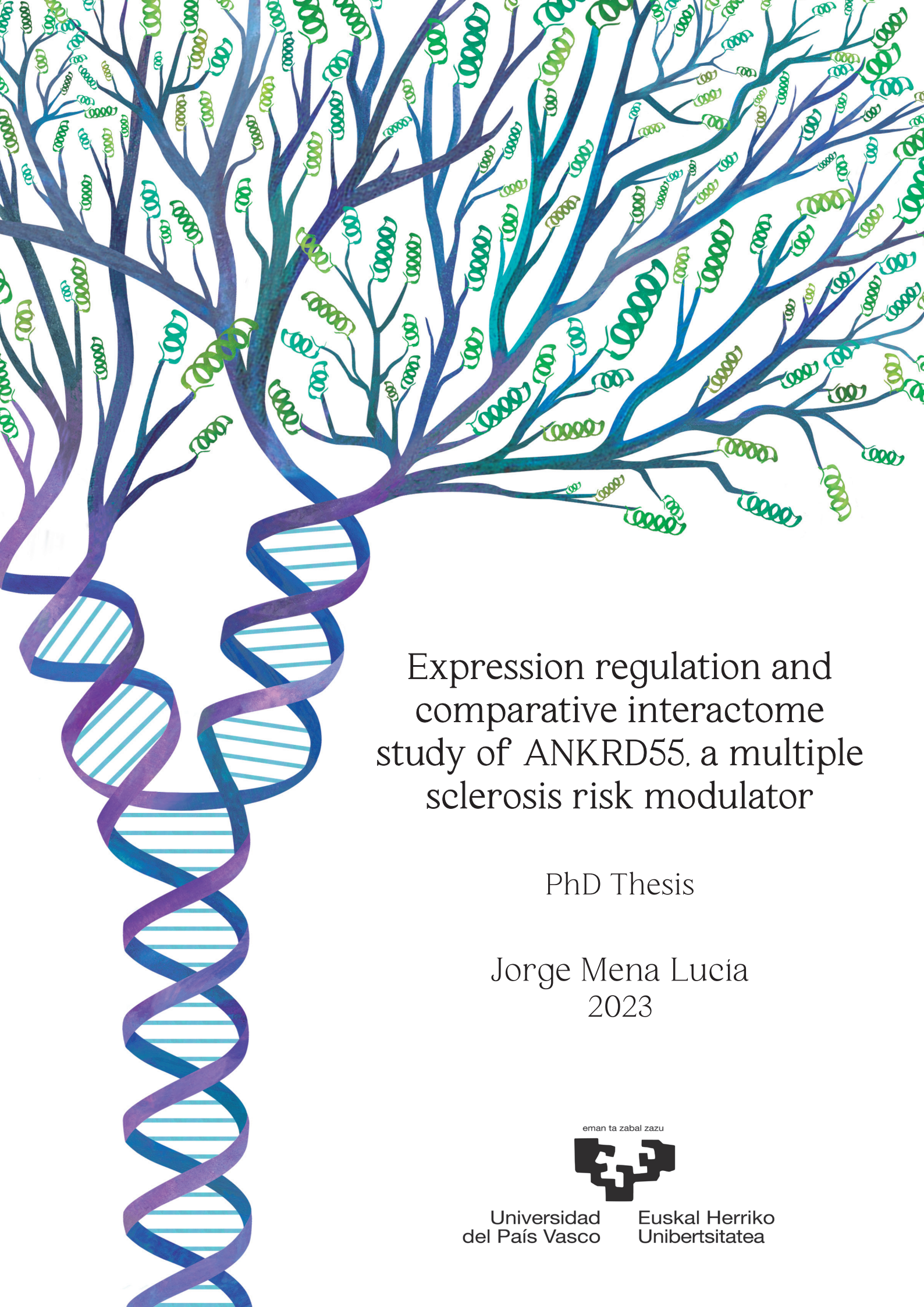


Expression regulation and comparative interactome study of ANKRD55, a multiple sclerosis risk modulator

PhD Thesis

Jorge Mena Lucía

2023

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Expression regulation and comparative interactome study of ANKRD55, a multiple sclerosis risk modulator

PhD Thesis

2023

Author: **Jorge Mena Lucía**

Director: **Koen Vandenbroeck**

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)
Facultad de Medicina y Enfermería. Departamento de Neurociencias

Diseño de portada: **Teresa Arroyo Corcobado**

“No sé lo que le podré parecer al mundo, pero a mí me parece que sólo he sido como un niño jugando en la orilla del mar, que se divierte al encontrar de vez en cuando un guijarro más liso o una concha más bonita de lo común, mientras el inmenso océano de la verdad se extiende sin descubrir frente a mí.”

