




Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



**Universitat Autònoma
de Barcelona**

Servei de Neurologia-Neuroinmunologia. Centre d'Esclerosi Múltiple
de Catalunya (Cemcat). Institut de Recerca Vall d'Hebron (VHIR).
Hospital Universitari Vall d'Hebron.

**Búsqueda de variantes genéticas causales
y estudio de sus implicaciones funcionales
en pacientes con esclerosis múltiple**

Elia Gil Varea

Directores: Manuel Comabella, Sunny Malhotra y Xavier Montalban

Tutor: Albert Selva O'Callaghan

Tesis doctoral

Departamento de Medicina

Programa de Doctorado en Medicina

Barcelona 2018



Los directores: el **Dr. Manuel Comabella**, jefe del laboratorio de Neuroinmunología Clínica del VHIR, el **Dr. Sunny Malhotra**, investigador principal en el laboratorio de Neuroinmunología Clínica del VHIR, y el **Dr. Xavier Montalban**, director del Cemcat, y el tutor: el **Dr. Albert Selva O'Callaghan**, jefe del Departamento de Medicina de la UAB,

CERTIFICAN:

Que la tesis titulada “Búsqueda de variantes genéticas causales y estudio de sus implicaciones funcionales en pacientes con esclerosis múltiple”, presentada por ELIA GIL VAREA para optar al grado de Doctor por la Universidad Autónoma de Barcelona, ha sido realizada bajo su dirección en el *Institut de Recerca Vall d'Hebron* y reúne las condiciones requeridas para su depósito y defensa ante el tribunal correspondiente.

Barcelona, septiembre de 2018

Dr. Manuel Comabella

Director

Dr. Sunny Malhotra

Director

Dr. Xavier Montalban

Director

Dr. Albert Selva O'Callaghan

Tutor

Elia Gil Varea

Doctorando

*La ciencia, muchacho, está hecha de errores, pero de errores útiles de cometer,
pues poco a poco, conducen a la verdad.*

Julio Verne

AGRADECIMIENTOS

Agradecimientos

Ya son cinco años en Barcelona, de los que casi cuatro los he pasado en el mundo de la tesis. Ahora, no me creo estar escribiendo esto, estar cerrando esta etapa. Es el momento de tomar consciencia de la cantidad de personas que han contribuido, de una forma o de otra, a que este trabajo haya sido posible; personas que me han apoyado y que han escuchado mis motivaciones y dramas científicos. Personas que han conseguido que desconecte cuando esta era la única solución y que me han hecho sentir que podía con ello, incluso cuando se me caía el mundo encima.

Gracias a mis tres directores. En primer lugar, a Manuel Comabella: gracias, Manolo, por conseguir que me sienta cómoda en un mundo tan desconocido y abrumador para mí cuando llegué: la genética. Gracias por haber confiado en mí y por hacerme un hueco en el grupo. Gracias a Sunny por tranquilizarme en los momentos de agobio, por comprenderme y por darme consejos sobre el camino a seguir para terminar la tesis en el momento previsto. Gracias a Xavier Montalban por haberme dado la oportunidad de desarrollar mi tesis en el Cemcat. Estoy muy orgullosa de haber formado parte de este centro y de haber experimentado cómo funciona un equipo multidisciplinar. Es admirable la conexión que existe entre todas las especialidades del centro para finalmente desembocar en el objetivo más importante: el paciente con esclerosis múltiple. Gracias a mi tutor, Albert Selva, por tu disposición para resolver mis dudas.

A todas las personas que integran el Cemcat. En especial, gracias a Luciana por la búsqueda de los datos de miles de pacientes. Gracias a Nuria, Neus, Jordi, Josep, Andrea, Dunia, Ilaria... por su trato siempre tan amable. Gracias a todos los pacientes: sin vuestra colaboración e implicación sería imposible seguir avanzando en la investigación de la enfermedad.

A la gente del laboratorio. Gracias, Mireia, porque desde el primer día me hiciste sentir integrada y cómoda. Me ha encantado compartir día tras día nuestras anécdotas. Gracias por ayudarme con la búsqueda de mis millones de muestras, tan raras y específicas, siempre con buen humor y eficacia. Gracias al Despachito Bonito 150: Laura y Clara, sois dos personas excepcionales con quienes he sentido una gran conexión. Gracias por escuchar cada día mis dramas y alegrías y por compartir conmigo los vuestros. Aunque tengamos proyectos muy distintos, siempre nos hemos intentado ayudar y he sentido vuestra implicación. Laura, me encanta poder hablar contigo de todo lo habido y por haber: siempre aprendo algo nuevo de ti. Tus anécdotas y tu forma de contarlas han hecho más amenos los días de trabajo. Clara, es genial tener cada día a un metro de distancia a una compañera que se interesa en ayudarte y, a la

Agradecimientos

vez, a una amiga a la que te hace ilusión contarle cada día tus batallitas, escuchar las tuyas, y desahogarnos mutuamente. Gracias a Carmen Espejo por escuchar siempre mis problemas científicos y por ayudarme con tus ideas. No sólo eso, sino que también hemos compartido conversaciones en el despacho y comidas donde nos hemos reído mucho. Herena, ¡la *crack* de la citometría!, mil gracias por enseñarme a diseñar experimentos de citometría, por ayudarme técnicamente y por hacerme sentir tan a gusto trabajando contigo: he tenido mucha suerte. A Nico, por tener siempre una sonrisa y alguna historia con la que alegrar el día. Gracias por hacerme sentir a gusto hablando contigo, por tu buen humor y por ese asado tan exquisito que te curraste. A Carme Costa: gracias por enseñarme a hacer inmunes y por explicarme tan bien todo. Tu optimismo, tu fuerza y la buena energía que transmites siempre ayudan. A Rucsanda, por traer al laboratorio una buena dosis de alegría, por demostrar tu interés por la investigación y por las conversaciones de gatos: ¡me encantan! A Laura Navarro, por los buenos momentos compartidos en el día a día. A las demás personas que han pasado por el grupo, especialmente a Jorge, Meri, Hartmut, Ona (¡gracias por ayudarme con el orden de las muestras de ADN!), Zoraida, Mireia, Gemma y Natalia.

A los Amigos del VHIR: gracias por confiar en los investigadores predoctorales. Sin vuestra ayuda habría proyectos que sería difícil sacarlos adelante. Vuestras visitas al laboratorio fueron muy gratificantes: siempre mostrasteis un gran interés por lo que hacíamos. Gracias también a las trabajadoras de mecenazgo, Alejandra, Sílvia, Marta, Laia y Rosa, por su implicación con los predoc y por su interés en el avance de nuestros proyectos.

A Nino Spataro, Elena Bosch y Arcadi Navarro, del Parc de Recerca Biomèdica de Barcelona, por su colaboración con la resecuenciación y el análisis bioinformático. Nino, gracias por ayudarme a resolver mis dudas respondiendo a mis emails eternos y reuniéndote conmigo.

A los grupos de la Dra. Fuencisla Matesanz y el Dr. Antonio Alcina, el Dr. Óscar Fernández y la Dra. Laura Leyva, la Dra. May Villar, la Dra. Elena Urcelay y el Dr. Rafael Arroyo, el Dr. Albert Saiz, el Dr. Lluís Ramió-Torrentà, el Dr. David Otaegui, el Dr. Koen Vandenberg y el Dr. Guillermo Izquierdo por su disponibilidad y amabilidad para colaborar mediante el envío de muestras imprescindibles para muchos estudios de esta tesis. Especialmente, gracias al grupo de May Villar y Amalia Tejada por la colaboración con la citometría.

Al grupo de Fuencisla Matesanz y Antonio Alcina, en Granada: gracias a los dos por darme la oportunidad de realizar la estancia. Gracias, Icíar e Iván, por vuestra acogida con los brazos abiertos desde el primer día. Gracias, María por tus mil consejos de laboratorio y de lugares donde hacer la compra (¡ecológica!), por nutrirme a base de nueces y chocolate y por tu grandísima participación en el ensayo de la luciferasa. Fuencisla, gracias no sólo por el periodo en Granada, sino por todo. En Granada, me hiciste sentir cómoda y me enseñaste muchísimas

cosas útiles para la tesis y para mi formación como investigadora. En el resto de mi etapa predoctoral, me has ayudado de principio a fin: me has resuelto miles de dudas en conversaciones telefónicas de horas, me has enseñado a sacar adelante experimentos y me has dado consejos e ideas sobre pasos a seguir. En definitiva, tu colaboración ha sido un pilar fundamental para esta tesis. Mil gracias por compartir tu conocimiento.

A los majos: desconectar de la rutina con vosotros siempre es renovador. Gracias, Marta, por ayudarme con mis cien mil preguntas pre-depósito y por tu constante buen humor. Tere, gracias por tu bondad y por hacer que tu compañía siempre sea de lo más agradable. Cristian, gracias por tu vitalidad, por estar siempre dispuesto a cualquier plan y por tus consejos sobre cine y series (¡muy necesarios!). Gracias, Aída, por llevar siempre contigo tan buena energía y saber contagiarla. Me has dado muy buenos consejos en esta etapa y has sabido reducir mi estrés. Josep, gracias por haberme hecho reír en tantas ocasiones y por abrirnos las puertas de tu *Royal Palace*, todavía con herederos del gremio de “biocosas”. Y, ya sabes el lema: ¡levítame! Gracias a Diego: hijo, desde el año del máster conectamos y hemos compartido grandes momentos, entre ellos el mítico guitareo. También hemos sido un apoyo mutuo muchas veces y siempre te agradeceré cómo me has escuchado. Gracias, Blanca, por hacer que la convivencia fuera perfecta durante el primer año en Barcelona. Aunque después te fuiste a tu Huesconsin, en todo momento he sabido que podía contar contigo y has escuchado mis chapas a distancia.

A Aina y Tito. Os he conocido en esta etapa de mi vida y me ha encantado compartir festivales, *room escapes*, cenas y también algún *mochi*... ¡jaja! Espero celebrar la tesis con un buen concierto de Catfolkin y una tarta de queso y chocolate ;)

A Aitziber. Gracias por mostrarme siempre tu cariño y por enseñarme que podemos estar meses sin vernos pero que, cuando quedamos, todo sigue igual.

A Salvador. Has seguido jueves tras jueves mi progreso. “- ¿Qué? ¿Cómo vas? ¿Ya tienes resultados? - La semana que viene.” Y, así, eternamente. Hasta que por fin te pude decir que ya lo tenía. Gracias por ayudarme no sólo con la voz, sino con mis preocupaciones y “paranoias”.

Al Equipo Viña Rock: gracias por hacer que cada puente de mayo sea perfecto. Durante la etapa de la tesis, han sido cuatro Viñas con vosotros que han servido para desconectar y volver con las pilas cargadas. En especial, muchas gracias a “Manu”: ha sido increíble conocerte. Me encanta hablar contigo y haber creado esta gran amistad. Gracias, Sandra, por tu alegría y simpatía. Gracias a los dos por compartir vuestra boda conmigo.

A mis amigas de siempre: Carla, Concha, Eva, Icía, Iris, María, Marta, Mónica y Sheila. Todas habéis estado presentes a lo largo de estos años para apoyarme. Icía, gracias por mostrar tanto interés hacia mi tesis, por tus palabras de ánimo y por confiar siempre en que podría con ello.

Agradecimientos

Iris, hemos crecido juntas y parte de lo que soy ahora lo he construido contigo. Hemos compartido horas y horas, siempre nos hemos entendido a la perfección y sé que siempre formarás parte de mi vida. Concha: tus ideas tan claras y tu sensatez te hacen una persona única, una amiga de la que no paro de aprender. Cuando he necesitado que me escuchen, siempre has estado. Cuando te he pedido algún consejo, nunca me has dicho lo que quería oír, sino la realidad. Te he notado muy presente durante la tesis para poder compartir los momentos de motivación y felicidad y los más duros y frustrantes. Marta, el buen rollo que transmites me han hecho sonreír en todo momento. Porque con nadie me “río en alto” como contigo. Con nadie tengo tantas cosas en común. Cuando nos vemos, consigues que te eche de menos poco después de despedirnos, pero tengo la suerte de saber que siempre estás. Concha, Marta: gracias por escuchar mis batallas y sentirnos tan cerca todos los días de la tesis.

A mis abuelos Nati y Nilo. Lo peor de esta etapa es que los dos os habéis ido en ella. Nilo, pudiste ver cómo me daban la beca para hacer la tesis, por lo que estoy segura de que te fuiste feliz porque tu única nieta consiguió su meta. Nati, gracias por enseñarme que se puede ser verdaderamente feliz a pesar de que la vida te ponga obstáculos. Eres ejemplo de vitalidad y fuerza. A mi abuela Amparo y a mi abuelo Boni, por su energía interminable: ¡yo quiero llegar así a los 91! A mis tías: Amparo, por amenizar los días eternos de tesis con las conversaciones en "Piso" y por estar siempre que lo necesito, y Ana, por acordarte de mí cada mañana, por preocuparte por mi trabajo y por tu constante buen humor. A mis tíos: Eduardo, por transmitirme tu interés hacia mí y hacia mi tesis, Michele, por cómo nos acogiste en Venecia, y Boni, por los buenos momentos compartidos. A Lara, porque aunque nos veamos poco siempre nos tenemos presentes.

A las dos personas que mejor me conocen, me cuidan y se desviven por mí: mis padres, Elia y Ángel. Sin vuestro apoyo y confianza en mí, no habría escrito esta tesis. Gracias por haber apostado por mí, por mi futuro, y por enseñarme que con esfuerzo, ilusión y constancia podría lograr mis objetivos. Gracias por aliviar mis momentos de estrés y angustia, por aconsejarme en los momentos de indecisión, por explicarme las cosas que no entendía y por hacer que volver a casa unos días sea tan renovador.

A Guille: tu apoyo incondicional me ha ayudado a no rendirme nunca. Ha sido una etapa muy intensa: días de euforia porque todo salía y días de lloros porque nada funcionaba. Y tú, todos los días, ahí. Hemos celebrado mis logros, me has dado ideas para solucionar los problemas de laboratorio, has sufrido conmigo mis momentos de desesperación y agobio con una paciencia y un buen humor increíbles y siempre has conseguido que siga. Gracias por cuidarme de una forma tan sincera cuando no tenía tiempo de cuidarme yo.

ABREVIACIONES

Abreviaciones

1,25(OH)₂-D: calcitriol	CDCV: modelo enfermedad común-variante común
25(OH)D: calcidiol	CDRV: modelo enfermedad común-variante rara
ACP: análisis de componentes principales	Cemcat: Centro de Esclerosis Múltiple de Cataluña
ADCC: citotoxicidad celular dependiente de anticuerpos	CMSP: células mononucleares de sangre periférica
ADN: ácido desoxirribonucleico	CNV: variación en el número de copias
ADNc: ácido desoxirribonucleico complementario	CXCR5: <i>C-X-C motif chemokine receptor 5</i>
AF647: Alexa Fluor® 647	CYP24A1: <i>cytochrome P450 family 24 subfamily A member 1</i>
APC: célula presentadora de antígeno/aloficocianina	DE: desviación estándar
APC·H7: aloficocianina H7	DHODH: enzima mitocondrial dihidroorotato deshidrogenasa
APP: proteína precursora amiloidea	DIS: diseminación en el espacio
ARN: ácido ribonucleico	DIT: diseminación en el tiempo
BB515: <i>Brilliant™ Blue 515</i>	DMEM: medio Eagle modificado por Dulbecco
BLC: <i>B lymphocyte chemoattractant</i>	EAE: encefalomiелitis autoinmune experimental
BCR: receptor de las células B	EDSS: escala del estatus de discapacidad expandido
BGI: <i>Beijing Genomics Institute</i>	EM: esclerosis múltiple
BHE: barrera hematoencefálica	EMPP: esclerosis múltiple primariamente progresiva
BOC: bandas oligoclonales	EMPR: esclerosis múltiple progresiva recurrente
BV605: <i>Brilliant Violet™ 605</i>	EMRR: esclerosis múltiple remitente recurrente
BV421: <i>Brilliant Violet™ 421</i>	
BWA: <i>Burrows-Wheeler Aligner</i>	
CD: célula dendrítica	
CDC: citotoxicidad dependiente del complemento	

Abreviaciones

EMSP: esclerosis múltiple secundariamente progresiva

ENCODE: *ENCyclopedia Of DNA Elements*

ETS1: *ETS proto-oncogene 1, transcription factor*

FBS: suero fetal bovino

FCRL1: *Fc receptor like 1*

FITC: isotiocianato de fluoresceína

FLI1: *Fli-1 proto-oncogene, ETS transcription factor*

FSC: dispersión de luz frontal

GAPDH: *glyceraldehyde-3-phosphate dehydrogenase*

GATK: *Genome Analysis Tool Kit*

GPCR: receptor acoplado a proteína G

GRCh37: *Genome Reference Consortium Human Build 37*

GWAS: estudio de asociación del genoma completo

Hg19: *Homo sapiens (human) genome assembly GRCh37*

HLA (o MHC): complejo mayor de histocompatibilidad (*human leukocyte antigen*)

HMM: modelo estadístico oculto de Márkov

IC: intervalo de confianza

IFN: interferón

Ig: inmunoglobulina

IL: interleuquina

IMSGC: *International Multiple Sclerosis Genetics Consortium*

iNKT: célula NKT invariante

Kb: kilobases

LCR: líquido cefalorraquídeo

LD: desequilibrio de ligamiento

Luc: luciferasa

MA: alelo minoritario

MAGNIMS: *Magnetic Resonance Network in Multiple Sclerosis*

MAF: frecuencia del alelo minoritario

MBP: proteína básica de la mielina

MOG: glicoproteína de mielina del oligodendrocitos

MR: resonancia magnética

MRI: imagen por resonancia magnética

NAbs: anticuerpos neutralizantes

NGS: secuenciación de nueva generación

NK: célula *natural killer*

NKT: célula T *natural killer*

OPC: célula precursora de oligodendrocitos

OR: odds ratio

Pb: pares de bases

PBS: tampón fosfato salino

PCR: reacción en cadena de la polimerasa

PE: ficoeritrina

PerCP•Cy5.5: peridinin clorofil proteína conjugado con cianina 5.5

PVE: potenciales visuales evocados

PWM: Matriz de peso por posición

RA: alelo de referencia

RGSI: regulator of G protein signaling 1

ROS: especies reactivas de oxígeno

Rpm: revoluciones por minuto

RPMI: medio *Roswell Park Memorial Institute*

RT-PCR: PCR a tiempo real

RXR: receptor *X retinoide*

SEM: error estándar de la media

SI: sistema inmune

SNC: sistema nervioso central

SNP: polimorfismo de un solo nucleótido

SSC: dispersión de luz lateral

TCR: receptor de las células T

Tfh: linfocito T colaborador folicular

Tfr: linfocito T regulador folicular

Th: linfocito T colaborador

TRAF3: *TNF receptor associated factor 3*

Treg: Linfocito T regulador

TSMF: Ts translation elongation factor, mitochondrial

TSS: sitio de inicio de la transcripción

UTR: región no traducida

VDR: receptor de la vitamina D

VDRE: Vitamin D response element

VEB: virus de Epstein-Barr

VHH-6: virus del herpes humano tipo 6

VR: variante rara

WT: *wild type*

XHMM: *eXome-Hidden Markov Model*

ÍNDICE

Índice

RESUMEN	25
1. INTRODUCCIÓN	31
1.1. La esclerosis múltiple: definición y etiología.....	33
1.2. Epidemiología.....	33
1.3. Formas clínicas	34
1.3.1. Esclerosis múltiple remitente recurrente (EMRR)	35
1.3.2. Esclerosis múltiple secundariamente progresiva (EMSP)	35
1.3.3. Esclerosis múltiple primariamente progresiva (EMPP)	36
1.3.4. Esclerosis múltiple progresiva recurrente (EMPR).....	36
1.4. Diagnóstico.....	37
1.4.1. MRI	37
1.4.2. Análisis del líquido cefalorraquídeo (LCR) – Bandas oligoclonales (BOC).....	38
1.4.3. Potenciales visuales evocados (PVE).....	38
1.5. Anatomía patológica	38
1.5.1. Desmielinización.....	39
1.5.2. Lesiones inflamatorias.....	39
1.5.3. Neurodegeneración y pérdida axonal.....	40
1.5.4. Gliosis reactiva	41
1.6. Inmunopatogénesis.....	41
1.6.1. Componente celular de la respuesta inmunitaria en la EM.....	41
1.6.2. Mecanismos inmunopatogénicos	49
1.7. Tratamiento.....	52
1.8. Complejidad de la EM	52
1.8.1. Factores ambientales	52
1.8.2. Factores genéticos.....	56
1.9. La genética de la esclerosis múltiple.....	56
1.9.1. Búsqueda de biomarcadores genéticos para la EM	56
1.9.2. El riesgo de recurrencia en la EM	57
1.9.3. Genes asociados con la EM	58
1.9.4. La heredabilidad en la EM.....	66
1.9.5. Variantes genéticas raras en la esclerosis múltiple	68
1.9.6. Las variantes genéticas estructurales	70

2. HIPÓTESIS Y OBJETIVOS.....	73
3. MATERIALES Y MÉTODOS.....	77
3.1. Metodología relacionada con el objetivo 1.....	79
3.1.1. Individuos y obtención de muestras de ADN genómico	79
3.1.2. Selección de los 14 genes candidatos	80
3.1.3. Proceso de resecuenciación de ADN.....	80
3.1.4. Mapeo fino de los datos crudos de la resecuenciación	82
3.1.5. Control de calidad de las muestras resecuenciadas	82
3.1.6. <i>Variant calling</i> y anotación funcional.....	83
3.1.7. Identificación de variantes genéticas comunes asociadas con la EM	84
3.1.8. Identificación de variantes genéticas raras asociadas con la EM	85
3.1.9. Identificación de variantes genéticas estructurales asociadas con la EM	86
3.2. Metodología relacionada con el objetivo 2.....	87
3.2.1. Genotipado de variantes comunes	87
3.2.2. Cuantificación de los niveles de ARNm de <i>CXCR5</i> en pacientes con EM con distintos genotipos.....	90
3.2.3. Estudio de la región genómica del polimorfismo validado en el gen <i>CXCR5</i>	92
3.2.4. Ensayo reportero dual-luciferasa para polimorfismos del gen <i>CXCR5</i>	93
3.2.5. Análisis por citometría de flujo de la expresión de <i>CXCR5</i> en poblaciones de CMSP de pacientes con EM con distintos genotipos.....	103
3.2.6. Determinación de los niveles de calcidiol (25(OH)D) y calcitriol (1,25(OH) ₂ -D) en suero de pacientes con EM con distintos genotipos del polimorfismo rs2762943 (<i>CYP24A1</i>).....	106
3.2.7. Inducción de la expresión de <i>CYP24A1</i> en respuesta a la estimulación con vitamina D en CMSP de pacientes con EM con distintos genotipos del polimorfismo rs2762943 (<i>CYP24A1</i>).....	108
3.3. Metodología relacionada con el objetivo 3.....	111
3.3.1. Cuantificación de la expresión de los genes que acumulan variantes raras en pacientes con EM	111
3.3.2. Evaluación del cambio en la expresión de <i>RGS1</i> en respuesta al tratamiento con interferón-β en pacientes con y sin variantes raras.....	113
3.3.3. Inmunofenotipado de las CMSP que expresan <i>RGS1</i>	115
3.4. Metodología relacionada con el objetivo 4.....	119
3.4.1. Validación experimental de la delección identificada en <i>FCRL1</i>	119
3.4.2. Caracterización de la delección identificada en el gen <i>FCRL1</i>	121

3.4.3. <i>Screening</i> de la delección de <i>FCRL1</i> en una cohorte independiente	121
4. RESULTADOS	125
4.1. Resultados relacionados con el objetivo 1.....	127
4.1.1. Cobertura de los datos resecuenciados	127
4.1.2. Análisis de la calidad de las muestras resecuenciadas.....	128
4.1.3. Variantes genéticas comunes asociadas con el riesgo de desarrollar EM	130
4.1.4. Variantes genéticas raras asociadas con el riesgo de desarrollar EM.....	135
4.1.5. Variantes genéticas estructurales identificadas	139
4.2. Resultados relacionados con el objetivo 2.....	141
4.2.1. Validación de las variantes comunes seleccionadas mediante genotipado en una cohorte independiente	141
4.2.2. Estudios relacionados con el polimorfismo rs10892307 (<i>CXCR5</i>).....	148
4.2.3. Estudios relacionados con el polimorfismo rs2762943 (<i>CYP24A1</i>).....	158
4.3. Resultados relacionados con el objetivo 3.....	162
4.3.1. Determinación de los niveles de ARNm de genes que acumulan variantes raras en CMSP de pacientes con EM.....	162
4.3.2. Estudios relacionados con la acumulación de variantes raras en el gen <i>RGS1</i> en pacientes con EM	167
4.4. Resultados relacionados con el objetivo 4.....	173
4.4.1. Validación experimental de la delección identificada en el gen <i>FCRL1</i> en un paciente con EM.....	173
4.4.2. Definición de los límites exactos de la delección del gen <i>FCRL1</i>	174
4.4.3. Búsqueda de nuevos individuos portadores de la delección del gen <i>FCRL1</i>	174
5. DISCUSIÓN.....	175
6. CONCLUSIONES	193
7. LÍNEAS DE FUTURO.....	197
8. BIBLIOGRAFÍA	201
9. ANEXO	221

RESUMEN

Resumen

Varios estudios de asociación del genoma completo (GWAS) han contribuido a la caracterización del componente genético de la esclerosis múltiple (EM). Estos estudios examinan cientos de miles o millones de polimorfismos de un solo nucleótido (SNPs) distribuidos por todo el genoma. Sin embargo, las variantes genéticas causales de la enfermedad y los mecanismos patogénicos por los cuales se asocian con el riesgo de EM se desconocen. Además, los GWAS analizan los SNPs comunes (frecuencia del alelo minoritario (MAF) $\geq 1\%$) y no consideran la contribución de las variantes raras. En este contexto, el objetivo de la presente tesis fue identificar variantes genéticas asociadas con la susceptibilidad de EM y estudiar sus implicaciones funcionales. La primera fase del estudio consistió en una resecuenciación de las regiones codificantes y reguladoras de 14 genes vinculados con el riesgo de EM en una cohorte de 524 pacientes y 546 controles sanos. El análisis de los datos se estructuró en tres bloques: a) la identificación de variantes comunes asociadas con el riesgo de EM; b) la detección de genes que tienden a acumular variantes raras en pacientes con EM; y c) el hallazgo de variantes estructurales asociadas con la susceptibilidad de desarrollar EM. Respecto al primer bloque, se identificaron 32 SNPs comunes diferencialmente distribuidos entre pacientes con EM y controles sanos, de las que se hizo una selección para su validación mediante genotipado en una cohorte independiente de 3.450 pacientes y 1.688 controles. Entre los validados, se escogieron dos para estudios adicionales: i) rs10892307 (*CXCR5*), para el que se identificó un SNP en alto desequilibrio de ligamiento (LD), rs11602393, cuyo alelo de riesgo se asoció con una disminución de la actividad del promotor de *CXCR5* mediante un ensayo reportero dual-luciferasa y con una reducción del porcentaje de células T reguladoras circulantes que expresaban *CXCR5* en su superficie mediante citometría de flujo; y ii) rs2762943 (*CYP24A1*), para el cual se encontró que los pacientes con EM portadores del alelo de riesgo mostraban niveles de calcitriol en suero disminuidos y tendían a niveles aumentados de la expresión génica de *CYP24A1*, codificante de una enzima que degrada la vitamina D, tras su estimulación *in vitro* con calcitriol en células mononucleares de sangre periférica (CMSP). Referente al segundo bloque, se detectaron enriquecimientos de variantes raras más frecuentemente en pacientes con EM en regiones concretas de los genes *FCRL1*, *RGS1*, *TRAF3* y *CYP24A1*, de los que el reportado en toda la secuencia de *RGS1* se asoció con una expresión génica de *RGS1* disminuida en CMSP, con un menor porcentaje de células B circulantes que expresan *RGS1* en su membrana, y con una falta de respuesta a la estimulación de CMSP *in vitro* con interferón- β , todo en pacientes con EM. El tercer bloque reveló una delección de los exones 4-10 del gen *FCRL1* en un paciente con EM pero, a pesar de validarla

Resumen

experimentalmente, no se identificaron más individuos portadores para asociarla sólidamente con la enfermedad. Los resultados obtenidos permiten concluir que la variante rs11602393 se asocia con una disminución de las células T reguladoras circulantes con capacidad migratoria hacia los centros germinales mediante la quimioatracción CXCR5-CXCL13; que el alelo de riesgo del polimorfismo rs2762943 podría relacionarse con niveles disminuidos en suero de calcitriol por una mayor actividad de la enzima CYP24A1; que la presencia de variantes raras asociadas al gen *RGS1* puede influenciar la capacidad migratoria de las células B al reducir la proporción de las mismas que lo expresan y alterar la respuesta de las CMSP al tratamiento con interferón- β , y que la delección del gen *FCRL1* todavía no puede asociarse con el riesgo de EM.

Abstract

Several genome-wide association studies (GWAS) have contributed to the characterization of the genetic component of multiple sclerosis (MS). These studies examine hundreds of thousands or millions of single nucleotide polymorphisms (SNPs) distributed throughout the whole genome. However, the causal genetic variants of the disease and the pathogenic mechanisms by which they are associated with the risk of developing MS remain unknown. In addition, GWAS are focused on common variants (minority allele frequency (MAF) $\geq 1\%$) and do not take into account the contribution of the rare genetic variants. In this context, the objective of this thesis was to identify genetic variants associated with MS susceptibility and to study their functional implications. DNA resequencing of the coding and regulatory regions of 14 genes linked to MS risk in a discovery cohort of 524 MS patients and 546 healthy controls comprised the first phase of the study. The analysis of the data was structured in three blocks: a) the identification of common variants associated with the risk of MS; b) the detection of genes that tend to accumulate rare variants in MS patients; and c) the finding of structural variants associated with the susceptibility to develop MS. Regarding the first block, 32 common SNPs differentially distributed between MS patients and healthy controls were identified, from which a selection was performed for their validation in an independent cohort of 3450 patients and 1688 controls by means of genotyping. Among the validated polymorphisms, two were chosen for additional studies: i) rs10892307 (*CXCR5*), for which a SNP was identified in high linkage disequilibrium (LD), rs11602393, whose MS risk allele was associated with a decreased *CXCR5* promoter activity by a dual-luciferase reporter assay and with a reduced proportion of circulating regulatory T cells expressing *CXCR5* on their surface by flow cytometry; and ii) rs2762943 (*CYP24A1*), for which it was found that MS patients carrying the risk allele showed diminished serum calcitriol levels and a trend for increased levels of *CYP24A1* gene expression, which codes for an enzyme responsible for vitamin D degradation, after its *in vitro* stimulation with calcitriol in peripheral blood mononuclear cells (PBMCs). In reference to the second block, rare variant enrichments were detected more frequently in MS patients within specific regions of the *FCRL1*, *RGS1*, *TRAF3* and *CYP24A1* genes, of which the one reported in the whole sequence of *RGS1* was associated with a decreased *RGS1* gene expression in PBMCs, with a lower percentage of circulating B cells expressing *RGS1* on their membrane, and with a lack of response to *in vitro* stimulation of PBMCs with interferon- β , all in MS patients. The third block uncovered a deletion of exons 4-10 in the *FCRL1* gene in a patient with MS but, despite its experimental validation, no additional carrier individuals could be identified to tightly associate it with MS. Overall, the complete set of results led to conclude that i) the variant rs11602393 is

Abstract

associated with a decrease in circulating regulatory T cells with CXCR5-mediated migratory capacity towards CXCL13 chemokine enriched germinal centers; ii) the risk allele of the rs2762943 polymorphism could be related to a decrease in serum calcitriol levels influenced by a higher CYP24A1 enzyme activity; iii) that the presence of rare variants associated with the *RGS1* gene can alter both the migratory capacity of B cells by reducing the proportion of them expressing RGS1 and the response to *in vitro* exposure of PBMCs to interferon- β treatment; and iv) that the *FCRL1* deletion still cannot be associated with the risk of MS.

1. INTRODUCCIÓN

1. Introducción

1.1. La esclerosis múltiple: definición y etiología

La esclerosis múltiple (EM) es una enfermedad inflamatoria crónica del sistema nervioso central (SNC) caracterizada por desmielinización, gliosis reactiva y daño neuronal. Cada uno de estos componentes contribuye en un grado determinado al fenotipo del paciente con EM, lo que conduce a que la enfermedad sea altamente heterogénea (1). Las manifestaciones clínicas de la EM son altamente variables e incluyen síntomas motores, sensoriales, visuales y cognitivos, sin ser ninguno de ellos específicos de la enfermedad (2).

La causa de la EM permanece desconocida. Sin embargo, la hipótesis más aceptada desde hace algunas décadas es la existencia de un proceso autoinmune en el SNC, el cual estaría condicionado tanto por una predisposición genética como por un componente ambiental (3). Los principales factores ambientales asociados con la manifestación de la EM son el tabaco, agentes infecciosos (como el virus de Epstein-Barr) y los niveles de vitamina D.

1.2. Epidemiología

La EM afecta principalmente a adultos jóvenes y se diagnostica típicamente entre los 20 y los 40 años de edad (4). Es una enfermedad que predomina en el sexo femenino, con una proporción de casi 3 mujeres por cada hombre (5).

La Figura 1 representa la prevalencia de la EM en las diferentes regiones geográficas a nivel mundial. El grupo étnico caucásico es el más afectado por la EM (6). En Europa, existe un gradiente decreciente de prevalencia de la enfermedad desde el norte hacia el sur del continente. Concretamente, en España 47.000 personas sufren EM, por lo que la prevalencia es de casi 100 pacientes por cada 100.000 habitantes. Además, según los datos de la Sociedad Española de Neurología (SEN), en nuestro país se diagnostican 1.800 casos nuevos cada año. Estas elevadas cifras suponen que la EM sea una enfermedad con un importante impacto social por ser la causa no traumática más frecuente de discapacidad neurológica en los adultos jóvenes en nuestro país (7). A nivel global, sin embargo, la prevalencia media de la EM es de 33 individuos afectados por cada 100.000 habitantes. Esta cifra reducida con respecto a España se debe a la baja prevalencia de la enfermedad en Asia oriental y en el África subsahariana (2 pacientes por cada 100.000 habitantes) (8).

Introducción

La disminución de la esperanza de vida de un paciente diagnosticado de EM es moderada; los pacientes con EM viven aproximadamente 7 años menos en comparación con la población general (75.9 años vs 83.4 años) (8,9).

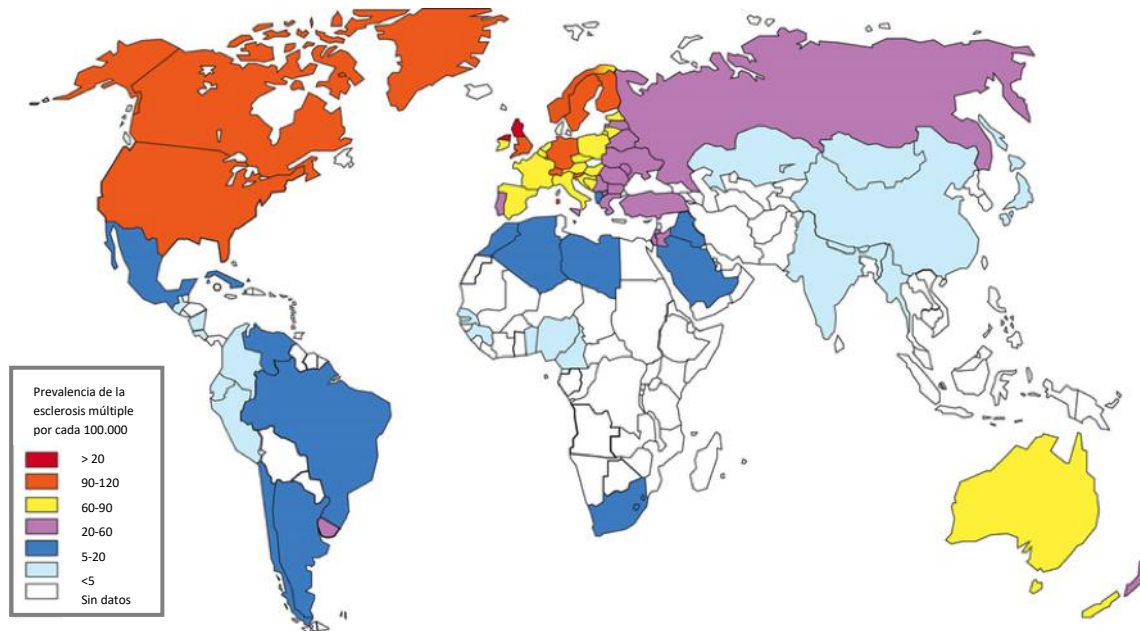


Figura 1. Distribución geográfica de la esclerosis múltiple a nivel mundial. Figura adaptada de la revisión de Shapira y col. (10).

1.3. Formas clínicas

En la mayoría de los casos, la EM cursa con la presencia de brotes que afectan a diferentes áreas del SNC (2). Un brote o exacerbación es un episodio clínico de al menos 24 horas de duración en el que aparecen nuevos síntomas de la enfermedad o un aumento de los ya presentes (11). Estos síntomas varían en función de la zona del SNC afectada y suelen ir acompañados de una mayor discapacidad; la mayoría de los pacientes con EM experimentan síntomas que incluyen movilidad, visión y función cognitiva disminuidas, así como depresión y la fatiga (12).

El curso de la EM no puede pronosticarse y es altamente heterogéneo. Lo mismo ocurre con la sintomatología. Sin embargo, se ha conseguido consensuar una clasificación de los pacientes con EM por patrones generales de presentación de la enfermedad, que han permitido que se reconozcan varios subtipos de EM (13). Las definiciones de los posibles cursos clínicos fueron

definidas por un comité internacional de expertos (14) y se revisaron recientemente para confirmar su validez (15).

El primer episodio clínico o brote con características sugestivas de EM se conoce como síndrome clínico aislado (CIS, por *clinically isolated syndrome*). El curso de la EM después del CIS es variable; el riesgo a largo plazo de desarrollar una EM clínicamente definida es del 60-80% cuando existen lesiones en la resonancia magnética (MR, por *magnetic resonance*) y del 20% cuando la imagen por resonancia magnética (MRI, por *magnetic resonance imaging*) es normal aparte de la lesión sintomática (16). De forma general, la EM puede presentarse o evolucionar siguiendo cuatro formas clínicas diferentes (Figura 2).

1.3.1. Esclerosis múltiple remitente recurrente (EMRR)

Es el tipo más frecuente. Afecta al 85-90% de los pacientes con EM (17). Este tipo de EM se caracteriza por la presencia de brotes sintomáticos seguidos de periodos de remisión en los que los síntomas mejoran o desaparecen (18). Sin embargo, a medida que avanza la enfermedad, la mejora experimentada en la remisión es menor (13). En la EMRR, los brotes coinciden con inflamación y desmielinización focalizada en el SNC.

1.3.2. Esclerosis múltiple secundariamente progresiva (EMSP)

Prácticamente el 80% de los pacientes con EMRR evolucionan hacia la forma progresiva secundaria (EMSP) una o dos décadas después del diagnóstico (13). El tratamiento preventivo con agentes modificadores de la enfermedad o inmunomoduladores ayuda a retrasar dicha progresión (18). En este tipo de EM, predomina el empeoramiento progresivo de la función neurológica respecto a los procesos inflamatorios, lo que supone una acumulación de discapacidad (*fuentes: <http://www.nationalmssociety.org>*).

Dentro de la EMSP se pueden apreciar episodios de actividad (con la presencia de brotes y/o evidencia de nuevas lesiones en la MRI) y episodios de no actividad de la enfermedad, así como episodios de progresión (evidencias de empeoramiento, con o sin brotes) y de no progresión (*fuentes: <http://www.nationalmssociety.org>*).

1.3.3. Esclerosis múltiple primariamente progresiva (EMPP)

Afecta al 10-15% de los pacientes de EM (19). Se caracteriza por un empeoramiento constante de los síntomas desde el diagnóstico. Los pacientes con EMPP no presentan brotes ni remisiones, pero en alguna ocasión sí pueden existir periodos de estabilidad (20). Curiosamente, en esta forma clínica no se observa predominancia de su incidencia en mujeres y los pacientes debutan más tarde (21).

1.3.4. Esclerosis múltiple progresiva recurrente (EMPR)

Se trata de una forma parecida a la EMPP pero sin periodos de estabilidad. El paciente presenta tanto una progresión constante como episodios de brotes. Es una forma clínica con una incidencia muy baja entre los pacientes. Cabe resaltar que existe cierta controversia sobre este término; de hecho, fue eliminado en la revisión de Lublin y col. (15).

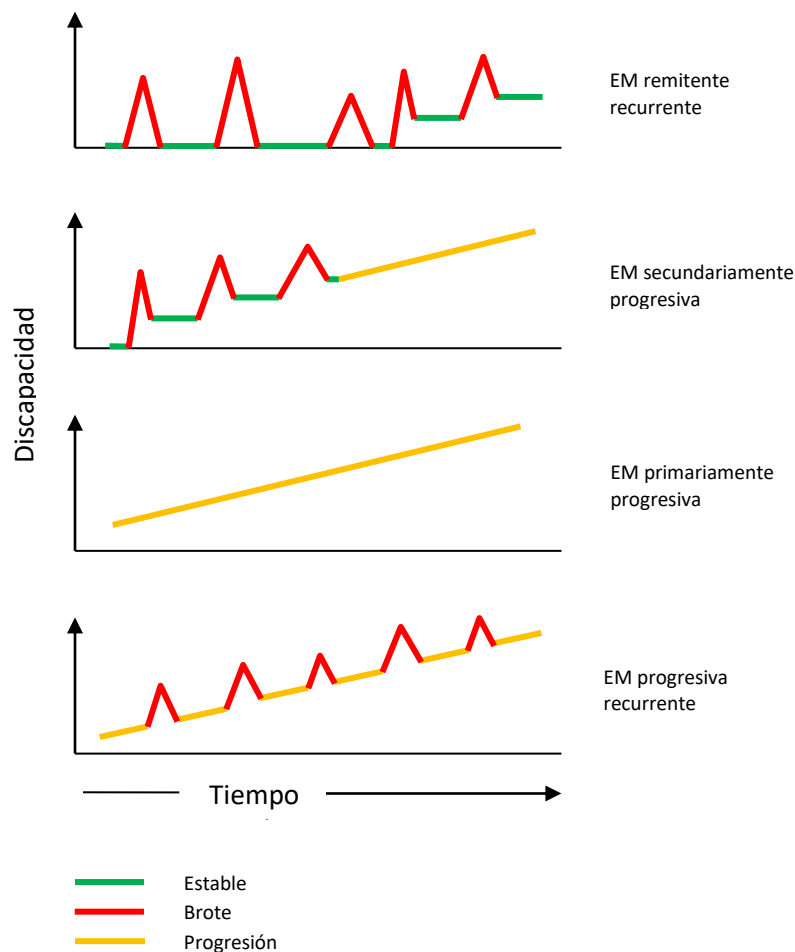


Figura 2. Diferentes formas clínicas de la esclerosis múltiple.

1.4. Diagnóstico

En el pasado, el diagnóstico de la EM era exclusivamente clínico (22). Desde que se establecieron los criterios diagnósticos de Poser y col. en 1983 se añadió la necesidad de demostrar diseminación de las lesiones desmielinizantes del SNC en el espacio (DIS, por *dissemination in space*) y en el tiempo (DIT, por *dissemination in time*) y de excluir otras enfermedades (23). Los criterios de Poser clasificaban a los pacientes con EM en función del número de brotes, las evidencias clínicas y paraclínicas, la presencia de bandas oligoclonales (BOC) de inmunoglobulina G (IgG) y/o un incremento del índice de IgG (IgG en el líquido cefalorraquídeo (LCR)/IgG en suero). Sin embargo, en 2001, McDonald y col. reevaluaron estos criterios y propusieron una serie de modificaciones, entre las que destaca la integración de la imagen por MRI en el diagnóstico (24). En 2010, el Grupo Europeo del MAGNIMS (*Magnetic Resonance Network in Multiple Sclerosis*) publicó una revisión de los criterios de McDonald 2001, donde se subrayó la posibilidad de establecer un diagnóstico de EM partiendo de una primera y única resonancia magnética (RM), sin la necesidad de un espacio de tiempo y de una segunda RM, lo que permitiría un diagnóstico precoz (25). La última revisión de los criterios de McDonald se realizó en 2017 por Thompson y col. (26). En ella, se realizaron los siguientes cambios: i) en pacientes con un CIS y una demostración clínica o de MR de DIS, la presencia de BOC permite un diagnóstico de EM; ii) las lesiones sintomáticas pueden utilizarse para demostrar la DIS o la DIT en pacientes con síndrome supratentorial, infratentorial o de médula espinal; y iii) las lesiones corticales pueden usarse para demostrar la DIS. Todo lo expuesto evidencia que los criterios puramente clínicos de McDonald están apoyados por pruebas diagnósticas complementarias, tales como las imágenes por resonancia magnética (MRI), el análisis del LCR y los potenciales visuales evocados (PVE) (27), las cuales se resumen a continuación.

1.4.1. MRI

La MR es una técnica altamente sensible para detectar lesiones desmielinizantes en el SNC (25) y crucial para el diagnóstico de la EM. Las imágenes por RM potenciadas en T2 muestran anomalías en la sustancia blanca en el 95% de los pacientes (28); sin embargo, su sola presencia no permite el diagnóstico de la enfermedad (29). La Figura 3 representa las características típicas de las lesiones encontradas en pacientes con EM mediante MRI.

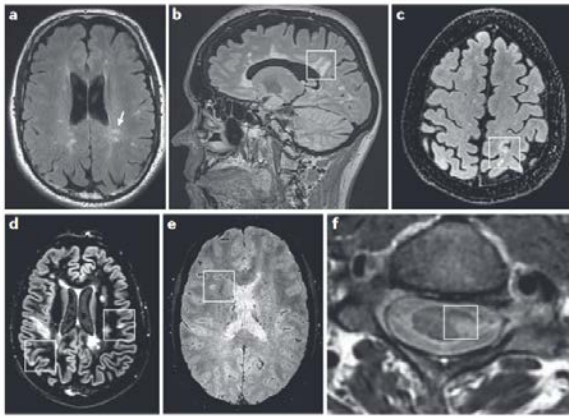


Figura 3. a y b| Lesiones periventriculares conocidas como dedos de Dawson (flechas y caja); **c|** Lesión yuxtacortical en forma de S (caja); **d|** Lesiones corticales (cajas); **e|** El signo de la vena central, típico en EM (caja); **f|** Lesión en la médula espinal (caja). Extraída de Geraldés y col. (30).

1.4.2. Análisis del líquido cefalorraquídeo (LCR) – Bandas oligoclonales (BOC)

Las bandas oligoclonales (BOC) son bandas de proteínas que se detectan en la región de las inmunoglobulinas por isoelectroenfoque y técnicas de inmunotransferencia. La presencia de BOC de inmunoglobulina G (IgG) en el líquido cefalorraquídeo (LCR), pero no en el suero, es una alteración inmunológica que sugiere la síntesis intratecal de inmunoglobulinas. Se observa en casi el 90% de los pacientes con EM (31) y cuanto mayor sea su número, peor pronóstico (32), mientras que su ausencia se asocia con un curso benigno de la EM (33).

1.4.3. Potenciales visuales evocados (PVE)

Los potenciales visuales evocados (PVE) son estudios no invasivos que determinan la respuesta electrofisiológica del sistema nervioso ante un estímulo visual externo. Se utilizan para evaluar las vías de conducción visuales a través del nervio óptico y el cerebro (34) y para localizar sitios afectados por la patología. Las respuestas se recogen colocando electrodos en el lóbulo occipital del paciente. Los potenciales visuales en algunas series son anormales en un alto porcentaje de los pacientes con diagnóstico definitivo de EM.

1.5. Anatomía patológica

Hoy en día, todavía hay quien defiende que la EM es principalmente un proceso neurodegenerativo con implicación secundaria del sistema inmune (35). Sin embargo, una inmensa evidencia sitúa a la EM como una enfermedad inmunomediada en primera instancia.

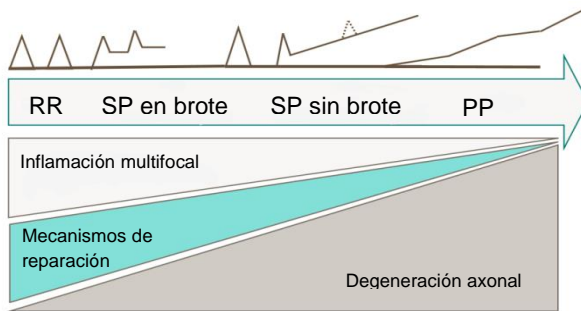


Figura 4. Balance entre los mecanismos fisiopatológicos que ocurren en la EM. RR: *EM remitente-recurrente*; SP: *EM secundariamente progresiva*; PP: *EM primariamente progresiva*. Adaptada de Grigoriadis y col. (36).

Los principales eventos patológicos que ocurren en el SNC de pacientes con EM son la desmielinización, la inflamación, la neurodegeneración y la gliosis reactiva. Los fenotipos que cursan con brotes (EMRR y EMSP) se caracterizan por el predominio de la inflamación multifocal. A medida que la EM progresa hacia fenotipos que cursan sin brotes, predomina el componente neurodegenerativo y fallan los mecanismos de reparación del SNC (Figura 4).

1.5.1. Desmielinización

La desmielinización es la pérdida de la vaina de mielina que recubre y protege a los axones de las neuronas. El principal signo patológico de la EM consiste en la presencia de placas de desmielinización multifocales en el SNC, con grados variables de inflamación, neurodegeneración y gliosis y caracterizadas por la infiltración de células inmunes, proceso ausente en el tejido nervioso normal (37). Estas placas se diseminan a lo largo de todo el SNC, pero predominan en la región cervical de la médula espinal, el nervio óptico, el tronco encefálico, el cerebelo y las regiones yuxtacortical y periventricular de la sustancia blanca (38,39). Además, aunque son más frecuentes en la sustancia blanca del SNC, también pueden encontrarse en la sustancia gris, donde son menos inflamatorias (40).

La mielina es destruida por los macrófagos, actividad que va precedida o acompañada de apoptosis de oligodendrocitos, los cuales pierden sus prolongaciones y, por tanto, su capacidad de mielinizar (41). Actualmente, se acepta la hipótesis de que la causa de la desmielinización es una reacción autoinmune contra la propia mielina y contra los oligodendrocitos productores de mielina (42).

1.5.2. Lesiones inflamatorias

La inflamación es una característica presente durante todas las etapas de la EM. Todos los pacientes con EM presentan infiltrados inflamatorios perivasculares, lo que sugiere que la desmielinización y la neurodegeneración podrían ser de naturaleza inflamatoria (39). Los infiltrados inflamatorios están formados principalmente por macrófagos activados/microglía y

linfocitos T CD8⁺ citotóxicos y, en menor medida, por linfocitos T CD4⁺ colaboradores, linfocitos B y células plasmáticas (43). Se ha observado que las lesiones desmielinizantes más activas muestran una mayor densidad de infiltrado inflamatorio (44). Además, la microglía activada produce especies reactivas de oxígeno (ROS, por *reactive oxygen species*), las cuales están implicadas en prácticamente todas las condiciones neurodegenerativas y neuroinflamatorias descritas (45) y causan disfunción mitocondrial.

1.5.3. Neurodegeneración y pérdida axonal

En las lesiones crónicas, el grado de neurodegeneración es variable (36,46), se correlaciona con el nivel de inflamación, especialmente en las formas progresivas de la EM (44), y finalmente tiene como consecuencia la pérdida total del axón mediante un proceso progresivo lento (47), el cual tiene repercusiones directas sobre la discapacidad neurológica que sufre el paciente (46). De hecho, la progresión de la EM depende en gran parte de la pérdida acumulativa de axones (2). Respecto a los mecanismos que subyacen al daño axonal, se ha observado que los defectos mitocondriales producidos por las ROS dan lugar a un estado de deficiencia energética que, a su vez, desencadena la degeneración axonal (48). Las mitocondrias también son importantes para el mantenimiento de la homeostasis del calcio, cuyo desequilibrio intra-axonal también puede resultar en neurodegeneración. Además, la proteína precursora amiloidea (APP, por *Amyloid Precursor Protein*), que es la molécula precursora del β -amiloide, se acumula en los axones dañados (49) y puede utilizarse como marcador de daño axonal, ya que no se detecta en tejido normal (46).

En los últimos años, se ha despertado el interés hacia los cambios degenerativos observados en la sustancia de gris de pacientes con EM, descritos incluso en las primeras fases de la enfermedad, los cuales son independientes de la desmielinización. Además, mientras que las lesiones de la sustancia blanca se caracterizan por la presencia de células inmunes infiltradas, en las lesiones de la sustancia gris apenas se observan estas células (50), hecho que obstaculiza aún más el estudio del vínculo entre la inflamación y la neurodegeneración en la EM. A día de hoy, aún se duda si las lesiones corticales son secundarias al daño axonal o son procesos patológicos primarios.

Por último, cabe resaltar que se postula que el fallo de la remielinización en la fase crónica de enfermedad se debe no a la destrucción de células precursoras de oligodendrocitos (OPCs) locales, sino a la incapacidad de dichas OPCs para proliferar y diferenciarse, al menos en las lesiones cerebrales (51).

1.5.4. Gliosis reactiva

La gliosis reactiva es la proliferación y activación de las células gliales como respuesta al daño del SNC. En la EM, es muy característica la presencia de astrogliosis. En ella, debido a la destrucción de neuronas próximas, los astrocitos cambian su patrón de expresión molecular, así como su morfología, y proliferan rodeando las placas de desmielinización. Como consecuencia, se forman cicatrices gliales (52).

1.6. Inmunopatogénesis

La inmunopatogénesis de la EM es un campo con todavía muchos aspectos por elucidar. Diversos estudios inmunológicos, genéticos e histopatológicos en pacientes con EM han demostrado que el sistema inmune (SI) juega un papel fundamental en la patogénesis de la enfermedad. La hipótesis de que la EM es una enfermedad autoinmune proviene principalmente de las similitudes observadas entre la EM y su modelo animal, la encefalomiелitis autoinmune experimental (EAE) (3).

1.6.1. Componente celular de la respuesta inmunitaria en la EM

En la EM, las células tanto del sistema inmune innato como del adaptativo están descompensadas.

1.6.1.1. Sistema inmunitario adaptativo o adquirido

El componente celular del sistema inmunitario adaptativo está formado por las células T y las células B. Las células T pueden encontrarse en las lesiones del SNC de pacientes con EM desde las primeras fases de la enfermedad (43). Las células B también tienen un papel importante en la EM como precursoras de células plasmáticas secretoras de anticuerpos y como células presentadoras de antígenos (APCs) para activar células T (53), aunque su abundancia en el infiltrado inflamatorio varía más a lo largo de la progresión de la enfermedad que en el caso de las células T (13).

i) Linfocitos T CD4⁺

Los linfocitos T CD4⁺ siempre se han considerado los principales mediadores de la EM en base a estudios con la EAE donde la inyección de células T CD4⁺ mielina-específicas induce la

enfermedad (54). En pacientes con EM las células T CD4⁺ reconocen concretamente a las proteínas derivadas de la mielina, como la proteína básica de la mielina (MBP, por *myelin basic protein*) o la glicoproteína de mielina del oligodendrocitos (MOG, por *myelin-oligodendrocyte glycoprotein*); pero este acontecimiento también ocurre en individuos sanos, por lo que se teoriza que el reconocimiento ocurre con mayor frecuencia o avidez en pacientes con EM (55). Otro hecho que sitúa a los linfocitos T CD4⁺ como principales impulsores de la EM es que los principales factores genéticos asociados a la EM se localizan en la región del complejo mayor de histocompatibilidad (MHC, por *major histocompatibility complex*) de clase II del genoma (56), que juega un papel fundamental en el desarrollo de la tolerancia central de las células T en el timo y en la presentación de antígenos a los linfocitos T CD4⁺ (57).

Cuando las APCs presentan un antígeno similar a algún autoantígeno del SNC a las células T CD4⁺ *naïve* en los tejidos linfoides secundarios, éstas se activan, proliferan y se diferencian a una de las poblaciones de células T CD4⁺ efectoras representadas en la Figura 5. Estas células pueden distinguirse por sus perfiles únicos de producción de citocinas, por su también única expresión de receptores de citocinas y por sus funciones exclusivas (58,59). Lo que determina que una célula T CD4⁺ *naïve* se diferencie hacia un subtipo celular u otro es el conjunto de genes que dichas células expresan, lo cual está a su vez influenciado por el conjunto de citocinas presentes en el ambiente producidas por las propias APCs (60).

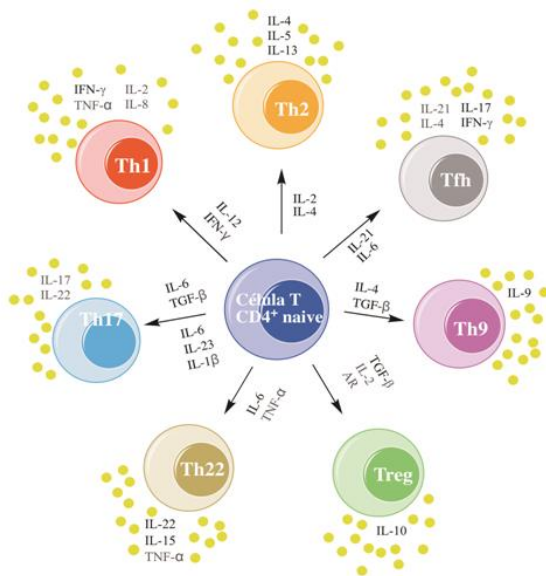


Figura 5. Diferenciación de las células T CD4⁺ naïve a las distintas poblaciones de células T CD4⁺ efectoras. En función de las citocinas predominantes en el ambiente, los linfocitos T CD4⁺ *naïve* se diferencian a un subtipo de células T CD4⁺ efectoras concreto. Una vez diferenciadas, cada tipo celular liberará citocinas específicas: proinflamatorias, antiinflamatorias y/o reguladoras. AR: ácido retinoico.

Las células Th1 y Th17 son los principales subtipos de células T CD4⁺ efectoras implicadas en la EM, aunque hay mucha controversia para decidir su importancia relativa en la patogénesis de la enfermedad y en sus diferentes etapas (61). Si el ambiente es rico en IL-12 e interferon-γ (IFN-γ), las células T CD4⁺ *naïve* se diferencian a células T efectoras 1 (**Th1**) (62), que secretan citocinas proinflamatorias, predominantemente IFN-γ, pero también TNF-α, IL-2 o IL-8 (63), y

median la respuesta inmune contra patógenos intracelulares víricos y bacterianos (64) y las reacciones alérgicas. Por otro lado, la combinación de IL-6 y TGF- β induce el desarrollo de células T efectoras 17 (**Th17**) a partir de las células T CD4⁺ *naïve* (65); además, la IL-23 es necesaria para la expansión y mantenimiento de estas células (66). Otro estudio observó que la presencia de IL-6, IL-23 e IL-1 β también induce la diferenciación a Th17 en ausencia de TGF- β (67). Las células Th17 secretan las citocinas proinflamatorias IL-17 y la reguladora IL-22 (68,69), mediante las que controlan infecciones extracelulares bacterianas y fúngicas (64). Cuando predominan IL-2 o IL-4, las células T CD4⁺ *naïve* se diferencian a células T efectoras 2 (**Th2**) (70), que producen citocinas antiinflamatorias, como IL-4, IL-5 o IL-13 (71), para eliminar parásitos extracelulares como los helmintos (72). Se ha observado que la IL-12, inductora de la diferenciación a Th1, bloquea la generación de células Th2 (73). Además, la relación Th1/Th2 es alta en pacientes con EM. Cuando en el ambiente hay una elevada concentración de IL-4 y TGF- β , las células T CD4⁺ *naïve* se diferencian en células T efectoras 9 (**Th9**) (74). Interesantemente, la presencia de únicamente IL-4 no induce la diferenciación a Th9, sino a Th2. Lo mismo ocurre con la presencia de sólo TGF- β , que daría lugar a células T reguladoras (**Treg**), y no a células Th9, con la IL-2 y el ácido retinoico (AR) contribuyendo a este proceso (75). Las células Treg producen citocinas reguladoras, como IL-10, y su función es mantener la homeostasis, activación y función linfocitaria (64), prevenir la autoinmunidad y controlar la magnitud de la respuesta inflamatoria (76). Recientemente, se ha descubierto que las células Treg promueven la remielinización en el SNC (77). Las células T CD4⁺ *naïve* se diferencian a células **Th22** en presencia de IL-6 y TNF α (78), células que principalmente secretan IL-22, citocina de naturaleza dual pro y antiinflamatoria, IL-15 y TNF α , ambas proinflamatorias (79). Se ha observado un aumento de las células Th22 mielina-específicas en la periferia y en el LCR de pacientes con EMRR. Además, se ha demostrado su resistencia al IFN- β (80). En presencia de IL-21 e IL-6, las células T CD4⁺ *naïve* se diferencian a células T auxiliares foliculares (**Tfh**, por *T-follicular helper*) (81,82), que expresan CXCR5, PD-1 y Bcl6 y producen principalmente IL-21 (83). Estas células se localizan en los órganos linfoides secundarios, donde proporcionan señales de proliferación, supervivencia y diferenciación a las células B de los centros germinales (84), lo que evidencia su papel crucial en la reacción del centro germinal (CG), esquematizada en la Figura 6. El objetivo de la reacción es que las células B maduras se diferencien a células B de memoria o a células plasmáticas productoras de anticuerpos de alta afinidad que aseguren una protección a largo plazo frente al mismo patógeno que ha desencadenado la reacción (85). Este desenlace va precedido de varias rondas de hipermutación somática, expansión clonal (ambos en la zona oscura del CG) y selección basada en su competitividad por la unión al antígeno (en la zona clara) (86). Se ha observado que las Tfh controlan la selección de las células B de alta afinidad en la zona clara y determinan si ya pueden diferenciarse o deben volver a la zona oscura y experimentar de nuevo hipermutación somática y expansión clonal (87).

Introducción

Un riesgo asociado a la reacción del GC es la posibilidad de que se produzcan mutaciones que generen clones de células B autorreactivas. Cabe añadir que se han observado formaciones de CGs ectópicos en tejidos no linfoides de pacientes con enfermedades autoinmunes (85).

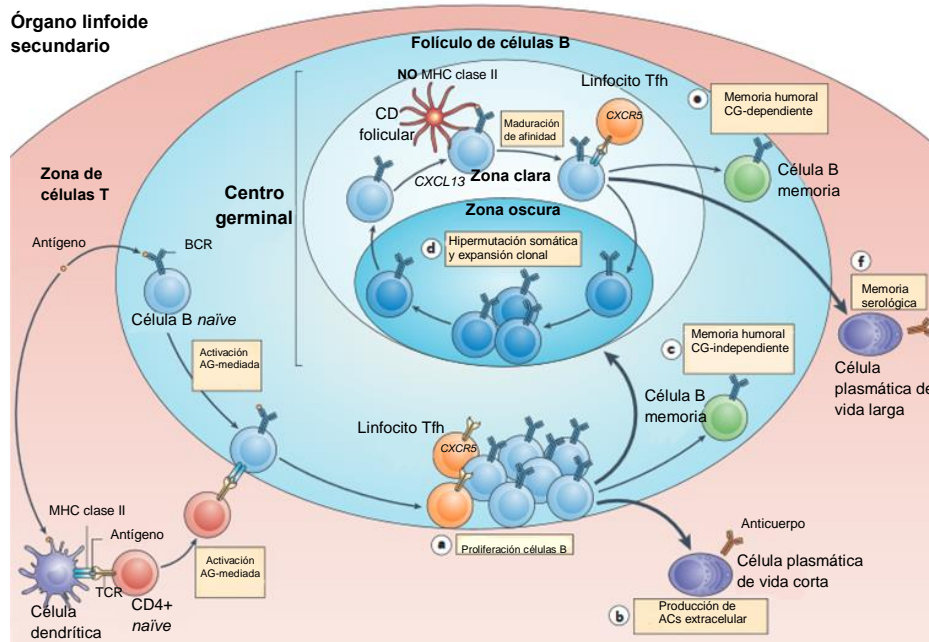


Figura 6. Reacción del centro germinal: la base de la inmunidad humoral dependiente de células T.

Los centros germinales (CGs) son estructuras especializadas dentro de los folículos de los órganos linfoides secundarios donde se produce una colaboración entre células B antígeno-específicas células Tfh y células dendríticas foliculares. Las células B *naíve* recirculan constantemente a través de los órganos linfoides secundarios en busca de infección. Cuando encuentran un antígeno, se congregan en la frontera entre la zona de células T y los folículos de células B (zona T-B) en busca de la ayuda de las células T, previamente activadas por el mismo antígeno presentado por CDs. La interacción entre las células T y las células B conlleva la proliferación de éstas con la ayuda de las Tfh, mediante interacción directa o secreción de citocinas. A partir de aquí, las células B pueden bien diferenciarse directamente a células B memoria o a células plasmáticas productoras de anticuerpos, o bien formar el CG, donde se mueven bidireccionalmente entre la zona clara y la zona oscura para experimentar varias rondas de hipermutación somática, expansión clonal y maduración de afinidad para diferenciarse a uno de los dos tipos mencionados. El número de rondas es controlado por las células Tfh. AC: anticuerpos; AG: antígeno; CD: célula dendrítica; CG: centro germinal. Adaptada de Kurosaki y col. (88).

