

**APLICAÇÕES DO SEQUENCIAMENTO DO EXOMA
NA INVESTIGAÇÃO DAS BASES BIOLÓGICAS DA
ESCLEROSE MÚLTIPLA (EM)**

SHEILA GARCIA

**Dissertação apresentada à Fundação Antônio
Prudente para obtenção do título de Mestre em
Ciências**

Área de Concentração: Oncologia

Orientador: Dr. Emmanuel Dias-Neto

Co-Orientadora: Dra. Diana Noronha Nunes

São Paulo

2015

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Garcia, Sheila

Aplicações do sequenciamento do exoma na investigação das bases biológicas da esclerose múltipla (EM) / Sheila Garcia – São Paulo; 2015.
101p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração: Oncologia.

Orientador: Emmanuel Dias-Neto

Descritores: 1. EXOMA. 2. SEQUENCIAMENTO DE NUCLEOTÍDEOS EM LARGA ESCALA. 3. ESCLEROSE MÚLTIPLA RECIDIVANTE-REMITENTE/genética. 4. MUTAÇÕES/genética. 5. RETROVIRUS ENDÓGENOS/genética.

“Quem elegeu a busca não pode recusar a travessia.”

Guimarães Rosa

DEDICATÓRIA

*“À Deus, que me permitiu vislumbrar essa oportunidade.
Ao meu pai, Wanderley, que me incentivou e sempre esteve ao meu lado
dando apoio total e tornou possível que eu chegasse aqui.
À minha mãe pelo suporte emocional e por ser um exemplo de mulher.
Ao Gil pelo incentivo, amor e paciência.”*

AGRADECIMENTOS

Ao Dr. Emmanuel Dias-Neto, pela oportunidade de crescer profissionalmente em um excelente centro de pesquisa e por não medir esforços para conseguir colaborações. Agradeço pela paciência e delicadeza ao me aconselhar e por ter acreditado no meu trabalho. Sabemos que não é fácil, mas vale a pena quando chegamos na reta final e sabemos que a excelência do nosso orientador nos faz acreditar no potencial do nosso trabalho e nos torna seguros em finalizar mais uma etapa da nossa jornada científica.

Aos meus irmãos, que são e sempre serão grandes exemplos. Obrigada por estarmos juntos sempre e por me dar sobrinhos lindos e carinhosos.

Ao Gil, meu amor, sou muito feliz por ter você como companheiro sempre me apoiando. Obrigada por todo incentivo, amor, agrados e paciência.

À Dra. Diana Noronha Nunes, a pessoa que abriu as portas do CIPE na última semana do processo seletivo e que, mesmo sem a presença do Emmanuel naquela semana, fez todo o possível para que eu pudesse ingressar no mestrado. Jamais me esquecerei desse dia. Obrigada por estar sempre disposta a ajudar/aconselhar.

Ao Prof. Keith Okamoto pela orientação, conselhos e amizade, durante o período que estive em seu laboratório.

À Aline Grazielle, pela amizade e companheirismo.

Aos meus amigos que mesmo distante sempre estão presentes com afetos e boas conversas.

À Dra. Márcia Santos, minha amiga e co-orientadora da iniciação científica e maior exemplo profissional a seguir. Agradeço por ter me dado a oportunidade de aprender ciência e por todo o incentivo.

Aos colegas de laboratório, Juliana Laino, Juliana Vasconcelos, Gabi, Ana, Andrew, em especial ao Frederico pelos momentos alegres diários, pelas sugestões, apoio e amizade durante esse período.

À Dra. Maria Galli de Amorim e à Melissa Pizzi por todo apoio, amizade e colaboração.

Aos colegas de Laboratório da Bioinformática, especialmente ao Renan Valieris e Dr. Jorge Stefano, pela ajuda nas análises dos dados.

Ao Dr. Fábio Marchi pela ajuda em vários momentos.

À Dra. Dirce Carraro e aos colegas do laboratório de Genômica e Biologia Molecular, em especial às Dras. Bruna Barros e Elisa Napolitano e pelo conhecimento transmitido e pela convivência.

Aos membros da banca de qualificação, especialmente ao Dr. Vladimir Lima, pelas sugestões e comentários no desenvolvimento deste trabalho.

Ao Dr. João Paulo Kitajima pelas análises de predição de patogenicidade no programa Abracadabra.

Ao Prof. Sergio Pena pela discussões quando das análises de parentesco dos indivíduos deste estudo.

Ao Dr. Paulo Sérgio Lopes de Oliveira e ao Msc. José Geraldo de Carvalho Pereira pelas análises, *in silico*, da estrutura de GALR2.

Ao Dr. Israel Tojal da Silva da Rockefeller University, New York USA, pela colaboração com as análises de retrovírus endógenos humanos.

Às meninas que recepcionam meu dia com muita alegria, Vanessa e Patrícia.

À toda a equipe de funcionários da Pós Graduação, em especial à Vanuza.

Aos colaboradores de Ribeirão Preto, Dr. Wilson da Silva Jr e em especial ao Dr. Julio C. C. Lorenzi, por todo seu suporte durante as coletas e entusiasmo com os resultados. Pela dedicação ao projeto e pela disponibilidade.

Aos colaboradores da Santa Casa, em especial ao Dr. Guilherme Olival e Dra. Juliana Amm Diniz.

Ao Dr. Victor Piana pela disponibilidade e acessibilidade.

À enfermeira Sandra Cardoso, pela ajuda nas coletas de sangue.

Ao Sr. Waldemar Benassi, pelo suporte financeiro que viabilizou este projeto e inúmeros outros que fizeram uso do sequenciador de DNA *Ion Proton* adquirido graças à sua generosa doação.

Aos Drs. Doralina Guimarães Brum (UNESP), Amilton Barrera (FMUSP, Ribeirão Preto) e Alessandro Farias (UNICAMP) e Camila Romano por cederem amostras e ajudarem em discussões do projeto.

À equipe da biblioteca do A.C.Camargo Cancer Center.

Em memória ao Dr. Ricardo Brentani que nos confiou a missão de planejar um projeto em esclerose múltipla, idealizou e concretizou o CIPE.

Ao CNPq, à CAPES e à FAPESP, pelo suporte financeiro indispensável para o desenvolvimento deste projeto.

RESUMO

Garcia S. **Aplicações do sequenciamento do exoma na investigação das bases biológicas da esclerose múltipla**. São Paulo; 2015. [Dissertação de Mestrado-Fundação Antônio Prudente].

Esclerose Múltipla (MS) é uma doença inflamatória degenerativa crônica, desmielinizante e de origem multifatorial. Esta doença afeta aproximadamente 2,3 milhões de adultos/jovens no mundo e sua etiologia ainda é pouco entendida. Neste estudo, empregamos sequenciamento completo do exoma de 31 indivíduos distribuídos em 9 trios/quartetos (pai-mãe-filho afetado e irmã/o) envolvendo pacientes com Esclerose Múltipla (EM). Para o estudo foram coletados leucócitos de sangue periférico e partimos para a construção de bibliotecas e sequenciamento completo na plataforma Ion Proton (Life Technologies). No total, geramos 1.935.266.861 de sequências com uma média de 61 milhões *reads*/indivíduo, permitindo que aproximadamente 90% do exoma tivesse uma cobertura > 20 vezes. A partir destes dados analisamos: a) o perfil de mutações *de novo* (ausentes nos pais/irmãos(ãs) presentes nos pacientes), encontrando 12 variações; b) inserções *de novo* de retrovírus (DNR) endógenos humanos dentro de genes codificadores de proteínas, encontrando um total de 1208 eventos, sendo 109 eventos compartilhados por pelo menos duas famílias independentes; c) a herança de polimorfismos raros de alelos em homozigose nos pacientes afetados (SRHH), onde encontramos um total de 24 SRHH. Dentre estes cabe ressaltar a identificação de um SNP raro em GALR2 (0,6% de frequência na população global), que aparentemente gera um receptor defectivo. Por sua importância em processos de mielinização, este achado pode indicar uma via importante para a gênese da EM podendo, inclusive, apontar alternativas terapêuticas importantes que incluam a utilização de análogos da galanina. Nossas análises estruturais de GALR2 usando modelagem por homologia sugerem que a alteração W249L não impede a ligação da galanina, mas possivelmente altera sua capacidade de sinalização intracelular. Esses achados expandem nossa compreensão sobre o panorama genético da EM e podem prover fonte de estudos futuros com aplicações diagnósticas e terapêuticas da EM.

SUMMARY

Garcia S. [**Applications of exome sequencing for the research of the biological bases of multiple sclerosis**]. São Paulo; 2015. [Dissertação de Mestrado-Fundação Antônio Prudente].

Multiple Sclerosis (MS) is a chronic, multifactorial, inflammatory, demyelinating and degenerative disease that affects approximately 2.3 million adults/young-age people in the world. The etiology of MS is still poorly understood and its heterogeneity is indicative of a combination of genetic and environmental factors leading to significant variability in its clinical presentation. In this study, we performed whole-exome sequencing of 31 MS individuals distributed in nine trios/quartets (father-mother-MS patient and a healthy sibling). We constructed gDNA libraries from peripheral blood leukocytes and performed massive sequencing on the *Ion Proton* (Life Technologies) platform. We generated a total of 1.935.266.861 reads, with an average of 61 million reads/individual, which enabled a coverage of >20X for at least 90% of the exome. From this we analyzed: i) *de novo* mutations (absent in either the father or mother, but present in the affected patient); ii) *de novo* retro-insertions of human endogenous virus, inside coding regions of genes, identifying a total of 1208 *de novo* retro-insertion events (DNR), of which 109 events were identified in at least two independent patients; iii) the homozygous inheritance of rare polymorphisms that segregate with the disease, which identified 24 events. Here we highlight the homozygous identification of a rare *GALR2* allele (global frequency of 0,6%), which may produce a defective receptor. Due to its important role in the neural regeneration and myelination processes, we believe that this *GALR2* variant may be involved in MS development, suggesting that the administration of galanin agonists may represent an putative new therapeutic approach for selected MS patients. Using modeling by homology our structural analysis of *GALR2* suggests that the W249L alteration does impair galanin binding, but can probably alter the intracellular signaling pathways. Our results expand the understanding of the genetic landscape of MS and indicate new exome analytical approaches might have diagnostic and therapeutic applications for MS.

LISTA DE FIGURAS

Figura 1	Prevalência de EM no mundo.....	3
Figura 2	Quantidade de DNA do EMRV entre controles (doadores de sangue) e pacientes com EM.....	17
Figura 3	O risco de ocorrência de esclerose múltipla é diretamente proporcional à proximidade genética com uma pessoa afetada.....	22
Figura 4	Exemplo do perfil de distribuição de tamanho após fragmentação enzimática do gDNA, visualizadas utilizando o equipamento bioanalyzer.....	35
Figura 5	Exemplo de perfil de tamanho das bibliotecas após amplificação, avaliada através do bioanalyzer.....	34
Figura 6	Análise do tamanho e quantificação da biblioteca final.....	35
Figura 7	Resultado do sequenciamento do exoma.....	37
Figura 8	<i>Pipeline</i> geral de análise de exoma para a determinação de bases variantes.....	38
Figura 9	Fluxograma do pipeline para detecção dos eventos de retroinserção de novo (DNR) obtidos através dos dados do sequenciamento do exoma.....	40
Figura 10	Sequenciamento do exoma de nove trios/quartetos utilizando a plataforma Ion Proton.....	44
Figura 11	SNVs de novo validados através da ferramenta Sanger.....	49
Figura 12	Localização dos eventos DNR dos grupos quatro (G4), cinco(G5) e seis (G6) no genoma.....	56
Figura 13	Localização dos eventos DNR dos grupos três (G3) e seis (G6) no genoma.....	58

Figura 14	Localização dos eventos DNR dos grupos quatro (G4) e seis (G6) no genoma.....	59
Figura 15	Localização dos eventos DNR dos grupos três (G3) e seis (G6) no genoma.....	59
Figura 16	Localização dos eventos DNR dos grupos oito(G8) e três (G3) no genoma.....	60
Figura 17	Cromatogramas de sequenciamento Sanger representativos das regiões de interesse, contendo os SNPs raros, herdados em homozigose, validados neste estudo.....	65
Figura 18	Validação pelo método Sanger do candidato SNP raro rs61729458 no gene LAMB2.....	67
Figura 19	Conservação da arginina entre as espécies na posição onde ocorre o SNP rs61729458 de LAMB2.....	68
Figura 20	Validação pelo método Sanger do candidato SNP raro rs45621431 no gene CYP3A43.....	69
Figura 21	Validação pelo método Sanger do SNP rs112318500 do gene BIN1 em homozigose.....	70
Figura 22	Expressão do gene BIN1 nos diversos tecidos cerebrais analisados pelo portal GTEX.....	70
Figura 23	Conservação da treonina entre as espécies na posição onde ocorre o SNP rs112318500.....	71
Figura 24	Dados de expressão expressão normalizada dos mRNAs codificadores de galanina (GAL).....	73
Figura 25	Os gráficos apresentam a expressão normalizada dos mRNAs codificadores de GALR2 em diferentes tecidos humanos.....	74
Figura 26	Conservação do aminoácido triptofano (W) entre as espécies na posição onde ocorre o SNP rs61745847.....	75

Figura 27	Validação pelo método Sanger do candidato SNP raro rs61745847	76
Figura 28	Representação do sequenciamento do <i>amplicon</i> GALR2 da amostra do metrô de São Paulo-SP.....	78
Figura 29	Comparação da frequência populacional de SNPs missense descritos no ExAC (count >2).....	79
Figura 30	Vista lateral (A e B) e vista superior (extracelular > intracelular) (C e D) da proteína GalR2 inserida na membrana celular.....	80
Figura 31	Análise estrutural do volume das cavidades após simulação da ligação da galanina.....	81

LISTA DE QUADROS E TABELAS

Quadro 1	Informações gerais dos pacientes e das famílias avaliadas.....	31
Quadro 2	Análise de discrepância de variantes em homozigose nas famílias participantes do estudo.....	46
Quadro 3	Candidatos <i>Single Nucleotide Variants</i> (SNVs) de novo.....	47
Quadro 4	Genes com eventos DNR encontrados exclusivamente em portadores de EM e compartilhados por pelo menos dois pacientes.....	53
Quadro 5	Análise de enriquecimento de vias por ontologia utilizando a ferramenta DAVID e validação in silico através da ferramenta KOBAS 2.0.....	61
Quadro 6	SNPs raros herdados em homozigose, localizados dentro de ORFs de genes humanos e com impacto na sequência das proteínas, exclusivos de pacientes com EM.....	63
Quadro 7	Genes, SNPs e características dos iniciadores usados na validação dos polimorfismos raros.....	64
Quadro 8	Frequências observadas para o alelo variante (T) do SNP rs61745847.....	77
Tabela 1	Sequências obtidas para cada indivíduo analisado neste estudo, incluindo dados de mapeamento.....	45
Tabela 2	Validação de candidatos SNVs de novo.....	48
Tabela 3	Análise de vias de doença utilizando a ferramenta DAVID.....	55

LISTA DE ABREVIATURAS

ACCCC	A.C.Camargo Cancer Center
AD	Alzheimer disease
AhR	Aryl hydrocarbon Receptor
CDNF	Fator neurotrófico de dopamina cerebral
Chr	Chromossomo
CIPE	Centro Internacional de Pesquisa
CIS	Síndrome clinicamente isolada
CNV	Copy number variation
CPO	Células progenitoras oligodendrocíticas
CPZ	Cuprizona
DIS	Disseminação no espaço
DIT	Disseminação no tempo
DNR	de novo retroinsertion
EAE	Encefalomielite autoimune experimental
EBV	Epstein Barr Virus
EDSS	Expanded disability status scale
EDTA	Ethylenediamine tetraacetic acid
EM	Esclerose múltipla
EMPP	Esclerose múltipla primária progressiva
EMRR	Esclerose múltipla remitente recorrente
EMRV	Esclerose múltipla associada a retrovírus
env	Envelope de EMRV
ERV	Endogenous retrovirus
FALS	Familial Amyotrophic Lateral Sclerosis
FDR	False discovery rate
GO	Gene ontology
GWAS	Genome Wide Association Study
HERV	Human endogenous retrovirus
HHV	Herpesvírus humano

HLA	Antígeno Leucocitário Humano
IFNB	Interferon beta
IGg	Imunoglobulina G
IL	Interleucina
LCR	Líquido cefalorraquidiano
LEMP	Leucoencefalopatia multifocal progressiva
lncRNA	Long noncoding RNA
LOAD	Late onset alzheimer disease
MAF	Minor allele frequency
MHC	Major Histocompatibility Complex
MKP	MAP Kinase Phosphatase
MLV	Vírus da leucemia murina
MMTV	Vírus tumoral mamário de camundongo
MOG	Glicoproteína oligodendrocítica
mRNA	RNA mensageiro
MSIF	Multiple Sclerosis International Federation
NFAT	Fator nuclear de células T ativadas
NGS	Next generation sequencing
NMO	Neuromielite óptica
NO	Óxido nítrico
ORF	Open reading frame
PBMC	Peripheral blood mononuclear cell
PP	Primária progressiva
RE	Retrovírus endógeno
RNM	Ressonância nuclear magnética
RNS	Espécie reativa de nitrogênio
ROS	Espécie reativa de oxigênio
SF	Sistemas funcionais
SNC	Sistema nervoso central
snoRNA	Small nucleolar RNA
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant

SRHH	SNP raro herdado em homozigose
TCLE	Termo de Consentimento Livre Esclarecido
Th	T helper
Tm	Temperatura de melting
tRNA	RNA transportador
TSS	Transcription start site
TTS	Transcription termination site
UV	Ultra violeta
VDR	Vitamin D receptor
VDRE	Elemento responsivo a vitamina D
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

ÍNDICE

1	INTRODUÇÃO.....	1
1.1	Aspectos Gerais da Esclerose Múltipla.....	1
1.2	Características Epidemiológicas	1
1.3	Aspectos diagnósticos e manifestações clínicas	4
1.4	Esclerose Múltipla e Neurodegeneração.....	7
1.5	Terapias utilizadas na EM.....	9
1.6	Etiologia.....	10
1.6.1	Autoimunidade e Esclerose Múltipla.....	10
1.6.2	Fatores Ambientais Associados à Esclerose Múltipla	11
1.6.2.1	Agentes Infecciosos	11
1.6.2.2	Exposição solar e Vitamina D	18
1.6.2.3	Consumo de tabaco	20
1.6.3	Fatores Genéticos Associados à Esclerose Múltipla.....	21
1.7	Análises de exomas como ferramenta de avaliação de doenças neuroológicas	24
1.7.1	Análises alternativas de dados de exoma.....	27
2	OBJETIVOS.....	29
2.1	Objetivo Geral.....	29
2.2	Objetivos Específicos.....	29
3	CASUÍSTICA, MATERIAIS E MÉTODOS.....	30
3.1	Casuística	30
3.2	Confirmação das relações de parentesco entre os indivíduos analisados	32
3.3	Obtenção do DNA.....	32
3.4	Construção das bibliotecas de exoma	33
3.5	Análises de Bioinformática.....	38
3.5.1	Pipeline geral do Exoma	38
3.5.1.1	Variantes nucleotídicas de novo (SNVs)	39

3.5.1.2	Polimorfismos raros de base única (SNPs raros)	39
3.5.1.3	Pipeline das retroinserções de novo (DNR).....	41
3.6	Validação dos Achados	41
3.6.1	Seleção dos Candidatos SNVs de novo	41
3.6.2	Seleção dos Candidatos SNPs raros.....	41
3.6.3	Amplificação dos genes de interesse para validação	42
3.6.4	Validação dos achados por sequenciamento Sanger.....	42
4	RESULTADOS E DISCUSSÃO	44
4.1	Mutações <i>de novo</i> em indivíduos portadores de EM	46
4.2	Retroinserções <i>de novo</i> em indivíduos portadores de EM.....	52
4.3	Avaliação da herança de alelos raros em homozigose –SNPs raros	62
5	CONCLUSÃO	84
6	PERSPECTIVAS FUTURAS	85
7	REFERÊNCIAS BIBLIOGRÁFICAS.....	86

ANEXOS

Anexo 1 Carta de aprovação do Comitê de Ética em Pesquisa-CEP da
Fundação Antônio Prudente e FMRP-USP

Anexo 2 Termo de Consentimento Livre e Esclarecido

1 INTRODUÇÃO

1.1 ASPECTOS GERAIS DA ESCLEROSE MÚLTIPLA (EM)

A esclerose múltipla (EM) é uma grave doença inflamatória, com um importante componente auto-imune e que leva a uma extensa degeneração do sistema nervoso central (SNC). Durante o curso da doença ocorre uma infiltração linfocitária localizada que danifica o revestimento de mielina neuronal através de uma reação exacerbada de linfócitos T. No início da sua manifestação a reação inflamatória é transitória e alguma remielinização ainda ocorre, o que permite uma recuperação de certas disfunções. No entanto, com o avanço da doença os processos que orquestram a re-mielinização perdem a efetividade e são insuficientes para frear a progressão dos sinais, o que leva a uma neurodegeneração extensa e crônica (revisto em COMPSTON e COLES 2008).

O dano na bainha de mielina afeta os neurônios por todo o SNC incluindo cérebro e medula espinhal, impactando negativamente a capacidade de transmissão de sinais elétricos entre neurônios, o que leva a sintomas diversos. A doença se manifesta de formas distintas e novos sinais podem surgir em eventos isolados (forma recorrente) ou de modo contínuo (forma progressiva). Infelizmente a EM tende a se agravar e sintomas neurológicos permanentes e debilitantes ocorrem com frequência (COMPSTON e COLES 2008).

Ainda não existe cura conhecida para a EM e os tratamentos atuais são parcialmente efetivos, aliviam alguns sintomas (melhorando de forma ainda modesta a função neurológica) e diminuem a severidade ou a ocorrência dos episódios, mas conferem uma variedade de efeitos adversos (COMPSTON e COLES 2008; MENGE et al. 2014).

1.2 CARACTERÍSTICAS EPIDEMIOLÓGICAS

A EM afeta cerca de 2,3 milhões de pessoas no mundo e tem uma prevalência média de cerca de 30/100.000 habitantes (TRISOLINI e HONEYCUTT 2010). No

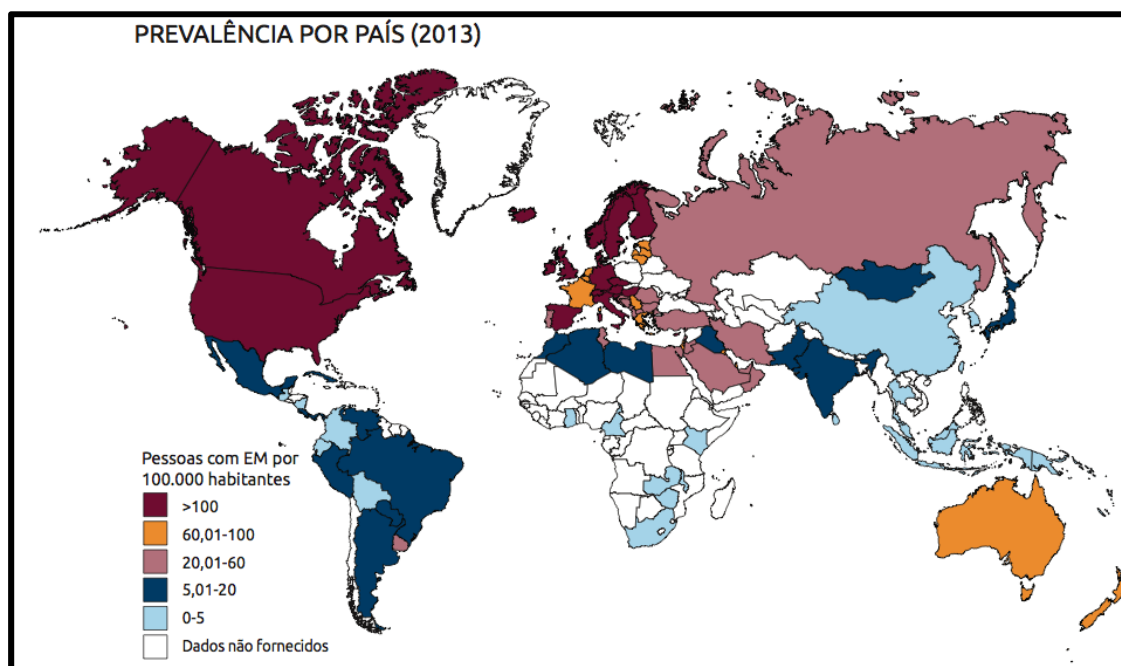
Brasil estima-se uma prevalência entre 15-18/100.000 habitantes (dados do estado de São Paulo e da cidade de Belo Horizonte) (CALLEGARO et al. 2001; LANA-PEIXOTO et al. 2012).

O início dos sintomas normalmente ocorre entre 20-40 anos de idade e, em maior frequência nas mulheres, com uma taxa de 3 mulheres para cada homem afetado (BOIKO et al. 2002; MILO e MILLER 2014).

Estudos epidemiológicos sugerem que a prevalência de EM aumenta com o aumento da latitude nos hemisférios norte e sul (VUKUSIC et al. 2007). A heterogeneidade genética entre as populações dos diferentes países, assim como fatores exógenos tais como exposição solar/deficiência de vitamina D, aspectos sócio-econômicos e qualidade dispare dos serviços de saúde pública (que afetam as taxas de diagnóstico) certamente influenciam na diversidade de prevalência reportada para os diferentes países (BULMAN e EBERS 1992).

Um estudo recente realizado na França (VUKUSIC et al. 2007), envolvendo aproximadamente 4 milhões de agricultores diagnosticou a EM em 2667 indivíduos (67/100.000). Ao estratificarem os indivíduos entre as regiões do país o estudo apontou uma prevalência de 100/100.000 no norte e 50/100.000 no sul da França (p-valor <0,01), reforçando o papel das latitudes na distribuição desigual da EM mesmo dentro de um único país.

A *Multiple Sclerosis International Federation* (MSIF) agrupa a distribuição mundial de acordo com 5 categorias de prevalência, sempre considerando a taxa de pacientes a cada 100.000 habitantes: 1) Áreas de alta prevalência: >100 casos; 2) áreas de prevalência média/alta, entre 60,01 a 100 casos; 3) áreas de média prevalência, entre 20,01 e 60 casos; 4) áreas de prevalência média/baixa, entre 5,01 e 20 casos e 5) áreas de baixa prevalência, entre 0 e 5 casos. De acordo com esta distribuição, o Brasil é considerado um país com prevalência média/baixa. Uma figura demonstrando a distribuição da prevalência mundial da esclerose múltipla, de acordo com dados da MSIF é apresentada a seguir (**Figura 1**).



Fonte: Multiple Sclerosis International Federation-MSIF (2015)

Figura 1 – Prevalência de EM no mundo.

Vários estudos com diferentes populações e fundamentados na hipótese de que a latitude seja um possível fator de risco em EM demonstraram, com resultados convergentes, o efeito da migração e da idade na migração na taxa de prevalência de EM. Esses estudos sugerem que, se um indivíduo migrar de uma área de alto risco para uma área de baixo risco com idade igual ou superior a 15 anos, o risco de desenvolver EM será mantido inalterado em relação ao seu local de nascimento. Por outro lado, se um indivíduo migrar de uma área de alto risco para uma região de baixo risco antes de completar 15 anos, a susceptibilidade de desenvolver EM será similar à de um nativo da região de migração (KURTZKE 2000). Essas análises enfatizam que a susceptibilidade para desenvolver EM deve ser resultado da combinação de fatores ambientais (presentes na regiões de risco) e de elementos endógenos (relacionados a componentes genéticos e também relacionados à “idade de susceptibilidade”), indicando a necessidade de um período de exposição para que a doença se estabeleça (KURTZKE et al. 1970; 1998; GALE e MARTYN 1995; KURTZKE 2000; CABRE et al. 2005).

1.3 ASPECTOS DIAGNÓSTICOS E MANIFESTAÇÕES CLÍNICAS

Ainda não existe um teste diagnóstico acurado para diagnosticar a EM em seu início. Em geral, o diagnóstico requer evidências de distúrbios neurológicos propagados no tempo (DIT) e no espaço (DIS) que afetem o sistema nervoso central e que permitam a exclusão de doenças com características similares à EM (tais como a neuromielite óptica (NMO) e doenças do espectro da NMO). Os sintomas/surtos surgem como consequência dos distúrbios neurológicos (POLMAN et al. 2011). SCHUMACKER et al. (1965) foram os primeiros a introduzir o conceito de DIT e DIS. Posteriormente, Poser adicionou a análise do líquido cefalorraquidiano (LCR), para auxiliar na identificação dos sinais da doença no SNC (POSER et al. 1983).

Um primeiro método para avaliar clinicamente o grau de prejuízo neurológico, foi baseado na combinação de um escore de debilidade - *Expanded Disability Status Scale* (EDSS) com a gradação do prejuízo em 8 sistemas funcionais (SFs), a saber: piramidal, cerebelar, tronco cerebral, sensorial, intestinal e urinário, visual, cerebral e outros. Essa gradação pode variar de 0, quando o indivíduo não tem danos neurológicos a 6, quando há prejuízo máximo nos SFs. O EDSS pode variar de 0-10 e depende de quantos sistemas funcionais são prejudicados e o grau de prejuízo desses sistemas. Essa avaliação clínica, para detectar atividade e severidade da EM, pela contagem de surtos e quantificação dos prejuízos neurológicos ainda é aplicada (KURTZKE 1983; RANSOHOFF et al. 2015).

Os critérios para diagnóstico clínico são: i) dois ou mais surtos e evidência clínica de duas ou mais lesões e ii) dois ou mais surtos e evidência clínica de uma lesão. Para demonstrar a DIS são necessárias ao menos duas lesões, por isso pode-se considerar uma lesão na medula espinhal como substituta de uma lesão no cérebro acoplada com análise de LCR.

- ✓ Um surto e duas ou mais lesões. Para demonstrar DIT pode-se fazer exames de ressonância nuclear magnética (RNM), mas deve-se aguardar um intervalo de três meses entre o evento clínico e o aparecimento de uma nova lesão.
- ✓ Um surto e uma lesão. DIT e DIS devem ser demonstrados com RNM e LCR respectivamente. Casos em que RNM e LCR não são feitos, a evidência clínica

de um segundo surto em outro sítio poderá ser utilizado como critério para DIT e DIS.

- ✓ Para diagnóstico de esclerose múltipla primária progressiva (EMPP), devido à ausência de surtos nesse subtipo, deve-se avaliar DIS usando RNM ou potencial evocado visual (PEV) e DIT usando RNM, ou confirmar a progressão de debilidade por um ano e ainda confirmar LCR positivo.

A revisão dos critérios de McDonald em 2010 teve como objetivo diferenciar pacientes pediátricos, asiáticos, latinos e ainda discriminar EM dos espectros da NMO. Paralelamente, a melhoria nas técnicas de ressonância magnética permitiu uma maior sensibilidade na detecção de lesões antes indetectáveis e simplificou os critérios diagnósticos, possibilitando o diagnóstico antecipado com alto grau de sensibilidade e especificidade. Para adequar o critério de RNM em todos os subtipos de EM, um estudo colaborativo europeu (MAGNIMS) analisou RNM em EM e comparou com o critério de RNM utilizado na prática clínica e demonstrou que pelo menos uma lesão em pelo menos duas de quatro regiões consideradas características de EM são suficientes para DIS. Para DIT, o grupo aceitou que o aumento de gadolínio sem aumento da lesão na RNM antes do tratamento pode substituir o seguimento por RNM e finalmente, para diagnóstico de EMPP foi sugerido utilizar o novo critério MAGNIMS, ou seja, pelo menos uma lesão em uma área característica de EM, ou pelo menos duas lesões na medula espinhal, ou ainda LCR positivo confirmado (POLMAN et al. 2011; CECCARELLI et al. 2012; HAGHIKIA et al. 2013).

O diagnóstico da EM é feito pela combinação do exame de RNM com os critérios de McDonald, exceto em localidades onde não há disponibilidade dos aparelhos de RNM. Neste caso, quando também não há alternativas para o exame de imagem, com LCR e PEV, o diagnóstico deve ser reportado como “possível EM” até que um novo evento ocorra para cumprir o critério clínico mínimo de dois surtos e duas lesões (MCDONALD et al. 2001; MILO e MILLER 2014).

Os sintomas mais frequentes em EM são fadiga, distúrbios visuais, e sensoriais, vertigem, depressão ou alteração no humor, alteração no padrão da marcha, redução/perda de força muscular, espasticidade, incontinência urinária, problemas intestinais e comprometimento cognitivo. A contínua atividade e progressão da

doença acumula debilidade ao longo do tempo, de modo que pacientes com uma média de 10 anos de doença provavelmente terão dificuldade para andar (EDSS4) e entre 15 e 20 anos provavelmente deverão usar um suporte unilateral para andar, quando atingem EDSS 6. A expectativa de vida de pacientes com EM é reduzida em uma média de 7 a 10 anos (MILO e MILLER 2014).

Alguns fatores parecem ter relação com um pior prognóstico, tais como o subtipo de EM (EMPP e EMSP), alta taxa de surto e em curto intervalo entre o primeiro e o segundo surto, aumento do EDSS em 2 e 5 anos, curto tempo para a progressão e vários sistemas funcionais envolvidos nesta escala de debilidade (DEGENHARDT et al. 2009).

Em 1996, a Sociedade Nacional de Esclerose Múltipla dos Estados Unidos caracterizou a EM em subtipos clínicos. A grande maioria dos pacientes pode ser classificada em um dos seguintes grupos:

- a. Esclerose Múltipla Remitente Recorrente (EMRR): afeta cerca de 85% dos indivíduos. Caracteriza-se por uma fase definida por surtos novos que ocorrem em média uma vez a cada biênio e recorrentes sintomas neurológicos que não progridem entre os surtos e ainda apresentam sinal de recuperação total ou parcial;
- b. Doença Secundária Progressiva (SP): Após 25 anos de doença aproximadamente 90% dos indivíduos com EMRR evoluem para a fase SP. A EMSP é diagnosticada com base na história natural da doença, em um contexto de piora gradual, associada a progressão com surtos ocasionais e poucas remissões, após o curso da doença inicial (EMRR). A resposta à imunoterapia nessa fase ocorre apenas quando há eventos inflamatórios residuais, ou surtos ou lesões identificadas por RNM. Atualmente, ainda não existe uma ferramenta diagnóstica que determine o ponto de transição entre os fenótipos EMRR e EMSP, mas sabe-se que essa transição é geralmente gradual (LUBLIN e REINGOLD 1996).
- c. Doença Primária Progressiva (PP): Aproximadamente 10% dos pacientes iniciam a EM em uma fase de progressão desde o início da doença e podem apresentar eventualmente pequenas melhorias, pouca ou nenhuma característica inflamatória e ausência de surtos anteriores à progressão, essa

fase é definida como EM primária progressiva (EMPP). Assim como EMSP, a EMPP também é diagnosticada pela história natural da doença e a evolução da progressão em ambas são parecidas. A ausência de surto na fase PP pode ser atribuída a uma lesão individual localizada em uma região silenciosa que culmina na debilidade (LUBLIN et al. 2014; RANSOHOFF et al. 2015).

Em 2011, os critérios clínicos postulados em 1996 foram revisados e foram sugeridas modificações que contemplassem a atividade e progressão da doença com uma reavaliação da atividade da doença (critérios clínicos e de imagem), pelo menos uma vez ao ano. Um dos aspectos explorados foi a inclusão da CIS como uma primeira manifestação clínica da doença e que exerce o critério de DIT. Outro aspecto explorado foi a incorporação da atividade e progressão da doença, como aspectos modificadores dos fenótipos básicos de EM. Por exemplo, pacientes com EMRR que apresentam novas lesões por RNM são classificados como EMRR ativos. Por outro lado, pacientes que não apresentam novas lesões ou cujas lesões parecem estabilizadas e não apresentam crescimento, são classificados como EMRR inativos (LUBLIN et al. 2014).