Isaac Newton

"I do not know what I may appear to the world; but to myself I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me."

Isaac Newton

Agradecimientos

Esta tesis doctoral es el resultado de años de esfuerzo y trabajo. No ha sido fácil ni corta ni ha estado exenta de obstáculos. Pero una de las ventajas de ser corredor de maratones es que te prepara para este tipo de caminos largos y duros. Durante esta tesis he pasado, entre otras muchas cosas, por dos mudanzas de laboratorio, dos estancias, una pandemia mundial, cambios personales y muchos retos propios de toda investigación científica. He descubierto que ser investigador implica también tener una parte de escritor, diseñador, ingeniero, informático… Y de todos los pasos que ha tenido este camino me llevo un gran aprendizaje. También, al igual que ocurre a menudo en las maratones, ha habido momentos en los que pensé que no sería capaz de terminarla. Pero de la misma forma que para conseguir completar los 42 km de una carrera muchas veces se necesita el ánimo del público, no habría sido capaz de terminar esta tesis sin la ayuda y el apoyo de mucha gente imprescindible que, durante todos estos años y de una u otra manera, me ha aportado tanto. Estas líneas son de agradecimiento para todos vosotros.

En primer lugar, gracias a Koen, mi director de tesis, por haberme dado la oportunidad y por confiar siempre en mí. Por tus consejos y ayuda en la ciencia. Y porque en estos años hemos aprendido a conocernos y entendernos mutuamente. A Iraide, por ser tan polifacética, por estar siempre dispuesta a colaborar y por ayudarme tanto en el laboratorio como en toda la parte administrativa y burocrática (menos divertida pero tan necesaria). Por supuesto también a todos los que han sido mis compañeros en este laboratorio. Desde los primeros, Paloma, Nerea, Haize, Cristina, Andoni, Pepe y Janire, hasta los más recientes, Raquel (pronto llegará también tu momento) y Ane. Gracias por la complicidad, la ayuda y las risas tan necesarias, porque han sido muchos momentos juntos. A todos los estudiantes de grado, máster o ciclos formativos, porque con vosotros he aprendido a enseñar. También agradezco la ayuda de toda la gente de la UPV/EHU, Biocruces y Achucarro. Habéis sido muchos los que he conocido durante este tiempo y todos tenéis un lugar aquí. Siempre recordaré con cariño a toda la gente de los laboratorios donde hice las estancias, tanto del grupo de Manuel Comabella, en el Vall d'Hebron de Barcelona, como del grupo de Stephen Eyre, en la Universidad de Manchester. Fueron experiencias inolvidables donde aprendí mucho. Y gracias también al laboratorio de Felix Elortza del CIC bioGUNE, en especial a Mikel, por toda la ayuda la parte de la proteómica.

Infinitas gracias a mis padres, Chema y Marisol, por haberme apoyado tanto en todos los aspectos. Por animarme, darme tantos buenos consejos, estar siempre ahí y confiar en mí. Sé que parecía que este momento no iba a llegar nunca, pero aquí está por fin. Y a mi hermanita

Inés, por todos los bonitos años compartidos juntos, en la vida y especialmente en la etapa de Bilbao. Por tener siempre una sonrisa, por ayudarme y por abrir el camino de los doctores de esta familia. También a Marc, amigo antes que familia política. Y todos mis mejores deseos para la pequeña Live.

A toda(s) mi(s) familia(s): "Los Lucía *et al.*", "Los Mena", "Los Bravo" y al "Equipo Salamandra" de primos de Madrid, por preguntarme siempre por mi investigación y animarme con mi labor de científico. Os siento siempre cerca a pesar de la distancia. También quiero dar las gracias a la familia de Sandra, por todos sus ánimos y por haberme tratado siempre tan bien. Y quiero recordar y dedicar especialmente esta tesis a los que durante estos años nos han dejado físicamente, aunque siempre estarán muy presentes en mí. A mis abuelos Honorio y Delfina, por todo el cariño y tantos años vividos junto a ellos. Porque sé lo orgullosos que habrían estado de esto. Y a mi primo Mario, que era mucho más que un primo. Todas las experiencias y risas que vivimos juntos jamás se olvidarán. Gracias por todas tus enseñanzas y por el amor compartido por la naturaleza y por la ciencia.