Recientemente, se han identificado unas células en la periferia que comparten propiedades funcionales con las Tfh de ganglios linfáticos y bazo: las Tfh circulantes (89). Éstas, expresan de forma constitutiva CXCR5, PD-1 sólo cuando están activadas y Bcl6 no lo expresan (90). Por último, cabe nombrar a las células T reguladoras foliculares (**Tfr**, por *T-follicular*

regulatory), un subtipo de células Treg diferenciadas a partir de células Treg convencionales y localizadas en los centros germinales, aunque también en la circulación. Las Tfr expresan FoxP3, CD25 y CXCR5 y se encargan de suprimir la reacción del CG mediante la inactivación tanto de las células B como de las Tfh (91). En pacientes con EM se ha observado una disminución de la frecuencia de las Tfr, así como de su función supresora (92).

La producción de citocinas proinflamatorias por las células T CD4⁺ autorreactivas efectoras, especialmente por las Th1 y las Th17, y la expresión de moléculas de adhesión, enzimas proteolíticas y ROS por las células inmunes de los pacientes con EM, facilitan la extravasación de los leucocitos al SNC (93). Una vez en el SNC, las APCs locales (macrófagos, microglía y células dendríticas) presentan los péptidos de la mielina en el contexto del MHC de clase II a los linfocitos T CD4⁺ mielina-específicos que han atravesado la BHE. Así, éstos se reactivan y se produce una respuesta inflamatoria en la que liberan mediadores solubles y citocinas proinflamatorias que ejercen varias acciones: por un lado, pueden romper más la BHE y estimular la quimiotaxis, lo que resulta en una segunda oleada de reclutamiento de células inflamatorias en el SNC (3). Por otro lado, pueden destruir la mielina por sí solas. Por último, pueden activar a las células de la microglía, con la consiguiente producción de otras citocinas y ROS que también dañan a la mielina.

ii) Linfocitos T CD8⁺

Aunque clásicamente se ha situado a las células T CD4⁺ como principales mediadoras de la EM, el análisis de las células T presentes en lesiones de cerebro de pacientes con EM ha aumentado el interés hacia la implicación de las células T CD8⁺ en la patogénesis de la enfermedad al detectar su presencia. De hecho, algunos estudios han demostrado que, al contrario de lo observado en el modelo animal, las células T CD8⁺ predominan sobre las células T CD4⁺ en las lesiones desmielinizantes de los pacientes con EM (94), hallazgo identificado por primera vez en el parénquima cerebral (95) y posteriormente observado en las acumulaciones leucocitarias de los espacios perivasculares de lesiones activas de pacientes con EMPP (96). Además, el número de células T CD8⁺ en las lesiones de pacientes con EM se correlaciona con el grado del daño axonal existente (97). Sin embargo, el mecanismo exacto mediante el cual atraviesan la BHE aún se desconoce (98).

El perfil citotóxico de las células T CD8⁺ refuerza su implicación en la EM; las células T CD8⁺ producen citocinas proinflamatorias, así como moléculas con efecto citolítico, como la perforina y la granzima, que lisan sus células diana en el contexto del MHC de clase I. Una vez resuelta la infección normal, los clones experimentan apoptosis y tan solo el 10% persisten como células

Introducción

memoria (99) que protegen contra una reinfección pero que también pueden ser patogénicas si reconocen autoantígenos, Otro hecho que apoya su papel en la patogénesis de la EM es que existe una evidencia de expansión oligoclonal de células T CD8⁺ en lesiones desmielinizantes del SNC y en LCR de pacientes con EM (100). Sin embargo, no se conoce el antígeno reconocido específicamente por las células T CD8⁺ patogénicas, aunque sí se ha observado que la transferencia adoptiva de células T CD8⁺ activadas contra la mielina induce EAE severa (101). Es interesante mencionar que también existe una subpoblación de células T CD8⁺ reguladoras con efecto supresor y/o citotóxico sobre las células T CD4⁺ patogénicas. La función reguladora de estas células se media a través de una combinación de regulación mediada por el IFN- γ y citotoxicidad mediada por la perforina (102). Se ha observado que en pacientes con EM esta subpoblación tiene la capacidad citolítica disminuida, favoreciendo la presencia de células T CD4⁺ autorreactivas (103). Por último, otro hallazgo que apoya la involucración de las células T CD8⁺ en la fisiopatología de la enfermedad es la asociación de varios polimorfismos en la región del MHC de clase I con la EM (104–106). Por todo ello, el desarrollo de tratamientos específicamente dirigidos hacia las células T CD8⁺ podría ser pertinente.

iii) Linfocitos B

La inmunidad humoral también juega un papel importante en la patogénesis de la EM. Se ha descrito expansión clonal de células B autorreactivas en lesiones del SNC, en LCR y en sangre de pacientes con EM (107,108). En el SNC, las células B forman parte del infiltrado inflamatorio presente en las lesiones activas de los pacientes con EM, donde principalmente se encuentran en el espacio perivascular, pero también en el parénquima (109). Además, se ha observado que las células B forman agregados de tipo folicular en el espacio subaracnoideo de las meninges de algunos pacientes con EM, cuya organización suele ser en forma de centros germinales. Estos agregados son característicos de formas progresivas de EM (110), aunque también se han observado en estadios tempranos de EMRR (111). Su permanencia en el SNC puede ser porque éste se ha definido como un ambiente propenso a fomentar la supervivencia y desarrollo de las células B, ya que las células residentes del SNC producen mediadores como CXCL12 (CXC-chemokine ligand 12, cuyo receptor CXCR4 está altamente expresado en células B), BAFF (*B cell activating factor*) o APRIL (*a proliferation-inducing ligand*) (112). Además, se observa una presencia continua de BOC de inmunoglobulina G (IgG) en el LCR de más del 95% de los pacientes con EM, un producto de la expansión clonal de células B (113). Respecto a la sangre de pacientes con EM, también se ha observado la acumulación de clones de células B autorreactivas, lo que puede deberse a que la función defectuosa de las células T

reguladoras reduce la tolerancia periférica a las células B (114), al igual que se describió para las células T autorreactivas.

La participación de las células B en la inmunopatogénesis de la EM también se apoya en la observación que el tratamiento con anticuerpos monoclonales anti-CD20, el cual depleciona las células B inmaduras y maduras pero no las células plasmáticas CD20⁺, reduce rápidamente la formación de nuevas lesiones en el SNC (53). Las principales funciones patogénicas de las células B en la EM son: i) la secreción de autoanticuerpos que, aunque aún no se conoce con certeza su especificidad, sí se han propuesto algunos autoantígenos como posible diana, tales como MOG (115); ii) la liberación de citocinas proinflamatorias (IL-6, LT, TNF, etc.), habiéndose demostrado que la secreción de IL-6 por células B es mayor en pacientes con EM que en controles sanos (116) y que el TNF puede inducir la apoptosis de oligodendrocitos (117); o iii) la presentación de antígenos a los linfocitos T CD4⁺ patogénicos via el MHC de clase II, con su consiguiente activación (118), campo aún poco estudiado.

1.6.1.2. Sistema inmunitario innato

El sistema inmune innato es la primera línea de defensa frente a patógenos. Varios estudios recientes sugieren que también participa en la patogénesis de la EM, principalmente mediante la influencia sobre la función efectora de las células T y B, las cuales a su vez expresan citocinas y marcadores de activación que activan aún más a las células del sistema inmune innato (119).

i) Células dendríticas

Las células dendríticas (CDs) son APCs profesionales cuya interacción con los linfocitos T CD4⁺ *naïve* es crítica para la diferenciación de éstos a células efectoras (120). Además, las CDs son capaces de activar a las células *natural killer* (NKs) mediante la liberación de IL-12 e IL-15 (121) y de modular su proliferación y función mediante la secreción de IL-27 (122). También se ha descrito que liberan osteopontina, una citocina proinflamatoria que afecta a la progresión de la EM (123) y que promueve la supervivencia de las células T activadas (124). En pacientes con EM, las CDs abundan en lesiones inflamatorias del SNC, en el LCR y en la circulación (125), donde además presentan alteraciones fenotípicas y/o funcionales respecto a individuos sanos (126): producen altos niveles de TNF- α , IFN- γ e IL-6 (citocinas proinflamatorias), lo que va acompañado de una mayor respuesta inflamatoria de las células T. Todo ello sugiere un estado activo y proinflamatorio de las CDs, contribuyendo así a la patogénesis de la EM (127).

ii) Fagocitos (macrófagos y microglía)

En pacientes con EM, la destrucción de la mielina en el SNC también se asocia con células con capacidad fagocítica o fagocitos (128), de los que existen dos tipos que difieren en su origen. En primer lugar, en el SNC existen macrófagos derivados de monocitos periféricos. Éstos presentan un fenotipo activado con una expresión aumentada de citocinas proinflamatorias y metaloproteinasas de matriz extracelular (MMPs, por *matrix metalloproteinases*) en pacientes con EM, lo que facilita su infiltración en el SNC, donde se diferencian a macrófagos que fagocitan la mielina. En segundo lugar, las células de la microglía residente activada de forma local también actúan como macrófagos en el SNC, los cuales presentan antígenos, producen citocinas (129,130) y se activan en respuesta a la neurodegeneración, infecciones, o inflamación (131). En la EAE se ha observado que los macrófagos derivados de monocitos parecen iniciar la desmielinización, mientras que los derivados de la microglía eliminan *debris* celulares y, por tanto, promueven la recuperación del tejido (132). Otro reciente estudio ha observado una dualidad funcional dentro de la microglía: por un lado, expresan moléculas antiinflamatorias, se encargan de la fagocitosis de *debris* celulares y liberan factores neurotróficos, lo que favorecería la remielinización; por otro lado, actúan como APCs, secretan moléculas proinflamatorias y fagocitan mielina, lo que potenciaría la desmielinización (133). Merece la pena destacar que los macrófagos son las células predominantes en las lesiones desmielinizantes y su número se corresponde con el daño tisular. En la EAE, la activación de los macrófagos mediada por células T es crítica para la desmielinización inflamatoria (134). De hecho, la microglía y las células T infiltradas en el SNC están muy próximas en las lesiones de pacientes con EM y en la EAE, lo que sugiere una interacción entre ellas. Se ha propuesto una estimulación recíproca: las células T activan a la microglía y la microglía ya activada es capaz de activar a las células T (135).

iii) Células natural killer

Las células *natural killer* (NK) son linfocitos con un papel clave en la conexión entre el sistema inmune innato y el adaptativo. En base a la expresión de CD56 en su superficie, se clasifican como células efectoras (CD56^{dim}) o reguladoras (CD56^{bright}). Las células NK CD56^{dim} muestran mayor actividad citotóxica que las CD56^{bright} (136) y expresan las enzimas proteolíticas perforina y granzima (137), mientras que las células NK CD56^{bright} expresan bajos niveles de dichas enzimas (138) y secretan grandes cantidades de citocinas inmunoregulatorias (136).

En pacientes con EM, en cuyas lesiones desmielinizantes se han observado células NK (139), se han atribuido tanto funciones patogénicas como protectoras a diferentes subpoblaciones de células NK en el SNC en función del subtipo celular y del ambiente. Por un lado, se ha

observado la actividad citotóxica de algunas NKs hacia oligodendrocitos, astrocitos y microglía durante la inflamación; por otro lado, existen células NK que participan en la protección y reparación del SNC mediante la producción de factores neurotróficos (119).

iv) Otras células del SI innato

Otras células del sistema inmune innato a destacar son los mastocitos, las células NKT invariantes (iNKT, por *invariant natural killer T cells*) y las células T $\gamma\delta$. Los mastocitos son células mieloides que contienen gránulos citoplasmáticos en cuyo interior se encuentra la histamina, así como otros mediadores químicos y citocinas (140) que, al ser secretados, actúan sobre CD4⁺, neutrófilos y linfocitos y pueden afectar a la diferenciación de los linfocitos T CD4⁺ *naïve* a efectores (141). En pacientes con EM se ha detectado la presencia de mastocitos en las placas de desmielinización inflamatoria del SNC (142). Las células iNKT son una población que comparte propiedades con las células T y con las NK caracterizada por expresar las cadenas $\alpha\beta$ invariantes del receptor de las células T (TCR) y moléculas de superficie presentes en las células NK (143). Estas células producen citocinas implicadas en la prevención de la autoinmunidad, como IL-10 y TGF- β (119). En pacientes con EM, el número total de células iNKT en sangre periférica está disminuido (144). Por último, las células T $\gamma\delta$ son un subtipo de linfocitos T que, a diferencia de las células T $\alpha\beta$, ejerce su función como parte del sistema inmune innato y se caracteriza por poseer un TCR diferente originado a partir de una reorganización distinta del gen que lo codifica (145). Se localizan en la piel y la mucosa (146), aunque también se han observado en la circulación (147). En pacientes con EM, se ha detectado su aumento en LCR (148), en lesiones activas y crónicas (149) y en sangre periférica (150).

1.6.2. Mecanismos inmunopatogénicos

No hay una respuesta indudable a la pregunta de si la EM se origina en la periferia o en el SNC, aunque sí está esclarecido que en ambos lugares ocurren procesos inmunopatogénicos que subyacen a la enfermedad.

1.6.2.1. Cambios inmunológicos periféricos

En la ontogenia de los linfocitos T en el timo, la mayoría de los autorreactivos son eliminados gracias a la tolerancia central (151). Pero la imperfección de este proceso conduce a que algunos se liberen a la periferia, donde ocurren los mecanismos de tolerancia periférica, como la

inactivación directa de los linfocitos T autorreactivos mediante delección clonal o inducción de anergia (152), o los mecanismos inmunosupresores ejercidos por las células Treg para que no se activen (153,154). En sangre periférica y en LCR de pacientes con EM se han encontrado linfocitos T autorreactivos específicos contra la mielina y, aunque también pueden encontrarse en individuos sanos, los presentes en los pacientes con EM muestran un fenotipo más activo, una mayor avidéz por la mielina (55) y una mayor producción de citocinas proinflamatorias (155). Estas diferencias conducen a pensar que en la EM se rompe la tolerancia periférica, lo que puede deberse a la alteración tanto de factores intrínsecos a las células T (mecanismos transcripcionales o epigenéticos, receptores de citocinas, etc.) como de factores extrínsecos (descompensación de las células Treg) (156,157). La hiperactivación de los linfocitos T autorreactivos va acompañada de una sobreexpresión de moléculas de adhesión que los hace más propensos a interactuar con la BHE y, así, a infiltrarse en el SNC, donde median la respuesta inflamatoria contra la mielina (3). Dicha infiltración ocurre junto con monocitos y linfocitos B activados (158). El mecanismo exacto de activación de los linfocitos T autorreactivos se desconoce, pero existen varias hipótesis. Una de ellas propone que la activación ocurre mediante mimetismo molecular: las células T CD4⁺ *naïve* que reconocen a un epítipo presente en un agente infeccioso se encuentran con las APCs que les presentan dicho patógeno en los tejidos linfoides secundarios, se activan y, posteriormente, se produce una reactividad cruzada contra un antígeno propio que comparte el epítipo desencadenante (159,160). Otra hipótesis sostiene que existen receptores de los linfocitos T (TCRs) degenerados que reconocen epítopos endógenos de la mielina presentes de forma constitutiva en los ganglios linfáticos cervicales (161–163), proceso conocido como *epitope spreading*. Una tercera hipótesis es que cuando se produce una infección por un virus irrelevante pero que desencadena la liberación de diferentes citocinas, éstas pueden provocar una activación *bystander* de células T autorreactivas (164). Por último, un estudio identificó en 2010 células T CD8⁺ que expresan un TCR dual específico tanto para la MBP como para antígenos virales, cuya activación también podría inducir reacciones autoinmunes (165).

1.6.2.2. Migración celular y cambios en el sistema nervioso central

Como se mencionó previamente, algunos autores proponen que la EM se inicia en el SNC, de modo que eventos intrínsecos al mismo, aún sin elucidar, podrían desencadenar el desarrollo de la enfermedad, con la infiltración de linfocitos autorreactivos como fenómeno secundario. Uno de los eventos postulados como candidatos es una reacción inflamatoria ante una infección viral en el SNC (158). A pesar de no existir una confirmación del lugar de origen de la enfermedad, sí

se sabe con certeza que la migración de linfocitos autorreactivos activados desde la periferia al SNC atravesando la BHE es un paso crucial para la formación de lesiones en la EM.

La BHE, que limita la entrada de proteínas y células del torrente sanguíneo al SNC, está formada por células endoteliales unidas mediante uniones estrechas (166). Los pericitos (167), las células microgliales perivasculares, los astrocitos y las neuronas contribuyen a la regulación funcional de las uniones estrechas y al mantenimiento de la integridad de la BHE (168). Mientras que los linfocitos T en reposo tienen capacidad limitada para invadir el SNC, los linfocitos T activados expresan moléculas (receptores de quimiocinas, citocinas inflamatorias, moléculas de adhesión celular (MACs), integrinas, MMPs y ROS) que promueven su extravasación al SNC mediante la afectación sobre la permeabilidad de la BHE (93). Las células endoteliales de la BHE expresan MACs que reconocen con los ligandos presentes en los leucocitos activados, cuya interacción media la trans migración de las células inmunes al SNC. Una vez en el SNC, los linfocitos T necesitan ser reactivados por APCs locales para ejercer el daño (155). A modo de resumen, la Figura 7 representa la inmunopatogénesis de la EM.

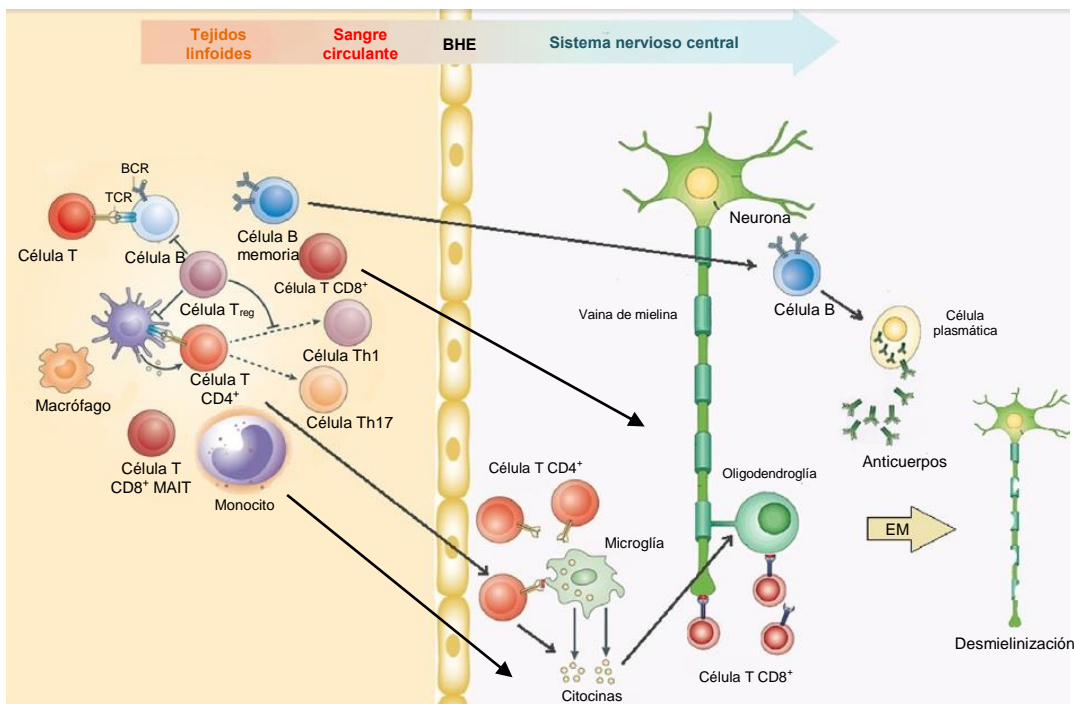


Figura 7. Inmunopatogénesis de la EM. Las células del sistema inmune son activadas en la periferia y atraviesan la BHE mediante su interacción con las moléculas de adhesión de las células endoteliales de la BHE. Ya en el SNC, la reactivación de las células inmunes por las APCs locales mediante la presentación de péptidos de la mielina desencadena la producción de citocinas proinflamatorias, que pueden romper más la BHE y atraer a nuevas células inmunes, pueden atacar a la mielina directamente, o pueden activar a las células de la microglía, las cuales también dañarían la mielina. Adaptada de Dolati y col. (169).

1.7. Tratamiento

El objetivo de los tratamientos es reducir la frecuencia y severidad de los brotes y el desarrollo de nuevas lesiones, así como ralentizar la progresión de la discapacidad de los pacientes. Algunos neurólogos hablan de poder “congelar” el estado patológico del paciente. Sin embargo, no existe ningún tratamiento eficaz que frene completamente la evolución de la EM. La alta heterogeneidad clínica y patológica de la EM implica que se requiera más de un enfoque a la hora de tratar la enfermedad, de los que el tratamiento con interferón-beta (IFN- β) es el más común.

Los interferones son citocinas que participan en la regulación de la respuesta inmune frente a patógenos (170). La acción del IFN- β desencadena los siguientes resultados: limitación de la migración leucocitaria a través de la BHE, incremento de la producción de células NK CD56^{bright}, estimulación de la producción de factores tróficos *in vitro*, como NGF, que median en la reparación de tejidos, aumento de la producción de citocinas antiinflamatorias y reducción de la producción de citocinas proinflamatorias (171). Además, un estudio observó que el IFN- β puede suprimir la proliferación de células T CD4⁺ y CD8⁺ autorreactivas en la periferia (172). Otros tratamientos contra la EM son el acetato de glatirámico (AG), el dimetil fumarato (DMF), la teriflunomida, los anticuerpos monoclonales (mAbs) anti-CD20, el natalizuma, el fingolimod o el alemtuzumab.

1.8. Complejidad de la EM

La EM es una enfermedad de etiología desconocida y de patogenia compleja que se desarrolla como resultado de la convergencia de factores genéticos y ambientales con eventos reguladores post-genómicos y epigenéticos, como reorganizaciones genéticas, *splicing* y edición del ARNm, metilación o microARNs.

1.8.1. Factores ambientales

En este apartado se abordarán los principales factores ambientales que se han descrito como factores de susceptibilidad para desarrollar EM.

1.8.1.1. La vitamina D

La vitamina D es una molécula liposoluble que desempeña un papel importante en la regulación de la función inmune y en la reducción de la inflamación (173). Se ha reportado su efecto protector para la EM, lo que se basa en una disminución del riesgo de desarrollar la enfermedad asociada a la exposición solar, ya que la vitamina D es sintetizada por la piel cuando ésta se expone al sol (174), y a la ingesta de suplementos de vitamina D (175). Además, altos niveles en suero de la forma 25(OH)-D, el precursor de la vitamina 1,25(OH)₂-D, se correlacionan con una reducción en la actividad de la EM observada mediante RM, con una tasa de progresión más lenta y con un menor número de brotes de los pacientes (176–178). También se ha observado que los niveles en suero tanto de 25(OH)-D como de la forma fisiológicamente activa de la vitamina D, 1,25(OH)₂-D, son menores en pacientes con EMRR que en controles sanos; sin embargo, curiosamente, no se observa esta diferencia entre pacientes con EMPP y controles sanos (179).

1.8.1.2. Gradiente latitudinal

Muchos estudios indican que la incidencia y la prevalencia de la EM en una población son mayores cuanto más alejada esté del ecuador (180), aunque otras investigaciones señalan que esta correlación podría ser un artefacto (181,182).

1.8.1.3. Tabaco

Diversos estudios han relacionado fumar tabaco con un mayor riesgo de desarrollar EM (183), asociación aún más fuerte en hombres que en mujeres (184). Varios autores indican que fumar tabaco también tiene una influencia negativa sobre la progresión de la EMRR hacia formas progresivas (185,186), aunque estos resultados no han sido uniformes en todos los estudios. Además, se ha observado que el riesgo de desarrollar anticuerpos neutralizantes (NAbs, por *neutralizing antibodies*) contra el IFN- β es mayor en fumadores que en no fumadores (187).

1.8.1.4. Dieta y obesidad

Varios estudios epidemiológicos indican que un alto consumo de grasa animal saturada se asocia con una mayor incidencia de EM y que la ingesta de suplementos de ácidos grasos omega-3 podría influir positivamente en el curso de la EM (188). Durante los últimos años, la

Introducción

dieta mediterránea está adquiriendo relevancia a la hora de combatir la EM y un mayor consumo de frutas y verduras se ha relacionado con un menor riesgo de desarrollar EM (189). Sin embargo, en este campo también hay resultados contradictorios (190) y muchos autores señalan que la dieta mediterránea simplemente favorece la salud de la población general, y no específicamente la de los pacientes con EM. La obesidad sí está más sólidamente asociada a la EM; varios estudios demuestran que la obesidad en la infancia y en la adolescencia son factores de riesgo para la susceptibilidad a la EM, asociación demostrada consistentemente en mujeres y de forma menos sólida en hombres (191). Además, un estudio reciente indica que un índice de masa corporal (IMC) genéticamente elevado se asocia con un mayor riesgo de EM, lo que proporciona otra evidencia del papel causal de la obesidad en la etiología de la EM (192).

1.8.1.5. Microbiota

La microbiota intestinal es el conjunto de bacterias, arqueas, hongos, eucariotas y virus que residen en las superficies mucosas del intestino, mientras que su material genético se conoce como microbioma. Actualmente, el estudio del papel de la microbiota comensal en la patogénesis de la EM está en expansión debido a que diversas investigaciones con EAE demuestran que la microbiota intestinal i) puede modular el sistema inmune del huésped, ii) altera la integridad y la función de barreras biológicas, tales como la barrera intestinal, y iii) tiene un efecto directo sobre las células residentes del SNC (193). Estudios con humanos han demostrado que los pacientes con EM poseen una disbiosis microbiana intestinal respecto a los individuos sanos. Este desbalance del equilibrio microbiano se caracteriza por un incremento en la abundancia de géneros con efecto proinflamatorio en pacientes con EM, mientras que los individuos sanos muestran una mayor abundancia de géneros con efecto regulador sobre los linfocitos T (194).

1.8.1.6. Higiene

La hipótesis de la higiene sugiere que la escasa exposición a patógenos durante la infancia puede aumentar el riesgo de desarrollar alergias o enfermedades autoinmunes en individuos susceptibles (195). Numerosos estudios han observado que la prevalencia de la EM en los países desarrollados excede la de las regiones caracterizadas por infecciones infantiles generalizadas o por una higiene deficiente. Específicamente, un estudio concluyó que existe una correlación negativa entre la incidencia de la EM y la presencia del helminto de distribución mundial *Trichuris trichiura*, el cual predomina en los lugares con una higiene menor (196). Además, dos recientes meta-análisis han demostrado que la infección con *Helicobacter pylori* se correlaciona

negativamente con la EM (197,198). También se ha hipotetizado que la infección con el parásito *Toxoplasma gondii* podría proteger contra la EM (199). A modo de conclusión, se puede decir que la higiene excesiva puede limitar la exposición a compuestos inmunomoduladores de los microorganismos comensales que ofrecen protección contra enfermedades inmunomediadas.

1.8.1.7. Infecciones virales

Cada vez hay más evidencias de que los virus pueden desempeñar cierto papel en la patogénesis de la EM, siendo el virus de Epstein-Barr (VEB) principalmente y el virus del herpes humano tipo 6 (VHH-6) los más sólidamente asociados (200,201). El hallazgo más importante que asocia al VEB con la EM es una mayor seroprevalencia y títulos más altos de anticuerpos contra el VEB en pacientes con EM respecto a individuos sanos (202). Se han propuesto varios posibles mecanismos biológicos que podrían explicar cómo el VEB puede inducir la aparición de EM, entre los que destacan dos: en primer lugar, el mimetismo molecular, es decir, la respuesta inmune de linfocitos autorreactivos que reconocen la MBP frente a péptidos compartidos entre ésta y el VEB, de modo que, ante una infección, dichos linfocitos se activarían y podrían atravesar la BHE, causando una reacción inflamatoria en el SNC al reaccionar contra la mielina (161,203). El segundo mecanismo sería la activación inespecífica (*bystander*) de células del sistema inmune VEB-específicas como consecuencia de la infección por dicho virus, causando un ataque al propio sistema inmune (204). Las principales células candidatas a mediar esta acción son los linfocitos T CD8⁺, ya que se ha demostrado que la frecuencia de células T CD8⁺ específicas para los antígenos líticos del VEB aumenta en la sangre durante el brote en pacientes con EMRR (205). En el caso del VHH-6, su ADN, ARN o proteína se encuentran aumentados en distintos fluidos corporales y en tejido cerebral de pacientes con EM comparados con controles sanos (206). También se han propuesto diferentes mecanismos de acción del VHH-6 en EM, como son el mimetismo molecular de algunos péptidos del VHH-6 con la MBP (162), la infección de astrocitos en el SNC con la consiguiente creación de un entorno proinflamatorio que estimula la infiltración de leucocitos (207), la promoción de la infiltración leucocitaria en el SNC por la quimioatracción desencadenada por la similitud entre proteínas codificadas por el VHH-6 y receptores humanos de quimiocinas o la infección de células endoteliales, aumentando su capacidad de secretar quimiocinas proinflamatorias y reclutando así células inmunes al SNC (208).

1.8.2. Factores genéticos

La variación genética constituye aproximadamente el 30% del riesgo total de la EM (158). Las numerosas investigaciones han identificado un número de variantes genéticas asociadas a la enfermedad que, en su conjunto, constituyen más de la mitad del componente genético de la EM. A continuación, se desarrollará un apartado donde se detalla el camino que ha seguido la búsqueda de los genes que determinan el riesgo de desarrollar EM a lo largo de los años, así como las regiones más fuertemente asociadas.

1.9. La genética de la esclerosis múltiple

1.9.1. Búsqueda de biomarcadores genéticos para la EM

Dado que la EM no sigue un patrón de herencia mendeliana, la búsqueda de variantes genéticas asociadas a la enfermedad es esencial para construir la estructura genética de la misma. El objetivo de ello es identificar variantes que puedan clasificarse en la categoría de biomarcadores para la EM, los cuales se definen como cualquier sustancia, estructura o proceso que puede medirse en el cuerpo, o sus productos, e influir o predecir la incidencia de un resultado o enfermedad (209). Toda nueva identificación de un biomarcador debe incluir una fase de descubrimiento en un número limitado de muestras y una de validación de los resultados (210).

En el caso de la EM, los biomarcadores se clasifican en cuatro categorías según su capacidad para i) predecir la susceptibilidad de desarrollar EM, ii) asegurar un diagnóstico correcto, iii) constituir un indicador de la forma clínica o la actividad de la enfermedad, e iv) identificar pacientes con riesgo de desarrollar efectos adversos ante un tratamiento (210). Los biomarcadores genéticos pueden incluirse dentro de cualquiera de las cuatro categorías y consisten en cualquier variación a nivel de la secuencia del ADN, aunque los polimorfismos de un solo nucleótido (SNPs, por *Single Nucleotide Polymorphisms*) son las variaciones más utilizadas debido a la disponibilidad de tecnologías de secuenciación masiva y genotipado de alto rendimiento (211). El objetivo final es identificar variantes genéticas con una frecuencia del alelo minoritario diferente entre dos grupos, en función del parámetro que se esté estudiando (susceptibilidad a la EM, respuesta al interferón- β , etc.). La importancia de los biomarcadores genéticos en una enfermedad tan heterogénea como la EM reside en su utilidad bien para determinar la susceptibilidad a la misma, lo que permitiría un diagnóstico temprano, o bien para encuadrar a cada paciente en un contexto patológico desde el inicio de la EM, pudiendo así estratificar a los pacientes y seleccionar tratamientos óptimos personalizados para cada uno.

1.9.2. El riesgo de recurrencia en la EM

Aunque la EM no es una enfermedad hereditaria, diversos estudios han demostrado que presenta cierta agregación familiar, pero sin un patrón de herencia mendeliana. A finales del siglo XX se publicaron los primeros trabajos que evidencian la contribución de los genes a la etiología de la EM (212). Como muestra la Figura 8, estos trabajos revelan un aumento en la incidencia de la EM en familiares de primer grado de consanguinidad con respecto a la población general. Además, diversos estudios genéticos han documentado una concordancia del 25-30% entre gemelos monocigóticos frente al ~5% en gemelos dicigóticos (213). También se ha observado que cuando ambos progenitores están afectados por la EM, el riesgo de que su descendencia desarrolle la enfermedad es mayor que si sólo está afectado uno de ellos (214). Además, el riesgo de recurrencia asociado a medios hermanos es aproximadamente la mitad que el asociado a hermanos completos (215). Por último, cabe resaltar que en estudios realizados con pacientes con EM que habían sido adoptados, no se observó una incidencia mayor en la familia adoptante respecto a la población general (216), lo que resalta el importante papel de los genes en la susceptibilidad de desarrollar EM y que no sólo participa el componente ambiental. Sin embargo, el hecho de que no siempre exista concordancia para desarrollar EM en individuos genéticamente idénticos (gemelos monocigóticos) confirma que existen otros factores de riesgo.

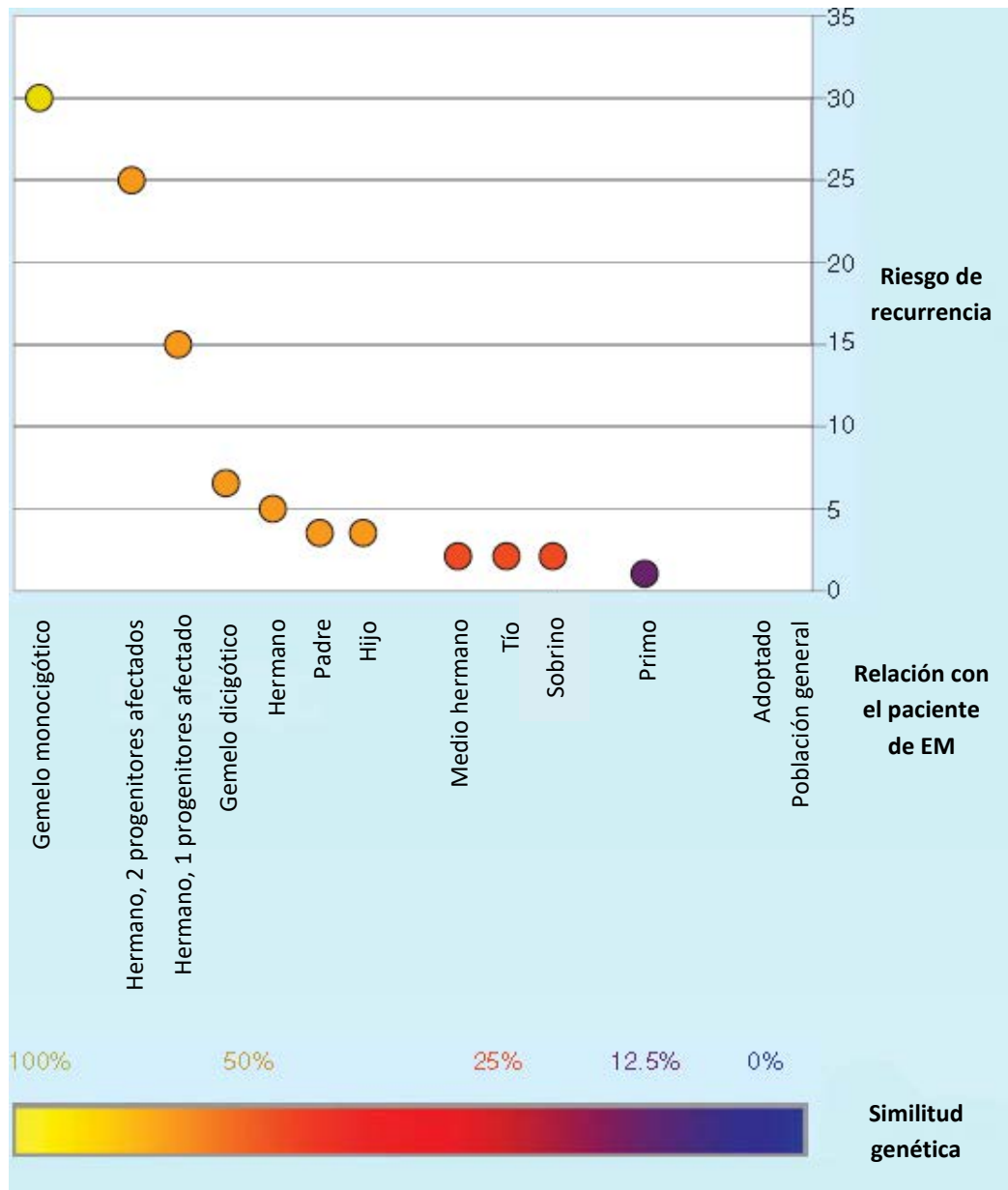


Figura 8. El riesgo de recurrencia en la EM. Se representa la probabilidad de desarrollar EM respecto a la relación familiar que se tenga con el paciente de EM. La barra inferior indica el porcentaje del componente genético compartido. Adaptada de Oksenberg y col. (213).

1.9.3. Genes asociados con la EM

El esclarecimiento de la existencia de un componente genético subyacente a la EM impulsó el inicio de numerosas investigaciones con el objetivo de identificar genes asociados a la enfermedad y elucidar la arquitectura genética de la EM. Entre los *loci* de susceptibilidad identificados, se observa una predominancia de genes relevantes para el sistema inmunológico.

1.9.3.1. Genes del complejo HLA

En humanos, la región del MHC se denomina sistema del antígeno leucocitario humano (HLA, por *human leukocyte antigens*). El sistema o complejo HLA se sitúa en el brazo corto del cromosoma 6 (*locus* 6p21.3), abarca ~3,5 megabases con más de 220 genes alojados y es la región con mayor densidad de genes del genoma humano (217), así como la que exhibe los niveles más altos de polimorfismo (218). Además, esta región se caracteriza por la existencia de un alto desequilibrio de ligamiento (LD, por *linkage disequilibrium*), cuya distribución no es homogénea a lo largo de toda la región MHC, sino en forma de grandes bloques haplotípicos.

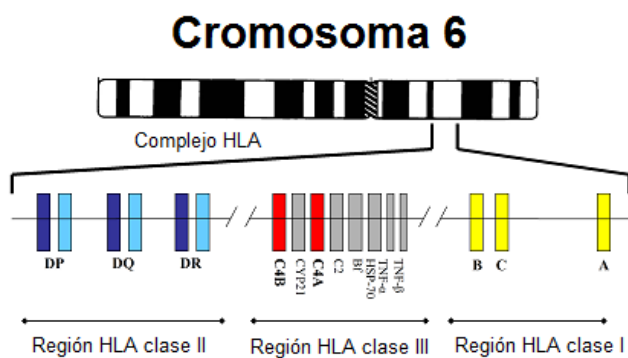


Figura 9. Representación esquemática del MHC en humanos: el denominado sistema HLA.

Como muestra la Figura 9, el sistema HLA se divide en las regiones HLA-II, HLA-III y HLA-I, de la más próxima al centrómero a la más próxima al telómero. A su vez, cada una de las regiones del complejo HLA se subdivide en distintos genes o *loci*. Los genes de las regiones HLA-II y HLA-I codifican glucoproteínas de la membrana plasmática implicadas en los mecanismos de presentación de antígenos a los linfocitos T, mientras que los genes de la región HLA-III codifican citocinas y proteínas del sistema del complemento, TNF- α y TNF- β , entre otros. El *WHO Nomenclature Committee for Factors of the HLA System* ha establecido una nomenclatura oficial para nombrar a las diferentes variantes alélicas que se identifican dentro del sistema HLA: el prefijo HLA siempre aparece y va seguido de un guion, el gen correspondiente y un asterisco, campo a partir del cual se especifica mediante distintos dígitos el grupo alélico o la proteína codificada. Para más información, consultar la página web <http://hla.alleles.org/nomenclature/index.html>.

Las primeras asociaciones genéticas con el riesgo de desarrollar EM se identificaron en la región del HLA y, durante varias décadas, fue el único *locus* asociado de forma consistente a la EM. En el año 1972, la EM se asoció con los alelos del HLA clase I *HLA-A*03* (219) y *HLA-B*07* (220), pero más adelante se descubrió que eran eventos secundarios a la asociación del alelo del HLA clase II *HLA-DR2* (221). Años después, se realizó un mapeo fino y se observó que el haplotipo que realmente confiere susceptibilidad a la EM es el HLA-DRB5*0101-HLA-DRB1*1501-HLA-DQA1*0102-HLA-DQB1*0602 (denominado DR2) en poblaciones del norte de Europa (222,223). El alto LD entre estos *loci* dificultó la identificación de un único *locus* causal, aunque más adelante se observó que estos 4 polimorfismos no siempre se

Introducción

encuentran en el mismo haplotipo, es decir, no siempre se heredan juntos. A día de hoy, se considera que el alelo *HLA-DRB1*1501* es el principal factor genético asociado a la EM (224–226), ya que ha sido vinculado a la enfermedad de forma consistente en prácticamente todas las poblaciones testadas. Además, se ha demostrado que la vitamina D es capaz de regular la expresión del alelo *HLA-DRB1*1501*, lo que supone una conexión entre uno de los principales factores ambientales asociados a la EM y su factor genético central (227). Como dato también interesante, la presencia de BOC en el LCR de pacientes con EM se ha correlacionado positivamente con la presencia del alelo *HLA-DRB1*1501* en diferentes poblaciones de Japón (228), Suecia (229), Australia (230) y España (231), hallazgo que conecta uno de los principales factores diagnósticos de la enfermedad de nuevo con su componente genético central. En el caso de Japón y Suecia, la presencia de BOC en el LCR de pacientes con EM se ha correlacionado negativamente con el alelo *HLA-DRB1*0405*.

Un hecho que aumenta aún más la complejidad de la asociación del sistema HLA con la EM es la existencia de numerosas interacciones epistáticas dentro de la región HLA de clase II, de modo que determinadas variantes genéticas anulan el efecto conferido por ciertos haplotipos del gen *HLA-DRB1*, mientras que otros genes pueden actuar como modificadores de la enfermedad (213).

1.9.3.2. Otros genes asociados

Dado que la EM es una enfermedad en la que se postula que un conjunto de genes contribuyen a la susceptibilidad a la misma, durante las últimas décadas se han llevado a cabo diversos tipos de estudios con la finalidad de identificar genes asociados con la EM fuera de la región HLA.

i) Estudios en familias

Tal y como se mencionó previamente, en la EM se puede observar cierta acumulación familiar de la enfermedad. Pero cuando afecta a una familia, lo hace de forma irregular. Por ello, no se puede concluir que se trata de una enfermedad con herencia mendeliana. Con la finalidad de identificar algún gen que contribuya al riesgo de desarrollar EM, se comenzaron a realizar estudios en familias con dos o más miembros afectados de EM pertenecientes al menos a dos generaciones distintas para identificar un componente genético común y desenmarcar marcadores genéticos que han sido heredados (Figura 10). Sin embargo, la aplicación de estos estudios para la identificación de genes asociados con la EM va acompañada de importantes limitaciones: (i) en primer lugar, son más útiles y concluyentes para enfermedades con herencia

monogénica que para enfermedades poligénicas, como la EM; (ii) además, la mayoría de los pacientes con EM no tienen ningún familiar con la enfermedad, por lo que no es fácil conseguir un tamaño muestral adecuado. Por todo ello, los estudios de ligamiento en familias sólo han reportado asociación de la región HLA con la EM una vez más, pero no han permitido identificar asociación fuera de ella (232).

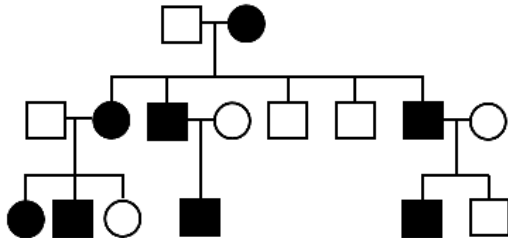


Figura 10. Familia válida para estudios de ligamiento. Diagrama representativo de una familia con varios miembros (color negro) afectados de una enfermedad.

ii) Estudios de asociación en genes candidatos

La primera fase para realizar estos estudios consiste en la elección de genes candidatos a ser asociados con la EM en base a diferentes hipótesis, como su previa asociación con otras enfermedades inmunomediadas, autoinmunes o desmielinizantes o su función biológica relativa a procesos potencialmente relacionados con la patogénesis de la EM (233). Una vez detectada una asociación, el polimorfismo podría ser el causal o estar en alto LD con el realmente causal. Gracias a estos estudios, se han identificado varios *loci* implicados en la EM, incluidos algunos genes en la región del HLA de clase I (219) y de clase II (234), el gen *IL7RA* (235), el gen *CIITA* (236) o el gen *SOCS1* (237). A pesar de su precisión, los estudios de genes candidatos han sido criticados debido a tres desventajas principales: la especificidad de sus resultados sobre una población concreta (238), la falta de replicación de los resultados obtenidos y su capacidad limitada para incluir todos los posibles genes y polimorfismos causales (239). Sin embargo, hoy en día se utiliza como abordaje el realizar meta-análisis de los resultados de varios estudios de asociación en genes candidatos para resolver sus limitaciones y para aumentar el poder estadístico. A modo de resumen, los meta-análisis de estudios de asociación de genes candidatos realizados hasta ahora han confirmado la asociación con la EM de polimorfismos situados en 15 genes, pero no lo han conseguido en el caso de 14 genes (238).

iii) Estudios de asociación del genoma completo (GWAS)

Los estudios de asociación del genoma completo (GWAS, por *genome-wide association studies*) han permitido solucionar parcialmente las limitaciones de los estudios en genes candidatos. Los GWAS se basan en examinar o genotipar de forma simultánea un gran número de polimorfismos genéticos, seleccionados sin partir de ninguna hipótesis, en grandes cohortes para determinar si existe asociación de alguno de ellos con el fenotipo de interés. En la EM, los GWAS se han utilizado para elucidar los factores genéticos asociados con la susceptibilidad a desarrollar la enfermedad, con aspectos clínicos (niveles de IgG, BOCs, lesiones cerebrales, edad de inicio, etc.), con el fenotipo del paciente o con la respuesta al tratamiento (233). Dada la existencia de grandes bloques de LD a lo largo del genoma, en los GWAS es posible utilizar un número de SNPs relativamente pequeño (*tag SNPs*), cada uno representativo de un grupo de SNPs más numeroso o haplotipo (240). El número de SNPs incluidos en los GWAS realizados hasta la fecha varía ampliamente (con un rango de 262 K a aproximadamente 600 K). Además, sólo los SNPs cuyos alelos minoritarios son comunes (frecuencia de alelo menor, o MAF, superior al 1%) se incluyeron en el diseño de dichos GWAS. Cabe destacar que en algunos estudios sólo se estudiaron los SNPs con una MAF superior al 5%. Todos los GWAS realizados constan de dos fases: una inicial de descubrimiento y una de replicación, en la que se validan los resultados en una cohorte independiente.

El primer GWAS en EM se publicó en el año 2007 e incluía 931 tríos familiares (formados por un paciente con EM y sus dos progenitores) y grupos caso-control (241). Los autores mostraron su decepción al observar en el *Manhattan Plot* que la única línea de puntos que superaba el umbral de asociación $p < 1 \times 10^{-7}$ aparecía en el cromosoma 6 (Figura 11), lo que supondría volver a reportar una asociación de la región HLA con la EM.

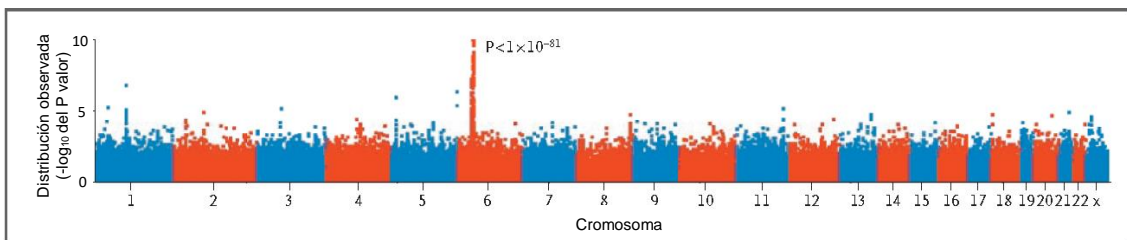


Figura 11. Esquema del primer GWAS desarrollado en EM. En él, se incluyeron 931 tríos familiares. Los p valores (mostrados como $-\log_{10}$) de los resultados se representan a lo largo de todas las cromosomas del genoma. El *locus* de riesgo clásico en el cromosoma 6p21, el HLA-DR, destaca con una fuerte significación estadística ($p < 1 \times 10^{-81}$) (241).

Sin embargo, para futuras investigaciones, se tuvieron en cuenta los “*hits*” identificados en otros cromosomas que, aunque no superaban el umbral de significación, podrían ser indicios de asociación. De hecho, la selección de 110 SNPs de este primer GWAS para su validación condujo a la identificación de genes de susceptibilidad a la EM fuera de la región HLA, como alelos del gen *IL2RA* o del gen *IL7RA*.

A partir del año 2007, se han realizado numerosos GWAS. En todos ellos, independientemente del resto de *loci* identificados, el *locus* HLA se reportó como asociado a la EM de forma sólida. El segundo GWAS en EM se realizó en el año 2008 e incluyó 242 pacientes con EM y 242 controles sanos (242). En este caso, se analizaron 500.000 SNPs y se seleccionaron 320 para ser genotipados en una cohorte de replicación independiente española caucásica. De ellos, los 8 que se validaron con el mayor poder estadístico se genotiparon en una segunda cohorte de replicación estadounidense caucásica para su confirmación y se observó que las asociaciones más fuertes correspondían a los SNPs rs3129934 (HLA, $p=1,4\times 10^{-5}$) y rs1327328 (13q31.3, $p=0,0018$), siendo la última región un nuevo *locus* de riesgo para la EM. Sin embargo, según estudios de simulación, los GWAS con menos de 2.000 casos y 2.000 controles en la fase de descubrimiento tendrán poco poder para detectar asociaciones típicas para variantes comunes (243); de ahí que estas asociaciones no alcancen el nivel de significación de GWAS ($p=5\times 10^{-8}$) (244). También en 2008, Aulchenko y col. publicaron el tercer GWAS en EM (245), en el que identificaron una nueva variante genética, rs10492972 (*KIF1B*), asociada con el riesgo de desarrollar EM. Inicialmente, detectaron este SNP en una pequeña cohorte holandesa de 45 pacientes con EM y 195 controles ($p=0,0004$) y posteriormente lo replicaron en tres cohortes independientes (procedentes de Holanda, Suiza y Canadá); el meta-análisis del total de 2.679 casos y 3.125 controles que constituyen estas tres cohortes resultó significativo ($2,5\times 10^{-10}$).

Fue en 2009 cuando se publicó el cuarto GWAS en EM, que testó 551.642 SNPs en 978 pacientes con EM y 883 controles sanos con alto grado de similitud demográfica, procedentes de Amsterdam, San Francisco y Suiza (246). En él, se identificaron 13 nuevos SNPs fuera de la región HLA asociados a la EM que superaron el umbral de significación $p=10^{-5}$, entre los que destaca una variante genética localizada en el gen *GPC5* (rs9523762). En el siguiente GWAS en EM, publicado en el año 2009, se examinaron 302.098 SNPs en una cohorte formada por 1.618 pacientes con EM y 3.413 controles sanos de ascendencia europea procedentes de Australia y Nueva Zelanda (247). Los SNPs se priorizaron para la fase de replicación y se seleccionaron 100 para intentar validarlos en una cohorte independiente formada por 2.256 casos y 2.310 controles sanos también de Australia y Nueva Zelanda. Como resultado, se asociaron dos nuevos *loci* al riesgo de desarrollar EM, situados en los cromosomas 12 y 20 (concretamente, *upstream* de *CD40*), los cuales ya estaban asociados a otras enfermedades autoinmunes.

Introducción

En el año 2010, se desarrolló un GWAS que incluía 68 casos con EM y 136 controles sanos procedentes de una región de Finlandia con alta prevalencia de la enfermedad (248). Las 27 variantes que resultaron más significativas ($p < 10^{-4}$) se testaron en una cohorte adicional formada por 711 casos de EM y 1029 controles sanos finlandeses. De este análisis, se seleccionaron los 2 SNPs más significativos fuera de la región HLA (rs744166, *STAT3*, $p = 0,0012$ y rs1364194, cromosoma 16, $p = 0,0047$) y se genotiparon en una muestra internacional procedente de 6 poblaciones diferentes y formada por 3.859 pacientes con EM y 9.110 controles sanos. El resultado fue la validación de la variante del gen *STAT3*, rs744166 ($p = 2,75 \times 10^{-10}$), cuyo alelo protector para la EM es de riesgo para la enfermedad de Crohn (249). Además, los autores evaluaron el bloque de LD que contiene este SNP y observaron que engloba el inicio del gen *STAT3* y parte de su promotor. También en 2010, Sanna y col. publicaron un GWAS donde genotiparon 6,6 millones de SNPs en 882 pacientes con EM-RR y 872 controles sanos, todos procedentes de una población de Cerdeña con alta prevalencia de EM (0,16%), e identificaron asociación del gen *CBLB* con la EM, lo que se confirmó en 1.775 casos con EM y 2.005 controles sanos también de Cerdeña. Los autores observaron que el SNP rs9657904 (*joint* $p = 1,60 \times 10^{-10}$) explicaba por completo la asociación de esta región a la EM (250). En el mismo año, 2010, se publicó otro GWAS donde se genotiparon ~300.000 SNPs en 590 pacientes con EM y 825 controles sanos de ascendencia europea y en el cual se hallaron 4 nuevos *loci* asociados a la EM, aunque ninguno de ellos alcanzó el nivel de significación de GWAS (251).

El siguiente GWAS en EM se reportó en el año 2011 y consistió en un meta-análisis de 7 GWAS que incluyó 5.545 casos y 12.153 controles, todos ellos de ascendencia europea, y que analizó 2.529.394 SNPs. El resultado fue la identificación de 3 nuevos *loci* de susceptibilidad a la EM (252). Posteriormente, también en 2011, el mayor GWAS hasta el momento fue desarrollado por una colaboración entre el IMSGC y el WTCCC2 (226). En él, participaron casi 10.000 pacientes con EM y más de 17.000 controles sanos, todos de ascendencia europea y recogidos por 23 grupos de investigación pertenecientes a 15 países diferentes. Además de la fuerte asociación de la región HLA y de replicar casi todas las asociaciones previamente reportadas, se identificaron 29 nuevos *loci* asociados a la EM (véase la Figura 12).

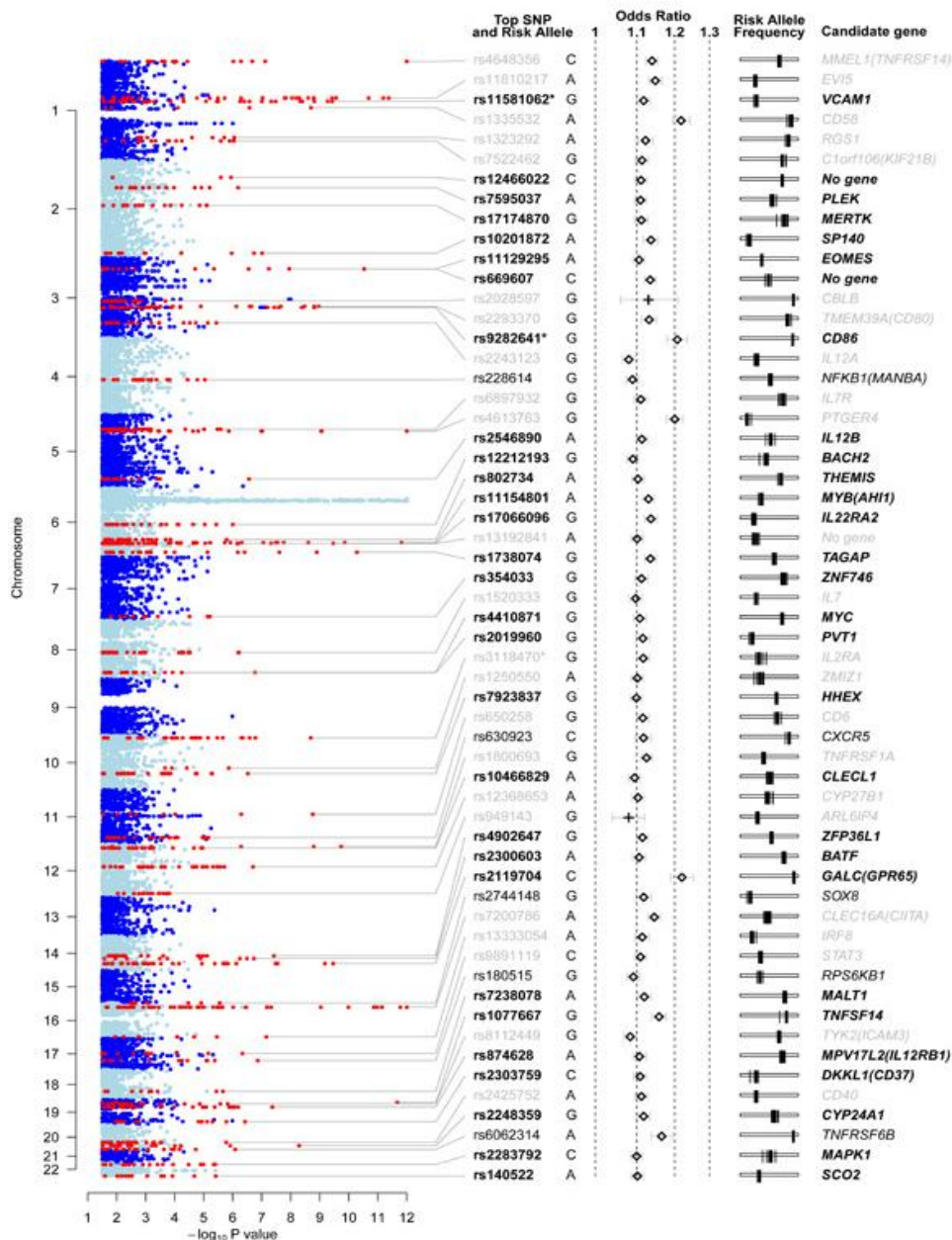


Figura 12. Regiones genómicas fuera del complejo HLA asociadas a la EM en el GWAS del año 2011. En negrita, aparecen los SNPs de nueva identificación; en negro, los SNPs fuertemente asociados; en gris, los previamente descritos. De cada uno, se indica su alelo de riesgo (*Risk Allele*), su Odds Ratio y la frecuencia del alelo de riesgo (*Risk Allele Frequency*). La columna *candidate gene* indica al mejor gen candidato de cada asociación identificada. Adaptada de Sawcer y col. (226).

En 2012, Matesanz y col. publicaron un GWAS realizado en una muestra formada por 296 casos con EM y 801 controles sanos de origen español. Además, hicieron un meta-análisis de los nuevos datos generados con los previamente publicados. El resultado más trascendental fue la confirmación de la asociación del *locus* 5p13.1 a la EM mediante la identificación de un nuevo SNP asociado al riesgo de desarrollar EM en dicha región (253). También en 2012 se publicó un GWAS en el que testaron 277.866 SNPs en 197 pacientes con EMPP y 234 controles

Introducción

sanos de ascendencia europea procedentes de Italia. En este caso, la validación de los SNPs seleccionados se realizó en una cohorte formada por 379 casos con EMPP y 398 controles sanos, también de ascendencia europea pero procedentes de Australia y de países nórdicos, consiguiendo evidenciar asociación de un nuevo *locus* en el cromosoma 7 con la EM, aunque no se alcanzó el nivel de significación de GWAS (254).

En este apartado es importante mencionar que en septiembre de 2013 se realizó un importante avance en la genética de la EM que afectó al diseño de los GWAS y meta-análisis realizados a partir de dicha fecha. Se trata de un estudio liderado por el IMSSGC en el que se empleó un *chip* customizado desarrollado por el *ImmunoChip Consortium*. Inicialmente, se analizaron 161.311 variantes autosómicas en 14.498 pacientes con EM y 24.091 controles sanos. Posteriormente, se replicaron los hallazgos en 14.802 sujetos con EM y 26.703 controles sanos utilizando datos de GWAS previos disponibles. Finalmente, se realizó un análisis conjunto de las fases de descubrimiento y replicación. Como resultado, las variantes genéticas fuera de la región MHC asociadas a la EM ascendieron a 110 (en 103 *loci*) gracias a las 48 nuevas identificadas en este estudio (255).

En 2016, se publicó un GWAS que constaba de una cohorte inicial formada por 4.888 casos con EM y 10.395 controles alemanes de ascendencia europea y de una cohorte de replicación constituida por 2.903 pacientes con EM y 3.323 controles sanos, todos de ascendencia sarda. Globalmente, se identificaron 5 nuevos *loci* asociados a la EM (256). Posteriormente, en el mismo año, se publicó un GWAS que también reveló 5 nuevos *loci* asociados al riesgo de desarrollar EM. En este caso, participaron 2.273 pacientes con EM y 2.148 controles sardos y se imputaron 574.519 SNPs. La cohorte de replicación la constituyeron más de 10.000 pacientes con EM y más de 10.000 controles sanos de ascendencia europea (257).

Por último, cabe destacar que, en 2017, se ha establecido el último mapa de referencia de la arquitectura genética de la EM, el cual incluye 200 variantes autosómicas fuera del MHC, una variante situada en el cromosoma X y 32 asociaciones independientes dentro del complejo MHC. Estos datos derivan del estudio realizado por el IMSSGC en el que se analizan los datos genéticos de 47.351 pacientes con EM y 68.284 controles sanos (258).

1.9.4. La heredabilidad en la EM

La heredabilidad es la proporción de la varianza de un rasgo en una población que es atribuible a la variación genética (es decir, a las diferencias en el genotipo). Matemáticamente, es la proporción de la varianza genética aditiva respecto a la varianza fenotípica. Durante las últimas

décadas, se han identificado numerosas variantes genéticas comunes ($MAF \geq 5\%$ o $MAF \geq 1\%$, según el criterio seleccionado) asociadas al riesgo de desarrollar EM; sin embargo, dado su pequeño tamaño de efecto y su baja odds ratio (OR) asociada, sólo son capaces de explicar el ~39% de la heredabilidad total de la EM, cifra a la que fundamentalmente contribuyen las variantes de la región del MHC, aunque podría extenderse al ~48% si se tuviesen en cuenta también las 192 variantes con efecto sugestivo de asociación (9%) (258). El desconocimiento de la mayor parte de la proporción de la heredabilidad, tanto de la EM como de otras enfermedades complejas, condujo al origen del término “heredabilidad perdida” (259).

Se han llevado a cabo numerosas estrategias científicas para tratar de identificar la procedencia de dicha heredabilidad perdida (260). Una de las hipótesis más competentes atribuye una gran responsabilidad a las variantes genéticas cuya baja frecuencia en la población impide que resulten significativas y, por tanto, tienden a ser desdeñadas. Sin embargo, a pesar de su baja frecuencia, se trata de variantes con alta penetrancia, siendo ésta la proporción de individuos que portan un determinado alelo de un gen y que, además, expresan el fenotipo asociado a dicho alelo. Por ello, por norma general, estas variantes sólo se encuentran en individuos afectados de la patología asociada y no en la población general. En la literatura ya existe la evidencia empírica de que las variantes de baja frecuencia y raras están asociadas a enfermedades complejas (261,262). En el caso de las variantes comunes, su frecuencia en la población general es alta, pero su efecto o su penetrancia son bajos, con una OR asociada rara vez superior a 1,20, de modo que es posible encontrar individuos portadores de la variante que no desarrollan la enfermedad en cuestión.

Pero la hipótesis expuesta no derriba por completo el clásico paradigma de “enfermedad común-variante común” (CDCV, por *common disease-common variant*), que sostiene que la susceptibilidad genética a una enfermedad común se debe a variantes comunes relativamente frecuentes en la población general y de baja penetrancia (263), sino que genera un nuevo modelo denominado “enfermedad común-variante rara” (CDRV, por *common disease-rare variant*) que podría ser compatible con el tradicional y que propone que existen muchas variantes raras, cada una de ellas con alta penetrancia, que predisponen a desarrollar una enfermedad común. Estas variantes tendrían un alto impacto funcional mediante el truncamiento de la función o producción proteica (264). Como ya se ha mencionado, los dos modelos opuestos no son excluyentes, ya que es posible que una combinación de variantes comunes y raras sea la causa de que un determinado individuo sea susceptible a la enfermedad.

Un concepto pertinente en este contexto es la selección natural. Quizás, el argumento más consistente para apoyar al modelo CDRV proviene de la teoría de la evolución. Cuando aparece una variante deletérea, la selección negativa actúa para evitar que la deriva genética fije dicha

variante, es decir, para que no alcance la frecuencia de 1%. De hecho, el objetivo final podría ser conseguir una frecuencia de 0% para que la variante desaparezca. Por lo tanto, la mayoría de los SNPs observados en la población son probablemente neutrales o levemente deletéreos (265). Esto podría explicar que las variantes que superan dicha frecuencia, es decir, las variantes comunes, tengan baja penetrancia a la hora de promover el desarrollo de una enfermedad y que una sola variante común no pueda ser la única causa genética que desencadene una patología. Dado que una enfermedad por sí sola se puede considerar como algo conceptualmente deletéreo respecto al *fitness* (266), las variantes que la promueven podrían haber sido seleccionadas en contra de la evolución natural; podría haber individuos portadores de variantes deletéreas, de alta penetrancia, que la selección natural no ha conseguido eliminar por completo pero cuya frecuencia es tan baja que se consideran raras.

1.9.5. Variantes genéticas raras en la esclerosis múltiple

Durante los últimos años, se ha potenciado el interés hacia las variantes genéticas raras. En la literatura, no existe un consenso sobre qué es una variante rara, es decir, sobre qué umbral de frecuencia establecer para considerar a una variante genética como rara. Lo más común es situar este umbral en $MAF \leq 0,5\%$ ó 1%. Algunos autores hablan de variantes de baja frecuencia ($0,5\% \leq MAF < 5\%$) y de variantes raras ($MAF \leq 0,5\%$).