1.4 ESCLEROSE MÚLTIPLA E NEURODEGENERAÇÃO

Acredita-se que a variabilidade de apresentações clínicas da EM acontece pela manifestação de diferentes fases de um mesmo processo patológico. Como exemplificado pela inflamação aguda, que recruta as células progenitoras oligodendrocíticas (CPOs), importantes no processo de remielinização. Entretanto, embora pareça contraditório, artigos demonstram que uma redução na resposta inflamatória também pode limitar o processo regenerativo e levar ao acúmulo de danos axonais e neuronais (LUCCHINETTI et al. 2000; FRANKLIN 2002).

A mielina, substância encontrada ao redor dos prolongamentos neuronais (chamados axônios), tem uma composição eminentemente lipídica, com grande capacidade de isolamento elétrico. Esta propriedade isolante faz da mielina um elemento fundamental na facilitação de uma rápida condução dos impulsos elétricos que se propagam pelos axônios, e também impede que estes deixem os prolongamentos neuronais. Como já discutido, na EM, processos inflamatórios

autoimunes danificam a bainha de mielina, dificultando e retardando a transmissão dos impulsos nervosos e levando ao desenvolvimento de sintomas diversos, que variam de acordo com a função, intensidade das lesões e a localização dos neurônios afetados. Deste modo, além dos aspectos clássicos de uma doença autoimune, causados pela atividade inflamatória aumentada, a desmielinização, os danos axonais, a menor capacidade de neuroregeneração e as contínuas perdas neuronais contribuem em conjunto para o estabelecimento e a progressão da doença (COMPSTON e COLES 2002).

Os processos inflamatórios e imunes estão importantemente ligados à patogênese da EM e envolvem a associação linfócitos e microgliócitos. A micróglia representa os principais fagócitos do sistema nervoso central (SNC), sendo constituída por uma variedade de macrófagos. A micróglia participa de todas as fases da EM através de seus receptores *Toll-like*, que interagem com moléculas endógenas e exógenas ao SNC. Estas interações modulam a apresentação de antígenos por estas células e contribuem para a iniciação e propagação de diversas outras respostas imunológicas, incluindo a inflamação e fagocitose, que culmina com a destruição da bainha de mielina. A fagocitose feita pela micróglia e outros tipos de macrófagos é uma das principais marcas da lesão da EM, que também ocorre graças à bateria de moléculas efetoras carregadas por estas células, tais como espécies reativas de nitrogênio, que também causam importantes danos aos oligodendrócitos (JACK et al. 2005). Um dos maiores desafios terapêuticos na EM tem sido a modulação da resposta da micróglia, de modo a limitar os danos causados ao SNC e permitir uma recuperação local. No entanto, além desta atividade neurotóxica na EM, a micróglia também possui grande importância na neuroproteção, através de mecanismos que envolvem a secreção de mediadores solúveis que levam ao reparo neuronal e regeneração, inclusive promovendo a remoção de fragmentos de células mortas e a indução de neurogênese (revisto em NAPOLI e NEUMANN 2010).

Recentes estudos têm demonstrado que a substância cinza cerebral pode ter um importante papel relacionado ao acúmulo de dano irreversível e progressiva debilidade, e ainda que os diferentes processos neurodegenerativos na substância branca e os danos acumulados na substância cinza sejam capazes de determinar a

heterogeneidade observada em pacientes com EM (LASSMANN 2012; SCHUTZER et al. 2013).

As primeiras lesões ocorrem em áreas focais, conhecidas como placas. A localização dessas placas implica em manifestações clínico-neurológicas variáveis, que de modo geral, é resultado da invasão de células imunes através da barreira hematoencefálica num processo contínuo de ativação de macrófagos (células migratórias do sistema imune inato) e micróglia (células residentes do sistema imune inato), com consequente neurodegeneração e desmielinização. Embora a EM seja considerada uma doença autoimune, onde as células T atacam seus próprios antígenos em indivíduos predispostos, o mecanismo de desmielinização, lesões inflamatórias e mudanças neuronais que melhor correlaciona com debilidade clínica, ainda é pouco entendido (WEINER 2008).

1.5 TERAPIAS UTILIZADAS NA EM

A partir da década de 90 a imunoterapia com interferon β (IFN β 1b) e acetato de glatirâmer se tornaram padrão para tratamento dos surtos e redução da taxa de progressão anual em pacientes com EMRR. Posteriormente, em 2003 surgiu o Natalizumabe, um anticorpo monoclonal humanizado que reconhece a molécula de adesão α 4-integrina. O Natalizumabe foi a primeira a demonstrar a possibilidade de alcançar remissão, com a ausência de sinais de EM detectável por RNM e com uma taxa de redução de surto em dois anos de tratamento de mais de 65% dos casos, além de mais de 90% de redução de novas lesões inflamatórias (POLMAN et al. 2006). No entanto, houve relatos do uso de Natalizumabe levando a uma leucoencefalopatia multifocal progressiva (POLMAN et al. 2006). Já em 2009 foi aprovada a droga oral Fingolimode, um imunomodulador que parece agir retendo os linfócitos no interior dos linfonodos e impedindo que estes se desloquem para o SNC onde promovem a inflamação e as reações autoimunes. Deste modo esta droga permite remielinização e inibe a astrogliose. Essa droga apresenta uma taxa de redução de surtos entre 55-60% e uma excelente redução de lesão ativa, mas pode acarretar complicações cardíacas em indivíduos com mais de 50 anos de idade (AKTAS et al. 2010; KAPPOS et al. 2010).

Os diversos mecanismos de ação dessas drogas, as diferentes vias e a relação risco/benefício, fazem com que o acompanhamento da evolução da doença, análise de fatores prognósticos, análise de RNM e histórico de tratamentos anteriores sejam a base para a escolha do tratamento mais adequado. No entanto, os tratamentos usados na EM têm efeito mínimo sobre a síndrome clinicamente isolada e nenhum efeito sobre taxa de declínio após o início da fase progressiva da doença (KAPPOS et al. 2009; HAGHIKIA et al. 2013).

1.6 ETIOLOGIA

Apesar de não se saber a causa da EM, fatores imunológicos, ambientais e genéticos têm sido relacionados à sua etiologia, conforme discutido a seguir.

1.6.1 Autoimunidade e EM

A hipótese da auto-imunidade como um componente importante na etiologia da EM é suportada por centenas de estudos de diversos grupos e tem como base a presença de infiltrados de células T ativas no SNC e a capacidade de modulação dos sintomas com tratamentos imunoterápicos (MARTIN et al. 1992). Durante o estabelecimento de doenças autoimunes, acredita-se que ocorra uma reativação de clones de células T auto-reativas que não tenham sido eliminados durante a maturação do timo, em um processo desencadeado pela exposição a algum antígeno que mimetize proteínas próprias do organismo hospedeiro. No caso da EM, uma das hipóteses é a de que a apresentação de antígenos que ocasione uma reação cruzada com proteínas associadas a processos relacionados à mielina que reveste os neurônios do SNC, iniciando assim uma reação imune, principalmente baseada em células T, contra esta proteína (GUGGENMOS et al. 2004).

Existem várias hipóteses para o desenvolvimento da autoimunidade em EM, uma delas é a de que nos pacientes algumas proteínas corporais que não são expressas no timo, como a glicoproteína oligodendrocítica da mielina (MOG) possivelmente desestabilizaria a manutenção da auto-tolerância, que poderia atuar por deleção de células T auto-reativas ou mesmo pela seleção de células T

regulatórias, podendo contribuir para o desenvolvimento de doenças autoimunes órgão-específicas (BRUNO et al. 2002).

1.6.2 Fatores Ambientais Associados à Esclerose Múltipla

Os potenciais fatores ambientais de risco mais consistentemente relacionados ao desenvolvimento da EM são: infecções, uso de tabaco e déficit de exposição solar causando carência de vitamina D (GIOVANNONI e EBERS 2007; PANTAZOU et al. 2015). Cada um destes fatores é discutido a seguir.

1.6.2.1 Agentes Infecciosos

O conceito de que agentes infecciosos podem estar envolvidos com a patogênese da EM surgiu inicialmente com o neurologista Pierre Marrie, sucessor do neurologista Jean Martin Charcot no Hospital Salpêtrière, em Paris (MILO e MILLER 2014). Neste contexto, um dos microorganismos primeiramente associados foi o vírus de Epstein-Barr (EBV).

Alguns achados sugerem que o risco de desenvolver EM aumenta em indivíduos infectados com EBV. No entanto, diante da alta prevalência de EBV na população geral (atingindo aproximadamente 95%) (POHL et al. 1996) e da pequena diferença de positividade entre grupos com ou sem EM (2-4% de diferença) (LEVIN et al. 2010; BELBASIS et al. 2015) torna-se difícil estabelecer um claro papel deste vírus no desenvolvimento da EM.

Todos os pacientes com EM tem anticorpos contra EBV. No entanto não se sabe se a infecção por EBV é causa ou consequência da EM (ASCHERIO et al. 2010). Um estudo com 62.439 mulheres encontrou um aumento substancial no título sérico anti-EBV antes do início da EM. Surpreendentemente, a maioria das mulheres soropositivas para EBV tiveram um início tardio da EM (ASCHERIO et al. 2001). A capacidade do EBV de causar danos neurológicos sérios em varias partes do sistema nervoso central já foi documentada, mas ainda não existem estudos que demonstraram a associação do EBV com a causa ou gravidade da EM (MAJID et al. 2002).

Devido à capacidade dos retrovírus em causar imunodeficiência e diversas doenças neurológicas, como a paraparesia espástica tropical e leucoencefalopatia

multifocal progressiva, alguns estudos tentam entender os mecanismos de ação dos retrovírus endógenos humano (REH). Estudos demonstraram uma associação entre infecções virais e doenças neurológicas crônicas, como por exemplo no caso de pacientes com panencefalite esclerosante subaguda, que apresentaram altas concentrações de anticorpos contra o vírus do sarampo no plasma e no LCR (CONNOLLY et al. 1967). A leucoencefalopatia multifocal progressiva (LEMP) é outra doença que está associada com uma infecção viral, no caso com o poliomavírus JC humano, que foi encontrado em oligodendrócitos de pacientes portadores desta doença. Interessantemente, uma das características da LEMP é a desmielinização progressiva do sistema nervoso central, com destruição de oligodendrócitos (PADGETT et al. 1971).

Algumas linhas de pesquisa indicam que a EM também possa ser uma imunopatologia associada a alguma infecção viral que participe do componente autoimune da doença (MARRIE 2004). Essa premissa é corroborada por vários achados epidemiológicos que indicam que a EM seja uma doença autoimune adquirida e que a herança genética não é preponderante para o desenvolvimento da mesma.

Mais de 90% dos pacientes com EM possuem altas concentrações de IgG no LCR e no plasma, o que sugere fortemente uma infecção relacionada com a doença, assim como descrito para todas as outras doenças com essas mesmas características (GILDEN et al. 1996). Nestas outras doenças, os altos títulos de IgG correspondiam a anticorpos específicos contra o agente causador da mesma, como no caso da panencefalite esclerosante subaguda, onde as IgGs oligoclonais representavam anticorpos contra o vírus do sarampo (VANDVIK et al. 1976). Já na EM, o alvo destas IgGs ainda não foi caracterizado (GILDEN 2005).

Existem hipóteses de que mais de um agente infeccioso poderia causar a EM, pois as características clínicas dos pacientes podem ser muito distintas, inclusive no que diz respeito a graus de desmielinização e inflamação. No entanto, em várias outras doenças, o mesmo agente infeccioso pode causar uma sintomatologia diversa, como no caso da neuro-sífilis, causada pelo *Treponema pallidum*. Assim, a genética do indivíduo, idade, estado nutricional, *background* genético, imunidade, entre outros aspectos podem contribuir para a variabilidade de sintomas e gravidade da

doença. Desta forma, a hipótese de infecção por apenas um organismo como causa da EM ainda é a mais plausível (GILDEN 2005). A baixa abundância de um microorganismo putativo causador da EM e sua presença em um estado de latência pode explicar por que o mesmo ainda não tenha sido descoberto. Um dos exemplos que justificaria essa hipótese é a presença do vírus da *Varicella zoster*, que se encontra latente em gânglios, com baixa carga viral e baixa expressão gênica (GILDEN 2005).

Hipóteses sobre vários organismos que podem ter alguma relação com a EM foram levantadas, como a *Chlamydia pneumoniae* (SRIRAM et al. 1999). Pacientes com EM e controles tiveram seu soro e LCR testados para esta bactéria gram-negativa e ambos os grupos foram positivos, porém o título de anticorpos contra esta bactéria no LCR se mostrou mais elevado nos pacientes com EM. No entanto, características clínicas da doença (como duração da doença, atividade da doença, etc) não foram correlacionadas com os altos títulos de anticorpo contra *C. pneumoniae* (KRAMETTER et al. 2001).

Dois vírus com possível associação com a EM vem sendo amplamente estudados, o herpesvírus humano 6 (HHV-6) e o vírus Epstein-Barr (EBV). Estes vírus ficam latentes em células-B (EBV) e células-T (HHV-6), e para ambos a soroconversão acontece logo antes ou depois da puberdade, concordantes com evidências do tempo de exposição a eventuais agentes causadores da EM (GILDEN 2005). Anticorpos contra esses vírus são comumente encontrados na população, no entanto é possível que exista uma ineficiência na resposta imune de pacientes com EM, que pode levar à reativação da infecção latente por HHV-6 (BLUMBERG et al. 2000). Os títulos virais de HHV-6 podem variar durante as fases de remissão e recorrência da doença, sendo que os títulos virais estavam mais elevados em pacientes na fase recorrente ($p=0,007$). Como a reativação viral pode ser baixa, é possível que ela não seja passível de detecção por PCR. Neste sentido a detecção viral é maior em sangue total do que em soro, e alguns pesquisadores sugerem que o HHV-6 possa infectar outras células que não os linfócitos, como os oligodendrócitos (CASELLI et al. 2002).

Alguns trabalhos mostram que a ativação de retrovírus humanos endógenos (HERVs) pode estar relacionada com a EM (CERMELLI e JACOBSON 2000).

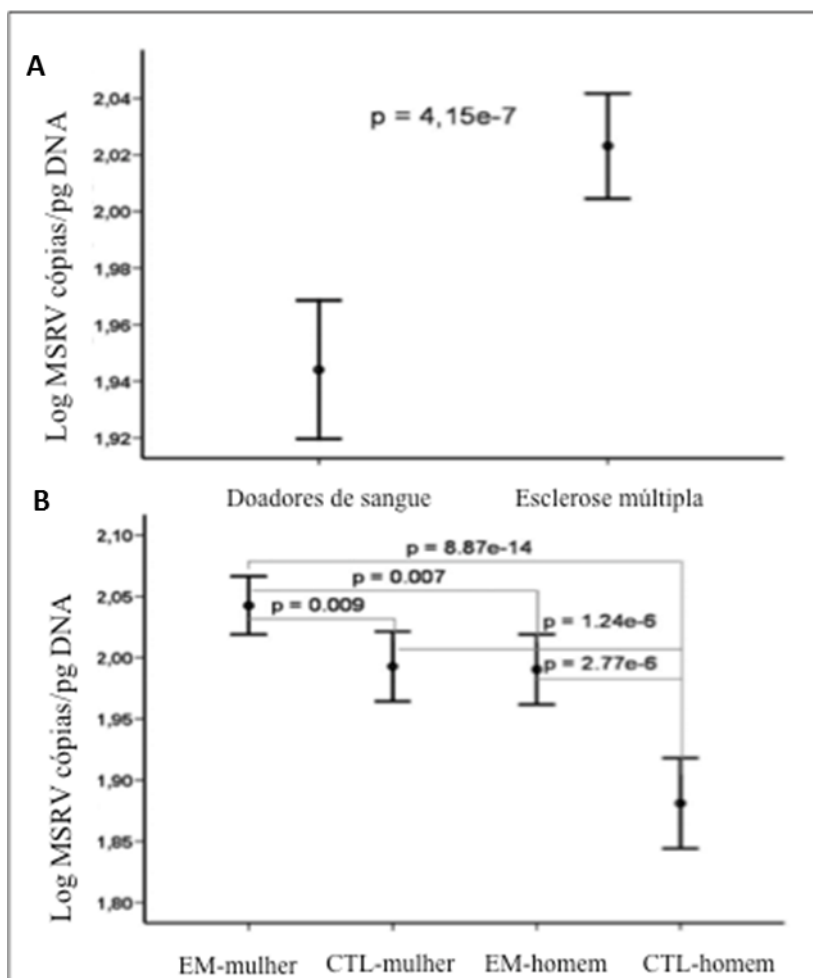
Retrovírus que infectaram humanos há muitos milhões de anos atrás foram incorporados e fazem parte do genoma humano (GIFFORD e TRISTEM 2003). Estas sequências geralmente são inativas, mas em determinadas circunstâncias elas podem ser traduzidas. Existem evidências de que os HERVs produzam proteínas neurotóxicas associadas com a progressão da EM e alguns laboratórios já isolaram proteínas de HERVs em lesões de pacientes com EM (ANTONY et al. 2011). Interessantemente os herpes vírus parecem ser capazes de ativar HERVs, e neste sentido EBV parece ser particularmente ativo nesta condição. O fato de EBV ser altamente prevalente em pacientes com EM torna a hipótese de correlação entre EM e HERVs ainda mais interessante (MAMELI et al. 2013). A expressão de HERVs está relacionada a ativação de ambos os sistemas imunes inatos e adaptativos, o que pode ser uma ligação a mais com as características inflamatórias e imunológicas de EM. Outro dado que corrobora a hipótese do envolvimento dos HERVs com a EM vem do fato de pacientes soropositivos para HIV e tratados com anti-retrovirais terem uma baixa incidência de EM. Somado a isso foi descrito um interessante relato de caso de um paciente portador de EM que se tornou HIV soropositivo e, após iniciar o tratamento contra o HIV com terapia anti-retroviral, teve uma melhora radical nos sintomas da EM (MARUSZAK et al. 2011). Em geral, pacientes com EM possuem altos níveis de expressão de HERVs, quando comparados a controles ou a pessoas que possuam outra doença neurológica. Neste sentido, foi iniciado em 2013 um ensaio clínico (número NCT01767701; www.clinicaltrials.org) no Reino Unido chamado INSPIRE, que avalia o efeito do anti-retroviral Raltegravir (Isentress, MerckSharp & Dohme Corp) em pacientes com EM. Esta droga também pode ser ativa contra os herpes vírus.

Os genes dos retrovírus HERV-W podem estar presentes em varias cópias no genoma, sendo que a maioria delas é truncada, ou não possui uma janela aberta de leitura (ORF). No entanto algumas ORFs permanecem, indicando o potencial de alguns destes retrovírus para produzirem proteínas completas ou truncadas. Estes genes estão dispersos em várias cópias em cromossomos distintos e podem ser expressos em tecidos específicos, durante certas fases do desenvolvimento (LÖWER et al. 1993, 1996). Um destes retrovírus, denominado *Multiple Sclerosis Associated Virus* (MSAV) é um membro da família HERV-W e foi associado à EM quando foi

originalmente isolado de células leptomenígeas presentes no LCR de um paciente com esta doença (PERRON et al. 1989). Sua presença foi detectada em vários pacientes com EM por vários grupos de pesquisa, tanto no sangue, quanto no LCR, sendo que sua detecção em pessoas sem EM é significativamente menor (DOLEI e PERRON 2009). Um estudo posterior de DOLEI et al. avaliou a presença de MSAV em uma população com alto risco de desenvolvimento de EM e observou a sua forma extracelular no plasma de 100% dos pacientes com doença ativa. A presença de MSAV também foi observada no LCR de pacientes com doença em evolução (DOLEI et al. 2002). Nos pacientes com doença estável não foi detectado MSAV no LCR ao contrário do que ocorre nos pacientes com doença mais severa (SOTGIU et al. 2002). MSAVs levam células gliais à apoptose, mas o mesmo efeito não ocorre em neurônios (MÉNARD et al. 1997). Células mononucleares de sangue periférico (PBMC) derivadas de indivíduos infectados por MSAV continuam a produzir vírions quando cultivados *in vitro*. No entanto esse mesmo achado não foi observado em PBMC de indivíduos não infectados por MSAV, nem espontaneamente, nem sob tratamento com agentes indutores, sugerindo que a expressão deste retrovírus não é ubíqua e que eles poderiam ser um membro exógeno dos HERV-W (SERRA et al. 2003). Como no genoma humano não foram identificadas cópias intactas de MSAV, uma hipótese para a produção de partículas virais viáveis endógenas poderia envolver trans-complementação de cópias dos genes necessários. Assim é possível, como descrito para o vírus MMTV (vírus tumoral mamário de camundongo) (BERLIOZ e DARLIX 1995), ou o vírus MLV (vírus da leucemia murina) (XU et al. 1996), que eles possam ter se formado a partir de uma cópia endógena não-ubíqua, por retrotransposição *de novo* ou recombinação (LINIAL e MILLER 1990).

Já a Sincitina-1 é uma proteína viral codificada pelo gene *env* de um outro HERV-W, o ERVWE-1, incompetente para replicação. Interessantemente, a Sincitina-1 possui uma identidade de 94% com o EMRV. Este gene está integrado no genoma humano no locus 7q21-22, região candidata à susceptibilidade para EM. Enquanto o EMRV é capaz de formar virions viáveis extracelulares com atividade de transcriptase reversa, a Sincitina-1 só foi encontrada intracelularmente ou na membrana plasmática. A Sincitina-1 é altamente expressa na glia de pacientes com EM e pode estar associada a alguns sintomas da EM (DOLEI 2005).

A maior prevalência da EM em mulheres já foi bem descrita na literatura (SADOVNICK 2009), e neste sentido a presença de cópias de HERV-W no cromossomo X poderia estar relacionada com este fenômeno. Em um trabalho recente de GARCIA-MONTOJO 2013, a estratificação de controles e portadores de EM e também a avaliação por sexo, mostrou que a quantidade de DNA do gene do envelope de MSAV (*Msav-env*) varia de acordo com o desenvolvimento da doença e com o sexo. Este trabalho mostra um aumento de DNA genômico em mulheres com EM em relação a mulheres controle. O mesmo achado se manteve na comparação entre homens com e sem EM e quando o sexo foi considerado entre pacientes com EM, o número de cópias de DNA de *Msav-env* foi maior em mulheres (**Figura 2**).



Fonte: GARCIA-MONTOJO et al. (2013).

Figura 2 – Quantidade de DNA do EMRV entre controles (doadores de sangue) e pacientes com EM. A) A quantidade de DNA do EMRV (avaliada pelo número de cópias por picograma de DNA) é maior nos pacientes com EM (N=178) quando comparada aos controles (N=124) (ANOVA; $p=4.15e-7$). B) Comparação da quantidade de EMRV entre homens e mulheres, portadores de EM ou controles (CTL). Os pontos indicam a média e as barras representam o intervalo de 95% de confiança.

Em 2000, CHRISTENSEN et al. identificaram a presença de partículas virais dos retrovírus RGH, que também fazem parte da mesma subseção dos oncovirinae do tipo C assim como os retrovírus HERV-W/EMRV, em pacientes com EM. Os RGH são variantes homólogas aos HERV-H e se mostraram presentes no sobrenadantes de 4 linhagens celulares de EM e também em 73% das amostras de plasma de 33 pacientes com EM. Interessantemente, todos os pacientes positivos para a presença

de RGH possuíam doença ativa. Desta forma, é possível que os retrovírus das HER-W e HER-H possam ter papéis similares em sua associação com a EM.

De modo geral, REs podem induzir anormalidades no genoma ou em produtos gênicos provocando alteração da sequência genômica normal com consequências funcionais. Paralelamente, os REVs fornecem elementos regulatórios que podem atuar em processos de ativação/inibição gênica e *splicing alternativo*. A *transcriptase* reversa de um REH pode ainda interagir com RNAs virais exógenos (ex. EBV) (MAGER 1999; LING et al. 2001; CUSICK et al. 2013).

Os REs são vírus patogênicos em doenças neurológicas mediadas por componentes imunológicos, como a EM. Explorar esses possíveis agentes etiológicos se faz necessário para um melhor entendimento das bases biológicas da EM.

1.6.2.2 Exposição solar e vitamina D

A exposição solar, associada à latitude parece ser um potencial fator relacionado à EM. Com o objetivo de fazer uma correlação direta entre exposição à radiação solar e a mortalidade por EM, foi desenvolvida uma investigação exploratória tipo caso-controle baseada em atestados de óbito emitidos nos EUA, levando em consideração características importantes que poderiam enviesar a análise, como idade, sexo, características sócio-econômicas e região da residência. Para isso, foram selecionados indivíduos que viviam em ambientes similares, mas expostos a diferentes níveis de radiação UV. Através da análise multivariada de regressão logística, os indivíduos que trabalhavam em ambiente de alta exposição à radiação solar foram negativamente associados à morte por EM (odds ratio (OR)=0,74), o mesmo ocorreu para os indivíduos que residiam em estados de alta exposição solar (OR=0,53) (FREEDMAN et al. 2000).

Os estudos de prevalência de EM em altas latitudes, levaram a hipótese de que a EM poderia estar associada a deficiência da vitamina D devido à baixa incidência de radiação UVB (ASCHERIO et al. 2010). A vitamina D (colecalfiferol) é um pré-hormônio sintetizado na pele como consequência da exposição à radiação UVB. No fígado e nos rins a vitamina D é processada para a forma ativa 1,25(OH)₂D₃ (chamada de calcitriol), importante substância imunomoduladora. A exposição excessiva à radiação UVB e consequente aumento na produção de

vitamina D não causa intoxicação, pois os metabólitos da vitamina D são convertidos em produtos inertes (HOLLIS 2005). No entanto, a diminuição da síntese cutânea de vitamina D pode ter consequências danosas, podendo ocorrer por diversos fatores, incluindo idade mais avançada e presença de maior quantidade de melanina na pele. A maior fonte de obtenção da vitamina D é através da radiação UVB, mas alguns alimentos, como óleo de fígado de bacalhau e óleo de peixe também são fontes de vitamina D (ASCHERIO et al. 2010; HAUSER et al. 2013).

O receptor da vitamina D (VDR) é expresso em células do sistema imune, incluindo células imunes inatas e adaptativas e tem um importante papel na modulação de processos imunes através da regulação transcricional ativada por seu ligante, o $1,25(\text{OH})_2\text{D}_3$ (VANAMERONGEN et al. 2004). Esses processos são principalmente relacionados à inibição da expansão das células T-helper 1 (envolvidas na patogênese de doenças auto-imune), através do bloqueio da produção da interleucina 12 (IL-12) derivada de macrófagos e células dendríticas, da inibição da maturação das células dendríticas, aumento da atividade de células T-regulatórias, aumento da capacidade fagocítica pelos macrófagos e monócitos entre outros. Esse receptor é encontrado, por exemplo, em células T e B ativas no cérebro e atua como fator de transcrição em genes responsivos ao seu metabólito ativo [$1,25 (\text{OH})_2 \text{D}_3$] (LEMIRE 1995; D'AMBROSIO et al. 1998; PONSONBY et al. 2005).

A transcrição dos genes responsivos ao $1,25 (\text{OH})_2 \text{D}_3$ é modulada pela heterodimerização do VDR ativo com o receptor nuclear retinóide X (VDR/RXR) que se liga a uma sequência de DNA na região promotora do gene responsivo à $1,25 (\text{OH})_2 \text{D}_3$, denominada de elemento responsivo a vitamina D (VDRE), regulada pelo calcitriol (VANAMERONGEN et al. 2004).

A síntese de vitamina D, que tem uma relação direta com exposição à luz solar, tem sido alvo de inúmeras pesquisas que buscam sua correlação com a patogênese da EM. O provável mecanismo de ação da vitamina D é o seu metabólito ativo, o calcitriol, e seu receptor nuclear (VDR) (O'GORMAN et al. 2012). Alguns dos genes responsivos ao calcitriol têm um importante papel em processos imunológicos relacionados às células T, como por exemplo, o fator nuclear de células T ativadas (*NFAT*), que tem um papel central na cascata de sinalização das células T e MAP quinase-fosfatase-5 (*MKP5*) que regula a resposta imune inata e

adaptativa (SHAW et al. 1988; ZHANG et al. 2004; WÖBKE et al. 2014). Assim, acredita-se que os efeitos imunomodulatórios da vitamina D tenham um papel na gênese da EM por mecanismos ainda não totalmente conhecidos.

1.6.2.3 Consumo de tabaco

Fortes indícios associam o consumo de tabaco (incluindo quantidade e tempo de consumo) ao maior risco de desenvolver a EM e também a uma redução no tempo de conversão da fase EMRR para a fase SP (WESTON e CONSTANTINESCU 2015).

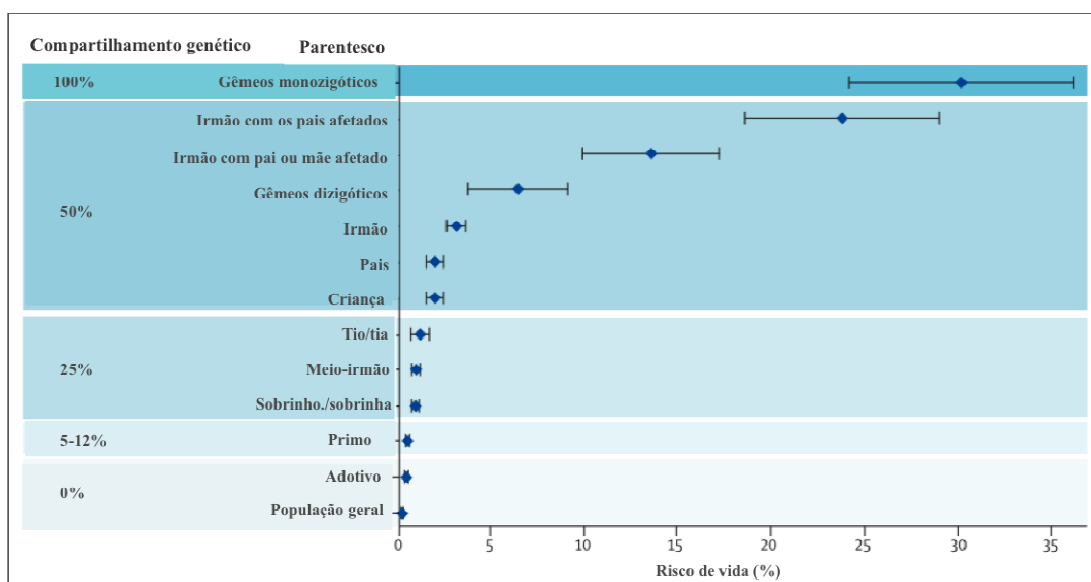
É sabido que os xenobióticos, como por exemplo os numerosos componentes da fumaça do cigarro, são fortes ligantes ao receptor aryl hidrocarboneto (AhR). O AhR é um fator de transcrição expresso em níveis elevados em células T-helper 17 (Th17) vinculadas a doenças autoimunes, como a EM. A ativação do AhR por esses xenobióticos induz dois processos importantes na autoimunidade. A via mais comum é a da regulação da expansão de células Th17 e subsequente impacto na produção da interleucina 22 (IL-22), importante em diversas doenças autoimunes e também na EM (revisto em PAN et al. 2013). Como as células Th17 são essenciais para o desenvolvimento de algumas doenças autoimunes, acredita-se que a via de ativação do receptor AhR, possa agravar esta condição (WESTON e CONSTANTINESCU 2015).

Apesar de estudos epidemiológicos terem sugerido fatores diversos, como os citados acima, que podem ter importância no desenvolvimento da EM, ainda não existe um pleno consenso quanto às causas mais relevantes para o desenvolvimento da EM. De modo importante, ainda não existem evidências científicas suficientes para comprovar que os fatores até hoje sugeridos sejam essenciais na etiologia da EM (MARRIE 2004; ASCHERIO e MUNGER 2007; ASCHERIO et al. 2010; KLEINewietfeld et al. 2013; BELBASIS et al. 2015).

1.6.3 Fatores Genéticos Associados à Esclerose Múltipla

A importância de fatores genéticos na gênese da EM é inequívoca. Uma prova disto é o fato de, como comprovado por estudos epidemiológicos diversos, quanto maior o grau de compartilhamento genético com um portador da doença,

maior a probabilidade de que um segundo indivíduo também desenvolva a doença. Enquanto a prevalência da EM na população que vive em áreas alto risco é de 0,1%, a taxa de recorrência familiar média é 200X maior, ou 20% [revisto em (EBERS et al. 1995)]. Estudos com famílias e com gêmeos têm demonstrado que a genética tem um papel preponderante na gênese da EM (**Figura 3**).



Fonte: Adaptado de COMPSTON e COLES (2002).

Figura 3 – O risco de ocorrência de esclerose múltipla é diretamente proporcional à proximidade genética com uma pessoa afetada

Uma das hipóteses mais estudadas pelos grupos que investigam a EM é apoiada pelo conceito de que combinações de certas alterações genéticas predisponentes, encontradas nos genomas de indivíduos saudáveis, quando transmitidas à geração seguinte, levariam ao desenvolvimento da doença. Neste modelo, um arranjo ao acaso das alterações genéticas existentes no genoma dos pais é transmitida aos filhos e estas, em conjunto com fatores ambientais diversos, seriam responsáveis por desencadear a doença. No entanto, cabe notar que a herança genética na EM não é absoluta. Por exemplo, quando um indivíduo é afetado pela EM, as chances do seu gêmeo monozigótico – e, portanto com um genoma teoricamente idêntico ao do irmão afetado – desenvolver a doença é de aproximadamente 35%. Se por um lado este risco é extremamente elevado quando comparado com a prevalência de EM na população em geral, este dado demonstra

que a identidade genética, no que se refere aos alelos herdados dos pais, não é capaz de explicar totalmente a gênese da EM (ZUVICH et al. 2009).

Uma hipótese a ser considerada sobre os aspectos genéticos da EM é de que múltiplas variantes em diversos genes podem fornecer contribuições diferentes e aditivas no que se refere à predisposição ao desenvolvimento da doença. É possível que alterações em vias biológicas específicas causem distintos níveis de predisposição relacionados a fatores biológicos relevantes para a doença tais como capacidade de síntese e metabolização de vitamina-D, predisposição a infecções virais, capacidade e intensidade de resposta imunológica, predisposição aumentada a fatores que levem a neurodegeneração ou que impeçam um reparo a danos neuronais. Tais alterações, quando avaliadas em conjunto com fatores exógenos, certamente podem contribuir muito para a maior compreensão da biologia da EM.

Devido ao difícil acesso ao tecido afetado pela EM, o uso de modelos animais de EM com encefalomyelite autoimune experimental (EAE) tem sido usado na tentativa de desvendar as vias neurodegenerativas associadas à EM, buscando esclarecer os mecanismos neuroaxonais de injúria ou proteção durante a inflamação crônica (FRIESE et al. 2006). O dano neuroaxonal é mediado por uma série de produtos neurotóxicos, como espécies reativas de oxigênio, glutamato, citocinas e quimiocinas secretados pelas células imunes. Esses produtos são importantes no processo de inibição da resposta inflamatória e para a dissolução de um quadro inflamatório já estabelecido, direcionam a resposta imune e alteram o metabolismo de neurônios e axônios. No entanto, a manutenção desses produtos de células imunes por um período prolongado é uma fonte intrínseca de estresse (FRIESE et al. 2014).

Evidências apontam que no citoplasma de neurônios derivados de lesões corticais ativas de EM há acúmulo de moléculas oxidadas e moléculas de DNA quebradas. As espécies reativas de oxigênio (ROS) e de nitrogênio (RNS), como o óxido nítrico (NO) estão intrinsecamente relacionadas a lesões em EM e EAE e seus produtos oxidativos estão envolvidos no processo inflamatório, sugerindo que macrófagos e micróglia estejam envolvidos no curso da neurodegeneração (HAIDER et al. 2011; NIKIĆ et al. 2011; FISCHER et al. 2013). Entretanto, ainda não está claro se vias antioxidativas tem um papel na proteção de neurônios contra o estresse oxidativo, mediado pelo processo inflamatório.

