed TP, Wood MA, et al. How data analysis affects power, reproducibility and biological insight of RNA-seq studies in complex datasets. *Nucleic Acids Res.* 2015 Sep 18;43(16):7664–74.
 177. Bush SJ, McCulloch MEB, Summers KM, Hume DA, Clark EL. Integration of quantitated expression estimates from polyA-selected and rRNA-depleted RNA-seq libraries. *BMC Bioinformatics.* 2017 Jun 13;18(1):301.
 178. Kriesel JD, Bhetariya PJ, Chan BK, Wilson T, Fischer KF. Enrichment of Retroviral Sequences in Brain Tissue from Patients with Severe Demyelinating Diseases. *J Emerg Dis Virol.* 2017 Aug;3(2).
 179. Balashov KE, Olek MJ, Smith DR, Khoury SJ, Weiner HL. Seasonal variation of interferon- γ production in progressive multiple sclerosis. *Ann Neurol.* 1998 Nov;44(5):824–8.
 180. Sun L, He C, Nair L, Yeung J, Egwuagu CE. Interleukin 12 (IL-12) family cytokines: Role in immune pathogenesis and treatment of CNS autoimmune

- disease. *Cytokine*. 2015 Oct;75(2):249–55.
181. Malekzadeh A, Van de Geer-Peeters W, De Groot V, Elisabeth Teunissen C, Beckerman H, Study Group T-A. Fatigue in Patients with Multiple Sclerosis: Is It Related to Pro- and Anti-Inflammatory Cytokines? *Dis Markers*. 2015;2015:1–7.
182. Wylezinski LS, Hawiger J. Interleukin 2 Activates Brain Microvascular Endothelial Cells Resulting in Destabilization of Adherens Junctions. *J Biol Chem*. 2016 Oct 28;291(44):22913–23.
183. Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med*. 2006/03/03. 2006;354(9):899–910.
184. Liu S, Thomas SM, Woodside DG, Rose DM, Kiosses WB, Pfaff M, et al. Binding of paxillin to $\alpha 4$ integrins modifies integrin-dependent biological responses. *Nature*. 1999 Dec 9;402(6762):676–81.
185. Kummer C, Petrich BG, Rose DM, Ginsberg MH. A Small Molecule That Inhibits the Interaction of Paxillin and $\alpha 4$ Integrin Inhibits Accumulation of Mononuclear Leukocytes at a Site of Inflammation. *J Biol Chem*. 2010 Mar 26;285(13):9462–9.
186. Feral CC, Rose DM, Han J, Fox N, Silverman GJ, Kaushansky K, et al. Blocking the $\alpha 4$ integrin paxillin interaction selectively impairs mononuclear leukocyte recruitment to an inflammatory site. *J Clin Invest*. 2006 Mar 1;116(3):715–23.
187. Rolland A, Jouvin-Marche E, Saresella M, Ferrante P, Cavaretta R, Créange A, et al. Correlation between disease severity and in vitro cytokine production mediated by MSR/V (multiple sclerosis associated retroviral element) envelope protein in patients with multiple sclerosis. *J Neuroimmunol*. 2005 Mar;160(1–2):195–203.
188. Kataria H, Alizadeh A, Shahriary GM, Saboktakin Rizi S, Henrie R, Santhosh KT, et al. Neuregulin-1 promotes remyelination and fosters a pro-regenerative inflammatory response in focal demyelinating lesions of the spinal cord. *Glia*. 2018 Mar;66(3):538–61.
189. Bahadori Z, Behmanesh M, Sahraian MA. Two functional promoter polymorphisms of neuregulin 1 gene are associated with progressive forms of multiple sclerosis. *J Neurol Sci*. 2015 Apr 15;351(1–2):154–9.

9. Anexos

Anexo 1. Aprovação do comitê de ética da Faculdade de medicina da Universidade de São Paulo.