Como se mencionó en el apartado anterior, los tests de asociación utilizados en los GWAS que emplean *SNP-chips* prediseñados no capturan variantes genéticas menos frecuentes o raras. Debido a su baja frecuencia, las variantes raras generalmente no se incluyen en los *arrays* de genotipado comerciales y se necesitan tamaños de muestra muy elevados para detectar una asociación significativa con un rasgo o enfermedad. Sin embargo, ahora son detectables gracias a las tecnologías de secuenciación de nueva generación (NGS, por *next-generation sequencing*), por lo que actualmente es posible estudiar su contribución a las enfermedades comunes mediante la resecuenciación de muestras de ADN de casos y controles. Una vez detectadas, se pueden llevar a cabo diferentes diseños y estrategias científicas para determinar si tienen un papel funcional en la enfermedad de interés o no. Dado que los métodos de asociación tradicionales empleados para las variantes comunes no pueden utilizarse para las variantes raras, se ha desarrollado un abordaje basado en la aplicación de una serie de tests conocidos en su conjunto como *collapsing methods* (267) para testar si el número de muestras que portan al menos una variante rara de un gen en concreto difiere entre el grupo de casos y el de controles (Figura 13). En otras palabras, proporcionan la información de si existe una acumulación de variantes raras de un gen en uno de los dos grupos respecto al otro, constituyendo así una pista

para encauzar los futuros estudios funcionales. Otra forma de estudiar su papel en las enfermedades comunes es mediante *arrays* de genotipado que incluyan variantes raras (268). De este modo, el investigador puede incluir los SNPs que considere interesantes candidatos.

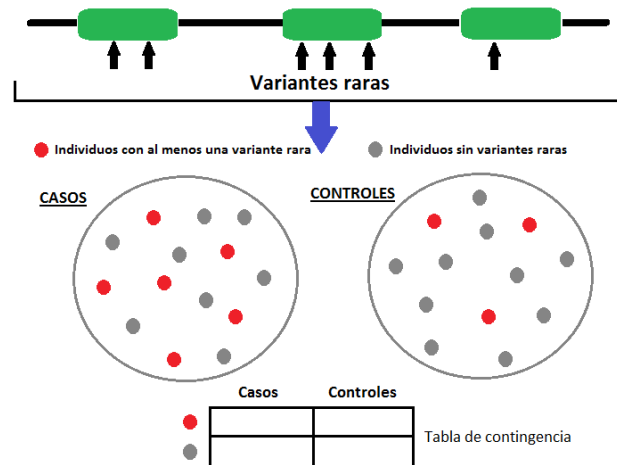


Figura 13. Explicación esquemática de la base lógica del conjunto de *collapsing methods*.

La contribución de las variantes raras en la EM aún no se ha dilucidado. Aunque se han descrito formas familiares de EM, hasta la fecha no se han reportado variantes raras y penetrantes asociadas de forma sólida. Cabe destacar un estudio donde se realizó una secuenciación completa de las regiones codificantes de genes candidatos de susceptibilidad a la EM (269) y se identificó una variante rara en el gen *CYP27B1* que causa la pérdida completa de la función del gen en un paciente con EM; después observaron que, en homocigosis, esta variante causaba raquitismo dependiente de vitamina D tipo 1, mientras que en heterocigosis causaba niveles bajos de calcitriol. El genotipado adicional de más de 12.500 individuos mostró que otras variantes raras de pérdida de función de *CYP27B1* también conferían un riesgo significativo de EM. En otro reciente estudio, se secuenció el exoma completo de dos hermanos afectados de EM (270). Ambos pacientes resultaron homocigotos para una variante rara predicha como patogénica (Gly578Ser) en el gen *NLRP1*. A continuación, se amplió la búsqueda de variantes patogénicas del gen *NLRP1* en 38 pacientes con EM familiar y 44 pacientes con EM esporádica y se identificaron otras dos familias con EM con un total de 5 variantes raras potencialmente patogénicas de *NLRP1*. La más interesante consiste en un desplazamiento del marco de lectura que introduce un codón de stop prematuro. En los pacientes con alguna variante rara de *NLRP1*, así como en controles sanos, se procedió a la inmunoestimulación específica de la ruta de señalización del gen *NLRP1* en células mononucleares de sangre periférica (CMSP), se midió la actividad de los efectores situados *downstream* de *NLRP1* y se secuenció el transcriptoma

(RNAseq). El perfil de expresión de genes después de la estimulación de la ruta de *NLRP1* mostró diferencias significativas en la expresión de genes efectores situados *downstream*, siendo un ejemplo interesante el aumento de la expresión de *IL1B* en pacientes con EM portadores de variantes raras de *NLRP1* respecto a controles sanos. El RNAseq evidenció una sobreactivación de los genes relacionados con el inflamasoma en los pacientes portadores de variantes raras de *NLRP1*.

Todo lo expuesto pone de manifiesto la posible implicación de variantes raras de genes de susceptibilidad a la EM en la patogénesis de la enfermedad y motiva la investigación en este campo.

1.9.6. Las variantes genéticas estructurales

Las variantes genéticas pueden clasificarse en dos categorías principales: SNPs y variantes estructurales o variaciones en el número de copias (CNVs, *por copy number variations*) (271). Las CNVs son mutaciones que pueden ser heredadas u ocurrir *de novo* e incluyen deleciones, inserciones y duplicaciones. Se trata de una forma de alteración cromosómica estructural en la que una región del ADN posee un número de copias distinto al de un genoma de referencia. Aparecen a lo largo de todo el genoma humano de manera muy común y en cientos de localizaciones (272). A veces, una CNV puede ser tan grande que afecta a medio millón de nucleótidos. Dado que muchas CNVs suponen cambios en los niveles de expresión génica, las CNVs pueden explicar una proporción significativa de la variación fenotípica normal (273). Funcionalmente, las variantes estructurales pueden ser benignas, tener influencias sutiles, predisponer o causar una enfermedad en la generación actual, o predisponer a una enfermedad en la siguiente generación (274).

Mediante el uso de *microarrays* de ADN, es posible evaluar pacientes con enfermedades con un componente genético y compararlos con individuos sanos para examinar qué variantes estructurales están asociadas con la patología y cuáles son comunes en la población general. Esta información puede contribuir a la identificación de nuevos genes asociados a una enfermedad (275). De hecho, ya existen estudios publicados donde se ha demostrado asociación entre alguna variante estructural y la diversidad fenotípica o la susceptibilidad a una enfermedad (276,277).

No hay mucha información sobre CNVs asociadas a la EM. La mayoría de los estudios reportados hasta la fecha están enfocados hacia fenotipos de EM muy concretos, como pacientes con EM con insuficiencia venosa cerebroespinal crónica, causada por malformación venosa

estenotante (278), o pacientes con EM pediátrica (279). En 2005, se publicó un estudio donde se investigó la asociación de la delección de los siete primeros exones del gen *ILT6* (*immunoglobulin-like transcript 6*) con la EM (280). La razón de escoger dicha delección fue doble: en primer lugar, se trata de una delección de 6.7 kbp descrita en humanos como una variante estructural frecuente (281); en segundo lugar, el gen *ILT6* codifica un inmunoreceptor soluble expresado predominantemente en células B y monocitos que actúa reconociendo antígenos en el contexto del MHC de clase I. Como resultado, se observó que la deficiencia homocigota de *ILT6* era significativamente más prevalente en pacientes con EM (7,1%) que en controles sanos (3,8%). Recientemente, se ha publicado otro estudio en pacientes con EM japoneses donde identifican variaciones estructurales, también en la categoría de delecciones, en las regiones genómicas que codifican las cadenas *alpha* y *gamma* del TCR. Entre ellas, se observó una alta prevalencia en la cohorte de EM de dos delecciones exónicas: una en *TRG* (16,40%) y otra en *TRA* (17,28%) (282).

Aun siendo pocos, los indicios de la implicación de determinadas variantes estructurales en la EM evidencian la necesidad de hacer énfasis en el diseño de estudios cuyo objetivo sea la identificación de nuevas CNVs asociadas a la EM y la posterior validación de los hallazgos en cohortes independientes.

2. HIPÓTESIS Y OBJETIVOS

2. Hipótesis y objetivos

- **HIPÓTESIS:**

La resecuenciación dirigida de ADN permitirá la identificación de variantes genéticas comunes, raras y estructurales diferencialmente distribuidas entre pacientes con EM y controles sanos. La presencia, bien de una variante común, o bien de un enriquecimiento de variantes raras, se asociará con un efecto funcional. Este efecto incrementará su valor como posible biomarcador aplicable en la práctica clínica del paciente con EM.

- **OBJETIVO PRINCIPAL:**

Identificar nuevas variantes genéticas asociadas con la susceptibilidad a la EM y estudiar sus posibles implicaciones funcionales.

- **OBJETIVOS SECUNDARIOS:**

1. Identificar variantes genéticas comunes, raras y estructurales asociadas con el riesgo de desarrollar EM en 14 genes candidatos mediante resecuenciación dirigida de ADN.
2. Validar una selección de variantes genéticas comunes asociadas con la susceptibilidad de desarrollar EM en la resecuenciación de ADN en una cohorte independiente de pacientes y controles sanos y realizar estudios adicionales relacionados con los polimorfismos validados.
3. Investigar la posible implicación funcional de la presencia de variantes genéticas raras de genes que las acumulan con mayor frecuencia en pacientes con EM.
4. Validar y caracterizar experimentalmente las variantes genéticas estructurales detectadas en la resecuenciación de ADN e identificar nuevos individuos portadores.

3. MATERIALES Y MÉTODOS

3. Materiales y métodos

3.1. Metodología relacionada con el objetivo 1

"Identificar variantes genéticas comunes, raras y estructurales asociadas con el riesgo de desarrollar EM en 14 genes candidatos mediante resecuenciación dirigida de ADN"

3.1.1. Individuos y obtención de muestras de ADN genómico

Participaron un total de 1.095 individuos, 545 pacientes con EM con formas de inicio con brote y 550 controles sanos, todos de origen europeo. Los 545 pacientes con EM se reclutaron en los siguientes centros: Barcelona–Cemcat (n=358), Málaga (n=75), Granada (n=14), Barcelona–Hospital Clinic (n=52), Girona (n=46). Todos los pacientes incluidos en el estudio cumplían los criterios de diagnóstico para la EM (24). Los 550 controles se obtuvieron de la misma región que los casos de EM [Barcelona-Cemcat (n=411), Málaga (n=79), Granada (n=14), Girona (n=46)]. Las características demográficas de la cohorte definitiva (n=1.070, teniendo en cuenta la eliminación de las muestras que no superaron el control de calidad) se resumen en la Tabla 1.

Características	Pacientes con EM	Controles sanos
Individuos, n	524	546
Mujeres/Hombres (%mujeres)	310/214 (59,2%)	316/230 (57,9%)
Edad, años	39,5 ± 10,5	40,1 ± 12,7
Forma clínica, %RR/SP	69,3%/30,7%	-
Edad de inicio, años	30,1 ± 9,7	-

Tabla 1. Características clínicas y demográficas de los pacientes con EM y de los controles sanos incluidos en la cohorte de resecuenciación tras el control de calidad. La edad se expresan como media ± desviación estándar (DE). RR: *esclerosis múltiple remitente recurrente*; SP: *esclerosis múltiple secundariamente progresiva*.

El ADN genómico de los 1.095 individuos se extrajo a partir de sangre periférica utilizando protocolos estándar en cada centro. En el caso de las muestras del Cemcat, la sangre completa se centrifugó, se recogieron los leucocitos y se lisaron los restos de eritrocitos mediante el tampón de lisis RBC (por *red blood cells*) (Sigma-Aldrich, St. Louis, MA, EEUU). A continuación, se lisaron los leucocitos mediante la incubación toda la noche a 37°C con tampón *Nuclei lysis* (Sigma-Aldrich) y con proteinasa K (Roche, Suiza) para degradar las proteínas. Al día siguiente,

Materiales y métodos - Objetivo 1

se añadió acetato de amonio (Sigma-Aldrich) y se centrifugó para eliminar los restos celulares. Se recogió la fase superior del tubo, donde se encuentra el ADN, y se añadió el mismo volumen de etanol absoluto (Sigma-Aldrich), con la consiguiente precipitación del ADN. Se realizó un lavado con etanol al 70% y se resuspendió el ADN en tampón Tris-EDTA (Sigma-Aldrich). La cuantificación del ADN obtenido se realizó con el espectrofotómetro NanoDrop® ND-1000 (Thermo Fisher Scientific, Wilmington, EEUU). El protocolo más detallado se encuentra en el Anexo. Se utilizaron 5 µg de ADN genómico por individuo para construir la biblioteca de ADN descrita más adelante.

3.1.2. Selección de los 14 genes candidatos

Se seleccionaron 14 genes candidatos asociados con la EM en los estudios GWAS descritos en el apartado de Introducción, en base a una expresión diferencial entre pacientes con EM y controles sanos en estudios de *microarrays* llevados a cabo por el grupo en células de sangre periférica, y/o que presentaran una función importante en la patogenia de la EM según la literatura. Los 14 genes seleccionados para la resecuenciación fueron los siguientes: *FCRL1*, *RGS1*, *TIMMDC1*, *HHEX*, *CXCR5*, *LTBR*, *TSM*, *GALC*, *TRAF3*, *STAT3*, *TNFSF14*, *IFI30*, *CD40* y *CYP24A1*.

3.1.3. Proceso de resecuenciación de ADN

3.1.3.1. Regiones genómicas resecuenciadas

Para cada uno de los 14 genes seleccionados, se resecuenciaron las siguientes regiones:

- Los **exones**, cuyas coordenadas se obtuvieron utilizando el software BioMart (Ensembl genes 67, GRCh37.p7). Además, todos ellos se extendieron 100 pares de bases hacia arriba y hacia abajo para incluir posibles sitios de *splicing*. Esta extensión se amplió en los casos donde se observó que se solapaba con elementos reguladores conocidos (datos descargados de Ensembl).
- Una región de 2000 pares de bases hacia arriba del sitio de inicio de la transcripción (TSS, por *transcription start site*), de modo que se capturó el **promotor** (extremo **5'-terminal**) de cada gen.
- El **extremo 3'-terminal** de cada gen: se consideraron los elementos reguladores ya descritos, concatenándolos si la distancia entre ellos era inferior a 120 pares de bases.

En total, se obtuvieron **141 fragmentos independientes** para un total de **140144 bases**. La versión del genoma empleado como referencia en esta tesis ha sido GRCh37/hg19.

3.1.3.2. Construcción de la biblioteca de ADN genómico o genoteca

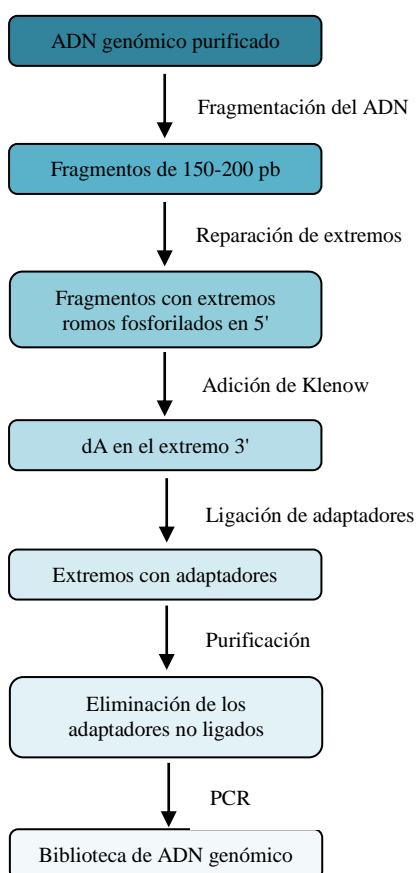


Figura 14. Pasos seguidos para la construcción de la biblioteca de ADN genómico o genoteca.

La Figura 14 esquematiza los pasos seguidos para la construcción de la genoteca. El ADN genómico de cada individuo se fragmentó de forma aleatoria mediante un instrumento de ultrasonicación de Covaris™, dando lugar a fragmentos entre 150 y 200 pares de bases. A continuación, se repararon los extremos salientes resultantes de la fragmentación y se convirtieron en extremos romos; para ello, se aprovechó la actividad exonucleasa 3'→5' del fragmento de Klenow para eliminar los salientes 3' y la actividad polimerasa de la enzima ADN polimerasa T4 para rellenar los salientes 5'. Posteriormente, se añadió una desoxiadenosina (dA, base A) al extremo 3' de los fragmentos romos de ADN fosforilados utilizando la actividad actividad exonucleasa 3'→5' del fragmento de Klenow. Así, se preparan los fragmentos de ADN para la ligación de los adaptadores, que tienen una timina (base T) en su extremo 3'.

Se ligaron las secuencias adaptadoras a ambos extremos de los fragmentos resultantes para poder identificar las muestras en la resecuenciación y se purificaron. Por último, se enriquecieron los fragmentos de ADN con adaptadores en sus extremos mediante amplificación por reacción en cadena de la polimerasa (PCR, por *polymerase chain reaction*).

3.1.3.3. Enriquecimiento de las secuencias diana y resecuenciación

La captura de las secuencias diana de ADN y la resecuenciación se realizó en BGI Hong Kong (ahora, BGI Tech Solutions). El enriquecimiento selectivo de los fragmentos genómicos ya unidos a moléculas adaptadoras en ambos extremos se realizó mediante los *arrays* de captura de NimbleGen, lo que permite secuenciar exclusivamente las regiones de interés y no el genoma completo. Para ello, se realizó una PCR mediada por ligación (LM-PCR, por *Ligation-mediated PCR*) donde se emplearon dos *primers* que hibridaban en los extremos de los adaptadores. Posteriormente, se purificó el ADN y se hibridó con la matriz personalizada de NimbleGen para el enriquecimiento. Se lavaron los fragmentos no hibridados y los productos de la LM-PCR no capturados y capturados se sometieron a PCR cuantitativa para estimar la magnitud del enriquecimiento. Cada biblioteca capturada se cargó en la plataforma Illumina HiSeq2000, y se realizó la resecuenciación de forma independiente para cada biblioteca para garantizar que cada muestra cumpliera con la cobertura promedio deseada.

3.1.4. Mapeo fino de los datos crudos de la resecuenciación

Los datos crudos de la resecuenciación se descargaron del servidor de BGI y se procesaron mediante la *pipeline* clásica de **GATK** (*Broadinstitute*; información más detallada sobre este protocolo en <https://www.broadinstitute.org/gatk/guide/best-practices>). Las lecturas recalibradas se analizaron utilizando **FASTQC**, una herramienta para evaluar el rendimiento del proceso de resecuenciación. De esta forma, los datos ya están listos para el análisis (Figura 15).

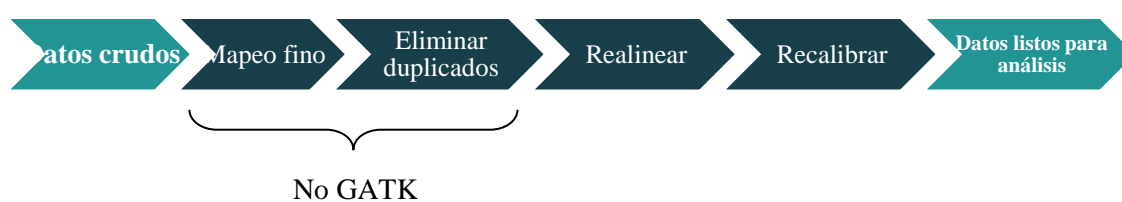


Figura 15. Representación esquemática de los pasos a seguir para obtener los resultados de la resecuenciación de ADN de forma apropiada para ser analizados a partir de los datos crudos. En primer lugar, el mapeo fino de las lecturas contra el genoma humano de referencia (GRCh37/hg19) se realizó utilizando la herramienta bioinformática *Burrows-Wheeler Aligner* (BWA). A continuación, se eliminaron los duplicados ópticos, se realinearon las lecturas alrededor de las indels conocidas, y se realizaron las recalibraciones pertinentes basadas en el control de calidad.

3.1.5. Control de calidad de las muestras resecuenciadas

3.1.5.1. Análisis de la cobertura

El número medio de lecturas que recibe cada una de las bases se denomina cobertura. El análisis de la cobertura de la resecuenciación se realizó mediante la herramienta bioinformática *DepthOfCoverage*, de GATK, con la finalidad de identificar tanto fragmentos como muestras completas con baja cobertura para prescindir de ellas en el estudio y eliminarlas antes de aplicar el llamado de variantes genéticas.

3.1.5.2. Pruebas de control de calidad de las muestras

Se aplicaron tres pruebas diferentes de control de calidad:

- **Número de *singletons* por muestra**, para identificar muestras portadoras de un elevado número de *singletons*, es decir, de fragmentos que no alinean al genoma humano de referencia, lo que hace probable que dichas muestras pertenecieran a una población distinta y deban ser eliminadas del estudio.
- **Análisis de componentes principales (ACP)**, para investigar la subestructura de la población dentro de los casos con EM y de los controles. El ACP se utilizó para identificar muestras con genotipos que difieren demasiado de los del resto de las muestras, de modo que dichas muestras probablemente pertenecen a una población diferente y es preferible descartarlas. El ACP se realizó mediante el *software* SmartPCA y se obtuvieron los 10 primeros componentes principales (-k10).
- **Número de diferencias genóticas**, para detectar pares de muestras demasiado similares entre ellas, lo que podría ser indicio de algún tipo de relación entre ellas o de tratarse de muestras repetidas, de modo que es conveniente eliminarlas del análisis.

3.1.6. Variant calling y anotación funcional

La identificación de variantes genéticas se realizó utilizando la herramienta bioinformática de GATK *Haplotype Caller*. La anotación funcional de las variantes detectadas en la resecuenciación se realizó utilizando la herramienta ANNOVAR. Las variantes se clasificaron en las siguientes categorías (parcialmente superpuestas):

- **Codificantes.**

- **Reguladoras.**
- **Alteradoras del código genético:** SNPs no sinónimos, mutaciones sin sentido e *indels* exónicos (con desplazamiento del marco de lectura o sin él).
- **Posiblemente patogénicas:** SNPs no sinónimos, mutaciones sin sentido e *indels* exónicos predichos como causantes de consecuencias patológicas por al menos un algoritmo después de introducirlas en las bases de datos SIFT, Polyphen, MutationTaster y MutationAssessor.

Las variantes codificantes y reguladoras se obtuvieron considerando si la variación se produce dentro de los fragmentos exónicos o reguladores descargados de Ensembl y utilizados para el diseño.

3.1.7. Identificación de variantes genéticas comunes asociadas con la EM

3.1.7.1. Obtención de bloques haplotípicos

No existe una definición estricta de “bloque haplotípico”, pero podría considerarse bien como una región genómica con diversidad de haplotipos limitada, o bien como un conjunto de sitios consecutivos entre los cuales hay poca o ninguna evidencia de recombinación histórica (283) y, por tanto, con alto LD entre ellos. Los bloques haplotípicos se calcularon descargando los datos de variación genética de la población de ascendencia europea (CEU) del Proyecto 1000 Genomas que comprendían los fragmentos genómicos incluidos en la resecuenciación. Para cada SNP, se obtuvo una lista de SNPs en LD ($r^2 > 0,8$). Los bloques parcialmente superpuestos se consideraron bloques independientes, mientras que los bloques que se encontraban dentro de otros bloques se eliminaron. La lista final de bloques se obtuvo considerando todos los bloques calculados que contenían al menos una variante común detectada en nuestro estudio. Además, las variantes comunes detectadas en nuestro estudio que no cayeron en ningún bloque calculado se consideraron bloques independientes.

3.1.7.2. Pruebas de asociación de variantes comunes a la EM

La asociación de variantes genéticas comunes ($MAF \geq 1\%$) con la EM se testó utilizando PLINK (284) y el comando `-assoc`. El test básico de asociación se aplica sobre un determinado rasgo de la enfermedad o sobre la enfermedad en sí misma y se basa en la comparación de las frecuencias alélicas entre casos y controles (en nuestro caso, casos con EM y controles sanos). La asociación de variantes comunes con la EM se analizó de tres formas diferentes: i) Mediante el

test exacto de Fisher; ii) Mediante la aplicación de una regresión logística considerando un modelo aditivo y iii) Mediante la aplicación de un test de asociación asintótico. Cada uno de estos test dio lugar a unos p-valores diferentes.

3.1.7.3. Análisis del desequilibrio de ligamiento entre variantes comunes

El grado de LD entre los SNPs comunes identificados en la resecuenciación y entre éstos y los SNPs ya descritos en la misma región genómica se calculó mediante la herramienta informática online SNAP (por *SNP Annotation and Proxy Search*), disponible en <http://archive.broadinstitute.org/mpg/snap/ldsearchpw.php>.

3.1.8. Identificación de variantes genéticas raras asociadas con la EM

Para evaluar la contribución de las variantes genéticas raras (VRs) a la patogénesis de la EM, se utilizaron diferentes tests estadísticos que en su conjunto se denominan *collapsing methods*. Éstos analizan si el número de muestras que portan al menos una VR difiere entre casos y controles. Los tests de asociación de VRs se aplicaron utilizando la herramienta *Variant Association Tools*, incluida en el *software Variant Tools* versión 2.0 (<http://varianttools.sourceforge.net/>). Concretamente, se aplicaron los siguientes tests: i) Data-adaptive Sum test (aSUM); ii) C-alpha; iii) Combined and Multivariate Collapsing (CMC) *one-tailed*; iv) CMC *two-tailed*; v) Kernel Based Adaptive Clustering (KBAC) *one-tailed*; vi) KBAC *two-tailed*; vii) Replication Based Test (RBT) *one-tailed*; viii) RBT *two-tailed*; ix) Variable Thresholds method (VT) *one-tailed*; x) VT *two-tailed* y xi) Weighted Sum Statistic (WSS).

En el presente estudio, consideramos como VRs aquellas con frecuencia del alelo menor (MAF, por *minor allele frequency*) $\leq 0,5\%$ dentro de nuestra población de 1070 individuos. Para su estudio, los datos se clasificaron en distintas categorías siguiendo uno de los dos criterios descritos en los subapartados desarrollados a continuación. Una vez establecidas las categorías, se aplicaron los *collapsing methods* a cada gen individualmente. La dirección de la asociación se infirió calculando el número promedio de VRs en casos y controles.

3.1.8.1. Criterio 1: clasificación de las variantes raras por tipo de variante

Siguiendo el primer criterio, las VRs se clasificaron en las siguientes categorías: i) “todas las variantes”, que representa a todas las VRs identificadas en un gen determinado; ii) “variantes no sinónimas”, categoría que incluye sólo a los SNPs localizados dentro de una secuencia codificante que modifican la cadena de aminoácidos resultante; y iii) “variantes posiblemente patogénicas”, que incluye las variantes no sinónimas detectadas predichas como funcionalmente relevantes en al menos una datos de anotación funcional.

3.1.8.2. Criterio 2: clasificación de las variantes raras por tipo de secuencia

A continuación, investigamos si las VRs se distribuían de forma diferencial entre casos con EM y controles sanos al considerar, por un lado, las secuencias reguladoras resecuenciadas y, por otro, las secuencias codificantes resecuenciadas. Además, se analizó dicha distribución diferencial en las regiones UTR y en el promotor. De nuevo, se aplicaron los *collapsing methods* teniendo en cuenta bien todos los genes por separado, o bien todos en conjunto.

3.1.9. Identificación de variantes genéticas estructurales asociadas con la EM

Las variantes estructurales o variaciones en el número de copias (CNVs) se detectaron utilizando el *software* XHMM (*eXome-Hidden Markov Model*), que utiliza el ACP y un modelo estadístico oculto de Márkov (HMM, por *Hidden Markov Model*) para detectar las CNVs presentes en los datos normalizados de la profundidad de la cobertura (en inglés, *depth of coverage* o, simplemente, *coverage*) de un proceso de secuenciación dirigida de última generación. El análisis de la cobertura permite detectar diferencias en el número de lecturas que alinean con el genoma de referencia utilizado; por tanto, la cobertura observada para un determinado fragmento de un individuo concreto indica el número de copias que porta dicho fragmento. Con la finalidad de considerar sólo las CNVs de alta calidad y fiables, las variantes estructurales detectadas con una puntuación de calidad Phred escalada menor a 60 se eliminaron.

3.2. Metodología relacionada con el objetivo 2

"Validar una selección de variantes genéticas comunes asociadas con la susceptibilidad de desarrollar EM en la resecuenciación de ADN en una cohorte independiente de pacientes y controles sanos y realizar estudios adicionales relacionados con los polimorfismos validados"

3.2.1. Genotipado de variantes comunes

3.2.1.1. Individuos y obtención de muestras de ADN genómico

Se incluyeron un total de 3.450 pacientes con EM y 1.688 controles sanos de origen europeo. Ninguno de los individuos estaba incluido en la cohorte de resecuenciación. Los 3.450 pacientes con EM se reclutaron en los siguientes ocho centros de EM nacionales: Madrid (n=1.260), Sevilla (n=790), Bilbao (n=426), Málaga (n=384), Barcelona–Cemcat (n=326), Barcelona– San Sebastián (n=250) y Hospital Clinic (n=104). Todos los pacientes cumplieron criterios de EM (24). Los 1.688 controles sanos provinieron de las siguientes regiones: Granada (n=1.007), Bilbao (n=397) y Barcelona–Cemcat (n=284). La Tabla 2 resume las características demográficas y clínicas basales de los individuos que componen la cohorte de genotipado.

Características	Pacientes con EM	Controles sanos
n	3.450	1.688
%Mujeres/%Hombres	69/31	62/38
Edad, años	39,9 ± 12,8	42,6 ± 13,0
Forma clínica (%RR/%SP/%PP)	83,0/8,6/8,4	-

Tabla 2. Características clínicas y demográficas de los pacientes con EM y de los controles sanos incluidos en la cohorte de validación. Los datos se calcularon utilizando las muestras cuya información estaba disponible. La edad se expresa como media ± desviación estándar. *RR: esclerosis múltiple remitente recurrente. SP: esclerosis múltiple secundariamente progresiva. PP: esclerosis múltiple primariamente progresiva.*

El ADN genómico de los 5.138 individuos se extrajo en cada centro utilizando protocolos estándar. En el caso de los individuos reclutados en el Cemcat, el ADN se obtuvo y se cuantificó tal y como se explicó en el apartado 3.1.1. para la resecuenciación. Las muestras se prepararon a una concentración 40 ng/μL y se enviaron y centralizaron en el Centro Nacional de Genotipado (CeGen-PRB2), en Santiago de Compostela.

3.2.1.2. Elección de los SNPs comunes

Para seleccionar las variantes genéticas comunes que se incluyeron en el genotipado, se hizo una priorización de las mismas en base a las siguientes pautas:

- A partir de los resultados de la resecuenciación, se seleccionaron aquellos SNPs comunes (frecuencia del alelo menos común (MAF)>1%) cuya asociación con la EM resultó estadísticamente significativa ($p < 0,05$).
- Posteriormente, se realizó un estudio bibliográfico y se utilizó la herramienta informática Ensembl (<http://grch37.ensembl.org/index.html>) para analizar si alguno de los SNPs ya estaba asociado con la EM en otros estudios.
- Una vez delimitada esta lista de genes, se utilizaron las herramientas bioinformáticas LDLink (<https://analysistools.nci.nih.gov/LDlink>) y SNAP (<http://archive.broadinstitute.org/mpg/snap/ldsearchpw.php>) para analizar el grado de LD existente entre las variantes comunes seleccionadas hasta el momento. Las variantes que se observaron en alto LD entre sí se consideraron como una misma señal, de forma que sólo se seleccionó una de ellas como representativa de todas (*tag* SNP). La selección del *tag* SNP se realizó en base a su *p-valor* (se eligió aquella con *p-valor* más bajo) y a su odds ratio (OR) (se prefirió aquella con la OR más alejada de 1,0).
- Una vez analizado el LD entre nuestros SNPs, se analizó el LD entre cada *tag* SNP seleccionado y los SNPs en la misma región publicados en el último meta-análisis como asociados a la EM (258), con la finalidad de evitar reportar asociación de una variante que, si se encontrase en alto LD con otra ya reportada, constituiría la misma señal de asociación.
- Para tener seguridad absoluta de que se están reportando nuevas asociaciones con la EM, también se seleccionaron para genotipar una serie de SNPs identificados en el meta-análisis previamente mencionado (258) localizados en la misma región que nuestros SNPs seleccionados, con el objetivo de llevar a cabo una regresión logística condicional y confirmar que son señales independientes en nuestra cohorte.
- Finalmente, se tuvo en cuenta la compatibilidad de los SNPs seleccionados entre sí en la reacción de PCR *multiplex* incluida en el proceso de genotipado.

Después de aplicar estos filtros, se incluyeron 9 variantes comunes halladas en la resecuenciación para su validación mediante genotipado y 9 procedentes del meta-análisis del IMSGC. Los detalles de dichas variantes se encuentran en el apartado 4.2.1.1. de Resultados.

3.2.1.3. Genotipado de SNPs

El genotipado de las 18 variantes comunes se realizó en el CeGen-PRB2 mediante la tecnología iPLEX® Gold y la plataforma MassArray, de Agena Bioscience (San Diego, CA, USA). Este método utiliza la espectrometría de masas MALDI-TOF (MALDI, por *Matrix-Assisted Laser Desorption/Ionization*, y TOF, por *Time-Of-Flight*) para detectar los productos de la reacción de discriminación alélica. El diseño del genotipado se realizó utilizando el software MassARRAY Assay Designer 4.1, de Agena Bioscience. Las 18 variantes se analizaron en una sola reacción.

El genotipado constó de tres fases: en la primera, se amplificaron a la vez todos los fragmentos de ADN (~100bp) que contienen los SNPs de interés mediante una PCR *multiplex*; en la segunda, los productos de la PCR se trataron con fosfatasa alcalina para neutralizar los dNTPs no incorporados; la tercera consiste en la discriminación alélica mediante *primer extension* o extensión de una única base (*Single Base Extension*, SBE) (285) utilizando una ADN polimerasa para extender un primer que alinea inmediatamente adyacente a la posición del SNP de interés. Todas las reacciones terminan tras una SBE que genera una diferencia de masa en el producto de dicha extensión, la cual es determinada por la base añadida. Las condiciones de las PCR de las fases 1 y 3 se detallan en el Anexo.

A continuación, se eliminaron las sales de los productos de reacción iPLEX Gold con *Clean Resin*, siguiendo el protocolo del fabricante. Los productos libres de sales se transfirieron de forma automatizada con un Nanodispensador RS1000 a la superficie de un *SpectroChip* de 384 pocillos cubierto por una matriz que permite la cristalización del producto a analizar. Después, un láser ioniza las moléculas del producto cristalizado. Los iones viajan a través de un tubo de vacío a una velocidad u otra en función de su masa hasta un detector acoplado al *software* MassARRAY Typer 4.0, que calcula su masa y determina así la base. Se analizaron el 5% de las muestras de forma aleatoria y por duplicado y la reproducibilidad fue del 100%.

3.2.1.4. Análisis estadístico de los resultados del genotipado

Para investigar las asociaciones entre los polimorfismos genotipados y la EM, se compararon las frecuencias alélicas de cada SNP entre los pacientes con EM y los controles sanos. Los datos crudos del genotipado se descargaron de la herramienta bioinformática SNPator (<http://www.snpator.org>) y se procesaron utilizando PLINK V.1.05. El equilibrio de Hardy-Weinberg para todos los SNPs bialélicos se evaluó aplicando un test exacto. Para evitar los falsos positivos debidos al *multiple testing* y considerando que todos los SNPs analizados no están en completo LD, se aplicó el método de Benjamini-Hochberg. Con el fin de determinar el

efecto y la independencia de la asociación de cada variante con la EM, se llevaron a cabo regresiones logísticas.

3.2.2. Cuantificación de los niveles de ARNm de *CXCR5* en pacientes con EM con distintos genotipos

3.2.2.1. Individuos

Se incluyeron un total de 33 pacientes con EM (formas de inicio con brote): 15 heterocigotos para la variante rs10892307 y 18 sin ella. Todos los pacientes incluidos se visitan regularmente en el Cemcat y ninguno había recibido tratamiento para la EM previo al estudio. La Tabla Suplementaria 1 (Anexo) resume las características demográficas de los pacientes incluidos.

3.2.2.2. Aislamiento de las células mononucleares de sangre periférica

Las CMSP de los 33 pacientes con EM se aislaron mediante centrifugación por gradiente de densidad Ficoll-Isopaque (Gibco BRL, Life Technologies LTD, Paisley, UK). Posteriormente, se realizaron dos lavados con suero fisiológico (Grifols, España) y las células se congelaron en nitrógeno líquido hasta su utilización en viales de al menos 5×10^6 células en 1 mL de medio RPMI completo (compuesto por RPMI 1640 (Gibco®, Invitrogen), suero bovino fetal (FBS) inactivado por calor al 10%, 2,0 mM de L-glutamina (Gibco®), 40U/mL de penicilina-streptomicina (Gibco®) y tampón Hepes 1M al 2% (Gibco®)) con un 10% de dimetil sulfóxido (DMSO).

3.2.2.3. Extracción y cuantificación del ARN total de CMSP

Las CMSP de los pacientes con EM se descongelaron parcialmente en el baño a 37°C e inmediatamente se añadió de forma paulatina medio RPMI completo precalentado a 37°C hasta su descongelación total. Después de un lavado con 10 mL de RPMI completo, el ARN total se extrajo mediante una técnica estándar utilizando el reactivo TRIzol™ (Invitrogen, UK). El protocolo detallado se encuentra en el Anexo. El ARN obtenido se cuantificó mediante el espectrofotómetro NanoDrop® ND-1000 (Thermo Fisher Scientific) utilizando 2 µL y la pureza se valoró mediante la evaluación del ratio 260/280 para descartar la contaminación con proteínas y del ratio 260/230 para descartar la existencia de restos de sales.

3.2.2.4. Síntesis de ADN complementario (ADNc)

Las muestras de ARN total se retrotranscribieron utilizando el kit High Capacity RT (Applied Biosystems, CA, EEUU). El material de partida fueron 400 ng de ARN total disueltos en 13 μ L de agua libre de RNasa. A este volumen se le añadieron 7 μ L de *máster mix* compuesta por: tampón RT 1 \times , 1 μ L de MultiscribeTM Reverse Transcriptase, 1 μ L de enzima inhibidora de la RNasa, 4 mM de dNTPs, random primers 1 \times y 0,2 μ L de agua libre de RNasa. Las reacciones se realizaron en tubos de 200 μ L. El equipo utilizado fue el termociclador GeneAmp[®] PCR System 2700 (Applied Biosistemas) y el programa de tiempos y temperaturas empleado fue el recomendado por el kit (Tabla 3). Por cada 1 μ g de ARN total añadido se obtiene 1 μ g de ADNc, por lo que se obtuvo el ADNc a una concentración de 20 ng/ μ L.

	Paso 1	Paso 2	Paso 3	Conservación
Temperatura (°C)	25	37	85	4
Tiempo	10 min	120 min	5 min	∞

Tabla 3. Programa de temperaturas y tiempos utilizado para la síntesis de ADN complementario a partir de ARN total.

3.2.2.5. Cuantificación de la expresión génica de CXCR5 por reacción en cadena de la polimerasa a tiempo real (*real-time* PCR)

La determinación de la expresión génica de *CXCR5* (*C-X-C motif chemokine receptor 5*) se realizó mediante PCR a tiempo real (*real-time* PCR) o PCR cuantitativa (qPCR, por *quantitative PCR*) utilizando ensayos de expresión génica TaqManTM. Cada ensayo TaqManTM consta de:

- Un par de primers específicos del gen de interés.
- Una sonda TaqManTM con un fluoróforo (FAMTM o VICTM) en el extremo 5', y un atenuador no fluorescente (NFQ, por *nonfluorescent quencher*) unido a un ligando de unión al surco menor (MGB, por *minor groove binder*) en el extremo 3'. El NFQ de las sondas TaqManTM-MGB asegura una menor señal de *background*, por lo que aumenta la precisión en la cuantificación.

Se utilizó un ensayo TaqManTM específico para *CXCR5* (Hs00540548_s1) (Applied Biosystems). Como control endógeno, se empleó el gen *GAPDH* (*glyceraldehyde-3-phosphate dehydrogenase*) (ensayo Hs99999905_m1) (Applied Biosystems). Cada reacción contenía 1 μ L de muestra de ADNc (20 ng), 10 μ L de TaqMan Universal PCR Master Mix, No AmpErase[®], UNG (Applied Biosystems), 8 μ L de agua libre de RNasa y 1 μ L de un ensayo TaqManTM. Las

muestras se determinaron por duplicado. Las reacciones se realizaron en placas ópticas de 384 pocillos y se ejecutaron en el termociclador ABI PRISM 7900HT *Sequence Detection* (Applied Biosystems). Los ajustes de la *real-time* PCR fueron los siguientes: 2 minutos a 50°C (activación de la enzima amperasa), 5 minutos a 95°C (desnaturalización del ADNc de doble cadena. Durante esta etapa, la señal de la sonda TaqMan™ fluorescente es atenuada por el NFQ) y 40 ciclos de 15 segundos a 95°C (hibridación de los *primers* y la sonda con sus secuencias diana específicas) y 1 minuto a 60°C (síntesis de nuevas hebras complementarias mediante la Taq-ADN polimerasa. Cuando la polimerasa alcanza una sonda TaqMan™, su actividad de nucleasa 5' escinde la sonda, separando el fluoróforo del NFQ y permitiendo así que se emita la señal fluorescente. En cada ciclo de la *real-time* PCR, se liberan más moléculas de fluoróforo, lo que resulta en un aumento en la intensidad de fluorescencia proporcional a la cantidad de amplicón sintetizado).

Los datos se analizaron mediante la aplicación de cuantificación relativa del software SDS 2.4 (Applied Biosystems) y el RQ Manager 1.2.1 (Applied Biosystems). A cada valor de ciclo umbral (Ct) del gen *CXCR5*, se le sustrajo el Ct del gen *GAPDH* (gen de referencia de expresión estable) para estimar la expresión génica mediante la normalización de los datos crudos. Los resultados se analizaron siguiendo el método $2^{-\Delta\Delta CT}$ (286) y se expresaron como *fold-change* de la expresión génica en pacientes con EM portadores de la variante rs10892307 en comparación con aquéllos sin ella.

3.2.2.6. Análisis estadístico de los resultados

El análisis estadístico se realizó utilizando el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Se aplicó la prueba de la U de Mann-Whitney para evaluar las diferencias en la expresión génica de *CXCR5* entre pacientes con EM con y sin la variante rs10892307.

3.2.3. Estudio de la región genómica del polimorfismo validado en el gen *CXCR5*

Para analizar la región genómica donde se localiza el polimorfismo rs10892307, se utilizó la aplicación bioinformática LDlink (<https://analysistools.nci.nih.gov/LDlink/>), cuyos datos proceden del Proyecto 1000 Genomas. Se utilizó el módulo LDproxy y se calculó el coeficiente de correlación al cuadrado, r^2 , para estimar el grado de desequilibrio de ligamiento entre rs10892307 y otros polimorfismos del genoma en poblaciones europeas. Se seleccionaron aquellos polimorfismos con un $r^2 > 0,5$ con el SNP rs10892307 para posteriores estudios. Una

vez seleccionados, se utilizó la base de datos Regulome DB (<http://www.regulomedb.org/>) para analizar su capacidad de afectar a la regulación de la transcripción de algún gen. Esta base de datos asigna a cada SNP una puntuación de 1 a 6 (con diferentes subcategorías) en función de la evidencia experimental existente sobre capacidad de regular la expresión génica, de modo que la puntuación refleja su relevancia como elemento regulador. Por último, se utilizó el navegador UCSC Genome Browser (<http://genome.ucsc.edu/>) para representar esquemáticamente la localización del SNP rs10892307 y de los seleccionados a partir de la información de Regulome DB, así como para analizar los elementos reguladores de la región (información generada a partir del Proyecto ENCODE (por *Encyclopedia of DNA Elements*)).

3.2.4. Ensayo reportero dual-luciferasa para polimorfismos del gen *CXCR5*

Para estudiar el efecto individual de los polimorfismos rs10892307, rs11602393, rs3176905 y rs55756957 en la expresión de *CXCR5* a nivel transcripcional, se utilizó el ensayo reportero dual-luciferasa. Esta parte del trabajo se desarrolló en el laboratorio del grupo "Genetic and molecular bases of multiple sclerosis", situado en el Instituto de Parasitología y Biomedicina López-Neyra (IPBLN, CSIC), en Granada, bajo la supervisión de los Doctores Fuencisla Matesanz, Antonio Alcina y Maria Fedetz. La Figura 16 esquematiza los pasos seguidos.

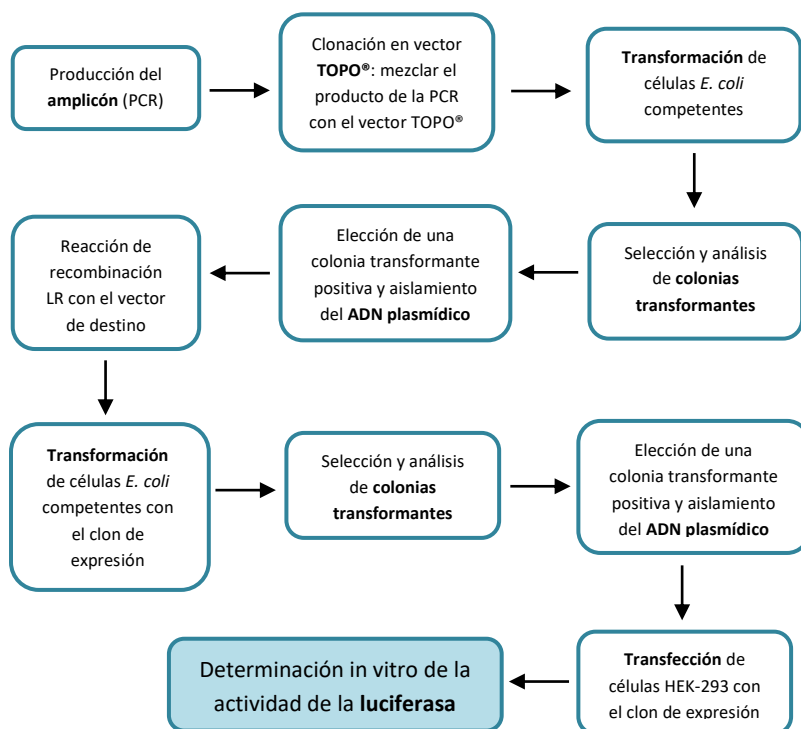


Figura 16. Diagrama de flujo descriptivo de las fases para estudiar la repercusión de las variantes rs10892307, rs11602393, rs3176905 y rs55756957 en la actividad transcripcional del promotor del gen *CXCR5*.

3.2.4.1. Diseño de los *primers* y amplificación de los fragmentos de interés

Se diseñaron 4 parejas de *primers*, una para cada polimorfismo, mediante el *software* Primer3. Cada pareja se diseñó para amplificar específicamente la región que contiene el polimorfismo de interés. Los *primers* fueron sintetizados por la compañía Metabion (Alemania). La Tabla 4 muestra las secuencias de todos los *primers*.

Polimorfismo	<i>Primer forward</i>	<i>Primer reverse</i>	Amplión
rs10892307	TCCACAGTGGGAGAGGATTC	CCTACTGCCTCAGGAGACGA	300 pb
rs11602393	CTGGAGGCTTGGAGAGAGTG	ATGCCCGGGTAGTTCTGTCT	389 pb
rs3176905	AGCAGGAGGGAGTCAGACAA	TTGGGTGGGCTAAGAAAATG	342 pb
rs55756957	CACAGCTCCCCTCTCGTTAG	GCTCCTCCATCAGAATCTCG	232 pb

Tabla 4. Parejas de *primers* utilizadas para amplificar cada una de las regiones reguladoras portadoras de los polimorfismos bajo estudio.

El material de partida fue el ADN genómico obtenido de una línea celular linfoblastoide (LCL) de un individuo del Proyecto 1000 Genomas heterocigoto para las 4 variantes bajo estudio. Se realizaron 4 PCR independientes, cada una de ellas para amplificar el fragmento de ADN donde se encuentra uno de los SNPs. En cada reacción, se añadieron 50 ng de ADN genómico (en 1 μ L) en un volumen final de 50 μ L, de los que 10 μ L eran Q5® Buffer (5 \times), 1 μ L dNTPs (10 nM), 2,5 μ L *primer forward* (10 μ M), 2,5 μ L *primer reverse* (10 μ M), 10 μ L Q5® GC (5x), 0,5 μ L enzima polimerasa Taq Q5® High-Fidelity (New England Biolabs, MA, EEUU) y 22,5 μ L H₂O doble destilada. Las reacciones se realizaron en tubos de 200 μ L y se ejecutaron en el termociclador MJ Mini™ Thermal Cycler (Bio-Rad Laboratories, Hercules, CA). Las condiciones de la PCR fueron las detalladas en la Tabla 5.

	Preamplificación	Desnaturalización	Hibridación	Elongación
Temperatura (°C)	98	98	60	72
Tiempo	30 s	10 s	20 s	20 s
Número de ciclos	1	←	32	→

Tabla 5. Protocolo de PCR diseñado para la amplificación de los fragmentos de ADN genómico portadores de los cuatro polimorfismos de interés.

3.2.4.2. Electroforesis en gel de agarosa

Tras la amplificación, se realizó una electroforesis en gel de agarosa al 2% para corroborar la amplificación de los 4 fragmentos. Se prepararon 2 g de agarosa por cada 100 mL de Tris-Acetate-EDTA (TAE) 1× disuelto en agua desionizada y se añadieron 5 µL de RedSafe™ (20.000×) (iNtRON Biotechnology, Seoul, Korea) por cada 100 mL. El ADN amplificado se cargó con un 16% de tampón de carga. Se utilizó el 100 bp DNA Ladder (Invitrogen) como marcador de peso molecular para estimar el tamaño de los fragmentos obtenidos en la PCR. La electroforesis se realizó a un voltaje de 100 voltios durante 45 minutos. La corriente se generó mediante el aparato PowerPac™ Basic (Bio-Rad Laboratories, Hercules, CA). El gel se analizó colocándolo bajo luz ultravioleta y realizando fotos para visualizar las bandas de ADN.

3.2.4.3. Purificación del producto de la PCR y adición de una cola de poli(A)

Una vez corroborados, los productos de la PCR se purificaron mediante el kit High Pure PCR Product Purification (Roche, Suiza) siguiendo las indicaciones del protocolo. El ADN se obtuvo en un volumen final de 50 µL de tampón de elución (10 mM Tris-HCl, pH 8,5) y se cuantificó mediante el espectrofotómetro NanoDrop™ 1000 (Thermo Fisher Scientific). A continuación, se añadieron desoxiadenosinas a los extremos romos 3' de los fragmentos de ADN amplificados para poder llevar a cabo la posterior ligación. Se realizaron 4 reacciones independientes, una para cada polimorfismo. En cada una, se añadieron 25 µL de ADN en un volumen final de 50 µL, de los cuales 5 µL eran (NH₄)₂SO₄ *buffer* (10×), 2 µL MgCl₂ *solution* (50mM), 5 µL dATP (1mM), 0,2 µL BIOTAQ™ *DNA Polymerase* (Bioline, Canadá) y 12,8 µL H₂O doble destilada. Las reacciones se incubaron a 72°C durante 20 minutos.

3.2.4.4. Ligación en pCR™8/GW/TOPO® – TA cloning

Para clonar el producto amplificado de la PCR en un vector, se utilizó el kit pCR™8/GW/TOPO® TA Cloning® Kit (Invitrogen). El vector TOPO® (Figura 17) consta de los siguientes sitios:

- Sitio de clonación TOPO®, para la clonación del producto de la PCR amplificado.
- Sitios EcoR I, que flanquean el sitio de clonación TOPO® para simplificar la escisión del producto de PCR clonado.

Materiales y métodos - Objetivo 2

- Sitios attL1 y attL2, para la transferencia del fragmento de interés al vector Gateway® mediante recombinación.
- Sitios de unión de primers específicamente diseñados dentro de los sitios attL1 y attL2 para facilitar la secuenciación utilizando los primers GW1 y GW2.
- Secuencias rrnB de terminación de la transcripción para evitar la expresión basal del producto de PCR de interés en *E. coli*.
- Marcador de resistencia a la espectinomicina, para la selección de las colonias de *E. coli* de interés.
- Origen de replicación pUC, para la replicación de un alto número de copias del plásmido en *E. coli*.

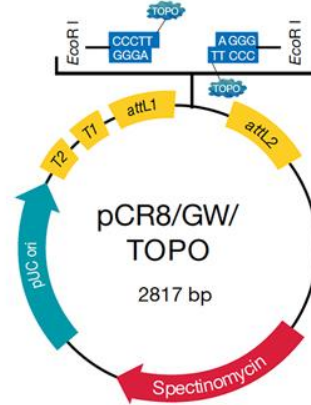


Figura 17. Esquema del vector pCR8/GW/TOPO®, diseñado para llevar a cabo el proceso de TA cloning® y, así, utilizar el vector resultante para múltiples aplicaciones.

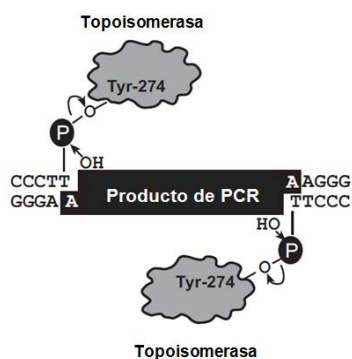


Figura 18. Representación gráfica de la reacción entre el vector TOPO® linearizado y el producto amplificado por PCR para su incorporación en el vector mediante TA cloning.

Como representa la Figura 18, la incorporación se produce gracias a que el vector TOPO® se suministra linearizado y con la secuencia específica 5'-CCCTT-3' en sus dos extremos. Se aprovecha la presencia de una desoxitimidina libre (T) en su extremo 3' para ligarse directamente con el producto amplificado con la cola de poli-A en su extremo 3' (TA cloning®). Esta reacción es catalizada por la topoisomerasa I del virus *Vaccinia*, unida de forma covalente a los dos extremos 3' del vector TOPO®. El producto de la ligación se denomina "clon de entrada". En la reacción TA cloning®, el inserto puede integrarse en sentido *forward* o *reverse* y, dado que el individuo de partida es heterocigoto para los cuatro polimorfismos, puede ser *wild type* o mutado. Todo ello depende del azar.

Se realizó una reacción para cada uno de los cuatro polimorfismos bajo estudio. En cada reacción de clonación TOPO®, se añadió 1 µL del producto de la PCR, 1 µL de solución salina (1,2 M NaCl; 0,06 M MgCl₂), 3 µL de H₂O doble destilada y 1 µL del vector TOPO®. La reacción se incubó 5 minutos a temperatura ambiente y se pasó rápidamente a hielo.

3.2.4.5. Transformación de células competentes (*E. coli*)

El producto de la reacción de clonación TOPO® se transformó en células competentes de *E. coli* (One Shot® Mach1™-T1^R Chemically Competent *E. coli* kit) (Invitrogen). Se llevaron a cabo cuatro transformaciones independientes, una para cada polimorfismo. Para ello, se añadió 1 µL del producto de la reacción TOPO® a 25 µL de células competentes de *E. coli*. Como control positivo para verificar la eficiencia de la transformación, se utilizó 1 µL (10 pg) del plásmido pUC19. Las células se incubaron en hielo durante 30 minutos. A continuación, las bacterias se sometieron a un choque térmico de 42°C durante 30 segundos e inmediatamente se transfirieron a hielo durante 2 minutos. Se añadieron 125 µL de medio S.O.C. a temperatura ambiente y las bacterias se incubaron a 37°C a 225 rpm durante 1 hora. Una vez transcurrido este tiempo, se sembraron 100 y 25 µL de células transformadas en dos placas de agar con LB precalentadas a temperatura ambiente con 100 µg/mL de espectinomicina (o ampicilina en el caso de pUC19) para la selección de las bacterias que habían incorporado el plásmido recombinante resistente al antibiótico. Las placas se incubaron invertidas toda la noche a 37°C.

3.2.4.6. Análisis de las colonias de bacterias transformantes

Se inocularon 24 colonias para cada polimorfismo en 100 µL de agua MiliQ cada una, con la finalidad de encontrar las cuatro combinaciones posibles orientación-alelo: *forward*+WT, *forward*+MA, *reverse*+WT y *reverse*+MA. El ADN plasmídico de todas las colonias inoculadas se amplificó mediante PCR para verificar que todas contienen el vector TOPO®. En cada reacción de amplificación, se añadieron 2 µL de la colonia inoculada a 8 µL de Máster Mix (25 µL de (NH₄)₂SO₄ *buffer* (10×), 5 µL de dNTP Mix (10µM), 12,5 µL del *primer forward* del SNP correspondiente (10 µM), 12,5 µL del *primer reverse* del SNP correspondiente (10 µM), 10 µL de MgCl₂ *solution* (50mM), 2 µL de BIOTAQ™ *DNA Polymerase* (Bioline, Canadá) y 133,5 µL de H₂O MiliQ). Las reacciones se realizaron en placas de 384 pocillos y se ejecutaron en el termociclador DNA Engine® (Bio-Rad Laboratories, Hercules, CA). Las condiciones de la PCR fueron las especificadas en la Tabla 6.

	Preamplificación	Desnaturalización	Hibridación	Elongación
Temperatura (°C)	94	94	60	72
Tiempo	3 min	20 s	20 s	20 s
Número de ciclos	1	←	30	→

Tabla 6. Condiciones de la PCR diseñada para amplificar el ADN plasmídico de las bacterias *E.coli* competentes constituyentes de diferentes colonias y previamente transformadas con los constructos de interés.

A continuación, se realizó una electroforesis en gel de agarosa al 2% para visualizar el producto amplificado y corroborar que cada colonia seleccionada incorporó el producto de la reacción de clonación TOPO®.

3.2.4.7. Crecimiento de las colonias positivas y extracción de plásmidos

Para cada SNP, se escogieron una media de 15 colonias en base a las bandas visualizadas en el gel de agarosa. Se inocularon 40 µL de cada colonia seleccionada en 5 mL de medio LB líquido con 100 µg/mL de espectinomicina y se incubaron en agitación a 37°C toda la noche para promover su crecimiento. Al día siguiente, se extrajo el ADN plasmídico de las colonias bacterianas crecidas mediante el kit PureLink® Quick Plasmid Miniprep Kit (Invitrogen). Se siguieron las indicaciones del protocolo incluido y se obtuvieron hasta 30 µg de ADN plasmídico en 75 µL de buffer TE.

3.2.4.8. Secuenciación y análisis de los plásmidos extraídos

Con el objetivo de encontrar las cuatro combinaciones posibles de cada polimorfismo, se secuenciaron los plásmidos con el inserto extraídos de todas las colonias para comprobar:

- a) Si el fragmento insertado es *wild type* o porta el alelo minoritario (MA, por *minor allele*) para el SNP de interés.
- b) Si la orientación del fragmento insertado es *forward* o *reverse*.

Para ello, se aprovechó la existencia del sitio de unión del primer GW1 en el vector TOPO®, especialmente diseñado para facilitar la secuenciación, cuya secuencia es la siguiente:

GW1: 5'-GTTGCAACAAATTGATGAGCAATGC-3'

Se añadieron una media de 600 ng del ADN plasmídico y 6,4 pmol del primer GW1 en un volumen final de 12 µL de H₂O MiliQ. La secuenciación se realizó en el Servicio de Secuenciación del IPBLN mediante el procedimiento de secuenciación cíclica utilizando la química BigDyeTerminator v3.1 y la electroforesis en un sistema multicapilar automático. El análisis de las secuencias se llevó a cabo mediante el software bioinformático DNASIS® MAX 3.0. Si una vez analizadas las secuencias no se consiguieron las cuatro combinaciones para un determinado polimorfismo, se volvió al punto 3.2.4.4. (ligación en el vector TOPO®) con dicho polimorfismo y así sucesivamente hasta tener los 4 vectores.

3.2.4.9. Transferencia del inserto al vector de expresión Gateway®

La tecnología de clonación Gateway® es un método basado en la recombinación sitio-específica del fago λ mediante el que se transfieren secuencias de ADN entre diferentes vectores, manteniéndose tanto el marco abierto de lectura (ORF, por *open reading frame*) como la orientación del inserto.

En el presente estudio, partimos de un vector TOPO®, denominado "clon de entrada", que contiene la secuencia de interés clonada flanqueada por dos sitios de recombinación: attL1 y attL2. Se realizó una reacción LR, en la que se transfirió el inserto a un vector Gateway®, conocido como "vector de expresión" o "vector destino". Para ello, los sitios attL1 y attL2 del "clon de entrada" recombinan con los sitios attR1 y attR2 del "vector destino" y dan lugar al "clon de expresión" (Figura 19). Esta reacción está catalizada por un conjunto de enzimas que contiene las proteínas Int, IHF y Xis.

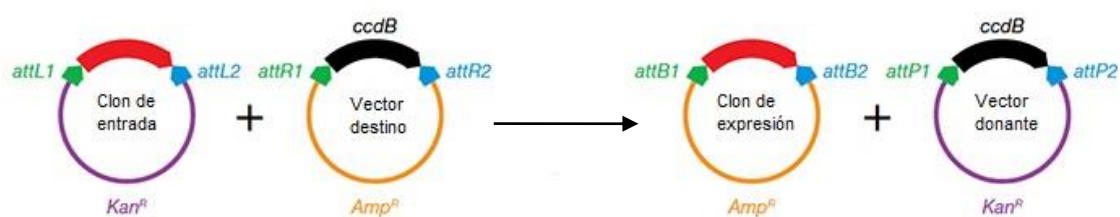


Figura 19. Ilustración descriptiva de la reacción LR para transferir el fragmento insertado en el clon de entrada (vector TOPO®) a un vector destino (vector Gateway®), dando lugar al clon de expresión, que contiene un gen reportero que codifica la enzima luciferasa. Adaptada de <https://www.thermofisher.com>.

En este trabajo, utilizamos el vector destino pGL4.23-GW, esquematizado en la Figura 20. A continuación, se mencionan los sitios de este vector que son de interés para nuestro estudio:

- Región de clonación, en la que se insertó el gen *ccdB* para realizar la selección negativa. El gen *ccdB* interfiere con la actividad ADN girasa de *E.coli*, inhibiendo así el crecimiento de las bacterias. Cuando ocurre la recombinación con el vector TOPO®, el gen *ccdB* es reemplazado por el inserto de interés. Las bacterias que incorporan los plásmidos que no tienen el inserto (y por tanto siguen teniendo *ccdB*), mueren.
- Promotor mínimo.
- El gen reportero *luc2*, que codifica la enzima luciferasa.
- Gen de resistencia a la Ampicilina.
- Región de unión del primer RV.

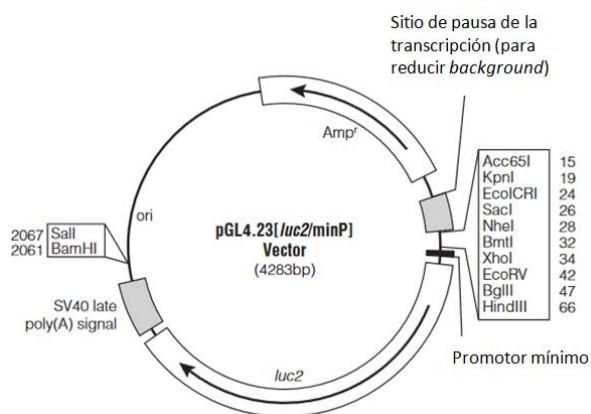


Figura 20. Explicación esquemática del vector destino pGL4.23-GW, de Promega. En el presente estudio, se utilizó una versión modificada del mismo (Addgene), constituida por las mismas regiones.

Para llevar a cabo la reacción LR, se utilizó el kit LR Clonase™ enzyme mix (Invitrogen). Se añadieron 100 ng del clon de entrada (en 1-3 μ L), 300 ng del vector destino pGL4.23-GW (en 2 μ L), 4 μ L de 5x LR Clonase™ reaction buffer LR y hasta 16 μ L de buffer TE a pH 8,0. A continuación, se añadieron 4 μ L de Clonase™ enzyme mix a cada muestra y se incubó la reacción a 25°C durante 1 hora. Transcurrido este tiempo, se añadieron 2 μ L de Proteinasa K (2 μ g/ μ L) a cada reacción y se incubaron a 37°C durante 10 minutos.

3.2.4.10. Transformación de células competentes (*E. coli*) con el clon de expresión y análisis de las colonias

El clon de expresión producto de la reacción LR se transformó en células competentes de *E. coli* (apartado 3.2.4.5.). Se llevaron a cabo 16 transformaciones independientes, 4 para cada polimorfismo (las 4 combinaciones posibles), utilizando 1 μ L del producto de la reacción LR y siguiendo el protocolo detallado en el apartado 3.2.4.6. En este caso, se sembraron 100 y 25 μ L de células transformadas en dos placas de agar con LB precalentadas a temperatura ambiente que contenían 100 μ g/mL de ampicilina para seleccionar las bacterias que habían incorporado el clon de expresión resistente a la ampicilina. Las placas se incubaron toda la noche a 37°C.

Al día siguiente, se inocularon 5 colonias de bacterias para cada uno de los 16 clones de expresión en 100 μ L de H₂O MiliQ. El ADN plasmídico de las colonias inoculadas se amplificó mediante PCR para verificar que contienen el clon de expresión, tal y como se detalló en el apartado 3.2.4.6. para el vector TOPO®. A continuación, se realizó una electroforesis en gel de agarosa al 2% como se indicó en el apartado 3.2.4.2. para visualizar el producto amplificado y verificar que cada colonia seleccionada era positiva para el vector de expresión.

3.2.4.11. Crecimiento de las colonias positivas para el clon de expresión y extracción del ADN plasmídico

Se escogió una colonia para cada vector de expresión (en total, 16 colonias) en base a las bandas visualizadas en el gel de agarosa. Se inocularon 40 μL de cada colonia seleccionada en 50 mL de medio LB líquido con 100 $\mu\text{g/mL}$ de ampicilina y se incubaron en agitación a 37°C toda la noche para promover su crecimiento.