As variantes de nucleotídeo único (SNPs), geralmente investigadas por estudos de *Genome Wide Association Studies* (GWAS) são importantes candidatos a fatores de risco em diversas doenças. Na doença de Parkinson, por exemplo, SNPs no gene codificador do fator neurotrófico da dopamina cerebral (CDNF) são considerados fatores de risco para a doença (Choi et al. 2011). Em EM, estudos de GWAS não medem esforços na tentativa de identificar fatores de susceptibilidade. Vários SNPs já foram associados à EM e o principal *locus* é o do MHC, no chr 6p21.3 (BARANZINI et al. 2009).

1.7 ANÁLISES DE EXOMAS COMO FERRAMENTA DE AVALIAÇÃO DE DOENÇAS NEUROLÓGICAS

Ao longo das décadas, houve grandes avanços no entendimento da EM, de modo que é possível delinear vias moleculares relevantes para a doença, além de sugerir alelos de risco e fatores ambientais envolvidos. Muitas terapias para tratar os sintomas (surtos) foram desenvolvidas como resultados de uma importante interação entre equipes multidisciplinares.

O desafio agora é compreender melhor os aspectos da doença, como vias de reparo das lesões e fatores de proteção neuronais que possam ser utilizados em terapias alvo, para tratar a fase progressiva da EM e possivelmente direcionar o tratamento.

Decifrar o conjunto de alterações genéticas de doenças neurológicas é um fator crucial para o entendimento de seus aspectos patológicos. As metodologias de nova geração para o sequenciamento completo do exoma (WES) têm sido bem sucedidas em identificar novos genes em doenças complexas. O poder da ferramenta está no dado produzido, onde é possível obter uma detalhada investigação das regiões codificadoras de proteínas.

O sequenciamento de exomas tem sido usado no estudo de diversas doenças neurodegenerativas. Em um estudo com variantes raras em pacientes com esclerose lateral amiotrófica familiar (FALS), foi identificada uma mutação no gene *TUBA4A*, codificador de proteína envolvida na estabilização de microtúbulos (SMITH et al. 2014). Em um estudo caso-controle, o grupo do GOLDSTEIN (2015) demonstrou

através do sequenciamento do exoma que dentre diversos genes de associação com ALS conhecidos, uma quinase não-canônica da família *Tank-Binding Kinase 1* (*TBK1*) foi identificada como um gene associado à doença (CIRULLI et al. 2015).

Na doença de Alzheimer, através do sequenciamento do genoma (WGS) e exoma foram identificados mais de 20 *loci* associados como fatores de risco (KARCH et al. 2014). Em um outro estudo de WES com 30 mulheres caucasianas com autismo, foram identificados SNPs em genes candidatos associados ao espectro de autismo, com alta probabilidade de causalidade (BUTLER et al. 2015).

Os estudos citados acima apontam que a investigação de alterações gênicas raras através de sequenciamento em larga escala de exomas em doenças neurodegenerativas podem ser informativas para identificar componentes genéticos de relevância para o desenvolvimento dessas doenças.

Em esclerose múltipla, os estudos de GWAS através das ferramentas WGS e WES, já identificaram mais de 100 variantes germinativas que conferem susceptibilidade, mapeadas em regiões estratégicas de genes com funções imunológicas (KEMPPINEN et al. 2014).

Inicialmente, o único componente genético associado à EM era a região do antígeno leucocitário humano (HLA). Nos últimos 5 anos de GWAS, 57 novos loci de susceptibilidade a EM foram identificados. As principais vias envolvidas são JAK-STAT, leucemia mielóide aguda e sinalização do receptor de células T (SAWCER et al. 2011; MULTIPLE e GENETICS 2013).

Na tentativa de encontrar padrões de variações genéticas e epigenéticas entre as células CD41T que justificassem a discordância da doença entre gêmeos monozigóticos em EM, o grupo do Dr. Stephens Kingsmore investigou pela primeira vez, em doença autoimune, o sequenciamento individual do genoma de gêmeos monozigóticos discordantes para EM (afetados e não afetados pela EM) totalizando três pares. O grupo observou que não houve diferença robusta do ponto de vista genético, epigenético e do transcriptoma entre os irmãos gêmeos e entre os irmãos dos três pares de gêmeos. Paralelamente, as sequências obtidas pelo sequenciamento do genoma também foram alinhadas com o genoma viral, mas nenhuma sequência alinhada no genoma viral diferenciou os indivíduos afetados dos indivíduos não afetados pela EM. O grupo concluiu que uma nova geração de ferramentas de

sequenciamento e analíticas pode ter um grande potencial no estudo de gêmeos monozigóticos discordantes, onde fatores de confusão genéticos e não-genéticos podem ser minimizados aumentando as chances de rastrear mecanismos envolvidos em doenças neurológicas complexas (BARANZINI et al. 2010).

RAMAGOPALAN et al. (2011) em um estudo de coorte familiar realizado no Canadá sequenciaram o exoma de 43 indivíduos não-relacionados portadores de EM e encontraram a variante rara rs118204009 em heterozigose, validada no mesmo estudo em um coorte maior (12,579 indivíduos com EM) com grande significância estatística. Esse polimorfismo raro é caracterizado por uma troca de arginina para leucina e induz a perda total da atividade da enzima 1-alfa-hidroxilase, codificada pelo gene *CYP27B1*. Essa enzima converte 25-dihidroxitamina D no metabólito ativo da vitamina D, de modo que a perda de função dessa enzima provoca a redução no níveis de calcitriol. Esse estudo reforça ainda mais a importância vitamina D deva influenciar a susceptibilidade à EM e que estudos com suplementação de vitamina D para a prevenção, ou mesmo no tratamento da EM, possam ser justificados (RAMAGOPALAN et al. 2011). Em 2012, o mesmo grupo ainda com análise de EM familiar, selecionou para o sequenciamento completo do exoma, uma família com 15 indivíduos afetados em 4 gerações. Após sequenciarem 4 indivíduos, esses autores identificaram uma variante no gene *TYK2* (rs55762744) encontrada em 72% dos casos (10/14) e 47% (28/60) de familiares não afetados. No mesmo estudo esta alteração foi avaliada em uma análise de 2104 trios, sendo encontrada em 0,8% dos pacientes, frequência idêntica à observada em indivíduos sem a doença (análises de nosso grupo no banco ExAC com cerca de 120 mil controles saudáveis – exac.broadinstitute.org). Os autores concluíram que esta variante, relacionada com a via de regulação de produção de citocinas, é uma variante de risco modesto para EM (DYMENT et al. 2012).

O primeiro sequenciamento de exoma em células únicas de pacientes com EM foi desenvolvido pelo grupo do pesquisador Stephen Sawcer, no Departamento de Neurologia da Universidade de Cambridge/UK em 2014. Eles sequenciaram linfócitos CD4+ do LCR de 21 pacientes nos estágios iniciais da EM para tentar identificar mutações somáticas associadas ao desenvolvimento de EM, para a referência foram usadas as células do sangue periférico dos mesmos pacientes.

Segundo os autores do artigo, devido aos artefatos gerados durante a etapa de amplificação do DNA genômico (gDNA) ou durante o sequenciamento, mais de 90% das diferenças observadas entre as células CD4⁺ e as células germinativas não são verdadeiras e, portanto não foi possível distinguir mutações somáticas autênticas. De acordo com esse estudo, as ferramentas de sequenciamento para análise de exoma de células únicas precisam ser aprimoradas para assegurar a detecção de mutações somáticas raras (KEMPPINEN et al. 2014).

1.7.1 Análises de dados de exoma

A disponibilidade de um exoma completo sequenciado permite diversas abordagens de análise alternativas. Uma delas é a busca por mutações *de novo*, ou seja, alterações encontradas no genoma do paciente não herdadas dos pais mas que surgem em algum momento na formação das células germinativas ou das fases iniciais do embrião.

Graças ao desenvolvimento de novas técnicas de sequenciamento de DNA que levaram à redução do custo, foi evidenciado que em humanos, a taxa de mutação média normal por geração está entre 7.6×10^{-9} e 2.2×10^{-8} nt/genoma/geração (LYNCH 2010). Deste modo, foi calculado que cada recém-nascido normal adquire em média de 50 a 100 novas mutações em seu genoma (ausentes no genoma de seus pais). Estas mutações espontâneas, quando se localizam em regiões importantes do genoma podem ter consequências funcionais. Em 2010 foi sugerido que estas mutações podem ajudar a explicar o porque de algumas doenças caracterizadas por severas deficiências reprodutivas continuarem sendo frequentes na espécie humana (VISSERS et al. 2010). Neste ano, o grupo do Dr. Joris Veltman (Universidade de Nijmegen, na Holanda) investigou a ocorrência de mutações *de novo* em pacientes com a síndrome de Schinzel-Giedion. Pacientes com esta síndrome são facilmente reconhecíveis, graças a certas características faciais próprias, severo retardo mental, múltiplas más-formações congênitas, além de uma excessiva formação de tumores, principalmente neoplasias neuroepiteliais. Após sequenciarem o exoma completo de 4 indivíduos não-relacionados, todos portadores desta doença, estes autores identificaram que todos os 4 pacientes apresentavam mutações importantes no mesmo gene, o *SETBP1*. Alterações diversas neste mesmo gene foram

posteriormente confirmadas em outros pacientes com a mesma doença, elucidando pela primeira vez as causas de uma doença ainda desconhecida, e abrindo possibilidades concretas de diagnóstico e tratamento (HOISCHEN et al. 2010).

No final do ano de 2010, este mesmo grupo demonstrou que mutações *de novo* podem ser uma das principais causas de retardo mental primário (VISSERS et al. 2010). Neste artigo, estes autores demonstram que mutações *de novo* puderam ser prontamente identificadas usando uma abordagem de sequenciamento do exoma em trios compostos por pai-mãe-filho afetado. Neste trabalho, foram sequenciados os exomas de 10 trios. Todos os pacientes incluídos tinham retardo mental moderado ou severo, sem histórico familiar da doença, sem qualquer diagnóstico sindrômico ou etiológico e sem alterações cromossômicas identificáveis. Dentre os 13 genes mutados identificados com maior grau de confiança, dez afetam direta ou indiretamente proteínas associadas ao retardo mental. Tais achados tiveram grande impacto na pesquisa genômica de doenças humanas e tem profundas implicações em aspectos que envolvem a gênese e o diagnóstico não apenas do retardo mental, mas também em diversas outras doenças humanas. Os autores concluem que o sequenciamento sistemático do exoma de trios envolvendo ambos os pais, além dos filhos afetados, pode revelar outros casos onde mutações *de novo* sejam capazes de explicar outras doenças humanas.

Em 2012 CHESI et al. sequenciaram o exoma de trios de 47 pacientes com esclerose lateral amiotrófica para identificar mutações *de novo* e encontraram alterações em genes que codificam reguladores da cromatina, como *CREST* e *FUS*, candidatos a contribuir com os mecanismos patogênicos dessa doença. O quadro descrito acima reforça a possibilidade de que a ocorrência de mutações *de novo* possa ser uma das possíveis causas do estabelecimento de doenças neurodegenerativas.

Análises de mutações *de novo* ainda não foram realizadas para pacientes portadores de EM, mas podem ser uma importante ferramenta, principalmente para o estudo da doença em sua principal apresentação, onde não há histórico familiar.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar alterações genômicas em indivíduos portadores de EM através do sequenciamento global do exoma.

2.2 OBJETIVOS ESPECÍFICOS

- o Avaliar alterações moleculares do tipo *de novo*;
- o Investigar retro-inserções *de novo* de retrovírus endógenos humanos;
- o Avaliar herança, em homozigose, de alelos raros de SNPs;
- o Validar pelo método de *Sanger* as alterações sugeridas pelos dados de sequenciamento.

3 CASUÍSTICA, MATERIAIS E MÉTODOS

3.1 CASUÍSTICA

Este estudo foi do tipo prospectivo e para viabilizá-lo contatamos diversos neurologistas do estado de São Paulo. Nos reunimos individualmente com cada um dos grupos participantes, quando explicitamos os objetivos do estudo e propusemos colaboração para o recrutamento dos pacientes e análise dos dados gerados. Neste processo esclarecemos o tipo de paciente a ser recrutado, de acordo com os critérios de inclusão e exclusão (apresentados abaixo). Os grupos que auxiliaram no recrutamento foram:

- Departamento de Neurologia, Hospital das Clínicas da Universidade de São Paulo, Ribeirão Preto, SP, por meio de colaboração com o Prof. Wilson Araújo da Silva Jr. e Dr. Amilton Antunes Barreira. Em Ribeirão Preto, contamos ainda com o apoio logístico de nosso colaborador local, Dr. Júlio César C. Lorenzi, onde coletamos 2 trios e 2 quartetos completos;
- Drs. Guilherme Sciascia Olival e da Dra. Juliana Amm Diniz, neurologistas de São Paulo, SP, onde obtivemos 3 trios completos;

Os dois quartetos restantes foram coletados em São Paulo através de contatos feitos pelo nosso grupo, totalizando 9 famílias coletadas (paciente-mãe-pai e irmão quando houve).

As amostras de outros pacientes, usadas em sequenciamento específico para investigar a ocorrência de um dos achados foram obtidas de nossos colaboradores Drs. Alessandro Farias (UNICAMP), Dra. Doralina Brum (UNIFESP), Dr. Amilton Barreira (USP Ribeirão Preto) e Dr. Guilherme Olival (Santa Casa de São Paulo).

Os critérios de inclusão adotados foram:

- ✓ Diagnóstico de esclerose múltipla do tipo remitente-recorrente, de acordo com os critérios revisados de McDonald (POLMAN et al. 2011);
- ✓ Pacientes com idade entre 18 e 60 anos e com ambos pais biológicos vivos e disponíveis no momento da coleta do sangue;

Quanto aos critérios de exclusão estes foram:

- ✓ Presença de comprometimento mental que não permita entender o termo de consentimento livre e esclarecido (TCLE);
- ✓ Presença de casos de EM, doenças neurodegenerativas ou psiquiátricas na família (parentesco de primeiro grau) (CARTA et al. 2014);
- ✓ Presença de alguma doença grave e que requeiram procedimento cirúrgico para acesso venoso;
- ✓ Pacientes com alguma outra doença sistêmica concomitante;
- ✓ Pais ou pacientes que sofreram transplante de órgãos ou transfusão sanguínea;
- ✓ Ausência de um ou de ambos pais biológicos, ou que estes não concordassem em participar do estudo.

Os indivíduos que se encaixaram nos critérios do estudo foram convidados a participar. Aqueles que aceitaram fizeram a leitura, concordaram e assinaram o TCLE (**Anexo 2**). A execução deste projeto foi aprovada pelos comitês de ética em pesquisa das instituições pertinentes (A.C.Camargo Cancer Center – 1748/13 e Faculdade de Medicina/USP, Ribeirão Preto) conforme o **Anexo 1**.

Deste modo, a casuística deste estudo foi composta por 31 indivíduos. Quinze indivíduos estavam distribuídos em 5 famílias, compostas por pai, mãe e paciente afetado pela EM (trios). Outros 16 indivíduos estavam distribuídos em 4 famílias contendo o trio pai-mãe-paciente e também um irmão biológico saudável (quartetos) (**Quadro 1**).

Quadro 1 – Informações gerais dos pacientes e famílias avaliadas.

Grupos	Trio	Sexo	Idade	Média idade	Idade ao surgimento dos primeiros sintomas	Média idade ao surgimento dos primeiros sintomas	Idade ao diagnóstico	Média de idade ao diagnóstico	Tempo de seguimento (anos)	EDSS	Média de EDSS	Tratamento na Coleta
Paciente	1	F	37	32,25	23	22,75	23	24,75	14	3,5	2,25	Fingolimod
	2	F	31		16		23		12	1		REBIF 44 (Betainterferon-1a)
	3	M	32		15		18		13	1		Tysabri\Natalizumab
	4	M	32		28		29		3,6	1		Avonex, 1A (Betainterferon)
	5	F	27		23		24		3	3		Copaxane/acetato de glatirâmer
	6	F	31		26		27		4	1		Avonex, 1A (Betainterferon)
	7	F	38		30		32		13	3,5		Tysabri\Natalizumab
	8	F	30		21		22		8	4		Tysabri\Natalizumab
Mãe	1	F	72	60,625	-		-		-	-		-
	2	F	63		-		-		-	-		-
	3	F	60		-		-		-	-		-
	4	F	60		-		-		-	-		-
	5	F	55		-		-		-	-		-
	6	F	56		-		-		-	-		-
	7	F	65		-		-		-	-		-
	8	F	54		-		-		-	-		-
Pai	1	M	67	61,375	-		-		-	-		-
	2	M	65		-		-		-	-		-
	3	M	64		-		-		-	-		-
	4	M	63		-		-		-	-		-
	5	M	56		-		-		-	-		-
	6	M	54		-		-		-	-		-
	7	M	67		-		-		-	-		-
	8	M	55		-		-		-	-		-
Irmã(o)	2	M	40	32	-		-		-	-		-
	4	F	30		-		-		-	-		-
	5	F	23		-		-		-	-		-
	6	F	35		-		-		-	-		-

Legenda – EDSS = *Expanded Disability Scale Score*.

Observação - Durante nossas análises descobrimos por análise dos dados de exoma que, para um dos trios, não há indícios de relação biológica entre os pais e o paciente (trio #9; vide a sessão de Resultados e Discussão). Deste modo, dados deste paciente não foram apresentados na tabela acima.

3.2 CONFIRMAÇÃO DAS RELAÇÕES DE PARENTESCO ENTRE OS INDIVÍDUOS ANALISADOS

Solicitamos a colaboração do Prof. Dr. Sérgio DJ Pena, do Instituto de Ciências Biológicas da UFMG, especialista em genética molecular e pioneiro nacional em análises de paternidade humana, para nos auxiliar a avaliar a discrepância de variantes em homozigose nas famílias participantes do estudo da nossa casuística com os dados do sequenciamento do exoma através de um protocolo adaptado (RYAN et al. 2013). Com base nas análises de todos os indivíduos do estudo, observamos as discrepâncias de SNPs em homozigose obtidas para todos os possíveis pares de pais x filhos e mães x filhos, entre irmãos e entre indivíduos não relacionados. Dentro das faixas de variação observadas pudemos avançar com as análises que determinaram presença ou ausência de parentesco.

3.3 OBTENÇÃO DO DNA

Como fonte do estudo, utilizamos leucócitos de sangue periférico, amostra de fácil obtenção que permite a extração de DNA de boa qualidade e quantidade adequada. Para isto coletamos 4mL de sangue periférico de cada indivíduo em tubos contendo EDTA. Este DNA foi usado para a investigação de alterações genômicas estruturais sub-microscópicas e mutações, ambas detectáveis e quantificáveis por protocolos de sequenciamento em larga-escala.

A extração de DNA foi feita conforme instruções do *Wizard Genomic DNA Purification kit* (Promega, Madison, WI, USA- catálogo A1120) seguindo as instruções para extração a partir de sangue total (protocolo 3.A. do *Wizard Genomic DNA Purification Kit* revisado 12/2010). Para algumas poucas amostras o DNA foi extraído diretamente com o kit *PureLink Genomic DNA minikit* (Invitrogen™, Carlsbad, CA, USA - Catálogo k1820-01), seguindo as instruções do fabricante.

3.4 CONSTRUÇÃO DAS BIBLIOTECAS DE EXOMA

Optamos pelo sistema de captura (*Ion TargetSeq™ Exome Kits* - Life Technologies, California, EUA). Apesar de ser um pouco mais trabalhoso, mais demorado e requerer mais DNA genômico, este kit oferece uma cobertura mais ampla do exoma quando comparado ao kit mais recente desta empresa, baseado em amplificação (36% a mais, ou 45,1Mb x 33Mb). Esta cobertura inclui exons e sítios doadores/aceptores de *splicing* (39,5 Mb), >99,4% dos CCDs (*consensus coding DNA*) e > 97% das variantes do COSMIC V56 (<100 bp). Além disto, o kit também captura a porção 3'-UTR destes genes, incluindo 44.100 sítios preditos de ligação a miRNA e 1,472 RNAs não-codificadores tais como miRNA, tRNA, snoRNA, lncRNA, totalizando os 45,1 Mb alvo.

Para a confecção de cada biblioteca de exoma, utilizamos 1µg de DNA genômico leucocitário quantificado através do equipamento Qubit (*Qubit® 2.0 Fluorometer*, *Invitrogen*). As bibliotecas foram feitas seguindo o protocolo recomendado para o *Ion Xpress™ Plus Fragment Library Kit*. A cada etapa do protocolo de construção de biblioteca do exoma ocorre uma etapa de purificação da amostra com as *beads* magnéticas AMPure XP(*Agencourt® AMPure® XP Reagent*, *Beckman Coulter*).

Inicialmente o DNA foi ligado a *beads* magnéticas AMPure XP com 1,8 vezes o volume da amostra, e fragmentado enzimaticamente utilizando o mix de enzimas provenientes do kit *Ion Shear™ Plus Enzyme Mix II* por 30 minutos a 37°C. Em seguida, o DNA fragmentado foi purificado com o reagente AMPure XP e teve seu perfil avaliado no aparelho Bioanalyzer 2100 (*Agilent® High Sensitivity DNA Kit* - Agilent Technologies, California, EUA), onde o esperado era quando obtivemos picos na faixa de 70-400 pares de bases, como exemplificado na **Figura 4**. Nessa análise a área dos picos é dividida em múltiplas regiões e o software determina a média do tamanho em pares de bases, percentagem do DNA total analisado, molaridade e concentração.

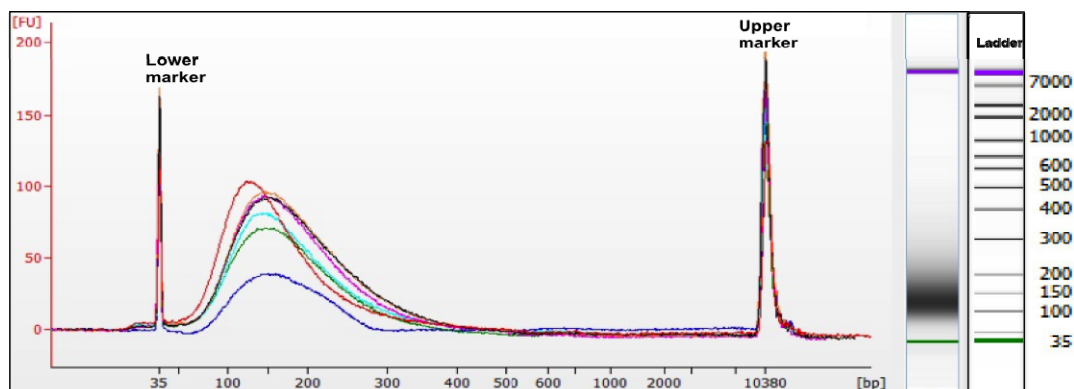


Figura 4 – Exemplo do perfil de distribuição de tamanho após fragmentação enzimática do gDNA, visualizadas utilizando o equipamento bioanalyzer.

Após a fragmentação foi feita a ligação dos adaptadores P1 e A, onde o adaptador A contém uma sequência de 10 bases únicas (*barcodes*) que permite a identificação de cada amostra após a corrida de sequenciamento. Essa etapa é seguida pela seleção por tamanho da região-alvo de interesse através do *E-Gel® SizeSelect™ Agarose Gel 2%*, quando recuperamos fragmentos na faixa de 275-295 pares de bases (fragmentos de DNA associados aos *barcodes*), conforme exemplificado na **Figura 5**. Cada biblioteca foi amplificada por oito ciclos de PCR com a enzima *Platinum® PCR SuperMix High Fidelity*. A seguir, cada biblioteca foi quantificada através do kit *Qubit® 2.0 Fluorometer* com o kit *Qubit® dsDNA HS Assay*, para verificar se havia massa necessária para prosseguir com o protocolo, e analisada no *Agilent® Bioanalyzer®* com o *Agilent® High Sensitivity DNA Kit*, a fim de avaliar o perfil de tamanho das bibliotecas.

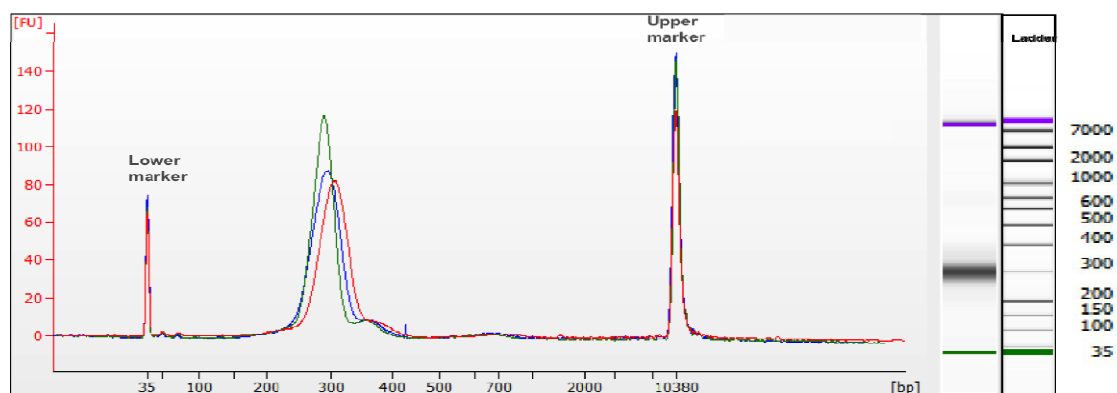


Figura 5 – Exemplo de perfil de tamanho das bibliotecas após amplificação, avaliada através do bioanalyzer.

A próxima etapa é a de captura das regiões exônicas, na qual utilizamos 500 ng de biblioteca amplificada para cada trio ou quarteto. A captura dessas regiões se dá pela hibridização das bibliotecas com sondas biotiniladas (*Ion TargetSeq™ Custom Probe Pool*), além disso, usamos bloqueadores dos adaptador P1 e A, com a finalidade inibir a captura em regiões inespecíficas. Em seguida, a reação de hibridização ocorreu por 72 horas em um termociclador (*Veriti 96 Well Thermal Cycler –AB Applied Biosystems*) a 47°C com a tampa aquecida a 57°C, a fim de evitar evaporação. Logo após o período de hibridização, o DNA capturado foi ligado a esferas de estreptavidina (*Dynabeads® M-270, Invitrogen*) em uma reação de 45 minutos a 47°C no termociclador, seguido por lavagens descritas no protocolo para eliminação das porções que não são alvo. O DNA capturado foi utilizado como molde para amplificação por PCR por oito ciclos. Nesse passo, a distribuição de tamanho da biblioteca enriquecida foi verificada no aparelho *Bioanalyzer* (**Figura 6A**) e quantificada pelo kit de amplificação por PCR quantitativa em tempo real (qPCR) (*Ion Library Quantitation Kit*). Esse kit de quantificação usa cinco diluições seriadas (1:10) partindo de 6.8pM de uma biblioteca de DNA genômico de *E. coli* (*E. coli DH10B Control Library*) para gerar uma curva padrão, que é então usada para cálculo da concentração da biblioteca enriquecida. No protocolo de preparo da amostra para quantificação, fazemos 3 diluições da biblioteca, quando é preciso que pelo menos uma das diluições fique entre os pontos da curva padrão (**Figura 6B**).

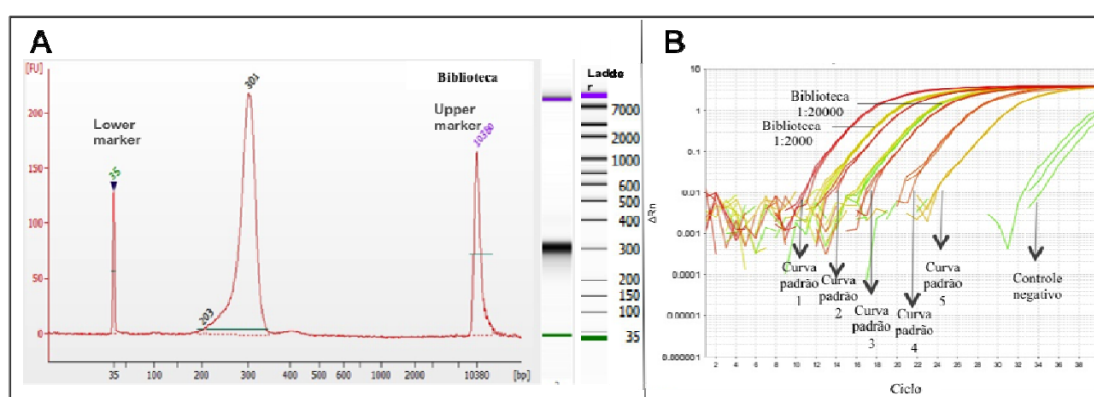


Figura 6 – Análise do tamanho e quantificação da biblioteca final. A) representa uma imagem da distribuição por tamanho da biblioteca final gerada pelo equipamento *bioanalyzer*, onde vemos o pico de 301pbs dentro do tamanho esperado. B) representa uma curva de qPCR contendo cinco pontos da curva padrão e as respectivas diluições (1:2000 e 1:20000) da biblioteca.

Uma vez prontas, fizemos a etapa de preparo da PCR em emulsão (ePCR) para cada trio ou quarteto (kit v3, Life Technologies, California, EUA). Essa etapa tem por objetivo fazer uma amplificação clonal do DNA capturado nas esferas. No preparo da emulsão, são necessários 100pM de material. O fator de diluição para cada biblioteca foi determinado pelo resultado do qPCR, onde o valor dado pelo equipamento foi multiplicado pela diluição da biblioteca que estava dentro da curva padrão, e dividido por 100. A diluição precisa da biblioteca favorece a distribuição de uma única molécula de DNA em cada nanopartícula com sequência complementar a sequência do adaptador P1 (*Ion Sphere Particle - ISP*) que será amplificada pelo método de PCR em emulsão.

Para checar a eficiência da PCR em emulsão, foi feita uma quantificação baseada em fluorescência. Para isso, sondas que se ligam especificamente aos adaptadores A (localizado na porção livre da molécula - fluorescência para 647 nm), e P1 (ligado às *ISPs* – fluorescência para 488nm), foram acrescentadas a uma fração da amostra (2uL do volume final), e suas fluorescências lidas no equipamento Qubit 2.0 (Life Technologies, California, EUA). A relação entre as duas sondas nos informa a proporção de *ISPs* que contêm moléculas da biblioteca e *ISPs* vazias.

Na etapa seguinte fizemos o enriquecimento da amostra utilizando o método de seleção pelo complexo estreptavidina-biotina, onde capturamos apenas as *ISPs* que contêm moléculas de DNA amplificado. O próximo passo foi o sequenciamento na plataforma *Ion Proton* (Ion Torrent, Life Technologies) que se baseia na detecção da alteração do pH causada pela liberação de um íon de hidrogênio toda vez que uma base nucleotídica é incorporada à cadeia nascente de DNA pela DNA polimerase. O sinal deste íon é convertido em um sinal digital e interpretado pelo equipamento como a incorporação de uma determinada base. No Núcleo de Genômica do CIPE/ACCCC temos gerado, em média, entre 12-15Gb de sequências mapeáveis por corrida. Para a reação de sequenciamento, a amostra é depositada em um *chip* (Ion PI Chip), que possibilitaria a leitura de até 100 milhões de sequências, dentro do limite teórico de uma única *ISP* por poço.

Previamente à deposição da amostra, o *chip* passou por diversos passos de lavagem antes da deposição da amostra. Posteriormente, adicionamos 5µL de *Ion PI™ Control Ion Sphere™* e 20 µL do *Ion PI Sequencing Primer* à biblioteca

enriquecida, e incubamos a 95°C por 2min e a 37°C por 2min. As *Ion Control Spheres* fornecem um controle de qualidade sobre o sequenciamento e as amostras sequenciadas. O DNA molde pronto foi depositado no *chip* e nos 5 minutos anteriores ao sequenciamento acrescentamos a enzima *Ion PI Sequencing Polymerase v3* (**Figura 7**).

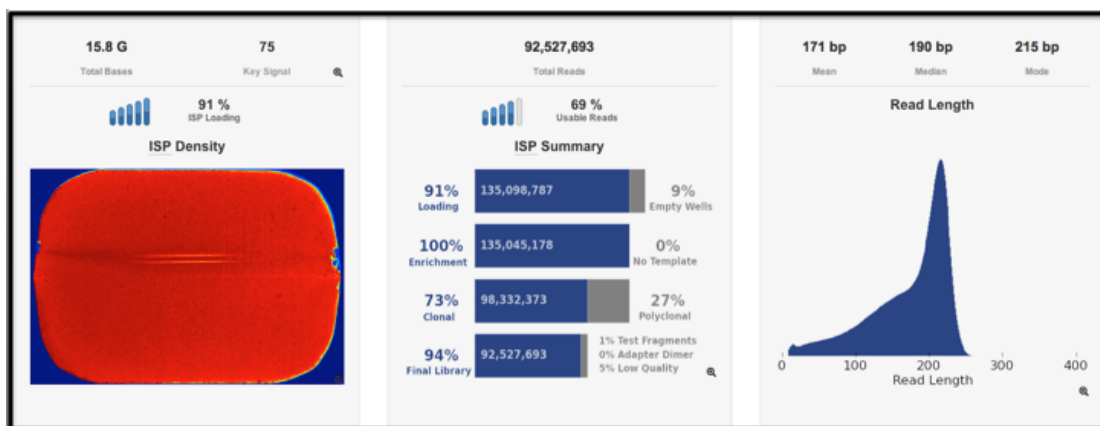


Figura 7 – Resultado do sequenciamento do exoma. Esse sequenciamento gerou aproximadamente 92,5 milhões de seqüências com uma média de 190 pares de base, e um total de 15,8 Gb de dados.

Para obtermos uma quantidade similar de *reads* para cada amostra, organizamos o sequenciamento da seguinte forma:

- 1) Trios (pai, mãe e filho afetado) - utilizamos 3 amostras por captura e sequenciamos 2 capturas em 3 chips.
- 2) Quartetos (pai, mãe, filho afetado e irmã/o) – selecionamos 4 amostras por captura e sequenciamos 1 captura em 2 chips.

Dessa forma obtivemos uma cobertura média de pelo menos 20x para 90% das bases sequenciadas em 95% das amostras.

3.5 ANÁLISES DE BIOINFORMÁTICA

3.5.1 Pipeline geral do Exoma

O *pipeline* de análise incluiu o alinhamento das sequências contra o genoma humano (Hg19) utilizando o *software* TMAP (*Torrent Mapper* 4.3.18 (6B3FD1B) 201408050144). Posteriormente fizemos a chamada de variantes utilizando o *software* TVC 4.0-5 (72041) – *Torrent Variant Caller*, seguindo o *pipeline* do *Ion Torrent* (mendel.iontorrent.com/ion-docs/) e revalidamos as variantes através do *pipeline* do GATK (www.broadinstitute.org/gatk/guide/best-practices?bpm=DNaseq). Para a anotação das variantes utilizamos o *software* SnpEff versão 3.5 (build2014-03-05) e um *script in house* para o alinhamento das variantes descobertas contra dados dos seguintes bancos públicos: dbNSFP versão 2.4; COSMIC v69; 1000genomes; ExAC (*Exome Aggregation Consortium-Broad Institute*); ESP versão ESP6500SI-V2; HapMap; dbSNP versão 138, Polyphen (*Polymorphism Phenotyping v2* – Harvard), Abracadabra (Kitajima JP & Schlesinger D. Abracadabra, *patent pending*) e SIFT (<http://sift.jcvi.org/> - J. Craig Venter Institute). Uma síntese deste *pipeline* encontra-se abaixo na **Figura 8**.