Gracias una y mil veces a mis amigos de toda la vida. A mis queridos mangurrinos Biza, Uro y Teresuki (muchas gracias por la portada). A Mariete, a Quique, a Cris. Porque por muchos años que pasen y kilómetros que nos separen, sé que siempre estáis ahí. Sois esenciales. A toda la demás gente de Cáceres, de Euskadi y de Cataluña. Tanto a los que ya estaban de antes como a los que he conocido durante este tiempo, porque todos habéis sido y sois también importantes. Una mención especial para la inolvidable "Bilbo peña", porque siempre recordaré los maravillosos años que compartimos en esa bonita tierra, a la que llegué sin conocer nada ni a nadie y en la que me acogisteis tan bien. Eskerrik asko benetan!

Y, sobre todo, gracias de corazón a Sandra, mi pareja. Por quererme, entenderme y cuidarme. Por apoyarme, motivarme y hacerme sentir importante. Por crecer conmigo y por creer en mí. Por estar siempre a mi lado, incluso (o sobre todo) en los malos momentos. Realmente, todas las cosas bonitas que siento por ti no caben en estas líneas. No dejes nunca de tener un corazón tan grande. Mil gracias por interesarte siempre por mi trabajo. Tanto, que ya eres toda una experta en genética y biología molecular. Siempre hablamos de las casualidades de la vida y, de hecho, de una forma casual, gracias a esta tesis nuestras historias se cruzaron aquel verano en Barcelona. Desde entonces han sido innumerables las experiencias preciosas que hemos vivido juntos. La tesis (por fin) acaba aquí, pero a nosotros aún nos queda mucho por compartir.