Al día siguiente, se extrajo el ADN plasmídico de las colonias crecidas mediante el kit EndoFree Plasmid Midi Kit (Qiagen, Santa Clarita, USA). Se siguieron las indicaciones del protocolo incluido y se obtuvieron hasta 100 μg de ADN plasmídico en 100 μL de buffer TE. Como control negativo se utilizó el plásmido pGL4.23 sin la región de clonación o el cassette del Gateway® (Promega, Madison, EEUU), el cual se preparó de la misma forma. Como valor de control interno para normalizar la expresión del gen reportero luciferasa de luciérnaga de los plásmidos de interés, se cotransfectó el vector pRL-TK (Promega), que expresa la luciferasa de renilla (*Rluc*) bajo el control del promotor constitutivo de la timidina quinasa (Figura 21).

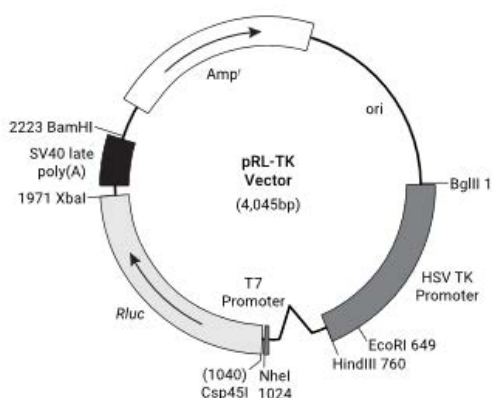


Figura 21. Mapa del vector pRL-TK codificante de la luciferasa de renilla, cuyo valor de expresión se utilizó como control interno para normalizar la de la luciferasa de luciérnaga influenciada por los insertos bajo estudio.

Se secuenciaron los plásmidos purificados para asegurar que se consiguieron los 16 constructos de interés. La secuenciación se realizó como se detalló en el apartado 3.2.4.8., pero en este caso utilizando el *primer* RVprimer3.

3.2.4.12. Transfección transitoria y determinación de la actividad luciferasa

Para los ensayos de transfección transitoria, se utilizó la línea celular HEK-293 (células embrionarias de riñón humano que crecen en suspensión). Las células HEK-293 se mantuvieron a 37°C, con CO₂ al 5%, en medio DMEM suplementado con 2 mM L-Glutamina, 10% suero

Materiales y métodos - Objetivo 2

fetal bovino descomplementado (FBS) y 1% aminoácidos no esenciales (todo de GIBCO-BRL). El medio se cambió cada 2 ó 3 días y el crecimiento se mantuvo hasta alcanzar un ~90% de confluencia, momento en el que se trasplantaban a un nuevo soporte.

La transfección se realizó mediante el reactivo jetPRIME® (Polyplus-Transfection, NY, EEUU), basado en polímeros, estable en presencia del suero y los antibióticos, y con baja citotoxicidad. Se llevaron a cabo 4 experimentos de transfección independientes para cada uno de los 16 constructos. Para cada experimento, se utilizaron placas de 96 pocillos y se sembraron 10.000 células HEK-293 en 100 μ L de medio completo por pocillo 24 horas antes del experimento. En el momento de la transfección, las células estaban a un 60-80% de confluencia.

Se prepararon 16 mezclas que contenían 50 μ L de jetPRIME® buffer, 1 μ g del plásmido correspondiente, 2 ng del plásmido utilizado como control interno pRL-TK y 2 μ L del reactivo JetPRIME®. Las mezclas se incubaron 10 minutos a temperatura ambiente y se añadieron 5 μ L de cada una de ellas a los pocillos correspondientes, evitando el contacto directo con las células. La placa se mezcló suavemente y se incubó a 37°C con un 5% de CO₂. Después de 24 horas, la luminiscencia fue analizada en los sobrenadantes de cada pocillo utilizando el kit Dual Luciferase® Reporter Assay System (Promega).

Las medidas obtenidas se analizaron comparando la relación entre la luminiscencia de la luciferasa de luciérnaga y la de la luciferasa de renilla en cada pocillo, procedimiento realizado para los 16 constructos y para el vector pGL4.23 vacío. Después, se comparó la actividad transcripcional del promotor de cada constructo con la del vector vacío.

3.2.4.13. Análisis estadístico de los resultados del ensayo reportero de luciferasa

Dado que se realizaron 4 experimentos independientes para cada uno de los 16 vectores construidos, se realizó la media de las 4 medidas de luminiscencia ya normalizadas respecto a la luminiscencia emitida por el vector pRL-TK (renilla) para cada constructo. Se aplicó el mismo procedimiento para el vector pGL4.23 vacío. Para evaluar la diferencia entre la actividad transcripcional del vector pGL4.23 vacío y la de los 16 vectores pGL4.23-GW construidos, así como entre diferentes vectores de un mismo SNP, se aplicó la U de Mann-Whitney.

3.2.5. Análisis por citometría de flujo de la expresión de CXCR5 en poblaciones de CMSP de pacientes con EM con distintos genotipos

3.2.5.1. Individuos

Se incluyeron un total de 20 pacientes con EM (formas de inicio con brote). De éstos, 10 eran portadores del alelo minoritario del SNP rs10892307 y 10 no portaban dicho alelo. Todos los pacientes incluidos se visitan regularmente en el Cemcat y ninguno de ellos había recibido ningún tratamiento para la EM previo a la extracción de la muestra del estudio. La Tabla Suplementaria 2 (Anexo) resume las características demográficas de los pacientes incluidos.

3.2.5.2. Obtención y mantenimiento de las CMSP

Las CMSP de los 20 pacientes se aislaron y congelaron como se explicó en el apartado 3.2.2.2. hasta el día de la utilización. El experimento constó de tres días experimentales, en los que se procesaron 8, 6 y 6 muestras con representación proporcional de los pacientes con el alelo minoritario y de los pacientes sin él. Las CMSP se descongelaron parcialmente en el baño a 37°C y se añadió de forma progresiva medio RPMI completo precalentado a 37°C hasta su descongelación total. Las CMSP se lavaron con 10 mL de medio completo precalentado y se resuspendieron en 1 mL de medio para su conteo con la cámara de Neubauer. Los cultivos de CMSP se realizaron en tubos de vidrio de 5 mL. En cada uno, se sembraron 2×10^6 células a una concentración de 2×10^6 células/mL en RPMI completo y se incubaron toda la noche a 37°C en una atmósfera con 5% de CO₂. Al día siguiente, se recogieron las células para proseguir con el marcaje para citometría de flujo.

3.2.5.3. Anticuerpos utilizados

Para el estudio de la expresión de CXCR5 en diferentes subpoblaciones de células T y células B periféricas se realizaron marcajes de superficie e intranucleares utilizando los anticuerpos detallados en la Tabla 7.

Anticuerpo	Fluorocromo	Isotipo	Clon	Volumen	Marcaje	Casa comercial
Anti-CD3 de humano	FITC	IgG1, κ de ratón	UCHT1	5 μL	Superficie	BD Biosciences
Anti-CD4 de humano	APC·H7	IgG1, κ de ratón	RPA-T4	0,25 μL	Superficie	BD Biosciences
Anti-CD45RO de humano	BV605	IgG2a, κ de ratón	UCHL1	0,5 μL	Superficie	BD Biosciences
Anti-CD25 de humano	BV421	IgG1, κ de ratón	BC96	4 μL	Superficie	Biolegend
Anti-FoxP3 de humano	PE	IgG2a, κ de rata	PCH101	5 μL	Intranuclear	eBioscience
Anti-CD19 de humano	PerCP·Cy5.5	IgG1, κ de ratón	HIB19	0,25 μL	Superficie	BD Biosciences
Anti-CD20 de humano	APC·H7	IgG2b, κ de ratón	2H7	0,25 μL	Superficie	BD Biosciences
Anti-CD27 de humano	BV605	IgG1, κ de ratón	L128	2,5 μL	Superficie	BD Biosciences
Anti-CD38 de humano	PE	IgG1, κ de ratón	HIT2	20 μL	Superficie	BD Biosciences
Anti-CD24 de humano	BB515	IgG2a, κ de ratón	ML5	2,5 μL	Superficie	BD Biosciences
Anti-CXCR5 de humano	Alexa Fluor 647	IgG2b, κ de rata	RF8B2	0,25 μL	Superficie	BD Biosciences

Tabla 7. Anticuerpos utilizados para el estudio de la expresión superficial de CXCR5 en las poblaciones seleccionadas de linfocitos T y linfocitos B circulantes. El volumen se corresponde con los μL añadidos a cada muestra.

3.2.5.4. Diseño de los paneles de citometría

Se diseñaron los tres paneles detallados en la Tabla 8. Para cada población señalada, se estudió la expresión de CXCR5. Para ello, se analizó el porcentaje de células que expresaban CXCR5 y la intensidad media de fluorescencia (MFI, por *mean fluorescence intensity*) para CXCR5.

	Población	Marcadores			
PANEL 1	Linfocitos T CD4 ⁺	CD3 ⁺	CD4 ⁺		
	Linfocitos T CD8 ⁺	CD3 ⁺	CD4 ⁻		
	Linfocitos T CD4 ⁺ <i>naïve</i>	CD3 ⁺	CD4 ⁺	CD45RO ⁻	
	Linfocitos T CD4 ⁺ memoria	CD3 ⁺	CD4 ⁺	CD45RO ⁺	
PANEL 2	Linfocitos T reguladores	CD3 ⁺	CD4 ⁺	CD25 ⁺	FoxP3 ⁺
PANEL 3	Linfocitos B	CD19 ⁺			
	Linfocitos B <i>naïve</i>	CD19 ⁺	CD20 ⁺	CD27 ⁻	
	Linfocitos B memoria	CD19 ⁺	CD20 ⁺	CD27 ⁺	
	Plasmablastos	CD19 ⁺	CD27 ^{high}	CD38 ^{high}	
	Linfocitos B reguladores	CD19 ⁺	CD24 ⁺	CD27 ^{high}	

Tabla 8. Poblaciones celulares estudiadas y marcadores utilizados para su identificación mediante citometría de flujo. En todas ellas, se analizó la expresión de CXCR5.

Cabe mencionar que no se analizó la expresión de CXCR5 por las subpoblaciones *naïve* y de memoria de los linfocitos T CD8⁺ debido a la baja expresión de dicha molécula.

3.2.5.5. Marcaje celular

El marcaje con los anticuerpos correspondientes se realizó en placas de 96 pocillos en un volumen final de 100 µL. En primer lugar, se añadieron 5×10^5 CMSP de cada paciente en un pocillo de la misma y se lavaron dos veces. A continuación, se realizó el marcaje de viabilidad para descartar las células muertas, bien con el Fixable Viability Dye eFluor™ 450 (eBioscience, Hatfield, UK) para los paneles 1 y 2, o bien con el Fixable Viability Stain 510 (BD Biosciences) para el panel 3. Después se bloquearon los receptores de superficie Fc para evitar las uniones inespecíficas, para lo que se utilizó el Human BD Fc Block™ (BD Biosciences). Posteriormente, las células se incubaron durante 30 minutos con los cócteles de anticuerpos correspondientes según el panel. Por último, tras un lavado con PBAzida, las células se resuspendieron en 200 µL de PBAzida para adquirirlas en el citómetro. El protocolo detallado se encuentra en el Anexo.

En el caso del panel 3, una vez finalizado el marcaje de superficie de los antígenos CD3, CD4 y CD25, se realizó el marcaje intranuclear para detectar FoxP3 utilizando el kit de permeabilización Anti-Human Foxp3 Staining Set PE (eBioscience) siguiendo las instrucciones del kit. El protocolo detallado se encuentra en el Anexo.

3.2.5.6. Adquisición y análisis de las muestras en el citómetro de flujo

Las células se adquirieron utilizando el citómetro de flujo CytoFLEX, de Beckman Coulter (Brea, CA, USA). Los datos obtenidos se analizaron con el *software* CytExpert 1.0 (Beckman Coulter). En cada panel, se llevó a cabo una estrategia de *gating* concreta para identificar a las poblaciones celulares bajo estudio. En primer lugar, y en todos los casos, se combinaron los parámetros FS (tamaño celular) y SS (complejidad celular) para definir morfológicamente la población linfocitaria. En segundo lugar, se combinaron diferentes parámetros en función del panel bajo análisis. La estrategia seguida para identificar las poblaciones de cada panel se detalla en el Anexo.

Para calcular la MFI de CXCR5 (marcado con un anticuerpo unido conjugado con Alexa Fluor 647), el valor obtenido para cada muestra se normalizó respecto al valor de la MFI de Alexa

Fluor 647 de un control negativo sin marcaje para dicho fluorocromo. De este modo, se evitó el sesgo entre los diferentes días de adquisición.

3.2.5.7. Análisis estadístico de los resultados de la citometría de CXCR5

El análisis estadístico se llevó a cabo utilizando el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Se aplicó la U de Mann-Whitney para evaluar si existen diferencias tanto en el porcentaje de diferentes poblaciones de células T y células B que expresan CXCR5, como en su MFI de CXCR5 entre pacientes con EM con y sin el alelo minoritario del SNP rs10892307.

3.2.6. Determinación de los niveles de calcidiol (25(OH)D) y calcitriol (1,25(OH)₂-D) en suero de pacientes con EM con distintos genotipos del polimorfismo rs2762943 (CYP24A1)

3.2.6.1. Individuos

i) Individuos incluidos en la determinación de la concentración de 25(OH)D en suero

Se incluyeron un total de 101 pacientes con EM (formas de inicio con brote). De éstos, 30 (29,7%) eran portadores del alelo minoritario del polimorfismo rs2762943 (gen *CYP24A1*) y 71 (70,3%) carecían de dicho alelo. Ninguno de estos pacientes había recibido tratamiento para la EM previo a la extracción de la muestra utilizada para este estudio. La Tabla Suplementaria 3 (Anexo) muestra las características demográficas de los pacientes con EM incluidos, así como la estación del año en la que se obtuvo la muestra de suero, dado que los niveles de vitamina D fluctúan con la misma.

ii) Individuos incluidos en la determinación de la concentración de 1,25(OH)₂-D en suero

Se incluyeron un total de 100 pacientes con EM (formas de inicio con brote), todos ellos independientes de los utilizados en el apartado anterior, de los cuales 26 (26%) eran portadores del alelo minoritario de la variante común rs2762943 y 74 (74%) no poseían dicho alelo. Ninguno de los pacientes había recibido tratamiento para la EM previo a la extracción de la

muestra de suero empleada en este estudio. La Tabla Suplementaria 4 (Anexo) resume las características demográficas de los pacientes con EM incluidos y la estación del año en la que se obtuvo la muestra.

3.2.6.2. Obtención de las muestras de suero

La sangre periférica de los pacientes con EM incluidos en el estudio se obtuvo mediante punción venosa. Las muestras se dejaron coagular durante 30 minutos a temperatura ambiente y posteriormente se centrifugaron a 3000 rpm durante 10 minutos. El suero se alicuotó y se conservó a -80°C hasta su utilización.

3.2.6.3. Determinación de los niveles de 25-OH vitamina D en suero

Los niveles de 25(OH)D en suero se determinaron utilizando el kit *LIAISON® 25 OH Vitamin D Total Assay*, de DiaSorin (Stillwater, MN, USA). Se trata de un inmunoensayo competitivo directo por quimioluminiscencia para la determinación cuantitativa del total de 25(OH)D en suero. Este ensayo se ejecutó en el autoanalizador *LIAISON® XL*. El protocolo detallado se encuentra en el Anexo.

El kit mide entre 4 y 150 ng/mL. Los resultados son la suma de la vitamina D₃ (colecalfiferol) y la vitamina D₂ (ergocalciferol), las formas más abundantes de vitamina D en la naturaleza, que en conjunto se denominaron 25-(OH)-D.

3.2.6.4. Determinación de los niveles de 1,25(OH)₂-vitamina D en suero

Para la determinación de los niveles de 1,25(OH)₂-D en suero se utilizó el kit *LIAISON® XL 1,25 Dihydroxyvitamin D Assay*, de DiaSorin (Stillwater). Esta tecnología consiste en un inmunoensayo quimioluminiscente no competitivo o tipo sándwich de 3 pasos modificado que utiliza una proteína de fusión obtenida mediante tecnología de ADN recombinante (RFP, de *Recombinant Fusion Protein*), que captura la molécula de 1,25(OH)₂-D y cambia su conformación, y un anticuerpo monoclonal murino conjugado con un derivado del isoluminol que reconoce específicamente el complejo RFP-1,25(OH)₂-D y se une a él (Figura Suplementaria 1). Una vez unido el anticuerpo al complejo, se agregan los reactivos inductores para que ocurra una reacción quimioluminiscente. La señal luminosa se midió en RLU con un

fotomultiplicador y es proporcional a la concentración de 1,25(OH)₂-D. El ensayo se ejecutó en el autoanalizador LIAISON® XL.

El límite inferior de cuantificación es 5,0 pg/mL. Según el fabricante, los límites inferior y superior del rango de referencia son 19,9 y 79 pg/ml. Los resultados son la suma de 1,25(OH)₂-D₂ y 1,25(OH)₂-D₃ y se denominaron 1,25(OH)₂-D.

3.2.6.5. Análisis estadístico de los resultados de la determinación de la vitamina D en suero

Para el análisis estadístico de las medidas de 25-(OH)-D y 1,25(OH)₂-D obtenidas, se utilizó el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Para evaluar si existían diferencias significativas en los niveles de ambas formas de vitamina D en suero entre pacientes con EM con y sin la variante común rs2762943, se aplicó un T test no pareado.

3.2.7. Inducción de la expresión de *CYP24A1* en respuesta a la estimulación con vitamina D en CMSP de pacientes con EM con distintos genotipos del polimorfismo rs2762943 (*CYP24A1*)

3.2.7.1. Individuos

Se incluyeron 16 pacientes con EM (formas de inicio con brote), de los que 8 portaban el alelo minoritario del polimorfismo rs2762943 y 8, utilizados como controles, carecían de dicho alelo. Ninguno de los individuos había recibido tratamiento para la EM antes de la extracción de la muestra utilizada en este estudio. La Tabla Suplementaria 5 (Anexo) resume las características demográficas de los pacientes.

3.2.7.2. Aislamiento de las CMSP

Las CMSP de los 16 pacientes con EM se aislaron mediante la técnica de centrifugación por gradiente de densidad Ficoll-Isopaque (Gibco BRL) y se congelaron en nitrógeno líquido tal y como se detalló en el apartado 3.2.2.2.

3.2.7.3. Estimulación de CMSP con 1,25(OH)₂-D₃.

Las CMSP se descongelaron y contaron tal y como se explicó en el apartado 3.2.5.2. Se sembraron 1.5×10^6 células en cada tubo (3 tubos por paciente) a una concentración final de 2×10^6 células/mL en medio completo y se incubaron toda la noche a 37°C en una atmósfera con 5% de CO₂. Al día siguiente, cada paciente se trató de la siguiente forma: el primer tubo no recibió ningún tipo de tratamiento; el segundo tubo se utilizó como control vehículo (etanol al 95% en agua libre de ribonucleasas); y el tercer tubo se trató con 1 nM de 1,25(OH)₂-D₃ (Sigma-Aldrich). Las células se introdujeron de nuevo en el incubador 24 horas.

3.2.7.4. Extracción y cuantificación del ARN total de CMSP y síntesis de ADNc

El ARN total de las muestras de CMSP recogidas a las 24 horas post-tratamiento se aisló utilizando las columnas QIAshredder® y el kit comercial RNeasy® Mini Kit (Qiagen). Se siguió el protocolo incluido en el kit. Brevemente, las células se lisaron con el buffer de lisis RLT (al que previamente se le añadieron 10 µl de β-Mercaptoetanol por cada 1 ml de buffer). El lisado se cargó en la columna QIAshredder y se centrifugó para eliminar los restos celulares y mejorar la manipulación de las muestras después de la lisis. A partir del homogeneizado de CMSP obtenido, se aisló el ARN total siguiendo las instrucciones del protocolo del kit RNeasy® Mini Kit. Las muestras de ARN total se eluyeron en un volumen final de 30 µL de agua libre de RNasas y se cuantificaron en el espectrofotómetro NanoDrop® ND-1000 (Thermo Fisher Scientific), tal y como se detalló previamente (apartado 3.2.2.3). La retrotranscripción de las muestras de ARN total obtenidas de las CMSP se realizó mediante el kit High Capacity RT (Applied Biosystems), tal y como se detalló en el apartado 3.2.2.4. Una vez finalizada la reacción, se obtuvo el ADNc a una concentración de 20 ng/µL en un volumen total de 20 µL.

3.2.7.6. Cuantificación de la expresión génica de *CYP24A1* por reacción en cadena de la polimerasa a tiempo real (*real-time* PCR)

La cuantificación relativa de la expresión génica de *CYP24A1* tanto en las muestras tratadas con vitamina D como en las muestras sin tratar se realizó mediante *real-time* PCR. La expresión se determinó en base a la cuantificación de la expresión endógena del gen de referencia *GAPDH*. Los ensayos TaqMan™ específicos utilizados fueron Hs00167999_m1, para el gen *CYP24A1*, y Hs02786624_g1, para el gen *GAPDH* (Applied Biosystems).

El protocolo seguido fue el detallado en el apartado 3.2.2.5. Los datos se analizaron con el software SDS 2.4 (Applied Biosystems) y el RQ Manager 1.2.1 (Applied Biosystems). Los

resultados se analizaron siguiendo el método $2^{-\Delta\Delta CT}$ (286) y se expresaron como *fold-change* de i) la expresión génica basal de *CYP24A1* y de ii) la expresión génica de *CYP24A1* tras el tratamiento con $1,25(OH)_2-D_3$ en pacientes con EM con el polimorfismo rs2762943 en comparación con aquéllos sin él. También se analizó el grado de inducción del gen *CYP24A1* tras recibir un tratamiento u otro en ambos grupos de pacientes.

3.2.7.7. Análisis estadístico de los resultados de la estimulación de *CYP24A1*

El análisis estadístico se realizó utilizando el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Para evaluar si existen diferencias en la expresión génica de *CYP24A1* tanto basal como tras haber recibido el tratamiento con $1,25(OH)_2-D_3$ entre pacientes con EM con y sin la variante rs2762943, se aplicó la prueba de la U de Mann-Whitney. El cambio en la expresión de *CYP24A1* tras su inducción con $1,25(OH)_2-D_3$ se analizó aplicando un T test pareado.

3.3. Metodología relacionada con el objetivo 3

“Validar y caracterizar experimentalmente las variantes genéticas estructurales detectadas en la resecuenciación de ADN e identificar nuevos individuos portadores”

3.3.1. Cuantificación de la expresión de los genes que acumulan variantes raras en pacientes con EM

3.3.1.1. Individuos

Se cuantificaron los niveles de ARN mensajero de cuatro genes diferentes en pacientes con EM (formas de inicio con brote) que acumulan VRs en la región específica del gen reportada como enriquecida. Todos estos pacientes forman parte de la cohorte de resecuenciación de ADN. Ninguno de ellos había recibido ningún tratamiento de la EM previo al estudio. Para cada gen, se incluyeron pacientes diferentes:

- Para los experimentos con el gen *CYP24A1* se incluyeron 11 pacientes con EM con al menos una VR en regiones reguladoras de dicho gen, 4 con al menos una VR en sus regiones UTR, 4 con al menos una VR en su promotor y 25 pacientes con EM sin VRs en el gen *CYP24A1*.
- En el caso del gen *RGS1*, el estudio englobó 6 pacientes con EM con una o más VRs de *RGS1* considerando todas las regiones del gen, 4 pacientes con EM con al menos una VR de *RGS1* considerando sólo las regiones reguladoras y 28 pacientes con EM sin VRs del mismo gen.
- Para el gen *FCRL1*, se incluyeron 6 pacientes con EM con una ó más VRs acumuladas en el promotor *FCRL1* y 15 pacientes con EM que carecían de VRs de dicho gen.
- Para cuantificar la expresión génica de *TRAF3*, se utilizaron muestras de 4 pacientes con EM con al menos una VR en las regiones UTR de dicho gen y 16 pacientes con EM sin VRs del mismo.

Las muestras de CMSP de estos pacientes procedieron de diferentes centros de EM: el Instituto de Investigación Biomédica de Málaga, el Hospital Dr. Josep Trueta (Girona), el Hospital Clinic (Barcelona) y el Cemcat (Barcelona). La Tabla Suplementaria 6 (Anexo) detalla la información demográfica de los pacientes, así como la procedencia de las muestras utilizadas para analizar la expresión de cada gen. Las muestras de CMSP de los pacientes se enviaron y centralizaron en nuestro instituto de investigación (Institut de Recerca Vall d’Hebron), donde se

procedió a la extracción del ARN total y a la síntesis del ADN complementario para llevar a cabo la *real-time* PCR.

3.3.1.2. Aislamiento de las CMSP, extracción de ARN total y síntesis de ADNc

Las CMSP de los pacientes con EM se obtuvieron en cada centro utilizando protocolos estándar. En el caso de los pacientes del Cemcat, se aislaron y se congelaron tal y como se detalló en el apartado 3.2.2.2. Todas las CMSP se centralizaron en nuestro laboratorio, donde se descongelaron y se les extrajo y cuantificó el ARN total mediante el reactivo TRIzol™ (Invitrogen) y el NanoDrop® ND-1000 (Thermo Fisher Scientific) (apartado 3.2.2.3). Las muestras de ARN total obtenidas se retrotranscribieron utilizando el kit High Capacity RT (Applied Biosystems) como se detalló en el apartado 3.2.2.4., obteniendo 400 ng de ADNc en 20 µL.

3.3.1.5. Cuantificación de la expresión génica por reacción en cadena de la polimerasa a tiempo real (*real-time* PCR)

Se cuantificó la expresión génica relativa de los siguientes genes que acumulan VRs en pacientes con EM: *CYP24A1* (*cytochrome P450 family 24 subfamily A member 1*), *RGS1* (*regulator of G protein signaling 1*), *FCRL1* (*Fc receptor like 1*) y *TRAF3* (*TNF receptor associated factor 3*). Como control endógeno, se utilizó el gen *GAPDH* (*glyceraldehyde-3-phosphate dehydrogenase*). La Tabla 9 muestra los ensayos TaqMan™ utilizados para cada gen (Applied Biosystems).

Gen	Ensayo TaqMan™
<i>CYP24A1</i>	Hs00167999_m1
<i>RGS1</i>	Hs01023772_m1
<i>FCRL1</i>	Hs00364705_m1
<i>TRAF3</i>	Hs00936781_m1
<i>GAPDH</i>	Hs99999905_m1

Tabla 9. Listado de los ensayos TaqMan™ utilizados en la cuantificación relativa de la expresión génica de los genes que acumulan VRs en pacientes con EM.

Las reacciones de amplificación y el análisis de los datos crudos obtenidos se realizaron tal y como se detalló en el apartado 3.2.2.5. En este caso, a cada valor de Ct del gen a cuantificar, se le sustrajo el valor de Ct del gen *GAPDH* (gen endógeno) para normalizar los datos crudos. Los valores obtenidos se analizaron siguiendo el método $2^{-\Delta\Delta CT}$ (286) y se expresaron como *fold-change* de la expresión génica en pacientes con EM portadores de al menos una VR para un gen en concreto en comparación con aquéllos sin VRs en el mismo gen.

3.3.1.6. Análisis estadístico de los resultados de la *real-time* PCR

El análisis estadístico se realizó mediante el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Para cada uno de los cuatro genes, se aplicó el test de Mann-Whitney para elucidar si existen diferencias significativas en los niveles de ARNm del gen en cuestión entre pacientes con EM con al menos una VR de dicho gen y aquéllos sin ninguna.

3.3.2. Evaluación del cambio en la expresión de *RGS1* en respuesta al tratamiento con interferón- β en pacientes con y sin variantes raras

3.3.2.1. Individuos

Se incluyeron un total de 6 pacientes con EM (formas de inicio con brote) de la cohorte utilizada en la resecuenciación de ADN. De ellos, 3 portaban al menos una VR del gen *RGS1*, mientras que los otros 3 carecían de VRs de dicho gen. Ninguno de los pacientes había recibido tratamiento de la EM antes de la extracción de la muestra. Uno de los pacientes sin VRs procedió de Girona, mientras que las demás correspondieron a pacientes que se visitan regularmente en el Cemcat. Las CMSP del paciente de Girona se enviaron a nuestro instituto de investigación (Institut de Recerca Vall d'Hebron), donde se realizaron los cultivos pertinentes. La información de los pacientes incluidos se detalla en la Tabla Suplementaria 7 (Anexo).

3.3.2.2. Aislamiento de las CMSP

Las CMSP de los 6 pacientes con EM se aislaron mediante la técnica de centrifugación por gradiente de densidad Ficoll-Isopaque (Gibco BRL) y se conservaron en nitrógeno líquido tal y como se detalló en el apartado 3.2.2.2. hasta su utilización.

3.3.2.3. Inducción de la expresión génica de *RGS1* en respuesta a la estimulación de CMSP con interferón- β

Para evaluar el cambio en la expresión génica de *RGS1* en respuesta a la estimulación con interferón- β (IFN- β), las CMSP de los pacientes con EM incluidos en el estudio se descongelaron y contaron del mismo modo que se explicó previamente (apartado 3.2.5.2). Se sembraron 1×10^6 células/pocillo en placas de 24 (Falcon, Beckton Dickinson, New Jersey, USA) a una concentración de 1×10^6 células/mL en medio completo y se incubaron toda la noche a 37°C en una atmósfera con 5% de CO₂. Al día siguiente, se trataron sólo los pocillos pertinentes con 100 UI/mL de IFN- β 1b (Betaferon®) y se volvieron a introducir en el incubador. Las células se recogieron 4 horas después de recibir la estimulación con IFN- β y se lavaron con PBS a 4°C.

3.3.2.4. Extracción y cuantificación del ARN total de CMSP y síntesis de ADNc

El ARN total de las muestras de CMSP recogidas a las 4 horas post-estimulación se aisló utilizando las columnas QIAshredder® y el kit comercial RNeasy® Mini Kit (Qiagen), ya descrito en el apartado 3.2.7.4. Las muestras de ARN total obtenidas se retrotranscribieron utilizando el kit High Capacity RT (Applied Biosystems) (apartado 3.2.2.4.). Como resultado, se obtuvo el ADNc a una concentración de 20 ng/ μ L en un volumen total de 20 μ L.

3.3.2.6. Cuantificación de la expresión génica por reacción en cadena de la polimerasa a tiempo real (*real-time* PCR)

Se cuantificó la expresión génica relativa del gen *RGS1* (*regulator of G protein signaling 1*) tanto en las muestras sin tratar como en las tratadas con IFN- β mediante *real-time* PCR. La expresión se determinó en base a la cuantificación de la expresión de *GAPDH*. Los ensayos TaqMan™ utilizados fueron Hs01023772_m1 (*RGS1*) y Hs02786624_g1 (*GAPDH*).

En este caso, cada reacción contenía 2 μ L de muestra de ADNc (10 ng/ μ L), 10 μ L de TaqMan Universal PCR Master Mix, No AmpErase®, UNG (Applied Biosystems), 7 μ L de agua libre de RNasa y 1 μ L del ensayo TaqMan™ correspondiente. Las reacciones de amplificación y el análisis de los datos crudos obtenidos se realizaron tal y como se detalló en el apartado 3.2.2.5. Los resultados se analizaron siguiendo el método $2^{-\Delta\Delta CT}$ (286) y se expresaron como el cambio en la expresión (*fold-change*) génica de *RGS1* tras la estimulación con IFN- β , tanto para los pacientes con EM con alguna VR de *RGS1* como para aquéllos sin VRs en este gen.

3.3.2.7. Análisis estadístico de los resultados

El análisis estadístico se realizó mediante el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. En los dos grupos de pacientes con EM bajo estudio, el cambio en la expresión génica de *RGS1* tras recibir el tratamiento de IFN- β se analizó aplicando un T test pareado.

3.3.3. Inmunofenotipado de las CMSP que expresan *RGS1*

3.3.3.1. Individuos

Se incluyeron 11 pacientes con EM (formas de inicio con brote), de los cuales 4 portaban al menos 2 VRs de *RGS1* y 7 no tenían ninguna VR de dicho gen (Tabla Suplementaria 8, Anexo). Todos los pacientes formaban parte de la cohorte de resecuenciación de ADN y se trata de pacientes que se visitan regularmente en el Cemcat. Ninguno de ellos había recibido tratamiento de la EM antes de la extracción de sangre.

3.3.3.2. Obtención y mantenimiento de las CMSP

Las CMSP de los 11 pacientes con EM incluidos en el estudio se aislaron y congelaron como se detalló en el apartado 3.2.2.2. hasta su utilización, momento en el que se enviaron en nieve carbónica al Servicio de Inmunología del Hospital Ramón y Cajal (Madrid) para determinar su inmunofenotipo. Allí, cada día experimental, se procesaron 4 muestras con representación proporcional de los pacientes con variantes raras y de los pacientes sin ellas. Las CMSP de los 11 pacientes se descongelaron en el baño a 37°C e inmediatamente se añadió de forma progresiva medio RPMI completo frío hasta su descongelación total. Las CMSP se traspasaron a 10 mL de medio completo frío para lavar las células 3 veces mediante centrifugación. Se aspiró el sobrenadante, se resuspendieron las células en 1 mL de medio completo y se contaron utilizando la cámara de Neubauer. Se añadieron 1×10^6 células a cada uno de los tubos de citómetro de polipropileno, repartiendo cada muestra de la siguiente forma:

- Tubo 1 control
- Tubo 1 problema
- Tubo 2 control
- Tubo 2 problema

3.3.3.3. Anticuerpos utilizados

Para analizar la proporción de células de distintas subpoblaciones de CMSP que expresan RGS1 se realizaron marcajes de superficie e intracelulares (en el caso de RGS1) utilizando los anticuerpos detallados en la Tabla 10.

Anticuerpo	Fluorocromo	Isotipo	Clon	Marcaje	Casa comercial
Anti-CD3 de humano	PerCP	IgG1, κ de ratón	SK7	Superficie	BD Biosciences
Anti-CD8 de humano	APC-H7	IgG1, κ de ratón	SK1	Superficie	BD Biosciences
Anti-CD14 de humano	FITC	IgG2b, κ de ratón	M ϕ P9	Superficie	BD Biosciences
Anti-CD19 de humano	PE-Cy7	IgG1, κ de ratón	SJ25C1	Superficie	BD Biosciences
Anti-CD25 de humano	APC	IgG1, κ de ratón	2A3	Superficie	BD Biosciences
Anti-CD56 de humano	APC	IgG2b, κ de ratón	NCAM16.2	Superficie	BD Biosciences
Anti-CD45 de humano	V500	IgG1, κ de ratón	HI30	Superficie	BD Biosciences
Anti-CD127 de humano	BV421	IgG1, κ de ratón	HIL-7R-M21	Superficie	BD Biosciences
Anti-RGS1 de humano	-	IgG de conejo policlonal	-	Intracelular	Novus Biologicals
Anti-IgG de conejo	PE	IgG policlonal de burro	-	Intracelular	Biolegend
Control IgG monoclonal de conejo	-	IgG de conejo	60024B	Control intracelular	R&D Systems

Tabla 10. Anticuerpos utilizados para analizar la expresión de RGS1 en la cara interna de la membrana plasmática de las poblaciones de CMSP seleccionadas. El volumen utilizado fue el recomendado por el fabricante en todos los casos.

3.3.3.4. Diseño de los paneles de citometría

Se diseñaron los dos paneles especificados en la Tabla 11. Para cada población, se estudió el porcentaje de células que expresaba RGS1 en la cara interna de su membrana plasmática.

Población		Marcadores				
PANEL 1	Monocitos	CD45 ⁺	CD14 ⁺	CD3 ⁻		
	Linfocitos B	CD45 ⁺	CD19 ⁺	CD3 ⁻		
	Células NK	CD45 ⁺	CD56 ⁺	CD3 ⁻		
	Células NKT	CD45 ⁺	CD56 ⁺	CD3 ⁺		
PANEL 2	Linfocitos T	CD45 ⁺	CD3 ⁺			
	Linfocitos T CD4 ⁺	CD45 ⁺	CD3 ⁺	CD8 ⁻		
	Linfocitos T CD8 ⁺	CD45 ⁺	CD3 ⁺	CD8 ⁺		
	Linfocitos Treg	CD45 ⁺	CD3 ⁺	CD8 ⁻	CD25 ⁺⁺	CD127 ⁻

Tabla 11. Poblaciones de CMSP estudiadas y marcadores utilizados para su identificación mediante citometría de flujo. En todas ellas, se analizó el porcentaje que expresaba RGS1.

3.3.3.5. Marcaje celular

El marcaje con los anticuerpos correspondientes se llevó a cabo en tubos de polipropileno. El marcaje intracelular se llevó a cabo para valorar la expresión intracelular de RGS1 por parte de las siguientes poblaciones: monocitos, linfocitos B, células NK, células NKT, linfocitos T CD3⁺, linfocitos T CD4⁺, linfocitos T CD8⁺ y Treg. Para ello, en primer lugar se realizó el marcaje de los antígenos de superficie y posteriormente el marcaje intracelular utilizando el kit de permeabilización BD Cytotfix/Cytoperm™ (BD Biosciences), siguiendo el siguiente protocolo:

1. Marcaje de los antígenos de superficie, siguiendo las recomendaciones del fabricante. Incubación 30' a 4°C en oscuridad.
2. Lavado con 3 mL de salino.
3. Bloqueo de los receptores de superficie Fc para evitar las uniones inespecíficas con 2µg/tubo de IgG humana purificada (Human BD Fc Block™, de BD Biosciences). Incubación 30' a 4°C en oscuridad.
4. Lavado con 3 mL de salino.
5. Fijación y permeabilización de las células con 200 µL del Buffer de Fijación (Cytotfix/Cytoperm). Incubación 30' a 4°C en oscuridad.
6. Lavado de las células con 3 mL del Buffer BD Perm/Wash™ (BD Biosciences) diluido 1:10.
7. Adición de 1µg/tubo de anticuerpo policlonal anti-RGS1 en los tubos problema (Tubos 1 y 2 problema) o de 1µg/tubo de IgG monoclonal de conejo (R&D Systems) en los tubos control (Tubos 1 y 2 control). Incubación 30' a 4°C en oscuridad.

8. Lavado de las células con 3 mL del Buffer BD Perm/Wash™ (BD Biosciences) diluido 1:10.
9. Adición de 31,25 µg/tubo de anticuerpo de burro anti-IgG de conejo marcado con ficoeritrina (Biolegend). Incubación 30' a 4°C en oscuridad.
10. Lavado de las células con 3 mL del Buffer BD Perm/Wash™ (BD Biosciences) diluido 1:10.

3.3.3.6. Adquisición y análisis de las muestras en el citómetro de flujo

Las células se adquirieron en el laboratorio de Inmunología del Hospital Universitario Ramón y Cajal utilizando el equipo de citometría de flujo FACSCanto II, de Becton Dickinson. Los valores medios de autofluorescencia se establecieron utilizando controles negativos de isotipo apropiados. Una vez adquiridas las muestras, se analizaron los datos obtenidos con el *software* FACSDiva V.8.0 (BD Biosciences). En cada panel, se llevó a cabo una estrategia de *gating* para identificar las poblaciones celulares bajo estudio. En todos los casos se realizó la siguiente estrategia de *gating*: en primer lugar se combinaron los parámetros FSC y SSC para definir la población mononuclear. En segundo lugar, se combinaron los parámetros FS en altura y área para seleccionar los singletes y excluir los posibles dupletes. Posteriormente, se seleccionó la población positiva de intensidad intermedia y alta para CD45. Por último se combinaron diferentes parámetros en función del panel bajo análisis y los resultados se expresaron como el porcentaje de células positivas para RGS1 dentro de la población de interés. La estrategia de *gating* para identificar cada población se detalla en el Anexo.

3.3.3.7. Análisis estadístico de los resultados de la citometría de RGS1

El análisis estadístico se llevó a cabo utilizando el paquete SPSS 17.0 (SPSS Inc., Chicago, EEUU) para MS-Windows. Se aplicó la U de Mann-Whitney para evaluar si existen diferencias significativas en el porcentaje de diferentes poblaciones de CMSP que expresan RGS1 entre pacientes con EM con y sin VRs en este gen.

3.4. Metodología relacionada con el objetivo 4

“Validar y caracterizar experimentalmente la delección hallada en el gen *FCRL1* de un paciente con EM e identificar nuevos individuos portadores de la misma”

3.4.1. Validación experimental de la delección identificada en *FCRL1*

3.4.1.1. Individuos

Se incluyeron 11 pacientes con EM, de los cuales uno (muestra 1295) era el paciente identificado como portador de una delección en heterocigosis en el gen *FCRL1* (exones 4 al 10, ambos incluidos) y 10 carecían de ella. Todos ellos se visitan regularmente en el Cemcat.

3.4.1.2. Obtención y cuantificación del ADN genómico

El ADN genómico de los 11 pacientes se extrajo a partir de sangre periférica y se cuantificó con el NanoDrop® ND-1000 (Thermo Fisher Scientific) como se detalló previamente (apartado 3.1.1.). El ADN genómico se diluyó en agua libre de RNasa a una concentración 5 ng/μL.

3.4.1.3. Determinación del número de copias de diferentes exones del gen *FCRL1* por PCR cuantitativa *dúplex*

Para confirmar el resultado obtenido *in silico*, se utilizaron cuatro ensayos de número de copias TaqMan™ prediseñados (*TaqMan™ Copy Number Assays*) (Applied Biosystems), mediante los cuales se evaluó si el paciente 1295 realmente sólo portaba una copia de los exones 4-10 del gen *FCRL1*. Cada ensayo consta de un par de *primers* específicos para un exón concreto y una sonda MGB con el fluoróforo FAM™. Se adquirieron ensayos que detectan los exones 4, 9 y 10, de los cuales el paciente 1295 sólo debería tener una copia, y el exón 3, del cual debería tener dos copias. La Tabla 12 muestra los ensayos utilizados.

Exón <i>FCRL1</i>	Ensayo TaqMan™
Exón 3	Hs00256453_cn
Exón 4	Hs01867260_cn
Exón 9	Hs00700948_cn
Exón 10	Hs02549152_cn

Tabla 12. Listado de los ensayos TaqMan™ utilizados para la detección de los exones 3, 4, 9 y 10 del gen *FCRL1* utilizando ADN genómico.

Materiales y métodos - Objetivo 4

Cada ensayo de número de copias TaqMan™ se ejecuta simultáneamente con un ensayo de referencia (*TaqMan™ Copy Number Reference Assay*) en una *real-time* PCR *dúplex*. El ensayo de referencia consta de un par de *primers* y una sonda TAMRA™ con el fluoróforo VIC® para detectar la secuencia de referencia en el ADN genómico. El ensayo de número de copias detecta la secuencia genómica de interés (en nuestro caso, un exón) y el ensayo de referencia detecta una secuencia que se sabe que existe en dos copias en un genoma diploide. En el presente trabajo, se utilizó RNase P como ensayo de referencia (Applied Biosystems).

La cantidad de ADN genómico purificado añadida en cada reacción fue de 10 ng (en 2 µL) en un volumen total de 10 µL, de los que 5 µL eran TaqMan Genotyping Master Mix (Applied Biosystems), 2 µL agua libre de RNasa, 0,5 µL el ensayo TaqMan™ de referencia RNase P y 0,5 µL el ensayo TaqMan™ de un exón concreto. Es decir, en cada pocillo se añadió la sonda de referencia y la de un determinado exón, de modo que para cada individuo a analizar se necesitaron 4 pocillos: exón 3, exón 4, exón 9 y exón 10. Las reacciones se realizaron en placas ópticas de 384 pocillos y se ejecutaron en el termociclador ABI PRISM 7900HT Sequence Detection (Applied Biosystems). El programa de tiempos y temperaturas de la *real-time* PCR fue el detallado en la Tabla 13.

	<i> Holding </i>	<i> Cycling </i>	
Temperatura (°C)	95°C	95°C	60°C
Tiempo	10 min	15 s	60 s
Número de ciclos	1	← 40 →	

Tabla 13. Programa de temperaturas y tiempos utilizado para la determinación del número de copias de diferentes exones del gen *FCRL1* por PCR cuantitativa *dúplex*.

Los datos se analizaron con el *software* CopyCaller™ (Applied Biosystems). La cantidad de copias de cada exón en cada muestra se determinó por cuantificación relativa utilizando el método comparativo C_T ($\Delta\Delta C_T$). Este método mide la diferencia de C_T (ΔC_T) entre la secuencia diana y la de referencia y, después, compara los valores ΔC_T entre las muestras que se sabe que tienen una copia y las que se sabe que tienen dos. El número de copias se calcula multiplicando por 2 la cantidad relativa.

3.4.2. Caracterización de la deleción identificada en el gen *FCRL1*

3.4.2.1. Muestra

Se utilizó el ADN genómico del paciente 1295, portador heterocigoto de la deleción bajo estudio. El ADN se extrajo a partir de sangre periférica siguiendo el protocolo estándar detallado en el anexo y se diluyó en agua libre de RNasa. Se utilizaron 600 ng de ADN.

3.4.2.2. Obtención de la secuencia del fragmento delecionado del gen *FCLR1*

Debido a que la aproximación bioinformática no permite conocer con precisión ni el tamaño concreto de la deleción identificada ni los extremos de misma, se procedió a su delimitación experimental. Para ello, se amplificó la región con diferentes combinaciones de oligonucleótidos hasta obtener una banda de ~1000 pb, que fue secuenciada utilizando el secuenciador 3130xl Genetic Analyzer (Applied Biosystems). Las parejas de oligonucleótidos flanqueantes de la deleción diseñadas fueron las indicadas en la Tabla 14.

Pareja	Primer forward	Primer reverse
1	CAGGCCCTGTCTTAAGCATT	GTCAAAGACCCAGCGTTCAC
2	ACACCAAGAAGCACAAAAGGT	CAGGTGATGGAGGGAGACAG
3	CACATAGCAATGGCAGGAAA	GCGTCCAAAAGTCTTGAGGAG
4	GGCAGAGAAAAGCACTTGGAC	GCAGACAACAGCATTCTCA

Tabla 14. Parejas de oligonucleótidos diseñadas para amplificar la región que contiene la deleción identificada en el gen *FCRL1*. Todas ellas flanquean la deleción bajo estudio.

Finalmente, la pareja que se utilizó fue la número 3.

3.4.3. Screening de la deleción de *FCRL1* en una cohorte independiente

3.4.3.1. Individuos

Para tratar de identificar más individuos portadores de la deleción identificada en el gen *FCRL1*, se incluyeron 256 pacientes con EM que se visitan regularmente en el Cemcaty 269 controles sanos procedentes de la misma región. La Tabla 15 describe las características demográficas y

clínicas basales de los individuos incluidos. También se incluyó al paciente 1295, portador de la delección, como control positivo.

Características	Pacientes con EM	Controles sanos
n	256	269
Mujeres/Hombres (%M)	164/92 (64,1%)	135/134 (50,2%)
Edad, años	45,5 ± 13,5	34,1 ± 12,8
Forma clínica (%RR/%SP/%PP)	35,9/8,2/55,5	-

Tabla 15. Características demográficas y clínicas basales de los pacientes con EM y de los controles sanos incluidos para tratar de identificar más portadores de la delección identificada en el gen *FCRL1*. La edad se expresa como media ± desviación estándar. *M*: mujeres. *RR*: esclerosis múltiple remitente recurrente. *SP*: esclerosis múltiple secundariamente progresiva. *PP*: esclerosis múltiple primariamente progresiva.

3.4.3.2. Diseño del ensayo TaqMan™

Una vez definidos los bordes de la delección, se diseñó un ensayo TaqMan™ que la detectaba específicamente y se hizo un *screening* de las muestras de los 524 individuos descritos en el apartado anterior para analizarla en un número más grande de pacientes con EM y controles sanos y así estudiar su asociación con la enfermedad. La estrategia seguida fue diseñar un ensayo TaqMan™ constituido por dos *primers* que hibridasen en la zona que flanquea la delección y una sonda que hibridase entre los dos *primers* y también en la zona flanqueante de la delección. El ensayo TaqMan™ diseñado fue el siguiente:

Primer forward: GCTGGGACTAAATTTGGTGTGC

Primer reverse: ACTCCATCTAGAGATCCCAATGTTT

Sonda TaqMan™: CGGAGGCTAATGTGCTT

Teniendo en cuenta que se trata de una delección de más de 8 kb, no es posible amplificar un fragmento tan grande. Por ello, cuando se produjo amplificación significó que el individuo en cuestión era portador de la delección. Por el contrario, cuando el fragmento delecionado estaba intacto, no se produjo amplificación.

Se realizó una PCR con las siguientes condiciones:

	Preamplificación	Desnaturalización	Hibridación	Elongación
Temperatura (°C)	50	95	95	60
Tiempo	2 min	5 min	15 s	1 min
Número de ciclos	1	←	40	→

Tabla 16. Condiciones de la PCR diseñada para detectar individuos portadores de la delección identificada en el gen *FCRL1* en un paciente con EM.

4. RESULTADOS

4. Resultados

4.1. Resultados relacionados con el objetivo 1

"Identificar variantes genéticas comunes, raras y estructurales asociadas con el riesgo de desarrollar EM en 14 genes candidatos mediante resecuenciación dirigida de ADN"

La primera fase se desarrolló con el fin de detectar variantes genéticas asociadas con el riesgo de desarrollar EM en 14 genes de riesgo (*FCLR1*, *RGS1*, *TIMMDC1*, *HHEX*, *CXCR5*, *LTBR*, *TSMF*, *GALC*, *TRAF3*, *STAT3*, *TNFSF14*, *IFI30*, *CD40* y *CYP24A1*), seleccionados según los criterios detallados en el apartado 3.1.2. de Materiales y métodos. Para ello, se realizó la resecuenciación de sus regiones codificantes y reguladoras. Tras el análisis de los datos crudos, los resultados se dividieron en tres categorías: i) variantes comunes, ii) variantes raras y iii) variantes estructurales. A continuación, se detallan los resultados obtenidos.

4.1.1. Cobertura de los datos resecuenciados

En la resecuenciación de ADN, la cobertura describe el número medio de lecturas que recibe cada una de las bases, es decir, el número de veces que ha sido resecuenciado cada nucleótido. Uno de los principales objetivos del proceso de resecuenciación es conseguir un alto número de lecturas de cada base, ya que cuantas más veces sea resecuenciada una base, más fiable es el resultado obtenido. En el presente estudio, la cobertura media tanto de las muestras como de los fragmentos resecuenciados fue de 85,64× (Figura 22A y B, respectivamente), la cual supera el umbral óptimo para poder llevar a cabo estudios de identificación de variantes genéticas comunes y raras.

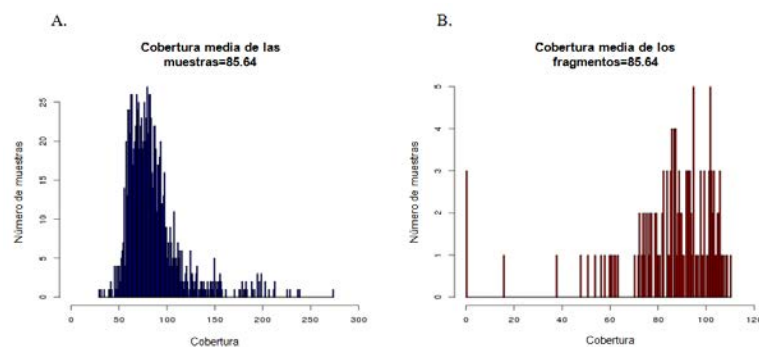


Figura 22. Cobertura media de (A) las muestras de ADN incluidas en la resecuenciación de ADN y de (B) los fragmentos resecuenciados.

Sólo 3 de los 141 fragmentos de ADN diana mostraron una cobertura media inferior a 5× en diferentes muestras y se eliminaron del análisis antes de realizar el llamado de variantes. Dichos 3 fragmentos con baja cobertura fueron los siguientes: chr14:88425798-88426157 (gen *GALC*), chr14:103275246-103275566 (gen *TRAF3*) y chr19:18289346-18289503 (gen *IFI30*).

4.1.2. Análisis de la calidad de las muestras resecuenciadas

El control de calidad previo al análisis de los resultados de la resecuenciación es necesario para obtener conclusiones válidas. En este apartado se expone el control de calidad aplicado a cada muestra para evaluar si pertenecen a poblaciones diferentes o si existen dos o más muestras demasiado similares entre ellas, hechos que supondrían la eliminación de dichas muestras para realizar el análisis.

4.1.2.1. Número de *singletons* por muestra

Un *tag* SNP es un polimorfismo representativo en una región del genoma con un alto LD formada por un grupo de polimorfismos constituyentes de un haplotipo. Un *singletone* es un *tag* SNP que sólo es representativo de él mismo y no de una región genómica, de modo que no está en alto LD con ningún SNP de esa región. Se evaluó el número de *singletones* por muestra para identificar muestras portadoras de un elevado número de los mismos, lo que podría ser indicativo de su pertenencia a otra población. Como se observa en la Figura 23, ninguna de las muestras mostró un exceso de *singletones* y sólo 4 de ellas portaban 5 *singletones*, cifra establecida como umbral a partir del cual las muestras se descartarían del análisis.



Figura 23. Número de *singletones* presentes en cada muestra incluida en la cohorte de resecuenciación. El eje de abscisas contiene todas las muestras incluidas en el estudio, mientras que el eje de ordenadas indica el número de *singletones*.

4.1.2.2. Análisis de componentes principales (ACP)

Sólo una de las muestras se identificó como *outlier* sobre el primer componente del ACP, mientras que 4 muestras resultaron *outliers* sobre el quinto componente (datos no mostrados). Todas ellas se eliminaron del análisis.

4.1.2.3. Número de diferencias genóticas o similitud genotípica

Se contó el número de diferencias de genotipo observadas entre cada posible par de muestras para todas las posiciones del genoma sometidas al llamado de variantes (Figura 24A-B). Un total de 42 parejas mostraron un número de diferencias menor a 55 y, para ellas, se descartó la muestra que mostraba la cobertura más baja con el fin de evitar la inclusión de muestras repetidas o con vínculo familiar.

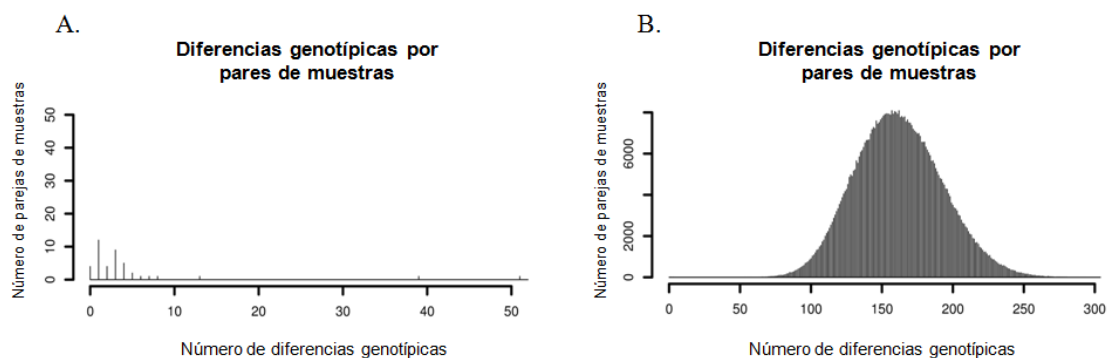


Figura 24. Número de diferencias genóticas existente entre cada posible pareja de muestras incluidas en la resecuenciación. (A) Representación de las 42 parejas de muestras con menos de 55 diferencias genóticas. (B) Representación de las parejas de muestras con más de 55 diferencias genóticas y que, por tanto, fueron incluidas en el análisis.

Tras los controles de calidad descritos en los apartados anteriores, se mantuvieron un total de 1070 muestras para su posterior análisis, las cuales comprenden 524 casos con EM (formas de inicio con brote) procedentes de 5 centros de EM nacionales y 546 controles sanos obtenidos de las mismas regiones (las características clínicas y demográficas se describen en el apartado 3.1.1. de Materiales y métodos).

4.1.3. Variantes genéticas comunes asociadas con el riesgo de desarrollar EM

La asociación de las variantes comunes ($MAF \geq 1\%$) con el riesgo de desarrollar EM se realizó en base a la comparación de las frecuencias alélicas entre pacientes con EM y controles sanos. En este estudio, la asociación se evaluó mediante tres pruebas estadísticas diferentes: i) la utilización del test exacto de Fisher, ii) la aplicación de una regresión logística considerando un modelo aditivo y iii) la aplicación de un test de asociación asintótico, ampliamente utilizado para grandes tamaños de muestra. Cada uno de estos test da lugar a unos p-valores diferentes. A continuación, se presentan los resultados obtenidos tras aplicar dichos test de asociación.

4.1.3.1. Se identificaron 32 variantes comunes significativamente asociadas con el riesgo de desarrollar EM

En la Figura 25 se representan tres *Manhattan plots*, cada uno de ellos resultante de una de las pruebas estadísticas aplicadas. La interpretación de los mismos es común en los tres casos:

- En todos ellos aparecen dos líneas paralelas al eje de abscisas:
 - i) La línea inferior, de color negro, representa el umbral de significación considerando un $p\text{-valor}=0,05$.
 - ii) La línea superior, de color rojo, representa el umbral de significación corregido para comparaciones múltiples. En este caso, surge al dividir el p-valor (0,05) entre el número de bloques de LD (bloques haplotípicos) en los que se localizan las variantes comunes identificadas en la resecuenciación. Los bloques haplotípicos se calcularon como se detalló en el apartado 3.1.7.1. de Materiales y métodos. Se obtuvieron un total de **131 bloques independientes**.
- Cada punto del *plot* representa el p-valor obtenido para una variante común específica.
- Cuando una variante supera algún umbral de significación, se representa en color naranja. El alelo minoritario de dicha variante puede hallarse más representado en los casos con EM (triángulo) o en los controles sanos (diamante).

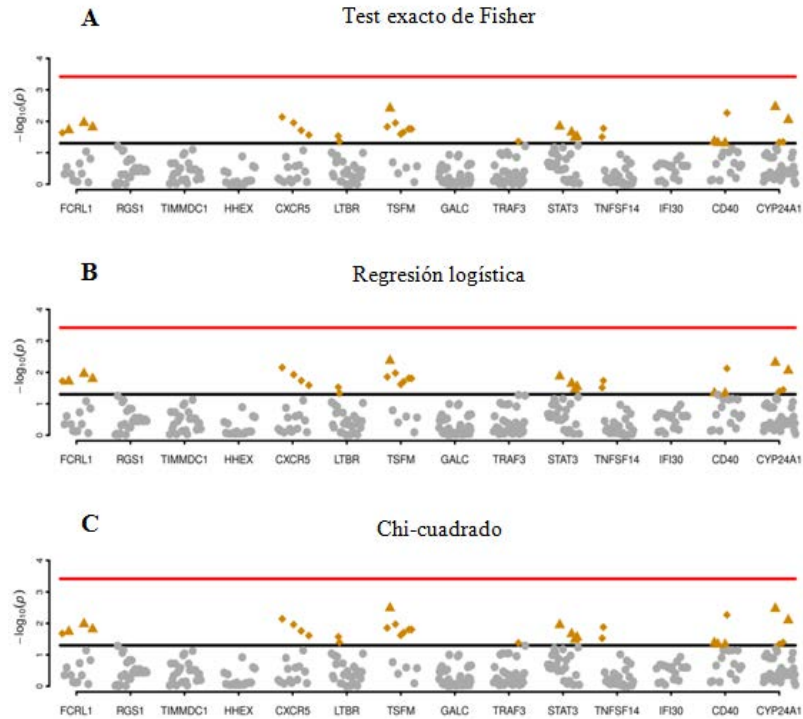


Figura 25. Manhattan plots representativos de las variantes genéticas comunes detectadas en la resecuenciación de ADN tras aplicar (A) el test exacto de Fisher, (B) una regresión logística considerando un modelo aditivo o (C) la prueba de chi-cuadrado.

Como se puede observar en la Figura 29, un total de 32 variantes comunes superaron el umbral de significación $p=0,05$ y, por tanto, se asociaron de forma significativa con la EM. Estas variantes se distribuyen en un total de 9 genes diferentes: *FCRL1*, *CXCR5*, *LTBR*, *TSFM*, *TRAF3*, *STAT3*, *TNFSF14*, *CD40* y *CYP24A1*. En la Tabla 17 se detalla la localización, frecuencias alélicas, p-valor y odds ratio (OR) de estas variantes. Gracias a la OR, se puede concluir si una variante se encuentra más representada en el grupo de casos con EM ($OR>1$) y, por tanto, aumenta el riesgo de desarrollar la enfermedad, o en el grupo de controles sanos ($OR<1$), constituyendo en este caso una variante protectora contra la EM. Se identificaron un total de 13 variantes que incrementan el riesgo de EM y 19 variantes protectoras para ella.

Resultados - Objetivo 1

Tabla 17. Variantes genéticas comunes diferencialmente distribuidas entre pacientes con EM y controles sanos ($p < 0,05$) identificadas mediante resecuenciación de ADN ($n=1070$).

Gen	SNP (Ref./Alt.)	MAF ¹ EM	MAF ¹ controles	P-valor exacto	P-valor logístico	P-valor asintótico	OR	L95	U95
<i>FCRL1</i>	rs1016849 (C/T)	0,17	0,21	0,023	0,019	0,021	0,77	0,62	0,96
<i>FCRL1</i>	rs3811024 (G/C/T)	0,24	0,20	0,019	0,019	0,018	1,28	1,04	1,57
<i>FCRL1</i>	rs3811023 (G/A)	0,25	0,20	0,011	0,011	0,010	1,30	1,06	1,60
<i>FCRL1</i>	rs34370597 (-/T) ²	0,25	0,21	0,015	0,016	0,015	1,28	1,05	1,57
<i>CXCR5</i>	rs10892307 (C/G)	0,13	0,18	0,0074	0,007	0,007	0,72	0,57	0,92
<i>CXCR5</i>	rs3922 (A/G)	0,40	0,45	0,011	0,012	0,011	0,80	0,67	0,95
<i>CXCR5</i>	rs581063 (T/C)	0,45	0,50	0,019	0,019	0,017	0,81	0,69	0,96
<i>CXCR5</i>	rs11217083 (G/A/C/T)	0,45	0,50	0,027	0,026	0,024	0,82	0,70	0,98
<i>LTBR</i>	rs3759323 (C/T)	0,13	0,17	0,030	0,030	0,027	0,76	0,60	0,97
<i>LTBR</i>	rs3759324 (T/C)	0,18	0,21	0,044	0,046	0,040	0,80	0,65	0,99
<i>TSFM</i>	rs11172335 (C/T)	0,22	0,26	0,015	0,014	0,014	0,78	0,64	0,95
<i>TSFM</i>	rs10747783 (T/C)	0,22	0,26	0,011	0,011	0,011	0,77	0,63	0,94
<i>TSFM</i>	rs34277814 (-/G) ²	0,25	0,20	0,004	0,004	0,003	1,35	1,11	1,66
<i>TSFM</i>	rs2014886 (C/T)	0,22	0,26	0,025	0,024	0,024	0,79	0,65	0,97
<i>TSFM</i>	rs1599932 (G/A)	0,22	0,26	0,022	0,020	0,020	0,79	0,65	0,96
<i>TSFM</i>	rs10783847 (G/A)	0,22	0,26	0,018	0,016	0,016	0,78	0,64	0,96

Gen	SNP (Ref./Alt.)	MAF ¹ EM	MAF ¹ controles	P-valor exacto	P-valor logístico	P-valor asintótico	OR	L95	U95
<i>TSM</i>	rs10783848 (G/A)	0,22	0,26	0,018	0,016	0,016	0,78	0,64	0,96
<i>TRAF3</i>	rs74082969 (A/G)	0,016	0,029	0,044	-	0,042	0,55	0,31	0,99
<i>STAT3</i>	rs141732716 (G/A)	0,02	0,009	0,014	0,013	0,011	2,53	1,20	5,32
<i>STAT3</i>	rs4796791 (T/C)	0,39	0,34	0,022	0,023	0,022	1,23	1,03	1,46
<i>STAT3</i>	rs3736164 (G/T)	0,29	0,25	0,033	0,035	0,033	1,23	1,02	1,49
<i>STAT3</i>	rs17884075 (C/-) ³	0,39	0,34	0,030	0,028	0,028	1,22	1,02	1,45
<i>TNFSF14</i>	rs141395077 (C/T)	0,013	0,027	0,031	0,031	0,030	0,50	0,26	0,94
<i>TNFSF14</i>	rs117732253 (G/A/C/T)	0,005	0,016	0,017	0,019	0,013	0,30	0,11	0,82
<i>CD40</i>	rs1883832 (C/T)	0,31	0,26	0,043	0,046	0,042	1,22	1,01	1,47
<i>CD40</i>	rs4810485 (G/A/T)	0,30	0,26	0,048	-	0,047	1,21	1,00	1,46
<i>CD40</i>	rs3746821 (G/A/T)	0,14	0,11	0,049	0,046	0,047	1,30	1,00	1,68
<i>CD40</i>	rs41282788 (G/C)	0,013	0,031	0,005	0,008	0,005	0,42	0,22	0,79
<i>CYP24A1</i>	rs6068816 (C/T)	0,14	0,10	0,003	0,005	0,003	1,48	1,14	1,93
<i>CYP24A1</i>	rs2259735 (C/T)	0,38	0,42	0,047	0,040	0,046	0,84	0,70	1,00
<i>CYP24A1</i>	rs2248137 (C/G)	0,36	0,40	0,045	0,036	0,041	0,83	0,70	0,99
<i>CYP24A1</i>	rs2762943 (G/T)	0,097	0,066	0,009	0,009	0,008	1,53	1,11	2,09

¹ MAF = frecuencia del alelo minoritario en la cohorte de resecuenciación de ADN. ² Se trata de una inserción. ³ Se trata de una delección. *Ref.:* alelo de referencia. *Alt.:* alelo alterado o minoritario. *OR=Odds Ratio. L95=Límite inferior del intervalo de confianza del 95% para la odds ratio. U95 = Límite superior del intervalo de confianza del 95% para la odds ratio.*

Sin embargo, no se encontró ninguna variante genética común con diferencias estadísticamente significativas en su frecuencia alélica entre casos con EM y controles sanos tras realizar la corrección ajustada al número de bloques haplotípicos independientes.