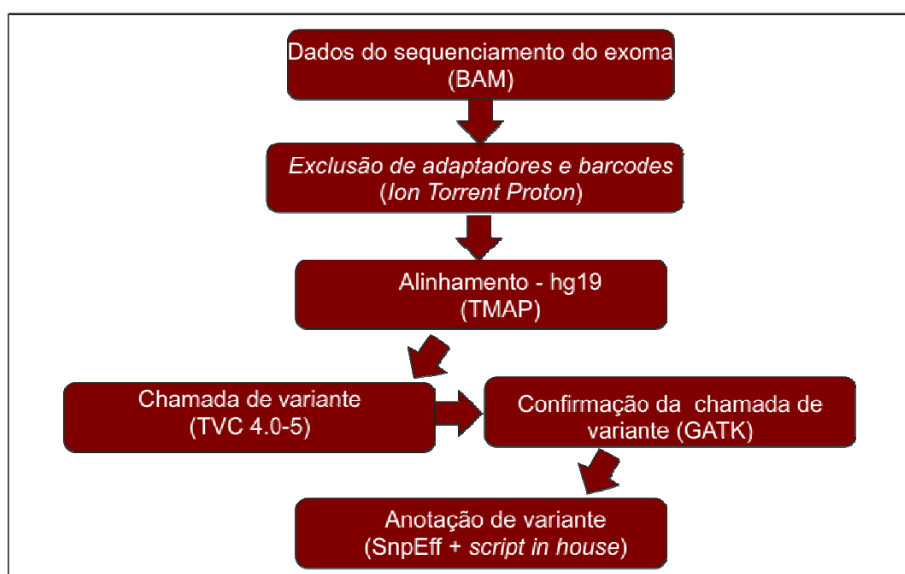


Figura 8 – *Pipeline* geral de análise de exoma para a determinação de bases variantes.

3.5.1.1 Variantes nucleotídicas únicas (SNVs) *de novo*

Inicialmente, devido ao pequeno número de alterações *de novo* observadas em cada grupo, reduzimos a estringência dos critérios utilizando os seguintes filtros para a identificação das SNVs *de novo*:

- Cobertura mínima da base possivelmente alterada de pelo menos 10x para cada indivíduo (pais e paciente);
- Frequência da variante $\geq 20\%$ no paciente;
- Para cobertura $\geq 20X$ permitimos no máximo 1 *read* variante de discordância entre os pais;
- Irmã/o (quando houver) deve ser homozigoto para a referência.
- As variantes deveriam contemplar *reads* na fita de DNA *forward* e *reverse*.
- Foram excluídas ainda aquelas alterações que ocorrem em homopolímeros.

3.5.1.2 Análise de inserções *de novo* de elementos retrovirais (DNR)

Com o objetivo de detectar potenciais eventos de retroinserção *de novo* (DNR), estabelecemos uma colaboração com o Dr. Israel Tojal da Silva (*Lab. of Molecular Immunology, The Rockefeller University*). Dr. Tojal da Silva tem experiência nesta área e recentemente desenvolveu uma abordagem para o estudo da integração viral no genoma humano para caracterizar a latência em células infectadas por HIV-1 (COHN et al. 2015). O *pipeline* usado neste artigo foi modificado e adaptado para a busca de eventos de retro-inserção *de novo* em nossos dados de exoma.

Como a nossa análise foi baseada em dados de sequenciamento completo do exoma (WES), utilizamos anotação de transcritos humanos obtida a partir do *UCSC Genome Browser website (track RefSeq genes - GRCh37/hg19)*. Para os genes com mais de uma isoforma, foi selecionada aquela mais longa (QUINLAN e HALL 2010). Inicialmente, recuperamos 3862 sequências de retrovírus endógenos a partir do repositório *Taxonomy Browser – NCBI/sessão Human Endogenous Retroviruses* (www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Tree&id=206037&lvl=3&p=7&lin=f&keep=1&srchmode=1&unlock). Estes retrovírus foram então anotados no genoma humano (*UCSC Genome Browser website (track RefSeq genes - GRCh37/hg19)*) com o *software* de alinhamento local

BWA (*Burrows - Wheeler Alignment Tool*) (LI e DURBIN 2009) e as coordenadas dessa anotação foram mantidas para posterior análise. Paralelamente, usamos o mesmo repositório de retrovírus para alinhar com as *reads* geradas pelo sequenciamento do exoma usando o software BWA. As *reads* que continham fragmentos derivados dos 3862 retrovírus endógenos foram detectadas pelo software. Na etapa seguinte, usando um *script* próprio foram mantidas apenas as *reads* com alinhamento parcial às sequências dos retrovírus. Estas *reads*, por sua vez, foram alinhadas às coordenadas do hg19 anotadas anteriormente via BWA. Somente as *reads* com alinhamento único ou definido pelo maior *score* de alinhamento foram mantidas. Em um passo adicional, as *reads* que compartilharam a mesma região anotada pelo banco de retrovírus no hg19 foram descartadas da análise. Finalmente, *reads* que passaram no filtro anterior e com as coordenadas genômicas localizadas até ± 1 Kb das bordas TSS (*transcription start site*) e TTS (*transcription termination site*) dos genes foram utilizadas para definir a localização dos potenciais eventos de retroinserção conforme ilustrado na **Figura 9**.

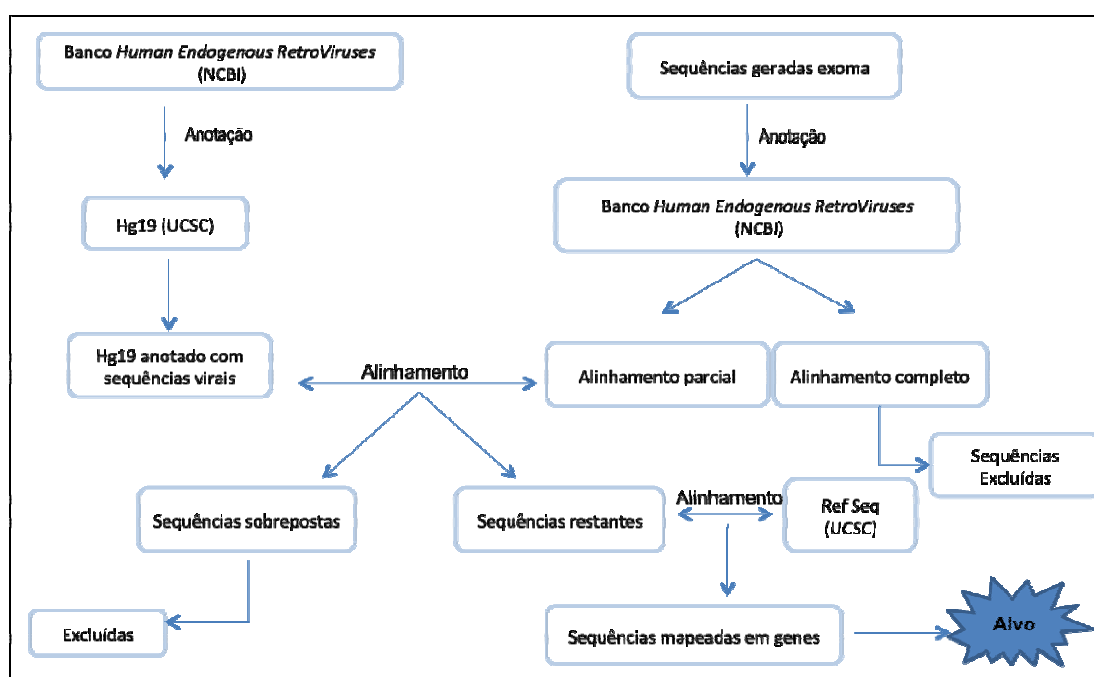


Figura 9 – Fluxograma do *pipeline* para detecção dos eventos de retroinserção de novo (DNR) obtidos através dos dados do sequenciamento do exoma.

3.5.1.3 Polimorfismos raros de base única (SNPs raros)

Com o objetivo de identificar alterações genômicas que possam ter envolvimento com esclerose múltipla e explorar os dados obtidos pelo sequenciamento do exoma, avaliamos a ocorrência de polimorfismos raros (*minor allele frequency* $\leq 5\%$), do tipo SNP, herdados em homozigose pelo paciente.

Os filtros utilizados foram:

- Cobertura mínima de 10x e 100% de frequência (homozigose) do alelo raro no paciente.
- Ambos os pais devem ser heterozigotos para o mesmo SNP;
- Os irmãos (quando houver) não podem ser homozigotos para a variante;
- Incluímos apenas alterações *missense* ou *nonsense*;
- MAF global (*Minor Allele Frequency*) $\leq 5\%$ de acordo com ExAC (exac.broadinstitute.org)
- As variantes deveriam contemplar *reads* na fita de DNA *forward* e *reverse*.

3.6 VALIDAÇÃO DOS ACHADOS

3.6.1 Seleção dos Candidatos a SNVs *de novo*

Visando foco inicial em alterações com impacto funcional mais óbvio, focamos a validação nos SNVs com localização em regiões codificadoras de proteína e que levassem a alteração da sequencia protéica. Além disto, priorizamos a validação de alterações localizadas em regiões com maior potencial de envolvimento com doenças do sistema nervoso central, resposta imune e metabolismo de vitamina D, por análise da literatura.

3.6.2 Seleção dos Candidatos a SNPs raros herdados em homozigose

Geramos uma lista de 24 candidatos, desses fizemos uma busca no PubMed e selecionamos aqueles SNPs raros em genes que tivesse envolvimento com doenças do SNC. Levamos ainda em consideração a possibilidade desta alteração ser ou não patogênica, com base em análise nos programas POLYPHEN (<http://genetics.bwh.harvard.edu/pph2/>), ABRACADABRA (Kitajima JP & Schlesinger D. Abracadabra, *patent pending*) e SIFT (<http://sift.jcvi.org/>).

3.6.3 Amplificação dos genes de interesse para validação

Pares de iniciadores foram desenhados de forma a permitirem a amplificação por PCR da região onde estão localizados os SNVs *de novo* e os SNPs raros selecionados para a validação. Para isto, utilizamos o programa *Primer3* levando em consideração a temperatura calculada de anelamento (T_m), uma distância mínima de 150 nt e máxima de 250 nt entre ambos os iniciadores e o nucleotídeo de interesse. O potencial de formação de homodímeros e de estruturas secundárias foi avaliado com a ferramenta *Oligotech* e para avaliar o potencial de formação de heterodímeros utilizamos a ferramenta *OligoAnalyzer* (*Integrated DNA Technologies – IDT*).

A amplificação por PCR foi feita inicialmente com 7,5 μ L do *GoTaqGreen Master Mix* (Promega, Wisconsin, EUA), para o volume final de 15 μ L, 5ng do gDNA, 2,5pmol dos iniciadores *forward* e *reverse* e água livre de nucleases utilizando uma etapa inicial de 95°C por 3 minutos para a ativação da enzima seguida de 35 ciclos de 95°C por 30 segundos para desnaturação, 57°C por 30 segundos para anelamento dos iniciadores e 72°C por 30 segundos para extensão. Ao final incluímos uma etapa de 72°C por 2 minutos para extensão.

Para as alterações nos genes *MSANTD1* e *GALR2* a PCR foi feita com um total de 40 ciclos, usando 30 segundos de anelamento a 60°C (*MSANTD1*) ou a 58°C (*GALR2*). Para ambas reações eliminamos a extensão com o objetivo de deixar a reação mais limpa e evitar múltiplas bandas de maior peso molecular. Na eletroforese em gel de agarose 1% utilizamos 5 μ L da reação de amplificação.

3.6.4 Validação dos achados por sequenciamento Sanger

Para o sequenciamento, purificamos os produtos de PCR com a enzima Exonuclease I e *Shrimp Alkaline Phosphatase* (ExoSAP – *ExoSAP-IT®* (Affymetrix, California, EUA) para remover os desoxirribonucleotídeos trifosfatados (dNTPs) excedentes, além de iniciadores não utilizados. A reação ocorreu a 37°C por 15 minutos e a 80°C por 15 minutos.

Para a reação de sequenciamento pelo método *Sanger* de terminação de cadeia, utilizamos 2,5 μ L do DNA purificado, 1,6 μ L do iniciador (2,5 μ L pmol/ μ L), 0,7 μ L do *BigDye Terminator v3.1 Ready Reaction Mix* (Life Technologies,

California, EUA), 2µL de *5X Sequencing Buffer* e 3,2µL de água livre de nucleases para um volume final de 10µL. A reação é colocada em um termociclador a 95°C por 2 minutos, seguida de 40 ciclos de 95°C por 18 segundos, 55°C de anelamento por 18 segundos e extensão de 60°C por 4 minutos. Após a ciclagem, o DNA amplificado foi precipitado com 1µL de NaOAc 3M pH 5,2, 1µL de EDTA 125mM pH 8,0 e 25µL de etanol 100% e incubado por 15 minutos no escuro a temperatura ambiente. Após centrifugação (centrífuga Eppendorf 5810R, rotor A-4-62) a 4000rpm por 30minutos a 4°C, o sobrenadante foi descartado e a placa invertida foi centrifugada a 650rpm por 1minuto com aceleração e desaceleração de 2. Ao adicionar 35µL de etanol 70% gelado a placa é novamente centrifugada a 4000rpm por 30min a 4°C. Seguimos novamente com um *spin* invertido a 650rpm com aceleração e desaceleração de 2 por 4 minutos. Posteriormente secamos o *pellet* por 20minutos no termociclador a 95°C e adicionamos 13µL de formamida Hi-Di (Life Technologies, California, EUA) seguido de desnaturação por 3 minutos a 95°C e incubação por 3 minutos no gelo. Seguimos com o sequenciamento no sequenciador capilar automático *3130xl Genetic Analyzer* (Life Technologies, California, EUA) e as sequencias geradas foram avaliadas no programa *CLC Genomics Workbench 6.0.3*.

4 RESULTADOS E DISCUSSÃO

Para avaliar eventos genômicos que ocorrem em pacientes diagnosticados com EM, nós construímos bibliotecas (exemplificado na **Figura 10**) e sequenciamos o exoma de 9 trios (pai, mãe e filho afetado – irmão/ã quando houve).. Dados globais do sequenciamento são mostrados a seguir, na **Tabela 1**.

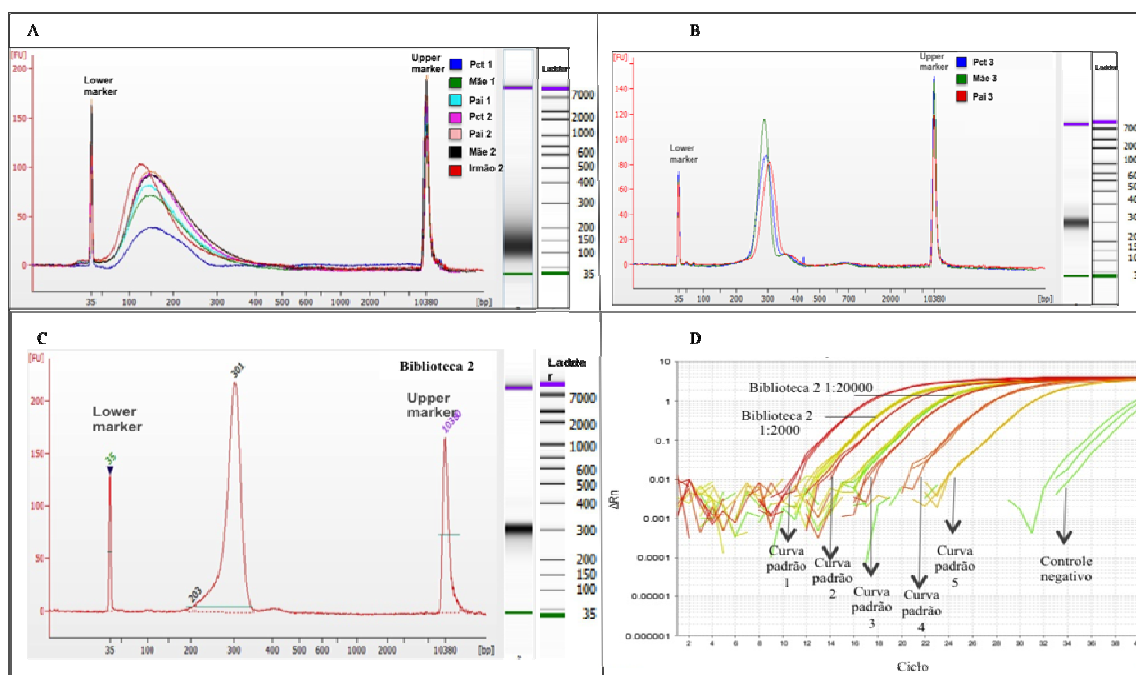


Figura 10 – Pontos de checagem da construção de biblioteca do exoma. a) perfil de distribuição de tamanho após fragmentação enzimática do gDNA das bibliotecas 1 e 2 contendo um trio e um quarteto (pct1, mãe1 e pai 1; pct2, pai2, mãe2 e irmão 2) respectivamente, avaliadas através do bioanalyzer; b) tamanho da biblioteca 3 após amplificação, contendo um trio (pct3, mãe3 e pai3), visualizada utilizando o equipamento bioanalyzer; c) distribuição por tamanho da biblioteca 2 final gerada pelo equipamento bioanalyzer e d) curva de qPCR contendo cinco curvas padrão e as respectivas diluições (1:2000 e 1:20000) da biblioteca 2. Pct = paciente.

Tabela 1 – Sequencias obtidas para cada indivíduo analisado neste estudo, incluindo dados de mapeamento.

Grupos	Amostras	#Total reads/amostra	Total de reads	#Reads mapeadas (hg19)/ amostra	Total de reads mapeadas	%	Média de cobertura onTarget (%)
1	Mãe	32.171.021	95.379.763	31.768.791	94.159.325	98.75	86.4
	Paciente	38.961.305		38.412.931		98.59	
	Pai	24.247.437		23.977.603		98.88	
2	Mãe	30.425.784	138.370.276	30.017.288	136.616.501	98.65	90.48
	Paciente	38.387.864		37.924.064		98.79	
	Pai	37.044.433		36.604.269		98.81	
3	Mãe	32.512.195	116.262.402	32.070.880	115.071.181	98.64	91.7
	Paciente	35.795.922		35.459.629		99.06	
	Pai	45.152.368		44.650.054		98.88	
4	Mãe	35.314.112	361.255.642	34.961.498	358.463.382	99.00	94.7
	Paciente	70.862.477		70.311.149		99.22	
	Pai	88.821.954		88.071.769		99.15	
5	Mãe	89.505.825	334.874.828	88.797.659	331.505.016	99.2	94.9
	Paciente	112.065.386		111.282.805		99.3	
	Pai	82.558.276		81.662.451		98.91	
6	Mãe	89.961.268	335.658.371	89.005.023	333.238.716	98.93	94.6
	Paciente	82.307.322		81.463.655		98.97	
	Pai	80.047.962		79.373.887		99.15	
7	Mãe	84.407.250	162.128.205	83.823.701	160.412.952	99.30	92.3
	Paciente	83.020.055		82.442.579		99.30	
	Pai	82.148.127		81.615.286		99.35	
8	Mãe	86.082.939	211.646.363	85.357.150	208.992.857	99.15	93.3
	Paciente	63.559.611		62.831.669		98.85	
	Pai	47.558.904		47.023.513		98.87	
9	Mãe	51.009.690	179.691.013	50.557.770	177.838.532	99.11	93.6
	Paciente	80.059.205		79.108.490		98.81	
	Pai	64.696.984		63.788.348		98.59	
	Mãe	66.890.174		66.096.019		98.81	
	Paciente	55.092.980		54.581.253		99.07	
	Pai	62.859.604		62.196.383		98.94	
	Mãe	61.738.429		61.060.896		98.90	
	Paciente						
	Pai						

Durante nossas análises observamos um alto grau de discrepâncias de SNPs nos indivíduos que compõe o nosso grupo #9. Em vários casos observamos, por exemplo, o pai e a mãe com o genótipo CC para um dado SNP e o filho com o genótipo AA para o mesmo. Tais achados são altamente sugestivos de troca de amostra, ou de ausência de relação de parentesco entre os indivíduos deste trio. Discutimos e analisamos estes dados junto com o Prof. Sérgio DJ Pena (ICB, UFMG), levantando o número de discrepâncias deste tipo observadas entre todos os indivíduos do estudo. Observamos entre pais e filhos dos demais trios a ocorrência de 7 a 294 posições discrepantes (devidas a mutações *de novo*, erros de sequenciamento, erros de mapeamento, etc). No caso do trio 9, em questão, observamos 2015 diferenças entre o filho e a mãe indicada e 2337 entre este mesmo indivíduo e o pai biológico indicado. Quando derivamos este número para todos os

indivíduos não-relacionados do estudo esse número variou entre 2033 e 2857. Estes dados confirmaram que: i) não houve troca de amostra, pois o paciente do grupo 9 não apresentou índices de discrepância baixos, compatíveis com parentesco, com nenhum outro indivíduo do estudo e ii) não há parentesco próximo entre este paciente e seus supostos pais biológicos. Deste modo, este trio foi excluído das análises *de novo* ou de SNPs raros em homozigose. Os dados do paciente foram usados apenas em algumas análises pontuais de frequência de achados dos demais casos. Os totais de discrepâncias observadas entre os indivíduos de cada grupo podem ser observados no **Quadro 2**, abaixo.

Quadro 2 – Análise de discrepância de variantes em homozigose nas famílias participantes do estudo.

Grupos	Número total de variantes	Mãe v Paciente	Paciente v Pai	Mãe v Pai	Irmão v Paciente	Irmão v Pai	Irmão v Mãe
grupo1	26678	11	10	2051	-	-	-
grupo2	25859	13	15	2033	627	244	10
grupo3	33129	7	294	2580	-	-	-
grupo4	36226	22	288	2857	838	8	21
grupo5	38018	8	10	2851	714	14	12
grupo6	33936	13	19	2657	611	12	8
grupo7	36679	17	12	2854	-	-	-
grupo8	33175	8	12	2390	-	-	-
grupo9	30281	2015	2337	2094	-	-	-

4.1 MUTAÇÕES *DE NOVO* EM PORTADORES DE EM

Durante a etapa de seleção dos candidatos para validação fizemos uma colaboração com o Dr. João Paulo Fumio Whitaker Kitajima e Dr. David Schlesinger, para checar a predição desses dos SNVs *de novo* e SNPs raros quanto a sua patogenicidade usando a ferramenta Abracadabra. Trata-se de uma ferramenta inovadora baseada em inteligência artificial, com alta sensibilidade para detecção de variantes patogênicas, em vias de ser patenteada (JP Kitajima, informação pessoal). Usamos ainda outros programas (Polyphen e SIFT), como anteriormente mencionado. Na etapa de análise dos dados, selecionamos os possíveis SNVs *de novo* (**Quadro 3**) de acordo com os filtros anteriormente apresentados.

Quadro 3 – Candidatos *Single Nucleotide Variants* (SNVs) *de novo*.

Trio	Coordenada genômica	Gene	Região	Descrição do gene	Classe de variantes	Referência/variante		Frequência da variante/cobertura				Efeitos/snp Eff	Predição Abracadabra	Predição Polyphen	Predição SIFT	Média de cobertura
						Nt	aa	paciente	mãe	pai	irmã/ão					
1	chr4:3257621	MSANTD1	CDS	<i>Myb/SANT-Like DNA-Binding Domain Containing 1</i>	missense	G/C	Val256Leu	22/18	0/18	0/23	-	moderado	D	B	B	20
2	chr7:20198265	MACC1	CDS	<i>Metastasis Associated In Colon Cancer 1</i>	missense	G/C	Phe573Leu	31,58/19	0/14	0/28	0/8	moderado	D	B	B	17
	chr3:9988171	PRRT3	CDS	<i>Proline-Rich Transmembrane Protein 3</i>	missense	C/T	Ala896Thr	50/58	0/55	0/50	0/64	moderado	B	B	B	57
3	chr17:79507110	C17orf70	3'UTR	<i>Chromosome 17 Open Reading Frame 70</i>	3'UTR	G/C	-	31,43/35	0/25	4,35/23	-	modificador	B	-	-	28
	chr1:153653692	NPR1	CDS	<i>Natriuretic Peptide Receptor</i>	missense	C/T	Pro320Ser	52,38/84	0/61	0/36	-	moderado	D	B	B	60
	chr5:95767783	PCSK1	Intronic	<i>Proprotein Convertase Subtilisin/Kexin Type 1</i>	intron	G/T	-	41,67/48	0/38	0/39	-	modificador	-	-	-	42
	chr16:29681008	SPN	3'UTR	<i>Sialophorin</i>	3'UTR	C/T	-	23,53/17	0/13	0/14	-	modificador	B	-	-	15
5	chr15:80215661	C15orf37, ST20	exon, intergenica	<i>Chromosome 15 Open Reading Frame 37, Suppressor Of Tumorigenicity 20</i>	exon, intergenica	G/A	-	38,46/13	0/19	0/16	0/17	modificador, modificador	B	-	-	16
	chr1:21502724	EIF4G3	5'UTR	<i>Eukaryotic Translation Initiation Factor 4 Gamma, 3</i>	5'UTR	T/C	-	50/16	0/26	0/27	0/23	modificador	B	-	-	23
	chr10:49400764	FRMPD2	CDS	<i>FERM And PDZ Domain Containing 2</i>	missense	C/G	p.Gly710Arg	20,78/77	0/69	0/58	0/70	moderado	B	B	B	68
	chr9:104170616	ZNF189	CDS	<i>Zinc Finger Protein 189</i>	stop gained	C/G	p.Ser189*	48,78/41	0/27	0/38	0/27	alto	D	PD	D	33
7	chr3:39142542	GORASP1	CDS	<i>Golgi Reassembly Stacking Protein 1, 65kDa</i>	frame shift	CTG/C	p.Thr133Ser	29,55/44	0/49	0/31	-	alto	D	PD	B	41

Legenda: As duas últimas colunas da tabela, as colunas de predição de patogenicidade representam: B = benigno; D= deletério e PD= possivelmente deletério.

Neste momento, para a seleção dos candidatos a validação optamos, primariamente, pelas alterações que levassem a uma alteração na sequência das proteínas. Adicionalmente, fizemos também uma análise na literatura que nos levou a seleção de alterações em genes com maior potencial de envolvimento em doenças do sistema nervoso central, resposta imune e metabolismo de vitamina D. Tivemos 60% de validação como apresentado na **Tabela 2** e na **Figura 11**.

Tabela 2 – Validação de candidatos SNVs *de novo*.

Grupo	Gene	Reads variante/cobertura	Iniciadores	Fragmento pbs	Predição Abracadabra	Predição Polyphen	Predição SIFT	Validação
1	MSANTD1	4/18	GGTGAAGAAGCGCAAGGTG CCCTGGGTGGCCTGAG	288	D	B	B	3
2	MACC1	6/19	GCCAGGCTATTTGCAGAAG GCTCCTTTGAAATCACCTTGAC	311	D	B	B	1
2	PRRT3	29/58	CACCTAGTGCGAGACAGTGTG GCTGGGCAACTCATCTAGGG	339	B	B	B	1
3	NPR1	44/84	GCAGGGCATAGGGTCCAG GGCAGAGGGTCATGGAGTG	315	D	B	B	2
3	SPN	4/17	GGATTACATG CGTGAGCCAC CAATCTTCCCTGCCTTCG	644	B	-	-	3
5	EIF4G3	8/16	CTTTGCCATGAGTTTACCGC GGACTGCGACGACGAGAAC	275	B	-	-	3
5	FRMPD2	16/77	CCAATTTGAGTCCTGCACAC GTACAGATCACTGCAAACCATC	255	B	B	B	2
5	ZNF189	20/41	CCAAACTCAGAAGAGAAATGCC GGCTGAAGCTCTGTTTACACTG	307	D	PD	B	1

Legenda – Na coluna validação os números representam: 1= validado; 2= Inconclusivo (repetir Sanger) e 3 = Não realizado (repetir PCR).

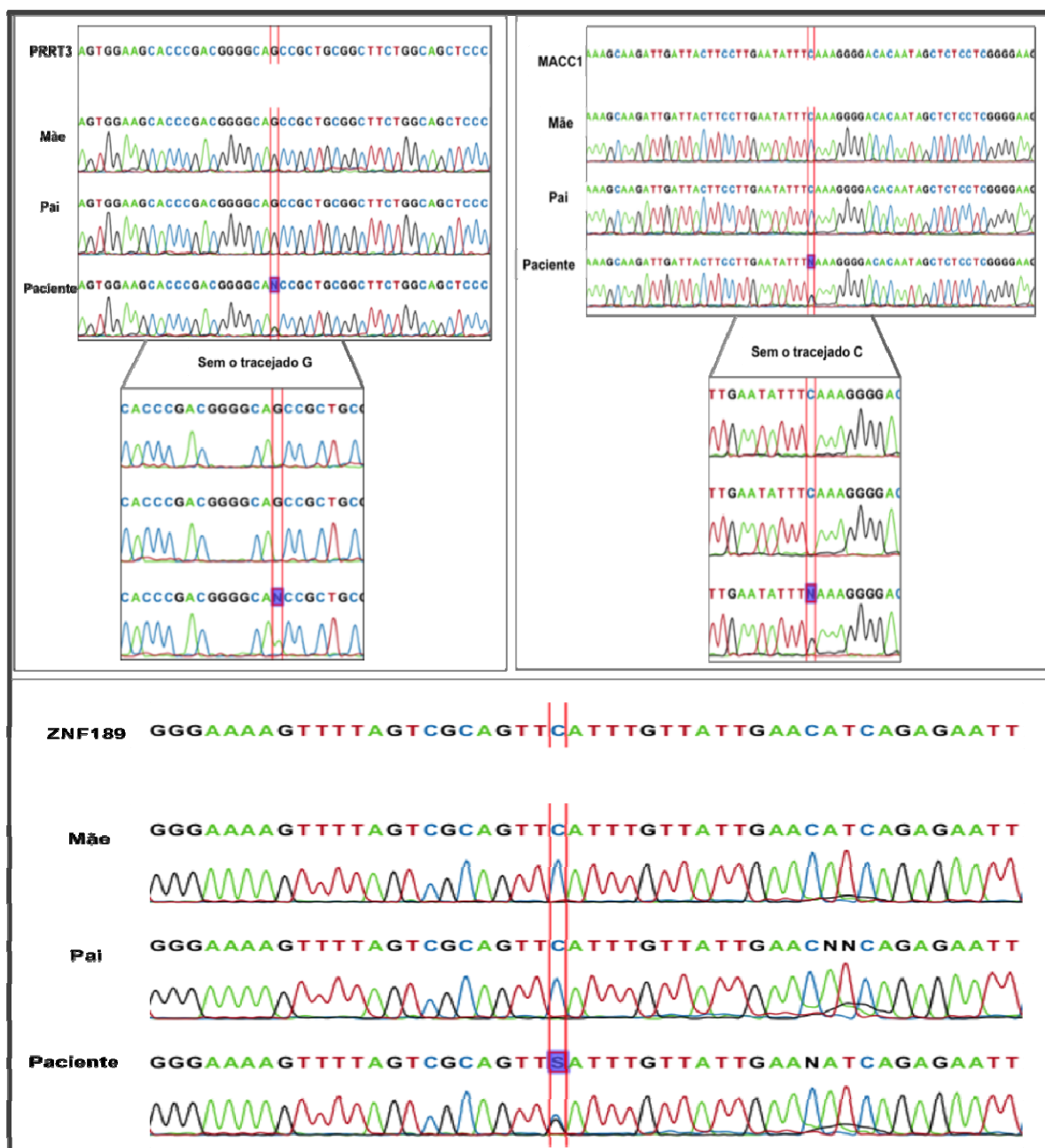


Figura 11 – SNVs *de novo* validados através da ferramenta Sanger.

Para a maioria dos genes, infelizmente, poucos estudos foram encontrados na literatura. Uma busca na literatura (PUBMED) com o gene *MACC1* apontou diversos estudos que avaliaram este gene em câncer. No entanto não encontramos dados mostrando seu envolvimento com processos do SNC. A troca de aminoácidos aqui observada (fenilalanina para leucina) foi predita como patogênica pela ferramenta Abracadabra, e como benigna pelas demais. Para gene *ZNF189* encontramos poucos

estudos funcionais. Em nossos dados de exoma identificamos uma mutação nesse gene que leva a uma proteína truncada predita como patogênica pelas ferramentas abracadabra e SIFT. Buscamos ainda alguma descrição de eventos com o gene *Proline-rich transmembrane protein 3 (PRRT3)* e infelizmente, ainda não há caracterização desta proteína transmembrana em termos de vias moleculares ou biológicas. Estudos funcionais no modelo animal de EM (EAE) ou em cultura de células são necessários para avaliar os efeitos destas alterações e se participam em processos envolvidos com a EM.

No entanto, alterações foram observadas em genes de maior interesse, discutidos a seguir:

- ✓ **NPR1** também conhecido como *Atrial Natriuretic Peptide Receptor A (ANPRA)* codifica para o receptor peptídeo natriurético atrial (guanilato ciclase). Este é um receptor ligado à membrana celular que funciona como receptor para hormônios peptídeos natriuréticos encontrados no cérebro, vasos sanguíneos e rins com uma função fisiológica de manutenção dos fluidos corporais e homeostase cardiovascular (PANDEY 2011). Atualmente, estudos têm avançado em demonstrar o papel dos peptídeos natriuréticos e seus receptores em vias de sinalização cGMP na plasticidade neural (PANDEY 2014).

O aumento da expressão de NPRA, provocado pelo receptor da vitamina D (via fortemente relacionada com a EM) presente em células do interior do ducto coletor medular, contribui com a excreção do sódio através do mecanismo dependente de cGMP. O aumento na expressão de NPRA pode contribuir para algum efeito benéfico no uso de 1,25(OH)2D3 em doenças do sistema cardiovascular (CHEN et al. 2007).

Além de ter um papel em processos cardíacos e vasculares, ZHAO e MA (2009) sugeriram que os peptídeos natriuréticos sejam uma nova classe de fatores extracelulares que regulam o processo axonal durante o desenvolvimento, através da sinalização de atividades relacionadas a formação de ramificações dos aferentes sensoriais centrais e crescimento axonal promovidas pelo gene C-type NP (CNP), cuja expressão ocorre unicamente na medula da espinha dorsal (ZHAO e MA 2009). Essa mutação *de novo* no gene NPR1 encontrada em pacientes com EMRR em nosso

estudo foi predita como patogênica pela ferramenta Abracadabra (Kitajima & Schlesinger) e pode ser um potencial alvo de estudos futuros com maior número de amostras.

- ✓ **FRMPD2** (*FERM and PDZ Domain Containing 2*) é um gene expresso em células epiteliais e, através do seu domínio PDZ 2 interage com diversas proteínas. Inicialmente, a proteína codificada por este gene foi descrita com ênfase em seu papel no estabelecimento e manutenção da polaridade celular e na adesão célula-célula, em um processo mediado por E-caderina (STENZEL et al. 2009). Uma publicação recente do grupo do Prof. Philip Rosenstiel (Universidade de Kiel, Alemanha) sugere que FRMPD2 possa atuar como uma molécula suporte que organiza espacialmente a formação da via de sinalização da resposta imune mediada por NOD2 (*nucleotide-binding oligomerization domain-2*) no intestino. Esse processo ocorre através da proteína Nod2 que reconhece o muramil dipeptídeo (MDP) derivado de bactérias e ativa a transcrição de NF-kB, onde FRMPD2 atua como um regulador positivo que interage com Nod2 através de repetições ricas em leucina restringindo a resposta imune mediada por Nod2 em células epiteliais polarizadas do intestino (STENZEL et al. 2009; LIPINSKI et al. 2012).
- ✓ **SPN** (Sialophorin) – As sialoforinas são as principais sialoglicoproteínas encontradas na superfície de células T, mas também estão em algumas células B e apresentam um importante papel nas propriedades físico-químicas da superfície dessas células, sugerindo um controle regulatório na resposta imune adaptativa e na progressão da resposta imune (MANJUNATH et al. 1995; PEDRAZA-ALVA et al. 1998). Em experimentos de EAE, o transporte de células T para o SNC é fundamental para que as células CD4 positivas possam mediar a doença, de maneira que a deficiência de CD43 (SPN) reduz a severidade da EAE, quando comparado com CD43 selvagem (FORD et al. 2003). Interessantemente, pacientes com EMRR tratados com acetato de glatiramer, interferon beta ou natilizumab apresentaram redução de células B1 CD43 positivas, sugerindo que essas células possam, de fato, fazer parte da patologia da EM (ROVITUSO et al. 2014).