<div style="display: flex; justify-content: space-between; align-items: center;"> <div> <p>FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO - FMUSP</p> </div> <div> </div> </div>								
PARECER CONSUBSTANCIADO DO CEP								
<p>DADOS DO PROJETO DE PESQUISA</p> <p>Título da Pesquisa: Perfil de expressão dos Retrovírus Endógenos Humanos da família W (HERV-W) em indivíduos com Esclerose Múltipla em condições de recorrência</p> <p>Pesquisador: Camila Malta Romano</p> <p>Área Temática:</p> <p>Versão: 2</p> <p>CAAE: 25239113.4.0000.0065</p> <p>Instituição Proponente: FUNDACAO FACULDADE DE MEDICINA</p> <p>Patrocinador Principal: Financiamento Próprio</p> <p>DADOS DO PARECER</p> <p>Número do Parecer: 620.741</p> <p>Data da Relatoria: 19/02/2014</p> <p>Apresentação do Projeto:</p> <p>Este trabalho visa estudar a atividade do HERVs (retrovirus endogenos) e sua relacao com os genes relacionados a inflamaCao e autoimunidade, entre os qauis a esclerose multipla. Eles utilizarao uma ferramenta que e o Next Generation Sequencing (NGS).</p> <p>Objetivo da Pesquisa:</p> <p>Comparar o transcriptoma de HERV-W de individuos com EM em momento de recorrencia com o perfil de expressao de individuos saudaveis.</p> <p>Avaliação dos Riscos e Benefícios:</p> <p>Riscos: minimos - coleta de sangue</p> <p>beneficios imediatos - nenhum para o participante da pesquisa</p> <p>Comentários e Considerações sobre a Pesquisa:</p> <p>Comentarios e Consideracoes sobre a Pesquisa:</p>								
<table style="width: 100%; border: none;"> <tr> <td style="border: none;">Endereço: DOUTOR ARNALDO 251 21º andar sala 36</td> <td style="border: none;">CEP: 01.246-903</td> </tr> <tr> <td style="border: none;">Bairro: PACAEMBU</td> <td style="border: none;"></td> </tr> <tr> <td style="border: none;">UF: SP</td> <td style="border: none;">Município: SAO PAULO</td> </tr> <tr> <td style="border: none;">Telefone: (11)3893-4401</td> <td style="border: none;">E-mail: cep.fm@usp.br</td> </tr> </table>	Endereço: DOUTOR ARNALDO 251 21º andar sala 36	CEP: 01.246-903	Bairro: PACAEMBU		UF: SP	Município: SAO PAULO	Telefone: (11)3893-4401	E-mail: cep.fm@usp.br
Endereço: DOUTOR ARNALDO 251 21º andar sala 36	CEP: 01.246-903							
Bairro: PACAEMBU								
UF: SP	Município: SAO PAULO							
Telefone: (11)3893-4401	E-mail: cep.fm@usp.br							

FACULDADE DE MEDICINA DA
UNIVERSIDADE DE SÃO
PAULO - FMUSP

Continuação do Parecer: 620.741

Interessante e bastante atual o assunto do estudo. A única consideração refere-se ao número da amostra, se é suficiente para resultados conclusivos. Cálculo amostral?

Considerações sobre os Termos de apresentação obrigatória:

em ordem, sem comentários

Recomendações:

sem recomendações

Conclusões ou Pendências e Lista de Inadequações:

Pendências anteriores - foram resolvidas

Inadequações - resolvidas.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Aprovado

SAO PAULO, 22 de Abril de 2014

Assinador por:
Paulo Euripedes Marchiori
(Coordenador)

Endereço: DOUTOR ARNALDO 251 21º andar sala 36
Bairro: PACAEMBU **CEP:** 01.246-903
UF: SP **Município:** SAO PAULO
Telefone: (11)3893-4401 **E-mail:** cep.fm@usp.br

Anexo 2. Aprovação do comitê de ética da Faculdade de Medicina da Santa Casa de Misericórdia de São Paulo

PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Perfil de expressão dos Retrovírus Endógenos Humanos da família W (HERV-W) em indivíduos com Esclerose Múltipla em condições de recorrência

Pesquisador: Camila Malta Romano

Área Temática:

Versão: 1

CAAE: 25239113.4.3001.5479

Instituição Proponente: FUNDACAO FACULDADE DE MEDICINA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 896.341

Data da Relatoria: 28/05/2014

Apresentação do Projeto:

Serão coletadas 20 amostras de sangue total sendo 10 de indivíduos com EM em condições de recorrência (GEM) e 10 indivíduos saudáveis (GC) as quais serão submetidas ao transcriptoma (RNAseq) pela plataforma Ion Proton. Para análise dos transcritos serão utilizadas como referência sequências de HERV e sequências de genes responsáveis pela expressão de citocinas pró e anti-inflamatórias. As análises das sequências obtidas pelo RNAseq serão realizadas com o auxílio de programas específicos como Geneious e CLC Genomics Workbench. Além disso, os níveis plasmáticos dessas citocinas também serão mensurados para que se possa estabelecer uma relação entre atividade/repressão dos genes associados e níveis plasmáticos encontrados no momento do surto.