Index

List of figures

List of tables

List of abbreviations

APC: Antigen-presenting cell

BBB: Blood-brain barrier

BSA: Bovine serum albumin

cDC: Conventional dendritic cell

CIS: Clinically isolated syndrome

CMV: Cytomegalovirus

CNS: Central nervous system

CSF: Cerebrospinal fluid

DAPI: 4',6-diamidino-2-phenylindole

DC: Dendritic cell

ddPCR: Droplet digital PCR

DMEM: Dulbecco's Modified Eagle's medium

DMTs: Disease modifying therapies

EAE: Experimental autoimmune encephalomyelitis

EBV: Epstein-Barr virus

EDSS: Expanded disability status scale

EGFP: Enhanced green fluorescent protein

ELISA: Enzyme-linked immunosorbent assay

eQTL: Expression quantitative trait loci

FACS: Fluorescence-activated cell sorting

FASP: Filter aided sample preparation

FBS: Fetal bovine serum

FGF-2: Fibroblast growth factor 2

FDR: False discovery rate

gDNA: Genomic DNA

GM-CSF: Granulocyte macrophage colony-stimulating factor

GWAS: Genome-wide associations study

HLA: Human leukocyte antigen

HSCT: Hematopoietic stem cell transplantation

HSV: Herpes simplex virus

IF: Immunofluorescence

IFN-β: Interferon beta

IFN-γ: Interferon gamma

IGF-1: Insulin-like growth factor 1

IgG: Immunoglobulin G

IL: Interleukin

IMhu-A: Immortalized human astrocytes-SV40

IMhu-M: Immortalized human microglia-SV40

IPA: Ingenuity pathway analysis

LD: Linkage disequilibrium

LPS: Lipopolysaccharide

MBP: Myelin basic protein

MHC: Major histocompatibility complex

miRNA: microRNA

moDC: Monocyte-derived dendritic cell

MOG: Myelin oligodendrocyte glycoprotein

MRI: Magnetic resonance imaging

MS: Multiple sclerosis

mTORC1: Mechanistic target of rapamycin complex 1

NK: Natural killer

NSAF: Normalized spectral abundance factor

PBMC: Peripheral blood mononuclear cell

PBS: Phosphate buffered saline

pDC: Plasmacytoid dendritic cell

PGE2: Prostaglandin E2

PLP: Myelin proteolipid protein

PPMS: Primary progressive multiple sclerosis

PRMS: Progressive-relapsing multiple sclerosis

qPCR: Quantitative polymerase chain reaction

RAR-α: Retinoic acid receptor alpha

RAR-γ: Retinoica cid receptor gamma

RNA-seq: RNA sequencing

ROI: Regions of interest

ROS: Reactive oxygen species

RPMI: Roswell Park Memorial Institute

RRMS: Relapsing-remitting multiple sclerosis

SEM: Standard error of the mean

SNP: Single nucleotide polymorphism

SPMS: Secondary progressive multiple sclerosis

STDEV: Standard deviation

TBS: Tris-buffered saline

Th: T helper cell

TLR: Toll-like receptor

TMEV-IDD: Theiler's murine encephalomyelitis virus-induced demyelinating disease

TPM: Transcripts per million

Treg: Regulatory T cell

UVR: Ultraviolet radiation

VDR: Vitamin D receptor

Resumen

El objetivo fundamental de este trabajo es el estudio de la expresión del gen *ANKRD55*, considerado un factor genético de riesgo de la esclerosis múltiple, así como el análisis en distintos tipos celulares del interactoma de la proteína codificada por dicho gen. La esclerosis múltiple es una patología crónica e inflamatoria que afecta al sistema nervioso central y constituye la primera causa de discapacidad de origen neurológico no traumática en adultos jóvenes. Se trata de una enfermedad autoinmune y neurodegenerativa que es más común en mujeres que en hombres y cuya manifestación suele comenzar en torno a los 30 años de edad. Se caracterizar por la destrucción de la vaina de mielina que recubre los nervios, con lo que la transmisión del impulso nervioso queda gravemente dañada, provocando finalmente una discapacidad física y un deterioro cognitivo severo en las personas que la padecen. Hasta la fecha no se ha logrado demostrar un único desencadenante de la esclerosis múltiple, si no que se piensa que se trata de una enfermedad multifactorial en la que influyen tanto componentes genéticos como ambientales.

El diagnóstico de la esclerosis múltiple se realiza mediante pruebas clínicas como la resonancia magnética y la detección de bandas oligoclonales en el líquido cefalorraquídeo. La enfermedad puede ser clasificada en varios subtipos. La mayoría de pacientes muestran la forma denominada remitente-recurrente (RR), caracterizada por la alternancia de episodios de recaídas en forma de brotes y períodos de estabilidad. A medida que la enfermedad avanza, muchos de estos pacientes alcanzan la forma secundaria progresiva (SP), en la cual se produce un proceso gradual e irreversible de daños neurológicos y un aumento de la discapacidad. Algunos pacientes, en cambio, presentan desde el comienzo la llamada forma primaria progresiva (PP), la cual no presenta recaídas agudas y tiene un desarrollo neurodegenerativo continúo más lento que la forma secundaria. Pese a los años de investigación científica, no existe aún una cura para la esclerosis múltiple. Sin embargo, gracias a los distintos tratamientos farmacológicos que se han ido desarrollando durante las últimas décadas, es posible paliar parcialmente algunos de los síntomas y retrasar el avance de la enfermedad, consiguiendo que los pacientes alcancen una mayor esperanza de vida con una calidad aceptable. Aun así, los estudios genéticos y moleculares siguen siendo claves para comprender la enfermedad y poder encontrar dianas terapéuticas más eficaces.

Existen numerosos factores, tanto ambientales como genéticos, que están fuertemente asociados con el riesgo de desarrollar la enfermedad. Entre los factores ambientales de riesgo se encuentran la infección por el virus de Epstein-Barr, el humo del tabaco, los niveles bajos de

vitamina D, la obesidad, la alteración de la microbiota y el estrés psicológico. En cuanto al componente genético, los estudios de asociación del genoma completo (GWAS) han permitido identificar hasta la fecha más de 230 variantes genéticas de riesgo asociadas a la susceptibilidad de padecer esclerosis múltiple. Muchas de estas variantes genéticas son polimorfismos de un solo nucleótido (SNP) relacionados con genes implicados en la respuesta inmune. Además de los SNPs, también se han identificado efectos epigenéticos, variantes genéticas raras, mecanismos de regulación génica e interacciones gen-ambiente relacionados con la enfermedad.