4.2 RETROINSERÇÕES *DE NOVO* EM INDIVÍDUOS PORTADORES DE EM

Cada evento de retroinserção mapeado em cada um dos 31 indivíduos deste estudo foi identificado de acordo com o gene que sofreu alteração. Encontramos um total de 13137 eventos distintos. A seguir criamos uma lista de genes alterados por eventos de retroinserção e para os pacientes consideramos apenas: i) genes portadores de eventos de retroinserção que pudessem ser considerados *de novo*, por estarem ausentes em seus pais biológicos e irmãos ii) eventos que fossem únicos de um dado paciente, ou seja, ausentes em todos os demais indivíduos saudáveis da nossa casuística. Enquanto a inclusão do primeiro filtro é mais óbvia, o segundo filtro visa excluir eventos de retroinserção que sejam comuns na população e provavelmente não tenham efeito sobre a EM. Considerando todos os pacientes tivemos um total de 1208 genes com DNR que permaneceram após a aplicação destes filtros. A exclusão de muitos eventos por estes filtros (>90%) sugere ser provável que vários dos eventos que permaneceram ainda podem ser achados casuais não relacionados com a doença. No entanto, é possível que ainda restem eventos relevantes dentro do contexto de nosso interesse. Para sermos mais rigorosos, mas ainda cientes de certas limitações, selecionamos para uma análise mais aprofundada apenas os genes que apresentaram eventos de DNR ausentes em todos os indivíduos saudáveis mas presentes em pelo menos dois pacientes diferentes. Deste modo, tivemos um *set* final de 109 genes (**Quadro 4**).

Quadro 4 – Genes com eventos DNR encontrados exclusivamente em portadores de EM e compartilhados por pelo menos dois pacientes.

Grupo	Gene	Descrição
g2 g5 g6	ARL15	ADP-ribosylation factor-like 15
g2 g5 g7	CP	Ceruloplasmin (ferroxidase)
g5 g6 g8	DCDC2	Doublecortin domain containing 2
g2 g5 g6	EPB41L1	Erythrocyte membrane protein band 4,1-like 1
g4 g7 g8	ICE1	Interactor of little elongation complex ELL subunit 1
g4 g5 g6	IL12A-AS1	Interleukin 12A (natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35)
g1 g2 g3	RNF44	Ring finger protein 44
g4 g5 g8	SLITRK6	SLIT and NTRK-like family, member 6
g3 g5 g6	TRAPPC12	Trafficking protein particle complex 12 [Homo sapiens]
g4 g5 g8	UMAD1	UBAP1-MVB12-associated (UMA) domain containing 1
g4 g6	ALG5	Asparagine-linked glycosylation 5, dolichyl-phosphate beta-glucosyltransferase homolog (S, cerevisiae)
g5 g6	AMPD3	Adenosine monophosphate deaminase (isoform E)
g5 g6	ANKRD28	Ankyrin repeat domain 28
g5 g8	ARMC4	Armadillo repeat containing 4
g3 g6	BCL2L1	BCL2-like 1
g2 g6	BCORL1	BCL6 co-repressor-like 1
g4 g6	BICC1	Bicaudal C homolog 1 (Drosophila)
g6 g8	C11orf70	Chromosome 11 open reading frame 70
g3 g5	C16orf96	Chromosome 16 open reading frame 96
g3 g5	C19orf35	Chromosome 19 open reading frame 35
g3 g8	C19orf57	Chromosome 19 open reading frame 57
g4 g8	C6orf89	Chromosome 6 open reading frame 89
g1 g5	C7orf60	Chromosome 7 open reading frame 60
g2 g4	C9orf41	Chromosome 9 open reading frame 41
g3 g4	CCT2	Chaperonin containing TCP1, subunit 2 (beta)
g4 g6	CDC23	Cell division cycle 23 homolog (S, cerevisiae)
g4 g5	CDH6	Cadherin 6, type 2, K-cadherin (fetal kidney)
g3 g4	CEP78	Centrosomal protein 78kDa
g4 g5	CERS3-AS1	Ceramide synthase 3
g2 g6	CLTC	Clathrin, heavy chain (Hc)
g1 g4	CSRNP3	Cysteine-serine-rich nuclear protein 3
g5 g6	CYP39A1	Cytochrome P450, family 39, subfamily A, polypeptide 1
g4 g6	DDI2	DDI1, DNA-damage inducible 1, homolog 2 (S, cerevisiae)
g4 g5	DDX52	DEAD (Asp-Glu-Ala-Asp) box polypeptide 52
g4 g8	DDX60	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60
g4 g5	DNM3	Dynamin 3
g2 g4	EFTUD2	Elongation factor Tu GTP binding domain containing 2
g2 g3	EIF5	Eukaryotic translation initiation factor 5
g4 g6	FAM227A	Family with sequence similarity 227, member A
g2 g6	FBXL13	F-box and leucine-rich repeat protein 13
g5 g8	FCAR	Fc fragment of IgA, receptor for
g4 g6	FFAR1	Free fatty acid receptor 1
g4 g5	FPGT	TNNI3 interacting kinase; fucose-1-phosphate guanylyltransferase
g3 g8	FRS2	Fibroblast growth factor receptor substrate 2
g4 g5	GABRA1	Gamma-aminobutyric acid (GABA) A receptor, alpha 1
g2 g3	GABRE	Gamma-aminobutyric acid (GABA) A receptor, epsilon
g2 g4	GAS7	Growth arrest-specific 7
g2 g3	GYS2	Glycogen synthase 2 (liver)
g4 g6	HLA-C	Major histocompatibility complex, class I, C; major histocompatibility complex, class I, B
g2 g6	HNRNPF	Heterogeneous nuclear ribonucleoprotein F
g3 g6	IFNGR1	Interferon gamma receptor 1
g5 g6	IL17RD	Interleukin 17 receptor D
g2 g6	ISL1	ISL LIM homeobox 1
g4 g5	KIAA0319	KIAA0319

Cont/ Quadro 4

Grupo	Gene	Descrição
g3 g5	LINC01169	Long intergenic non-protein coding RNA 1169
g2 g3	LMX1B	LIM homeobox transcription factor 1, beta
g5 g6	LOC100506990	Uncharacterized LOC100506990
g3 g6	LOC101927815	Uncharacterized LOC101927815
g4 g5	LOC102800447	Uncharacterized LOC102800447
g4 g5	LOC339539	Uncharacterized LOC339539
g4 g8	MDH1B	Malate dehydrogenase 1B, NAD (soluble)
g3 g8	MLXIPL	MLX interacting protein-like
g4 g6	MORN4	MORN repeat containing 4
g6 g8	MYNN	Myoneurin
g2 g8	NCBP1	Nuclear cap binding protein subunit 1, 80kDa
g4 g7	NDRG4	NDRG family member 4
g4 g5	NEK7	NIMA (never in mitosis gene a)-related kinase 7
g3 g3	NOTCH1	Notch homolog 1, translocation-associated (Drosophila)
g3 g4	PALM2	A kinase (PRKA) anchor protein 2; paralemm 2; PALM2-AKAP2 readthrough transcript
g2 g6	PANK2	Pantothenate kinase 2
g6 g8	PDGFRB	Platelet-derived growth factor receptor, beta polypeptide
g2 g4	PEX19	Peroxisomal biogenesis factor 19
g5 g8	PIBF1	Progesterone immunomodulatory binding factor 1
g3 g4	PNLDC1	Poly(A)-specific ribonuclease (PARN)-like domain containing 1
g4 g8	POU4F1	POU class 4 homeobox 1
g3 g8	PPP2R5A	Protein phosphatase 2, regulatory subunit B', alpha isoform
g3 g7	PRDM4	PR domain containing 4
g2 g8	PRDM9	PR domain containing 9
g2 g5	PSMC3	Proteasome (prosome, macropain) 26S subunit, ATPase, 3
g3 g5	PTPRR	Protein tyrosine phosphatase, receptor type, R
g2 g5	RAD54L	RAD54-like (S, cerevisiae)
g4 g6	RASSF2	Ras association (RalGDS/AF-6) domain family member 2
g6 g8	RIMS1	Regulating synaptic membrane exocytosis 1
g4 g5	RIMS3	Regulating synaptic membrane exocytosis 3
g2 g6	RNF212B	Ring finger protein 212B
g2 g6	RNF7	Ring finger protein 7
g5 g8	RPL3L	Ribosomal protein L3-like
g2 g3	RRN3	RRN3 RNA polymerase I transcription factor homolog (S, cerevisiae)
g3 g4	RRN3P2	RRN3 RNA polymerase I transcription factor homolog (S, cerevisiae) pseudogene
g4 g6	SFMBT2	Scm-like with four mbt domains 2
g2 g6	SGSH	N-sulfoglucosamine sulfohydrolase
g3 g8	SPG20	Spastic paraplegia 20 (Troyer syndrome)
g3 g6	STAC	SH3 and cysteine rich domain
g5 g6	STIM2	Stromal interaction molecule 2
g4 g8	STX3	Syntaxin 3
g3 g6	TBCE	Tubulin folding cofactor E
g4 g5	TMEFF2	Transmembrane protein with EGF-like and two follistatin-like domains 2
g5 g6	TMEM186	Transmembrane protein 186
g5 g6	TNC	Tenascin C
g2 g5	TNFAIP8L1	Tumor necrosis factor, alpha-induced protein 8-like 1
g4 g8	TOP3B	Topoisomerase (DNA) III beta
g5 g8	UBAP1L	Ubiquitin associated protein 1-like
g3 g4	USP47	Ubiquitin specific peptidase 47
g2 g6	VSIG4	V-set and immunoglobulin domain containing 4
g4 g5	WDR90	WD repeat domain 90
g3 g5	ZC3HAV1	Zinc finger CCCH-type, antiviral 1
g3 g8	ZFP1	ZFP1 zinc finger protein
g4 g5	ZNF197	Zinc finger protein 197
g2 g4	ZNF382	Zinc finger protein 382

Com o objetivo de explorar o possível papel destes eventos de um modo mais abrangente, fizemos análise funcional dos 109 genes através da ferramenta DAVID *Bioinformatic database*

(http://david.abcc.ncifcrf.gov/knowledgebase/DAVID_knowledgebase.html).

Inicialmente buscamos avaliar se esta lista de genes teria uma distribuição preferencial em alguma via associada a alguma doença humana. Como corte usamos os top 2 da lista com o menor p-valor, mas não utilizamos outros parâmetros, como o p-valor corrigido ou FDR. Essa análise resultou em 2 vias: “infecção por “*H. pylori*” e “Multiple sclerosis” (**Tabela 3**). O achado da via patológica de esclerose múltipla, em uma análise não dirigida ou supervisionada reforça o conceito ainda inexplorado de que eventos do tipo DNR podem ter uma relação direta com o desenvolvimento da EM.

Tabela 3 – Análise de vias de doença utilizando a ferramenta DAVID.

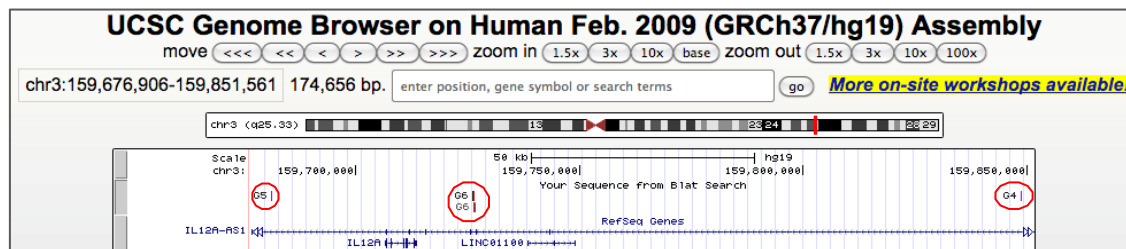
Termo	Count	P Valor	Genes	Fold Enrichment	Benjamini	FDR (%)
H, pylori infection	3	0,013	IFNGR1, IL12A, HLA-C	15,93	0,92	15,2
Multiple sclerosis	5	0,013	IFNGR1, BCL2L1, IL12A, HLA-C, SPG20	4,9	0,72	15,45

Legenda: O banco de dados de doença utilizado foi o *Genetic Association DB Disease-NIH*. Selecionamos as vias top 2 com menor p-valor.

A seguir avaliamos alguns dos genes onde identificamos DNRs e que fazem parte destas vias metabólicas, buscando dados da literatura que viessem a reforçar seu envolvimento com a EM. Cabe ressaltar a presença nesta lista de alguns genes de grande interesse na patogênese da EM, conforme discutido a seguir:

IL12A-AS1 (*IL12A antisense RNA 1*) – O transcrito canônico deste gene possui 2425 bases distribuídas em 10 exons. *IL12A-AS1* se localiza no cromossomo 3q25.33, em um bloco genômico de cerca de 300 mil bases na fita oposta do gene

codificador de IL12A, que ocupa uma pequena porção deste bloco, indicada na **Figura 12**, onde também indicamos as regiões onde observamos as DNRs.



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

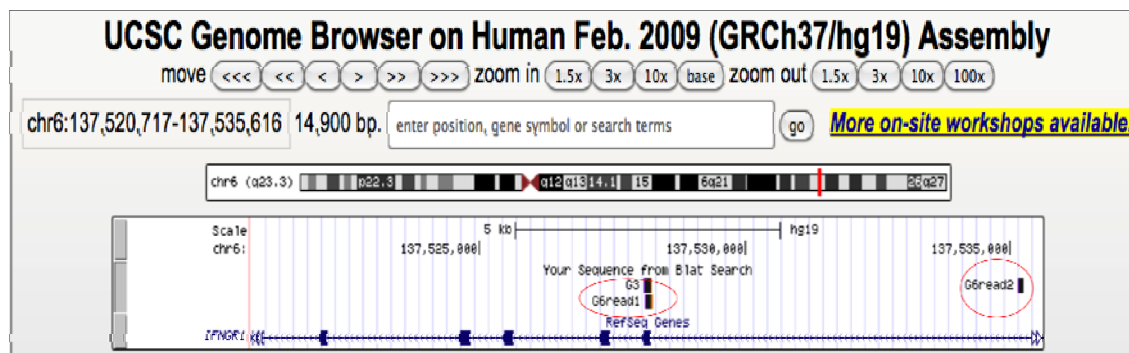
Figura 12 – Localização dos eventos DNR dos grupos quatro (G4), cinco (G5) e seis (G6) no genoma.

Não encontramos na literatura nenhuma descrição da atividade específica de IL12A-AS1. No entanto, a sua localização na fita oposta de IL12A pode ter um impacto funcional na expressão desta interleucina. Diversos artigos têm demonstrado que em muitos casos os transcritos anti-senso agem fazendo uma regulação negativa dos genes da fita oposta, por mecanismos diversos que incluem edição de RNA, RNA de interferência, interferência transcricional e outros (revisto em (WIGHT e WERNER 2013). Neste modelo, se a presença de um DNR em IL12-AS1, reduzir a atividade deste antisense, teríamos uma quantidade mais elevada de IL-12 nestes pacientes.

A relevância de IL-12 para a EM pode ser avaliada pelo resultado da busca no PubMed pelas palavras “IL-12” e “Multiple Sclerosis”, a qual retorna 436 artigos (Abril de 2015). IL-12 tem um importante papel na estruturação da resposta inflamatória durante infecções e na geração de células Th, que contribuem para a formação de uma resposta imune do tipo celular (HSIEH et al. 1993). O nível aumentado de IL-12 tem sido sugerido como um marcador de EM e a sua produção excessiva está envolvida na patogênese desta doença (BALASHOV et al. 1997; SEGAL et al. 2008). Recentemente, JANA et al. (2014), demonstraram que IL-12 induz a expressão de IL-7 na micróglia e em macrófagos. Por sua vez, IL-7 é uma citocina fundamental para o desenvolvimento da EM e de outras doenças autoimunes do sistema nervoso central e está aumentada nas lesões cerebrais de pacientes portadores de EM (KREMLEV et al. 2008).

IFNGR1 (*Interferon gamma receptor 1*) - Interferons estão envolvidos em processos diversos tais como defesa contra infecção viral, regulação do crescimento e resposta imune (SERRA et al. 2003; STÜRNER et al. 2014). Os interferons gama (IFNG) são produzidos por linfócitos T ou pelas células *natural killer* (DE ARAÚJO-SOUZA et al. 2015). Em nosso estudo, identificamos eventos DNR em *IFNGR1* nos pacientes dos grupos 3 e 6, como representado na **Figura 13**.

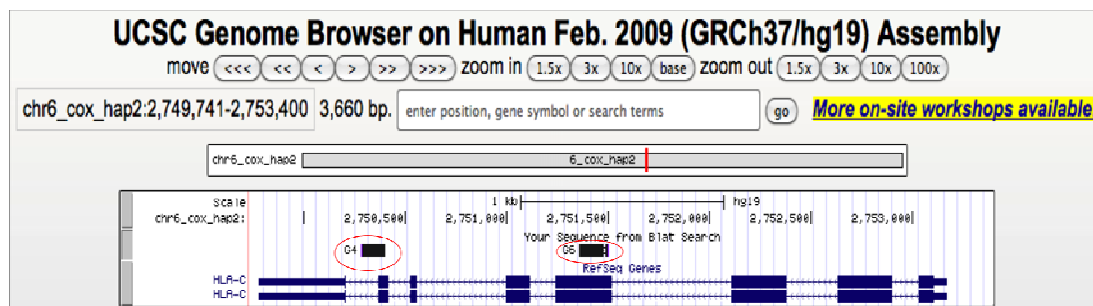
Em um trabalho desenvolvido pelo pesquisador Robert Axtell em 2010, foi avaliado o grau de resposta ao tratamento com interferon β no modelo animal EAE induzido com células Th1 e Th17 e em pacientes com EMRR. Em células T CD4+ estimuladas com interferon β (IFNB) foi demonstrado que a diminuição da frequência de IL-17 expressas em células T, com a neutralização do IFNG é reduzida de 60% (IFNB + IFNG) para 24% (apenas IFNB). Além disso, a abundância de células CD4+ IL-17+ é aumentada quando o IFNG é neutralizado, sugerindo seu papel na inibição da diferenciação de Th17. Posteriormente, para avaliar se o tratamento com IFNB requer IFNG para suprimir EAE, dois grupos de camundongos com EAE, um grupo controle e um grupo *Ifngr1*^{-/-}, foram tratados com Rebif (IFNB). No grupo controle houve redução dos sintomas após sete dias de tratamento e após a suspensão do tratamento a doença progrediu. No grupo de animais *Ifngr1*^{-/-} o tratamento não provocou efeito. Foi demonstrado ainda, que o soro de pacientes com EMRR não responsivos tem alta concentração de IL-17F em comparação com o grupo de pacientes responsivos ao tratamento (AXTELL et al. 2010). Interessantemente, identificamos evento DNR no receptor 1 do IFNG em dois grupos de pacientes com EMRR. Se esta inserção levar à uma perda de função haveria menor sinalização do IFNG, o que reduziria a inibição de IL-17. Corroborando o estudo citado acima, uma hipótese seria a de que pacientes que tem evento DNR nesse gene sejam menos responsivos ao tratamento com IFNB ou ainda que a evolução no quadro clínico seja mais acentuada devido à disfunção causada na via relacionada ao *IFNGR1*.



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 13 – Localização dos eventos DNR dos grupos três (G3) e seis (G6) no genoma.

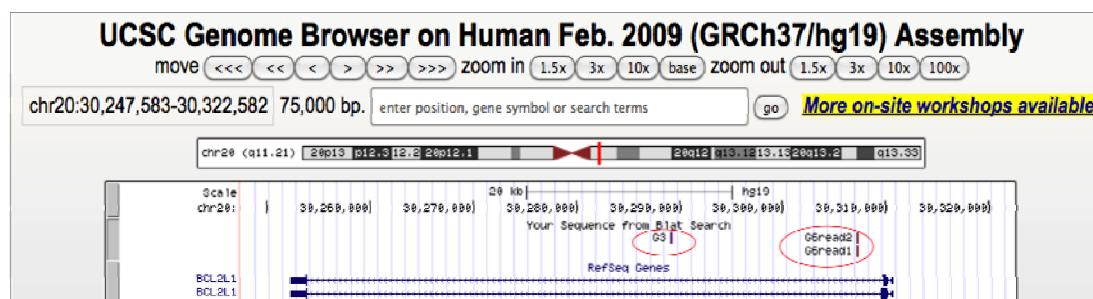
HLA-C (*Human Leukocyte Antigen C; Major Histocompatibility Complex, Class 1, C*). O conceito de que pelo menos um gene do complexo principal de histocompatibilidade (MHC) localizado no cromossomo 6p21 está associado à susceptibilidade em EM é bem estabelecido. Diferenças nos antígenos leucocitários humanos (HLA) entre indivíduos saudáveis e pacientes com EM e entre diferentes grupos étnicos vêm sendo identificadas há muitos anos, de modo que o extenso conteúdo de genes no MHC, uma região extremamente polimórfica e com alto desequilíbrio de ligação dificulta a associação de alelos que conferem susceptibilidade a doenças, como a EM (BERTRAMS et al. 1974; DYMENT et al. 2005). O antígeno leucocitário humano C (HLA-C) atua no sistema imune inato através da ligação com o receptor KIR ativando ou inibindo células *natural killer* e células T onde esse receptor é expresso. YEO et al. em 2007 demonstraram que o gene HLA-C ou um *locus* com baixo desequilíbrio de ligação com ele conferem efeito protetor (P valor = $3,3 \times 10^{-5}$) em EM. Corroborando com o trabalho acima, identificamos eventos de DNR no gene HLA-C nos pacientes dos grupos 4 (exon 4) e do grupo 6 (exon 7) (**Figura 14**), o que poderia alterar a expressão deste gene e talvez reduzisse um eventual efeito protetor.



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 14 – Localização dos eventos DNR dos grupos quatro (G4) e seis (G6) no genoma.

BCL2-L1 (*B-cell CLL/Lymphoma 2 Like 1*). A depleção seletiva de certos linfócitos no início do desenvolvimento de células T e B visa eliminar células auto-reativas, um processo importante na manutenção da homeostase do sistema imune e que parece ser a base de certas doenças autoimunes. Neste sentido, a família Bcl-2 tem um importante papel na modulação da apoptose e a abundância da proteína anti-apoptótica BCL2-L1 está associada ao direcionamento de células T para morte celular induzida por ativação (AICD) (BEHRENS e MUELLER 1997). Outrossim, altos níveis dessa proteína em pacientes com EM estão associados com resistência a AICD quando comparado com um grupo controle (WAICZIES et al. 2002). Em nosso grupo conseguimos identificar um mesmo evento DNR no intron 2 de BCL2-L1 em pacientes de dois grupos (grupo 3 e grupo 6), indicado na **Figura 15**.

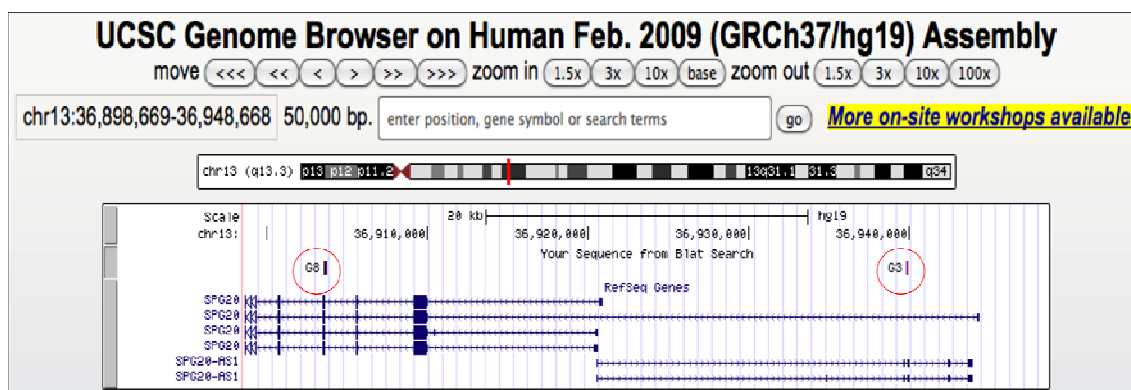


Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 15 – Localização dos eventos DNR dos grupos três (G3) e seis (G6) no genoma.

SPG20 (*Spastic paraplegia 20*) é um gene fundamental no desenvolvimento da síndrome de Troyer, uma doença autossômica recessiva caracterizada

principalmente por atraso mental, fraqueza espástica, disartria e baixa estatura. Esse gene codifica uma proteína multifuncional, com papel no tráfico intracelular do receptor fator de crescimento epidermal (EGFR), no *turnover* de corpos lipídicos, citocinese, regulação mitocondrial e sinalização de BMP (*bone morphogenetic protein*) por exemplo. Interessantemente, mutações em *SPG20* e elevada sinalização de BMP levam à neurodegeneração e ainda, reguladores do receptor BMP nas vias de tráfico endossomal e endocitose estão associados com doenças neurodegenerativas, incluindo a EM (HAFLER et al. 2007; KIM et al. 2010; BLACKSTONE 2012; NAHM et al. 2013). Neste gene, identificamos possíveis eventos de DNR no introns 1 (grupo-3) e no exon 4 (grupo-8) de portadores de EM (**Figura 16**).



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 16 – Localização dos eventos DNR dos grupos oito(G8) e três (G3) no genoma.

Para investigarmos vias específicas que pudessem ser impactadas por estes eventos avaliamos os 109 genes da **Quadro 5** no banco *Gene Ontology* (GO) também utilizando a ferramenta DAVID (categoria GOTERM_BP_ALL). Como corte utilizamos o P valor $\leq 0,01$ e o critério FDR (*False Discovery Ratio*) $\leq 5\%$, o que resultou em 6 processos biológicos associados. Com os mesmos 109 genes e mantendo apenas o critério do p-valor $\leq 0,01$, usamos a ferramenta KOBAS 2.0 (**KEGG Orthology Based Annotation System**, banco de dados GO) (<http://kobas.cbi.pku.edu.cn/program.run.do>) para validar *in silico* os processos biológicos encontrados anteriormente no GO. Assim, validamos *in silico* três

processos importantes para a EM: desenvolvimento de neurônios, desenvolvimento de projeções neuronais e organização de projeções celulares (**Quadro 5**).

Quadro 5 – Análise de enriquecimento de vias por ontologia utilizando a ferramenta DAVID e validação *in silico* através da ferramenta KOBAS 2.0.

Termo	Count (DAVID- NIAID/NIH)	Pvalor (DAVID- NIAID/NIH)	Genes (DAVID-NIAID/NIH)	Benjamini (DAVID- NIAID/NIH)	FDR (%) (DAVID- NIAID/NIH)	Count (Kobas 2.0)	Pvalor (Kobas 2.0)	Genes (Kobas 2.0)	Benjamini (Kobas 2.0)
GO:0048699 Generation of neurons	11	5,75E-04	NOTCH1, STX3, LMX1B, TBCE, PTPRR, DCDC2, POU4F1, ISL1, SLITRK6, GAS7, FRS2	0,44	0,9	14	0,034	NOTCH1, STX3, LMX1B, TNC, POU4F1, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1, FRS2	0,5
GO:0022008 Neurogenesis	11	0,001	NOTCH1, STX3, LMX1B, TBCE, PTPRR, DCDC2, POU4F1, ISL1, SLITRK6, GAS7, FRS2	0,4	1,5	14	0,05	NOTCH1, STX3, LMX1B, TNC, POU4F1, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1, FRS2	0,5
GO:0048666 Neuron development*	8	0,001	NOTCH1, STX3, LMX1B, TBCE, POU4F1, ISL1, SLITRK6, GAS7	0,46	2,8	13	0,007	NOTCH1, STX3, LMX1B, TNC, POU4F1, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1	0,45
GO:0030182 Neuron differentiation	9	0,001	NOTCH1, STX3, LMX1B, TBCE, PTPRR, POU4F1, ISL1, SLITRK6, GAS7	0,37	2,9	13	0,03	STX3, LMX1B, TNC, POU4F1, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1, NOTCH1	0,5
GO:0031175 Neuron projection development*	7	0,002	NOTCH1, STX3, TBCE, POU4F1, ISL1, SLITRK6, GAS7	0,34	3,2	12	0,006	STX3, TNC, POU4F1, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1, NOTCH1	0,43
GO:0030030 Cell projection organization*	8	0,002	DNM3, NOTCH1, STX3, TBCE, POU4F1, ISL1, SLITRK6, GAS7	0,38	4,4	14	0,016	STX3, TNC, ARMC4, NDRG4, TBCE, DCDC2, SPG20, KIAA0319, GAS7, SLITRK6, ISL1, DNM3, NOTCH1, POU4F1	0,5

Legenda: * Representa as vias que foram validadas.

Apesar de resultados promissores, cabe reforçar aqui que não houve tempo hábil para validarmos os eventos de DNR sugeridos pelas análises de bioinformática. Esta validação, a ser feita em genes de maior interesse, será concluída antes de submetermos para publicação os artigos desta dissertação.

4.3 AVALIAÇÃO DA HERANÇA DE ALELOS RAROS EM HOMOZIGOSE – SNPS RAROS

Hoje é consensual o conceito de que múltiplos fatores, genéticos e ambientais atuam em conjunto para o estabelecimento das doenças complexas. Deste modo, é claro que apenas a herança genética não é capaz, isoladamente, de explicar a etiologia destas doenças, incluindo a EM. Neste sentido, apesar do empenho de diversos grupos e de amplos estudos com grande casuística, utilizando a abordagem de GWAS (*do Inglês Genome Wide Association Study*) ou mesmo WGS (*do Inglês Whole Genome Sequencing*), ainda não é possível explicar totalmente a herdabilidade da maioria das doenças humanas. Sendo assim, é possível que uma combinação de variantes genômicas raras, com frequência muito baixa na população saudável (o que faz com que a maior parte das mesmas não seja incluída do painel de SNPs que compõe os estudos de GWAS) possa ter algum papel na gênese de doenças complexas e multifatoriais.

Apesar de este projeto ter sido inicialmente desenhado para o estudo de alterações de DNA *de novo*, ao analisarmos os nossos dados vislumbramos a possibilidade de avaliar a ocorrência de SNPs raros que, ao serem herdados em homozigose, pudessem ter alguma associação com a EM. É sabido que a maior parte dos SNPs humanos não-sinônimos (~45%) possui uma baixa frequência nas diversas populações (YUE e MOULT 2006), o que faz com que a ocorrência em homozigose, para alguns destes seja esperada, especialmente quando todo o exoma é avaliado. No entanto, a análise cuidadosa destes eventos, excluindo os que se encontram em homozigose nos pais saudáveis, e com especial atenção naqueles localizados dentro da região codificadora de proteínas e com potencial de alteração funcional de moléculas que façam sentido na clínica da doença, pode ser de interesse para o estudo da EM. Após analisarmos nossos dados buscando este tipo de evento, conforme descrito em Materiais e Métodos, identificamos um total de 24 SNPs raros herdados em homozigose (SRHH), com ambos os pais e irmãos heterozigotos, localizados nas regiões codificadoras e com potencial impacto funcional (**Quadro 6** – SNPs raros herdados em homozigose, localizados dentro da CDS de genes humanos e com impacto na sequência das proteínas, exclusivos de pacientes com EM. Dados gerados pela análise do sequenciamento do exoma de trios/quartetos em larga escala.

Quadro 6 - SNPs raros herdados em homozigose, localizados dentro de CDSs de genes humanos e com impacto na sequência das proteínas, exclusivos de pacientes com EM.

Trio	Coordenada genômica	Referência variante		dbSNP	MAF (100genomes)	ExAC (Broad Institute)	Efeitos/snpEff	Frequência da variante/cobertura				Predição AbraCadaBra	Predição PolyPhen (ExAC)	Predição SIFT (ExAC)	Média de cobertura
		nt	aa					paciente	mãe	pai	irmão				
2	chr6:29555738	C/A	p.SerTyr	rs61744949	0.0226	0,013	Moderado	100/49	42,37/59	51,61/93	57,14/28	B	B	B	62
3	chr1:110302450	C/T	p.Met35Ile	rs17598321	0.0112	0,0236	Moderado	100/34	50/36	42,31/26	-	B	PD	B	32
	chr11:95571347	C/G	p.Glu202Gln	rs61735578	0.0423	0,0361	Moderado	100/36	44,44/45	55,32/47	-	B	B	B	43
	chr19:13035722	T/C	p.Gln341Arg	rs35087277	0.0076	0,0211	Moderado	100/59	46/50	47,73/44	-	B	B	D	51
	chr19:13041442	G/A	p.Ala60Val	rs33925420	0.0038	0,0133	Moderado	100/44	57,89/38	56,82/44	-	B	B	B	42
	chr3:188925207	G/C	p.Ala12Pro	rs14775571	0.0018	0,0034	Moderado	100/31	47,37/38	35/40	-	B	B	B	36
	chr7:42005007	G/A	p.Pro122Ser	rs18149040	0.0024	0,0028	Moderado	100/53	29,03/31	35/40	-	B	B	B	35
	chr7:39454482	G/A	p.Met270Ile	rs45621451	0.0106	0,021	Moderado	100/40	37,14/35	61,64/39	-	B	PD	D	38
5	chr3:54925426	C/T	p.Thr732Met	rs112362895	0.0038	0,0076	Modificador, moderado	100/81	50,65/77	42,62/61	50/58	B	PRD	D	70
6	chr6:68713815	A/C	p.Met269Leu	rs36038900	0.0024	0,0035	Moderado	100/82	44,71/85	45,68/81	51,55/97	B	PD	B	87
7	chr10:95111492	T/A	p.Tyr1198Phe	rs12256834	0.0455	0,0178	Moderado	100/46	43,69/103	50,56/89	-	B	B	B	86
	chr19:17844227	G/A	p.Arg1005His	rs77093026	0.0272	0,0137	Moderado	100/48	51,28/78	40,98/61	-	B	PRD	B	63
	chr2:127808076	G/A	p.Thr532Met	rs112318500	0.0339	0,0091	Moderado	100/19	34,62/26	45,45/22	-	B	PD	D	23
	chr2:218682497	C/T	p.Ala1416Thr	rs61746065	0.0441	0,014	Moderado	100/52	41,67/48	55,56/54	-	B	B	B	52
	chr3:49159603	G/A	p.Arg1592Tyr	rs61729458	0.0126	0,0038	Moderado	100/17	51,61/31	37,5/40	-	B	PRD	D	30
	chr6:24828449	C/T	p.Val882Ile	rs61741706	0.0228	0,01	Moderado	100/60	30,16/63	50/68	-	B	B	B	64
	chr7:30830902	C/T	p.Ser262Leu	rs17159453	0.0367	0,0109	Modificador, moderado	100/44	38,18/55	49,02/51	-	B	B	B	50
	chr7:45125486	T/C	p.Glu68Val	rs116795021	0.0266	0,0066	Moderado	100/66	38,57/70	52,78/72	-	B	PRD	D	69
	chr9:35043935	C/G	p.Leu437Val	rs35523761	0.0321	0,0093	Moderado	100/28	43,9/41	52,78/23	-	B	B	B	31
8	chr17:74018597	T/C	p.Tyr168Cys	rs10445216	0.0082	0,01798	Moderado	100/18	40,62/32	43,48/23	-	B	PD	D	24
	chr17:74019453	G/A	p.Arg134Cys	rs150149800	0.0082	0,0146	Moderado	100/36	57,14/63	54/50	-	B	PD	D	50
	chr17:74073094	G/T	p.Tyr249Leu	rs61745847	0.0026	0,0063	Moderado	100/28	56,1/41	51,52/33	-	D	PRD	D	34
	chr3:160804167	C/T	p.Asp126Asn	rs2321257	0.0519	0,0456	Moderado	100/73	35,62/73	41,3/92	-	B	B	B	79
	chr6:152871865	T/G	p.Lys3874Thr	rs13210127	0.0543	0,0363	Moderado	100/95	40,54/74	55,10/98	-	B	B	D	89

Legenda: O asterisco (*) indica os genes selecionados para validação. Nas duas últimas colunas da tabela, as colunas de predição de patogenicidade representam: B = benigna; D= deletério; PRD= provavelmente deletério e PD= possivelmente deletério.