Objetivo da Pesquisa:

Objetivo Primário:

Comparar o transcriptoma de HERV-W de indivíduos com EM em momento de recorrência com o perfil de expressão gênica de indivíduos saudáveis.

Objetivo Secundário:

Realizar o sequenciamento em larga escala dos RNA mensageiros obtidos de células mononucleares do sangue periférico dos indivíduos do estudo; Identificar e analisar, através de

Endereço: SANTA ISABEL

Bairro: VILA BUARQUE

CEP: 01.221-010

UF: SP

Município: SAO PAULO

Telefone: (11)2176-7689

Fax: (11)2176-7688

E-mail: cepsc@santacasasp.org.br

Continuação do Parecer: 896.341

ferramentas de bioinformática, a presença, distribuição, nível e frequência de transcritos de HERVs, com ênfase na família W e de genes expressos de determinadas citocinas pró e anti-inflamatórias; Identificar os provírus que possuem nível de transcrição diferenciado entre os grupos analisados; Caracterizar as regiões promotoras responsáveis pela transcrição dos HERV-W; Investigar citocinas anti-inflamatórias e pró-inflamatórias no plasma obtido de pacientes com EM e comparar com dados de expressão gênica.

Avaliação dos Riscos e Benefícios:

Riscos:

Por ser um estudo observacional, ou seja, sem qualquer tipo de intervenção nos regimes terapêuticos aos quais os pacientes estão submetidos, esse estudo não oferece riscos para os voluntários envolvidos. A coleta de sangue consiste no único procedimento invasivo, o que pode, em casos extremos, acarretar em hematoma no local da punção, que deve desaparecer em um ou dois dias após a coleta.

Benefícios:

Os participantes não terão despesas pessoais por participar desse estudo, e nem qualquer obrigação. Também não haverá compensação financeira pela participação. A princípio, esse estudo não trará nenhum benefício imediato a aos participantes. Entretanto, as informações obtidas de sua participação poderão beneficiar no futuro outras pessoas nas mesmas condições clínicas de outros indivíduos com Esclerose Múltipla.

Comentários e Considerações sobre a Pesquisa:

Pesquisa com temática importante.

Considerações sobre os Termos de apresentação obrigatória:

De acordo.

Recomendações:

Conclusões ou Pendências e Lista de Inadequações:

Aprovado.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: SANTA ISABEL
 Bairro: VILA BUARQUE CEP: 01.221-010
 UF: SP Município: SAO PAULO
 Telefone: (11)2176-7689 Fax: (11)2176-7688 E-mail: cepsc@santacasasp.org.br

Continuação do Parecer: 896.341

Considerações Finais a critério do CEP:

SAO PAULO, 03 de Dezembro de 2014

Assinado por:
Nelson Keiske Ono
(Coordenador)

Endereço: SANTA ISABEL
Bairro: VILA BUARQUE **CEP:** 01.221-010
UF: SP **Município:** SAO PAULO
Telefone: (11)2176-7689 **Fax:** (11)2176-7688 **E-mail:** cepsc@santacasasp.org.br

Página 03 de 03

Anexo 3. Termo de Consentimento Livre e Esclarecido fornecido aos voluntários da pesquisa

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME:

DOCUMENTO DE IDENTIDADE Nº : SEXO : M ☐ F ☐

DATA NASCIMENTO:/...../.....

ENDEREÇO Nº APTO:

BAIRRO: CIDADE:

CEP:.....TELEFONE: DDD (.....)

2. RESPONSÁVEL LEGAL

NATUREZA (grau de parentesco, tutor, curador etc.)

DOCUMENTO DE IDENTIDADE :SEXO: M ☐ F ☐

DATA NASCIMENTO:/...../.....

ENDEREÇO:.....Nº.....APTO:

BAIRRO: CIDADE:

CEP:.....TELEFONE: DDD ().....