4.1.3.2. Se hallaron un total de 18 señales nuevas e independientes asociadas con el riesgo de desarrollar EM

De las 32 variantes comunes identificadas, 2 (rs4796791, en *STAT3*, y rs4810485, en *CD40*) estaban previamente descritas como asociadas con la EM como resultado de los dos últimos estudios a gran escala que se habían llevado a cabo por otros autores hasta la fecha de nuestra investigación (226,255) y que podían considerarse representativos de todos los publicados previamente. El primero de ellos (226) consistió en un GWAS realizado en 9.772 casos con EM y 17.376 controles sanos, mientras que el segundo (255) utilizó el *array* de genotipado personalizado *ImmunoChip* para testar 161.311 variantes autosómicas en 14.498 pacientes con EM y 24.091 controles sanos, con un posterior meta-análisis de sus resultados y datos previos de 14.802 casos con EM y 26.703 controles. Por tanto, se identificaron un total de 30 nuevos SNPs asociados con la EM.

A continuación, se buscó qué variantes comunes ya estaban asociadas con la EM en los dos estudios mencionados en la misma región de cada uno de los 30 nuevos SNPs identificados en la resecuenciación. Los resultados de esta búsqueda se resumen en la Tabla 18.

Gen	Variante	p-valor	OR	Estudio
<i>FCRL1</i>	rs2050568	$1,5 \times 10^{-10}$	1,08	Nat Genet. 2013
<i>CXCR5</i>	rs630923	$3,0 \times 10^{-7}$	1,12	Nature 2011
	rs9736016	$3,0 \times 10^{-15}$	1,10	Nat Genet. 2013
	rs523604	$6,2 \times 10^{-15}$	1,10	Nat Genet. 2013
<i>LTBR</i>	rs12296430	$7,2 \times 10^{-14}$	1,12	Nat Genet. 2013
<i>TSMF</i>	rs201202118	$9,0 \times 10^{-22}$	1,13	Nat Genet. 2013
<i>TRAF3</i>	rs12148050	$1,3 \times 10^{-10}$	1,10	Nat Genet. 2013
<i>STAT3</i>	rs9891119	$2,0 \times 10^{-10}$	1,10	Nature 2011
	rs4796791	$3,7 \times 10^{-20}$	1,12	Nat Genet. 2013
<i>TNFSF14</i>	rs1077667	$9,0 \times 10^{-14}$	1,16	Ambos*
<i>CD40</i>	rs2425752	$5,0 \times 10^{-10}$	1,11	Nature 2011
	rs4810485	$7,7 \times 10^{-16}$	1,11	Nat Genet. 2013
<i>CYP24A1</i>	rs2248359	$3,0 \times 10^{-11}$	1,12	Ambos*

Tabla 18. Variantes de los genes candidatos del presente estudio ya asociadas con la EM bien en los últimos estudios a gran escala desarrollados por otros autores hasta la fecha de nuestro estudio (226,255).

Posteriormente, se analizó el grado de LD en poblaciones europeas entre los 30 nuevos SNPs identificados y entre dichos 30 nuevos SNPs y los 13 ya descritos en la misma región en otros estudios (Tabla 18) para evaluar si el LD entre alguno de ellos es tan elevado ($r^2 > 0,8$) que constituyen la misma señal de asociación. El objetivo de este análisis fue determinar el número de señales de asociación nuevas identificadas en este estudio. Como resultado, los 30 nuevos SNPs identificados se redujeron a un total 18 nuevas señales de asociación (Tabla 19).

Variantes constituyentes de la señal				
Gen	Nº de señales nuevas	Señal 1	Señal 2	Señal 3
<i>FCRL1</i>	3	rs1016849	rs3811023 rs3811024	rs34370597
<i>CXCR5</i>	2	rs10892307	rs3922 rs581063 rs11217083	-
<i>LTBR</i>	1	rs3759323 rs3759324	-	-
<i>TSMF</i>	2	rs11172335 rs10747783*	rs34277814	-
<i>TRAF3</i>	1	rs74082969	-	-
<i>STAT3</i>	3	rs141732716	rs3736164	rs17884075
<i>TNFSF14</i>	2	141395077	rs117732253	-
<i>CD40</i>	2	rs3746821	rs41282788	-
<i>CYP24A1</i>	2	rs6068816	rs2762943	-
TOTAL	18	-	-	-

Tabla 19. Resumen de las señales de asociación con la EM identificadas como nuevas e independientes.

*Señal formada por los siguientes 6 polimorfismos: rs11172335, rs10747783, rs2014886, rs1599932, rs10783847 y rs10783848.

4.1.4. Variantes genéticas raras asociadas con el riesgo de desarrollar EM

Con el objetivo de abarcar un rango más amplio de variantes genéticas potencialmente asociadas con la EM, se analizaron las variantes raras detectadas en la resecuenciación. El umbral de frecuencia establecido para considerar una variante genética como rara fue $MAF=0,5\%$. Del total de 2.082 variantes polimórficas bialélicas detectadas en el análisis (1.903 SNPs y 179 indeles), el 76,3% se encontraron con una frecuencia del alelo menor $\leq 0,5\%$, de modo que se consideraron raras. Dada su baja frecuencia, no es posible realizar estudios estadísticos concluyentes con una sola variante rara; por tanto, tal y como se detalló en Materiales y Métodos, se evaluó si el número de muestras que portan al menos una variante rara para el gen o la región específica bajo análisis difiere entre los casos con EM y los controles sanos. Así, se determinó si existe enriquecimiento de variantes raras de un gen o región en alguno de los dos grupos. Para ello, se aplicaron once test de asociación diferentes específicos

Resultados - Objetivo 1

para variantes raras (ver apartado 3.1.8. de Materiales y Métodos), denominados en su conjunto *Collapsing methods*.

Los once tests se aplicaron varias veces, cada una de ellas clasificando los datos de una forma diferente siguiendo uno de los dos criterios detallados en el apartado 3.1.8. de Materiales y Métodos y resumidos a continuación:

1. En primer lugar, se analizaron los datos para evaluar si hay distribución diferencial de variantes raras entre pacientes con EM y controles sanos considerando los distintos tipos posibles de variantes. Se presentan los resultados de tres formas, considerando: i) todas las variantes, ii) sólo las variantes no sinónimas o iii) sólo las variantes posiblemente patogénicas.
2. En segundo lugar, investigamos si las variantes raras se encuentran distribuidas de forma diferencial entre casos con EM y controles sanos considerando: i) sólo las regiones codificantes resecuenciadas, ii) sólo las regiones reguladoras resecuenciadas, iii) sólo las regiones UTR o iv) sólo los promotores.

Los resultados obtenidos se representaron en diferentes *Manhattan plots*, cuya interpretación tiene aspectos comunes:

- En cada *plot*, se han representado dos líneas horizontales paralelas al eje de abscisas:
 - i) La línea inferior, de color negro, delimita el umbral de significación considerando un $p\text{-valor}=0,05$.
 - ii) La línea superior, de color rojo, representa el umbral de significación corregido para comparaciones múltiples. El valor delimitado por esta línea varía en función del criterio de clasificación aplicado, lo cual se detallará en cada apartado.
- Cada punto del *plot* representa uno de los once *collapsing methods* utilizados y su localización en el *plot* indica el $p\text{-valor}$ obtenido para el enriquecimiento de variantes raras al aplicar dicho *collapsing method*.
- Si un enriquecimiento de variantes raras supera el umbral de significación $p\text{-valor}=0,05$, se representa en color naranja. Si el enriquecimiento supera el umbral de significación corregido para comparaciones múltiples, se representa en color granate. Los puntos grises representan un $p\text{-valor}>0,05$.

- Cuando el enriquecimiento de variantes raras de un gen se observe en los casos con EM, se representa con un triángulo, mientras que si aparece en los controles sanos, con un diamante.

A continuación, se detallan los resultados obtenidos en las dos rondas de tests de asociación.

4.1.4.1. Enriquecimiento de variantes raras según el tipo de variante

Como muestra la Figura 26, al tener en cuenta todos los tipos de variantes, se identificó una acumulación significativa de variantes raras en el gen *RGS1* en pacientes con EM cuando se compararon con los controles sanos; en 8 de los 11 tests utilizados se obtuvo un p -valor $\leq 0,05$ y en 4 de ellos un p -valor significativo tras aplicar la corrección para comparaciones múltiples ($p=0,05/14$, siendo 14 el número de genes testados en este análisis). Además, también considerando todos los tipos de variantes, se observó un enriquecimiento de variantes raras del gen *HHEX* en controles sanos respecto a los pacientes con EM (2 tests: $p \leq 0,05$), lo que podría ser indicativo del papel protector contra la EM de dicho enriquecimiento.

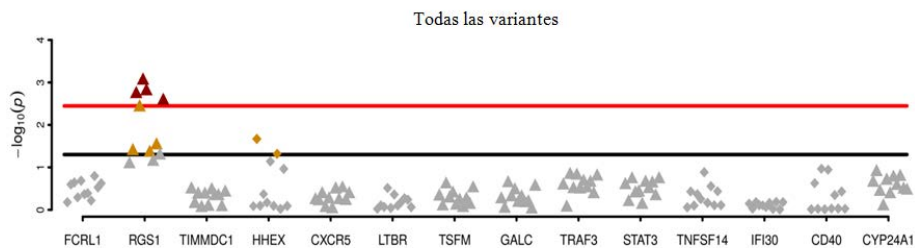


Figura 26. Representación de los p -valores obtenidos tras aplicar los 11 *collapsing methods* sobre cada uno de los genes teniendo en cuenta todos los tipos de variantes raras detectadas. Los puntos grises indican un p -valor $> 0,05$, los naranjas un p -valor $< 0,05$ pero superior al p -valor obtenido tras la corrección de comparaciones múltiples, y los granates un p -valor que supera la corrección de comparaciones múltiples. Los triángulos muestran un exceso de variantes raras en pacientes y los diamantes un exceso en controles.

Cuando se analizó la acumulación de variantes raras no sinónimas, ningún test determinó diferencias significativas entre pacientes con EM y controles sanos para ninguno de los genes ($p > 0,05$). Del mismo modo, la distribución de las variantes raras posiblemente patogénicas resultó equivalente entre los pacientes con EM y los controles sanos para los 14 genes incluidos en el estudio ($p > 0,05$) (datos no mostrados).

4.1.4.2. Enriquecimiento de variantes raras según el tipo de secuencia

Para explorar si existe alguna acumulación significativa de variantes raras específicamente en alguna región de cada gen bajo análisis, bien en pacientes con EM o bien en controles sanos, se clasificaron las variantes raras en función de si se hallan en regiones codificantes o reguladoras del genoma. Como muestra la Figura 27A, un test identificó un enriquecimiento de variantes raras situadas en regiones codificantes del gen *TIMMDC1* en pacientes con EM respecto a los controles sanos ($p < 0,05$). Igualmente, otro test determinó un enriquecimiento de variantes raras de regiones codificantes del gen *FCRL1* en controles sanos cuando se compararon con los pacientes con EM ($p < 0,05$). Respecto a las regiones reguladoras de los 14 genes bajo estudio, tal y como muestra la Figura 27B, un total de 5 tests determinaron la acumulación de variantes raras del gen *RGS1* en el grupo de pacientes con EM en comparación con el grupo de controles ($p < 0,05$), mientras que un total de 9 tests sugirieron un enriquecimiento de variantes raras del gen *CYP24A1* en pacientes con EM respecto a los controles sanos ($p < 0,05$). Ninguno de los hallazgos superó la corrección para comparaciones múltiples.

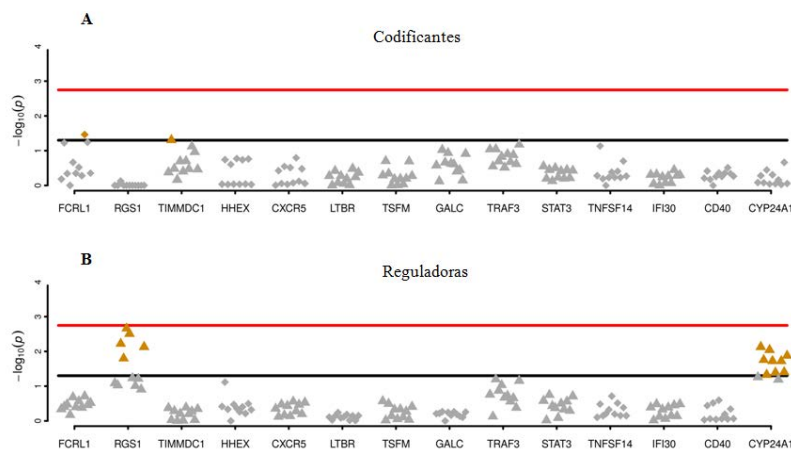


Figura 27. Enriquecimientos de variantes raras obtenidos tras aplicar los *collapsing methods* sobre cada gen clasificando las variantes como (A) codificantes o (B) reguladoras. Los puntos grises indican un p -valor $> 0,05$ (delimitado por la línea negra) y los naranjas un p -valor $< 0,05$ pero sin alcanzar la corrección de comparaciones múltiples (delimitada por la línea roja). Los triángulos muestran un exceso de variantes raras en casos con EM y los diamantes un exceso en controles.

A continuación, estudiamos el promotor y las regiones UTR de cada gen. Se compararon los pacientes con EM con controles sanos. Como muestra la Figura 28A, se observó una acumulación de variantes raras en el promotor del gen *FCRL1* en pacientes con EM, hallazgo determinado por 2 tests ($p < 0,05$). De igual modo, 6 tests sugirieron un exceso de variantes raras en el promotor de *CYP24A1* en pacientes con EM ($p < 0,05$). Cuando se analizaron las regiones

UTR (Figura 28B), 2 tests sugirieron una acumulación de variantes raras del gen *TRAF3*, mientras que todos los tests determinaron una acumulación en *CYP24A1*, en ambos casos en pacientes con EM ($p < 0,05$).

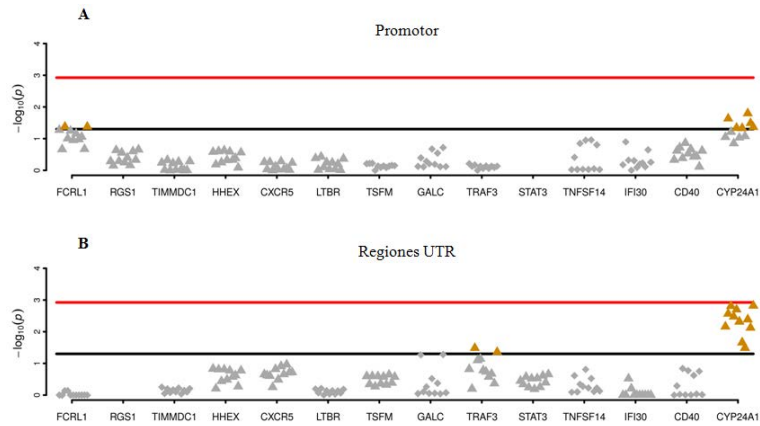


Figura 28. Enriquecimientos de variantes raras de cada gen localizadas (A) en el promotor o (B) en las regiones UTR. La interpretación es común a la de la Figura 31.

4.1.5. Variantes genéticas estructurales identificadas

4.1.5.1. Se identificó una deleción de 8,46 Kb en el gen *FCRL1* de un paciente con EM

Para evaluar si un determinado fragmento genómico se encuentra delecionado, se analizó la cobertura de la resecuenciación para dicho fragmento, es decir, el número de lecturas del fragmento que alinean con el genoma de referencia utilizado (GRCh37/hg19). Una disminución de la cobertura en cierta región del genoma significa que dicha región se halla delecionada. En el presente estudio, se detectó una bajada de la cobertura de resecuenciación de la región chr1:157764094-157772554, de 8,46 Kb, que incluye desde el exón 4 hasta el exón 10 del gen *FCRL1*, en un paciente con EM (Figura 29), lo que sugiere la deleción de la misma. Sin embargo, mediante esta metodología no es posible conocer con exactitud los límites superior e inferior de la deleción.

Resultados - Objetivo 1

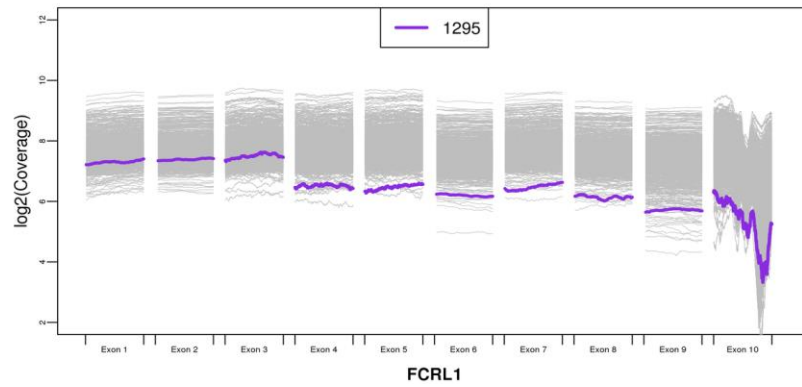


Figura 29. Representación de la cobertura de la resecuenciación para el gen FCRL1 del paciente con EM 1295. En cada exón, las líneas moradas se corresponden con la cobertura o el número de veces que se leyó cada posición en el paciente 1295, mientras que las grises son las del resto de individuos incluidos en el estudio.

4.2. Resultados relacionados con el objetivo 2

"Validar una selección de variantes genéticas comunes asociadas con la susceptibilidad de desarrollar EM en la resecuenciación de ADN en una cohorte independiente de pacientes y controles sanos y realizar estudios adicionales relacionados con los polimorfismos validados"

Una vez abordado el objetivo de identificar variantes genéticas comunes asociadas con la EM, se procedió a investigar si alguna de dichas variantes implica una variación funcional. Para ello, en primer lugar, se procedió a validar mediante genotipado en una cohorte independiente la asociación de las variantes más interesantes. Seguidamente, se desarrollaron distintos estudios diseñados en base a la función del gen asociado a cada variante validada. A continuación, se exponen los resultados obtenidos.

4.2.1. Validación de las variantes comunes seleccionadas mediante genotipado en una cohorte independiente

4.2.1.1. Variantes genéticas seleccionadas para su validación

Con el fin de validar los hallazgos resultantes del proceso de resecuenciación de ADN, se seleccionaron 9 de las 32 variantes comunes asociadas con la EM en la resecuenciación, elección realizada en base a los criterios detallados en el apartado 3.2.1.2. de Materiales y Métodos. Además, se seleccionaron un total de 9 variantes adicionales asociadas a la EM en el último estudio publicado por el IMSGC (258) con el objetivo de poder aplicar una regresión condicional cuando fuera conveniente y evaluar así la dependencia entre dos polimorfismos que se asocien a la EM de forma significativa. La Tabla 20 describe brevemente las 18 variantes seleccionadas.

La validación se llevó a cabo en una cohorte formada por 3.450 pacientes con EM procedentes de 8 centros nacionales y 1.688 controles sanos obtenidos de las mismas regiones (características clínicas y demográficas en el apartado 3.2.1.1. de Materiales y métodos).

Resultados - Objetivo 2

Tabla 20. Descripción y consecuencias funcionales de las variantes comunes asociadas a la EM seleccionadas para su validación mediante genotipado en una cohorte independiente.

Gen	Nombre oficial	SNP (Ref./Alt.)	Localización (hg19)	Procedencia	Localización (<i>Ensembl</i>)
<i>CXCR5</i>	C-X-C motif chemokine receptor 5	rs10892307 (C/G)	chr11:118754649	Resequenciación	UTR5'
		rs3922 (A/G)	chr11:118765600	Resequenciación	UTR3'
		rs6589706 (A/G)	chr11:118747813	IMSGC 2017	Región reguladora
<i>PHLDB1</i>	pleckstrin homology like domain family B member 1	rs34026809 (G/C)	chr11:118480695	IMSGC 2017	Intrónica
<i>BCL9L</i>	B cell CLL/lymphoma 9 like	rs149114341 (G/A)	chr11:118783424	IMSGC 2017	Intrónica
<i>LTBR</i>	lymphotoxin beta receptor	rs3759323 (C/T)	chr12:6485436	Resequenciación	Intrónica
		rs2364485 (C/A)	chr12:6514963	IMSGC 2017	<i>Downstream</i>
<i>TNFRSF1A</i>	TNF receptor superfamily member 1A	rs1800693 (T/C)	chr12:6440009	IMSGC 2017	Exón no codificante
<i>TSFM</i>	Ts translation elongation factor, mitochondrial	rs1599932 (G/A)	chr12:58177943	Resequenciación	Exón (cambio de sentido)
<i>EEF1AKMT3</i>	EEF1A lysine methyltransferase 3	rs701006 (A/G)	chr12:58106836	IMSGC 2017	Intrónica
<i>TRAF3</i>	TNF receptor associated factor 3	rs74082969 (A/G)	chr14:103375540	Resequenciación	UTR3'
		rs12588969 (C/G)	chr14:103230758	IMSGC 2017	Región reguladora
<i>STAT3</i>	signal transducer and activator of transcription 3	rs141732716 (G/A)	chr17:40469832	Resequenciación	Intrónica
		rs1026916 (A/G)	chr17:40529835	IMSGC 2017	Intrónica
<i>TNFSF14</i>	TNF superfamily member 14	rs117732253 (G/C/T)	chr19:6661631	Resequenciación	<i>Downstream</i>

Gen	Nombre oficial	SNP (Ref./Alt.)	Localización (hg19)	Procedencia	Consecuencia (<i>Ensembl</i>)
<i>CYP24A1</i>	cytochrome P450 family 24 subfamily A member 1	rs6068816 (C/T)	chr20:52781091	Resequenciación	Exón (sinónima)
		rs2762943 (G/T)	chr20:52790786	Resequenciación	<i>Upstream</i>
		rs2585447 (T/C)	chr20:52744437	IMSGC 2017	Intergénica

Se presentan todas las variantes comunes incluidas en la fase de genotipado. En negrita se muestran los 9 polimorfismos seleccionados a partir de los resultados obtenidos tras la resecuenciación de ADN. Los otros 9 polimorfismos descritos constituyen aquéllos asociados con la EM en el análisis a gran escala publicado en 2017 por el IMSGC, incluidos para futuros análisis de regresión condicional. La tabla también muestra el gen asociado a cada una de las variantes, su localización en el genoma humano y su consecuencia funcional determinada por la plataforma *Ensembl*. *Ref.*: alelo de referencia; *Alt.*: alelo alterado.

4.2.1.2. Variantes genéticas asociadas a la EM validadas

La Tabla 21 muestra los resultados del análisis estadístico de los datos derivados del genotipado para las 9 variantes seleccionadas de la cohorte de la resecuenciación de ADN. Cabe mencionar que todos los SNPs estaban en equilibrio de Hardy-Weinberg. Se validaron un total de tres variantes comunes, lo que significa que se confirma la asociación con la EM identificada en la resecuenciación. Estas tres variantes fueron las siguientes:

- i. **rs10892307** (*CXCR5* - *C-X-C motif chemokine receptor 5*; $p=1,37 \times 10^{-6}$; OR=0,77): se trata de un SNP situado en la región UTR5' del gen *CXCR5*. En el presente estudio, se encontró una mayor frecuencia de su alelo minoritario (G) en controles sanos que en pacientes con EM, mientras que el alelo de referencia o mayoritario (C) se halló de forma más frecuente en pacientes con EM que en controles. Por tanto, se consideró una variante protectora para la EM. El gen *CXCR5* codifica un receptor cuyo ligando es la quimiocina CXCL13 o BLC (*B lymphocyte chemoattractant*). Las células que expresan *CXCR5* en su superficie, como las células B, las células Tfh o las células Tfr, son capaces de migrar a los órganos linfoides secundarios gracias a la atracción hacia el ligando BLC, secretado por las células dendríticas foliculares y por las células estromales de dichos órganos.

Debido a que en esta región resultaron significativas otras dos variantes del estudio del IMSGC (258) [rs149114341 ($p=0,0002$; OR=0,65) y rs6589706 ($p=0,002$; OR=1,14)], se realizó una regresión logística condicional para evaluar el número de señales independientes. Como se observa en la Tabla 22, cuando se realiza la regresión condicionada al SNP rs10892307, la asociación del SNP rs149114341 se mantiene, pero se pierde la del SNP rs6589706, por lo que éste es dependiente de rs10892307. Sin embargo, al realizar la regresión condicionada al SNP rs6589706, la señal del SNP rs10892307 se mantiene, por lo que rs10892307 explica mejor la asociación de la señal con la EM que rs6589706 y, por ello, es el mejor candidato a ser la variante causal. Por último, al condicionar al SNP rs149114341, los SNPs rs10892307 y rs6589706 siguen resultando significativos, lo que confirma la independencia de rs149114341.

- ii. **rs1599932** (*TSM* - *Ts translation elongation factor, mitochondrial*; $p=0,001$; OR=0,85): este SNP es una variante exónica clasificada como mutación con cambio de sentido, de modo que provoca la aparición de un codón que codifica para un aminoácido diferente. En el genotipado, su alelo minoritario (A) se identificó de forma más frecuente en controles sanos que en pacientes con EM, en quienes era más frecuente el alelo de referencia (G), por lo que se trata de una variante protectora para la enfermedad. El gen *TSM* codifica una enzima que cataliza el intercambio de

nucleótidos de guanina en el factor de elongación Tu durante la elongación en la traducción de proteínas mitocondriales.

Dado que las otra variante situada en esta región e incluida en el genotipado no resultó significativa, no fue necesario realizar regresión condicional. Sin embargo, no se profundizó más en el estudio de este polimorfismo debido a que se descubrió que pertenecía a un gran bloque de LD ya asociado con la EM (287), de modo que supondría redundancia.

- iii. **rs2762943** (*CYP24A1* - *cytochrome P450 family 24 subfamily A member 1*; $p=0,047$; $OR=1,16$): se trata de un SNP localizado 261 pb hacia arriba del sitio de inicio de la transcripción del gen *CYP24A1*. Su alelo minoritario (T) resultó más frecuente en pacientes con EM que en controles sanos en la cohorte de genotipado, de forma que se consideró como una variante de riesgo, mientras que el alelo mayoritario (G) apareció de forma más frecuente en controles. El gen *CYP24A1* codifica una enzima con actividad 24-hidroxilasa capaz de degradar tanto la forma biológicamente activa de la vitamina D (1,25(OH)₂-D o calcitriol), como la forma precursora (25(OH)-D o calcidiol), en ambos casos mediante la incorporación de un grupo OH en la posición C-24 de la cadena lateral de las mismas. Así, esta enzima, inducible por el propio calcitriol, controla el balance entre la síntesis y el catabolismo de la vitamina D.

En este caso, también fue la única variante de la región que resultó significativa, luego no fue necesario realizar regresión condicional.

Tabla 21. Resultados del análisis estadístico de los datos del genotipado para las 9 variantes comunes seleccionadas a partir de los resultados de la rescuenciación de ADN.

Gen	SNP (Ref./Alt.)	MAF		Test exacto de Fisher		
		EM (%)	CS (%)	OR (DE)	IC (95%)	p-valor
<i>CXCR5</i>	rs10892307 (C/G)	15,8	19,6	0,77 (0,05)	0,69-0,85	1,37×10⁻⁶
	rs3922 (A/G)	41,3	42,0	0,97 (0,04)	0,90-1,06	0,547
<i>LTBR</i>	rs3759323 (C/T)	15,2	15,4	0,98 (0,06)	0,88 - 1,10	0,792
<i>TSM</i>	rs1599932 (G/A)	20,9	23,6	0,85 (0,05)	0,77 - 0,94	0,001
<i>TRAF3</i>	rs74082969 (A/G)	30,9	29,6	1,05 (0,12)	0,82- 1,33	0,713
<i>STAT3</i>	rs141732716 (G/A)	1,8	1,5	1,19 (0,17)	0,86 - 1,65	0,289
<i>TNFSF14</i>	rs117732253 (G/C/T)	1,0	1,2	0,80 (0,20)	0,54 - 1,18	0,261
<i>CYP24A1</i>	rs6068816 (C/T)	11,6	12,3	0,93 (0,06)	0,82 - 1,06	0,299
	rs2762943 (G/T)	9,1	7,9	1,16 (0,08)	1,00 - 1,35	0,047

En negrita se representan las variantes comunes validadas. La columna “MAF” indica la frecuencia del alelo minoritario en los casos con EM y en los controles sanos (CS) constituyentes de la cohorte de genotipado. Una odds ratio (OR)>1 informa de una mayor frecuencia del alelo minoritario (MA) en casos con EM, mientras que una OR<1 indica una mayor frecuencia del MA en controles. Ref.: alelo de referencia. Alt.: alelo alterado o minoritario. MAF: frecuencia del alelo minoritario. OR: Odds Ratio. DE: desviación estándar. IC: intervalo de confianza del 95% para el odds ratio.

Tabla 22. Regresión logística de las variantes de la región del SNP rs10892307 que resultaron diferencialmente distribuidas entre pacientes con EM y controles sanos en el genotipado.

SNP (Ref./Alt.)	Condicionada* a rs10892307			Condicionada* a rs6589706			Condicionada* a rs149114341		
	p-valor	OR (DE)	IC (95%)	p-valor	OR (DE)	IC (95%)	p-valor	OR (DE)	IC (95%)
rs10892307 (C/G)	-	-	-	0,0001	0,79 (0,06)	0,62-0,88	7,84×10 ⁻⁵	0,80 (0,05)	0,72-0,89
rs6589706 (G/A)	0,260	1,05 (0,05)	0,96-1,16	-	-	-	0,008	1,12 (0,04)	1,03-1,22
rs149114341 (G/A)	0,009	0,73 (0,12)	0,58-0,93	0,0008	0,68 (0,05)	0,60-0,78	-	-	-

La variante rs10892307 procede de la cohorte de resecuenciación, mientras que las variantes rs6589706 y rs149114341 proceden del último meta-análisis del IMMSGC (258). *Regresión logística condicional ajustada a un modelo aditivo para analizar la contribución dependiente o independiente de las diferentes variantes al riesgo de desarrollar EM. *Ref.: alelo de referencia. Alt.: alelo alterado o minoritario. OR: Odds Ratio. DE: desviación estándar. IC: intervalo de confianza del 95% para la odds ratio.*

4.2.2. Estudios relacionados con el polimorfismo rs10892307 (CXCR5)

En base a los resultados obtenidos tras analizar los datos del genotipado, se procedió a profundizar más en la posible asociación del SNP rs10892307 con cambios en el gen *CXCR5*. Para ello, se realizaron los estudios detallados en los siguientes apartados.

4.2.2.1. Se identificaron tres polimorfismos en alto LD con la variante rs10892307, todos ellos con actividad *enhancer*

Con el objetivo de evaluar si el polimorfismo del gen *CXCR5* validado, rs10892307, estaba en alto LD con algún otro SNP en población europea, se analizó el parámetro r^2 que mantenía este polimorfismo con el resto de los identificados en el genoma humano. Se seleccionaron aquéllos con un valor de $r^2 > 0,5$, lo que implica que en más del 50% de los casos esos polimorfismos aparecen juntos en el genoma de un individuo. Como resultado, un total de tres SNPs superaron este filtro y se consideraron en alto LD con rs10892307 (Tabla 23).

A continuación, se investigó para cada una de las variantes seleccionadas, así como para la ya validada, la puntuación que le asignaba la base de datos Regulome DB. Los resultados obtenidos también se muestran en la Tabla 23.

rs10892307				
SNP (Ref./Alt.)	r^2	D'	Localización (hg19)	Regulome DB score
rs10892307 (C/G)	1,0	1,0	chr11:118754648	4
rs3176905 (T/C)	0,99	1,0	chr11:118755630	5
rs11602393 (T/A)	0,94	0,97	chr11:118750967	6
rs55756957 (A/G)	0,54	0,94	chr11:118747576	2b

Tabla 23. Esta tabla muestra los tres SNPs seleccionados, además del validado, tras aplicar el filtro de LD $r^2 > 0,5$. Se muestran los valores de r^2 y D', así como su localización genómica y su puntuación según la base de datos Regulome DB.

Como se evidencia en la Tabla 23, la puntuación más baja (2b) la presenta la variante rs55756957. Esta puntuación denota que la presencia del polimorfismo probablemente afecte a la unión de factores de transcripción. Concretamente, podría alterar la unión de la proteína PPARalpha:RXRalpha, dado que esta variante se localiza en su secuencia consenso y, además,

experimentalmente siempre se ha encontrado a su alelo mayoritario (A) el sitio de unión de esta proteína, por lo que parece imprescindible (Figura 30).

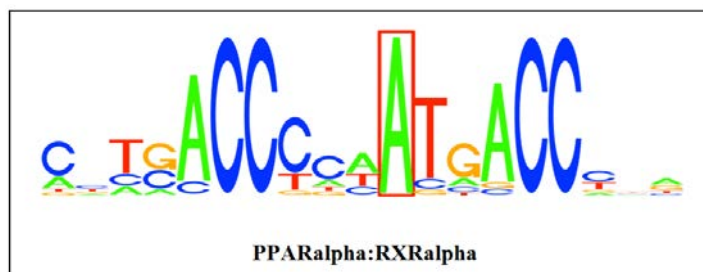


Figura 30. Resultados del análisis del SNP rs55756957 mediante Regulome DB. Se representa la PWM (*position-weight matrix*) del motivo de unión del heterodímero PPARalpha:RXRalpha identificado mediante la técnica de *DNA footprinting*. El nucleótido que aparece dentro del recuadro rojo es la localización del polimorfismo rs55756957. La altura relativa de cada nucleótido es proporcional a la frecuencia en la que se ha encontrado en esa posición experimentalmente.

La puntuación asignada al resto de variantes (4, 5 y 6) también demuestra que se ha observado unión de proteínas reguladoras en la región del SNP, de modo que también se postulan como *enhancers* o potenciadores de los niveles de transcripción génica, aunque con menor evidencia experimental. La Figura 31 representa de forma esquemática la localización relativa de las cuatro variantes y su situación en regiones reguladoras.

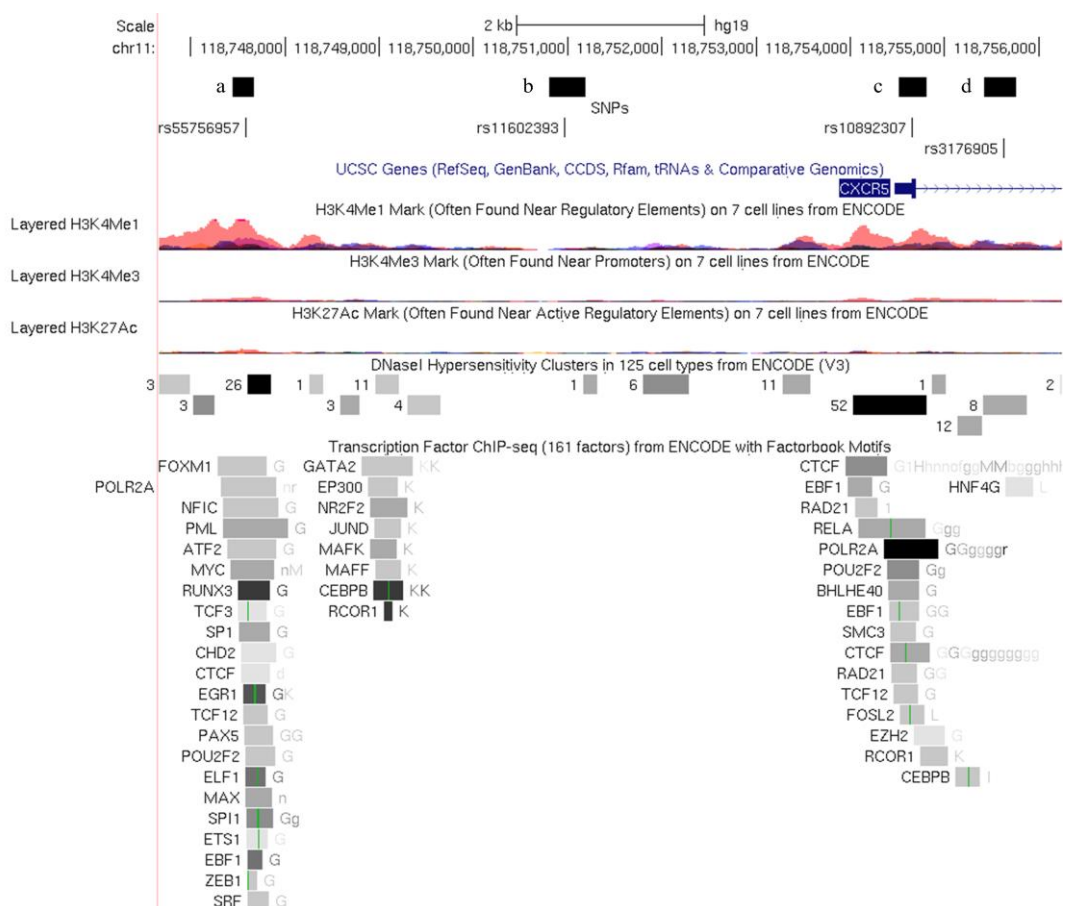


Figura 31. Representación esquemática que muestra la actividad *enhancer* de la variante rs10892307 y de las tres variantes en LD con ella (rs55756957, rs11602393 y rs3176905), todas ellas localizadas en regiones potencialmente reguladoras (cajas negras a-d), tal y como se indica en ENCODE. Los tracks *Layered H3K4Me1*, *Layered H3K4Me3* y *Layered H3K27Ac* muestran sitios de modificación de histonas, cuyo color varía en función del tipo celular donde se identificaron. El track *DNaseI Hypersensitivity Clusters* muestra zonas donde la cromatina es hipersensible a la enzima DNasa I. El track *Transcription Factor ChIP-Seq* representa las regiones del ADN donde se unen los factores de transcripción listados, identificadas mediante inmunoprecipitación de cromatina seguida de secuenciación del ADN precipitado.

4.2.2.2. La expresión génica basal de *CXCR5* en CMSP es similar entre pacientes con EM portadores y no portadores del alelo minoritario de rs10892307

Con el objetivo de evaluar si el polimorfismo rs10892307 estaba asociado a cambios en la expresión génica de *CXCR5*, se determinaron los niveles de ARNm de dicho gen en CMSP de pacientes con EM que no habían recibido tratamiento clasificados por su tenencia o carencia del alelo minoritario del polimorfismo rs10892307 mediante *real-time* PCR. El análisis de los datos obtenidos reveló que no existían diferencias estadísticamente significativas entre ambos grupos de pacientes (Figura 32).

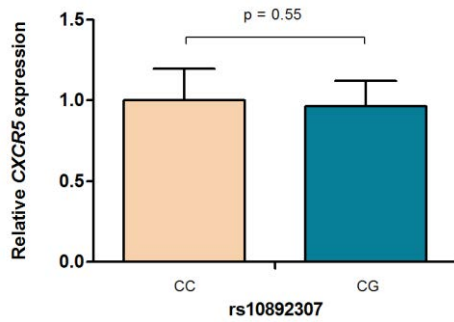


Figura 32. Diagrama de barras representativo del nivel de expresión génica de *CXCR5* determinado por *real-time* PCR en CMSP de pacientes con EM con presencia y ausencia del alelo minoritario (G) del SNP rs10892307. Los resultados se expresan como la media \pm el error estándar de la media (SEM) de la expresión de los pacientes CC respecto a los CG.

4.2.2.3. El polimorfismo rs11602393, que está en alto LD con rs10892307, aumenta la actividad transcripcional del promotor de *CXCR5*

Los resultados negativos de la *real-time* PCR nos invitaron a preguntarnos si este polimorfismo tiene un efecto modesto que no se puede percibir en el conjunto de CMSP. Por ello, se decidió realizar un estudio más específico y preciso y evaluar el efecto individual a nivel transcripcional no sólo del polimorfismo rs10892307, sino también de aquéllos que presentan un alto LD con el mismo (rs55756957, rs11602393 y rs3176905), mediante un ensayo reportero de luciferasa. En esta técnica, la secuencia reguladora bajo estudio, que contiene el SNP de interés, se clona *upstream* del gen de la luciferasa de un promotor mínimo en un plásmido artificial. Para cada SNP se construyeron cuatro plásmidos, cada uno de ellos portador de un fragmento con el alelo minoritario (MA, por *minor allele*) o con el de referencia (RA, por *reference allele*), y en orientación *forward* o *reverse* (Figura 33). Por tanto, en total se obtuvieron 16 constructos. Cada uno de ellos se transfectó en células HEK-293, donde el gen de la luciferasa se transcribirá proporcionalmente a como lo haría el gen diana si el individuo tuviera el polimorfismo de interés. Como control interno para normalizar los valores de expresión del constructo de interés, se utiliza un plásmido con un gen luciferasa distinto al previamente usado (en vez de ser de luciérnaga, de renilla) con un promotor constitutivo *upstream*.

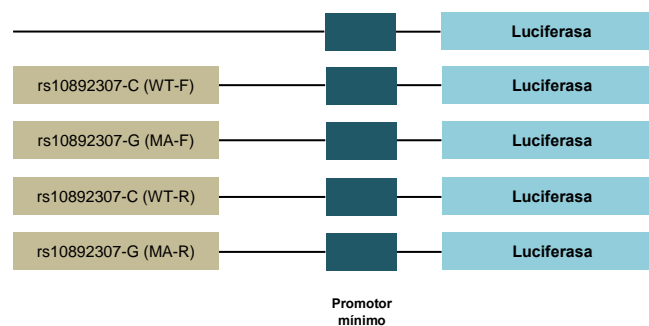


Figura 33. Ejemplo de los constructos resultantes para el polimorfismo rs10892307. Este procedimiento se realizó para los cuatro SNPs seleccionados. MA: alelo minoritario; WT: alelo de referencia; F: *forward*; R: *reverse*.

A continuación, se desglosan los resultados en función del polimorfismo bajo estudio:

- Como ilustra la Figura 34A, se observó que el constructo portador del fragmento con el RA (A) del polimorfismo rs55756957 mostraba una actividad de la luciferasa similar a la del vector vacío en ambas orientaciones, lo que evidencia su falta de capacidad para alterar la actividad transcripcional del promotor de *CXCR5*. Cuando el inserto contenía el MA (G) de rs55756957, sí se percibió un aumento de la actividad de la luciferasa respecto a la del vector vacío en la orientación *forward* ($p=0,005$), hecho que revela su capacidad como *enhancer* transcripcional. Este efecto del MA no se observó en orientación *reverse*. Sin embargo, cuando se comparó la actividad de la luciferasa entre el RA y el MA no se identificaron diferencias significativas en ninguna de las dos orientaciones, lo que indica que la tenencia del MA en vez de del RA no implica un cambio en la transcripción de *CXCR5*.
- La Figura 34B plasma los resultados respecto al polimorfismo rs11602393. El vector con el fragmento que contenía el RA (T) mostró actividad *enhancer* respecto al vector vacío en orientación *reverse* ($p=0,0009$), pero no en *forward*. Sin embargo, el MA (A) de este SNP confiere actividad *enhancer* respecto al vector vacío independiente de la orientación, ($p=0,03$ en orientación *forward*; $p=0,002$ en orientación *reverse*). De forma interesante, cuando se comparó la actividad de la luciferasa en orientación *forward* entre el vector portador del fragmento con el RA y el portador del fragmento con el MA, se observó que el MA confería un incremento estadísticamente significativo de la actividad transcripcional de 2,16 veces respecto al RA ($p=0,02$).
- Tal y como representa la Figura 34C, los cuatro constructos para la variante rs10892307 mostraron actividad *enhancer* respecto al vector vacío ($p\leq 0,02$ en todos los casos), por lo que este polimorfismo es un potenciador de la actividad transcripcional de manera orientación-independiente y alelo-independiente. Sin embargo, no se hallaron diferencias significativas al comparar la actividad transcripcional entre los constructos con el RA (C) y con el MA (G) en ninguna de las dos orientaciones.
- Por último, como refleja la Figura 34D, se evidenció que el RA (T) del polimorfismo rs3176905 no confería un aumento estadísticamente significativo de la actividad del promotor cuando se comparó con el vector vacío en ninguna de las orientaciones, pero sí lo hacía el MA (C), aunque sólo en orientación *reverse* ($p=0,01$). Sin embargo, al comparar la actividad luciferasa entre los vectores con el RA y con el MA, no se encontraron diferencias significativas en ninguna orientación.

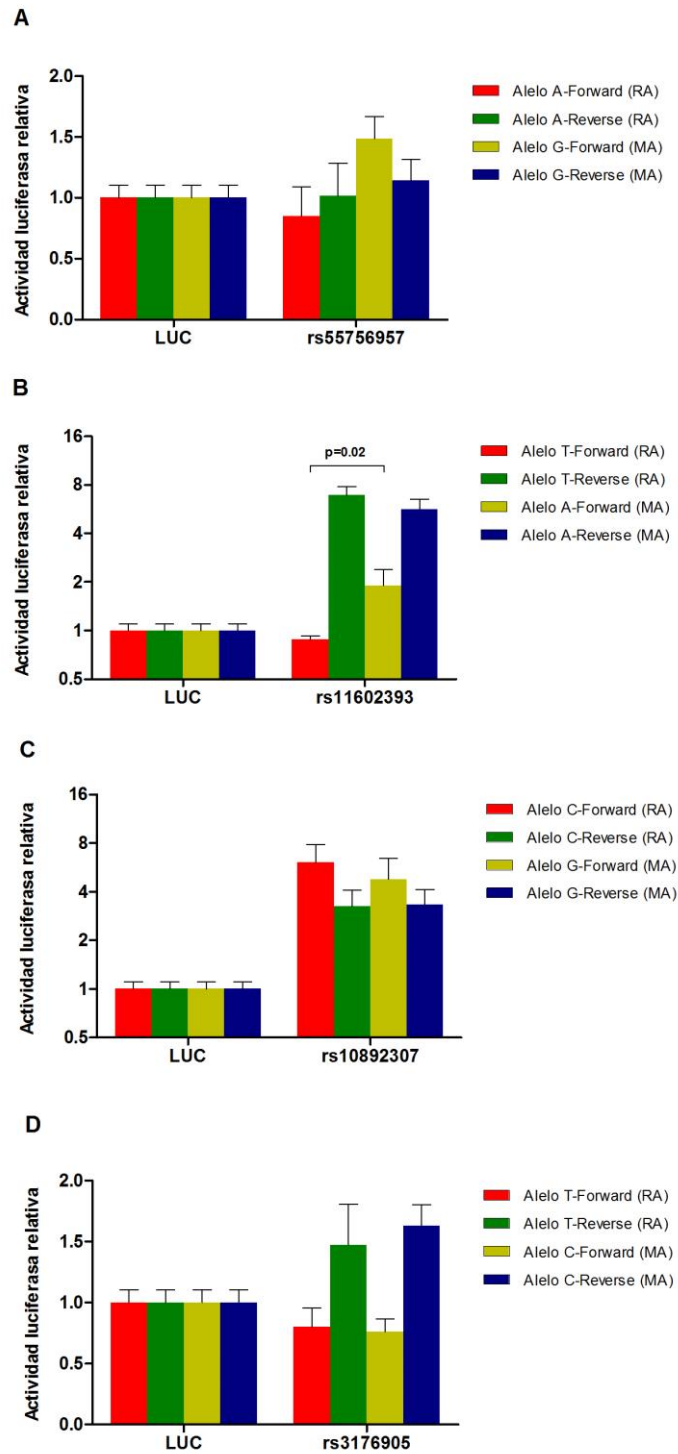


Figura 34. Actividad relativa de la luciferasa de los diferentes constructos, portadores de fragmentos potencialmente reguladores, respecto al vector vacío (LUC). Para cada uno de los 16 vectores construidos, se representa la media de 4 experimentos de transfección independientes. En todos los casos, los niveles de actividad de luciferasa se expresan respecto al nivel del plásmido control que contiene sólo el promotor básico y el gen que codifica la luciferasa de renilla. Sólo se indica el p-valor relevante, que es al comparar para una misma orientación diferentes alelos. RA: *alelo de referencia*. MA: *alelo minoritario*.

4.2.2.4. El porcentaje de linfocitos T reguladores que expresan CXCR5 a nivel basal está disminuido en los pacientes con EM con el alelo minoritario de la variante rs10892307

Con el objetivo de estudiar si el incremento de la actividad transcripcional asociado al SNP rs11602393 podría percibirse a nivel de proteína, se analizó la expresión basal de CXCR5 en la superficie de diferentes poblaciones de células T y células B de pacientes con EM con presencia y ausencia del alelo minoritario del SNP rs10892307, en total LD con el SNP rs11602393 ($r^2=0,94$), mediante citometría de flujo. La finalidad era evaluar si dicho incremento afectaba a alguna población celular concreta. Se determinaron tanto el porcentaje de células que expresan CXCR5 en su superficie, como la intensidad media de fluorescencia (MFI, por *mean fluorescence intensity*), que estima la cantidad de moléculas de CXCR5 presentes en el conjunto de células de cada población.

La Tabla 24 resume los resultados obtenidos para las poblaciones de linfocitos T analizadas. Tal y como ilustra la Figura 35I, se identificó que el porcentaje de linfocitos T reguladores que expresaban CXCR5 en su superficie era mayor en el grupo de pacientes con EM portadores del alelo minoritario del polimorfismo rs10892307 respecto al grupo de pacientes con EM que carecían de dicho alelo ($p=0,04$), aunque no se observaron diferencias en los niveles de expresión representados por los valores de la MFI. Respecto a las demás poblaciones estudiadas, la expresión de CXCR5 resultó similar entre ambos grupos de pacientes (Figura 35).

	%CXCR5 ⁺ (DE)			MFI CXCR5 (DE)		
	RA rs10892307	MA rs10892307	p-valor	RA rs10892307	MA rs10892307	p-valor
Linfocitos T CD4 ⁺	12,5 (2,2)	12,1 (3,2)	0,80	61,4 (13,2)	75,1 (37,3)	0,91
Linfocitos T CD8 ⁺	1,5 (0,8)	2,1 (1,2)	0,20	45,0 (10,3)	58,8 (23,4)	0,06
Linfocitos T CD4 ⁺ naïve	1,2 (0,6)	1,8 (1,3)	0,39	24,2 (4,0)	28,3 (11,5)	0,63
Linfocitos T CD4 ⁺ memoria ¹	17,1 (3,2)	18,9 (6,0)	0,63	66,2 (12,7)	82,7 (38,5)	0,48
Linfocitos T reguladores²	19,7 (4,9)	26,8 (9,8)	0,04	13,8 (2,8)	15,5 (5,3)	0,57

Tabla 24. Resumen del porcentaje de células de diferentes poblaciones de linfocitos T que expresan CXCR5 en su superficie y de la MFI para CXCR5 de dichas poblaciones en pacientes con EM con ausencia o presencia del alelo minoritario del SNP rs10892307. *DE: desviación estándar. RA: alelo de referencia. MA: alelo minoritario.* ¹ El hecho de que una célula T CD4⁺ de memoria exprese CXCR5 en su superficie supone que se trate de un linfocito T colaborador folicular (Tfh); asimismo, ² un linfocito T regulador que expresa CXCR5 en su superficie se denomina linfocito T regulador folicular (Tfr).

La Tabla 25 refleja los resultados obtenidos para las diferentes poblaciones de linfocitos B.

	%CXCR5 ⁺ (DE)			MFI CXCR5 (DE)		
	RA rs10892307	MA rs10892307	p-valor	RA rs10892307	MA rs10892307	p-valor
Linfocitos B	97,6 (2,1)	97,7 (1,6)	0,97	265,8 (69,3)	266,6 (45,6)	0,85
Linfocitos B <i>naïve</i>	99,5 (0,4)	99,5 (0,2)	0,54	280,5 (72,8)	278,3 (48,8)	0,80
Linfocitos B memoria	99,0 (0,9)	99,0 (1,1)	0,85	221,0 (72,4)	216,6 (57,6)	0,91
Plasmablastos	15,9 (6,5)	19,7 (11,9)	0,90	205,9 (90,0)	149,6 (59,8)	0,11
Linfocitos Breg	99,1 (0,9)	99,1 (0,5)	0,85	279,0 (48,1)	272,6 (42,8)	0,91

Tabla 25. Porcentaje de células de diferentes poblaciones de linfocitos B que expresan CXCR5 en su superficie y de la MFI para CXCR5 de las mismas poblaciones en pacientes con EM portadores o no portadores del alelo minoritario del SNP rs10892307. *DE: desviación estándar. RA: alelo de referencia. MA: alelo minoritario.*

Como se puede observar tanto en la Tabla 25 como en la Figura 36, la gran mayoría de los linfocitos B, B *naïve* y B memoria presentaban CXCR5 en su superficie, independientemente del genotipo de los pacientes para el polimorfismo bajo estudio. Por el contrario, el porcentaje de plasmablastos que expresaban CXCR5 era muy inferior (Tabla 25, Figura 36). Al clasificar a los pacientes con EM en función de si tenían o no el alelo minoritario del SNP rs10892307, no se observaron diferencias estadísticamente significativas ni en el porcentaje de los diferentes subtipos de células B que expresaban CXCR5 en su superficie, ni en el nivel de expresión de CXCR5 representado por la MFI (Figura 36).

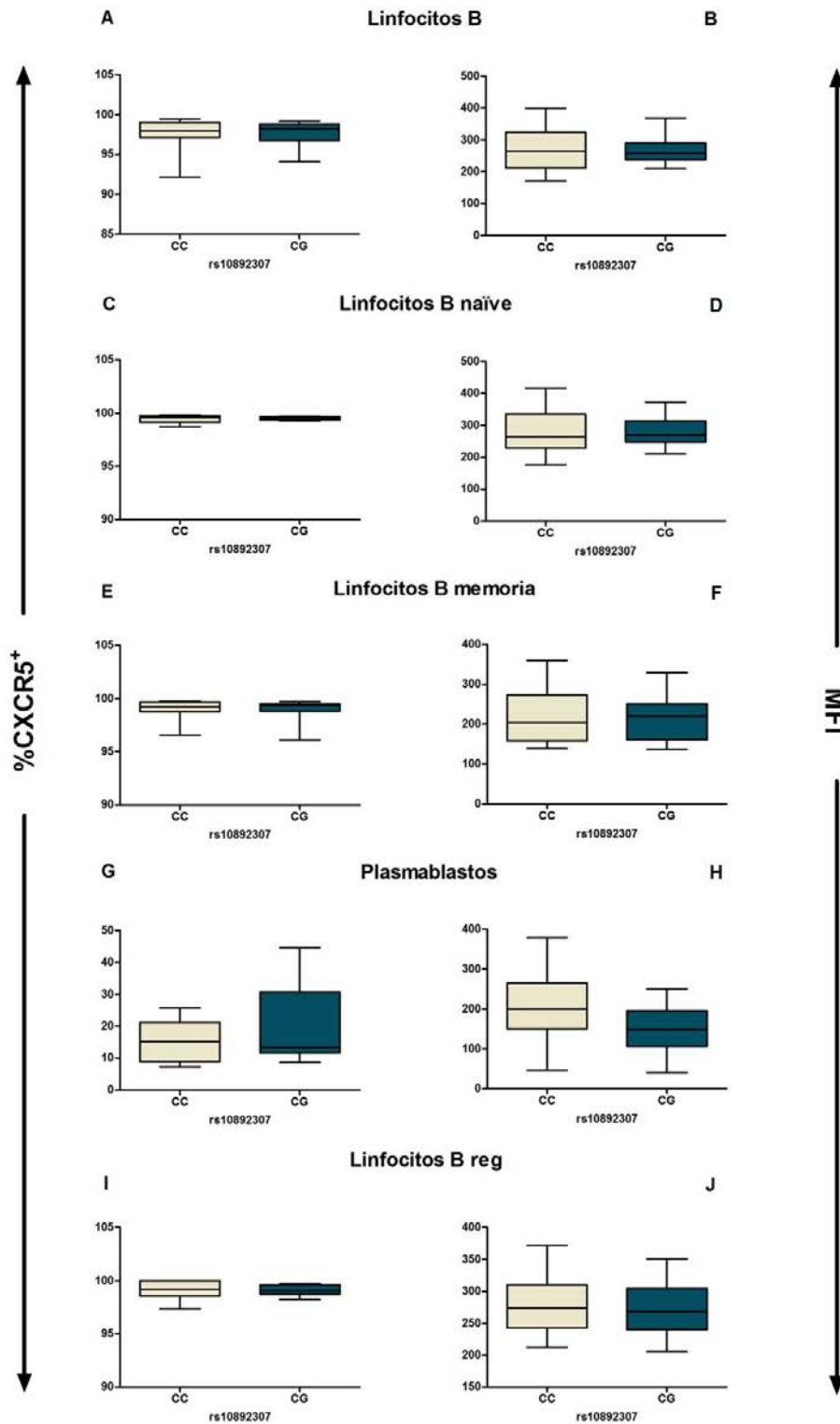


Figura 36. Expresión de CXCR5 en la superficie celular de las poblaciones estudiadas de linfocitos B de pacientes con EM con presencia o ausencia del alelo minoritario del SNP rs10892307. En los gráficos de la columna izquierda se representa el porcentaje de células que expresan CXCR5 en su superficie, mientras que en los de la columna derecha se representa la MFI de cada población para CXCR5. *MFI: intensidad media de fluorescencia.*

4.2.3. Estudios relacionados con el polimorfismo rs2762943 (*CYP24A1*)

Del mismo modo que para el SNP rs10892307, se diseñaron diferentes estudios funcionales para estudiar el posible efecto del polimorfismo del gen *CYP24A1* validado: rs2762943.

4.2.3.1. El polimorfismo rs2762943 no está en desequilibrio de ligamiento con ningún SNP en población europea y tiene actividad *enhancer*

Con el objetivo de evaluar si el polimorfismo del gen *CYP24A1* validado, rs2762943, estaba en alto LD con algún otro SNP en población europea, se analizó el coeficiente de correlación al cuadrado (r^2) que mantenía este polimorfismo con el resto de los presentes en el genoma humano utilizando LDLink. Dado que ninguna variante mantenía un valor superior al umbral de r^2 establecido ($r^2=0,5$) con el SNP rs2762943, se consideró que éste no estaba en alto LD con ningún polimorfismo en población europea.

A continuación, se comprobó la puntuación asignada al polimorfismo rs2762943 por la base de datos Regulome DB, cuyo rango es del 1 al 6. Cuanto menor sea el valor de la puntuación, mayor es la evidencia experimental que sitúa a una variante como implicada en la regulación de la expresión de un gen. Como resultado, se obtuvo una puntuación de 2b, la cual supone que puede afectar a la unión de factores de transcripción, hallazgo determinado mediante ensayos de *footprinting* y picos de sensibilidad a DNasa entre otros. En particular, se afectarían los motivos de unión de las proteínas ETS1 y FLI1, ya que el SNP rs2762943 se localiza en su zona de reconocimiento. En ambos casos, tal y como muestra la Figura 37, experimentalmente se ha encontrado de forma más frecuente al alelo minoritario (T) en el sitio de unión de estas proteínas que al alelo mayoritario (G).

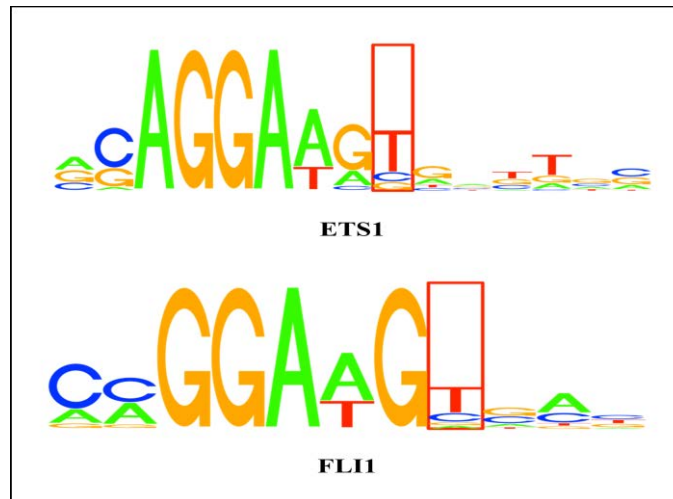


Figura 37. Resultados del análisis del SNP rs2762943 mediante Regulome DB. Se representa la PWM de los motivos de unión de ETS1 (arriba) y FLI1 (abajo), identificados mediante la técnica de *footprinting*, que se utiliza para determinar la secuencia de ADN a la que se une una proteína. Los nucleótidos que aparecen dentro del recuadro rojo ubican la localización del polimorfismo rs2762942 y la altura relativa de cada uno es proporcional a la frecuencia en la que se ha encontrado en esa posición experimentalmente.

4.2.3.2. Los niveles en suero de vitamina 1,25(OH)₂-D están disminuidos en pacientes con EM portadores del alelo de riesgo de rs2762943, pero este polimorfismo no afecta a la concentración de vitamina 25(OH)-D

Dado que el gen *CYP24A1* codifica una enzima cuya función es la inactivación tanto del metabolito biológicamente activo de la vitamina D como de su precursor, se analizó si existía alguna relación entre el polimorfismo rs2762943 del gen *CYP24A1* y los niveles de vitamina D en suero de pacientes con EM que no habían recibido tratamiento para la enfermedad en el momento de la obtención de la muestra.

Como muestra la Figura 38A, se observó una disminución significativa de los niveles en suero de la forma biológicamente activa de la vitamina D, el calcitriol o 1,25(OH)₂-D, en los pacientes con EM portadores del alelo de riesgo del polimorfismo rs2762943 respecto a aquéllos sin dicho alelo ($p=0,004$). No obstante, los niveles de su precursor, el calcidiol o 25(OH)-D, resultaron similares entre ambos grupos de pacientes con EM (Figura 38B).

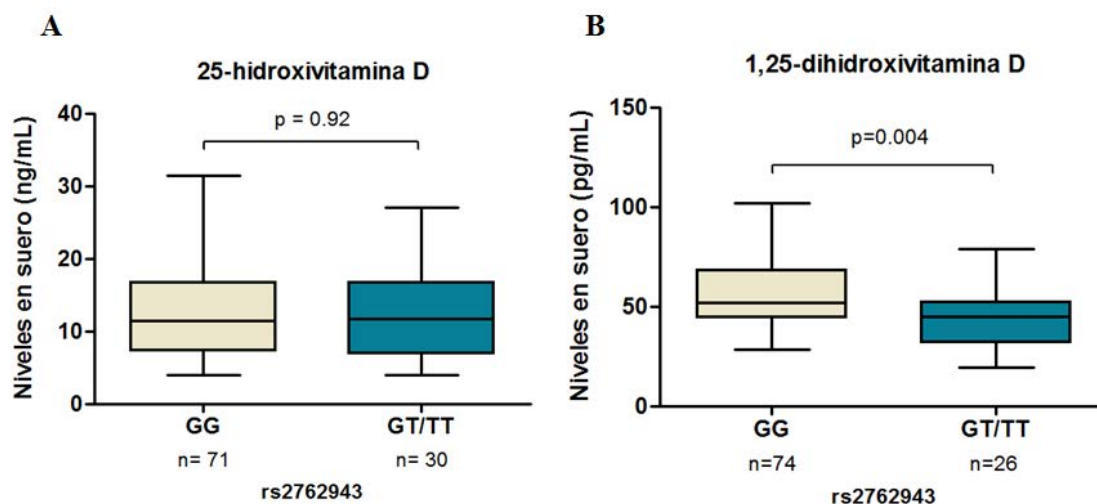


Figura 38. Concentración en suero de pacientes con EM de (A) 1,25-dihydroxivitamina D (1,25(OH)₂-D) y de (B) 25-hidroxitamina D (25(OH)-D) determinada mediante ensayos de quimioluminiscencia. Los pacientes se clasificaron en función de su genotipo para el polimorfismo rs2762943. Los resultados se expresan como media \pm desviación estándar (DE).

4.2.3.3. Los niveles basales de ARNm de *CYP24A1* son similares entre los pacientes con EM con y sin el alelo de riesgo de rs2762943

Con el objetivo de investigar si existía una diferencia en la expresión génica basal de *CYP24A1* entre pacientes con EM influenciada por su genotipo para el SNP rs2762943, se determinaron los niveles de ARNm en ambos grupos en CMSP obtenidas antes de haber recibido tratamiento para la enfermedad. Tal y como muestra la Figura 39, la expresión basal de *CYP24A1* resultó comparable entre ambos grupos.

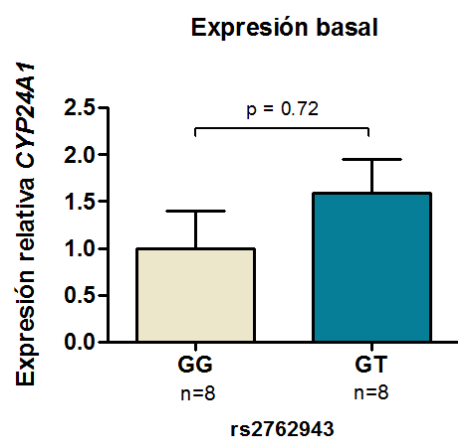


Figura 39. Niveles de expresión génica de *CYP24A1* en CMSP de pacientes con EM clasificados en función de la presencia o ausencia del alelo minoritario del SNP rs2762943. Los resultados se determinaron por *real-time* PCR y se representan como la expresión media \pm SEM.

4.2.3.4. El polimorfismo rs2762943 no afecta a la inducción de la expresión génica de *CYP24A1* mediante el tratamiento con 1,25(OH)₂-D₃ en CMSP de pacientes con EM

Las CMSP de pacientes con EM agrupados en función de su genotipo para el polimorfismo rs2762943 se extrajeron a nivel basal y se incubaron en presencia o ausencia de vitamina 1,25(OH)₂-D₃ durante 24 horas. Una vez pasado este tiempo, se determinó la expresión génica de *CYP24A1* mediante *real-time* PCR, dado que está descrito que el tratamiento con la vitamina 1,25(OH)₂-D₃ es capaz de incrementarla.