Posteriormente, tendo como base: i) a cobertura de sequências indicando os eventos; ii) a atividade predita e/ou a associação dos genes alterados com processos do SNC; iii) a classificação dos eventos como patogênicos em pelo menos uma das ferramentas de predição de patogenicidade (SIFT ou Polyphen) e iv) as atividades biológicas já conhecidas para os genes, que os tornavam mais promissores como relevantes para a EM, selecionamos 7 destes 24 SRHHs para validação pelo método *Sanger*. Os genes selecionados, assim como as sequências dos iniciadores e os tamanhos esperados dos *amplicons* são apresentados no **Quadro 7**.

Quadro 7 – Genes, SNPs e características dos iniciadores usados na validação dos polimorfismos raros.

Grupo	Coordenadas genômicas	Gene	Iniciadores	Fragmento pbs	Validação	Predição Abracadabra	Predição Polyphen (ExAC)	Predição SIFT (ExAC)
3	Chr11:95571347	MTMR2/rs61735578	GAGTTGCCAGAATTGCCTAAG CCACTCATGAACCTAGCAGAC	609bp	1	B	B	B
3	Chr7:99454482	CYP3A43/ rs45621431	CAGGAGGGGTGCTTAGGAC CTGCAATGTGGAAGAAATTCTC	268bp	1	B	PD	D
7	Chr2:127808076	BIN1/ rs112318500	CAGCAAGGTGGCATCTGAG GAAGCGGGTATGGGGAC	498bp	1	B	PD	D
7	Chr2:218682497	TNS1/rs61746065	CACCAGACTCCGACG CTTCATCAGCAGGAGAGGTG	263bp	2	B	B	B
7	Chr3:49159603	LAMB2/rs61729458	CTCCATCCCAGCTTCAGC CTGCCTTCTGTTTCTCATCCTC	262bp	1	B	PRD	D
8	Chr17:74073094	GALR2/ rs61745847	GTTCTCGGCTGACCTACG CAGCGCGTAAACGATGG	279bp	1	D	PRD	D
8	Chr6:152671865	SYNE1/ rs13210127	CATCTCTTGACATCCATGGG CT CCTTACCTTTCTCTGCTGC	247bp	1	B	B	D

Legenda: 1- validado; 2- inconclusivo (repetir Sanger).

Dentre os sete SRHH selecionados para validação pelo método Sanger, seis foram validados totalizando 85,7% (**Figura 17**). Um deles, o rs61746065/*TNS1* apesar de sugerir a validação por *Sanger*, o sequenciamento do paciente foi inconclusivo, por isso iremos repeti-lo. Os genes e seus SRHHs validados de maior interesse serão discutidos a seguir.

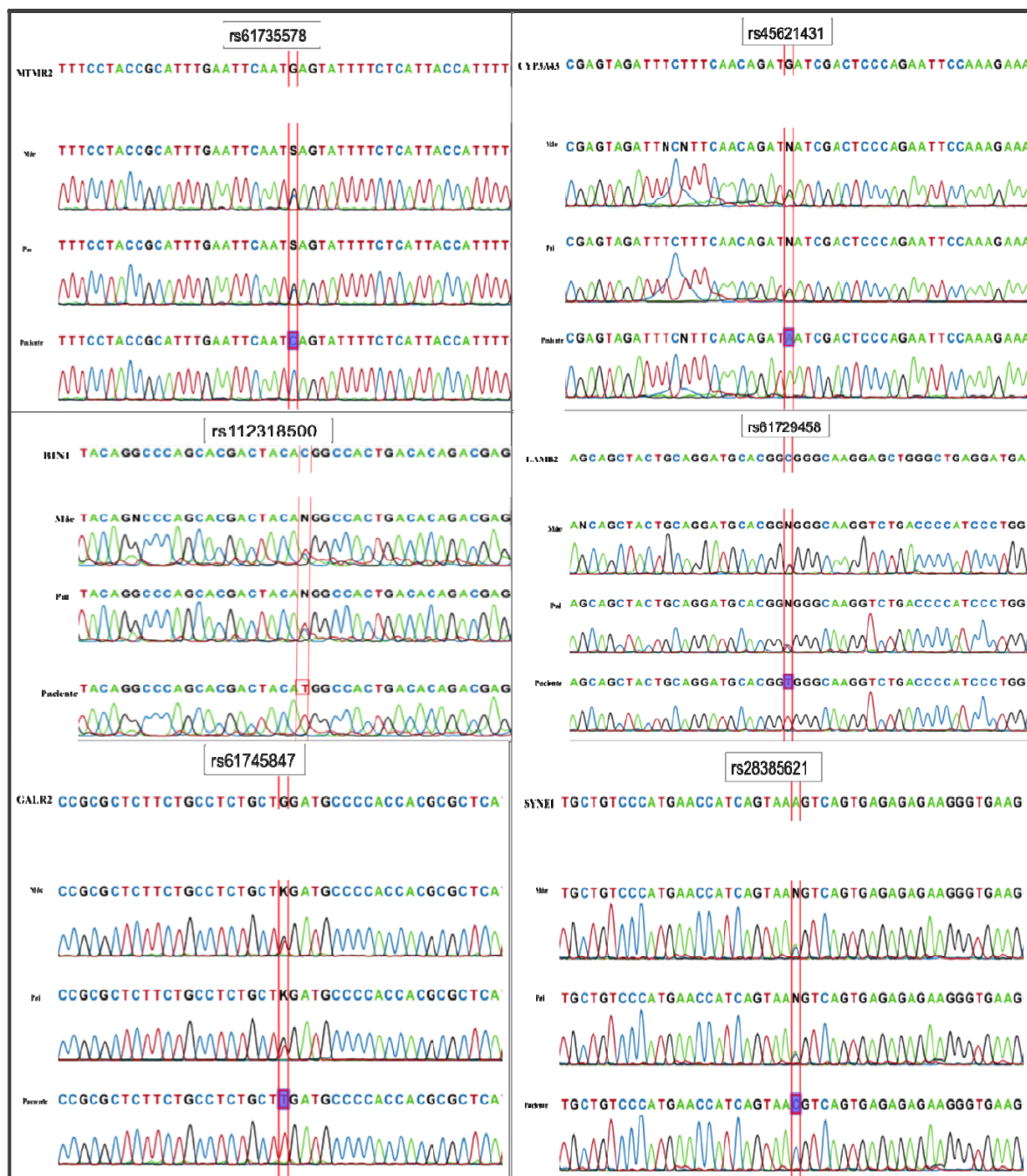


Figura 17 – Cromatogramas de sequenciamento Sanger representativos das regiões de interesse, contendo os SNPs raros, herdados em homozigose, validados neste estudo.

LAMB2 (Laminina B2; rs61729458) - Lamininas são glicoproteínas da membrana basal de diversos tipos celulares, caracterizadas por formar heterotrímeros compostos de três produtos gênicos independentes - as cadeias α , β e γ . Essas glicoproteínas da matriz extracelular podem se ligar a diversas classes de proteínas de membrana. O gene humano *LAMB2* (laminin β 2), localizado no chr3p21 codifica lamininas de cadeia β 2 com expressão já descrita na membrana basal sináptica. Essas lamininas tem um papel importante na integridade estrutural das sinapses da junção neuromuscular (sistema nervoso periférico) e nas sinapses do terminal fotorreceptor e do hipocampo (LOPES et al. 1985; LIBBY et al. 1999; NISHIMUNE et al. 2004; EGLES et al. 2007).

A retina é um modelo de sistema experimental bem estabelecido para avaliar os processos do SNC, de produção de diferentes tipos celulares, estabelecimento e manutenção de conexões entre essas células. A análise de sinapses em fotorreceptores da retina de camundongos depletados de lamininas de cadeia β 2 demonstra uma redução de 50% nos fotorreceptor. Ademais, a distribuição irregular de proteínas associadas à morfofisiologia sináptica na retina e a presença de má-formações de estruturas associadas a vesículas sinápticas sugerem um desequilíbrio na anatomia dos elementos sinápticos (LIBBY et al. 1999). Paralelamente, células hipocâmpais de camundongos depletados de *Lamb2* possuem alterações pré- e pós-sinápticas, reforçando o papel desta laminina no desenvolvimento e manutenção da função sináptica (EGLES et al. 2007; LIBBY et al. 1999). Em nosso estudo, observamos no paciente do grupo 7 o SRHH rs61729458 no gene *LAMB2* (**Figura 18**).

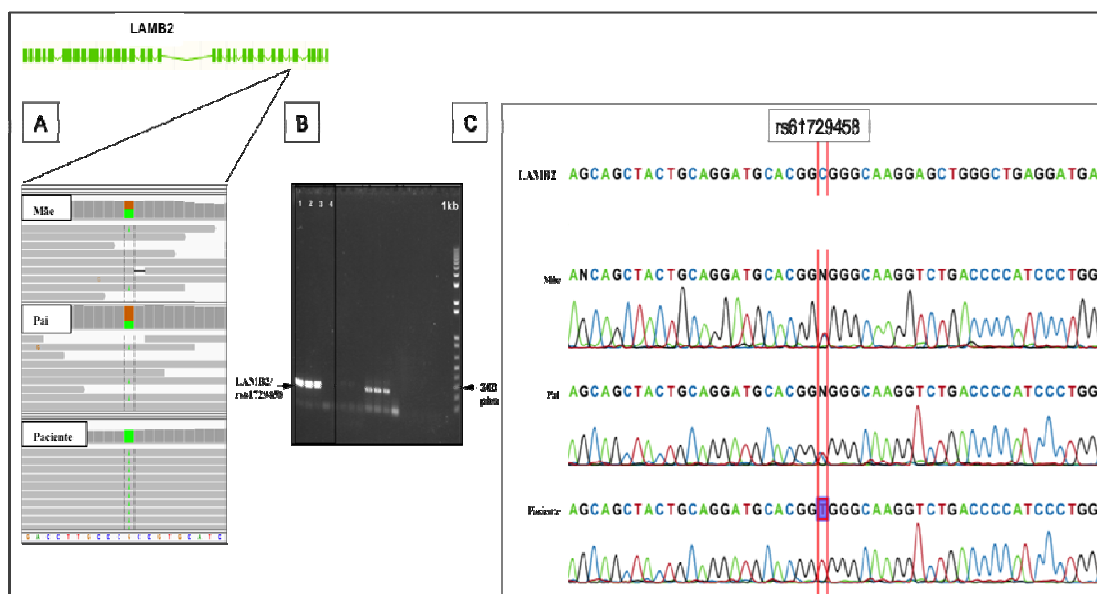
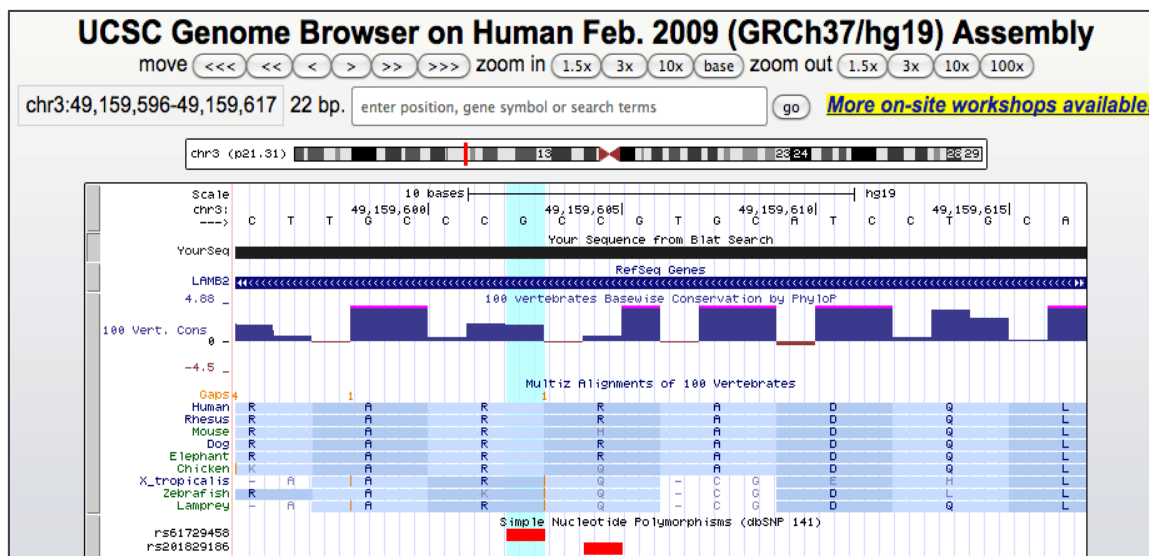


Figura 18 – Validação pelo método Sanger do candidato SNP *raro* rs61729458 no gene *LAMB2*. A) representa as *reads* obtidas pelo sequenciamento do exoma de trios (mãe, pai e paciente, respectivamente). Nota-se que somente as *reads* do paciente estão apresentadas em uma única cor (verde), mostrando a homozigose T/T causada por uma variante *rara*. B) Eletroforese em gel de agarose 1% mostrando a amplificação por PCR do fragmento onde está a alteração no gene *LAMB2*. 1, 2 e 3 representam o paciente a mãe e o pai e 4 representa o controle negativo. C) Cromatogramas obtidos pelo método Sanger validando as alterações sugeridas por NGS. No paciente observamos alteração em homozigose e nos pais em heterozigose.

Nesse polimorfismo, ocorre uma troca de arginina para triptofano no resíduo 1592 em uma região conservada em mamíferos de acordo com o UCSC (**Figura 19**) com uma frequência alélica de 1% na população global de acordo com o *1000 genomes*. Similarmente, no ExAC esse polimorfismo foi predito como patogênico (SIFT e Polyphen) com uma frequência alélica de apenas 0,38% (462/120710) e identificado em homozigose em apenas oito indivíduos em 60.355 indivíduos avaliados. Cabe notar que estes oito indivíduos homozigotos são de origem Africana.



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 19 – Conservação da arginina entre as espécies na posição onde ocorre o SNP rs61729458 de *LAMB2*

CYP3A43 (Cytochrome P450, Family 3, Subfamily A, Polypeptide 43; rs45621431) - As proteínas citocromo P450 pertencem à família de enzimas metabolizadoras de diversas drogas. Em nosso estudo observamos o SRHH rs61729458 no gene *CYP3A43* no paciente do grupo 3 (**Figura 20**), onde ocorre uma troca de aminoácido metionina para isoleucina no resíduo 275 predito como patogênico pelo ExAC, com uma frequência alélica de 1% no *1000genomes* e 2,18% (2649/121286) no ExAC. Esse membro da subfamília CYP3A foi caracterizado recentemente e está expresso em níveis relativamente elevados no cérebro em relação ao fígado de indivíduos com desordens neurológicas ou psiquiátricas. Essa enzima ainda é pouco estudada e o acesso à sua expressão no cérebro é dificultado devido à necessidade de autópsia, mas acredita-se que sua função esteja relacionada à metabolização de drogas no SNC (DOMANSKI et al. 2001; AGARWAL et al. 2008).

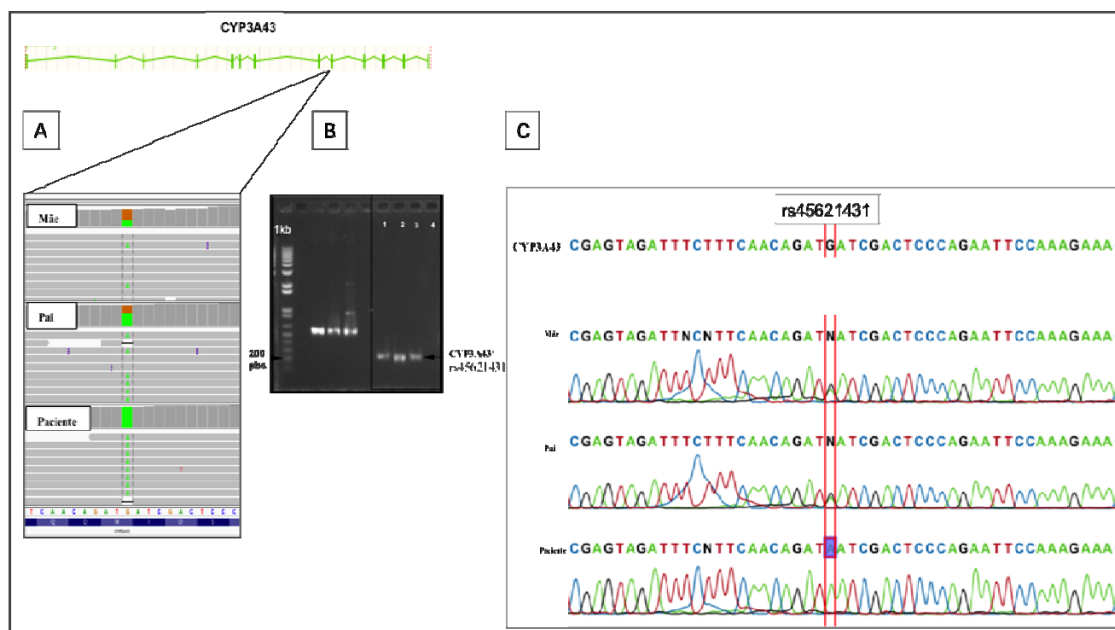


Figura 20 – Validação pelo método Sanger do candidato SNP *raro* rs45621431 no gene *CYP3A43*. A) representação as *reads* obtidas pelo sequenciamento do exoma de trios (mãe, pai e paciente, respectivamente). Nota-se que somente as *reads* do paciente estão apresentadas em uma única cor (verde), mostrando a homozigose A/A. B) Eletroforese em gel de agarose 1% mostrando a amplificação por PCR do fragmento onde está a alteração no gene *BIN1*. As canaletas 1, 2 e 3 indicam as bandas obtidas para o paciente, a mãe e o pai. Em 4 temos o controle negativo. C) Cromatograma *Sanger* mostrando a validação das sequências contendo a variante em homozigose, no paciente e em heterozigose nos pais.

BIN1 (Briding Integrator 1; rs112318500) - Observamos em nosso estudo o SRHH rs112318500 no gene *BIN1 (Briding Integrator 1)* no paciente do grupo 7, onde ocorre uma troca de bases G/A alterando o aminoácido de treonina para metionina no resíduo 532 (**Figura 21**). Esse gene codifica isoformas de proteínas adaptadoras nucleoplasmáticas que quando expressas no SNC podem interagir com dinaminas, anifisinas 1, endofilinas, sinaptojaninas entre outras (TAKEI et al. 1999; PICAS et al. 2014). Sua expressão é elevada em músculo esquelético e também nos diversos tecidos cerebrais analisados pelo portal GTEX (<http://www.gtportal.org/>) (**Figura 22**).

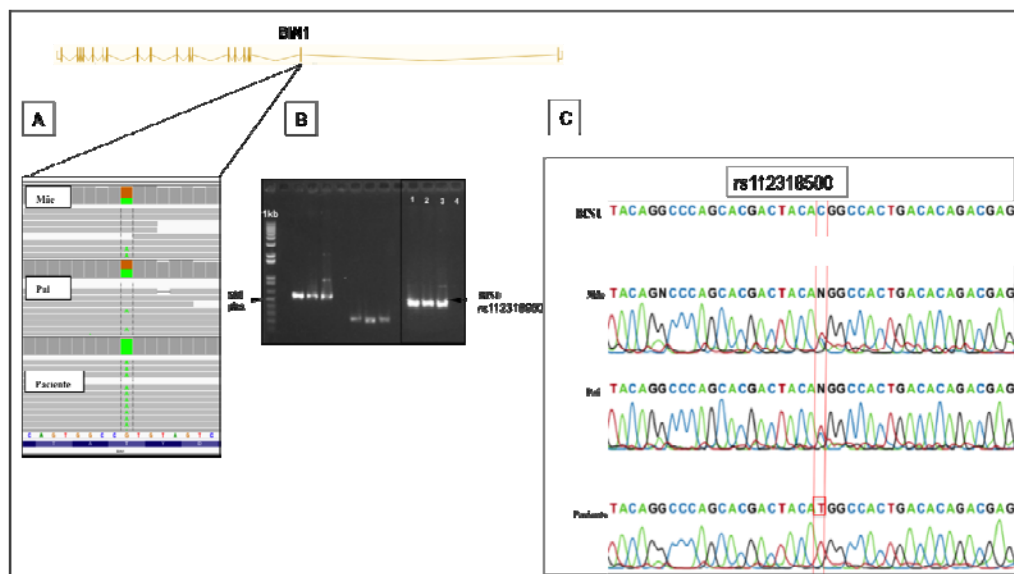


Figura 21 – Validação pelo método Sanger do SNP rs112318500 do gene *BIN1* em homozigose. A) representação das *reads* obtidas pelo sequenciamento do exoma de trios (mãe, pai e paciente, respectivamente). Nota-se que somente as *reads* do paciente estão apresentadas em uma única cor (verde), mostrando a homozigose A/A. B) Eletroforese em gel de agarose 1% mostrando a amplificação por PCR do fragmento contendo a alteração no gene *BIN1*. As canaletas 1, 2 e 3 indicam as bandas obtidas para o paciente a mãe e o pai e 4 representa o controle negativo. C) Cromatograma *Sanger* mostrando a validação dos achados de NGS onde temos as sequências contendo a variante em homozigose, no paciente e em heterozigose nos pais.

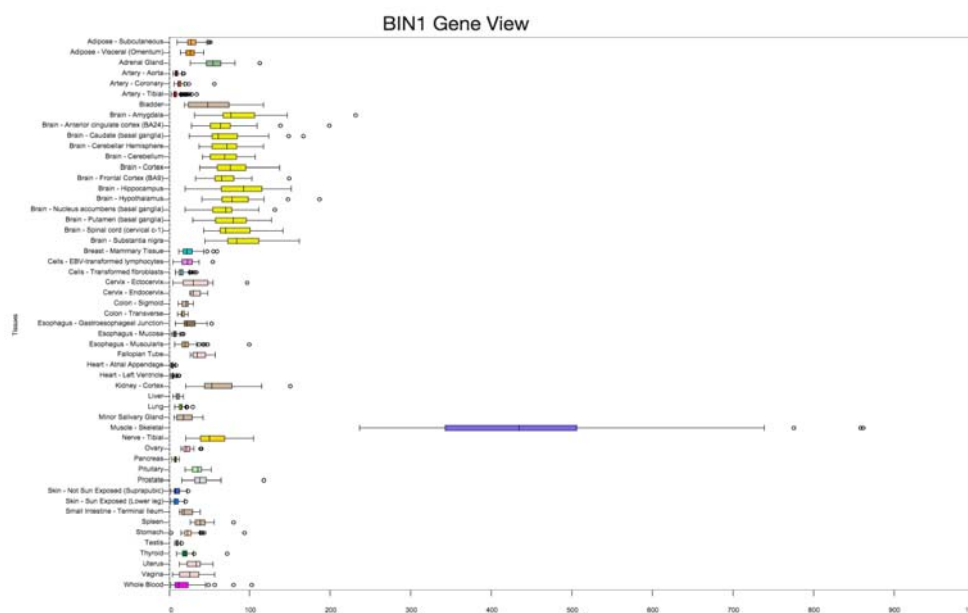
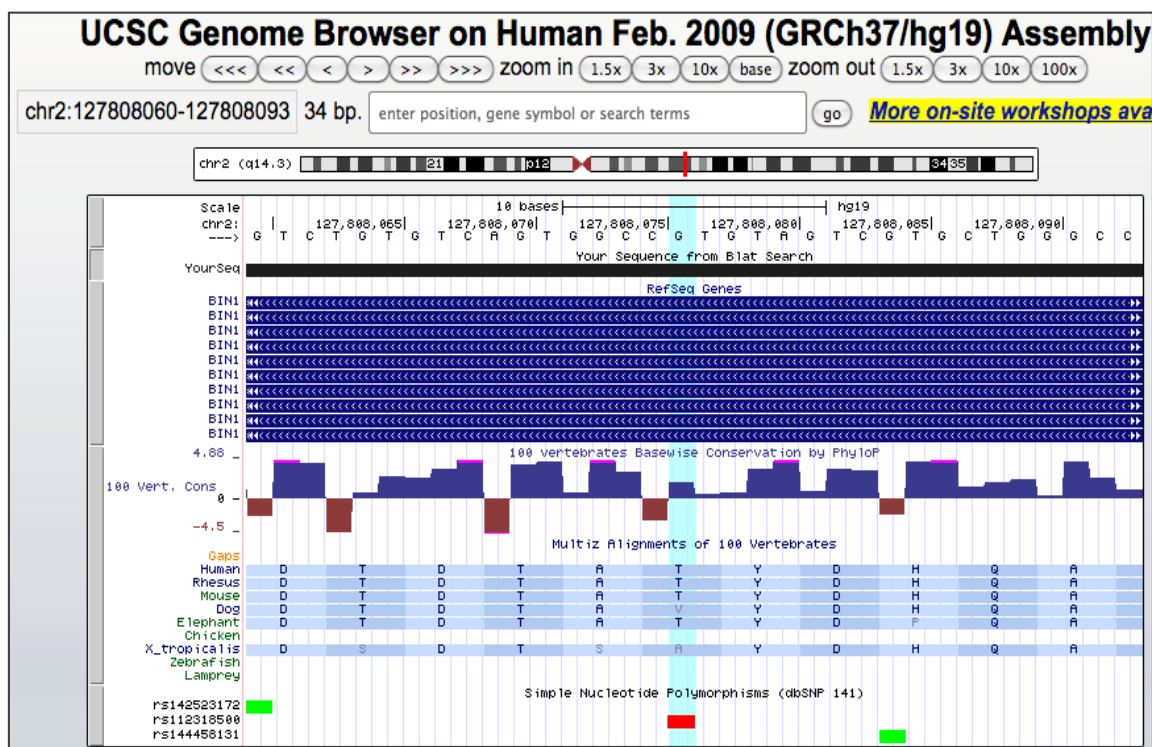


Figura 22 - Expressão do gene *BIN1* nos diversos tecidos analisados pelo portal GTEX (<http://www.gtexportal.org/>).

Esse polimorfismo tem uma baixa frequência alélica 3,3% e 0,9% no 1000genomes e ExAC respectivamente. Através da ferramenta ExAC, observamos ainda que o transcrito que representa essa alteração foi predito como deletério pelo SIFT e possivelmente deletério pelo Polyphen. A análise de conservação de sequências entre as espécies, usando o UCSC, não demonstra nenhuma espécie com uma metionina na região de interesse e ainda demonstra que a treonina é conservada na maioria dos mamíferos (**Figura 23**).



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 23 - Conservação da treonina entre as espécies na posição onde ocorre o SNP rs112318500.

BIN1 e seu parálogo *AMPH* codificam duas proteínas similares, anfifisina 2 e 1, respectivamente. As anfifisinas são altamente concentradas nos nervos terminais, especialmente na substância branca, formando homo ou heterodímeros que são capazes de interagir com proteínas envolvidas no processo de endocitose das vesículas sinápticas. Anfifisinas podem conectar os processos de formação de lipossomos em túbulos, propiciado pela dinamina com a formação de estruturas vesiculares revestidas por clatrina. Essa ponte entre clatrina e dinamina gerada pela

anfifisina é possivelmente dependente de interações proteína-proteína e de importantes interações da anfifisina com os lipídios da membrana celular (TAKEI et al. 1999; PICAS et al. 2014). Para o gene *BIN1* já foram descritos 970 SNPs, sendo apenas 11 não sinônimos, além disso é um forte *locus* associado a doença de Alzheimer de início tardio (LOAD) com possível mecanismo associado à função sináptica. O polimorfismo rs112318500 foi avaliado em dois estudos de associação com a AD. Em um desses estudos esse SNP não foi observado na casuística (HOLTON et al. 2013) e no outro não mostrou associação (MASOODI et al. 2013). Similarmente ao que ocorre com a miopatia centronuclear, onde a interação entre a dinamina e a anfifisina é necessária para o funcionamento normal dos músculos e posicionamento dos núcleos (NICOT et al. 2007), nossos achados reforçam a hipótese de que a alteração em homozigose desse gene no SNC provoca distúrbios na formação/ciclagem de vesículas sinápticas que podem estar envolvidas em processos neurodegenerativos (CARRASQUILLO et al. 2011; HU et al. 2011; HOLTON et al. 2013).

GALR2 (Galanin Receptor 2 - dbSNP rs61745847) – A galanina é um pequeno neuropeptídeo produzido no sistema nervoso central (essencialmente na hipófise e no hipotálamo – **Figura 24**). Sua presença em grânulos neurosecretórios sugere seu possível papel como neurotransmissor (KAPLAN et al. 1988). A galanina é um importante mensageiro entre o sistema nervoso central e o periférico e seu papel já foi descrito na modulação de diversas funções fisiológicas que incluem a regulação da liberação do hormônio do crescimento, da prolactina e de dopamina, a inibição da liberação de insulina e somatostatina, a contração da musculatura lisa gastro-intestinal e genito-urinária. A galanina parece estimular a ingestão de alimentos e a sinalização por seus receptores parece disparar processos relacionados com o desenvolvimento de diabetes e com a percepção de dor (MECHENTHALER 2008; YU et al. 2013).

As atividades da galanina são mediadas por seus três receptores conhecidos: GALR1, GALR2 e GALR3, membros da família de *G-protein coupled receptors* (GPCRs), do tipo *rhodopsin-like*. Estes receptores apresentam alta conservação de sequência inter-específica e são expressos em diferentes populações neuronais. A maior parte dos efeitos conhecidos da galanina é mediada pelo seu receptor GALR2

que é bastante expresso no tecido gastro-intestinal e no cérebro tem sua maior expressão no hipotálamo (**Figura 25**). Interessantemente, o hipotálamo é uma das áreas cerebrais com maior quantidade de lesões em pacientes com esclerose múltipla. Qiu et al. 2011, usaram ressonância magnética para avaliar os cérebros de 105 pacientes com EM e 12 pacientes com mielopatia longitudinal extensa (MLE) e encontraram lesões hipotalâmicas em 13,3% dos pacientes com EM e em nenhum caso de MLE (QIU et al. 2011). Posteriormente, HAIDER et al. (2014) avaliaram em autópsias os cérebros de 75 pacientes com EM e 12 controles. As regiões cerebrais com maior frequência de encontro de lesões desmielinizantes locais foram o núcleo caudato e o hipotálamo, o que pode ajudar a explicar o desequilíbrio do eixo hipotalâmico-pituitário-adrenal na esclerose múltipla (LYUBETSKA et al. 2015).

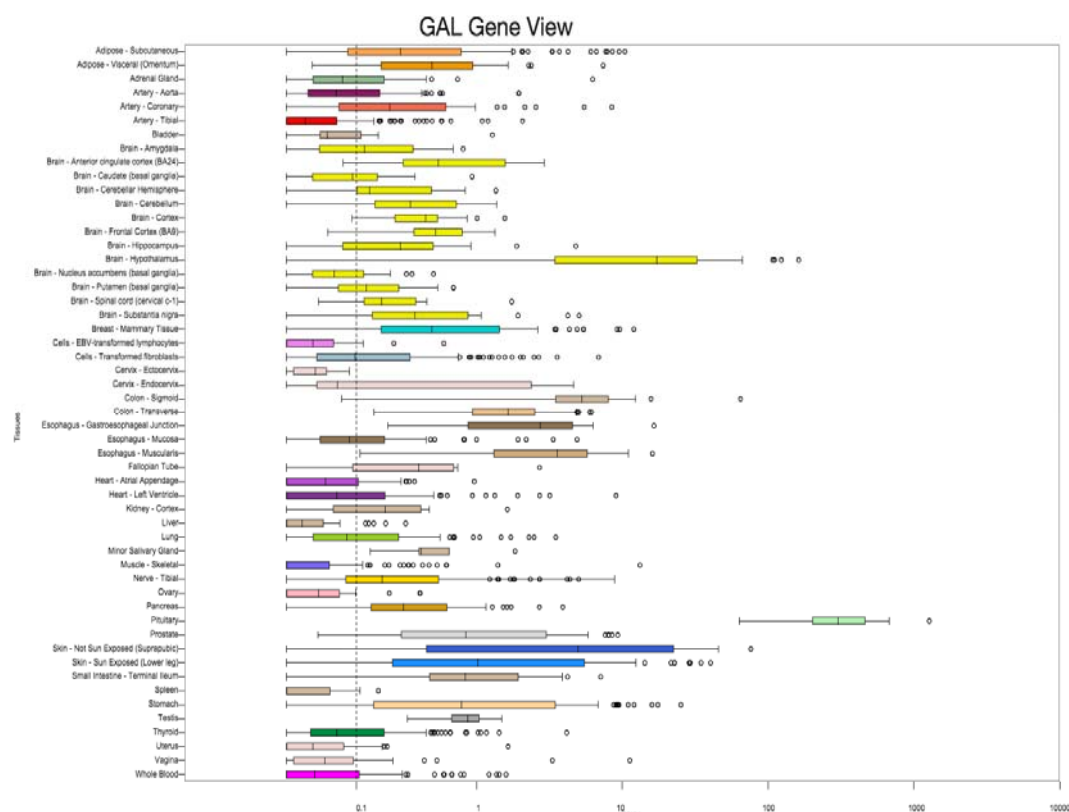


Figura 24 – Dados de expressão expressão normalizada dos mRNAs codificadores de galanina (GAL). A expressão foi determinada por RNA-Seq a partir de 53 tecidos saudáveis obtidos de 900 diferentes doadores em um total de 1421 amostras sequenciadas. Os dados foram obtidos a partir do *Genotype-Tissue Expression Project (GTEx)*, conduzido e disponibilizado pelo BROAD Institute (<http://www.gtexportal.org/home/>).