DADOS SOBRE A PESQUISA

1. TÍTULO DO PROTOCOLO DE PESQUISA: **Perfil de expressão dos Retrovírus Endógenos Humanos da família W (HERV-W) em indivíduos com Esclerose Múltipla em condições de recorrência.**

PESQUISADOR : **Dra. CAMILA MALTA ROMANO**

CARGO/FUNÇÃO: **PESQUISADOR CIENTÍFICO**

UNIDADE DO HCFMUSP: **LABORATÓRIO DE INVESTIGAÇÃO MÉDICA LIM-**

52

2. AVALIAÇÃO DO RISCO DA PESQUISA:

RISCO MÍNIMO	X	RISCO MÉDIO	☐
RISCO BAIXO	☐	RISCO MAIOR	☐

3. DURAÇÃO DO ESTUDO : 4 ANOS

4. DURAÇÃO DO ACOMPANHAMENTO DOS PACIENTES: **corte transversal**

1 – Desenho do Estudo e Objetivo

Estamos realizando um estudo de caráter observacional em indivíduos com Esclerose Múltipla, esse estudo visa analisar as variações dos níveis de RNA mensageiro de um gene específico, presente em nosso DNA, parecido com o gene de um retrovírus, conhecido como HERV-W. Esse gene está presente em todos os humanos. Isso está sendo proposto pelo fato de que esse gene aparentemente pode apresentar alguma relação com a doença, para isso também avaliaremos a expressão de citocinas (fatores imunológicos) responsáveis pelo início ou fim da inflamação. O objetivo deste termo é convidá-lo a participar desse estudo, que irá comparar o perfil do nível dos genes citados de indivíduos com Esclerose Múltipla em momento de recorrência com o perfil do nível dos genes de indivíduos saudáveis. É importante mencionar que nenhum tipo de interferência será feita por parte do estudo nos regimes terapêuticos que os voluntários estejam recebendo.

Dessa forma, de acordo com a decisão do seu médico, você pode ser incluído no seguinte grupo:

- 1) Pessoas que tem Esclerose Múltipla que estejam em um momento de recorrência (retorno da doença) da doença;

Caso você não possua Esclerose Múltipla e nem possua histórico da doença na família, você será incluído no seguinte grupo:

- 2) Pessoas que não possuam Esclerose Múltipla.

2 – Procedimentos de Pesquisa

Após ter sido esclarecido sobre o estudo, você poderá decidir se deseja ou não participar. Se decidir participar, você será solicitado a assinar o **Termo de Consentimento Livre e Esclarecido** ou colocar suas impressões digitais em frente a uma testemunha de sua confiança. Uma cópia deste documento lhe será fornecida. O seu acompanhamento e tratamento prosseguirão normalmente a cargo dos médicos que lhe acompanham e não será modificado de maneira alguma caso você decida participar ou não da pesquisa. Além disso, você poderá optar por não participar mais do estudo a qualquer momento, sem que exista qualquer implicação no seu tratamento médico. Caso você aceite participar, serão solicitadas duas amostras sangue a serem coletadas nos dias em que estiver sob os cuidados médicos. O seu material e os dados coletados durante este projeto serão utilizados apenas para o estudo da presença do material genético dos pacientes, e poderá ser armazenado para posteriores estudos que envolvam a investigação de outros vírus. Entretanto, o paciente deverá ser contatado para permitir o uso da amostra em outro (s) estudo (s).

3 - Benefícios e Obrigações do Participante

Você não terá despesas pessoais por participar desse estudo, e nem qualquer obrigação. Também não haverá compensação financeira por sua participação. A princípio, esse estudo não trará nenhum benefício imediato a você. Entretanto, as informações obtidas de sua participação poderão beneficiar no futuro outras pessoas nas mesmas condições que você.

4 – Riscos para o participante

Por ser um estudo observacional, ou seja, sem qualquer tipo de intervenção nos regimes terapêuticos aos quais os pacientes estão submetidos, esse estudo não oferece riscos para os voluntários envolvidos. A coleta de sangue consiste no único procedimento invasivo, o que pode, em casos extremos, acarretar em hematoma no local da punção, que deve desaparecer em um ou dois dias após a coleta.

5- Garantia de Acesso aos Pesquisadores

Você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. Um deles é Camila M. Romano, que pode ser encontrada na Rua Eneas de Carvalho Aguiar, 470, prédio IMTI, 2º andar, SP. Tel: 3061-7020, ou pelo endereço eletrônico cmromano@usp.br. Ou ainda, com o aluno de pós graduação Luiz Henrique da Silva Nali no mesmo endereço e telefone email luiznali@usp.br.