Uno de los genes cuyas variantes se han asociado con un aumento del riesgo de esclerosis múltiple, así como de otras enfermedades autoinmunes, es *ANKRD55*. Este gen, de función todavía desconocida, se localiza en el cromosoma 5q11.2 y se expresa en diversos tejidos humanos y tipos celulares. De entre ellos, los glóbulos blancos presentan uno de los mayores niveles de expresión de *ANKRD55*. Estas células, así como otras del sistema inmune, juegan un papel crucial en la patogénesis de la esclerosis múltiple debido a que atacan al propio sistema nervioso causando graves lesiones. En este proceso, son de especial relevancia tanto los linfocitos T $CD4^+$, que secretan citoquinas proinflamatorias con efectos dañinos, como las células dendríticas, encargadas de presentar los antígenos y activar a su vez a las células T $CD4^+$.

En este sentido, la primera parte de este trabajo se centra en el estudio de la regulación de la expresión génica de los genes *ANKRD55* y su vecino *IL6ST* en diversas células del sistema inmune, poniendo el foco principalmente en los linfocitos T $CD4^+$ y en las células dendríticas derivadas de monocitos (moDC) procedentes de pacientes con esclerosis múltiple y de controles sanos. En primer lugar, se examinó la expresión génica de *ANKRD55* y sus genes vecinos *IL6ST* (incluyendo tres de sus isoformas solubles: sgp130, sgp130-RAPS y sgp130E-10), *IL31RA*, *SLC38A9* y *DDX4* en cinco subpoblaciones diferentes de células mononucleares de sangre periférica: linfocitos T $CD4^+$, T $CD8^+$, linfocitos B, monocitos y células *natural killers* (NK). Tanto *ANKRD55* como *IL6ST* y sus isoformas solubles mostraron una mayor expresión en células T $CD4^+$, mientras que *IL31RA* únicamente se detectó en monocitos $CD14^+$ y *SLC38A9* tenía una expresión bastante similar en todas las subpoblaciones. El gen *DDX4* no se pudo detectar en ninguna de estos tipos celulares, por lo que se descartó en los posteriores análisis.

La expresión de *ANKRD55* y de tres de sus genes contiguos también se evaluó en células moDC inmaduras. En un análisis realizado a lo largo de los seis días que dura el proceso de diferenciación de monocitos a células dendríticas, se observó que la expresión de *ANKRD55* aumentaba a partir del segundo día, mientras que los niveles de *SLC38A9* alcanzaban su máximo en el primer día, los de *IL6ST* no variaban significativamente y los de *IL31RA* disminuían desde el

día uno hasta ser prácticamente indetectables. A continuación, se estudió el efecto de varios estímulos tolerogénicos sobre la expresión génica en moDC inmaduras y se encontró que el compuesto AM580, un agonista del receptor alfa del ácido retinoico (metabolito activo de la vitamina A), inducía fuertemente la expresión del gen *ANKRD55*. El efecto del AM580 resultó ser significativo en tres de las isoformas de este gen: *ANKRD55-201*, *ANKRD55-202* y *ANKRD55-204*. Por otro lado, se midió la expresión de los genes de interés en células moDC en distintos estados de maduración. Los resultados mostraron que el grado de maduración de las moDC, cuyo máximo nivel se observó al combinar interferón gamma junto con lipopolisacárido (IFN-γ/LPS), era inversamente proporcional al nivel de expresión de *ANKRD55*. De forma similar a los datos obtenidos por el análisis del ARNm del gen, la proteína ANKRD55 resultó ser prácticamente indetectable en monocitos mediante microscopía de fluorescencia, pero altamente visible en moDC inmaduras y menos abundante en moDC maduras. Además, se observó una colocalización de ANKRD55 con algunos marcadores de los llamados *nuclear speckles*, estructuras dinámicas del núcleo relacionadas con la expresión génica y el metabolismo del ARN.

Los siguientes análisis se centraron en la relación entre la expresión génica de *ANKRD55* e *IL6ST* y los genotipos de tres SNPs localizados en intrones de *ANKRD55*. Dos de ellos, rs6859219 y rs7731626, están asociados con el riesgo de sufrir esclerosis múltiple, mientras que el tercero, rs13186299, se relaciona con el porcentaje de monocitos en los glóbulos blancos y tiene un desequilibro de ligamiento parcial con los otros dos. En linfocitos T $CD4^+$ de donantes sanos se encontró que la expresión de *ANKRD55* era significativamente mayor en los homocigotos para los alelos de riesgo de los tres SNPs (CC en rs6859219, GG en rs7731626 y GG en rs13186299) en comparación con los homocigotos para los alelos protectores. La misma tendencia, aunque en menor medida, se observó en la expresión de *IL6ST* y sus isoformas solubles. Por el contrario, al realizar estos estudios en moDC inmaduras de donantes sanos, la mayor expresión de *ANKRD55* e *IL6ST* se encontró en los portadores de los alelos protectores de los tres SNPs (A para rs6859219, A para rs7731626 y C para rs13186299).