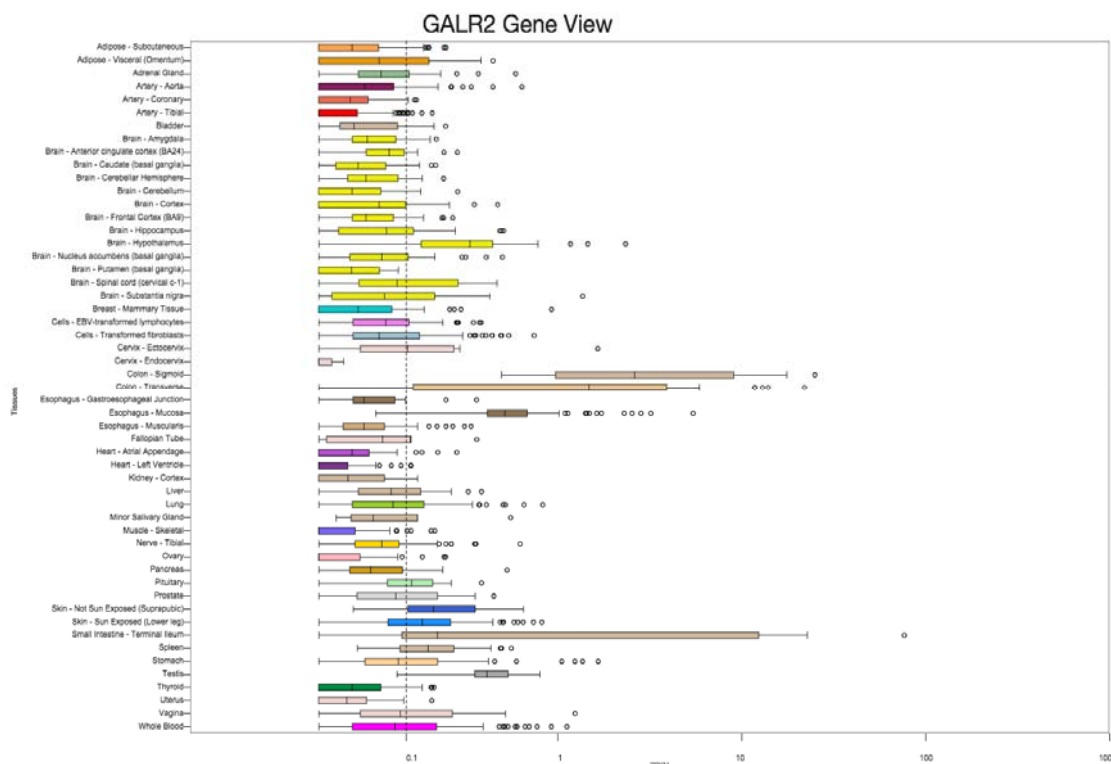
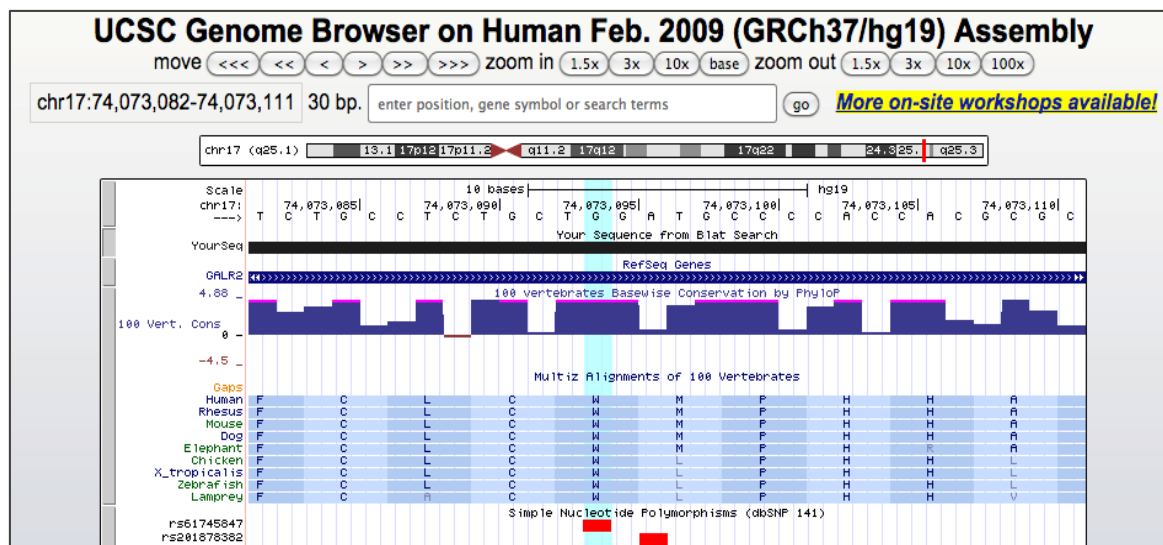


Figura 25 – Os gráficos apresentam a expressão normalizada dos mRNAs codificadores DE GALR2 em diferentes tecidos humanos. A expressão foi determinada por RNA-Seq a partir de 53 tecidos saudáveis obtidos de 900 diferentes doadores em um total de 1421 amostras sequenciadas. Os dados foram obtidos a partir do *Genotype-Tissue Expression Project* (GTEx), conduzido e disponibilizado pelo BROAD Institute (<http://www.gtexportal.org/home/>).

O SNP rs61745847: validação, consequências e frequência em diferentes populações - Em nossas análises de SRHHs identificamos um SNP raro (rs61745847) localizado no gene *GALR2*. Esta alteração promove a troca do códon TGG-249-TTG (Trp249Leu), uma região de alta conservação protéica que sugere a importância da sequência primária para a manutenção atividade funcional deste receptor. Uma consequência disto é que uma alteração em homozigose dessa região talvez tenha efeitos funcionais (**Figura 26**). De fato, predições de impacto funcional pelos programas Abracadabra (Kitajima & Schlesinger) e SIFT (<http://sift.jcvi.org/>) avaliam esta troca como patogênica, enquanto que a ferramenta Polyphen (<http://genetics.bwh.harvard.edu/pph2/>) a indica como provavelmente danosa.



Fonte: UCSC Genome Browser Home (genome.ucsc.edu/).

Figura 26 – Conservação do aminoácido triptofano (W) entre as espécies na posição onde ocorre o SNP rs61745847.

Inicialmente buscamos validar este SRHH por *Sanger* de modo a avaliar se o mesmo não se tratava de um artefato do sequenciamento NGS. Como pode ser visto na **Figura 27**, a validação foi bem sucedida e confirmou o status do SNP nos três indivíduos deste trio.

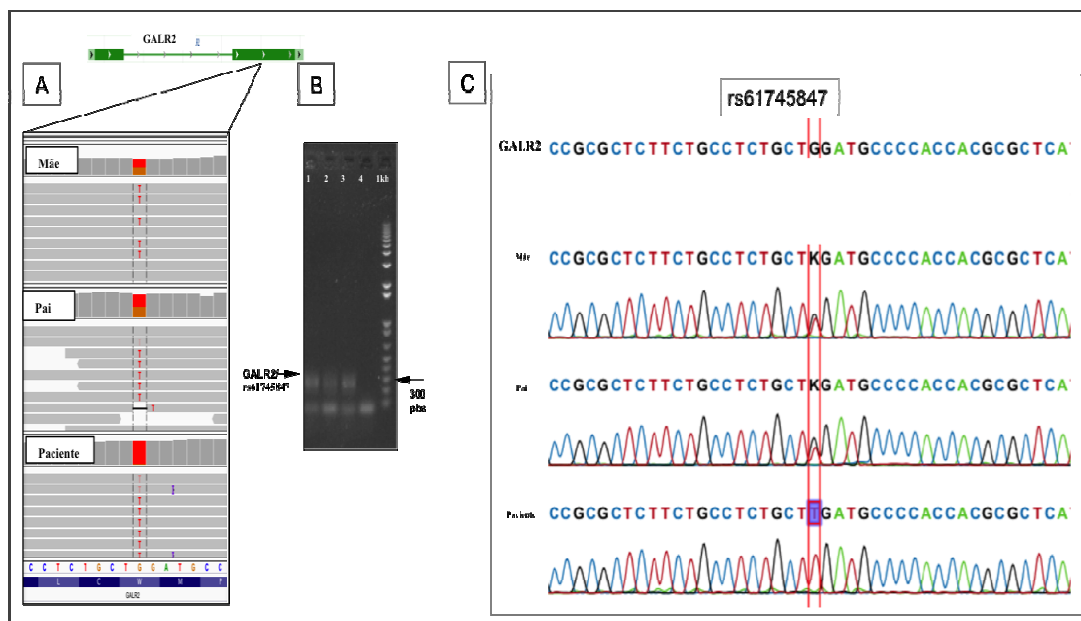


Figura 27 – Validação pelo método Sanger do candidato SNP *raro* rs61745847. A) representação gráfica localização e do conteúdo das sequências obtidas pelo sequenciamento NGS do exoma deste trio (mãe, pai e paciente, respectivamente). Nota-se que somente as *reads* do paciente estão apresentadas em uma única cor (vermelho), mostrando a homozigose T/T causada por uma variante *rara*. B) eletroforese em gel de agarose 1% mostrando os produtos de PCR da região de interesse que contém o SNP e interesse. Respectivamente, as canaletas 1-4 indicam a posição do gel com amplicons do paciente, da mãe e do pai e do controle negativo (sem DNA molde). c) Cromatogramas *Sanger* representando fragmentos da sequência de interesse contendo a variante em homozigose, no paciente, e sequências em heterozigose em seus pais.

A seguir avaliamos a frequência deste polimorfismo em diversos bancos de dados, e observamos uma frequência do alelo variante inferior a 1% para todas as populações em todos os bancos. As frequências observadas encontram-se abaixo, no **Quadro 8**.

Quadro 8 – Frequências observadas para o alelo variante (T) do SNP rs61745847.

<i>Banco</i>	<i>Frequência observada (%)</i>	<i>Alelos avaliados</i>	<i>Número indivíduos com genótipos TT</i>	<i>Observação</i>
1000Genomes	0,52	2491	0	Inclui todos os indivíduos do projeto
	0,15	1322	0	<i>Africanos</i>
	0,43	694	0	<i>Americanos</i> ²
	0	1008	0	<i>Ásiaticos</i>
dbSNP	0,4	4446	-	<i>NHLBI Exome Sequencing Project</i>
ExAC Browser ¹	0,63%	116266	4	-

Legenda: 1-Exome Agregation Consortium (ExAC), Cambridge, MA, USA (exac.broadinstitute.org) [acessado em 07/05/2015]; 2 - A população de Americanos do 1000genomes Project é composta por Colombianos, Mexicanos, Peruanos e Porto-Riquenhos.

Uma possível crítica para o achado seria uma presença elevada deste SNP na população brasileira. Buscando avaliar a frequência populacional em um grande N amostral e de modo rápido e viável, coletamos DNA humano, anonimizado, a partir de amostras de *swab* do metrô de São Paulo. O *swab* foi colhido usando o sistema da Oragene (Oracollect OC-100 –DNAgenotech-Ottawa (ON), Canada) logo após o horário de maior pico (20:00) em um dia regular de semana. Amostras foram coletadas em 4 carros distintos, a partir de barras usadas para apoio dos passageiros durante o deslocamento das composições. Cada um destes carros fazia parte de uma diferente composição do metrô representando duas diferentes linhas (verde e azul). *Swabs* foram obtidos ainda de corrimãos das escadas de quatro estações diferentes. A seguir, fizemos a extração de DNA do *swab*, usando instruções do fabricante, e geramos os amplicons que foram sequenciados pela plataforma *Ion Torrent PGM v2 kit 400*. Esse sequenciamento gerou 288,700 *reads* mapeadas no hg19 (software TMAP) com o *amplicon* de GALR2 (**Figura 28**). Considerando a qualidade das *reads* e das bases ≥ 10 , observamos 243821 *reads* na coordenada genômica chr17:74073094, das quais 779 *reads* contém a variante rs61745847 (G/T) representando 0,3% do total de *reads* nessa posição.

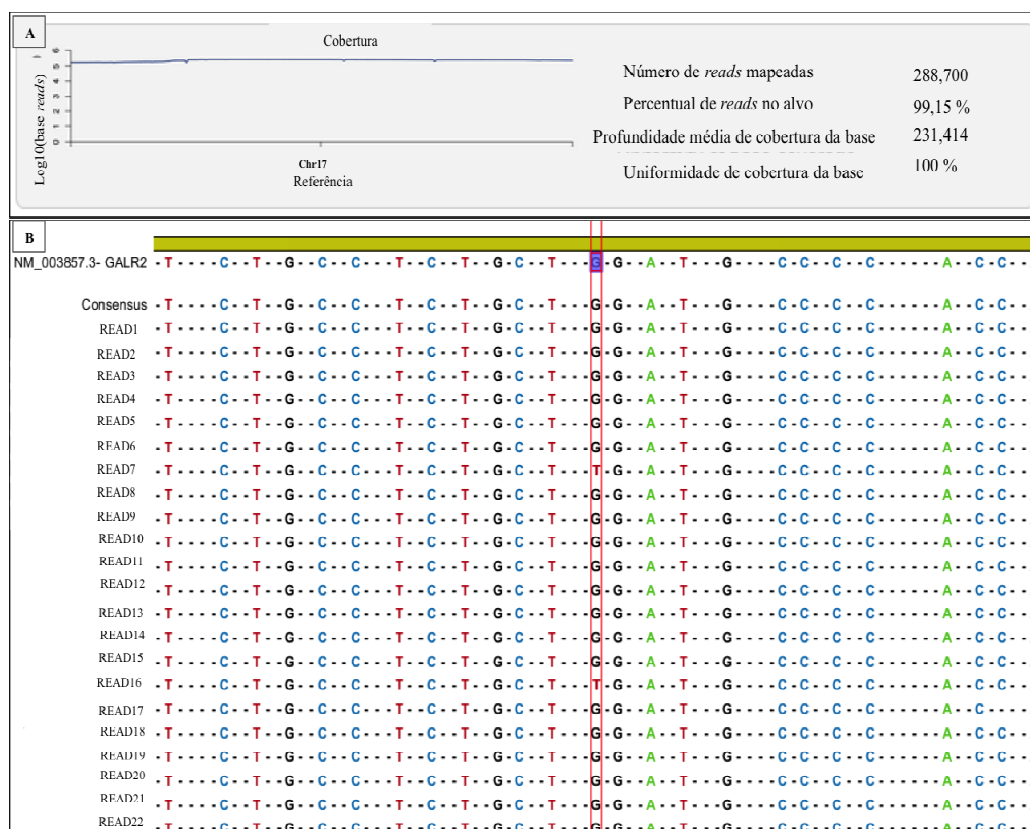


Figura 28 - Representação do sequenciamento do *amplicon* GALR2 da amostra do metrô de São Paulo-SP. (A) Resultado do sequenciamento do *amplicon* GALR2 que gerou 288,700 *reads* mapeadas, destas aproximadamente 99% estavam no alvo (*amplicon* GALR2) com 100% de uniformidade. (B) *Reads* alinhadas com o gene referência GALR2 destacando a posição do SNP rs61745847(chr17:74073094) onde ocorre uma alteração G/T.

Utilizamos o mesmo filtro de qualidade referido acima e consideramos ainda um *count* > 2 e apenas alterações *missense*, para comparar as frequências populacionais das variantes descritas no ExAC distribuídas na coordenada genômica do *amplicon* GALR2 (Chr17:74072952-74073230) com a frequência de *reads* obtidas pelo sequenciamento *amplicon* GALR2 da amostra do metrô contendo as mesmas variantes conforme apresentado na **Figura 29**.

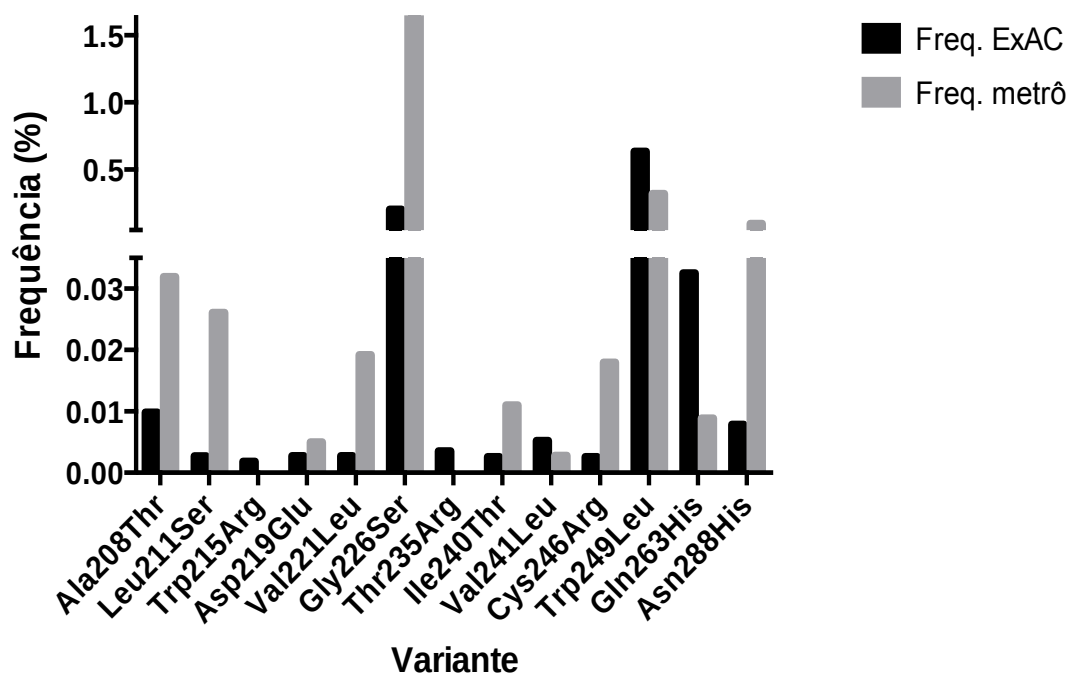


Figura 29 – Comparação da frequência populacional de SNPs *missense* descritos no ExAC (*count* >2), localizados dentro da *ORF* do gene humano *GALR2* distribuídos entre as coordenadas genômicas Chr17:74072952-74073230 e com impacto na sequência da proteína, com a frequência de *reads* das mesmas variante observadas na mostra de DNA do metrô de São Paulo. Dados gerados pela análise do sequenciamento do *amplicon* *GALR2* através da plataforma Ion PGM.

Estimativas do metrô de São Paulo indicam que cerca de 4 milhões de pessoas usam o metrô diariamente como meio de transporte (Metrô de São Paulo: Companhia do Metropolitano completa 45 anos com reconhecimento internacional). Se por um lado a nossa amostragem não nos permite dizer com exatidão o número de pessoas amostradas, acreditamos que este número tenha sido elevado, já que nos permitiu observar SNPs muito raros (com frequência abaixo de 0,1% no banco ExAC). A representação de todos estes SNPs raros, descritos no banco ExAC, com frequências similares em nossa amostragem, sugere que a coleta não foi enviesada e indica que o SNP de *GALR2* aqui discutido não tem uma frequência distintamente elevada em nossa população.

Apesar de não termos encontrado este SNP em outras amostras de nosso limitado grupo de EM, avançamos na tentativa de compreender seu possível impacto funcional. Assim propusemos uma colaboração com o Dr. Paulo SL Oliveira

especialista em modelagem molecular e mecânica molecular de proteínas e do MSc. José Geraldo C. Pereira, ambos ligados ao Laboratório de Biologia Computacional do Laboratório Nacional de Biociências (LNBio), os quais fizeram uma análise estrutural do mutante W249L de GALR2, com o objetivo de avaliar, *in silico*, se esta troca poderia acarretar em alteração funcional desta proteína.

A princípio foi feita a modelagem de GALR2 através de modelos estruturais construídos e resolvidos pelos programas de modelagem Yasara (KRIEGER et al. 2009) usando a estrutura do nociceptin/orphanin FQ receptor (NOP ou ORL-1, PDB ID: 4EA3) que possui 35% de identidade com GALR2 e cuja estrutura cristalográfica foi determinada no ano de 2013 (THOMPSON et al. 2012). Podemos observar que a diferença entre as cadeias laterais do triptofano e leucina não parece causar modificações significativas no volume da cavidade do sítio de ligação. No entanto, esta substituição parece alterar o volume da pequena cavidade adjacente localizada na porção intracelular (melhor visualizado nas **Figuras 30A-D**), cavidade esta que permite que o tripotafano rotacione para baixo (**Figura 31**).

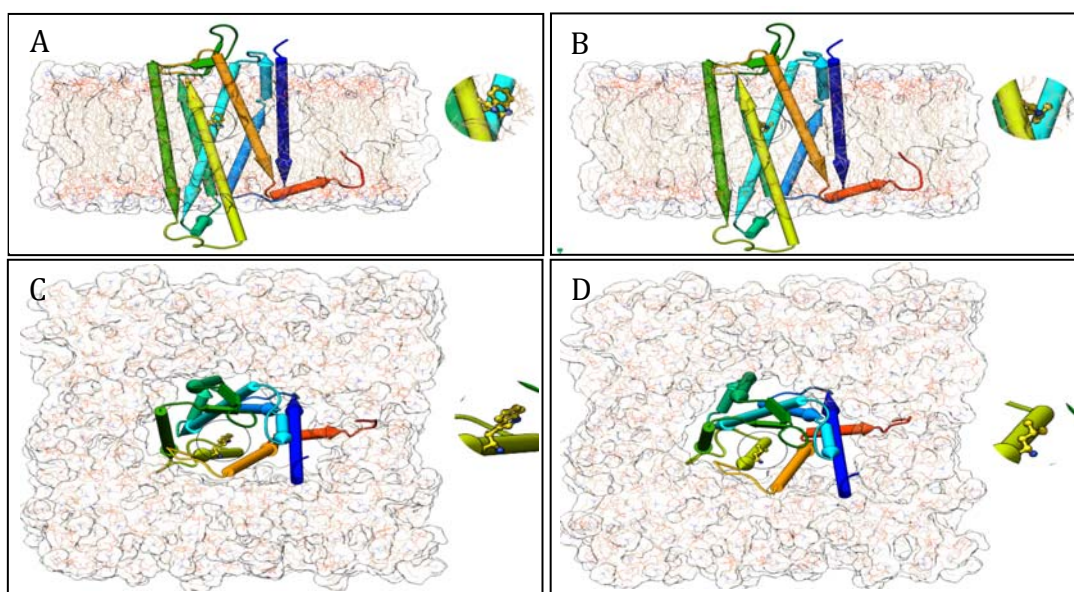


Figura 30– Vista lateral (A e B) e vista superior (extracelular > intracelular) (C e D) da proteína GalR2 inserida na membrana celular. A parte superior da membrana corresponde à porção extracelular. Em destaque está a localização de (A) estrutura de proteína selvagem contendo W249 e (B) da estrutura alterada pela presença de L249 na alça transmembrana 6 (TM6). A posição 249 ocupa o centro do feixe de hélices onde é formado o *pocket* de ligação da galanina.

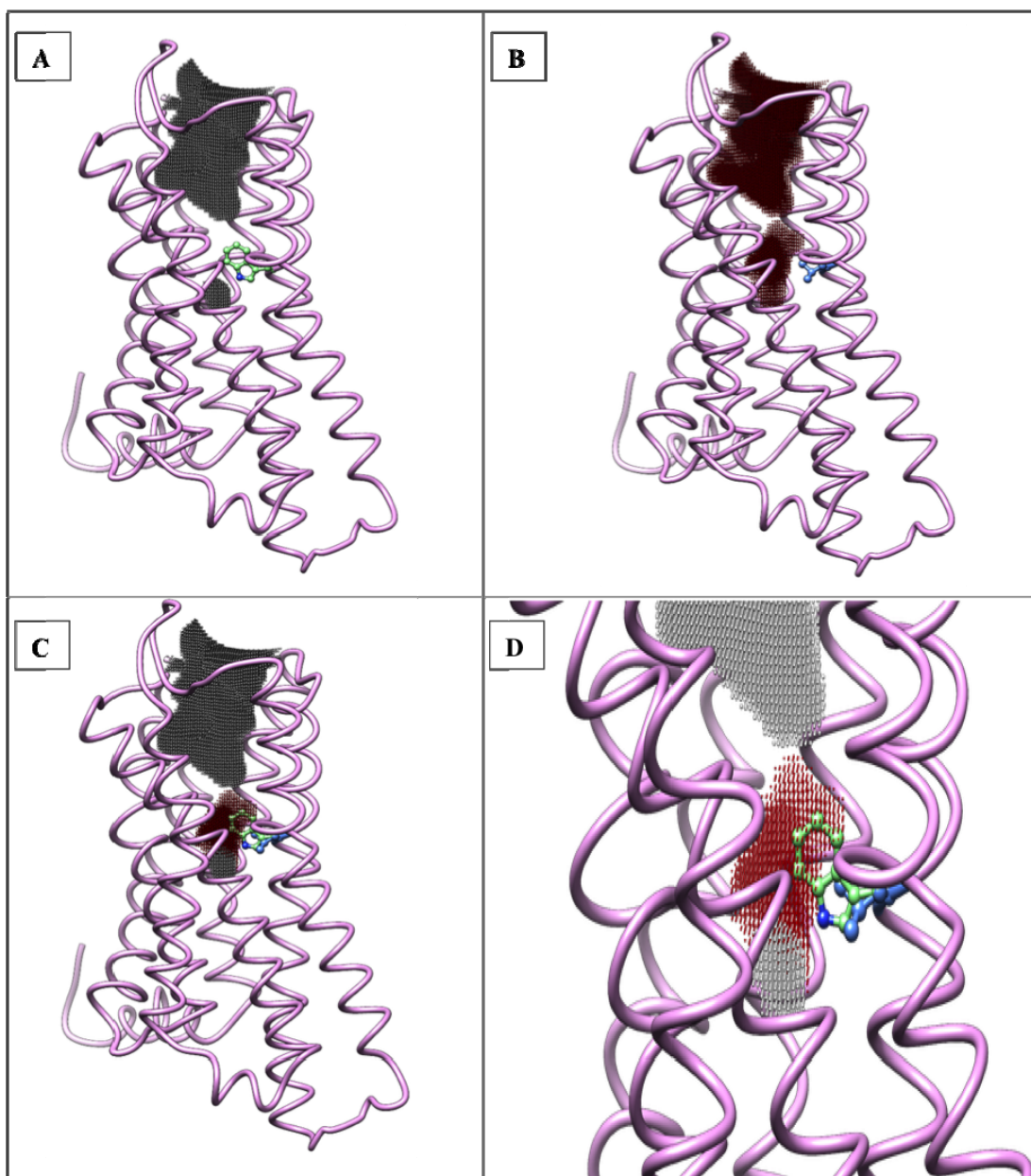


Figura 31 – Análise estrutural do volume das cavidades após simulação da ligação da galanina. (A) Cavidade de ligação da galanina GalR2 wt (W249) e cavidade adjacente (inferior ao *pocket*); (B) Cavidade de ligação da galanina GalR2 *mut* (L249) e cavidade adjacente; (C) Diferença entre as cavidades; (D) Visão mais detalhada das diferenças do *pocket* na região do resíduo 249.

Assim como outras GPCRs, GALR2 também possui 7 α -hélices transmembrana, que convertem sinais desencadeados por ligantes específicos, em respostas intra-celulares. Nestas hélices, existem diversos resíduos altamente conservados, que auxiliam a manter a estrutura e desencadeiam a sinalização

intracelular. Um dos resíduos importantes da hélice TM6, onde se localiza a alteração da proteína ocasionada por este SNP, é exatamente o triptofano, (denominado W6.48 conforme numeração de Ballesteros e Weinstein), que se localiza dentro do motivo altamente conservado CWxP (STODDART et al. 2014). Dentre as GPCRs classe-A *rhodopsin-like* 9 famílias (incluindo aminas, canabinóides, CAPA e várias outras, que totalizam 430 proteínas) este triptofano tem uma conservação de 100%. Diversos grupos sugerem que esta conservação se deva à capacidade de rotação da cadeia lateral do W6.48, que é uma parte fundamental do mecanismo de ativação de alguns receptores acoplados à proteína G (GPCRs) (SCHWARTZ et al. 2006; NYGAARD et al. 2009; SHI et al. 2002; KATRITCH et al. 2013).

Assim, de acordo com a literatura e com a análise de nossos colaboradores, parece que após a rotação, a cadeia adjacente seja ocupada realizando novas interações que permitirão novas mudanças conformacionais na porção intracelular do receptor. Na proteína mutante (W249L) a substituição do triptofano por uma leucina torna a cavidade adjacente maior na direção da cavidade de ligação da galanina. No entanto, devido ao volume da cadeia lateral da leucina, ela não seria capaz de preencher a cavidade adjacente e de realizar interações semelhantes ao Trp249 (quando rotacionado) que podem ser essenciais na ativação da GalR2. Através dessa análise, acreditamos que a ligação da galanina a GalR2 W249L será similar a ligação da galanina a GalR2 wt. Entretanto, a transdução do sinal para o interior da célula resultante da ativação do receptor provavelmente deve estar comprometida devido à incapacidade da Leucina de desencadear as mudanças conformacionais necessárias.

Galanina, GALR2 e Esclerose Múltipla – Em 2009 o grupo do Prof. David Wynick da *University of Bristol*, Reino Unido, demonstrou que a expressão de galanina se mostrou marcadamente aumentada na micróglia de lesões desmielinizantes da EM. Neste trabalho o grupo demonstrou que a super-expressão de galanina em animais transgênicos aboliu a doença em um modelo animal de encefalomielite autoimune enquanto que mutações com perda de função na galanina ou em GALR2 aumentaram a severidade da doença. O artigo discute que os efeitos marcantes da alteração da galanina endógena ou de mutações em GALR2 devem refletir um efeito neuroprotetor direto deste peptídeo pela ativação de GALR2 e

sugere que o uso de agonistas de galanina ou GALR2 podem ter implicações terapêuticas para a EM (WRAITH et al. 2009).

Posteriormente, um grupo da Universidade de Manitoba, no Canadá, utilizou o modelo de desmielinização da cuprizona (CPZ), para mimetizar a EM e demonstrou novamente os efeitos neuroprotetores da galanina no que se refere a mielinização. Animais transgênicos com expressão elevada de galanina tiveram uma menor perda de mielina o que novamente sugere o uso terapêutico de agonistas desta via pois agentes que protegem da perda de mielina e/ou aumentam a remielinização tem um grande potencial de reduzir o dano axonal (ZHANG et al. 2012). Em 2015, este mesmo grupo Canadense novamente demonstrou os fortes efeitos mielinizantes da galanina. Usando o modelo animal de expressão aumentada de galanina, os autores mostraram uma taxa acelerada de produção de mielina no dia 10 pós-natal, de modo independente de um aumento nas células precursoras de oligodendrócitos (LYUBETSKA et al. 2015).

Para avaliar o impacto dessa alteração faremos ensaios de mutagênese sítio dirigida e avaliação da sinalização intracelular, através de colaboração já estabelecida com o grupo do Prof. Stephen Hill (Universidade de Nottingham, Inglaterra) e do Dr. Kleber Franchini (Laboratório Nacional de Biociências, LNBio).

Acreditamos que esta seja a primeira demonstração de uma mutação em homozigose no gene GALR2 em esclerose múltipla. Nossos dados, em conjunto com a literatura, reforçam o envolvimento da galanina e de seu receptor GALR2 nos processos biológicos que levam à desmielinização que ocorre na EM e reforçam que o uso de agonistas da galanina seja um importante aliado no tratamento desta doença.

5 CONCLUSÃO

Diante dos resultados apresentados concluímos que:

- Algumas das alterações moleculares SNVs do tipo *de novo* validadas, ocorrem em genes relevantes do ponto de vista do SNC quando avaliadas individualmente. Porém nenhuma dessas alterações foi encontrada em mais de um paciente. Talvez devido ao pequeno número de amostras ou porque essas alterações não estão intrinsecamente associadas à etiologia da doença.
- A análise de eventos de retroinserção *de novo* (DNR) em genes revelou o enriquecimento da via de EM através da análise *in silico*. Além disso, os eventos DNR estão ranqueados em vias associadas ao SNC, para os quais três vias foram também validadas utilizando diferentes ferramentas de análise *in silico*.
- Observamos através da análise de SRHH alguns polimorfismos em genes com papéis importantes no SNC, considerados deletérios por ferramentas de predição *in silico*. Esses SNPs foram selecionados e validados através do Sanger.
- O SNP rs61745847 que leva a uma importante alteração no receptor do neuropeptídeo galanina –*GALR2*, foi observado em homozigose em uma paciente de nossa casuística.
- Este SNP é raro (<0.7%) em todas as populações mundiais, incluindo a brasileira, e análises de modelagem molecular indicam que este SNP leva a uma alteração da estrutura do receptor com implicações na via de sinalização intracelular.

Esperamos com os resultados apresentados nesta dissertação ter contribuído para o melhor entendimento dos fatores genéticos associados à EM. Para isso, continuaremos trabalhando com as variantes mais relevantes e tentaremos colaborações para aumentar o número de amostras e tornar essas análises mais representativas.

6 PERSPECTIVAS FUTURAS

Os achados observados nesse estudo abriram novas perspectivas de análises para o melhor entendimento da etiologia da EM, conforme segue abaixo:

- Validar os eventos DNR e correlacionar os próprios eventos nos respectivos genes, avaliando se eles se repetem em genes diferentes ou nos mesmos genes em diferentes indivíduos e checar na literatura algum descritivo desses eventos com EM;
- Avaliar a presença de variações em número de cópias (CNVs) *de novo* através da ferramenta Cytoscan TM Assay (Affymetrix) em colaboração com a Profa. Dra. Silvia Rogatto. Os grupos serão selecionados através de uma análise prévia *in silico* utilizando a ferramenta EXCAVATOR (dados não apresentados nesta dissertação);
- Validar as alterações SNVs *de novo* e SRHH em uma casuística independente e maior de portadores de EM;
- Fazer ensaios de mutação sítio dirigida para o SRHH rs61745847.

7 REFERÊNCIAS BIBLIOGRÁFICAS

Agarwal V, Kommaddi RP, Valli K, et al. Drug metabolism in human brain: high levels of cytochrome P4503A43 in brain and metabolism of anti-anxiety drug alprazolam to its active metabolite. **PLoS One** 2008; 3:e2337.

Aktas O, Küry P, Kieseier B, Hartung H-P. Fingolimod is a potential novel therapy for multiple sclerosis. **Nat Rev Neurol** 2010; 6:373–82.

Antony JM, Deslauriers AM, Bhat RK, Ellestad KK, Power C. Human endogenous retroviruses and multiple sclerosis: innocent bystanders or disease determinants? **Biochim Biophys Acta** 2011; 1812:162–76.

Ascherio A, Munger KL. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. **Ann Neurol** 2007; 61:504–13.

Ascherio A, Munger KL, Lennette ET, et al. Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. **JAMA** 2001; 286:3083–8.

Ascherio A, Munger KL, Simon KC. Vitamin D and multiple sclerosis. **Lancet Neurol** 2010; 9:599–612.

Axtell RC, De Jong B A, Boniface K, et al. T helper type 1 and 17 cells determine efficacy of interferon-beta in multiple sclerosis and experimental encephalomyelitis. **Nat Med** 2010; 16:406–412.

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; 94:599–603.

Baranzini SE, Mudge J, Van Velkinburgh JC, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. **Nature** 2010; 464:1351–6.

Baranzini SE, Wang J, Gibson R A, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. **Hum Mol Genet** 2009; 18:767–78.

- Behrens TW, Mueller DL. Bcl-x and the regulation of survival in the immune system. **Immunol Res** 1997; 16:149–60.
- Belbasis L, Bellou V, Evangelou E, Ioannidis JP A, Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. **Lancet Neurol** 2015; 14:263–273.
- Berlioz C, Darlix JL. An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. **J Virol** 1995; 69:2214–22.
- Bertrams J, Höher PG, Kuwert E. Letter: HL-A antigens in multiple sclerosis. **Lancet** 1974; 1:1287.
- Blackstone C. Cellular pathways of hereditary spastic paraplegia. **Annu Rev Neurosci** 2012; 35:25–47.
- Blumberg BM, Mock DJ, Powers JM, et al. The HHV6 paradox: ubiquitous commensal or insidious pathogen? A two-step in situ PCR approach. **J Clin Virol** 2000; 16:159–178.
- Boiko A, Vorobeychik G, Paty D, Devonshire V, Sadovnick D. Early onset multiple sclerosis: a longitudinal study. **Neurology** 2002; 59:1006–10.
- Bruno R, Sabater L, Sospedra M, et al. Multiple sclerosis candidate autoantigens except myelin oligodendrocyte glycoprotein are transcribed in human thymus. **Eur J Immunol** 2002; 32:2737–47.
- Bulman, DE & Ebers G. **The geography of multiple sclerosis reflects genetic susceptibility.pdf** London, Ontario, Canada, 1992; .
- Butler MG, Rafi SK, Hossain W, Stephan DA, Manzardo AM. Whole exome sequencing in females with autism implicates novel and candidate genes. **Int J Mol Sci** 2015; 16:1312–35.
- Cabre P, Signate A, Olindo S, et al. Role of return migration in the emergence of multiple sclerosis in the French West Indies. **Brain** 2005; 128:2899–910.
- Callegaro D, Goldbaum M, Morais L, et al. The prevalence of multiple sclerosis in the city of São Paulo, Brazil, 1997. **Acta Neurol Scand** 2001; 104:208–13.