6 – Direito de Confidencialidade

A confidencialidade das informações obtidas nesse estudo será garantida e os dados obtidos serão analisados em conjunto com os de outros voluntários, não sendo divulgada a identificação do voluntário. Sua participação neste estudo é confidencial e sigilosa. Você poderá ter acesso aos resultados dos seus exames colhidos durante o estudo, bem como terá acesso, se assim desejar, às publicações pertinentes.

7 - Autorização para armazenamento e utilização futura das suas amostras de sangue para novas pesquisas.

7.1 - Quais são as razões para guardar suas amostras obtidas durante esta pesquisa?

Toda pesquisa tem como objetivo buscar resposta para algum aspecto ainda não bem conhecido de determinadas doenças. Contudo, com a rápida evolução da medicina, novos fatos e novas dúvidas estão sempre surgindo. As amostras clínicas obtidas durante uma determinada pesquisa, se guardadas de forma segura e adequada, podem

ser usadas pelos investigadores para responder novas questões de forma rápida e eficiente, dispensando novas coletas.

7.2 - O que envolve autorizar o armazenamento de amostras clínicas para uso futuro?

Como parte deste projeto de pesquisa você deverá colher amostras de sangue para avaliar os níveis de HERV-W e os fatores imunológicos citado anteriormente. Estas amostras, depois de colhidas e examinadas no laboratório, serão armazenadas em vez de serem desprezadas.

4.3 - O que acontecerá com as suas amostras?

Suas amostras irão para o Repositório de Amostras do Laboratório de Virologia do Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo (LIM HC-FMUSP) e serão armazenadas de forma segura e adequada. Estas amostras ficarão disponíveis para pesquisas futuras que poderão incluir avaliações laboratoriais extras na sua amostra, juntamente com informações clínicas obtidas durante a pesquisa original durante a qual foram coletadas estas amostras. Pode ser que estas amostras sejam utilizadas somente daqui a muitos anos, mas elas podem ajudar muitos pacientes no futuro. Qualquer projeto de pesquisa a ser realizada no futuro utilizando suas amostras deverá ser submetido à Comissão de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, que irá avaliar se o estudo proposto justifica a sua utilização. Além disso, antes de utilizar estas amostras, você será contatado para dar uma autorização específica para a nova pesquisa que está sendo proposta.

A autorização para guardar suas amostras no Repositório de Amostras Clínicas do Laboratório de Virologia do HC-FMUSP é inteiramente de sua escolha. Você pode decidir não autorizar o armazenamento de suas amostras. Mesmo que você decida fornecer as amostras, você pode retirar sua permissão a qualquer momento, por qualquer razão. Se você decidir retirar sua permissão suas amostras serão destruídas. No entanto, a informação já obtida de suas amostras até a sua retirada será utilizada.

Se você não der permissão para a utilização futura de suas amostras, isto não afetará de nenhuma forma sua participação no presente estudo ou seu cuidado médico atual.

Seu médico pode retirar sua autorização para o Projeto de Repositório de Amostras Clínicas por qualquer razão, como por exemplo, se for difícil coletar seu sangue.

CONSENTIMENTO

Eu, abaixo assinado, afirmo que li na íntegra e entendi completamente as informações. Concordo que as informações a respeito de minha condição médica e os resultados obtidos através de meu sangue podem ser utilizados neste estudo. Minha participação é voluntária e livre de qualquer tipo de pressão ou coação. Eu entendo que estas informações serão confidenciais e que meu nome não será mencionado em qualquer publicação deste estudo.

Nome completo do voluntário/representante legal

Data

Assinatura do voluntário/representante legal

Data

Declaro que obtive de forma apropriada e voluntária o **Termo de Consentimento Livre e Esclarecido** deste paciente ou representante legal para a participação neste estudo

Pesquisadora responsável: Camila Malta Romano

Data

Médico responsável:

Data

CONSENTIMENTO***Permissão de armazenamento***

Eu, abaixo assinado, afirmo que li na íntegra e entendi completamente as informações. Concordo que as informações a respeito de minha condição médica e os resultados obtidos através de meu sangue podem ser armazenados para que seja possível estudá-los em projetos futuros. Entendo que os pesquisadores deverão entrar em contato comigo caso minha amostra seja utilizada para outro estudo além deste. Minha participação é voluntária e livre de qualquer tipo de pressão ou coação.

Eu entendo que estas informações serão confidenciais e que meu nome não será mencionado em qualquer publicação deste estudo.