Posteriormente se analizaron moDC y linfocitos T $CD4^+$ obtenidos de pacientes de esclerosis múltiple para determinar también en ellos la posible asociación entre la expresión génica de *ANKRD55* e *IL6ST* y los genotipos de los tres SNPs de interés. En este caso, los resultados obtenidos revelaron que en los linfocitos T $CD4^+$ de los pacientes la única asociación significativa se daba entre *ANKRD55* y el SNP rs7731626. En cambio, en las células moDC inmaduras de estos mismos pacientes no se encontró ninguna asociación estadísticamente significativa en ninguno de los análisis realizados. También se observó que las moDC de pacientes homocigóticos para el alelo de riesgo del SNP rs13186299 (genotipo GG) expresaban

mayores niveles de *ANKRD55* e *IL6ST* en comparación con las de donantes sanos que tenían el mismo genotipo. Sin embargo, al analizar las moDC de portadores del alelo protector para este mismo SNP no se encontraron diferencias de expresión génica entre pacientes y controles. Estos resultados sugieren que puede existir un condicionamiento previo en los monocitos de los pacientes de esclerosis múltiple causado por este alelo de riesgo.

Por otro lado, se compararon los niveles de expresión génica en las moDC de los individuos sanos, considerándolos a todos como un único grupo control independientemente de su genotipo, frente a los distintos grupos de enfermos, clasificados según la forma de esclerosis múltiple. Los resultados mostraron que la expresión de *ANKRD55* era significativamente mayor en los pacientes con la forma PP que en los controles sanos. En cambio, la expresión de *IL6ST* alcanzaba niveles significativamente más elevados en los individuos sanos en comparación con los distintos grupos de pacientes.

Más adelante, el interés del estudio se dirigió hacia la proteína gp130, producto del gen *IL6ST* y que constituye la subunidad beta del receptor de la interluekina-6 (IL-6). La forma soluble de esta proteína, denominada sgp130, actúa como inhibidor de una de las rutas de la señalización de la IL-6 y se ha demostrado que está relacionada en diversas patologías. Por ello, tras los resultados obtenidos en los análisis de expresión de *IL6ST*, se estudió el posible efecto de cada uno de los tres SNPs de interés sobre la secreción de la proteína sgp130 en pacientes con esclerosis múltiple. Se observó que la concentración de sgp130 en suero, medida mediante la técnica de ELISA, era significativamente superior en los pacientes homocigóticos para el alelo protector (AA) del SNP rs7731626 en comparación con los heterocigóticos y con los homocigóticos para el alelo de riesgo. En los otros dos SNPs se observó efecto similar, aunque sin llegar a ser estadísticamente significativo.

La segunda parte de este trabajo pone el foco en la proteína ANKRD55, cuya función exacta aún no se ha demostrado completamente. La estructura de esta proteína contiene nueve repeticiones de la secuencia de ankirina. Estas repeticiones de ankirina conforman uno de los motivos proteicos más abundantes en la naturaleza y muchas de las proteínas que las contienen suelen mediar en las interacciones proteína-proteína, implicadas a su vez en procesos biológicos como la regulación de la transcripción o el control del ciclo celular. Una de las estrategias utilizadas para tratar de comprender las funciones de una proteína es identificar su interactoma proteico, definido como el conjunto de todas las proteínas que interaccionan con ella. Siguiendo este planteamiento, en el presente trabajo se caracterizó el interactoma proteico de ANKRD55 en cuatro tipos celulares. Para ello se usaron distintas líneas celulares humanas como modelo

de astrocitos (IM-huA), microglías (IMhu-M), neuroblastoma (SH-SY5Y) y monocitos (THP-1), todas ellas de interés para entender la patogénesis de la esclerosis múltiple. Los astrocitos y las microglías forman parte de la llamada neuroglia, la cual se encarga de dar soporte al sistema nervioso mediante la regulación de la homeostasis, la protección de las sinapsis y la regeneración de los nervios. En cambio, durante la patogénesis de la esclerosis múltiple, tanto los astrocitos como las microglías activadas contribuyen a la inflamación, la desmielinización, la neurodegeneración y el daño al tejido nervioso. Por otra parte, la línea celular SH-SY5Y de neuroblastoma utilizada en el estudio posee características similares a las neuronas, mientras que las células THP-1 sirvieron como modelo de monocitos.