- Carrasquillo MM, Belbin O, Hunter TA, et al. Replication of BIN1 association with Alzheimer's disease and evaluation of genetic interactions. **J Alzheimers Dis** 2011; 24:751–8.
- Carta MG, Moro MF, Lorefice L, et al. The risk of Bipolar Disorders in Multiple Sclerosis. **J Affect Disord** 2014; 155:255–60.
- Caselli E, Boni M, Bracci A, et al. Detection of antibodies directed against human herpesvirus 6 U94/REP in sera of patients affected by multiple sclerosis. **J Clin Microbiol** 2002; 40:4131–7.
- Ceccarelli A, Bakshi R, Neema M. MRI in multiple sclerosis: a review of the current literature. **Curr Opin Neurol** 2012; 25:402–9.
- Cermelli C, Jacobson S. Viruses and multiple sclerosis. **Viral Immunol** 2000; 13:255–67.
- Chen S, Olsen K, Grigsby C, Gardner DG. Vitamin D activates type A natriuretic peptide receptor gene transcription in inner medullary collecting duct cells. **Kidney Int** 2007; 72:300–6.
- Christensen T, Dissing Sørensen P, Riemann H, et al. Molecular characterization of HERV-H variants associated with multiple sclerosis. **Acta Neurol Scand** 2000; 101:229–38.
- Cirulli ET, Lasseigne BN, Petrovski S, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. **Science** 2015; 347:121–12.
- Cohn LB, Silva IT, Oliveira TY, et al. HIV-1 Integration Landscape during Latent and Active Infection. **Cell** 2015; 160:420–432.
- Compston A, Coles A. Multiple sclerosis. **Lancet** 2002; 359:1221–31.
- Compston A, Coles A. Multiple sclerosis. **Lancet** 2008; 372:1502–17.
- Connolly JH, Allen I V, Hurwitz LJ, Millar JH. Measles-virus antibody and antigen in subacute sclerosing panencephalitis. **Lancet** 1967; 1:542–4.
- Cusick MF, Libbey JE, Fujinami RS. Multiple sclerosis: autoimmunity and viruses. **Curr Opin Rheumatol** 2013; 25:496–501.

- D'Ambrosio D, Cippitelli M, Cocciolo MG, et al. Inhibition of IL-12 production by 1,25-dihydroxyvitamin D₃. Involvement of NF-kappaB downregulation in transcriptional repression of the p40 gene. **J Clin Invest** 1998; 101:252–62.
- De Araújo-Souza PS, Hanschke SCH, Viola JPB. Epigenetic Control of Interferon-Gamma Expression in CD8 T Cells. **J Immunol Res** 2015; 2015:849573.
- Degenhardt A, Ramagopalan S V, Scalfari A, Ebers GC. Clinical prognostic factors in multiple sclerosis: a natural history review. **Nat Rev Neurol** 2009; 5:672–82.
- Dolei A., Serra C, Mameli G, et al. Multiple sclerosis-associated retrovirus (MSRV) in Sardinian MS patients. **Neurology** 2002; 58:471–3.
- Dolei A. MSRV/HERV-W/syncytin and its linkage to multiple sclerosis: the usability and the hazard of a human endogenous retrovirus. **J Neurovirol** 2005; 11:232–5.
- Dolei A, Perron H. The multiple sclerosis-associated retrovirus and its HERV-W endogenous family: a biological interface between virology, genetics, and immunology in human physiology and disease. **J Neurovirol** 2009; 15:4–13.
- Domanski TL, Finta C, Halpert JR, Zaphiropoulos PG. cDNA cloning and initial characterization of CYP3A43, a novel human cytochrome P450. **Mol Pharmacol** 2001; 59:386–92.
- Dyment D A, Cader MZ, Chao MJ, et al. Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. **Neurology** 2012; 79:406–11.
- Dyment D A, Herrera BM, Cader MZ, et al. Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance. **Hum Mol Genet** 2005; 14:2019–26.
- Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. **Nature** 1995; 377:150–1.
- Egles C, Claudepierre T, Manglapus MK, Champlaud M-F, Brunken WJ, Hunter DD. Laminins containing the beta2 chain modulate the precise organization of CNS synapses. **Mol Cell Neurosci** 2007; 34:288–98.

- Fischer MT, Wimmer I, Hoftberger R, et al. Disease-specific molecular events in cortical multiple sclerosis lesions. **Brain** 2013; 136:1799–1815.
- Ford ML, Onami TM, Sperling AI, Ahmed R, Evavold BD. CD43 modulates severity and onset of experimental autoimmune encephalomyelitis. **J Immunol** 2003; 171:6527–6533.
- Franklin RJM. Why does remyelination fail in multiple sclerosis? **Nat Rev Neurosci** 2002; 3:705–14.
- Freedman DM, Dosemeci M, Alavanja MC. Mortality from multiple sclerosis and exposure to residential and occupational solar radiation: a case-control study based on death certificates. **Occup Environ Med** 2000; 57:418–421.
- Friese M A, Montalban X, Willcox N, Bell JI, Martin R, Fugger L. The value of animal models for drug development in multiple sclerosis. **Brain** 2006; 129:1940–52.
- Friese M A, Schattling B, Fugger L. Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. **Nat Rev Neurol** 2014; 10:225–238.
- Gale CR, Martyn CN. Migrant studies in multiple sclerosis. **Prog Neurobiol** 1995; 47:425–48.
- Garcia-Montojo M, Dominguez-Mozo M, Arias-Leal A, 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; 8:e53623.
- Gifford R, Tristem M. The evolution, distribution and diversity of endogenous retroviruses. **Virus Genes** 2003; 26:291–315.
- Gilden DH. Infectious causes of multiple sclerosis. **Lancet Neurol** 2005; 4:195–202.
- Gilden DH, Devlin ME, Burgoon MP, Owens GP. The search for virus in multiple sclerosis brain. **Mult Scler** 1996; 2:179–83.
- Giovannoni G, Ebers G. Multiple sclerosis: the environment and causation. **Curr Opin Neurol** 2007; 20:261–8.

Guggenmos J, Schubart AS, Ogg S, et al. Antibody cross-reactivity between myelin oligodendrocyte glycoprotein and the milk protein butyrophilin in multiple sclerosis. **J Immunol** 2004; 172:661–8.

Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. **N Engl J Med** 2007; 357:851–62.

Haghikia A, Hohlfeld R, Gold R, Fugger L. Therapies for multiple sclerosis: translational achievements and outstanding needs. **Trends Mol Med** 2013; 19:309–19.

Haider L, Fischer MT, Frischer JM, et al. Oxidative damage in multiple sclerosis lesions. **Brain** 2011; 134:1914–24.

Hauser SL, Chan JR, Oksenberg JR. Multiple sclerosis: Prospects and promise. **Ann Neurol** 2013; 74:317–27.

Hoischen A, Van Bon BWM, Gilissen C, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. **Nat Genet** 2010; 42:483–5.

Hollis BW. Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for establishing a new effective dietary intake recommendation for vitamin D. **J Nutr** 2005; 135:317–22.

Holton P, Ryten M, Nalls M, et al. Initial assessment of the pathogenic mechanisms of the recently identified Alzheimer risk Loci. **Ann Hum Genet** 2013; 77:85–105.

Hsieh CS, Macatonia SE, Tripp CS, Wolf SF, O’Garra A, Murphy KM. Development of TH1 CD4⁺ T cells through IL-12 produced by Listeria-induced macrophages. **Science** 1993; 260:547–9.

Hu X, Pickering E, Liu YC, et al. Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer’s disease. **PLoS One** 2011; 6:e16616.

Jack C, Ruffini F, Bar-Or A, Antel JP. Microglia and multiple sclerosis. **J Neurosci Res** 2005; 81:363–73.

Janssens W, Decramer M, Mathieu C, Korf H. Vitamin D and chronic obstructive pulmonary disease: Hype or reality? **Lancet Respir Med** 2013; 1:804–812.

- Kaplan LM, Spindel ER, Isselbacher KJ, Chin WW. Tissue-specific expression of the rat galanin gene. **Proc Natl Acad Sci U S A** 1988; 85:1065–9.
- Kappos L, Freedman MS, Polman CH, et al. Long-term effect of early treatment with interferon beta-1b after a first clinical event suggestive of multiple sclerosis: 5-year active treatment extension of the phase 3 BENEFIT trial. **Lancet Neurol** 2009; 8:987–97.
- Kappos L, Radue E-W, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. **N Engl J Med** 2010; 362:387–401.
- Karch CM, Cruchaga C, Goate AM. Alzheimer's disease genetics: from the bench to the clinic. **Neuron** 2014; 83:11–26.
- Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. **Annu Rev Pharmacol Toxicol** 2013; 53:531–56.
- Kemppinen A K, Baker A, Liao W, et al. Exome sequencing in single cells from the cerebrospinal fluid in multiple sclerosis. **Mult Scler** 2014; 20:1564–8.
- Kim S, Wairkar YP, Daniels RW, DiAntonio A. The novel endosomal membrane protein Ema interacts with the class C Vps-HOPS complex to promote endosomal maturation. **J Cell Biol** 2010; 188:717–34.
- Kleinewietfeld M, Manzel A, Titze J, et al. Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. **Nature** 2013; 496:518–22.
- Krametter D, Niederwieser G, Berghold A, Birnbaum G, Hartung H. Chlamydia pneumoniae in multiple sclerosis: humoral immune responses in serum and cerebrospinal fluid and correlation with disease activity marker. 2001; 13–18.
- Kremlev SG, Gournier-Hausser AL, Del Valle L, Perez-Liz G, Dimitrov S, Tuszynski G. Angiocidin promotes pro-inflammatory cytokine production and antigen presentation in multiple sclerosis. **J Neuroimmunol** 2008; 194:132–42.
- Krieger E, Joo K, Lee J, et al. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. **Proteins** 2009; 77 Suppl 9:114–22.

- Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). **Neurology** 1983; 33:1444–52.
- Kurtzke JF. Multiple sclerosis in time and space--geographic clues to cause. **J Neurovirol** 2000; 6 Suppl 2:S134–40.
- Kurtzke JF, Dean G, Botha DP. A method for estimating the age at immigration of white immigrants to South Africa, with an example of its importance. **S Afr Med J** 1970; 44:663–9.
- Kurtzke JF, Delasnerie-Lauprêtre N, Wallin MT. Multiple sclerosis in North African migrants to France. **Acta Neurol Scand** 1998; 98:302–9.
- Lana-Peixoto MA, Frota ERC, Campos GB, Monteiro LP. The prevalence of multiple sclerosis in Belo Horizonte, Brazil. **Arq Neuropsiquiatr** 2012; 70:102–7.
- Lassmann H. Cortical lesions in multiple sclerosis: inflammation versus neurodegeneration. **Brain** 2012; 135:2904–5.
- Lemire JM. Immunomodulatory actions of 1,25-dihydroxyvitamin D3. **J Steroid Biochem Mol Biol** 1995; 53:599–602.
- Levin LI, Munger KL, O'Reilly EJ, Falk KI, Ascherio A. Primary infection with the Epstein-Barr virus and risk of multiple sclerosis. **Ann Neurol** 2010; 67:824–30.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics** 2009; 25:1754–60.
- Libby RT, Lavalley CR, Balkema GW, Brunken WJ, Hunter DD. Disruption of laminin beta2 chain production causes alterations in morphology and function in the CNS. **J Neurosci** 1999; 19:9399–411.
- Ling V, Wu PW, Finnerty HF, et al. Assembly and annotation of human chromosome 2q33 sequence containing the CD28, CTLA4, and ICOS gene cluster: analysis by computational, comparative, and microarray approaches. **Genomics** 2001; 78:155–68.
- Linial ML, Miller AD. Retroviral RNA packaging: sequence requirements and implications. **Curr Top Microbiol Immunol** 1990; 157:125–52.

- Lipinski S, Grabe N, Jacobs G, et al. RNAi screening identifies mediators of NOD2 signaling: implications for spatial specificity of MDP recognition. **Proc Natl Acad Sci U S A** 2012; 109:21426–31.
- Lopes JD, Dos Reis M, Brentani RR. Presence of laminin receptors in *Staphylococcus aureus*. **Science** 1985; 229:275–7.
- Löwer R, Boller K, Hasenmaier B, et al. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. **Proc Natl Acad Sci U S A** 1993; 90:4480–4.
- 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; 93:5177–84.
- Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. **Neurology** 1996; 46:907–11.
- Lublin FD, Reingold SC, Cohen J A, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. **Neurology** 2014; 83:278–86.
- Lucchinetti C, Brück W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. **Ann Neurol** 2000; 47:707–17.
- Lynch M. Rate, molecular spectrum, and consequences of human mutation. **Proc Natl Acad Sci U S A** 2010; 107:961–8.
- Lyubetska H, Zhang L, Kong J, Vrontakis M. An elevated level of circulating galanin promotes developmental expression of myelin basic protein in the mouse brain. **Neuroscience** 2015; 284:581–9.
- Mager DL. Human endogenous retroviruses and pathogenicity: genomic considerations. **Trends Microbiol** 1999; 7:431; author reply 431–2.
- Majid A, Galetta SL, Sweeney CJ, et al. Epstein-Barr virus myeloradiculitis and encephalomyeloradiculitis. **Brain** 2002; 125:159–65.

- Mameli G, Madeddu G, Mei A, et al. Activation of MSRV-type endogenous retroviruses during infectious mononucleosis and Epstein-Barr virus latency: the missing link with multiple sclerosis? **PLoS One** 2013; 8:e78474.
- Manjunath N, Correa M, Ardman M, Ardman B. Negative regulation of T-cell adhesion and activation by CD43. **Nature** 1995; 377:535–538.
- Marrie RA. Environmental risk factors in multiple sclerosis aetiology. **Lancet Neurol** 2004; 3:709–18.
- Maruszak H, Brew BJ, Giovannoni G, Gold J. Could antiretroviral drugs be effective in multiple sclerosis? A case report. **Eur J Neurol** 2011; 18:e110–1.
- Masoodi TA, Al Shammari S A, Al-Muammar MN, Alhamdan A A, Talluri VR. Exploration of deleterious single nucleotide polymorphisms in late-onset Alzheimer disease susceptibility genes. **Gene** 2013; 512:429–37.
- McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. **Ann Neurol** 2001; 50:121–7.
- Mechenthaler I. Galanin and the neuroendocrine axes. **Cell Mol Life Sci** 2008; 65:1826–35.
- [MSIF] Multiple Sclerosis International Federation. **Atlas 2013: epidemiology of MS**. Available from: <URL:<http://www.msif.org/wp-content/uploads/2014/09/Atlas-of-MS.pdf>> [2015 jan 12]
- Ménard A, Amouri R, Michel M, et al. Gliotoxicity, reverse transcriptase activity and retroviral RNA in monocyte/macrophage culture supernatants from patients with multiple sclerosis. **FEBS Lett** 1997; 413:477–85.
- Menge T, Stüve O, Kieseier BC, Hartung H. Alemtuzumab: the advantages and challenges of a novel therapy in MS. **Neurology** 2014; 83:87–97.
- Milo R, Miller A. Revised diagnostic criteria of multiple sclerosis. **Autoimmun Rev** 2014; 13:518–24.

- Multiple I, Genetics S. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. **Am J Hum Genet** 2013; 92:854–65.
- Nahm M, Lee M-J, Parkinson W, et al. Spartin regulates synaptic growth and neuronal survival by inhibiting BMP-mediated microtubule stabilization. **Neuron** 2013; 77:680–95.
- Napoli I, Neumann H. Protective effects of microglia in multiple sclerosis. **Exp Neurol** 2010; 225:24–8.
- Nicot A-S, Toussaint A, Tosch V, et al. Mutations in amphiphysin 2 (BIN1) disrupt interaction with dynamin 2 and cause autosomal recessive centronuclear myopathy. **Nat Genet** 2007; 39:1134–1139.
- Nikić I, Merkler D, Sorbara C, et al. A reversible form of axon damage in experimental autoimmune encephalomyelitis and multiple sclerosis. **Nat Med** 2011; 17:495–499.
- Nishimune H, Sanes JR, Carlson SS. A synaptic laminin-calcium channel interaction organizes active zones in motor nerve terminals. **Nature** 2004; 432:580–7.
- Nygaard R, Frimurer TM, Holst B, Rosenkilde MM, Schwartz TW. Ligand binding and micro-switches in 7TM receptor structures. **Trends Pharmacol Sci** 2009; 30:249–59.
- O’Gorman C, Lucas R, Taylor B. Environmental risk factors for multiple sclerosis: a review with a focus on molecular mechanisms. **Int J Mol Sci** 2012; 13:11718–52.
- Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. **Lancet** 1971; 1:1257–60.
- Pandey KN. The functional genomics of guanylyl cyclase/natriuretic peptide receptor-A: perspectives and paradigms. **FEBS J** 2011; 278:1792–807.
- Pandey KN. Guanylyl cyclase/natriuretic peptide receptor-A signaling antagonizes phosphoinositide hydrolysis, Ca²⁺ release, and activation of protein kinase C. **Front Mol Neurosci** 2014; 7:1–14.

Pantazou V, Schluep M, Du Pasquier R. Environmental factors in multiple sclerosis. **Presse Med** 2015;

Pedraza-Alva G, Mérida LB, Burakoff SJ, Rosenstein Y. T cell activation through the CD43 molecule leads to Vav tyrosine phosphorylation and mitogen-activated protein kinase pathway activation. **J Biol Chem** 1998; 273:14218–14224.

Perron H, Geny C, Laurent A, et al. Leptomeningeal cell line from multiple sclerosis with reverse transcriptase activity and viral particles. **Res Virol** 1989; 140:551–61.

Picas L, Viaud J, Schauer K, et al. BIN1/M-Amphiphysin2 induces clustering of phosphoinositides to recruit its downstream partner dynamin. **Nat Commun** 2014; 5:5647.

Polman CH, O'Connor PW, Havrdova E, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. **N Engl J Med** 2006; 354:899–910.

Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. **Ann Neurol** 2011; 69:292–302.

Ponsonby A-L, Lucas RM, Van der Mei IAF. UVR, vitamin D and three autoimmune diseases--multiple sclerosis, type 1 diabetes, rheumatoid arthritis. **Photochem Photobiol** 2005; 81:1267–75.

Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. **Ann Neurol** 1983; 13:227–31.

Qiu W, Raven S, Wu J-S, et al. Hypothalamic lesions in multiple sclerosis. **J Neurol Neurosurg Psychiatry** 2011; 82:819–22.

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. **Bioinformatics** 2010; 26:841–2.

Ramagopalan S V, Dymment D A, Cader MZ, et al. Rare variants in the CYP27B1 gene are associated with multiple sclerosis. **Ann Neurol** 2011; 70:881–6.

Ransohoff RM, Hafler D A, Lucchinetti CF. Multiple sclerosis-a quiet revolution. **Nat Rev Neurol** 2015; 11:134–142.

- Rovituso D, Heller S, Schroeter M, Kleinschnitz C, Kuerten S. B1 cells are unaffected by immune modulatory treatment in relapsing-relapsing multiple sclerosis patients. **J Neuroimmunol** 2014; 272:86–90.
- Ryan A, Baner J, Demko Z, et al. Informatics-based, highly accurate, noninvasive prenatal paternity testing. **Genet Med** 2013; 15:473–7.
- Sadovnick A D. European Charcot Foundation Lecture: the natural history of multiple sclerosis and gender. **J Neurol Sci** 2009; 286:1–5.
- Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. **Nature** 2011; 476:214–9.
- Schumacker GA, Beebe G, Kibler RF, et al. Problems of Experimental Trials of Therapy in Multiple Sclerosis: Report by the Panel on the Evaluation of Experimental Trials of Therapy in Multiple Sclerosis. **Ann N Y Acad Sci** 1965; 122:552–68.
- Schutzer SE, Angel TE, Liu T, et al. Gray matter is targeted in first-attack multiple sclerosis. **PLoS One** 2013; 8:e66117.
- Schwartz TW, Frimurer TM, Holst B, Rosenkilde MM, Eling CE. Molecular mechanism of 7TM receptor activation--a global toggle switch model. **Annu Rev Pharmacol Toxicol** 2006; 46:481–519.
- Segal BM, Constantinescu CS, Raychaudhuri A, Kim L, Fidelus-Gort R, Kasper LH. Repeated subcutaneous injections of IL12/23 p40 neutralising antibody, ustekinumab, in patients with relapsing-relapsing multiple sclerosis: a phase II, double-blind, placebo-controlled, randomised, dose-ranging study. **Lancet Neurol** 2008; 7:796–804.
- Serra C, Mameli G, Arru G, Sotgiu S, Rosati G, Dolei A. In vitro modulation of the multiple sclerosis (MS)-associated retrovirus by cytokines: implications for MS pathogenesis. **J Neurovirol** 2003; 9:637–43.
- Shaw JP, Utz PJ, Durand DB, Toole JJ, Emmel EA, Crabtree GR. Identification of a putative regulator of early T cell activation genes. **Science** 1988; 241:202–5.

- Shi L, Liapakis G, Xu R, Guarnieri F, Ballesteros J A, Javitch J A. Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. **J Biol Chem** 2002; 277:40989–96.
- Smith BN, Ticozzi N, Fallini C, et al. Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. **Neuron** 2014; 84:324–31.
- Sotgiu S, Serra C, Mameli G, et al. Multiple sclerosis-associated retrovirus and MS prognosis: an observational study. **Neurology** 2002; 59:1071–3.
- Sriram S, Stratton CW, Yao S-Y, et al. Chlamydia pneumoniae infection of the central nervous system in multiple sclerosis. **Ann Neurol** 1999; 46:6–14.
- Stenzel N, Fetzer CP, Heumann R, Erdmann KS. PDZ-domain-directed basolateral targeting of the peripheral membrane protein FRMPD2 in epithelial cells. **J Cell Sci** 2009; 122:3374–84.
- Stoddart L A, Kellam B, Briddon SJ, Hill SJ. Effect of a toggle switch mutation in TM6 of the human adenosine A₁ receptor on Gi protein-dependent signalling and Gi-independent receptor internalization. **Br J Pharmacol** 2014; 171:3827–44.
- Stürner KH, Borgmeyer U, Schulze C, Pless O, Martin R. A multiple sclerosis-associated variant of CBLB links genetic risk with type I IFN function. **J Immunol** 2014; 193:4439–47.
- Takei K, Slepnev VI, Haucke V, De Camilli P. Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis. **Nat Cell Biol** 1999; 1:33–9.
- Thompson A A, Liu W, Chun E, et al. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. **Nature** 2012; 485:395–9.
- Trisolini M, Honeycutt A WJ E LS. Global Economic Impact of Multiple Sclerosis. 2010;
- VanAmerongen BM, Dijkstra CD, Lips P, Polman CH. Multiple sclerosis and vitamin D: an update. **Eur J Clin Nutr** 2004; 58:1095–109.
- Vandvik B, Norrby E, Nordal HJ, Degré M. Oligoclonal measles virus-specific IgG antibodies isolated from cerebrospinal fluids, brain extracts, and sera from patients

with subacute sclerosing panencephalitis and multiple sclerosis. **Scand J Immunol** 1976; 5:979–92.

Vissers LELM, De Ligt J, Gilissen C, et al. A de novo paradigm for mental retardation. **Nat Genet** 2010; 42:1109–12.

Vukusic S, Van Bockstael V, Gosselin S, Confavreux C. Regional variations in the prevalence of multiple sclerosis in French farmers. **J Neurol Neurosurg Psychiatry** 2007; 78:707–9.

Waiczies S, Weber A, Lünemann JD, Aktas O, Zschenderlein R, Zipp F. Elevated Bcl-X(L) levels correlate with T cell survival in multiple sclerosis. **J Neuroimmunol** 2002; 126:213–20.

Weiner HL. A shift from adaptive to innate immunity: a potential mechanism of disease progression in multiple sclerosis. **J Neurol** 2008; 255 Suppl :3–11.

Weston M, Constantinescu CS. What role does tobacco smoking play in multiple sclerosis disability and mortality? A review of the evidence. **Neurodegener Dis Manag** 2015; 5:19–25.

Wight M, Werner A. The functions of natural antisense transcripts. **Essays Biochem** 2013; 54:91–101.

Wöbke TK, Sorg BL, Steinhilber D. Vitamin D in inflammatory diseases. **Front Physiol** 2014; 5:244.

Wraith DC, Pope R, Butzkueven H, et al. A role for galanin in human and experimental inflammatory demyelination. **Proc Natl Acad Sci U S A** 2009; 106:15466–71.

Xu L, Wrona TJ, Dudley JP. Exogenous mouse mammary tumor virus (MMTV) infection induces endogenous MMTV gag expression. **Virology** 1996; 215:113–23.

Yeo TW, De Jager PL, Gregory SG, et al. A second major histocompatibility complex susceptibility locus for multiple sclerosis. **Ann Neurol** 2007; 61:228–236.

Yu M, Fang P, Shi M, et al. Galanin receptors possibly modulate the obesity-induced change in pain threshold. **Peptides** 2013; 44:55–9.

Zhang L, Yu W, Schroedter I, Kong J, Vrontakis M. Galanin transgenic mice with elevated circulating galanin levels alleviate demyelination in a cuprizone-induced MS mouse model. **PLoS One** 2012; 7:e33901.

Zhang Y, Blattman JN, Kennedy NJ, et al. Regulation of innate and adaptive immune responses by MAP kinase phosphatase 5. **Nature** 2004; 430:793–7.

Zhao Z, Ma L. Regulation of axonal development by natriuretic peptide hormones. **Proc Natl Acad Sci U S A** 2009; 106:18016–21.

Zuvich RL, McCauley JL, Pericak-Vance MA, Haines JL. Genetics and pathogenesis of multiple sclerosis. **Semin Immunol** 2009; 21:328–33.

Anexo 1 – Carta de aprovação do Comitê de Ética em Pesquisa-CEP da Fundação Antônio Prudente e FMRP-USP

 **A.C. Camargo
Cancer Center**

Comitê de Ética em
Pesquisa - CEP

São Paulo, 29 de outubro de 2013.

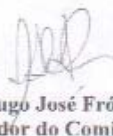
Ao
Dr. Emmanuel Dias-Neto.

Ref.: Projeto de Pesquisa nº. 1748/13
“Investigação de mutações e de *novo* como possível base biológica da esclerose múltipla”.

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antônio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 22/10/2013, tomaram conhecimento dos seguintes documentos:

- Solicitação de mudança de nível do estudo acima mencionado, de Departamental para Mestrado da aluna *Sheila Garcia*, em carta datada de 18 de setembro de 2013;
- Projeto de Mestrado da aluna *Sheila Garcia*, com conteúdo inalterado.

Atenciosamente,


Dr. Antônio Hugo José Fróes Marques Campos
2º Vice-Coordenador do Comitê de Ética em Pesquisa

Rua Professor Antônio Prudente, 211 • Liberdade • São Paulo / SP • CEP 01509-900
(11) 2189-5000 • www.accamargo.org.br



HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO



Ribeirão Preto, 12 de abril de 2013

Projeto de pesquisa: "Investigação de mutações de novo como possível base biológica da Esclerose Múltipla"

Pesquisadores responsáveis: Emmanuel Dias-Neto

Instituição Proponente: Fundação Antônio Prudente - Hospital do Câncer - AC Camargo

"O CEP do HC e da FMRP-USP concorda com o parecer ético emitido pelo CEP da Instituição Proponente, que cumpre as Resoluções Éticas Brasileiras, em especial a Resolução CNS 196/96. Diante disso, o HCFMRP-USP, como instituição co-participante do referido projeto de pesquisa, está ciente de suas co-responsabilidades e de seu compromisso no resguardo da segurança e bem-estar dos sujeitos desta pesquisa, dispondo de infra-estrutura necessária para a garantia de tal segurança e bem-estar.

Ciente e de acordo:

Drª Marcia Guimarães Villanova
Coordenadora do Comitê de Ética
em Pesquisa - HCFMRP-USP

Prof. Dr. José Alexandre de Souza Crippa
Coordenador Técnico Científico da Unidade
de Pesquisa Clínica - HCFMRP-USP



Termo de Consentimento Livre e Esclarecido (TCLE)

DADOS DE IDENTIFICAÇÃO DO PACIENTE OU RESPONSÁVEL LEGAL

Nome do Paciente: _____

Documento de identidade nº: _____ Sexo: M/F Nascimento: __/__/____

Endereço: _____ nº ____ Compl. ____ CEP _____

Cidade: _____ Estado: _____ Telefone: _____

DADOS SOBRE O ESTUDO

O presente estudo, intitulado “Investigação de mutações de novo como possível base biológica da Esclerose Múltipla (EM)”, é um projeto de pesquisas do Centro Internacional de Pesquisas e Ensino do AC Camargo Câncer Center em conjunto com o Departamento de Neurociências e Ciências do Comportamento da Faculdade de Medicina de Ribeirão Preto. Este estudo foi aprovado pelo Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antônio Prudente – Hospital do Câncer – AC Camargo – SP.

EXPLICAÇÕES DO PESQUISADOR AO PACIENTE OU SEU REPRESENTANTE LEGAL

Justificativas da pesquisa

Todas as células de nosso corpo contém em seu interior as informações necessárias para seu funcionamento. Essas informações estão codificadas nas moléculas de DNA presentes no núcleo das células. As informações contidas no DNA tem sido cada vez mais estudadas e já foi demonstrado que pequenos erros nessas moléculas, como as chamadas mutações, podem causar doenças como a Esclerose Múltipla. Neste projeto nosso objetivo é verificar todo o conteúdo presente nas moléculas de DNA, buscando todos os possíveis erros que possam contribuir para a Esclerose Múltipla, a fim de melhor entender essa doença. Os resultados deste estudo não trarão benefícios diretos a você, mas talvez sejam úteis no futuro, permitindo um diagnóstico mais informativo e um maior conhecimento a respeito desta doença.

Objetivos da pesquisa

Avaliar se as mutações encontradas no DNA dos pacientes com Esclerose Múltipla tem algum papel no desenvolvimento desta doença.

Procedimentos que serão utilizados e propósitos

Para participar do estudo você (o paciente) deverá:

- Ter diagnóstico de Esclerose Múltipla ou
- Ser pai ou mãe biológico de portador de Esclerose Múltipla

Caso você concorde em participar deste estudo, será feita uma coleta de quinze mililitros (15mL) de sangue por um profissional treinado. Esta coleta deverá ocorrer quando você for fazer coleta de sangue para outros exames solicitados pelo seu médico, portanto você não terá despesas adicionais decorrentes da participação nesta pesquisa, e deste modo não será ressarcida de gastos com transporte ou alimentação no dia da coleta de sangue. Este será um procedimento absolutamente normal de coleta de sangue, geralmente tem uma duração de menos de 3 minutos e não há necessidade de uso de qualquer tipo de anestesia. O sangue coletado será processado para a isolamento do DNA presente nas células do sangue. O DNA extraído ficará armazenado até a sua utilização, em freezers apropriados, localizados no Centro de Pesquisas (sob responsabilidade do Dr. Emmanuel Dias-Neto, PhD). Havendo sobra de material, este poderá ser utilizado para outros estudos futuros visando a maior compreensão do câncer humano. Você pode optar, neste momento, pelo direito de ser contactado para fornecer autorização para cada nova pesquisa que venha a fazer uso do seu material biológico ou autorizar o uso futuro do mesmo sem a necessidade de um novo contato.

Você tem a absoluta liberdade de participar ou não deste estudo sem que isto acarrete qualquer comprometimento ou alteração em seu tratamento. Caso você concorde em participar, não haverá nenhum gasto extra para você ou para a sua seguradora de saúde. Do mesmo modo, a sua participação é uma colaboração voluntária, e não haverá nenhuma forma de remuneração. Você pode desistir de participar deste estudo em qualquer momento, mesmo que você já tenha assinado este termo. Sua desistência em participar do estudo, assim como seu pedido de retirada de qualquer material armazenado, podem ser feitos a qualquer momento, e não irão trazer qualquer prejuízo ao seu tratamento. Todas as informações obtidas são sigilosas e não serão usadas para nenhum outro fim que não os objetivos desta pesquisa. Sua identidade será sempre preservada.

Se você concordar em participar, este Termo de Consentimento deverá ser assinado em 2 vias, sendo que uma via permanecerá com você e a outra com o pesquisador responsável.

Complicações e riscos esperados

Os riscos e complicações físicas associadas com a participação neste estudo são decorrentes da coleta de sangue, que será feita por profissional treinado seguindo procedimento padrão. É pouco provável que a coleta de sangue cause desconforto prolongado. Porém, se isto ocorrer, este deve ser bastante leve conseqüente apenas de uma picada de uma agulha. Todas as perguntas aqui feitas, e também o método de coleta que será utilizado neste estudo foram previamente empregados em outros projetos feitos neste e em outros hospitais do Brasil e do exterior. Todos os dados aqui levantados serão usados apenas para pesquisa científica, visando a melhor compreensão do câncer de mama. Todos os dados serão mantidos confidenciais e sua identidade será preservada em qualquer forma de divulgação científica dos mesmos.

Benefícios que poderão ser obtidos

A sua participação neste estudo não irá te trazer nenhum benefício pessoal direto. Contudo, os resultados podem contribuir para o maior entendimento dos fatores que determinam o início, a progressão ou a agressividade da Esclerose Múltipla, e desta forma podem ajudar outras pessoas no futuro.

Por favor, responda sim ou não às questões a seguir:

Foi garantido a você o direito de receber resposta a qualquer pergunta ou esclarecimento acerca dos procedimentos, riscos, benefícios e outros assuntos relacionados com a pesquisa e o tratamento?

SIM NÃO

Ficou esclarecida a você a liberdade de retirar o seu consentimento a qualquer momento, sem que isso traga prejuízo quanto à continuidade do seu tratamento neste hospital?

SIM NÃO

Ficou esclarecido que não haverá ressarcimento ou qualquer tipo de pagamento referente à realização deste estudo?

SIM NÃO

Ficou esclarecido que não haverá indenização além das previstas pela lei, em reparação a dano imediato ou tardio, causado pela pesquisa em questão?

SIM NÃO

Ficou esclarecido que a sua identidade será preservada, mantendo-se todas as informações em caráter confidencial, sob guarda do pesquisador principal, e que os dados serão utilizados somente para análises estatísticas e posteriormente para publicações científicas, sem qualquer informação que possa identificar o participante?

SIM NÃO

Ficou esclarecido a você os possíveis riscos e complicações associadas com a participação neste estudo, decorrentes da coleta de sangue?

SIM NÃO

CONSENTIMENTO PÓS-INFORMADO

De acordo com a Resolução nº 441/2011 do Conselho Nacional de Saúde, item 5: *“o consentimento livre e esclarecido referente à coleta, depósito, armazenamento e utilização de material biológico humano em Biobanco é formalizado através de TCLE, por meio do qual o sujeito da pesquisa deve se manifestar expressamente quanto às seguintes alternativas, excludentes entre si: I - necessidade de novo consentimento a cada pesquisa; e II - dispensa de novo consentimento a cada pesquisa”*.

Declaro que, após ter sido convenientemente esclarecida dos riscos e benefícios deste estudo, consinto em participar, na qualidade de paciente, deste projeto de pesquisa, e opto pela alternativa marcada abaixo:

() Quero ser contactado para fornecer novo consentimento a cada pesquisa futura que venha a fazer uso do meu material biológico.

() Dispensio ser contactado para fornecer novo consentimento a cada pesquisa futura que venha a fazer uso do meu material biológico .

Ainda de acordo com esta mesma resolução, item 5b: *“o TCLE deve conter a garantia expressa da possibilidade de acesso pelo sujeito da pesquisa, inclusive a(s) forma(s) de contato para tal, ao conhecimento dos resultados obtidos com a utilização do seu material biológico e às orientações quanto as suas implicações, incluindo aconselhamento genético quando aplicável, a qualquer tempo”*.

Deste modo, garantimos que você terá acesso, a qualquer momento, a todos os resultados gerados a partir das análises de sua amostra de sangue. Do mesmo modo é seu direito obter orientação, quando cabível, incluindo aconselhamento genético, se achados aplicáveis forem obtidos a partir da sua amostra.

Em casos de dúvidas e para ter acesso aos resultados da pesquisa, assim como às orientações quanto as implicações dos resultados obtidos, contatar o pesquisador principal, Emmanuel Dias-Neto, do Laboratório de Genômica Médica, do AC Camargo Câncer Center, no telefone 11 2189-5023. Se o pesquisador principal não fornecer as informações e/ou esclarecimentos suficientes, por favor, entre em contato com o Comitê de Ética em Pesquisa em Seres Humanos do AC Camargo Câncer Center, pelo telefone 11 2189-5000 ramal 2069, de segunda-feira à sexta-feira das 8 horas às 17 horas.

São Paulo, de

de 2015

Assinatura do sujeito da pesquisa ou do responsável legal



Assinatura do pesquisador