Como resultado, se observó que tanto los pacientes con el alelo de riesgo del SNP rs2762943 como aquéllos sin él exhibieron la misma capacidad de aumentar la expresión génica de *CYP24A1* tras estimular a las CMSP con 1,25(OH)₂-D₃ (Figura 40A). Por otro lado, al evaluar los niveles ARNm de *CYP24A1* de los pacientes con EM sin el alelo de riesgo de rs2762943 respecto a los pacientes con él, se detectó una expresión 4,03 veces mayor en los pacientes portadores del alelo de riesgo respecto a los no portadores tras haber recibido el tratamiento con 1,25(OH)₂-D₃, aunque este resultado no alcanzó la significancia estadística (Figura 40B).

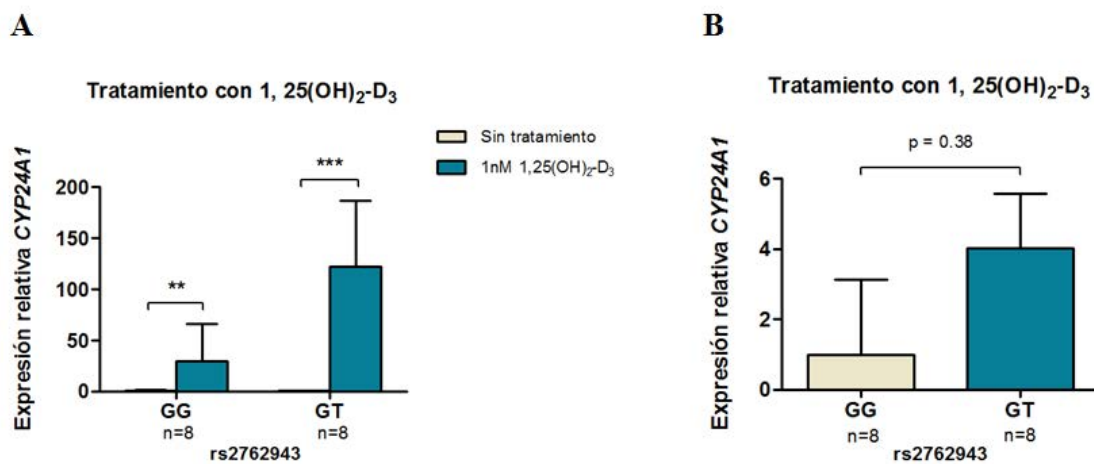


Figura 40. Diagramas de barras representativos de la respuesta del gen *CYP24A1* ante el tratamiento con 1,25(OH)₂-D₃. Se representan los niveles relativos de expresión génica de *CYP24A1* (A) antes y después de haber recibido la dosis de 1,25(OH)₂-D₃ en ambos grupos de pacientes y (B) solamente después de haber recibido el tratamiento comparando un grupo de pacientes respecto al otro.

4.3. Resultados relacionados con el objetivo 3

“Investigar la posible implicación funcional de la presencia de variantes genéticas raras de genes que las acumulan con mayor frecuencia en pacientes con EM”

En base a los resultados obtenidos en la resecuenciación de ADN, se procedió a desarrollar estudios adicionales para investigar si la acumulación de variantes raras (VRs) en genes de riesgo que se identificó de forma más frecuente en pacientes con EM que en controles sanos tiene algún efecto funcional.

4.3.1. Determinación de los niveles de ARNm de genes que acumulan variantes raras en CMSP de pacientes con EM

Mediante PCR a tiempo real, se determinaron los niveles de ARN mensajero de los genes *FCRL1*, *RGS1*, *TRAF3* y *CYP24A1* en las CMSP de pacientes con EM que portaban al menos una VR en la región identificada *in silico* como enriquecida y se compararon con los de pacientes con EM sin ninguna VR en la región correspondiente para evaluar el posible efecto de las mismas.

4.3.1.1. La expresión génica de *FCRL1* resultó similar en los pacientes con EM con y sin variantes raras en su promotor

Con el fin de investigar el posible impacto de las VRs en el promotor de *FCRL1*, se estratificaron los pacientes con EM que no habían recibido tratamiento para la enfermedad según su tenencia o ausencia de alguna VR en el promotor del gen *FCRL1*. Como representa la Figura 41, se observó que la expresión génica de *FCRL1* en CMSP era igual en los dos grupos de pacientes ($p=0,61$).

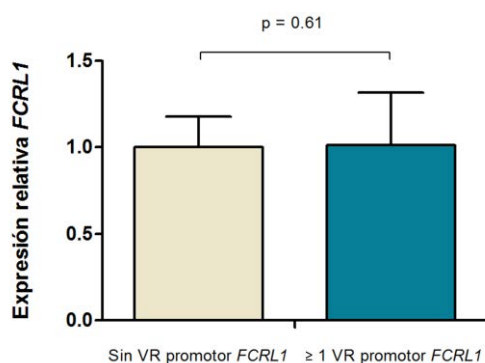


Figura 41. Diagrama de barras mostrando el nivel de expresión génica de *FCRL1* determinado por *real-time* PCR en CMSP de pacientes con EM con presencia y ausencia de VRs en el promotor del gen *FCRL1*. Los resultados se expresan como la media ± el error estándar de la media (SEM) de la expresión de los pacientes con VRs respecto a aquéllos sin ellas.

4.3.1.2. La expresión génica basal de *RGS1* está disminuida en pacientes con EM con variantes raras en dicho gen

Además, con el objetivo de evaluar si las VRs de *RGS1* tenían algún efecto a nivel transcripcional, se cuantificó la expresión génica de *RGS1* en CMSP de pacientes con EM con alguna VR de dicho gen y en pacientes con EM sin ninguna VR del mismo, todos ellos sin haber recibido tratamiento para la EM. Como resultado, se observó que la expresión génica de *RGS1* estaba disminuida de forma estadísticamente significativa en los pacientes con EM con VRs de *RGS1* cuando se compararon con aquellos pacientes sin ellas ($p=0,01$, Figura 42A). La Tabla 26 muestra las VRs del gen *RGS1* presentes en cada uno de los 6 pacientes con EM con VRs incluidos en el estudio, así como su descripción básica.

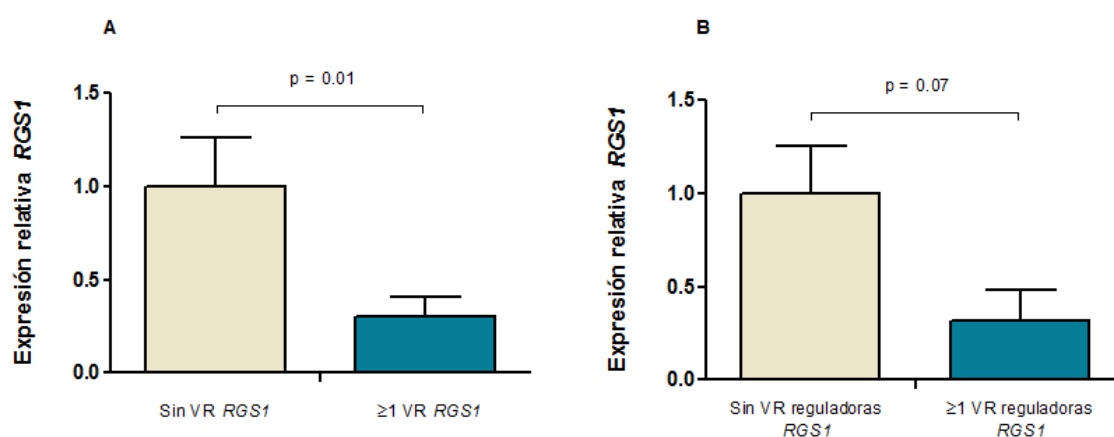


Figura 42. Diagramas de barras mostrando los niveles de ARNm del gen *RGS1* determinados por *real-time* PCR en CMSP de pacientes con EM con al menos una VR (A) a lo largo de todo el gen *RGS1* o (B) específicamente en regiones reguladoras del mismo, en ambos casos respecto a pacientes con EM sin ninguna VR en la región bajo estudio. Los resultados se expresan como la media \pm SEM de la expresión génica relativa de *RGS1* en pacientes con EM con VRs respecto a aquéllos sin VRs.

Por otro lado, tal y como muestra la Figura 42B, no se observaron diferencias estadísticamente significativas en la expresión génica de *RGS1* entre pacientes con EM con y sin VRs específicamente en regiones reguladoras del gen ($p=0,07$).

Dado su posible efecto funcional, se diseñó la Figura 43 para percibir una visualización de la distribución relativa todas las VRs del gen *RGS1* detectadas en la resecuenciación de ADN.

Resultados - Objetivo 3

Tabla 26. Descripción y consecuencias funcionales de los SNPs raros presentes en los 6 pacientes con EM portadores de VRs incluidos en la determinación de los niveles de expresión génica de *RGS1* por *real-time* PCR.

SNP (Ref./Alt.)	Pacientes EM	Localización (hg19)	MAF (alelo minoritario) ¹	Localización (<i>Ensembl</i>)
rs12065772 (C/T)	EM1 EM4 EM5 EM6	1:192550461	0,005 (T)	<i>Downstream</i>
rs142034366 (TAGG/T)	EM1 EM5	1:192545249	0,004 (T)	Exón no codificante
rs74130688 (G/A)	EM1 EM5	1:192543002	0,004 (A)	<i>Upstream</i>
rs4658058 (T/C)	EM1 EM5	1:192548930	0,003 (C)	UTR3'
rs60719978 (C/T)	EM1 EM5 EM6	1:192548245	0,003 (T)	Intrónica
rs12094893 (G/A)	EM1 EM5	1:192549871	0,003 (A)	<i>Downstream</i>
N/A (G/C)	EM2	1:192550276	0,0005 (C)	Intergénica
rs16834436 (A/G)	EM3	1:192546355	0,003 (G)	Exón no codificante
rs77617886 (T/A)	EM3	1:192546577	0,003 (A)	Exón no codificante
rs74130689 (G/A)	EM4	1:192548794	0,001 (A)	UTR3'
rs3806353 (C/T)	EM5	1:192543623	0,004 (T)	<i>Upstream</i>
rs3806354 (A/C)	EM5	1:192543513	0,004 (C)	<i>Upstream</i>
rs76219467 (A/T)	EM5	1:192546429	0,003 (T)	Exón no codificante
rs113090598 (T/A)	EM6	1: 192548461	0,0005 (A)	UTR3'
rs149919199 (G/C)	EM6	1: 192549635	0,0005 (C)	<i>Downstream</i>

¹ MAF=frecuencia del alelo minoritario en la cohorte de resecuenciación. Los 6 pacientes se enumeran del 1 al 6 y aparecen con el prefijo "EM". La tabla muestra la localización de las VRs en el cromosoma número 1, la frecuencia de su alelo minoritario (MAF) en la cohorte de resecuenciación y su consecuencia funcional determinada por la plataforma *Ensembl*. *Ref.:* alelo de referencia. *Alt.:* alelo alterado.

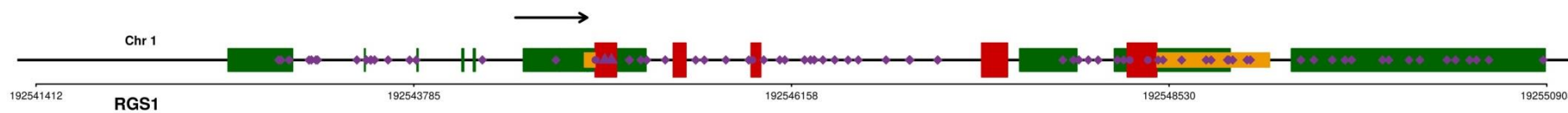


Figura 43. Representación esquemática de la distribución en el espacio genómico de las VRs identificadas en la resecuenciación de ADN en el gen *RGS1*. Los círculos representan las VRs no sinónimas, los triángulos las VRs sinónimas y los diamantes el resto de VRs. Las cajas rojas se corresponden con los exones del gen, las verdes con las secuencias reguladoras y las amarillas con las regiones UTR.

4.3.1.2. La presencia de variantes raras en las regiones UTR del gen *TRAF3* no afecta a su nivel de expresión génica en pacientes con EM

La cuantificación de la expresión génica del gen que codifica para la proteína TRAF3 en CMSP de pacientes con EM sin tratamiento y bien con alguna VR en las regiones UTR de dicho gen, o bien sin ninguna VR del mismo, determinó que no existían diferencias significativas entre ambos grupos ($p=0,32$, Figura 44).

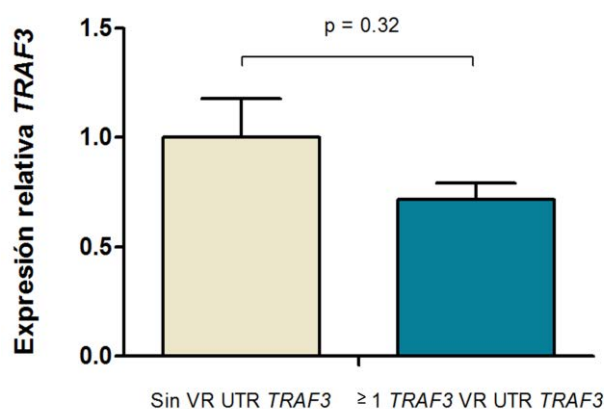


Figura 44. Niveles de expresión génica de *TRAF3* en CMSP de pacientes con EM con presencia de al menos una VR en las regiones UTR de *TRAF3* respecto a aquéllos con ausencia de las mismas. Los resultados se determinaron por *real-time* PCR y se expresan como la expresión media \pm SEM.

4.3.1.4. La acumulación de variantes raras en las regiones específicas del gen *CYP24A1* no afecta a su expresión génica en pacientes con EM

Para investigar si la presencia de VRs en las regiones del gen *CYP24A1* con tendencia a su acumulación en pacientes con EM tenía algún efecto sobre la expresión génica del mismo, se compararon los niveles de ARNm de *CYP24A1* entre varios grupos de pacientes con EM: i) pacientes con al menos una VR localizada en una región reguladora de *CYP24A1* y pacientes sin VRs en regiones reguladoras; ii) pacientes con alguna VR en las regiones UTR de *CYP24A1* y pacientes sin ninguna VR en dichas regiones; iii) pacientes con una o más VRs en el promotor de *CYP24A1* y pacientes sin VRs en él. Ninguno de los pacientes incluidos en los análisis había recibido tratamiento para la EM en el momento de la extracción de la muestra. Como refleja la Figura 45A-C, no se observaron diferencias estadísticamente significativas en la expresión génica de *CYP24A1* entre ambos grupos de pacientes en ninguno de los casos ($p>0,05$).

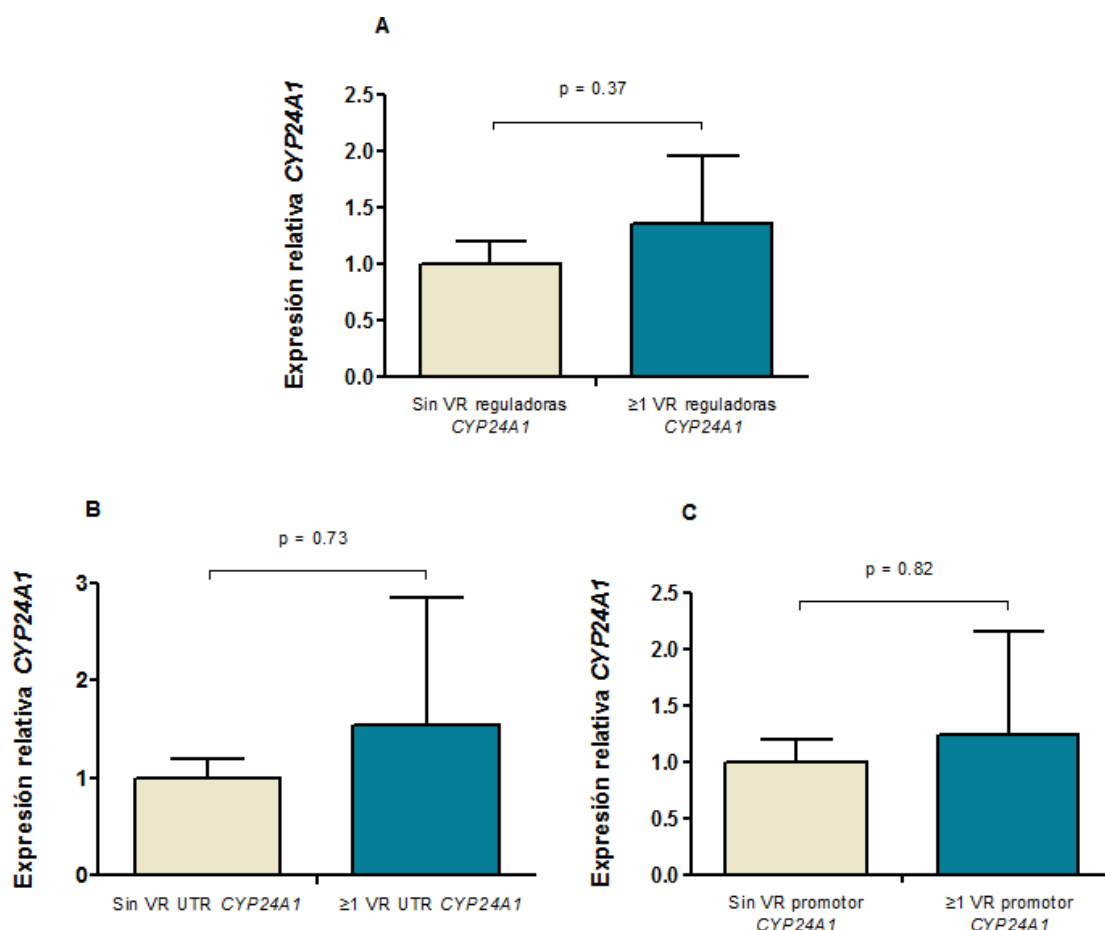


Figura 45. Diagrama de barras mostrando el nivel de expresión génica de *CYP24A1* determinado por *real-time* PCR en CMSP de pacientes con EM con VRs (A) en regiones reguladoras del mismo, (B) en sus regiones UTR o (C) específicamente en el promotor del gen, en todos los casos respecto a pacientes con EM sin VRs en la región bajo estudio. Los resultados se expresan como la media \pm SEM.

4.3.2. Estudios relacionados con la acumulación de variantes raras en el gen *RGS1* en pacientes con EM

4.3.2.1. Los pacientes con EM con la expresión génica de *RGS1* más baja comparten variantes raras

Con el fin de profundizar más en el estudio del efecto significativo de las VRs de *RGS1* sobre su expresión génica, se analizó el perfil de VRs de los pacientes incluidos en el estudio. Curiosamente, se observó que los dos pacientes que mostraban la expresión génica de *RGS1* más disminuida (EM1 y EM5) compartían 5 SNPs raros (ver Tabla 26) y una *indel* (por inserción-delección) rara (rs142034366). La Figura 46 representa esquemáticamente la localización relativa de los 6 polimorfismos compartidos en el gen *RGS1*.

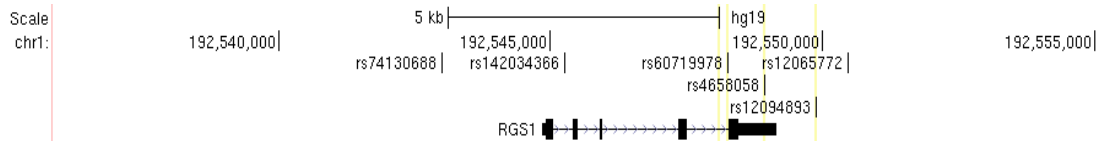
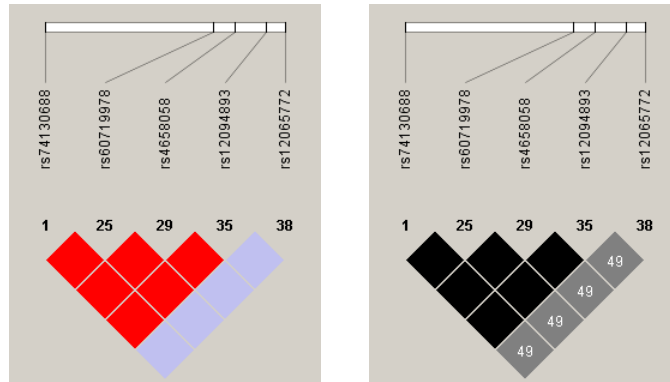


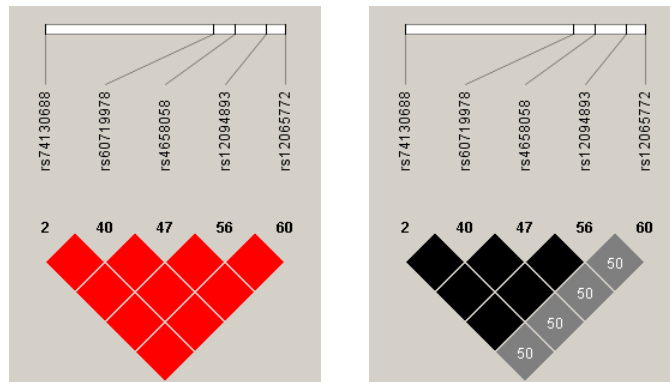
Figura 46. Ilustración esquemática de la localización de los 6 polimorfismos compartidos por los dos pacientes con EM con la expresión génica más baja de *RGS1*. Los polimorfismos se representan en la región del gen *RGS1*, tal y como indica ENCODE. Las cajas negras simbolizan los exones de *RGS1*.

A continuación, se analizó el grado de LD entre los 5 SNPs compartidos por estos dos pacientes para evaluar su frecuencia de recombinación en diferentes poblaciones utilizando los datos del Proyecto 1000 Genomas. Se calculó tanto el valor de D' , como el de r^2 , las dos medidas utilizadas en estudios genéticos para determinar el LD entre dos SNPs. Como se puede observar en la Figura 47, se trata de un haplotipo más frecuente en la población íbera que en el resto de Europa. Curiosamente, en la población africana el LD entre los 5 SNPs es totalmente diferente; de hecho, la frecuencia de estos alelos es alta en África si se compara con la de España y Europa.

Población íbera



Población europea



Población africana

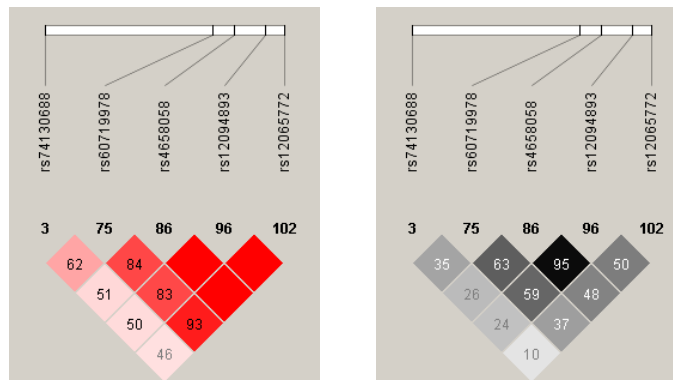


Figura 47. Mapa del LD existente entre los 5 SNPs en la región del gen *RGS1* compartidos por los dos pacientes con EM con la expresión de dicho gen más baja. Se representa el LD en las poblaciones íbera, europea y africana. Los mapas de la izquierda (en rojo) representan el valor de D' , mientras que los mapas de la derecha (en negro) representan el valor de r^2 . Los números en negrita asignados a cada SNP indican su posición en la lista de SNPs generada por HaploView. Cada cuadrado denota la fuerza del LD entre una pareja de SNPs, que será mayor cuanto más alto sea el número y más intenso sea el color.

4.3.2.2. La inducción de la expresión génica de *RGS1* mediante el tratamiento con IFN- β en CMSP está alterada en los pacientes con EM con variantes raras en dicho gen

Dado que se ha demostrado que el tratamiento con IFN- β es capaz de aumentar la expresión génica de *RGS1* en CMSP (288), se exploró si la presencia de al menos una VR a lo largo de la secuencia del mismo afectaba a su inducción. Para ello, se determinaron los niveles de ARNm del gen *RGS1* en CMSP de pacientes con EM con ausencia o presencia de VRs de dicho gen antes y 4 horas después de recibir tratamiento *in vitro* con IFN- β . Como muestra la Figura 48, la dosis de 100 UI/mL de IFN- β indujo de forma discreta la expresión génica de *RGS1* de forma significativa en pacientes con EM que carecen de VR ($p=0,007$), tal y como cabría esperar en condiciones normales. Sin embargo, en los pacientes con EM con al menos una VR en *RGS1* se observó un defecto en su inducción, viéndose la expresión de *RGS1* prácticamente inalterada antes y después del tratamiento ($p>0,05$).

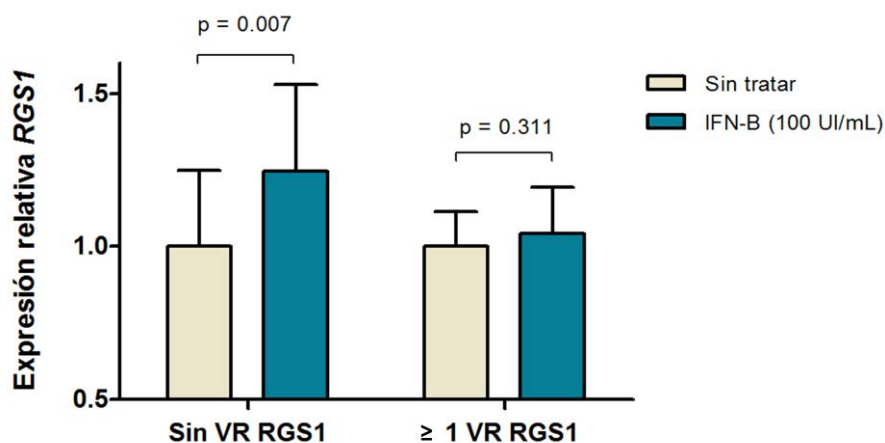


Figura 48. Expresión génica relativa de *RGS1* en CMSP de pacientes con EM, clasificados según la presencia o ausencia de VRs de *RGS1*, antes (sin tratar) y 4 horas después de haber recibido tratamiento con IFN- β *in vitro* (IFN- β).

4.3.2.3. La expresión de *RGS1* en linfocitos B está disminuida a nivel basal en los pacientes con EM con variantes raras de *RGS1*

Con el objetivo de investigar si el potencial efecto funcional de la acumulación de VRs del gen *RGS1* en pacientes con EM afectaba específicamente a alguna población celular, se analizó la expresión de *RGS1* mediante citometría de flujo en diferentes poblaciones de CMSP de pacientes con EM a nivel basal, clasificados en función de su tenencia o no de VRs de este gen.

La expresión de RGS1 se evaluó mediante la determinación individual del porcentaje de células T (CD3⁺, CD4⁺, CD8⁺, T reguladoras e NKT), células NK, células B y monocitos que son positivos para la molécula bajo estudio, es decir, que expresan RGS1 en la cara interna de su membrana plasmática. Tal y como se observa en la Figura 49A, la presencia de variantes raras en el gen *RGS1* se asoció con una disminución significativa del porcentaje de linfocitos B periféricos que expresan *RGS1* en pacientes con EM que no habían recibido tratamiento ($p=0,006$).

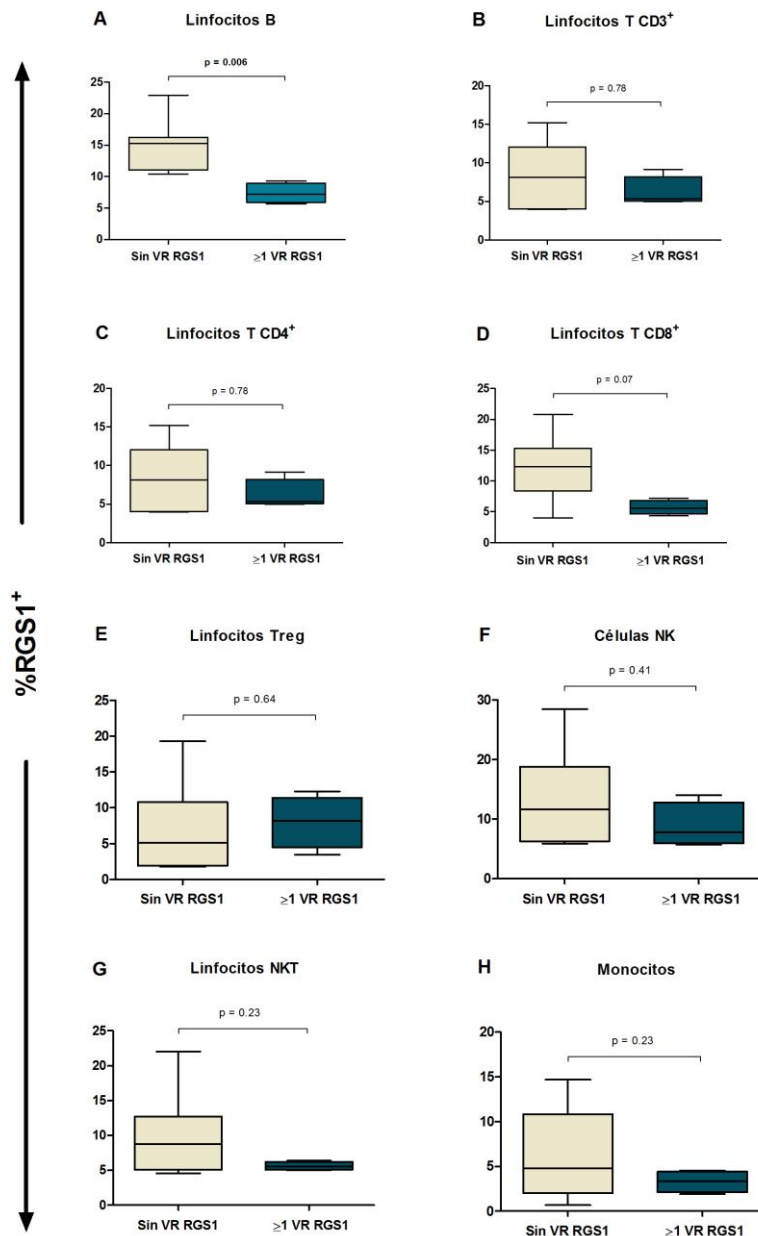


Figura 49. Porcentaje de células de diferentes poblaciones de CMSP provenientes de pacientes con EM con y sin variantes genéticas raras en *RGS1* que presentan RGS1 en la cara interna de su membrana plasmática.

Resultados - Objetivo 3

La Tabla 27 detalla el porcentaje de cada subpoblación celular que resultó positivo para RGS1 en cada grupo de pacientes con EM.

	% Células <i>RGS1</i> ⁺ (DE)		
	Sin VR <i>RGS1</i>	≥ 1 VR <i>RGS1</i>	p-valor
Linfocitos CD3 ⁺	8,6 (3,6)	5,2 (0,5)	0,23
Linfocitos CD4 ⁺	8,4 (4,37)	6,2 (2,0)	0,78
Linfocitos CD8 ⁺	11,8 (5,4)	5,7 (1,2)	0,07
Linfocitos T reguladores	7,3 (6,4)	8,0 (3,6)	0,64
Linfocitos NK	13,7 (8,5)	8,8 (3,7)	0,41
Linfocitos NKT	10,5 (6,0)	5,6 (0,6)	0,23
Linfocitos B	15,1 (4,1)	7,3 (1,6)	0,006
Monocitos	6,8 (5,1)	3,3 (1,2)	0,23

Tabla 27. Datos de los porcentajes de células de distintas poblaciones de CMSP, obtenidas de pacientes con EM con al menos una variante rara en *RGS1* o sin ninguna, que expresan RGS1. *DE*: desviación estándar.

4.4. Resultados relacionados con el objetivo 4

“Validar y caracterizar experimentalmente las variantes genéticas estructurales detectadas en la resecuenciación de ADN e identificar nuevos individuos portadores”

Dado que en el presente estudio sólo hemos reportado la identificación de una variante estructural, ésta será la única diana del cuarto y último objetivo. Con el fin de poder realizar estudios adicionales, se trató de identificar más individuos portadores de la misma.

4.4.1. Validación experimental de la delección identificada en el gen *FCRL1* en un paciente con EM

El paciente con EM 1295 fue aquél identificado *in silico* como portador de una delección heterocigota que abarcaba del exón 4 al exón 10, ambos incluidos, del gen *FCRL1*. Para validar experimentalmente este hallazgo, se procedió a evaluar el número de copias de los exones 4, 9 y 10 presentes en el paciente 1295 y en otros dos pacientes con EM de la cohorte de resecuenciación utilizados como controles negativos. También se analizó el número de copias del exón 3 para confirmar su presencia en el paciente 1295. Para ello, se emplearon ensayos de número de copias de TaqMan™ y se utilizó la técnica *real-time* PCR.

Como resultado, se confirmó que el paciente 1295 era portador de tan solo una copia de los exones 4, 9 y 10, mientras que tenía las dos copias del exón 3 del gen *FCRL1*; por el contrario, los dos pacientes utilizados como controles negativos portaban las dos copias de los cuatro exones testados (Figura 50).

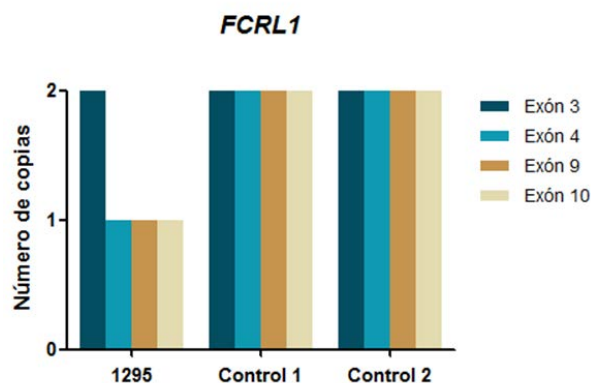


Figura 50. Diagrama de barras representativo de los resultados de la validación experimental de la presencia o ausencia de la delección en heterocigosis de los exones 4-10 del gen *FCRL1*.

4.4.2. Definición de los límites exactos de la delección del gen *FCRL1*

Con el objetivo de delimitar con precisión la localización genómica de la delección identificada, se realizó una amplificación del ADN de la región que la englobaba y se secuenció el producto obtenido para obtener las coordenadas. Curiosamente, se observó que no se trataba de una sola delección, sino de dos diferentes separadas por un pequeño fragmento de ADN de 105 pb (Figura 51). Las coordenadas de las dos delecciones son las siguientes (versión GRCh37/hg19 del genoma humano):

1. chr1:157763319-157768450

2. chr1:157768555-157773252

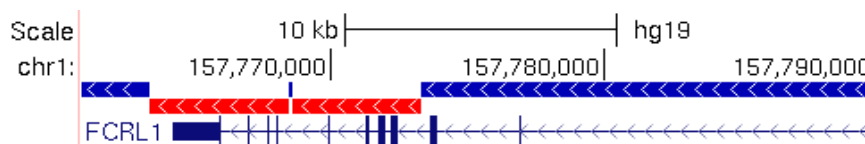


Figura 51. Esquema de la localización exacta en el gen *FCRL1* de la delección identificada en un individuo en el proceso de resecuenciación de ADN. Los dos fragmentos rojos, separados por un pequeño fragmento azul, representan las fracciones delecionadas, mientras que los fragmentos azules representan las fracciones genómica presentes. Las flechas blancas sobre estos fragmentos indican la dirección de la transcripción y, por tanto, el orden de los exones.

4.4.3. Búsqueda de nuevos individuos portadores de la delección del gen *FCRL1*

Una vez conocidas las coordenadas de las delecciones, el objetivo fue detectar nuevos individuos portadores para tratar de asociarla con la EM y para poder realizar estudios adicionales. Para ello, se diseñó un ensayo TaqMan™ constituido por dos *primers* y una sonda que hibridaban específicamente en la zona flanqueante de las delecciones, de modo que sólo se produjo amplificación del ADN genómico en los individuos portadores de las mismas, dado que en aquéllos con el gen intacto se producía un fallo en el proceso al tratarse de un amplicón tan grande.

Como resultado de analizar 256 pacientes con EM y 269 controles sanos, no se identificó ningún individuo adicional portador de las delecciones identificadas en el gen *FCRL1*.

5. DISCUSIÓN

5. Discusión

Durante los últimos años, numerosos estudios de asociación del genoma completo, o GWAS, han contribuido de forma significativa a la caracterización del componente genético de la EM. Estos estudios se emplean para evaluar la posible asociación entre variantes genéticas comunes distribuidas a lo largo de todo el genoma y el riesgo de desarrollar una enfermedad (289), así como para vincular variantes comunes con un determinado rasgo fenotípico o diferentes aspectos de una patología. Para ello, se requiere un gran número de casos y controles, por lo que es necesaria la formación de grandes consorcios y la participación de numerosos centros y países (290). Los GWAS utilizan *arrays* de genotipado prediseñados (SNP chips) que generalmente contienen desde 200.000 hasta más de 2.000.000 de SNPs (291). El diseño de los *arrays* tiene en cuenta el concepto de LD con la finalidad de proporcionar cobertura de la totalidad del genoma. Por tanto, se seleccionan una serie de SNPs, denominados *tag* SNPs, representativos de todos aquéllos en alto LD. Como consecuencia de trabajar con una selección, los polimorfismos cuyas frecuencias alélicas difieren entre los grupos bajo análisis podrían no ser la variante causal de la asociación, por lo que si se procediese a realizar estudios adicionales habría que considerar también aquéllas variantes del mismo bloque de LD o bloque haplotípico para captar la posible implicación funcional del mismo.

En el ámbito de la EM, hace 15 años se consolidó un consorcio dedicado a elucidar la arquitectura genética de la EM: el *International Multiple Sclerosis Genetics Consortium* (IMSGC). Además, varios grupos también han centrado su investigación en la identificación de nuevas variantes genéticas asociadas con la EM o con una de las diferentes formas clínicas de la enfermedad. Sin embargo, las variantes identificadas hasta la fecha, teniendo en cuenta la región del MHC, sólo explican el ~48% de la heredabilidad de la EM (258). Además, para la mayoría de los genes asociados de forma sólida con la EM, todavía no se han identificado las variantes genéticas causales de la enfermedad ni los mecanismos patogénicos por los cuales dichos genes se asocian con un mayor riesgo de desarrollar EM. Estos aspectos configuraron dos de los objetivos principales de la presente tesis doctoral: (i) detectar nuevas variantes genéticas asociadas con la susceptibilidad de desarrollar EM e (ii) identificar las variantes genéticas causales de dichas asociaciones mediante estudios adicionales.

Para la identificación de nuevas variantes genéticas comunes asociadas con el riesgo de EM se utilizó como abordaje la resecuenciación dirigida de 14 genes candidatos en una cohorte formada por 524 pacientes con EM (formas de inicio con brote: RR y SP) y 546 controles sanos.

Discusión

La aplicación de la secuenciación de ADN, ya sea de su totalidad o de regiones específicas, se ha utilizado ampliamente para identificar variantes asociadas a diferentes enfermedades autoinmunes, como el lupus eritematoso sistémico (292), la artritis reumatoide (293) o la propia EM (294). De las 32 variantes genéticas comunes ($MAF \geq 1\%$ en nuestra cohorte) identificadas en nuestro estudio como diferencialmente distribuidas entre pacientes con EM y controles sanos ($p < 0,05$), seleccionamos 9 para una validación mediante genotipado en una cohorte independiente, así como otras 9 adicionales procedentes del último estudio del IMMSGC (258) para abarcar la posibilidad de realizar una regresión logística condicional en el caso de que alguno de nuestros SNPs se validase; así, se puede determinar si se trata de asociaciones independientes o no. Es importante destacar que no se restringió la inclusión de los pacientes con EM por su forma clínica de la enfermedad a la hora de diseñar el genotipado, de modo que en este caso sí hay individuos afectados por formas PP y PR, además de RR y SP. Quizás es llamativo el hecho de que en la resecuenciación sólo participasen pacientes con formas de inicio con brote y la validación se realizase en una cohorte que incluye un grupo de pacientes parcialmente diferente. La razón de este abordaje se fundamenta en el diseño experimental de los GWAS desarrollados periódicamente por el IMMSGC, donde incluyen todo tipo de pacientes con EM, sin exclusión por su forma clínica. Nosotros decidimos crear una cohorte más homogénea para la fase de *screening* con el fin de capturar variantes genéticas que quizás el IMMSGC no detectó como consecuencia de la mayor heterogeneidad de su grupo de pacientes. Una vez capturadas, quisimos validarlas en una cohorte que ya sí estuviese constituida por toda la población de pacientes con EM. La limitación más arriesgada de este diseño fue la posibilidad de que una variante sí esté asociada con las formas de inicio con brote y realmente no lo esté con las formas progresivas y, por tanto, no consiguiésemos validarla debido a la pérdida de un valor de odds ratio (OR) relativamente alejado de 1.

A pesar de que se incluyeron 9 polimorfismos asociados con el riesgo de EM en la fase de resecuenciación, sólo 3 de ellos consiguieron superar la fase de genotipado. Estos resultados parecen indicar que la asociación de las otras 6 variantes con la EM podría comprender falsos positivos, aunque no descartamos que pueda tratarse de una falta de poder estadístico asociada a la cohorte de validación. De hecho, alguna variante extraída del análisis del IMMSGC tampoco sobrevivió a la fase de validación de nuestro estudio, lo que hace más plausible la última explicación que proponemos. En resumen, todo ello evidencia la importancia de emplear técnicas independientes de validación y de trabajar con el mayor tamaño de muestra posible en estudios genéticos.

De las 3 variantes genéticas validadas, una de ellas (rs1599932, *TSM*) se localizaba formando parte de un gran bloque de LD constituido por 9 genes que ya estaba asociado con la EM y caracterizado por otros autores (287), por lo que decidimos focalizar nuestra investigación hacia

las otras dos variantes para evitar reportar resultados redundantes. A pesar de que precisamente la variante rs1599932 (*TSFM*) era la única exónica de las tres, la mayoría de los SNPs asociados con el riesgo de EM se localizan fuera de regiones codificantes, hecho que no resta valor a su potencial efecto funcional; de hecho, muchos de ellos residen en el promotor o en regiones *enhancer*. Por tanto, nos propusimos profundizar en el estudio de las otras dos variantes genéticas validadas: rs10892307 (*CXCR5*) y rs2762943 (*CYP24A1*). En este contexto, cabe destacar que los análisis de regresión logística condicional de los SNPs del gen *CXCR5* en nuestra cohorte indicaron cierta dependencia entre las variantes rs6589706, asociada con el riesgo de EM por el IMSGC, y rs10892307, asociada con la susceptibilidad a la EM en nuestro estudio. Sin embargo, un aspecto que confiere valor a nuestra variante es que al condicionar el análisis al SNP rs10892307, la señal asociada a rs6589706 desaparece, pero al condicionar al SNP rs6589706, la asociación del SNP rs10892307 se mantiene, por lo que rs10892307 es la variante que mejor explica la asociación con el riesgo de EM. No sólo eso, sino que antes de realizar la regresión logística condicional, los valores de p y del OR ya sugerían la mayor potencia del SNP rs10892307. Curiosamente, la dirección de la asociación, indicada por el OR, es diferente entre ambos SNPs (rs6589706: OR=1,14; rs10892307: OR=0,77). La explicación podría ser que el alelo de riesgo en un caso es el minoritario, mientras que en el otro caso es el mayoritario. Por último, cabe mencionar que en poblaciones europeas el valor de r^2 entre los SNPs rs6589706 y rs10892307 es 0.158, cifra que evidencia la imposibilidad de afirmar un alto LD entre ellos.

➤ Efecto funcional de la variante rs10892307 (*CXCR5*)

El polimorfismo rs10892307 se localiza en la región UTR5' del gen *CXCR5* (*C-X-C Motif Chemokine Receptor 5*), que codifica un receptor de quimiocinas de la familia CXC localizado en la cara externa de la membrana plasmática. Bajo condiciones fisiológicas, este receptor se expresa altamente en células B maduras y es un antígeno de superficie clave para definir a las poblaciones de células T colaboradoras foliculares (Tfh, por *T follicular helper cells*), tanto a las localizadas en los órganos linfoides secundarios, como a las circulantes (89). El ligando de *CXCR5* es la quimiocina *CXCL13* (también conocido como BLC, por *B lymphocyte chemoattractant*), que se expresa de forma constitutiva en los órganos linfoides secundarios y que atrae a las células que expresan *CXCR5* en su superficie (295).

Diversos estudios con el modelo experimental de la EM, la EAE, han propuesto una asociación entre *CXCL13* y la EM. En uno de ellos, se observó una sobreexpresión génica de *CXCL13* en el SNC, así como la presencia de células *CXCL13*⁺ incluidas en grandes agregados anormales de células B formados en las meninges con organización de centro germinal (296). Otro de ellos

detectó ARNm y proteína de CXCL13 en la médula espinal de ratones con EAE, pero no en los controles *naïve* (297). En este último estudio, Bagaeva y col. también observaron que los ratones con EAE deficientes en CXCL13 exhibían un curso más leve de la enfermedad y que CXCL13 parecía ser fundamental para el establecimiento de lesiones crónicas de la sustancia blanca. También se han desarrollado estudios de pacientes con EM que vinculan al eje CXCR5-CXCL13 con la enfermedad. Uno de ellos reveló una sobreexpresión génica de *CXCL13* en áreas de lesiones desmielinizantes activas de pacientes con EM, la presencia de la proteína CXCL13 en el infiltrado perivascular y en las células infiltradas en el parénquima, y niveles elevados de CXCL13 en el LCR de pacientes con EMRR (298). Además, la mayoría de las células B presentes en el LCR de pacientes con EM son positivas para CXCR5 (299). La asociación de CXCR5 con la EM también se fundamenta en datos genéticos. La primera vez que se asoció al gen *CXCR5* con el riesgo de EM fue a través del polimorfismo rs630923 (C/A), situado en su promotor, cuyo alelo común (C) se relacionó con un mayor riesgo de EM. Esta asociación se identificó en el 2011, pero no alcanzaba el nivel de significación de GWAS (226), motivo por el cual en el 2013 el IMSSG trató de validarlo mediante genotipado en una cohorte independiente de ~20000 individuos y un posterior meta-análisis de los datos de los dos estudios (300). La combinación de los resultados evidenció su asociación superando el nivel de significación de GWAS. A partir de este momento, se han identificado más variantes asociadas a *CXCR5* vinculadas con el riesgo de desarrollar EM (255,258), lo que fortalece aún más la relación de este gen con la enfermedad.

La bien documentada asociación del gen *CXCR5* con la EM suscitó nuestro interés por descifrar si el alelo minoritario del polimorfismo rs10892307 podría afectar a la regulación de su expresión y, quizás, alterar así la capacidad migratoria de las células de los individuos portadores de la variante. En primer lugar, decidimos diseñar un ensayo reportero dual-luciferasa para evaluar su posible efecto a nivel transcripcional. Pero, debido a que mediante análisis bioinformático identificamos otros tres polimorfismos en alto LD ($r^2 > 0,5$) con el mismo en poblaciones europeas y, según los datos de ENCODE, todos ellos se situaban en regiones *enhancer*, no podíamos concluir que nuestro polimorfismo era la variante causal de la asociación de esta región con la EM. Por ello, si sólo se realizase el ensayo para estudiar este polimorfismo, posiblemente no se encontrase ningún efecto funcional y simplemente se haya asociado con la EM por su alto LD con el SNP causal real. Por tanto, con el objetivo de identificar la variante genética causal de la asociación de rs10892307 con la EM, optamos por realizar el ensayo reportero dual-luciferasa sobre los cuatro polimorfismos en alto LD: rs55756957, rs11602393, rs10892307 y rs3176905. Así, pudimos medir la activación del promotor influenciada por cuatro fragmentos, cada uno de ellos portador de uno de los polimorfismos bajo estudio, clonados de forma independiente *upstream* del promotor mínimo

presente en el vector pGL4.23 utilizado. Para garantizar que el efecto observado era debido exclusivamente a un polimorfismo, nos aseguramos de que en el resto de cada fragmento clonado no existía ninguna otra mutación que pudiese afectar a la actividad transcripcional. La elección de esta metodología para tratar de hallar una variante causal se realizó en base a la ventaja que supone un sistema que ha sido modificado para eliminar la mayoría de los sitios de unión de factores de transcripción y los elementos reguladores crípticos, que son secuencias inactivas en su localización genómica nativa pero que adquieren capacidad funcional si se sitúan adyacentes a algún gen mediante reorganización del ADN (301). El vector pGL4.23 simplemente posee un promotor mínimo con los elementos básicos para que ocurra la transcripción y es inactivo a menos que existan elementos reguladores situados *upstream* capaces de incrementar la actividad del promotor, lo que reduce el *background* y el riesgo de artefactos en nuestros resultados.

En el presente estudio, observamos que el fragmento que incluía la región del polimorfismo rs11602393 (T/A) incrementaba la actividad transcripcional del promotor en orientación *reverse*, pero este incremento era similar entre el fragmento portador del alelo de referencia (T) y el portador del alelo minoritario (A). Sin embargo, resultó curioso que en orientación *forward* el alelo de referencia de este SNP no tenía ningún efecto sobre la actividad transcripcional del promotor, pero el alelo minoritario sí mostraba actividad *enhancer*. La diferencia estadísticamente significativa entre la actividad de la luciferasa bajo el control del alelo de referencia y dicha actividad bajo el control del alelo minoritario del SNP rs11602393 en orientación *forward*, unida al hecho de que no se observó ningún efecto asociado al resto de polimorfismos analizados, nos condujo a hipotetizar que el polimorfismo rs11602393 es la variante causal de la asociación del SNP rs10892307 con la EM identificada en nuestro estudio. Esta teoría cobra fuerza si recordamos que el LD entre los polimorfismos rs11602393 y rs10892307 es prácticamente total en poblaciones europeas ($r^2=0,94$; $D'=0,97$), un hecho que realza la fuerte evidencia histórica de su baja frecuencia de recombinación. Posiblemente resulte paradójico que sea rs11602393 la variante que altera la regulación transcripcional cuando ésta se localiza a -3684 pb del sitio de inicio de la transcripción (TSS, por *transcription start site*) y no lo sea rs10892307, localizada en la región UTR5' a -2 pb del TSS. La explicación es que el ADN se estructura formando *loops* que ocasionan que un *enhancer* y un promotor lejanos entre sí se aproximen, lo que permite la interacción directa entre las proteínas activadoras unidas al *enhancer* y aquellas unidas al promotor formando parte del complejo de inicio de la transcripción (302). Cabe mencionar que, fundamentado en los datos de ENCODE, el polimorfismo rs55756957 (A/G) parecía el más prometedor a la hora de detectar un efecto funcional vinculado a su alelo minoritario, ya que es del que existía una mayor evidencia

Discusión

experimental respaldando que su localización cromosómica era un sitio de unión de factores de transcripción.

En este ámbito, otros autores ya han realizado ensayos con el gen reportero luciferasa para estudiar polimorfismos asociados con el riesgo de EM (287). De hecho, en uno de ellos, Ryu y col. estudiaron la actividad transcripcional del primer polimorfismo de *CXCR5* asociado con la susceptibilidad de desarrollar EM y mencionado previamente (rs630923), pero no observaron ninguna diferencia en la actividad del promotor entre el alelo de referencia y el minoritario (303). Más adelante, Mitkin y col. también utilizaron un ensayo de luciferasa para analizar la modulación de la actividad del promotor de *CXCR5* influenciada por el SNP rs630923 (304). En este caso, observaron que el alelo minoritario (A) creaba un sitio de unión del factor de transcripción MEF2C que suponía una disminución de la actividad transcripcional en células Raji activadas con LPS, PMA e ionomicina.

Basándonos en los resultados obtenidos en los ensayos reporteros dual-luciferasa, quisimos ahondar en el posible efecto del SNP rs11602393 realizando estudios adicionales con muestras de CMSP de pacientes con EM. Se consideró oportuno que todos los individuos incluidos en estos estudios fueran pacientes con EM y no controles sanos porque, ya que los pacientes con EM sufren alteraciones de algunas subpoblaciones de CMSP, de este modo la única diferencia entre ellos sería el genotipo. Como criterio de selección de los pacientes portadores y no portadores del alelo minoritario de la variante rs11602393 utilizamos su genotipo para la variante rs10892307 debido a que ésta es la que genotipamos y, por tanto, de la que disponemos información para cada individuo incluido en el estudio. Además, el alto LD entre ambas variantes supone que sean *proxy* SNPs (305). La determinación de los niveles de expresión génica de *CXCR5* en CMSP de pacientes con EM con presencia y ausencia del alelo minoritario de rs11602393 mediante *real-time* PCR no reveló diferencias entre ambos grupos. Este resultado pudo ser debido por un lado a la falta de poder estadístico, ya que es muy complicado identificar una eQTL partiendo de un tamaño de muestra pequeño, sobre todo cuando se trata de un efecto moderado como el que observamos mediante el ensayo de luciferasa. Por otro lado, pudo deberse a que el efecto del SNP sobre la actividad transcripcional del promotor de *CXCR5* es específico de un tipo celular concreto que podría quedar diluido en el conjunto de todas las CMSP. Esta última teoría ganó veracidad cuando estudiamos la expresión del receptor *CXCR5* en la membrana plasmática de diferentes subpoblaciones de linfocitos T y B y observamos un aumento significativo del porcentaje de células T reguladoras que expresaban *CXCR5* en su superficie en los pacientes con EM portadores del alelo minoritario de rs11602393 (A) respecto a aquéllos sin él. Esto implicaría que el alelo común de esta variante (T), asociado con un mayor riesgo de EM, acarrea una disminución de las células T reguladoras *CXCR5*⁺ respecto al alelo minoritario (A).

Las células T reguladoras que expresan CXCR5⁺ se identificaron recientemente y se definieron como células T reguladoras foliculares (Tfr). La presencia de CXCR5 en su superficie les confiere la capacidad de migrar a los centros germinales (CGs) mediante quimioatracción hacia CXCL13 (306). En ellos tiene lugar la reacción del CG, donde las células B previamente activadas por la presentación antigénica mediada por las células dendríticas, experimentan varias rondas de hipermutación somática, expansión clonal, maduración de afinidad y conmutación isotópica para diferenciarse a células B memoria o a células plasmáticas productoras de anticuerpos específicos (85). Esta reacción está arbitrada por las células Tfh mediante señales coestimuladoras directas y secreción de citoquinas (307), pero también requiere una regulación adicional muy precisa para evitar la producción de autoanticuerpos y el desarrollo de enfermedades autoinmunes desempeñada por las células Tfr, que suprimen a las células B foliculares e inhiben el ciclo celular de las Tfh para bloquear la producción de anticuerpos mediada por ellas (308). Gracias a que las células Tfr no sólo se encuentran en los órganos linfoides, sino también en la circulación (309), y a que se ha demostrado que la capacidad supresora *in vitro* de ambas poblaciones es comparable en humanos (92), en nuestro estudio consideramos que el análisis de la población circulante de células Tfr constituía un recurso representativo y válido para evaluar las alteraciones de los procesos que ocurren en los CGs. La única diferencia es que las células Tfr residentes muestran un rol efector, mientras que las circulantes tienen un fenotipo de memoria y necesitan ser reactivadas por CDs mediante presentación antigénica para dirigirse a los órganos linfoides secundarios e interactuar con las células Tfh y con las células B, proporcionando una respuesta rápida ante una reexposición antigénica (310). En nuestro estudio, la disminución significativa del porcentaje de células T reguladoras CXCR5⁺ circulantes observada en los pacientes con EM portadores del alelo de riesgo del SNP rs11602393 (T) implicaría una menor proporción de células reguladoras con capacidad migratoria hacia el CG. Este hecho podría ocasionar una inhibición defectuosa de la reacción del CG, con el consiguiente descontrol de la activación de las células B y de la producción de anticuerpos. En experimentos con ratones, se ha observado que la ausencia de células Tfr supone una excrecencia de clones de células B antígeno-inespecíficas en los centros germinales (311). Este dato es especialmente interesante en el campo de las enfermedades autoinmunes, como la EM. De hecho, aunque la cantidad de estudios sobre Tfr circulantes en enfermedades autoinmunes es limitada, existe alguna publicación que sugiere que la desregulación de las Tfr podría ocasionar alguna de ellas. Por ejemplo, Ma y col. detectaron una disminución del número de células Tfr circulantes en pacientes con lupus eritematoso sistémico respecto a controles sanos (312). De forma similar, los resultados de Zheng y col. mostraron una reducción en el número de células Tfr circulantes en pacientes con colangitis biliar primaria, una enfermedad autoinmune que causa daño hepático, respecto a controles sanos (313). En el ámbito de la EM, Dhaeze y col. han publicado recientemente que tanto la frecuencia como la

capacidad supresora de las células Tfr circulantes están disminuidas en los pacientes con EM cuando se comparan con individuos sanos (92). Según nuestros resultados, la frecuencia de estas células sería especialmente baja en aquellos pacientes con EM portadores del alelo de riesgo de la variante rs11602393 (T), lo que nos invita a teorizar que el alelo de riesgo se asocia con un menor limitación de la reacción del GC por parte de las Tfr y, por tanto, con una mayor probabilidad de que se generen anticuerpos contra antígenos propios.

➤ **Implicación funcional de la variante rs2762943 (CYP24A1)**

El polimorfismo rs2762943 se localiza en la región UTR5' del gen *CYP24A1* (*Cytochrome P450 Family 24 Subfamily A Member 1*). Éste codifica una enzima involucrada en la ruta metabólica de la vitamina D. La vitamina D₃ se obtiene a través de la dieta o se sintetiza en la piel a partir del 7-dehidrocolesterol mediante la acción de la luz UV. A partir de aquí, es transportada a través de la sangre unida a la proteína transportadora DBP (por *vitamin D binding protein*) hacia el hígado. En él, experimenta la primera hidroxilación, que tiene lugar en el C-25 y es catalizada por alguna enzima con actividad 25-hidroxilasa, de las que *CYP2R1* es la más importante (314), dando lugar a la 25-hidroxivitamina-D₃ o calcidiol (25(OH)D). Ésta es transportada a través de la sangre, también unida a la DBP, hacia el riñón, donde es internalizada mediante endocitosis y experimenta la segunda hidroxilación, mediada por la enzima *CYP27B1*, dando lugar a la 1,25-dihidroxivitamina-D₃ o calcitriol (1,25(OH)₂-D), la forma biológicamente activa de la vitamina D. Ésta ejerce su acción mediante la unión al receptor nuclear VDR (por *vitamin D receptor*). La enzima *CYP24A1* actúa en el riñón bien sobre la forma 25(OH)D o bien sobre la 1,25(OH)₂-D. Su función consiste en degradar ambas formas de vitamina D mediante la incorporación de un grupo OH en el C-24, lo que en última instancia limita la cantidad del metabolito fisiológicamente activo por un lado mediante la producción de 24,25-(OH)₂-D, disminuyendo así la cantidad del precursor 25(OH)D disponible para dar lugar a la forma activa y, por otro, degradándola directamente a 1,24,25-(OH)₃-D (315).

Es ampliamente conocido que la vitamina D juega un papel importante no sólo en la regulación de la homeostasis del calcio, sino también en la modulación del sistema inmune (316). Numerosos estudios avalan el vínculo entre los niveles disminuidos de la vitamina D y diversas enfermedades autoinmunes, como la diabetes tipo 1 (317), el lupus eritematoso sistémico (318), la artritis reumatoide (319) o la propia EM. Además, se ha publicado que la deficiencia de vitamina D en suero está relacionada con un mayor riesgo de sufrir un brote en pacientes con EMRR (176,320). Pero la vitamina D no sólo contribuye al riesgo de EM como factor ambiental, sino también como factor genético. Se han publicado varias investigaciones que asocian polimorfismos en el gen que codifica el receptor de la vitamina D (VDR) y en genes

que participan en su ruta metabólica con el riesgo de desarrollar EM. Respecto al receptor, diversos estudios han reportado asociaciones directas del alelo minoritario de polimorfismos del gen *VDR* con el riesgo de EM, como los SNPs rs2228570 (321–323) y rs731236 (324). Para el gen *CYP27B1*, el alelo minoritario de las variantes rs12368653, rs10876994, rs703842 (247) y rs118204009 (269), se ha asociado de forma significativa con el riesgo de EM mediante GWAS o secuenciación. En cuanto al gen *CYP24A1*, la primera vez que un estudio lo señaló como gen de riesgo para la EM fue mediante la asociación del alelo minoritario del SNP rs2248359 en poblaciones caucásicas (226), a pesar de que otros autores ya lo habían intentado antes sin éxito (322). Una vez publicada la asociación de este SNP, otro grupo de investigadores reportó que el alelo de riesgo de este polimorfismo (C) estaba fuertemente asociado con una expresión aumentada del gen *CYP24A1* en el lóbulo frontal, aunque no en la materia blanca (325). Sin embargo, más adelante, Karaky y col. trataron de correlacionar la expresión génica de *CYP24A1* con la variante rs2248359 y con todas aquellas situadas 1Mb alrededor de *CYP24A1* en una línea celular linfoblastoide y obtuvieron un resultado negativo (326). Hasta donde sabemos, no se han realizado más estudios funcionales con polimorfismos del gen *CYP24A1* en el terreno de la EM, pero sí se han asociado nuevos polimorfismos con la enfermedad, tales como rs2248137, rs6068816, rs2585447 y rs209967 (258). Cabe resaltar que los dos primeros conseguimos identificarlos en la fase de resecuenciación de nuestro estudio. Estas nuevas asociaciones consolidan aún más el vínculo genético entre *CYP24A1* y la EM, lo que incrementó nuestro interés por reportar alguna variante nueva.

Como se mencionó previamente, en el presente estudio conseguimos identificar y validar que el alelo minoritario del SNP rs2762943 (T) incrementa el riesgo de desarrollar EM. Este polimorfismo se sitúa a -672 pb del TSS del gen *CYP24A1* y, según los datos del Proyecto ENCODE, probablemente altere la actividad *enhancer* del fragmento donde se localiza. Este dato nos llevó a conjeturar que el genotipo de este *locus* podría incrementar el riesgo de EM mediante la regulación de la expresión de su gen asociado: *CYP24A1*. Además, mediante búsqueda bioinformática observamos que, teniendo en cuenta un umbral de $r^2=0,5$, este SNP no se encuentra en alto LD con ningún otro en poblaciones europeas, por lo que en el caso de llevar a cabo algún estudio funcional exitoso, él sería la variante causal del efecto encontrado. A este dato se suma el hecho de que en la fase de genotipado no conseguimos validar la asociación con la EM de ningún otro SNP de los incluidos. Toda esta información constituyó un incentivo para proceder a realizar estudios funcionales con el objetivo de perfilar más la asociación. En primer lugar, la determinación de los niveles de vitamina 25(OH)D y de vitamina 1,25(OH)₂-D en el suero de pacientes con EM con el alelo de riesgo del SNP rs2762943 y sin él reveló una menor concentración de la forma 1,25(OH)₂-D, la fisiológicamente activa, en aquellos pacientes portadores del alelo minoritario, aunque los niveles de 25(OH)D resultaron similares entre los

Discusión

dos grupos de pacientes. Cabe aclarar que, con el fin de evitar resultados falsamente positivos influenciados por la variación estacional de los niveles de vitamina D, se incluyeron el mismo porcentaje de muestras de suero obtenidas en primavera+verano y en otoño+invierno en ambos grupos de pacientes. Como primera impresión, puede resultar curioso que no hayamos encontrado diferencias en la determinación de la forma precursora, la vitamina 25(OH)D, ya que es la utilizada en clínica para evaluar el estado de vitamina D de un paciente al representar las reservas de la misma (327). De hecho, es la que siempre se reporta como disminuida en pacientes con EM. Una de las posibles explicaciones de este hallazgo es que el sustrato predilecto de la enzima *CYP24A1* es la vitamina 1,25(OH)₂-D (328) y, por tanto, es en este metabolito donde existen más posibilidades de observar un efecto cuando éste es moderado. Otra justificación posible es que, dado que la forma 1,25(OH)₂-D se determina en pg/mL y la 25(OH)D en ng/mL, el rango de referencia de la forma activa es más estrecho y sería más fácil detectar una diferencia moderada.

Estos resultados nos invitaron a teorizar que la causa de la disminución de los niveles en suero de la forma activa de la vitamina D en los pacientes con el alelo de riesgo podría ser una mayor actividad degradadora de la enzima *CYP24A1* conferida por dicho polimorfismo. Para evaluar esta hipótesis, determinamos mediante *real-time* PCR los niveles de expresión génica basal de *CYP24A1* en CMSP de pacientes con EM con presencia y ausencia del alelo de riesgo de rs2762943, pero no detectamos diferencias. Debido a que observamos que se trata de un gen con una baja expresión basal en CMSP, decidimos evaluar su respuesta tras estimular a las células con 1,25(OH)₂-D₃, ya que está descrita su capacidad de incrementar la expresión de *CYP24A1* para mantener su propia homeostasis y evitar la hipercalcemia (329,330). Nuestros resultados evidenciaron niveles de ARNm de *CYP24A1* cuatro veces mayores en los pacientes con el alelo de riesgo del SNP rs2762943 respecto a aquéllos sin él tras haber recibido el tratamiento con 1,25(OH)₂-D₃, una cifra considerable y que permite intuir una diferencia entre grupos, pero que no condujo a alcanzar la significancia estadística, posiblemente debido al tamaño de muestra limitado o al alto error estándar de la media del grupo sin el alelo de riesgo. Aun así, pensamos que una estimulación de la expresión génica de *CYP24A1* cuatro veces superior en el grupo con el alelo de riesgo concuerda con los niveles de 1,25(OH)₂-D, su molécula diana, reducidos en el suero de dicho grupo de pacientes y que podría tratarse de un indicio de que el alelo minoritario sí envuelve algún efecto funcional. Dado que un exceso de la enzima *CYP24A1* podría causar una deficiencia de la forma activa de la vitamina D y, por tanto, inactivar la ruta de señalización de la misma, sería interesante investigar la expresión génica de algún gen en cuyo promotor exista un elemento de respuesta a la vitamina D (VDRE, por *Vitamin D response element*), lo que significa que dicho gen es regulado por la unión del complejo VDR/RXR (*retinoid X receptor*) a dicho elemento, con la previa activación del VDR mediada por la 1,25(OH)₂-D₃.