Nome completo do voluntário/representante legal

Data

Assinatura do voluntário/representante legal

Data

Declaro que obtive de forma apropriada e voluntária o **Termo de Consentimento Livre e Esclarecido** deste paciente ou representante legal para a participação neste estudo

Pesquisadora responsável: Camila Malta Romano

Data

Médico responsável:

Data

Anexo 4. Produção do aluno**Artigos publicados durante o período do doutoramento**

1. do Olival GS, Faria TS, **Nali LH**, de Oliveira AC, Casseb J, Vidal JE, Cavenaghi VB, Tilbery CP, Moraes L, Fink MC, Sumita LM, Perron H, Romano CM. *Genomic analysis of ERVWE2 locus in patients with multiple sclerosis: absence of genetic association but potential role of human endogenous retrovirus type W elements in molecular mimicry with myelin antigen*. Front Microbiol. 2013 Jun 26;4:172. doi:10.3389/fmicb.2013.00172.
2. do Olival GS, Lima BM, Sumita LM, Serafim V, Fink MC, **Nali LH**, Romano CM, Thomaz RB, Cavenaghi VB, Tilbery CP, Penalva-de-Oliveira AC. *Multiple sclerosis and herpesvirus interaction*. Arq Neuropsiquiatr. 2013 Sep;71(9B):727-30. doi:10.1590/0004-282X20130160. Review. PubMed PMID: 24141514.
3. **Nali LH**, Moraes L, Fink MC, Callegaro D, Romano CM, Oliveira AC. *Natalizumabe treatment for multiple sclerosis: updates and considerations for safer treatment in JCV positive patients*. Arq Neuropsiquiatr. 2014 Dec;72(12):960-5. doi:10.1590/0004-282X20140142. Epub 2014 Dec 2. PubMed PMID: 25465776.
4. Urbano PR, **Nali LH**, Bicalho CS, Pierrotti LC, David-Neto E, Pannuti CS, Romano CM. *New findings about trichodysplasia spinulosa-associated polyomavirus (TSPyV)--novel qPCR detects TSPyV-DNA in blood samples*. Diagn Microbiol Infect Dis. 2016 Feb;84(2):123-4. doi: 10.1016/j.diagmicrobio.2015.10.011. Epub 2015 Oct 21. PubMed PMID: 26602950.
5. Bicalho CS, Oliveira RR, Pierrotti LC, Fink MC, Urbano PR, **Nali LH**, Luna EJ, Romano CM, David DR, David-Neto E, Pannuti CS. *Pre-transplant shedding of BK virus in urine is unrelated to post-transplant viruria and viremia in kidney transplant recipients*. Clin Transplant. 2016 Jul;30(7):796-801. doi: 10.1111/ctr.12752. Epub 2016 May 29. PubMed PMID: 27101526.

6. **Nali LH**, Fink MC, do Olival GS, Moraes L, Callegaro D, Tilbery CP, Vidal JE, Sumita LM, de Oliveira AC, Romano CM. *Polyomavirus detection in multiple sclerosis patients under Natalizumabe therapy: Profile and frequency of urinary shedding*. J Med Virol. 2017 Mar;89(3):528-534. doi: 10.1002/jmv.24646. Epub 2016 Aug 9. PubMed PMID: 27464945.

7. **Nali LH**, Fujita DM, Salvador FS, Fink MC, de Andrade HF Jr, Pannuti CS, Luna EJ. *Potential measles transmission risk in mass gatherings: Are we safe for the Olympic games--Rio 2016?* J Travel Med. 2016 May 13;23(4). pii: taw026. doi: 10.1093/jtm/taw026. Print 2016 Apr. PubMed PMID: 27178161.

8. Fujita DM, **Nali LH**, Urbano PR, Soeiro DM, Andrade HF Jr. *The fast transmission of infectious diseases around the world - a new concern to the public health*. Braz J Infect Dis. 2016 Sep-Oct;20(5):513-5. doi: 10.1016/j.bjid.2016.06.003. Epub 2016 Jul 26. PubMed PMID: 27473892.

9. **Nali LH**, Oliveira AC, Alves DO, Caleiro GS, Nunes CF, Gerhardt D, Succi RC, Romano CM, Machado DM. *Expression of human endogenous retrovirus K and W in babies*. Arch Virol. 2017 Mar;162(3):857-861. doi: 10.1007/s00705-016-3167-2. Epub 2016 Nov 24. PubMed PMID: 27885560.