Uno de los candidatos más atrayentes es el alelo HLA DRB1*1501, altamente asociado con el riesgo de EM y portador de un VDRE en su promotor (331).

Toda la investigación de la presente tesis en relación con la vitamina D tiene como último objetivo aportar nueva información sobre el posible mecanismo de acción de la vitamina D para ejercer su efecto inmunomodulador, lo cual sería interesante para ayudar a identificar nuevas dianas terapéuticas.

➤ **Implicación de las variantes raras en la EM**

En esta tesis doctoral, proponemos una posible contribución de las variantes genéticas raras a la base genética que subyace a la EM. Para ello, se consideró a una variante genética como rara cuando la frecuencia de su alelo minoritario era inferior al 0,5% en la cohorte de resecuenciación. Dado que la única variante genética estructural reportada en la presente tesis sólo se identificó en un paciente con EM, puede incluirse en esta categoría. Se trata de la delección en heterocigosis de los exones 4 al 10, con un pequeño fragmento de ~100 pb no delecionado entre medias, del gen *FCRL1* (*Fc receptor like 1*). Con los datos que obtuvimos en la resecuenciación, no fue posible establecer una asociación entre esta variante y la EM. De hecho, los únicos aspectos que apoyan que esta delección esté relacionada con el riesgo de EM son la asociación genética publicada entre el gen *FCRL1* y la enfermedad (258) y la posible importancia de su función biológica en la patogénesis de la EM, dado que se ha descrito una expresión génica aumentada de *FCRL1* en células B periféricas de pacientes con EM y con otras enfermedades autoinmunes respecto a controles sanos (332). También se ha observado que los pacientes con EM con baja neurodegeneración evidenciada mediante MRI mostraban niveles de ARNm de *FCRL1* en CMSP aumentados respecto a aquéllos con alta neurodegeneración (333). Nosotros, para poder aportar alguna evidencia funcional adicional, necesitábamos identificar más individuos portadores de la delección sujetos de futuros estudios; no obstante, a pesar de realizar un *screening* con muestras de ADN genómico de individuos no incluidos en la cohorte de resecuenciación, no logramos identificar ningún individuo portador nuevo.

La baja frecuencia de las variantes raras en la población pone en evidencia la dificultad de encontrar más de un individuo portador de una variante rara concreta y, por tanto, de realizar estudios funcionales para elucidar su posible implicación en un proceso fisiopatológico. Otro factor limitante derivado de su baja frecuencia es la imposibilidad de obtener resultados estadísticamente significativos de asociación con una característica fenotípica o con una enfermedad. Por todo ello, para poder evaluar de algún modo el impacto de las variantes raras en la EM, decidimos trabajar con variaciones de un solo nucleótido y llevar a cabo una

estrategia basada en la identificación de regiones genómicas enriquecidas en variantes raras de forma más frecuente bien en el grupo de pacientes, o bien en el de controles sanos (334). Una vez identificadas, decidimos dirigir el trabajo hacia aquellos genes con acumulación de variantes raras en pacientes con EM. Con el objetivo de esclarecer su posible repercusión a nivel funcional, se determinaron los niveles de ARNm en CMSP de los cuatro genes que resultaron enriquecidos en variantes raras en alguna región concreta en casos con EM: *FCRL1*, *TRAF3*, *CYP24A1* y *RGS1*. Respecto a los tres primeros, la ausencia de diferencias en la expresión génica en CMSP entre pacientes con y sin variantes raras acumuladas en regiones específicas no resta relevancia al hallazgo ni descarta su posible papel en la patogénesis de la EM, ya que el patrón transcripcional difiere entre tejidos incluso más que entre individuos (335). Algunas variantes genéticas podrían ser capaces de alterar la expresión génica de su gen asociado no en CMSP, pero sí en otros tejidos afectados en la EM, como el cerebro o algún órgano del sistema inmune. Sin embargo, lamentablemente, no disponemos de muestras *postmortem* de otros tejidos de pacientes con EM con la región genómica de los genes bajo estudio secuenciada. También contemplamos la posibilidad de no haber detectado un efecto debido a la baja potencia estadística influenciada por dos limitaciones fundamentales: en primer lugar, al tratarse de variantes genéticas raras, no disponíamos de muchos individuos portadores de las mismas que pudiesen incluirse en nuestros estudios; en segundo lugar, la necesidad de trabajar con CMSP extraídas cuando el paciente aún no había recibido tratamiento para la EM también restringía el repertorio de muestras que podíamos incluir.

Respecto a *RGS1*, el hallazgo de una disminución de su expresión génica en pacientes con EM con presencia de variantes raras respecto a aquéllos sin ellas lo situó como diana de los posteriores estudios funcionales, dado que a pesar del pequeño tamaño de muestra del grupo de individuos con variantes raras (n=6), se obtuvieron resultados estadísticamente significativos.

➤ **Consecuencia funcional del enriquecimiento de variantes raras del gen *RGS1***

El gen *RGS1* (*Regulator of G-protein signalling 1*) no sólo está asociado con la EM, sino también con otras enfermedades inflamatorias, como la celiaquía (336), o autoinmunes, como la diabetes tipo 1 (337). Su papel en la patogénesis de la EM se explica en gran parte mediante las asociaciones genéticas publicadas. La primera vez que se desenmascaró al gen *RGS1* como nuevo *locus* con evidencia de asociación con la EM fue en 2009 mediante un meta-análisis de diversos GWAS y una replicación de los resultados en una cohorte independiente. El SNP asociado fue rs2760524, aunque no alcanzó el nivel de significación de GWAS (338). Por ello, el IMSGC se propuso validarlo mediante genotipado en una nueva cohorte y, en este caso, el SNP rs2760524 se asoció sólidamente con la EM superando el nivel de significancia de GWAS

($p=3.55 \times 10^{-9}$) (339). Posteriormente, en un GWAS a gran escala colaborativo que incluyó 9.772 casos con EM y 17376 controles sanos, se identificó otro SNP asociado con *RGS1* que confería susceptibilidad a la EM y en alto LD con el anterior. Se trata del SNP rs1323292 (226), el cual fue validado recientemente en el último meta-análisis del IMSGC (258). Por último, se identificó el SNP rs1359062 como tercer *locus* de *RGS1* asociado con el riesgo de EM mediante la combinación de los datos procedentes de GWAS previos (14802 casos con EM y 26.703 controles sanos) y del ImmunoChip, que incluía 14.498 pacientes con EM y 24.091 controles sanos. Este polimorfismo también estaba en alto LD con los anteriores. Sin embargo, hasta la fecha, no existía evidencia de que las variantes raras de *RGS1* estuvieran asociadas con EM, lo que hace que nuestros resultados sean más novedosos.

El gen *RGS1* se expresa en linfocitos, monocitos, células dendríticas (340) y microglía (341). Cabe destacar que los linfocitos B presentes en los centros germinales de los folículos de los órganos linfoides secundarios expresan niveles especialmente elevados de *RGS1* (342). Este gen codifica una proteína localizada en la cara citosólica de la membrana plasmática que actúa como regulador negativo de la señalización mediada por receptores acoplados a la proteína G (GPCRs, por *G-protein-coupled receptors*). Los GPCRs interactúan con proteínas G heterotriméricas constituidas por una subunidad α , una β y una γ , y se ha descrito su participación en la patogénesis de la EM. *RGS1* se une a la subunidad α activa, unida a GTP, e incrementa la actividad GTPasa de dicha subunidad, aumentando así la velocidad de la conversión de GTP a GDP. Esta reacción conduce a la unión de las subunidades α con heterodímeros constituidos por las subunidades β y γ formando heterotrimeros inactivos, lo que constituye el final de la señalización. El hecho de que todos los receptores de quimiocinas estén acoplados a proteínas G heterotriméricas (343) promovió la investigación del papel de *RGS1* en la respuesta a la estimulación por quimiocinas, lo que permitió descubrir que las células B de ratones *RGS1*^{-/-} respondían de forma extrema a las quimiocinas CXCL12 (ligando del GPCR CXCR4) y CXCL13 (ligando del GPCR CXCR5) (344); por tanto, se deduce que *RGS1* participa en la desensibilización de la señalización mediada por los receptores CXCR4 y CXCR5 en las células B. Interesantemente, el seguimiento de células B *RGS1*^{-/-} y *wild type* transferidas de forma adoptiva a ratones experimentales reveló que la ausencia de *RGS1* fomenta la entrada de las células B en los nódulos linfáticos y aumenta la motilidad de dichas células en los folículos (345). Uno de los principales hallazgos derivados de nuestra investigación fue la disminución significativa del número de células B periféricas que expresan *RGS1* en pacientes con EM con al menos una variante rara de dicho gen respecto a aquéllos sin ellas. Este número reducido de células B circulantes *RGS1*⁺ implica que la presencia de variantes raras se asocie con un mayor número de células B en la periferia con carencia de *RGS1* en la cara interna de su membrana plasmática, lo que resultaría en una entrada más rápida

Discusión

de las mismas a los folículos de los órganos linfoides secundarios, consecuencia compatible con la observación de que las células B *RGS1*^{-/-} están disminuidas en sangre periférica respecto a las células *wild type* en los ratones que experimentaron la transferencia adoptiva explicada previamente (345). En este contexto, nosotros teorizamos que el defecto en la expresión de *RGS1* en las células B de pacientes con EM con variantes genéticas raras de *RGS1* podría facilitar su capacidad de migrar a través de la barrera hematoencefálica (BHE) para acceder al SNC por dos razones fundamentales: i) en primer lugar, parece que la posesión de un genotipo *RGS1*^{-/-} promueve una mayor migración hacia la quimiocina CXCL12, que no sólo se produce en los centros germinales (346), sino que también es secretada por las células residentes del SNC (112); ii) asimismo, el incremento de la motilidad de las células B que no expresan *RGS1* podría asociarse con un fenotipo más activo de dichas células que podría ir acompañado de una sobreexpresión de moléculas de adhesión que faciliten la extravasación al SNC. Además, el potencial aumento del número de células B que ingresan en los centros germinales en los pacientes con EM con variantes raras de *RGS1* incrementa la cantidad de células B susceptibles de experimentar la reacción del centro germinal y de diferenciarse a células B de memoria o a células plasmáticas productoras de anticuerpos de alta afinidad. Debido a que nos estamos refiriendo a pacientes con EM, se trataría de un aumento de la producción de anticuerpos patogénicos.

Otra respuesta celular ante la estimulación por quimiocinas es el incremento de los niveles de calcio intracelular (347). En relación a este aspecto, Moratz y col. determinaron los niveles de Ca²⁺ intracelular en células B foliculares *RGS1*^{-/-} y *wild type* tras su exposición a CXCL12 y observaron un mayor incremento del mismo en las células B *RGS1*^{-/-} respecto a las *wild type* (344), lo que apoya la respuesta exagerada a las quimiocinas de las células B deficientes en *RGS1*. Cabe destacar que también se ha reportado que el incremento de los niveles intracelulares de Ca²⁺ promueve una serie de procesos intracelulares, tales como la migración celular (348) y la adhesión (349), que son cruciales para la determinación del destino de las células B durante la respuesta inmune humoral (350). Además, el aumento del Ca²⁺ intracelular es una señal clave para el desarrollo de una respuesta de células B frente a un antígeno (351). En este contexto, y volviendo a nuestros resultados, las células B de pacientes con EM portadores de variantes genéticas raras de *RGS1* experimentarían un aumento excesivo de los niveles de Ca²⁺ intracelular después de exponerse a la quimiocina CXCL12, lo que podría desencadenar una mayor migración celular, una mayor activación de sus moléculas de adhesión o una respuesta inmune aberrante.

Otro vínculo ya descrito entre el gen *RGS1* y la EM es la capacidad del IFN- β , uno de los tratamientos más prescritos para formas de EM que cursan con brote (352), de incrementar los niveles de ARNm del gen *RGS1* en CMSP, monocitos, linfocitos T y linfocitos B de individuos

sanos tratados *in vitro*. En este mismo estudio, Tran y col. también determinaron los niveles de expresión génica de *RGS1* en CMSP de pacientes con EM que estaban recibiendo tratamiento con IFN- β y reportaron que los niveles de ARNm de *RGS1* aumentaban significativamente a las 4 horas de recibir el tratamiento (288). Basándonos en estos datos, nos interesamos por elucidar si las CMSP de pacientes con EM con ausencia y presencia de variantes genéticas raras asociadas a *RGS1* respondían de la misma forma ante la estimulación *in vitro* con 100UI/mL de IFN- β durante 4 horas. Tal y como cabría esperar, nuestros resultados evidenciaron un incremento significativo de los niveles de expresión génica de *RGS1* en aquellos pacientes sin variantes raras tras haber recibido el tratamiento. Sin embargo, curiosamente, se reveló una falta de inducción de la expresión génica de *RGS1* en las CMSP de los pacientes con EM con variantes raras de dicho gen después de administrarles el IFN- β . La asociación encontrada entre la tenencia de variantes genéticas raras de *RGS1* y la ausencia de su inducción mediada por el tratamiento con IFN- β en CMSP podría llevarnos a especular que los pacientes con EM portadores de al menos una variante rara de *RGS1* no responden correctamente al tratamiento con IFN- β , hecho que ocurre en una proporción relativamente elevada de pacientes con EM (353). A día de hoy, no se conoce con exactitud el motivo por el que se produce esta falta de respuesta al tratamiento con IFN- β . De hecho, a pesar de haberse identificado numerosos efectos beneficiosos desencadenados por el IFN- β , tales como la promoción de un ambiente antiinflamatorio (354) en detrimento de uno proinflamatorio (355), la limitación del tráfico de leucocitos a través de la BHE o la reparación del SNC mediante la producción del factor de crecimiento nervioso (NGF, por *nerve growth factor*) (171), no se conoce por completo su mecanismo de acción.

Nuestros resultados relacionados con gen *RGS1* alegan la necesidad de identificar biomarcadores genéticos, como polimorfismos, que permitan predecir bien el curso de la EM en base al conocimiento de la población celular que podría hallarse más alterada, o bien si un determinado paciente responderá al tratamiento con IFN- β .

6. CONCLUSIONES

6. Conclusiones

- La resecuenciación dirigida de los 14 genes de riesgo para la EM ha permitido identificar 32 variantes genéticas comunes diferencialmente distribuidas entre pacientes con EM y controles sanos, así como enriquecimientos de variantes raras en regiones genómicas específicas asociados al grupo de casos con EM o al de controles sanos.
- Los polimorfismos rs10892307 (*CXCR5*) y rs2762943 (*CYP24A1*) se validaron como variantes asociadas con el riesgo de EM en una cohorte independiente mediante el genotipado de una selección de SNPs.
- El alelo de riesgo de la variante rs11602393 (T), en alto LD con el polimorfismo rs10892307, disminuye la actividad transcripcional del promotor de *CXCR5* respecto al alelo protector en orientación *forward*. Dicha disminución parece afectar específicamente a la expresión del receptor *CXCR5* en la superficie de las células T reguladoras circulantes, lo que sugiere un menor control de la reacción del centro germinal.
- El alelo de riesgo del SNP rs2762943 (T) se asocia con niveles disminuidos de calcitriol, la forma activa de la vitamina D, en suero de pacientes con EM. Esta asociación está posiblemente relacionada con una expresión génica de *CYP24A1* incrementada en CMSP.
- La presencia de variantes genéticas raras asociadas al gen *RGS1* se relaciona con una disminución del porcentaje de linfocitos B periféricos que lo expresan en su membrana, lo que podría alterar la migración de los mismos, y con una falta de inducción de su expresión génica en CMSP tras el tratamiento con IFN- β .
- La delección en heterocigosis de los exones 4 al 10 del gen *FCRL1* identificada en un paciente con EM se puede considerar una variante rara, pero nuestros resultados no sugieren una asociación con la enfermedad.

7. LÍNEAS DE FUTURO

7. Líneas de futuro

De forma global, los resultados derivados de la presente tesis contribuyen a la elucidación de la arquitectura genética que subyace a la EM mediante el desenmascaramiento de nuevos polimorfismos asociados de forma significativa con el riesgo de desarrollar EM. Además, el hecho de haber conseguido identificar *loci* genéticos con un efecto funcional incrementa el valor de las asociaciones reportadas y permite especular sobre su posible implicación causal en la patogénesis de la EM. No obstante, es necesario verificar en futuros estudios si la variante rs11602393 podría afectar a la reacción del centro germinal, si el polimorfismo rs2762943 es capaz de incrementar la actividad de la enzima CYP24A1, si la presencia de variantes raras asociadas al gen *RGS1* puede influenciar la capacidad migratoria de las células B y alterar la respuesta al tratamiento con IFN- β , y si realmente la delección del gen *FCRL1* no está asociada con el riesgo de EM. Por tanto, las líneas futuras a seguir serían las siguientes:

1. Genotipar específicamente la variante genética común rs11602393 en una cohorte independiente de las incluidas en la presente tesis para identificar nuevos individuos portadores de la variante de riesgo y, así, poder incrementar el tamaño de muestra del inmunofenotipado para verificar la asociación del alelo de riesgo con una disminución de la proporción de células T reguladoras circulantes CXCR5⁺.
2. Aumentar el número de individuos incluidos en el tratamiento con calcitriol *in vitro* de CMSP de pacientes con EM portadores y no portadores del alelo minoritario de la variante rs2762943 para tratar de alcanzar la significancia estadística en la estimulación de la expresión génica de *CYP24A1*. Dado que se trata de una variante común pero con una baja frecuencia en la población general (MAF=3%), no disponemos de más muestras de CMSP basales de los individuos con genotipo para esta variante, por lo que habría que genotipar pacientes adicionales.
3. Realizar ensayos de migración de las células B para comprobar si la presencia de variantes genéticas raras asociadas a *RGS1* ocasiona una mayor migración de dichas células. Para ello, se podría secuenciar la región del gen *RGS1* en nuevos individuos y, una vez identificados aquéllos con presencia o ausencia de variantes raras, obtener una muestra de sus linfocitos B, cultivarlos *in vitro* y desarrollar un ensayo de quimiotaxis hacia un gradiente de CXCL12 y CXCL13 mediante la técnica de dos cámaras.

Líneas de futuro

4. Analizar si los individuos con variantes raras del gen *RGS1* exhiben una falta de respuesta al tratamiento con IFN- β o una deficiencia en la misma, lo que podría ayudar a detectar de forma precoz a los pacientes que no responderán a esta medicación y, así, a evitar iniciar un tratamiento que no será efectivo.

5. Examinar la presencia o ausencia de la delección identificada en el gen *FCRL1* en una nueva cohorte de pacientes con EM y controles sanos, bien para descartar definitivamente su asociación con la susceptibilidad a desarrollar la enfermedad, o bien para conseguir detectar nuevos individuos portadores y proseguir con estudios funcionales.

En definitiva, a pesar de haber cooperado al esclarecimiento del componente genético de la EM, nuestros hallazgos no constituyen el cierre de ningún capítulo, sino que conforman una pista sobre el camino a seguir para acercarse a la tan deseada medicina personalizada, basada principalmente en la identificación de biomarcadores objetivamente mensurables. Entre ellos, la presencia de una variante de riesgo.

8. BIBLIOGRAFÍA

8. Bibliografía

1. Vidal-Jordana A, Montalban X. Multiple sclerosis: epidemiologic, clinical, and therapeutic aspects. *Neuroimaging Clin N Am*. 2017;27(2):195–204.
2. Compston A, Coles A. Multiple sclerosis. *Lancet*. 2008;372(9648):1502–17.
3. Comabella M, Houry SJ. Immunopathogenesis of multiple sclerosis. *Clin Immunol*. 2012;142(1):2–8.
4. Sawcer S, Franklin RJM, Ban M. Multiple sclerosis genetics. *Lancet Neurol*. 2014;13(7):700–9.
5. Wallin MT, Culpepper WJ, Coffman P, Pulaski S, Maloni H, Mahan CM, et al. The Gulf War era multiple sclerosis cohort: age and incidence rates by race, sex and service. *Brain*. 2012;135(Pt 6):1778–85.
6. Simpson SJ, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. *J Neurol Neurosurg Psychiatry*. 2011;82(10):1132–41.
7. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21(10):718–79.
8. Leray E, Moreau T, Fromont A, Edan G. Epidemiology of multiple sclerosis. *Rev Neurol (Paris)*. 2016;172(1):3–13.
9. Marrie RA, Elliott L, Marriott J, Cossoy M, Blanchard J, Leung S, et al. Effect of comorbidity on mortality in multiple sclerosis. *Neurology*. 2015;85(3):240–7.
10. Shapira Y, Agmon-Levin N, Shoenfeld Y. Defining and analyzing geoepidemiology and human autoimmunity. *J Autoimmun*. 2010;34(3):J168–77.
11. Rae-Grant A, Ontaneda D. Management of acute exacerbations in multiple sclerosis. *Ann Indian Acad Neurol*. 2009;12(4):264.
12. Steelman AJ. Infection as an environmental trigger of multiple sclerosis disease exacerbation. *Front Immunol*. 2015;6:520.
13. Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol*. 2015;15(9):545–58.
14. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology*. 1996;46(4):907–11.
15. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. 2014;83(3):278–86.
16. Miller DH, Weber T, Grove R, Wardell C, Horrigan J, Graff O, et al. Fingertegrast for relapsing remitting multiple sclerosis: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2012;11(2):131–9.
17. Jongen PJ. Health-related quality of life in patients with multiple sclerosis: impact of disease-modifying drugs. *CNS Drugs*. 2017;31(7):585–602.
18. Goldenberg MM. Multiple sclerosis review. *P T*. 2012;37(3):175–84.
19. Miller DH, Leary SM. Primary-progressive multiple sclerosis. *Lancet Neurol*. 2007;6(10):903–12.
20. Thompson AJ, Montalban X, Barkhof F, Brochet B, Filippi M, Miller DH, et al. Diagnostic criteria for primary progressive multiple sclerosis: A position paper. *Ann Neurol*. 2000;47(6):831–5.
21. Antel J, Antel S, Caramanos Z, Arnold DL, Kuhlmann T. Primary progressive multiple sclerosis: part of the MS disease spectrum or separate disease entity? *Acta Neuropathol*. 2012;123(5):627–38.
22. Schumacher GA, Beebe G, Kibler RF, Kurland LT, Kurtzke JF, McDowell F, 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*. 2006;122(1):552–68.

Bibliografía

23. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol*. 1983;13(3):227–31.
24. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001;50(1):121–7.
25. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann Neurol*. 2011;69(2):292–302.
26. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162–73.
27. Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. *Lancet*. 2017;389(10076):1336–46.
28. Rovira-Cañellas A, Alonso-Farré J, Río-Izquierdo J. [Magnetic resonance in the clinical and therapeutic follow-up of multiple sclerosis]. *Rev Neurol*. 2000;30(10):980–5.
29. Wattjes MP, Steenwijk MD, Stangel M. MRI in the diagnosis and monitoring of multiple sclerosis: an update. *Clin Neuroradiol*. 2015;25 Suppl 2(S2):157–65.
30. Geraldes R, Ciccarelli O, Barkhof F, De Stefano N, Enzinger C, Filippi M, et al. The current role of MRI in differentiating multiple sclerosis from its imaging mimics. *Nat Rev Neurol*. 2018;14(4):199–213.
31. Álvarez-Cermeño JC, Villar LM. Oligoclonal bands—a useful tool to avoid MS misdiagnosis. *Nat Rev Neurol*. 2013;9(6):303–4.
32. Avasarala JR, Cross AH, Trotter JL. Oligoclonal band number as a marker for prognosis in multiple sclerosis. *Arch Neurol*. 2001;58(12):2044–5.
33. Zeman AZJ, Kidd D, McLean BN, Kelly MA, Francis DA, Miller DH, et al. A study of oligoclonal band negative multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 1996;60(1):27–30.
34. Sharma R, Joshi S, Singh KD, Kumar A. Visual evoked potentials: normative values and gender differences. *J Clin Diagn Res*. 2015;9(7):CC12–5.
35. Stys PK, Zamponi GW, van Minnen J, Geurts JGG. Will the real multiple sclerosis please stand up? *Nat Rev Neurosci*. 2012;13(7):507–14.
36. Grigoriadis N, van Pesch V. A basic overview of multiple sclerosis immunopathology. *Eur J Neurol*. 2015;22:3–13.
37. Sospedra M, Martin R. Immunology of multiple sclerosis. *Semin Neurol*. 2016;36(2):115–27.
38. Marcus JF, Waubant EL. Updates on clinically isolated syndrome and diagnostic criteria for multiple sclerosis. *The Neurohospitalist*. 2013;3(2):65–80.
39. Popescu BFG, Pirkó I, Lucchinetti CF. Pathology of multiple sclerosis: where do we stand? *Continuum (Minneapolis)*. 2013;19(4 Multiple Sclerosis):901–21.
40. L. B, G. GJJ, J. MS, P. V. Grey matter pathology in multiple sclerosis. *Acta Neurol Scand*. 2006;113(s183):48–50.
41. Prineas JW, Parratt JDE. Oligodendrocytes and the early multiple sclerosis lesion. *Ann Neurol*. 2012;72(1):18–31.
42. Patel J, Balabanov R. Molecular mechanisms of oligodendrocyte injury in multiple sclerosis and experimental autoimmune encephalomyelitis. *Int J Mol Sci*. 2012;13(8):10647–59.
43. Popescu BFG, Lucchinetti CF. Pathology of demyelinating diseases. *Annu Rev Pathol Mech Dis*. 2012;7(1):185–217.
44. Frischer JM, Bramow S, Dal-Bianco A, Lucchinetti CF, Rauschka H, Schmidbauer M, et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain*. 2009;132(5):1175–89.
45. Roy A, Jana A, Yatish K, Freidt MB, Fung YK, Martinson JA, et al. Reactive oxygen species up-regulate CD11b in microglia via nitric oxide: Implications for

- neurodegenerative diseases. *Free Radic Biol Med.* 2008;45(5):686–99.
46. Silber E, Sharief MK. Axonal degeneration in the pathogenesis of multiple sclerosis. *J Neurol Sci.* 1999;170(1):11–8.
 47. Dutta R, Trapp BD. Pathogenesis of axonal and neuronal damage in multiple sclerosis. *Neurology.* 2007;68(Issue 22, Supplement 3):S22–31.
 48. Mahad DJ, Ziabreva I, Campbell G, Lax N, White K, Hanson PS, et al. Mitochondrial changes within axons in multiple sclerosis. *Brain.* 2009;132(5):1161–74.
 49. Gehrman J, Banati RB, Cuzner ML, Kreutzberg GW, Newcombe J. Amyloid precursor protein (APP) expression in multiple sclerosis lesions. *Glia.* 1995;15(2):141–51.
 50. Prins M, Schul E, Geurts J, van der Valk P, Drukarch B, van Dam A-M. Pathological differences between white and grey matter multiple sclerosis lesions. *Ann N Y Acad Sci.* 2015;1351(1):99–113.
 51. Wolswijk G. Oligodendrocyte precursor cells in the demyelinated multiple sclerosis spinal cord. *Brain.* 2002;125(2):338–49.
 52. Sofroniew M V. Molecular dissection of reactive astrogliosis and glial scar formation. *Trends Neurosci.* 2009;32(12):638–47.
 53. Lehmann-Horn K, Kronsbein HC, Weber MS. Targeting B cells in the treatment of multiple sclerosis: recent advances and remaining challenges. *Ther Adv Neurol Disord.* 2013;6(3):161–73.
 54. Hemmer B, Kerschensteiner M, Korn T. Role of the innate and adaptive immune responses in the course of multiple sclerosis. *Lancet Neurol.* 2015;14(4):406–19.
 55. Bielekova B, Sung M-H, Kadom N, Simon R, McFarland H, Martin R. Expansion and functional relevance of high-avidity myelin-specific CD4⁺ T cells in multiple sclerosis. *J Immunol.* 2004;172(6):3893–904.
 56. Kaskow BJ, Baecher-Allan C. Effector T cells in multiple sclerosis. *Cold Spring Harb Perspect Med.* 2018;8(4):a029025.
 57. Mangalam A, Rodriguez M, David C. Role of MHC class II expressing CD4⁺ T cells in proteolipid protein91–110-induced EAE in HLA-DR3 transgenic mice. *Eur J Immunol.* 2006;36(12):3356–70.
 58. Zhu J, Paul WE. CD4 T cells: fates, functions, and faults. *Blood.* 2008;112(5):1557–69.
 59. Elyaman W, Khoury SJ. Th9 cells in the pathogenesis of EAE and multiple sclerosis. *Semin Immunopathol.* 2017;39(1):79–87.
 60. Kaiko GE, Horvat JC, Beagley KW, Hansbro PM. Immunological decision-making: how does the immune system decide to mount a helper T-cell response? *Immunology.* 2008;123(3):326–38.
 61. Frisullo G, Nociti V, Iorio R, Patanella AK, Marti A, Caggiula M, et al. IL17 and IFN γ production by peripheral blood mononuclear cells from clinically isolated syndrome to secondary progressive multiple sclerosis. *Cytokine.* 2008;44(1):22–5.
 62. Lovett-Racke AE, Yang Y, Racke MK. Th1 versus Th17: are T cell cytokines relevant in multiple sclerosis? *Biochim Biophys Acta - Mol Basis Dis.* 2011;1812(2):246–51.
 63. Assaf AM, Al-Abbassi R, Al-Binni M. Academic stress-induced changes in Th1- and Th2-cytokine response. *Saudi Pharm J.* 2017;25(8):1237–47.
 64. Zhu J, Paul WE. Peripheral CD4⁺ T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunol Rev.* 2010;238(1):247–62.
 65. Kimura A, Kishimoto T. IL-6: Regulator of Treg/Th17 balance. *Eur J Immunol.* 2010;40(7):1830–5.
 66. Morishima N, Mizoguchi I, Takeda K, Mizuguchi J, Yoshimoto T. TGF- β is necessary for induction of IL-23R and Th17 differentiation by IL-6 and IL-23. *Biochem Biophys Res Commun.* 2009;386(1):105–10.
 67. Ghoreschi K, Laurence A, Yang X-P, Tato CM, McGeachy MJ, Konkel JE, et al. Generation of pathogenic TH17 cells in the absence of TGF- β signalling. *Nature.* 2010;467(7318):967–71.
 68. Korn T, Bettelli E, Oukka M, Kuchroo VK. IL-17 and Th17 Cells. *Annu Rev Immunol.* 2009;27(1):485–517.
 69. Sanjabi S, Zenewicz LA, Kamanaka M, Flavell RA. Anti-inflammatory and pro-

- inflammatory roles of TGF- β , IL-10, and IL-22 in immunity and autoimmunity. *Curr Opin Pharmacol.* 2009;9(4):447–53.
70. Carbo A, Hontecillas R, Andrew T, Eden K, Mei Y, Hoops S, et al. Computational modeling of heterogeneity and function of CD4⁺ T cells. *Front Cell Dev Biol.* 2014;2:31.
 71. Kelly BL, Locksley RM. Coordinate regulation of the IL-4, IL-13, and IL-5 cytokine cluster in Th2 clones revealed by allelic expression patterns. *J Immunol.* 2000;165(6):2982–6.
 72. Hewitson JP, Grainger JR, Maizels RM. Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Mol Biochem Parasitol.* 2009;167(1):1–11.
 73. Figueiredo AS, Schumacher A. The T helper type 17/regulatory T cell paradigm in pregnancy. *Immunology.* 2016;148(1):13–21.
 74. Veldhoen M, Uytendhoeve C, van Snick J, Helmbj H, Westendorf A, Buer J, et al. Transforming growth factor- β “reprograms” the differentiation of T helper 2 cells and promotes an interleukin 9–producing subset. *Nat Immunol.* 2008;9(12):1341–6.
 75. Lee PW, Severin ME, Lovett-Racke AE. TGF- β regulation of T cells in multiple sclerosis. *Eur J Immunol.* 2017;47(3):446–53.
 76. Mills KHG. Regulatory T cells: friend or foe in immunity to infection? *Nat Rev Immunol.* 2004;4(11):841–55.
 77. Dombrowski Y, O’Hagan T, Dittmer M, Penalva R, Mayoral SR, Bankhead P, et al. Regulatory T cells promote myelin regeneration in the central nervous system. *Nat Neurosci.* 2017;20(5):674–80.
 78. Duhon T, Geiger R, Jarrossay D, Lanzavecchia A, Sallusto F. Production of interleukin 22 but not interleukin 17 by a subset of human skin-homing memory T cells. *Nat Immunol.* 2009;10(8):857–63.
 79. Fard NA, Azizi G, Mirshafiey A. The potential role of T helper cell 22 and IL-22 in immunopathogenesis of multiple sclerosis. *Innov Clin Neurosci.* 2016;13(7–8):30–6.
 80. Rolla S, Bardina V, De Mercanti S, Quaglino P, De Palma R, Gned D, et al. Th22 cells are expanded in multiple sclerosis and are resistant to IFN- β . *J Leukoc Biol.* 2014;96(6):1155–64.
 81. Nurieva RI, Chung Y, Hwang D, Yang XO, Kang HS, Ma L, et al. Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. *Immunity.* 2008;29(1):138–49.
 82. Crotty S. T follicular helper cell differentiation, function, and roles in disease. *Immunity.* 2014;41(4):529–42.
 83. Asrir A, Aloulou M, Gador M, Pérals C, Fazilleau N. Interconnected subsets of memory follicular helper T cells have different effector functions. *Nat Commun.* 2017;8(1):847.
 84. Zotos D, Coquet JM, Zhang Y, Light A, D’Costa K, Kallies A, et al. IL-21 regulates germinal center B cell differentiation and proliferation through a B cell–intrinsic mechanism. *J Exp Med.* 2010;207(2):365–78.
 85. Gatto D, Brink R. The germinal center reaction. *J Allergy Clin Immunol.* 2010;126(5):898–907.
 86. Oprea M, Perelson AS. Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J Immunol.* 1997;158(11):5155–62.
 87. Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell.* 2010;143(4):592–605.
 88. Kurosaki T, Kometani K, Ise W. Memory B cells. *Nat Rev Immunol.* 2015;15(3):149–59.
 89. Morita R, Schmitt N, Bentebibel S-E, Ranganathan R, Bourdery L, Zurawski G, et al. Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity.* 2011;34(1):108–21.
 90. Schmitt N, Bentebibel S-E, Ueno H. Phenotype and functions of memory Tfh cells in human blood. *Trends Immunol.* 2014;35(9):436–42.

91. Cha Z, Gu H, Zang Y, Wang Z, Li J, Huang W, et al. The prevalence and function of CD4 + CXCR5 + Foxp3 + follicular regulatory T cells in diffuse large B cell lymphoma. *Int Immunopharmacol.* 2018;61:132–9.
92. Dhaeze T, Peelen E, Hombrouck A, Peeters L, Van Wijmeersch B, Lemkens N, et al. Circulating follicular regulatory T cells are defective in multiple sclerosis. *J Immunol.* 2015;195(3):832–40.
93. Larochelle C, Alvarez JI, Prat A. How do immune cells overcome the blood-brain barrier in multiple sclerosis? *FEBS Lett.* 2011;585(23):3770–80.
94. Denic A, Wootla B, Rodriguez M. CD8 + T cells in multiple sclerosis. *Expert Opin Ther Targets.* 2013;17(9):1053–66.
95. Booss J, Esiri MM, Tourtellotte WW, Mason DY. Immunohistological analysis of T lymphocyte subsets in the central nervous system in chronic progressive multiple sclerosis. *J Neurol Sci.* 1983;62(1–3):219–32.
96. Hauser SL, Bhan AK, Gilles F, Kemp M, Kerr C, Weiner HL. Immunohistochemical analysis of the cellular infiltrate in multiple sclerosis lesions. *Ann Neurol.* 1986;19(6):578–87.
97. Haines JD, Inglese M, Casaccia P. Axonal damage in multiple sclerosis. *Mt Sinai J Med.* 2011;78(2):231–43.
98. Ifergan I, Kebir H, Alvarez JI, Marceau G, Bernard M, Bourbonniere L, et al. Central nervous system recruitment of effector memory CD8+ T lymphocytes during neuroinflammation is dependent on 4 integrin. *Brain.* 2011;134(12):3560–77.
99. Murali-Krishna K, Altman JD, Suresh M, Sourdive DJD, Zajac AJ, Miller JD, et al. Counting antigen-specific CD8 T cells: a reevaluation of bystander activation during viral infection. *Immunity.* 1998;8(2):177–87.
100. Friese MA, Fugger L. Pathogenic CD8 + T cells in multiple sclerosis. *Ann Neurol.* 2009;66(2):132–41.
101. Ford M, Evavold B. Specificity, magnitude, and kinetics of MOG-specific CD8+ T?cell responses during experimental autoimmune encephalomyelitis. *Eur J Immunol.* 2005;35(1):76–85.
102. Ortega SB, Kashi VP, Tyler AF, Cunnusamy K, Mendoza JP, Karandikar NJ. The disease-ameliorating function of autoregulatory CD8 T cells is mediated by targeting of encephalitogenic CD4 T cells in experimental autoimmune encephalomyelitis. *J Immunol.* 2013;191(1):117–26.
103. Correale J, Villa A. Isolation and characterization of CD8+ regulatory T cells in multiple sclerosis. *J Neuroimmunol.* 2008;195(1–2):121–34.
104. Fogdell-Hahn A, Ligiers A, Gronning M, Hillert J, Olerup O. Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens.* 2000;55(2):140–8.
105. Didonna A, Oksenberg JR. Genetic determinants of risk and progression in multiple sclerosis. *Clin Chim Acta.* 2015;449:16–22.
106. Harbo HF, Lie BA, Sawcer S, Celius EG, Dai K-Z, Oturai A, et al. Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis. *Tissue Antigens.* 2004;63(3):237–47.
107. Baranzini SE, Jeong MC, Butunoi C, Murray RS, Bernard CC, Oksenberg JR. B cell repertoire diversity and clonal expansion in multiple sclerosis brain lesions. *J Immunol.* 1999;163(9):5133–44.
108. Wurth S, Kuenz B, Bsteh G, Ehling R, Di Pauli F, Hegen H, et al. Cerebrospinal fluid B cells and disease progression in multiple sclerosis - A longitudinal prospective study. Jacobson S, editor. *PLoS One.* 2017;12(8):e0182462.
109. Meinel E, Krumbholz M, Hohlfeld R. B lineage cells in the inflammatory central nervous system environment: Migration, maintenance, local antibody production, and therapeutic modulation. *Ann Neurol.* 2006;59(6):880–92.
110. Serafini B, Rosicarelli B, Magliozzi R, Stigliano E, Aloisi F. Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. *Brain Pathol.* 2004;14(2):164–74.

Bibliografia

111. Lucchinetti CF, Popescu BFG, Bunyan RF, Moll NM, Roemer SF, Lassmann H, et al. Inflammatory cortical demyelination in early multiple sclerosis. *N Engl J Med*. 2011;365(23):2188–97.
112. Krumbholz M, Derfuss T, Hohlfeld R, Meinl E. B cells and antibodies in multiple sclerosis pathogenesis and therapy. *Nat Rev Neurol*. 2012;8(11):613–23.
113. Huttner HB, Schellinger PD, Struffert T, Richter G, Engelhorn T, Bassemir T, et al. MRI criteria in MS patients with negative and positive oligoclonal bands: equal fulfillment of Barkhof's criteria but different lesion patterns. *J Neurol*. 2009;256(7):1121–5.
114. Kinnunen T, Chamberlain N, Morbach H, Cantaert T, Lynch M, Preston-Hurlburt P, et al. Specific peripheral B cell tolerance defects in patients with multiple sclerosis. *J Clin Invest*. 2013;123(6):2737–41.
115. Pröbstel A-K, Sanderson NSR, Derfuss T. B cells and autoantibodies in multiple sclerosis. Kleinschnitz C, Meuth S, editors. *Int J Mol Sci*. 2015;16(7):16576–92.
116. Barr TA, Shen P, Brown S, Lampropoulou V, Roch T, Lawrie S, et al. B cell depletion therapy ameliorates autoimmune disease through ablation of IL-6-producing B cells. *J Exp Med*. 2012;209(5):1001–10.
117. Akassoglou K, Bauer J, Kassiotis G, Pasparakis M, Lassmann H, Kollias G, et al. Oligodendrocyte apoptosis and primary demyelination induced by local TNF/p55TNF receptor signaling in the central nervous system of transgenic mice: models for multiple sclerosis with primary oligodendroglialopathy. *Am J Pathol*. 1998;153(3):801–13.
118. Hoffmann F, Meinl E. B cells in multiple sclerosis: good or bad guys? *Eur J Immunol*. 2014;44(5):1247–50.
119. Gandhi R, Laroni A, Weiner HL. Role of the innate immune system in the pathogenesis of multiple sclerosis. *J Neuroimmunol*. 2010;221(1–2):7–14.
120. Ganguly D, Haak S, Sisirak V, Reizis B. The role of dendritic cells in autoimmunity. *Nat Rev Immunol*. 2013;13(8):566–77.
121. Ferlazzo G, Pack M, Thomas D, Paludan C, Schmid D, Strowig T, et al. Distinct roles of IL-12 and IL-15 in human natural killer cell activation by dendritic cells from secondary lymphoid organs. *Proc Natl Acad Sci*. 2004;101(47):16606–11.
122. Laroni A, Gandhi R, Beynon V, Weiner HL. IL-27 imparts immunoregulatory function to human NK cell subsets. Mosley RL, editor. *PLoS One*. 2011;6(10):e26173.
123. Chabas D. The influence of the proinflammatory cytokine, osteopontin, on autoimmune demyelinating disease. *Science*. 2001;294(5547):1731–5.
124. Hur EM, Youssef S, Haws ME, Zhang SY, Sobel RA, Steinman L. Osteopontin-induced relapse and progression of autoimmune brain disease through enhanced survival of activated T cells. *Nat Immunol*. 2007;8(1):74–83.
125. Huang Y-M, Xiao B-G, Özenci V, Kouwenhoven M, Teleshova N, Fredrikson S, et al. Multiple sclerosis is associated with high levels of circulating dendritic cells secreting pro-inflammatory cytokines. *J Neuroimmunol*. 1999;99(1):82–90.
126. Nuyts AH, Lee WP, Bashir-Dar R, Berneman ZN, Cools N. Dendritic cells in multiple sclerosis: key players in the immunopathogenesis, key players for new cellular immunotherapies? *Mult Scler J*. 2013;19(8):995–1002.
127. Dargahi N, Katsara M, Tselios T, Androutsou M-E, de Courten M, Matsoukas J, et al. Multiple sclerosis: immunopathology and treatment update. Bernitsas E, editor. *Brain Sci*. 2017;7(7):78.
128. Lassmann H, Brück W, Lucchinetti CF. The immunopathology of multiple sclerosis: an overview. *Brain Pathol*. 2007;17(2):210–8.
129. Almolda B, González B, Castellano B. Are microglial cells the regulators of lymphocyte responses in the CNS? *Front Cell Neurosci*. 2015;9:440.
130. Chastain EML, Duncan DS, Rodgers JM, Miller SD. The role of antigen presenting cells in multiple sclerosis. *Biochim Biophys Acta*. 2011;1812(2):265–74.
131. Ponomarev ED, Shriver LP, Maresz K, Dittel BN. Microglial cell activation and proliferation precedes the onset of CNS autoimmunity. *J Neurosci Res*. 2005;81(3):374–89.
132. Yamasaki R, Lu H, Butovsky O, Ohno N, Rietsch AM, Cialic R, et al. Differential roles

- of microglia and monocytes in the inflamed central nervous system. *J Exp Med*. 2014;211(8):1533–49.
133. Luo C, Jian C, Liao Y, Huang Q, Wu Y, Liu X, et al. The role of microglia in multiple sclerosis. *Neuropsychiatr Dis Treat*. 2017;13:1661–7.
 134. Tierney JB, Kharkrang M, La Flamme AC. Type II-activated macrophages suppress the development of experimental autoimmune encephalomyelitis. *Immunol Cell Biol*. 2009;87(3):235–40.
 135. Strachan-Whaley M, Rivest S, Yong VW. Interactions between microglia and T cells in multiple sclerosis pathobiology. *J Interf Cytokine Res*. 2014;34(8):615–22.
 136. Cooper MA, Fehniger TA, Caligiuri MA. The biology of human natural killer-cell subsets. *Trends Immunol*. 2001;22(11):633–40.
 137. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S. Functions of natural killer cells. *Nat Immunol*. 2008;9(5):503–10.
 138. Jacobs R, Hintzen G, Kemper A, Beul K, Kempf S, Behrens G, et al. CD56bright cells differ in their KIR repertoire and cytotoxic features from CD56dim NK cells. *Eur J Immunol*. 2001;31(10):3121–6.
 139. TRAUGOTT U, RAINE CS. Further lymphocyte characterization in the central nervous system in multiple sclerosis. *Ann N Y Acad Sci*. 1984;436(1 Multiple Scle):163–78.
 140. Amin K. The role of mast cells in allergic inflammation. *Respir Med*. 2012;106(1):9–14.
 141. Christy AL, Brown MA. The multitasking mast cell: positive and negative roles in the progression of autoimmunity. *J Immunol*. 2007;179(5):2673–9.
 142. Kruger PG. Mast cells and multiple sclerosis: a quantitative analysis. *Neuropathol Appl Neurobiol*. 2001;27(4):275–80.
 143. Haygreen E. NKT Cells : Invariant. British Society for Immunology. 2012. Disponible en: <https://www.immunology.org/public-information/bitesized-immunology/cells/nkt-cells-invariant>
 144. Van Kaer L, Wu L, Parekh V V. Natural killer T cells in multiple sclerosis and its animal model, experimental autoimmune encephalomyelitis. *Immunology*. 2015;146(1):1–10.
 145. Brenner MB, McLean J, Dialynas DP, Strominger JL, Smith JA, Owen FL, et al. Identification of a putative second T-cell receptor. *Nature*. 1986;322(6075):145–9.
 146. Wohler JE, Smith SS, Barnum SR. $\gamma\delta$ T cells: the overlooked T-cell subset in demyelinating disease. *J Neurosci Res*. 2010;88(1):1–6.
 147. Komori HK, Meehan TF, Havran WL. Epithelial and mucosal $\gamma\delta$ T cells. *Curr Opin Immunol*. 2006;18(5):534–8.
 148. Blink SE, Miller SD. The contribution of gammadelta T cells to the pathogenesis of EAE and MS. *Curr Mol Med*. 2009;9(1):15–22.
 149. Murzenok PP, Matusevicius D, Freedman MS. gamma/delta T cells in multiple sclerosis: chemokine and chemokine receptor expression. *Clin Immunol*. 2002;103(3 Pt 1):309–16.
 150. Stinissen P, Vandevyver C, Medaer R, Vandegaer L, Nies J, Tuyls L, et al. Increased frequency of gamma delta T cells in cerebrospinal fluid and peripheral blood of patients with multiple sclerosis. Reactivity, cytotoxicity, and T cell receptor V gene rearrangements. *J Immunol*. 1995;154(9):4883–94.
 151. Millar DG, Ohashi PS. Central tolerance: what you see is what you don't get! *Nat Immunol*. 2016;17(2):115–6.
 152. Xing Y, Hogquist KA. T-cell tolerance: central and peripheral. *Cold Spring Harb Perspect Biol*. 2012;4(6):a006957.
 153. Corthay A. How do regulatory t cells work? Vol. 70, *Scandinavian Journal of Immunology*. 2009. p. 326–36.
 154. Anderton SM. Avoiding autoimmune disease-T cells know their limits. *Trends Immunol*. 2006;27(5):208–14.
 155. Legroux L, Arbour N. Multiple sclerosis and T lymphocytes: an entangled story. *J Neuroimmune Pharmacol*. 2015;10(4):528–46.
 156. Gonsette RE. Self-tolerance in multiple sclerosis. *Acta Neurol Belg*. 2012;112(2):133–40.
 157. Toker A, Huehn J. To be or not to be a Treg cell: lineage decisions controlled by

- epigenetic mechanisms. *Sci Signal*. 2011;4(158):pe4.
158. Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol*. 2015;15:545–58.
 159. Fujinami RS, Oldstone MBA. Amino acid homology between the encephalitogenic site of myelin basic protein and virus: mechanism for autoimmunity. *Science*. 1985;230(4729):1043–5.
 160. Harkiolaki M, Holmes SL, Svendsen P, Gregersen JW, Jensen LT, McMahon R, et al. T cell-mediated autoimmune disease due to low-affinity crossreactivity to common microbial peptides. *Immunity*. 2009;30(3):348–57.
 161. Lang HLE, Jacobsen H, Ikemizu S, Andersson C, Harlos K, Madsen L, et al. A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nat Immunol*. 2002;3(10):940–3.
 162. Tejada-Simon M V., Zang YCQ, Hong J, Rivera VM, Zhang JZ. Cross-reactivity with myelin basic protein and human herpesvirus-6 in multiple sclerosis. *Ann Neurol*. 2003;53(2):189–97.
 163. Furtado GC, Marcondes MCG, Latkowski J-A, Tsai J, Wensky A, Lafaille JJ. Swift entry of myelin-specific T lymphocytes into the central nervous system in spontaneous autoimmune encephalomyelitis. *J Immunol*. 2008;181(7):4648–55.
 164. McCoy L, Tsunoda I, Fujinami RS. Multiple sclerosis and virus induced immune responses: Autoimmunity can be primed by molecular mimicry and augmented by bystander activation. *Autoimmunity*. 2006;39(1):9–19.
 165. Ji Q, Perchellet A, Goverman JM. Viral infection triggers central nervous system autoimmunity via activation of CD8+ T cells expressing dual TCRs. *Nat Immunol*. 2010;11(7):628–34.
 166. Obermeier B, Daneman R, Ransohoff RM. Development, maintenance and disruption of the blood-brain barrier. *Nat Med*. 2013;19(12):1584–96.
 167. Bell RD, Winkler EA, Sagare AP, Singh I, LaRue B, Deane R, et al. Pericytes control key neurovascular functions and neuronal phenotype in the adult brain and during brain aging. *Neuron*. 2010;68(3):409–27.
 168. Luissint A-C, Artus C, Glacial F, Ganeshamoorthy K, Couraud P-O. Tight junctions at the blood brain barrier: physiological architecture and disease-associated dysregulation. *Fluids Barriers CNS*. 2012;9(1):23.
 169. Dolati S, Babaloo Z, Jadidi-Niaragh F, Ayromlou H, Sadreddini S, Yousefi M. Multiple sclerosis: Therapeutic applications of advancing drug delivery systems. *Biomed Pharmacother*. 2017;86:343–53.
 170. Pestka S, Krause CD, Walter MR. Interferons, interferon-like cytokines, and their receptors. *Immunol Rev*. 2004;202(1):8–32.
 171. Kieseier BC. The mechanism of action of interferon- β in relapsing multiple sclerosis. *CNS Drugs*. 2011;25(6):491–502.
 172. Zafranskaya M, Oschmann P, Engel R, Weishaupt A, van Noort JM, Jomaa H, et al. Interferon- γ therapy reduces CD4 + and CD8 + T-cell reactivity in multiple sclerosis. *Immunology*. 2007;121(1):29–39.
 173. Correale J, Gaitan MI. Multiple sclerosis and environmental factors: the role of vitamin D, parasites, and Epstein-Barr virus infection. *Acta Neurol Scand*. 2015;132(199):46–55.
 174. van der Mei IAF, Ponsonby A-L, Blizzard L, Dwyer T. Regional variation in multiple sclerosis prevalence in Australia and its association with ambient ultraviolet radiation. *Neuroepidemiology*. 2001;20(3):168–74.
 175. Munger KL, Zhang SM, O'Reilly E, Hernán MA, Olek MJ, Willett WC, et al. Vitamin D intake and incidence of multiple sclerosis. *Neurology*. 2004;62(1):60–5.
 176. Simpson S, Taylor B, Blizzard L, Ponsonby A-L, Pittas F, Tremlett H, et al. Higher 25-hydroxyvitamin D is associated with lower relapse risk in multiple sclerosis. *Ann Neurol*. 2010;68(2):193–203.
 177. Pierrot-Deseilligny C, Rivaud-Péchoux S, Clerson P, de Paz R, Souberbielle J-C. Relationship between 25-OH-D serum level and relapse rate in multiple sclerosis patients before and after vitamin D supplementation. *Ther Adv Neurol Disord*.

- 2012;5(4):187–98.
178. Ascherio A, Munger KL, White R, Köchert K, Simon KC, Polman CH, et al. Vitamin D as an early predictor of multiple sclerosis activity and progression. *JAMA Neurol.* 2014;71(3):306–14.
 179. Correale J, Ysrraelit MC, Gaitan MI. Immunomodulatory effects of Vitamin D in multiple sclerosis. *Brain.* 2009;132(5):1146–60.
 180. Berg-Hansen P, Moen SM, Harbo HF, Celius EG. High prevalence and no latitude gradient of multiple sclerosis in Norway. *Mult Scler.* 2014;20(13):1780–2.
 181. Berg-Hansen P, Moen S, Harbo H, Celius E. High prevalence and no latitude gradient of multiple sclerosis in Norway. *Mult Scler J.* 2014;20(13):1780–2.
 182. Koch-Henriksen N, Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet Neurol.* 2010;9(5):520–32.
 183. Hernan MA. Cigarette smoking and incidence of multiple sclerosis. *Am J Epidemiol.* 2001;154(1):69–74.
 184. Riise T, Nortvedt MW, Ascherio A. Smoking is a risk factor for multiple sclerosis. *Neurology.* 2003;61(8):1122–4.
 185. Healy BC, Ali EN, Guttmann CRG, Chitnis T, Glanz BI, Buckle G, et al. Smoking and disease progression in multiple sclerosis. *Arch Neurol.* 2009;66(7):858–64.
 186. Roudbari SA, Ansari MM, Yousefzad A. Smoking as a risk factor for development of secondary progressive multiple sclerosis: a study in IRAN, Guilan. *J Neurol Sci.* 2013;330(1–2):52–5.
 187. Hedström AK, Ryner M, Fink K, Fogdell-Hahn A, Alfredsson L, Olsson T, et al. Smoking and risk of treatment-induced neutralizing antibodies to interferon β -1a. *Mult Scler.* 2014;20(4):445–50.
 188. Schwarz S, Leweling H. Diet and multiple sclerosis. *Nervenarzt.* 2005;76(2):131–42.
 189. Sedaghat F, Jessri M, Behrooz M, Mirghotbi M, Rashidkhani B. Mediterranean diet adherence and risk of multiple sclerosis: a case-control study. *Asia Pac J Clin Nutr.* 2016;25(2):377–84.
 190. Yadav V, Marracci G, Kim E, Spain R, Cameron M, Overs S, et al. Low-fat, plant-based diet in multiple sclerosis: a randomized controlled trial. *Mult Scler Relat Disord.* 2016;9:80–90.
 191. Gianfrancesco MA, Barcellos LF. Obesity and multiple sclerosis susceptibility: a review. *J Neurol neuromedicine.* 2016;1(7):1–5.
 192. Mokry LE, Ross S, Timpson NJ, Sawcer S, Davey Smith G, Richards JB. Obesity and multiple sclerosis: a Mendelian randomization study. Muraro PA, editor. *PLOS Med.* 2016;13(6):e1002053.
 193. Calvo-Barreiro L, Eixarch H, Montalban X, Espejo C. Combined therapies to treat complex diseases: the role of the gut microbiota in multiple sclerosis. *Autoimmun Rev.* 2018;17(2):165–74.
 194. Kim D, Zeng MY, Núñez G. The interplay between host immune cells and gut microbiota in chronic inflammatory diseases. *Exp Mol Med.* 2017;49(5):e339.
 195. Wendel-Haga M, Celius EG. Is the hygiene hypothesis relevant for the risk of multiple sclerosis? *Acta Neurol Scand.* 2017;136 Suppl:26–30.
 196. Fleming JO, Cook TD. Multiple sclerosis and the hygiene hypothesis. *Neurology.* 2006;67(11):2085–6.
 197. Yao G, Wang P, Luo X-D, Yu T-M, Harris RA, Zhang X-M. Meta-analysis of association between *Helicobacter pylori* infection and multiple sclerosis. *Neurosci Lett.* 2016;620:1–7.
 198. Jaruvongvanich V, Sanguankeo A, Jaruvongvanich S, Upala S. Association between *Helicobacter pylori* infection and multiple sclerosis: a systematic review and meta-analysis. *Mult Scler Relat Disord.* 2016;7:92–7.
 199. Stascheit F, Paul F, Harms L, Rosche B. *Toxoplasma gondii* seropositivity is negatively associated with multiple sclerosis. *J Neuroimmunol.* 2015;285:119–24.
 200. 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(6):824–30.

Bibliografia

201. Simpson S, Taylor B, Dwyer DE, Taylor J, Blizzard L, Ponsonby A-L, et al. Anti-HHV-6 IgG titer significantly predicts subsequent relapse risk in multiple sclerosis. *Mult Scler J*. 2012;18(6):799–806.
202. Ascherio A, Munger KL, Lennette ET, Spiegelman D, Hernán MA, Olek MJ, et al. Epstein-Barr virus antibodies and risk of multiple sclerosis. *JAMA*. 2001;286(24):3083.
203. Wekerle H, Hohlfeld R. Molecular mimicry in multiple sclerosis. *N Engl J Med*. 2003;349(2):185–6.
204. Serafini B, Rosicarelli B, Franciotta D, Magliozzi R, Reynolds R, Cinque P, et al. Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. *J Exp Med*. 2007;204(12):2899–912.
205. Angelini DF, Serafini B, Piras E, Severa M, Coccia EM, Rosicarelli B, et al. Increased CD8+ T cell response to Epstein-Barr virus lytic antigens in the active phase of multiple sclerosis. *Evavold BD, editor. PLoS Pathog*. 2013;9(4):e1003220.
206. Virtanen JO, Jacobson S. Viruses and multiple sclerosis. *CNS Neurol Disord Drug Targets*. 2012;11(5):528–44.
207. Cone RW, Huang M-LW, Hackman RC. Human herpesvirus 6 and pneumonia. *Leuk Lymphoma*. 1994;15(3–4):235–41.
208. Reynaud JM, Horvat B. Human Herpesvirus 6 and neuroinflammation. *ISRN Virol*. 2013;2013:1–11.
209. World Health Organization, International Programme on Chemical Safety. Biomarkers in risk assessment: validity and validation. *Environ Heal*. 2001.
210. Paul A, Comabella M, Gandhi R. Biomarkers in multiple sclerosis. *Cold Spring Harb Perspect Med*. 2018;a029058.
211. Ziegler A, Koch A, Krockenberger K, Großhennig A. Personalized medicine using DNA biomarkers: a review. *Hum Genet*. 2012;131(10):1627–38.
212. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. *Am J Med Genet*. 1988;29(3):533–41.
213. Oksenberg JR. Decoding multiple sclerosis: an update on genomics and future directions. *Expert Rev Neurother*. 2013;13(sup2):11–9.
214. Ebers GC, Yee IM, Sadovnick AD, Duquette P. Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. *Canadian Collaborative Study Group. Ann Neurol*. 2000;48(6):927–31.
215. Ebers GC, Sadovnick AD, Dymont DA, Yee IML, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *Lancet (London, England)*. 2004;363(9423):1773–4.
216. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. *Canadian Collaborative Study Group. Nature*. 1995;377(6545):150–1.
217. Muers M. Ups and downs at the MHC. *Nat Rev Genet*. 2011;12(7):456–7.
218. Cullen M, Perfetto SP, Klitz W, Nelson G, Carrington M. High-resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am J Hum Genet*. 2002;71(4):759–76.
219. NAITO S, NAMEROW N, MICKEY MR, TERASAKI PI. Multiple Sclerosis: association with HL-A3. *Tissue Antigens*. 2008;2(1):1–4.
220. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. *Lancet (London, England)*. 1972;1(7762):1240–1.
221. Terasaki P, Park M, Opelz G, Ting A. Multiple sclerosis and high incidence of a B lymphocyte antigen. *Science (80-)*. 1976;193(4259):1245–7.
222. Fogdell A, Hillert J, Sachs C, Olerup O. The multiple sclerosis- and narcolepsy-associated HLA class II haplotype includes the DRB5 * 0101 allele. *Tissue Antigens*. 1995;46(4):333–6.
223. Dymont DA, Ebers GC, Dessa Sadovnick A. Genetics of multiple sclerosis. *Lancet Neurol*. 2004;3(2):104–10.
224. Lill CM. Recent advances and future challenges in the genetics of multiple sclerosis. *Front Neurol*. 2014;5:130.
225. Barcellos LF, Oksenberg JR, Begovich AB, Martin ER, Schmidt S, Vittinghoff E, et al.

- HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. *Am J Hum Genet.* 2003;72(3):710–6.
226. Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476(7359):214–9.
 227. Ramagopalan S V., Maugeri NJ, Handunnetthi L, Lincoln MR, Orton S-M, Dymment DA, et al. Expression of the multiple sclerosis-associated MHC Class II allele HLA-DRB1*1501 is regulated by vitamin D. Roopenian DC, editor. *PLoS Genet.* 2009;5(2):e1000369.
 228. Kikuchi S, Fukazawa T, Niino M, Yabe I, Miyagishi R, Hamada T, et al. HLA-related subpopulations of MS in Japanese with and without oligoclonal IgG bands. *Neurology.* 2003;60(4):647–51.
 229. Imrell K, Landtblom A-M, Hillert J, Masterman T. Multiple sclerosis with and without CSF bands: clinically indistinguishable but immunogenetically distinct. *Neurology.* 2006;67(6):1062–4.
 230. Wu J-S, Qiu W, Castley A, James I, Joseph J, Christiansen FT, et al. Presence of CSF oligoclonal bands (OCB) is associated with the HLA-DRB1 genotype in a West Australian multiple sclerosis cohort. *J Neurol Sci.* 2010;288(1–2):63–7.
 231. Romero-Pinel L, Martínez-Yélamos S, Bau L, Matas E, Gubieras L, María Pujal J, et al. Association of HLA-DRB1*15 allele and CSF oligoclonal bands in a Spanish multiple sclerosis cohort. *Eur J Neurol.* 2011;18(10):1258–62.
 232. Sawcer S, Ban M, Maranian M, Yeo TW, Compston A, Kirby A, et al. A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet.* 2005;77(3):454–67.
 233. Bashinskaya V V., Kulakova OG, Boyko AN, Favorov A V., Favorova OO. A review of genome-wide association studies for multiple sclerosis: classical and hypothesis-driven approaches. *Hum Genet.* 2015;134(11–12):1143–62.
 234. COMPSTON D. B-lymphocyte alloantigens associated with multiple sclerosis. *Lancet.* 1976;308(7998):1261–5.
 235. Lundmark F, Duvefelt K, Iacobaeus E, Kockum I, Wallstrom E, Khademi M, et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet.* 2007;39(9):1108–13.
 236. Rasmussen HB, Kelly MA, Clausen J. Genetic susceptibility to multiple sclerosis: detection of polymorphic nucleotides and an intron in the 3' untranslated region of the major histocompatibility complex class II transactivator gene. *Hum Immunol.* 2001;62(4):371–7.
 237. Vandembroek K, Alvarez J, Swaminathan B, Alloza I, Matesanz F, Urcelay E, et al. A cytokine gene screen uncovers SOCS1 as genetic risk factor for multiple sclerosis. *Genes Immun.* 2012;13(1):21–8.
 238. Tizaoui K. Multiple sclerosis genetics: results from meta-analyses of candidate-gene association studies. *Cytokine.* 2018;106:154–64.
 239. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet.* 2002;3(5):391–7.
 240. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. *Nat Genet.* 2001;29(2):229–32.
 241. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med.* 2007;357(9):851–62.
 242. Comabella M, Craig DW, Camiña-Tato M, Morcillo C, Lopez C, Navarro A, et al. Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. Gwinn K, editor. *PLoS One.* 2008;3(10):e3490.
 243. Spencer CCA, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLOS Genet.* 2009;5(5):e1000477.
 244. Fadista J, Manning AK, Florez JC, Groop L. The (in)famous GWAS P-value threshold

- revisited and updated for low-frequency variants. *Eur J Hum Genet.* 2016;24(8):1202–5.
245. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan S V., Broer L, Jafari N, Hillert J, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nat Genet.* 2008;40(12):1402–3.
 246. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet.* 2009;18(4):767–78.
 247. Bahlo M, Booth DR, Broadley SA, Brown MA, Foote SJ, Griffiths LR, et al. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet.* 2009;41(7):824–8.
 248. Jakkula E, Leppä V, Sulonen A-M, Varilo T, Kallio S, Kempainen A, et al. Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. *Am J Hum Genet.* 2010;86(2):285–91.
 249. Zhang J, Wu J, Peng X, Song J, Wang J, Dong W. Associations between STAT3 rs744166 polymorphisms and susceptibility to ulcerative colitis and Crohn’s disease: a meta-analysis. Hold GL, editor. *PLoS One.* 2014;9(10):e109625.
 250. Sanna S, Pitzalis M, Zoledziewska M, Zara I, Sidore C, Murru R, et al. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet.* 2010;42(6):495–7.
 251. Nischwitz S, Cepok S, Kroner A, Wolf C, Knop M, Muller-Sarnowski F, et al. Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. *J Neuroimmunol.* 2010;227(1–2):162–6.
 252. Patsopoulos NA, de Bakker PIW. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol.* 2011;70(6):897–912.
 253. Matesanz F, Gonzalez-Perez A, Lucas M, Sanna S, Gayan J, Urcelay E, et al. Genome-wide association study of multiple sclerosis confirms a novel locus at 5p13.1. *PLoS One.* 2012;7(5):e36140.
 254. Martinelli-Boneschi F, Esposito F, Brambilla P, Lindström E, Lavorgna G, Stankovich J, et al. A genome-wide association study in progressive multiple sclerosis. *Mult Scler J.* 2012;18(10):1384–94.
 255. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet.* 2013;45(11):1353–60.
 256. Andlauer TFM, Buck D, Antony G, Bayas A, Bechmann L, Berthele A, et al. Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. *Sci Adv.* 2016;2(6):e1501678.
 257. Steri M, Orrù V, Idda ML, Pitzalis M, Pala M, Zara I, et al. Overexpression of the cytokine BAFF and autoimmunity risk. *N Engl J Med.* 2017;376(17):1615–26.
 258. Patsopoulos N, Baranzini SE, Santaniello A, Shoostari P, Cotsapas C, Wong G, et al. The Multiple Sclerosis Genomic Map: role of peripheral immune cells and resident microglia in susceptibility. *bioRxiv.* 2017.
 259. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747–53.
 260. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Vol. 11, *Nat Rev Genet.* 2010;11(6):446–50.
 261. Gudmundsson J, Sulem P, Gudbjartsson DF, Masson G, Agnarsson BA, Benediksdottir KR, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. *Nat Genet.* 2012;44(12):1326–9.
 262. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson P V., Bjornsson S, et al. A mutation in APP protects against Alzheimer’s disease and age-related cognitive decline. *Nature.* 2012;488(7409):96–9.
 263. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet.* 2001;17(9):502–10.
 264. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for

- complex diseases. *Curr Opin Genet Dev.* 2009;19(3):212–9.
265. Sunyaev S. Impact of selection, mutation rate and genetic drift on human genetic variation. *Hum Mol Genet.* 2003;12(24):3325–30.
 266. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet.* 2012;13(2):135–45.
 267. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet.* 2008;83(3):311–21.
 268. Saint Pierre A, Genin E. How important are rare variants in common disease? *Brief Funct Genomics.* 2014;13(5):353–61.
 269. Ramagopalan S V., Dymant DA, Cader MZ, Morrison KM, Disanto G, Morahan JM, et al. Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Ann Neurol.* 2011;70(6):881–6.
 270. Maver A, Lavtar P, Ristić S, Stopinšek S, Simčič S, Hočevkar K, et al. Identification of rare genetic variation of NLRP1 gene in familial multiple sclerosis. *Sci Rep.* 2017;7(1):3715.
 271. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet.* 2009;10(4):241–51.
 272. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. *Nat Genet.* 2004;36(9):949–51.
 273. Freeman JL. Copy number variation: new insights in genome diversity. *Genome Res.* 2006;16(8):949–61.
 274. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet.* 2006;7(2):85–97.
 275. Lobo I. Copy number variation and genetic disease. *Nat Educ.* 2008;1(1):65.
 276. Ionita-Laza I, Rogers AJ, Lange C, Raby BA, Lee C. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics.* 2009;93(1):22–6.
 277. Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature.* 2010;464(7289):713–20.
 278. Ferlini A, Bovolenta M, Neri M, Gualandi F, Balboni A, Yuryev A, et al. Custom CGH array profiling of copy number variations (CNVs) on chromosome 6p21.32 (HLA locus) in patients with venous malformations associated with multiple sclerosis. *BMC Med Genet.* 2010;11(1):64.
 279. McElroy J, Krupp L, Johnson B, McCauley J, Qi Z, Caillier S, et al. Copy number variation in pediatric multiple sclerosis. *Mult Scler J.* 2013;19(8):1014–21.
 280. Koch S, Goedde R, Nigmatova V, Epplen JT, Müller N, de Seze J, et al. Association of multiple sclerosis with ILT6 deficiency. *Genes Immun.* 2005;6(5):445–7.
 281. Torkar M, Haude A, Milne S, Beck S, Trowsdale J, Wilson MJ. Arrangement of the ILT gene cluster: a common null allele of the ILT6 gene results from a 6.7-kbp deletion. *Eur J Immunol.* 2000;30(12):3655–62.
 282. Sato S, Yamamoto K, Matsushita T, Isobe N, Kawano Y, Iinuma K, et al. Copy number variations in multiple sclerosis and neuromyelitis optica. *Ann Neurol.* 2015;78(5):762–74.
 283. Wall JD, Pritchard JK. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet.* 2003;4(8):587–97.
 284. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007;81(3):559–75.
 285. Ann-Christine S. From gels to chips: “Minisequencing” primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat.* 1999;13(1):1–10.
 286. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods.* 2001;25(4):402–8.
 287. Alcina A, Fedetz M, Fernández Ó, Saiz A, Izquierdo G, Lucas M, et al. Identification of a functional variant in the KIF5A-CYP27B1-METTLL1-FAM119B locus associated with

- multiple sclerosis. *J Med Genet.* 2013;50(1):25–33.
288. Tran T, Paz P, Velichko S, Cifrese J, Belur P, Yamaguchi KD, et al. Interferon β -1b induces the expression of RGS1 a negative regulator of G-protein signaling. *Int J Cell Biol.* 2010;2010:529376.
289. Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health.* 2010;31(1):9–20.
290. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 2008;9(5):356–69.
291. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet.* 2017;101(1):5–22.
292. Kawasaki A, Furukawa H, Nishida N, Warabi E, Kondo Y, Ito S, et al. Association of functional polymorphisms in interferon regulatory factor 2 (IRF2) with susceptibility to systemic lupus erythematosus: a case-control association study. *Montgomery CG, editor. PLoS One.* 2014;9(10):e109764.
293. Diogo D, Kurreeman F, Stahl EA, Liao KP, Gupta N, Greenberg JD, et al. Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am J Hum Genet.* 2013;92(1):15–27.
294. Dymnt DA, Cader MZ, Chao MJ, Lincoln MR, Morrison KM, Disanto G, et al. Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. *Neurology.* 2012;79(5):406–11.
295. Stein J V., Nombela-Arrieta C. Chemokine control of lymphocyte trafficking: a general overview. Vol. 116, *Immunology.* 2005. p. 1–12.
296. Magliozzi R, Columba-Cabezas S, Serafini B, Aloisi F. Intracerebral expression of CXCL13 and BAFF is accompanied by formation of lymphoid follicle-like structures in the meninges of mice with relapsing experimental autoimmune encephalomyelitis. *J Neuroimmunol.* 2004;148(1–2):11–23.
297. Bagaeva L V, Rao P, Powers JM, Segal BM. CXC chemokine ligand 13 plays a role in experimental autoimmune encephalomyelitis. *J Immunol.* 2006;176(12):7676–85.
298. Krumbholz M, Theil D, Cepok S, Hemmer B, Kivisäkk P, Ransohoff RM, et al. Chemokines in multiple sclerosis: CXCL12 and CXCL13 up-regulation is differentially linked to CNS immune cell recruitment. *Brain.* 2006;129(1):200–11.
299. Eggers EL, Michel BA, Wu H, Wang S, Bevan CJ, Abounasr A, et al. Clonal relationships of CSF B cells in treatment-naïve multiple sclerosis patients. *JCI Insight.* 2017;2(22):e92724.
300. Lill CM, Schjeide B-MM, Graetz C, Ban M, Alcina A, Ortiz MA, et al. MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain.* 2013;136(6):1778–82.
301. Wu K, Malik K, Tian L, Hu M, Martin T, Foster E, et al. Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, tCUP. *Mol Genet Genomics.* 2001;265(5):763–70.
302. Rao S. Embryonic stem cells: a perfect tool for studying mammalian transcriptional enhancers. *J Stem Cell Res Ther.* 2012;Suppl 10(S10).
303. Ryu J, Woo J, Shin J, Ryoo H, Kim Y, Lee C. Profile of differential promoter activity by nucleotide substitution at GWAS signals for multiple sclerosis. *Med (United States).* 2014;93(28):e281.
304. Mitkin NA, Muratova AM, Schwartz AM, Kuprash D V. The A allele of the single-nucleotide polymorphism rs630923 creates a binding site for MEF2C resulting in reduced CXCR5 promoter activity in B-cell lymphoblastic cell lines. *Front Immunol.* 2016;7:515.
305. de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet.* 2005;37(11):1217–23.
306. Linterman MA, Pierson W, Lee SK, Kallies A, Kawamoto S, Rayner TF, et al. Foxp3+

- follicular regulatory T cells control the germinal center response. *Nat Med.* 2011;17:975.
307. Sage PT, Sharpe AH. T follicular regulatory cells. *Immunol Rev.* 2016;271(1):246–59.
 308. Wallin EF, Jolly EC, Suchanek O, Bradley JA, Espeli M, Jayne DRW, et al. Human T-follicular helper and T-follicular regulatory cell maintenance is independent of germinal centers. *Blood.* 2014;124(17):2666–74.
 309. Sage PT, Francisco LM, Carman C V, Sharpe AH. The receptor PD-1 controls follicular regulatory T cells in the lymph nodes and blood. *Nat Immunol.* 2013;14(2):152–61.
 310. Sage PT, Alvarez D, Godec J, von Andrian UH, Sharpe AH. Circulating T follicular regulatory and helper cells have memory-like properties. *J Clin Invest.* 2014;124(12):5191–204.
 311. Linterman MA, Pierson W, Lee SK, Kallies A, Kawamoto S, Rayner TF, et al. Foxp3(+) follicular regulatory T cells control T follicular helper cells and the germinal center response. *Nat Med.* 2011;17(8):975–82.
 312. Ma L, Zhao P, Jiang Z, Shan Y, Jiang Y. Imbalance of different types of CD4(+) forkhead box protein 3 (FoxP3)(+) T cells in patients with new-onset systemic lupus erythematosus. *Clin Exp Immunol.* 2013;174(3):345–55.
 313. Zheng J, Wang T, Zhang L, Cui L. Dysregulation of circulating Tfr/Tfh ratio in primary biliary cholangitis. *Scand J Immunol.* 2017;86(6):452–61.
 314. Cheng JB, Levine MA, Bell NH, Mangelsdorf DJ, Russell DW. Genetic evidence that the human CYP2R1 enzyme is a key vitamin D 25-hydroxylase. *Proc Natl Acad Sci U S A.* 2004;101(20):7711–5.
 315. Christakos S, Ajibade D V, Dhawan P, Fechner AJ, Mady LJ. Vitamin D: metabolism. *Endocrinol Metab Clin North Am.* 2010;39(2):243–53.
 316. Aranow C. Vitamin D and the Immune System. *J Investig Med.* 2011;59(6):881–6.
 317. Boucher BJ. Vitamin D insufficiency and diabetes risks. *Curr Drug Targets.* 2011;12(1):61–87.
 318. Borba VZC, Vieira JGH, Kasamatsu T, Radominski SC, Sato EI, Lazaretti-Castro M. Vitamin D deficiency in patients with active systemic lupus erythematosus. *Osteoporos Int.* 2009;20(3):427–33.
 319. Welsh P, Peters MJL, Sattar N. Is vitamin D in rheumatoid arthritis a magic bullet or a mirage? The need to improve the evidence base prior to calls for supplementation. *Arthritis Rheum.* 2011;63(7):1763–9.
 320. Runia TF, Hop WCJ, de Rijke YB, Buljevac D, Hintzen RQ. Lower serum vitamin D levels are associated with a higher relapse risk in multiple sclerosis. *Neurology.* 2012;79(3):261–6.
 321. Partridge JM, Weatherby SJM, Woolmore JA, Highland DJ, Fryer AA, Mann CLA, et al. Susceptibility and outcome in MS: associations with polymorphisms in pigmentation-related genes. *Neurology.* 2004;62(12):2323–5.
 322. Simon KC, Munger KL, Yang X, Ascherio A. Polymorphisms in vitamin D metabolism related genes and risk of multiple sclerosis. *Mult Scler.* 2010;16(2):133–8.
 323. Tajouri L, Ovcarić M, Curtain R, Johnson MP, Griffiths LR, Csurhes P, et al. Variation in the vitamin D receptor gene is associated with multiple sclerosis in an Australian population. *J Neurogenet.* 2005;19(1):25–38.
 324. Cox MB, Ban M, Bowden NA, Baker A, Scott RJ, Lechner-Scott J. Potential association of vitamin D receptor polymorphism Taq1 with multiple sclerosis. *Mult Scler.* 2012;18(1):16–22.
 325. Ramasamy A, Trabzuni D, Forabosco P, Smith C, Walker R, Dillman A, et al. Genetic evidence for a pathogenic role for the vitamin D3 metabolizing enzyme CYP24A1 in multiple sclerosis. *Mult Scler Relat Disord.* 2014;3(2):211–9.
 326. Karky M, Alcina A, Fedetz M, Barrionuevo C, Potenciano V, Delgado C, et al. The multiple sclerosis-associated regulatory variant rs10877013 affects expression of CYP27B1 and VDR under inflammatory or vitamin D stimuli. *Mult Scler.* 2016;22(8):999–1006.
 327. Souberbielle J-C, Body J-J, Lappe JM, Plebani M, Shoenfeld Y, Wang TJ, et al. Vitamin D and musculoskeletal health, cardiovascular disease, autoimmunity and cancer:

- Recommendations for clinical practice. *Autoimmun Rev.* 2010;9(11):709–15.
328. Shinki T, Jin CH, Nishimura A, Nagai Y, Ohyama Y, Noshiro M, et al. Parathyroid hormone inhibits 25-hydroxyvitamin D₃-24-hydroxylase mRNA expression stimulated by 1 alpha,25-dihydroxyvitamin D₃ in rat kidney but not in intestine. *J Biol Chem.* 1992;267(19):13757–62.
329. Veldurthy V, Wei R, Campbell M, Lupicki K, Dhawan P, Christakos S. 25-Hydroxyvitamin D₃ 24-Hydroxylase: A Key Regulator of 1,25(OH)₂D₃ Catabolism and Calcium Homeostasis. *Vitam Hormon.* 2016;100:137–50.
330. Veldurthy V, Wei R, Oz L, Dhawan P, Jeon YH, Christakos S. Vitamin D, calcium homeostasis and aging. *Bone Res.* 2016;4:16041.
331. Handunnetthi L, Ramagopalan S V, Ebers GC. Multiple sclerosis, vitamin D, and HLA-DRB1*15. *Neurology.* 2010;74(23):1905–10.
332. Baranov KO, Volkova OI, Mechetina L V, Chikaev NA, Reshetnikova ES, Nikulina GM, et al. Expression of human B-cell specific receptor FCRL1 in normal individuals and in patients with autoimmune diseases. *Mol Biol (Mosk).* 2012;46(3):500–7.
333. Comabella M, Canto E, Nurtidinov R, Rio J, Villar LM, Picon C, et al. MRI phenotypes with high neurodegeneration are associated with peripheral blood B-cell changes. *Hum Mol Genet.* 2016;25(2):308–16.
334. San Lucas FA, Wang G, Scheet P, Peng B. Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinformatics.* 2012;28(3):421–2.
335. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. Human genomics. The human transcriptome across tissues and individuals. *Science.* 2015;348(6235):660–5.
336. Hunt KA, Zhernakova A, Turner G, Heap GAR, Franke L, Bruinenberg M, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet.* 2008;40(4):395–402.
337. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, Yang JHM, et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med.* 2008;359(26):2767–77.
338. De Jager PL, Jia X, Wang J, de Bakker PIW, Ottoboni L, Aggarwal NT, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet.* 2009;41(7):776–82.
339. International Multiple Sclerosis Genetics Consortium. IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. *Genes Immun.* 2010;11(5):397–405.
340. Bansal G, Druey KM, Xie Z. R4 RGS proteins: Regulation of G-protein signaling and beyond. *Pharmacol Ther.* 2007;116(3):473–95.
341. Atwood BK, Lopez J, Wager-Miller J, Mackie K, Straiker A. Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. *BMC Genomics.* 2011;12:14.
342. Moratz C, Kang VH, Druey KM, Shi CS, Scheschonka A, Murphy PM, et al. Regulator of G protein signaling 1 (RGS1) markedly impairs Gi alpha signaling responses of B lymphocytes. *J Immunol.* 2000;164(4):1829–38.
343. Ansel KM, Cyster JG. Chemokines in lymphopoiesis and lymphoid organ development. *Curr Opin Immunol.* 2001;13(2):172–9.
344. Moratz C, Hayman JR, Gu H, Kehrl JH. Abnormal B-Cell Responses to Chemokines, Disturbed Plasma Cell Localization, and Distorted Immune Tissue Architecture in Rgs1^{-/-} Mice. *Mol Cell Biol.* 2004;24(13):5767–75.
345. Han S-B, Moratz C, Huang N-N, Kelsall B, Cho H, Shi C-S, et al. Rgs1 and Gnai2 Regulate the Entrance of B Lymphocytes into Lymph Nodes and B Cell Motility within Lymph Node Follicles. *Immunity.* 2005;22(3):343–54.
346. Allen CDC, Ansel KM, Low C, Lesley R, Tamamura H, Fujii N, et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol.* 2004;5(9):943–52.

347. Turner MD, Nedjai B, Hurst T, Pennington DJ. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochim Biophys Acta*. 2014;1843(11):2563–82.
348. Dixit N, Simon SI. Chemokines, selectins and intracellular calcium flux: temporal and spatial cues for leukocyte arrest. *Front Immunol*. 2012;3:188.
349. Sjaastad MD, Nelson WJ. Integrin-mediated calcium signaling and regulation of cell adhesion by intracellular calcium. *BioEssays*. 1997;19(1):47–55.
350. Scharenberg AM, Humphries LA, Rawlings DJ. Calcium signalling and cell-fate choice in B cells. *Nat Rev Immunol*. 2007;7(10):778–89.
351. Baba Y, Kurosaki T. Role of calcium signaling in B cell activation and biology. *Curr Top Microbiol Immunol*. 2016;393:143–74.
352. Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol*. 1996;39(3):285–94.
353. Mahurkar S, Moldovan M, Suppiah V, Sorosina M, Clarelli F, Liberatore G, et al. Response to interferon-beta treatment in multiple sclerosis patients: a genome-wide association study. *Pharmacogenomics J*. 2017;17(4):312–8.
354. Liu Z, Pelfrey CM, Coteur A, Lee JC, Rudick RA. Immunomodulatory effects of interferon beta-1a in multiple sclerosis. *J Neuroimmunol*. 2001;112(1–2):153–62.
355. Galligan CL, Pennell LM, Murooka TT, Baig E, Majchrzak-Kita B, Rahbar R, et al. Interferon-beta is a key regulator of proinflammatory events in experimental autoimmune encephalomyelitis. *Mult Scler*. 2010;16(12):1458–73.

9. ANEXO

Financiación

Este trabajo ha sido financiado por el Fondo de Investigación Sanitaria del Instituto de Salud Carlos III, Ministerio de Economía y Competitividad (PI12/02229). La doctorando ha sido financiada por la ayuda predoctoral del Institut de Recerca Vall d'Hebron. El resto de la financiación ha procedido de fondos propios del Cemcat.

Individuos incluidos en los experimentos

1. Cuantificación de los niveles de ARNm de *CXCR5* en pacientes con EM con distintos genotipos

Características	Con rs10892307 (CG)	Sin rs10892307 (GG)
n	15	18
Mujeres/Hombres (% mujeres)	9/6 (60,0%)	10/8 (55,6%)
Edad, años	43,4 ± 14,2	43,6 ± 11,0

Tabla Suplementaria 1. Características demográficas de los pacientes con EM con diferente genotipo para el polimorfismo rs10892307 incluidos para determinar la expresión génica de *CXCR5* en CMSP mediante PCR a tiempo real. La edad se expresa como media ± desviación estándar.

2. Análisis por citometría de flujo de la expresión de *CXCR5* en poblaciones de CMSP de pacientes con EM con distintos genotipos

Características	Con rs10892307 (CG)	Sin rs10892307 (GG)
n	10	10
Mujeres/Hombres (% mujeres)	7/3 (70,0%)	9/1 (90,0%)
Edad, años	34,6 ± 8,4	39,0 ± 8,7

Tabla Suplementaria 2. Características demográficas de los pacientes con EM incluidos para cuantificar la expresión de *CXCR5* en la superficie de diferentes poblaciones de CMSP mediante citometría de flujo. La edad se expresa como media ± desviación estándar.

3 Individuos incluidos en la determinación de la concentración de 25(OH)D en suero

Características	Con rs10892307 (CG)	Sin rs10892307 (GG)
n	30	71
Estación de extracción (n)		
Primavera + Verano	11 (36,7%)	26 (36,6%)
Otoño + Invierno	19 (63,3%)	45 (63,4%)
Mujeres/Hombres (% mujeres)	16/14 (53,3%)	43/28 (60,6%)
Edad, años	34,9 ± 8,9	36,0 ± 10,9

Tabla Suplementaria 3. Características demográficas y estación de extracción de la muestra de suero de los pacientes con EM incluidos para determinar los niveles de la forma 25(OH) de la vitamina D. La edad se expresa como media ± desviación estándar.

4. Individuos incluidos en la determinación de la concentración de 1,25(OH)₂-D en suero

Características	Con rs10892307 (CG)	Sin rs10892307 (GG)
n	26	74
Estación de extracción (n)		
Primavera + Verano	11 (42,3%)	36 (48,6%)
Otoño + Invierno	15 (57,7%)	38 (51,4%)
Mujeres/Hombres (% mujeres)	19/7 (73,1%)	55/19 (74,3%)
Edad, años	38,9 ± 11,7	37,9 ± 11,9

Tabla Suplementaria 4. Características demográficas y estación de extracción de la muestra de suero de los pacientes con EM incluidos para determinar los niveles de la forma fisiológicamente activa (1,25(OH)₂) de la vitamina D. La edad se expresa como media ± desviación estándar.

5. Inducción de la expresión de *CYP24A1* en respuesta a la estimulación con vitamina D en CMSP de pacientes con EM con distintos genotipos del polimorfismo rs2762943 (*CYP24A1*)

Características	Con rs2762943 (GT)	Sin rs2762943 (GG)
n	8	8
Mujeres/Hombres (% mujeres)	6/2 (75,0%)	5/3 (62,5%)
Edad, años	34,7 ± 8,1	36,7 ± 9,0

Tabla Suplementaria 5. Descripción demográfica de los pacientes con EM incluidos para analizar la expresión del gen *CYP24A1* después de la estimulación con 1,25(OH)₂-D. Los pacientes se agruparon en función de la presencia o ausencia del polimorfismo rs2762943. La edad se expresa como media ± desviación estándar.

6. Cuantificación de la expresión génica de los genes que acumulan variantes raras en pacientes con EM mediante *real-time* PCR

	<i>CYP24A</i>				<i>RGS1</i>			<i>FCRL1</i>		<i>TRAF3</i>	
Características	Raras reguladoras	Raras UTR	Raras promotor	Sin raras	Raras	Raras reguladoras	Sin raras	Raras promotor	Sin raras	Raras UTR	Sin raras
N	11	4	4	25	6	4	28	6	15	4	16
M/H (%M)	6/5 (54,5%)	2/2 (50%)	2/2 (50%)	14/11 (56%)	4/2 (66,6%)	2/2 (50%)	18/10 (64,3%)	4/2 (66,6%)	7/8 (46,6%)	4/0 (100%)	4/12 (25%)
Edad ± DE	36,9±9,5	36,5±12,1	42,2±9,1	40,1±11,4	40,5±10,4	37,5±10,2	37,5±8,5	32,5±9,1	40,7±7,1	41,2±8,0	40,2±8,3
Cemcat (n)	11	4	4	25	4	3	10	6	15	3	5
Málaga (n)	-	-	-	-	2	1	13	-	-	1	11
Girona (n)	-	-	-	-	-	-	4	-	-	-	-
Hospital Clinic (n)	-	-	-	-	-	-	1	-	-	-	-

Tabla Suplementaria 6. Características demográficas de los individuos con EM incluidos para estudiar la expresión de los genes que acumulan variantes raras en pacientes con EM respecto a los controles sanos. Los pacientes se clasificaron en función de la presencia o ausencia de al menos una variante rara en un gen o en una región concreta de dicho gen. Todos los pacientes incluidos presentan formas de inicio con brote. Se detalla la procedencia de las muestras de CMSP incluidas. La edad se expresa como media ± DE. DE: desviación estándar. M: mujeres. H: hombres.

7. Evaluación del cambio en la expresión de *RGS1* en respuesta al tratamiento con interferón β en pacientes con y sin variantes raras

Características	Con VRs <i>RGS1</i>	Sin VRs <i>RGS1</i>
n	3	3
Mujeres/Hombres (%mujeres)	1/2 (33,3%)	0/3 (0,0%)
Edad, años	38,7 \pm 13,0	34,0 \pm 4,4

Tabla Suplementaria 7. Características demográficas de los pacientes con EM incluidos para analizar la estimulación de la expresión génica de *RGS1* en respuesta al tratamiento con IFN- β . Los pacientes se agruparon teniendo en cuenta la presencia o ausencia de al menos una VR del gen *RGS1*. La edad se expresa como media \pm desviación estándar.

8. Inmunofenotipado de las CMSP que expresan *RGS1*

Características	Con VRs <i>RGS1</i>	Sin VRs <i>RGS1</i>
n	4	7
Mujeres/Hombres (%mujeres)	3/1 (75,0%)	1/6 (14,3%)
Edad, años	45,5 \pm 9,3	34,4 \pm 5,2

Tabla Suplementaria 8. Características demográficas de los pacientes con EM incluidos para estudiar la expresión proteica de *RGS1* en la cara interna de la membrana plasmática de diferentes subpoblaciones de CMSP mediante citometría de flujo. Los pacientes se agruparon teniendo en cuenta la presencia o ausencia de al menos una VR del gen *RGS1*. La edad se expresa como media \pm desviación estándar.

Protocolos

1. OBTENCIÓN Y CUANTIFICACIÓN DE ADN GENÓMICO A PARTIR DE SANGRE PERIFÉRICA

➤ **Extracción del ADN**

Día 1:

1. Centrifugar los tubos de sangre con EDTA a 2000 rpm, 10 minutos, a T^a ambiente.
2. Descartar el plasma y recoger la parte superior del sedimento, donde están los leucocitos y los eritrocitos.
3. Lisar los restos de eritrocitos mediante la adición del tampón de lisis RBC (por *red blood cells*) (NH₄Cl 14mM, NaCOH₃ 1mM, Na₂EDTA 1mM) (Sigma-Aldrich, St. Louis, MA, EEUU) hasta 50 mL.
4. Poner los tubos en agitación durante 20 minutos.
5. Centrifugar a 3500 rpm, 20 minutos, a T^a ambiente.
6. Decantar el sobrenadante y añadir de nuevo tampón RBC hasta 50 mL.
7. Poner los tubos en agitación durante 5 minutos.
8. Centrifugar a 3500 rpm, 20 minutos, a T^a ambiente.
9. Recoger el sobrenadante y añadir 9 mL de tampón *Nuclei lysis* (Sigma-Aldrich) (NaCl 0,2 M, Tris·HCl 10mM, Na₂EDTA 2mM, 0.2% de dodecilsulfato sódico (SDS)) para lisar los leucocitos y 1 mg/mL de proteinasa K (Roche, Suiza) para degradar las proteínas.
10. Incubar los tubos toda la noche a 37°C para lisar los leucocitos.

Día 2:

11. Añadir 3.5 mL de acetato de amonio a 7,5 M (Sigma-Aldrich) y dejar 15 minutos a 4°C para que precipite.
12. Centrifugar 3500 rpm, 20 minutos, a T^a ambiente para eliminar los restos celulares.
13. Recoger la fase superior del tubo, donde se encuentra el ADN.
14. Añadir el mismo volumen de etanol absoluto y vortear.
15. Centrifugar 3500 rpm, 10 minutos, a T^a ambiente.
16. Decantar el sobrenadante, añadir 10 mL de etanol 70% y poner en agitación 10 minutos.
17. Centrifugar 3500 rpm, 10 minutos, a T^a ambiente.
18. Decantar el sobrenadante y secar el tubo con papel para eliminar los restos de etanol.

19. Resuspender el ADN en tampón Tris-EDTA (Sigma-Aldrich).

➤ **Cuantificación del ADN**

En todos los casos, la cuantificación del ADN obtenido se realizó en el espectrofotómetro NanoDrop® ND-1000 (Thermo Fisher Scientific, Wilmington, EEUU) en 2 µL de muestra. La pureza del ADN se estimó mediante la evaluación del ratio 260/280, para descartar la contaminación con proteínas, fenol residual, guanidina u otro reactivo utilizado en el protocolo de extracción que absorba a 280 nm, y del ratio 260/230, para descartar la existencia de restos de EDTA, carbohidratos o fenol, que absorben a 230 nm. Las muestras se almacenaron a -20°C hasta su utilización.

2. EXTRACCIÓN DEL ARN TOTAL A PARTIR DE CMSP MEDIANTE TRIZOL™

1. El día anterior a la extracción, traspasar las CMSP de nitrógeno líquido a -80°C.
2. Lavar cada vial de células con 10 mL de PBS 1×.
3. Centrifugar a 1500 rpm, 10 minutos, a 4°C.
4. Aspirar el sobrenadante y vortear el *pellet*.
5. Añadir 1 mL de TRIzol™ Reagent (Invitrogen).
6. Vortear 1 minuto y dejar 5 minutos a T^a ambiente.
7. Añadir 200 µL de cloroformo, vortear 20 segundos y dejar 3 minutos a T^a ambiente.
8. Centrifugar a 11200 rpm, 15 minutos, a 4°C.
9. Recoger la fase superior y trasladarla a un *ependorf* nuevo.
10. Añadir 500 µL de 2-propanol y dejar 7 minutos a T^a ambiente.
11. Centrifugar a 11200 rpm, 10 minutos, a 4°C.
12. Decantar el sobrenadante y añadir 1 mL de etanol 70%.
13. Centrifugar a 11200 rpm, 10 minutos, a 4°C.
14. Decantar el sobrenadante y secar los restos de etanol con papel de filtro.
15. Añadir 20 µL de agua RNasa-free, resuspender y cuantificar en el NanoDrop® ND-1000 del mismo modo que el ADN.

3. FASES DEL GENOTIPADO DE SNPs SELECCIONADOS

Fase 1. Amplificación simultánea de todos los fragmentos de ADN (~100bp) que contienen los SNPs de interés mediante una PCR *multiplex*. Las reacciones de PCR se establecieron en un volumen de 5 μ l y contenían 20 ng de ADN, PCR buffer 1 \times , MgCl₂ 2 mM, dNTPs 500 μ M y 1 U de enzima. Se utilizó un *pool* de *primers* a una concentración final de 100 nM para cada uno de ellos (IDT, Newark, EEUU). Las condiciones de la reacción PCR *multiplex* se resumen en la Tabla Suplementaria 9.

	Inicio	Desnaturalización	Alineamiento	Elongación	Elongación final
Temperatura (°C)	95	95	56	72	72
Tiempo	2 min	30 s	30 s	1 min	5 min
Número de ciclos	1	←	45	→	1

Tabla Suplementaria 9. Condiciones de la reacción PCR *multiplex* constituyente de la primera fase del proceso de genotipado de las variantes comunes seleccionadas.

Fase 2. Tratamiento de los productos de la PCR con 1.7 U de la enzima fosfatasa alcalina (SAP, por *Shrimp Alkaline Phosphatase*) a 37°C durante 40 minutos, seguido de la inactivación enzimática por calentamiento a 85°C durante 5 minutos, para neutralizar los dNTPs no incorporados.

Fase 3. Discriminación alélica. Las reacciones de extensión se establecieron en un volumen final de 9 μ l y contenían 0.222x de iPLEX buffer Plus, 0.222x de iPLEX *Termination mix* y 1.35 U/reacción de enzima iPLEX. En la *extension primer mix*, la concentración final de cada primer fue entre 0.52 μ M y 1.57 μ M (IDT, Integrated DNA technologies, Newark, EEUU). Las condiciones del termociclador fueron las descritas en la Tabla Suplementaria 10. Todas las reacciones terminan tras una extensión de una base que genera una diferencia de masa en el producto determinada por la base añadida.

	Inicio	Desnaturalización	Alineamiento	Elongación	Elongación final
Temperatura (°C)	95	90	52	80	72
Tiempo	30 s	5 s	5 s	5 s	3 min
Número de ciclos	1	←	40	→	1

Tabla Suplementaria 10. Condiciones de la reacción PCR de discriminación alélica integrante de la tercera y última fase del proceso de genotipado.

4. MARCAJE CELULAR PARA EL ANÁLISIS MEDIANTE CITOMETRÍA DE FLUJO

➤ Marcaje de superficie directo

1. Lavado de las células con 200 μ L de PBS 1 \times .
2. Marcaje de viabilidad: se empleó el Fixable Viability Dye eFluor™ 450 (eBioscience) o con Fixable Viability Stain 510 (BD Biosciences) a una concentración 1:2000 o 1:1000, respectivamente, para la discriminación entre células vivas y muertas. Se siguió el protocolo indicado por el fabricante de los reactivos.
3. Bloqueo de los receptores de superficie Fc para evitar las uniones inespecíficas: se utilizó el Human BD Fc Block™ (BD Biosciences), que evita la potencial unión de los anticuerpos de manera antígeno-independiente a dichos receptores. Incubación 10' a temperatura ambiente.
4. Adición de los anticuerpos específicos y de los controles de isotipo conjugados con los fluorocromos seleccionados. Incubación 30' a 4°C en oscuridad.
5. Lavado con 200 μ L de PBA-azida (PBS + BSA 1% + azida sódica 0,1%).
6. Resuspensión de las células en 200 μ L de PBA-azida para adquirirlas en el citómetro.

➤ Marcaje intranuclear directo

1. Adición de 200 μ L de Buffer de Fijación/Permeabilización a las células tras el último lavado del marcaje de superficie. Incubación 30' a temperatura ambiente y en oscuridad.
2. Centrifugación de las muestras a 1500 rpm, 5', y descarte del sobrenadante.
3. Dos lavados con 200 μ L de Buffer de Permeabilización 1 \times .
4. Adición de 100 μ L de Buffer de Permeabilización 1 \times y 2 μ L de suero de rata al 2% para bloquear las uniones inespecíficas intracelulares. Incubación 15' a temperatura ambiente.
5. Adición del anticuerpo anti-FoxP3. Incubación 30' a temperatura ambiente y en oscuridad.
6. Dos lavados con 200 μ L de Buffer de Permeabilización 1 \times .
7. Resuspensión de las células en 200 μ L de PBA-azida para adquirirlas en el citómetro.

5. DETERMINACIÓN DE LOS NIVELES DE CALCIDIOL Y CALCITRIOL EN SUERO

➤ Determinación de los niveles de calcidiol en suero

Las muestras de suero se descongelaron progresivamente, primero a -20°C y después a 4°C . El volumen mínimo necesario es $175\ \mu\text{L}$ de muestra. En primer lugar, se dispensaron las micropartículas magnéticas (con anticuerpos de captura en su superficie para ligar el analito a determinar) y el tampón del ensayo (etanol al 10%, surfactantes y ProClin®300 al 0,2%) en el módulo de la reacción y se añadió la muestra, el control o el calibrador. Durante una primera incubación, la 25-OH-D se disocia de su proteína transportadora y se une a su anticuerpo específico en la fase sólida. Transcurridos 10 minutos, se añadió el trazador de la 25-OH-D, que es un antígeno unido a un compuesto fluorescente (vitamina D unida a un derivado de isoluminol) que competirá con el antígeno objeto de la medición (25-OH-D). Después de una segunda incubación de 10 minutos, el material libre se eliminó en un ciclo de lavado. A continuación se agregaron los reactivos iniciadores para dar lugar a una reacción quimioluminiscente rápida. A mayor concentración de 25-OH-D en la muestra, menor será la cantidad de trazador unido al anticuerpo. La señal luminosa se midió en unidades lumínicas relativas (RLU, por *relative light units*) con un fotomultiplicador y es inversamente proporcional a la concentración de 25-OH-D presente en calibradores, controles o muestras de los pacientes.

➤ Determinación de los niveles de calcitriol en suero

Las muestras de suero se descongelaron progresivamente, primero a -20°C y después a 4°C . El volumen de suero requerido para la prueba es de $75\ \mu\text{L}$ por muestra más $150\ \mu\text{L}$ de volumen muerto. En primer lugar, se dispensaron el tampón de ensayo (Tampón fosfato con bloqueantes, IgG de ratón, albúmina sérica bovina, Tween20, sulfato de gentamicina al 0,01% y ProClin®300 al 0,2%) y el agente reductor TCEP (Tris (2-carboxietil) fosfina). A continuación, se añadieron los calibradores, los controles o las muestras y se dispuso la RFP, que cuando se une a la molécula de $1,25\text{-(OH)}_2\text{-D}$, cambia su conformación. Tras una primera incubación, se dispensaron partículas magnéticas revestidas con el anticuerpo monoclonal de ratón específico (fase sólida) y se procedió a una segunda incubación para inmovilizar al complejo RFP- $1,25\text{-(OH)}_2\text{-D}$. Después de lavar para eliminar el material no unido, se añadió el conjugado (anticuerpo monoclonal de ratón conjugado con un derivado del isoluminol, que reconoce al complejo RFP- $1,25\text{-(OH)}_2\text{-D}$ ya unido a la fase sólida) y se volvió a incubar. El material no unido se elimina con un segundo ciclo de lavado. Después, se agregaron los reactivos inductores y tuvo lugar una reacción quimioluminiscente instantánea. La señal luminosa se midió en RLU

con un fotomultiplicador y es proporcional a la concentración de 1,25-(OH)₂-D presente en los calibradores, los controles y las muestras del paciente.

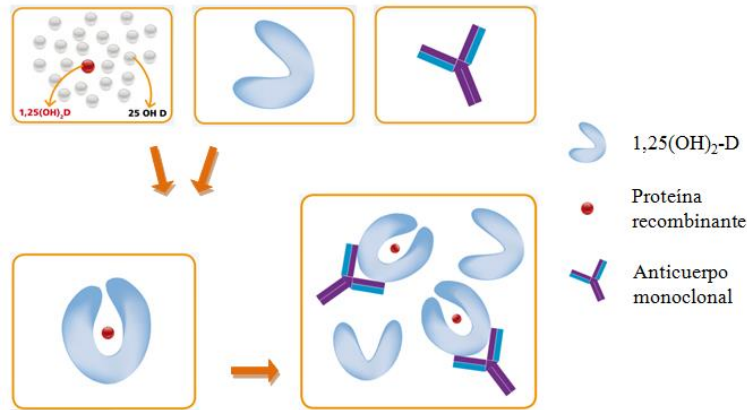
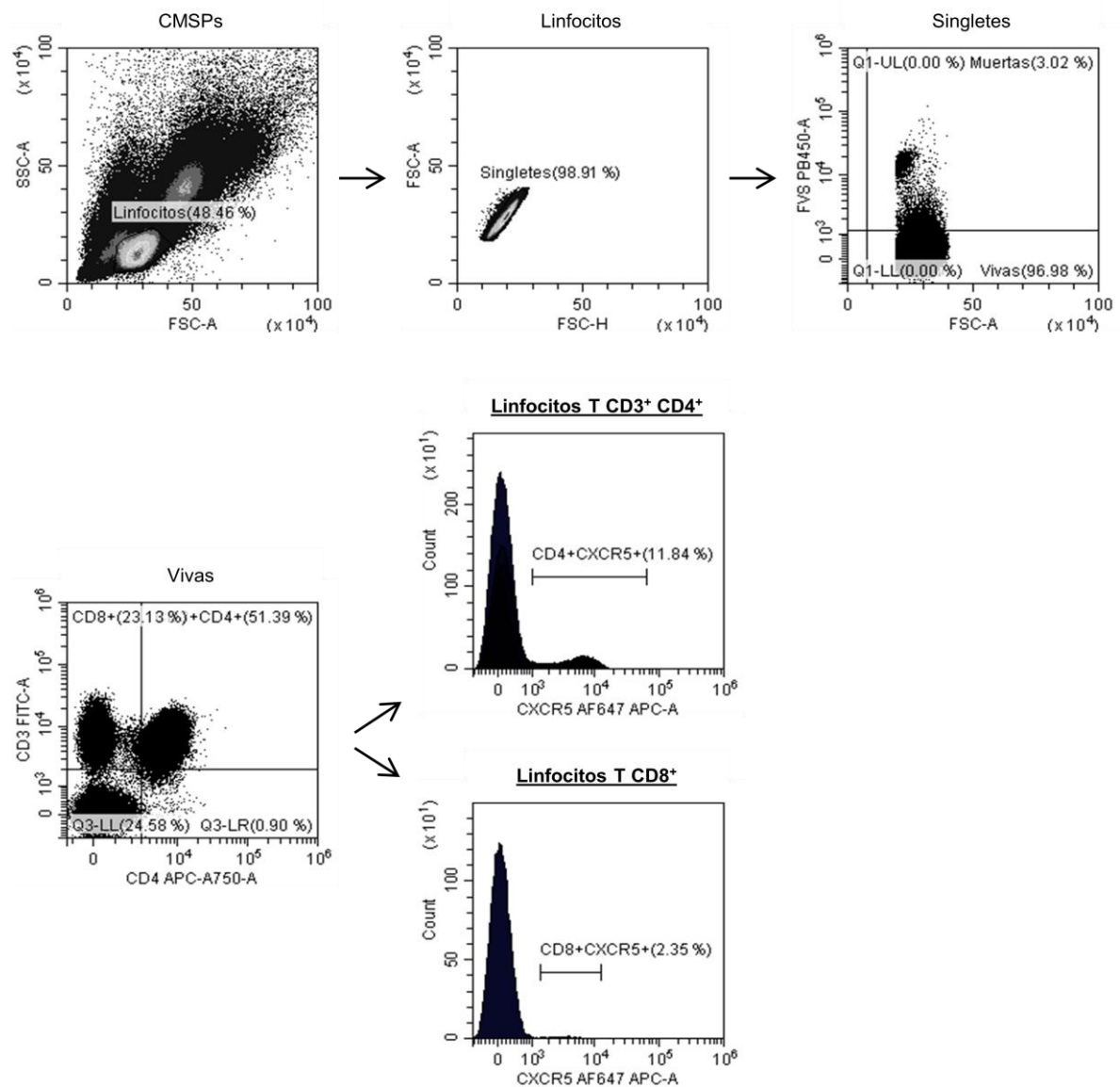


Figura Suplementaria 1. Explicación gráfica del principio básico del kit *LIAISON® XL 1,25 Dihydroxyvitamin D Assay*, de DiaSorin, para determinar los niveles de 1,25(OH)₂-D en suero. El ensayo consiste en un inmunoensayo quimioluminiscente de tipo sándwich en 3 pasos que utiliza una proteína quimérica (RFP) para capturar la molécula de 1,25(OH)₂-D y un anticuerpo monoclonal murino que reconoce el complejo RFP-1,25(OH)₂-D. Adaptada de <https://www.diasorin.com>.

Citometría de flujo para CXCR5: estrategia de *gating*

1. Panel 1: células T (excepto T reguladoras)

La estrategia de *gating* empleada para identificar las poblaciones seleccionadas de células T que expresan el receptor CXCR5 en su superficie fue la representada mediante la Figura Suplementaria 2.



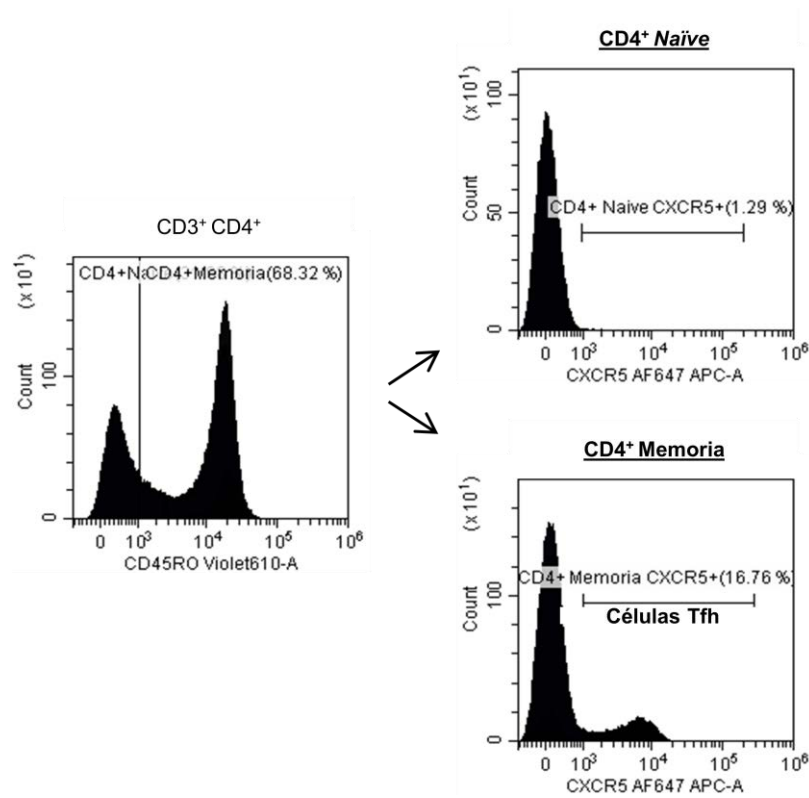


Figura Suplementaria 2. Estrategia de *gating* para analizar por citometría de flujo la expresión superficial de CXCR5 en las siguientes poblaciones de células T: linfocitos T CD4⁺, linfocitos T CD8⁺, linfocitos T CD4⁺ *naïve*, y linfocitos T CD4⁺ memoria. En el caso de que una célula T CD4⁺ memoria sea positiva para CXCR5, se considera una célula Tfh.

2. Panel 2: células T reguladoras

La estrategia de *gating* empleada para la detección de las células T reguladoras que expresaban el receptor de quimiocinas CXCR5 en su superficie fue la representada mediante la Figura Suplementaria 3.

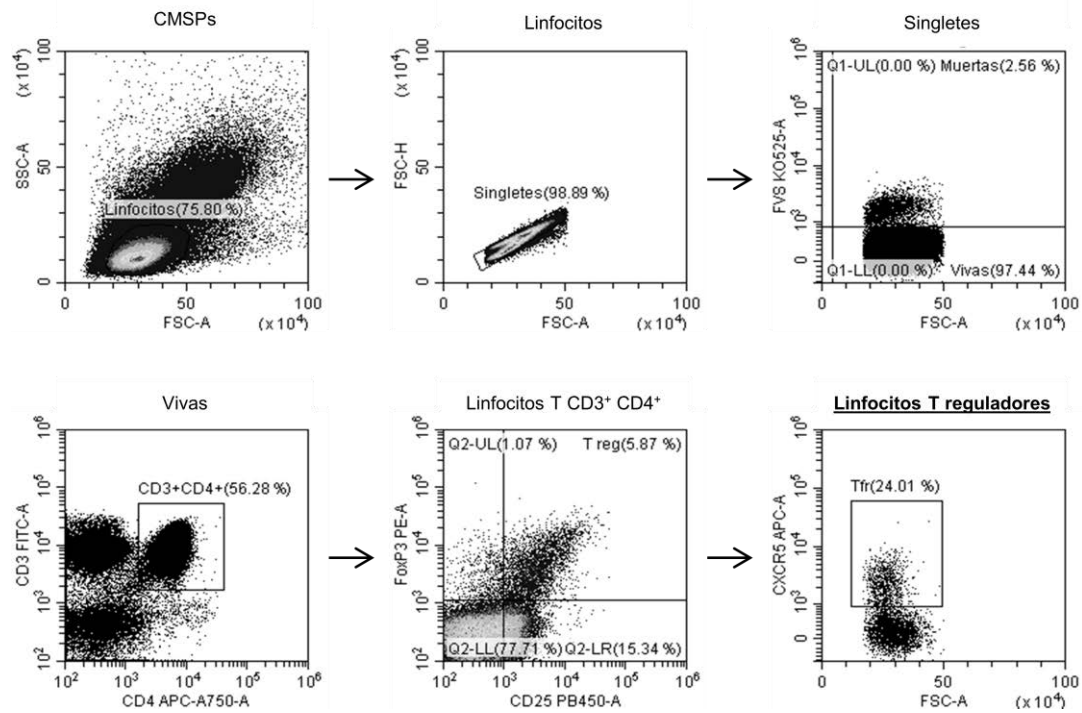
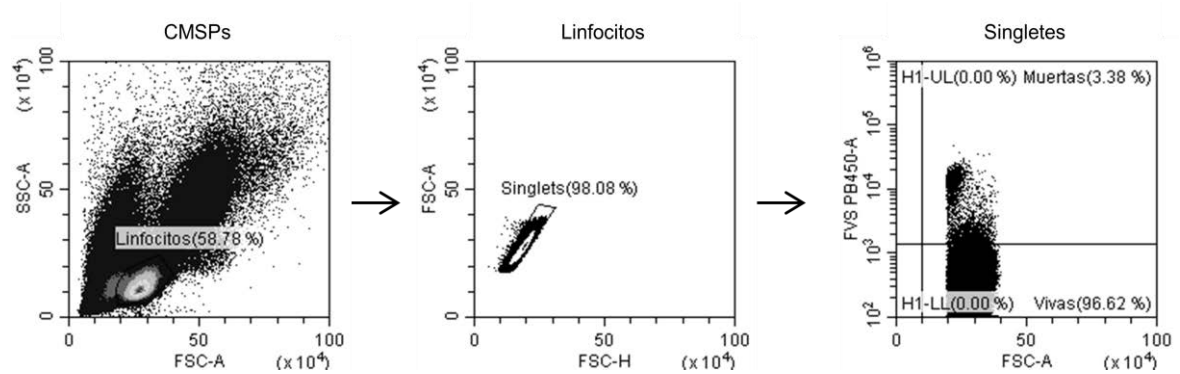


Figura Suplementaria 3. Estrategia de *gating* por citometría de flujo para la detección de los linfocitos T reguladores que expresan CXCR5 en su superficie y que, por tanto, son células T reguladoras foliculares.

3. Panel 3: células B

La estrategia de *gating* seguida para la identificación de las diferentes poblaciones de células B que expresaban CXCR5 en su superficie fue la ilustrada en la Figura Suplementaria 4.



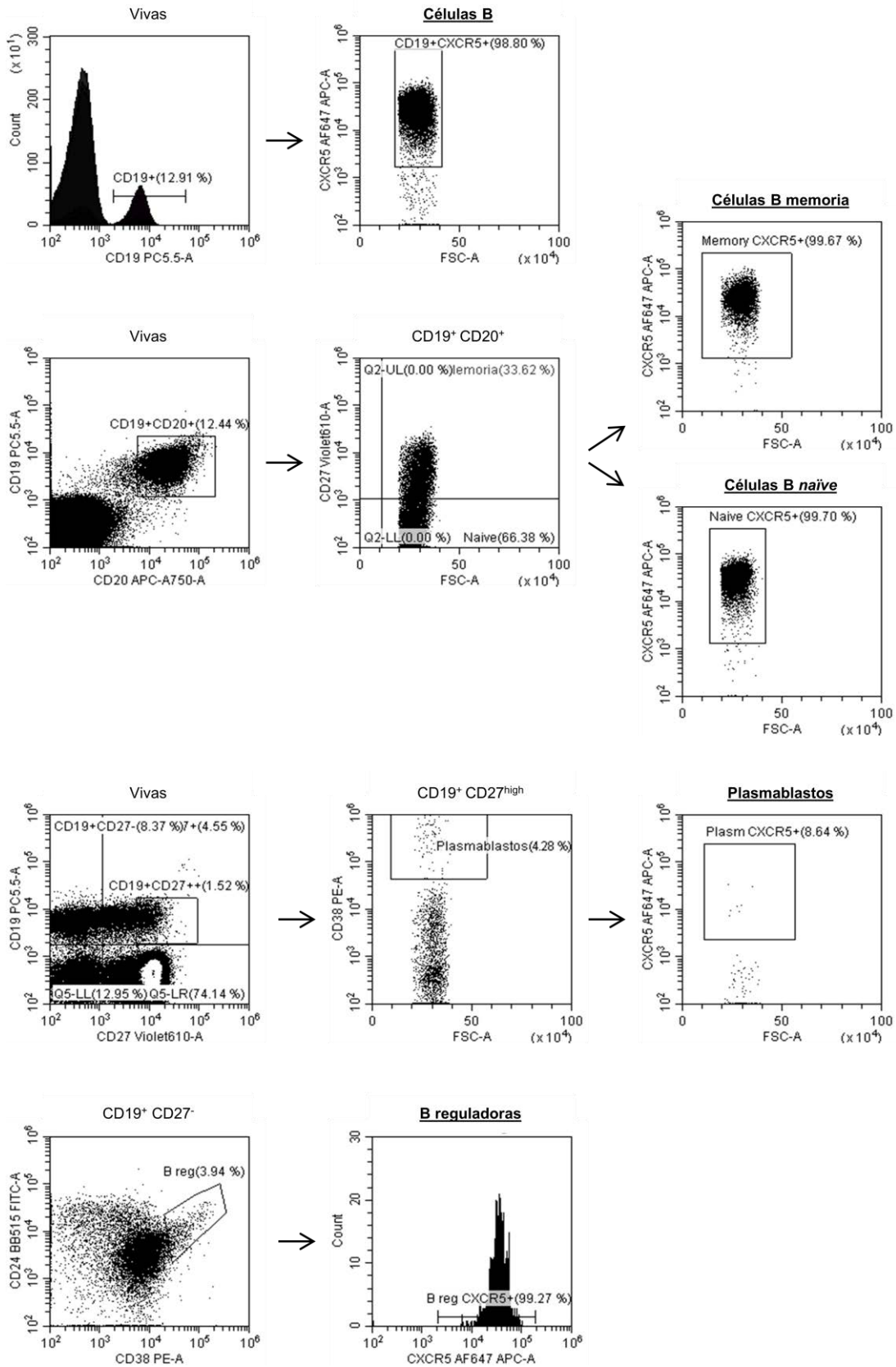
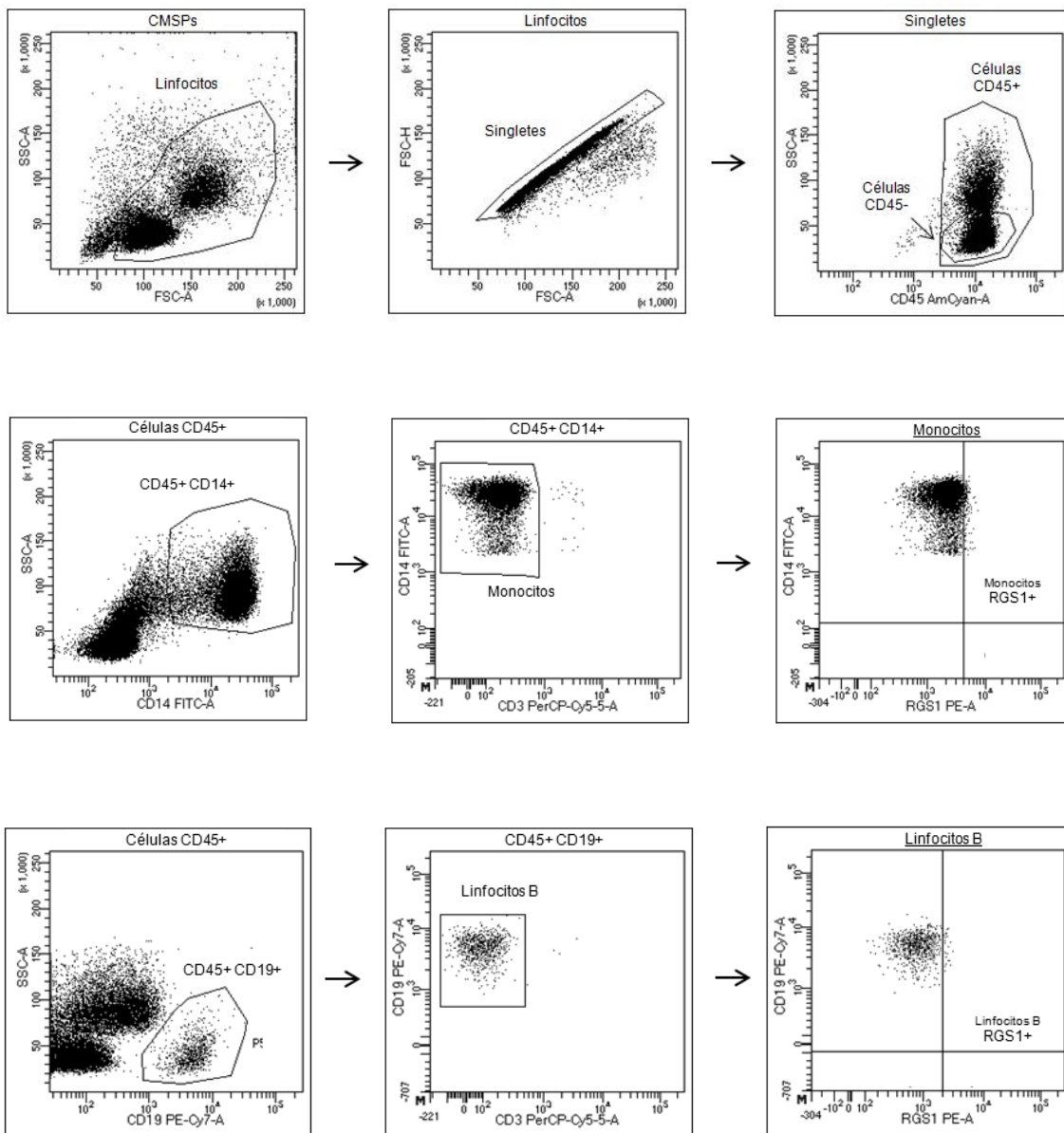


Figura Suplementaria 4. Estrategia de *gating* para delimitar mediante citometría de flujo las células B CXCR5⁺, las células B memoria CXCR5⁺, las células B *naïve* CXCR5⁺, los plasmablastos CXCR5⁺ y las células B reguladoras CXCR5⁺.

Citometría de flujo para RGS1: estrategia de *gating*

1. Panel 1: monocitos, linfocitos B, células NK y células NKT

La estrategia de *gating* utilizada para identificar las poblaciones de CMSP incluidas en el panel 1 que expresan el RGS1 en la cara interna de su membrana plasmática fue la representada mediante la Figura Suplementaria 5.



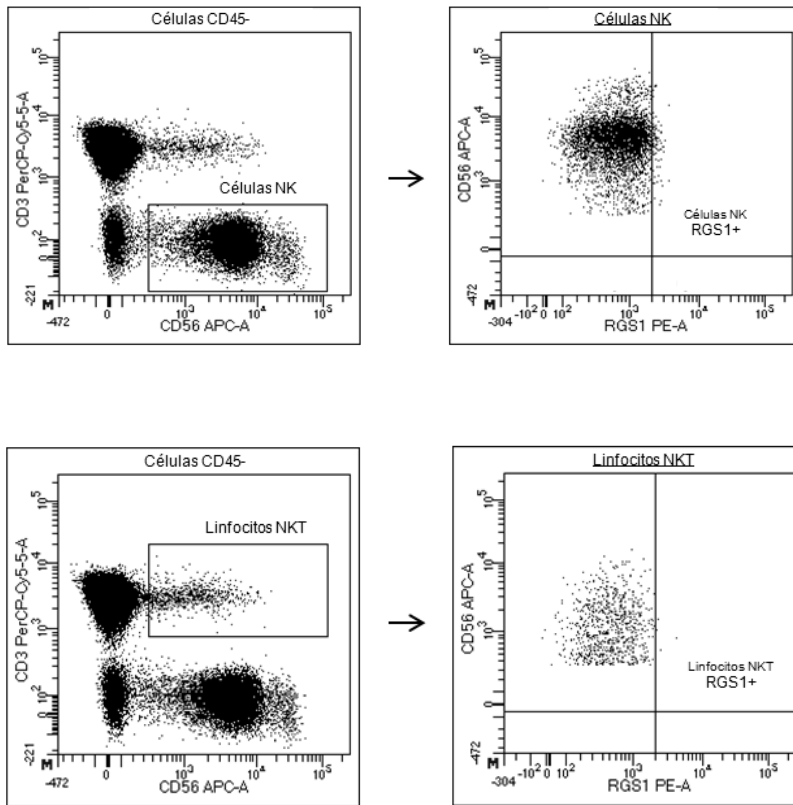
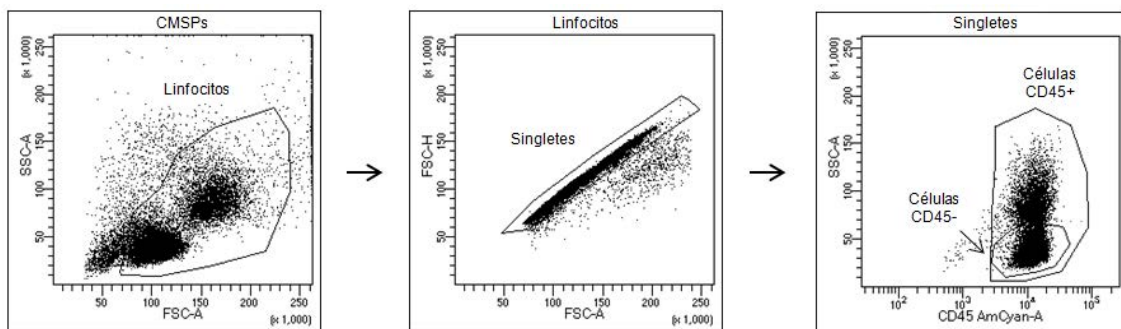


Figura Suplementaria 5. Estrategia de *gating* utilizada para delimitar mediante citometría de flujo los monocitos, linfocitos B, células NK y células NKT que expresan RGS1 en la cara interna de su membrana plasmática.

2. Panel 2: linfocitos T, CD4⁺, CD8⁺ y Treg

La estrategia de *gating* utilizada para identificar las poblaciones de CMSP incluidas en el panel 2 que expresan el RGS1 en la cara interna de su membrana plasmática fue la esquematizada en la Figura Suplementaria 6.



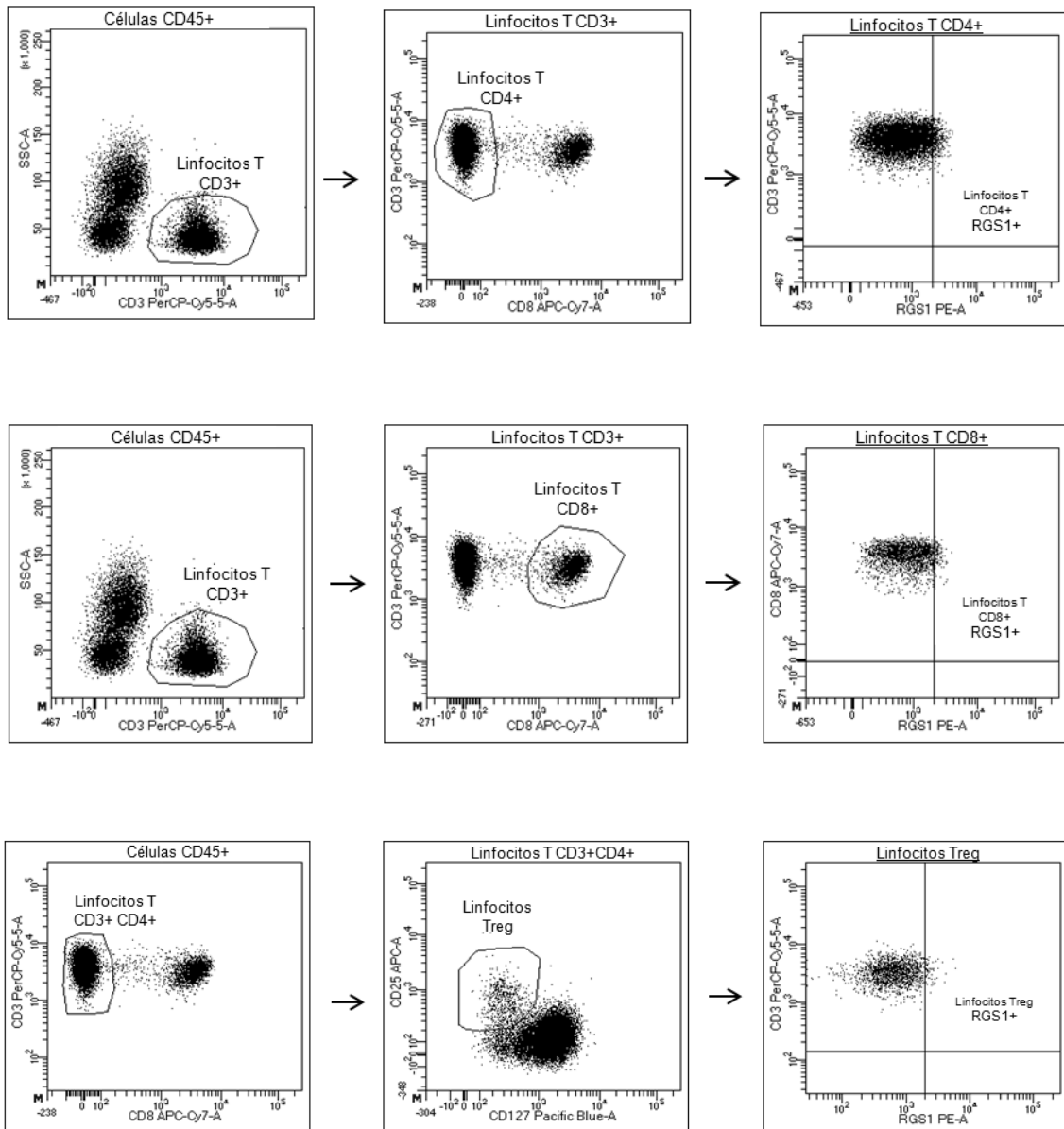


Figura Suplementaria 6. Estrategia de *gating* seguida para delimitar mediante citometría de flujo los linfocitos T, CD4⁺, CD8⁺ y Treg que expresan RGS1 en la cara interna de su membrana plasmática.

Beginnings are usually scary and endings are usually sad, but it's everything in between that makes it all worth living.

Bob Marley