Debido a los bajos niveles de expresión de la proteína ANKRD55 en estas cuatro líneas celulares, se optó por transfectarlas usando un ARNm con la secuencia del gen *ANKRD55* a fin de obtener cantidades suficientes de la proteína para realizar los estudios del interactoma. La identificación de las proteínas de cada interactoma se realizó mediante la técnica de espectrometría de masas de alta sensibilidad y los resultados se analizaron de manera bioinformática para detectar rutas y procesos celulares comunes. La mayoría de proteínas identificadas resultaron ser específicas de cada tipo celular. No obstante, se detectó un conjunto de proteínas compartido entre los interactomas de ANKRD55 de las cuatro líneas celulares analizadas. Dicho grupo constituía un "núcleo" común constituido por tres miembros de la misma familia de proteínas: 14-3-3η, 14-3-3β/α y 14-3-3θ. Las proteínas 14-3-3 se encuentran presentes en todas las células eucariotas y llevan a cabo funciones regulatorias de procesos biológicos fundamentales y también del sistema inmune. Estas proteínas abundan especialmente en el cerebro y su desregulación se ha asociado con enfermedades neurodegenerativas. Otros dos miembros de esta misma familia, 14-3-3γ y 14-3-3ε, fueron detectados en los interactomas de IMhu-A, IMhu-M y SH-SY5Y. Se observó que en estos tres tipos celulares ANKRD55 interaccionaba con la proteína BAG-2, una co-chaperona involucrada en la patogénesis de enfermedades neurodegenerativas. Entre el resto de proteínas compartidas por varios de los interactomas analizados se encontraron algunas que forman parte de los ribosomas y de las mitocondrias, otras implicadas en la síntesis de pirimidinas y varias proteínas de transporte intraflagelar.

Con el objetivo de hallar rutas moleculares y procesos biológicos en los que estuvieran implicados las proteínas resultantes de cada uno de los interactomas de ANKRD55, se realizó un análisis bioinformático usando el software "IPA" (Qiagen). Los resultados mostraron las líneas celulares IMhu-A, IMhu-M y SHSY5Y compartían varias de las rutas y procesos identificados con mayor significancia, entre las cuales estaban la regulación del punto de control de daños en el

ADN, así como diversas vías de señalización (HIPPO, ERK5, IGF-1 y PKA) implicadas en el control del ciclo y la proliferación celular, y en el desarrollo neurológico. En la línea THP-1, estas mismas rutas aparecían con un nivel de asociación mucho menor, lo cual implica que ANKRD55 podría estar ejerciendo una función diferente en estas células en comparación con los otros tres tipos celulares. En un estudio más detallado, se observó que la mayoría de las rutas con mayor significancia estadística contenían un alto porcentaje y variedad de las isoformas de 14-3-3, a excepción de la línea THP-1. Además, al comparar los datos de abundancia de las proteínas del interactoma de las microglías con los datos de transcriptómica obtenidos del RNA-seq en estas mismas células, se encontró una correlación positiva entre ambos conjuntos de datos, debida principalmente a las proteínas 14-3-3.

Para comprender el papel que jugaban el resto de proteínas diferentes a las 14-3-3, se realizó un análisis bioinformático utilizando la base de datos STRING y excluyendo del mismo a todas las isoformas de 14-3-3. En este análisis, el proceso de acetilación apareció como el término más asociado al combinar los datos de los cuatro interactomas de ANKRD55 estudiados. En este sentido, cabe mencionar que la desregulación de la acetilación está relacionada con el desarrollo de enfermedades neurodegenerativas. Entre el resto de los procesos biológicos que aparecieron compartidos por varias de las líneas celulares tras este análisis se encontraban algunos relacionados con el tráfico de exosomas extracelulares o con la unión al ARN. En el caso del interactoma de ANRKD55 procedente de las microglías, el término asociado con mayor significancia (excluyendo las proteínas 14-3-3) fue el denominado “partícula de transporte intraciliar b”. Dicho término, que no aparecía en el resto de líneas celulares analizadas, estaba definido por ocho proteínas de transporte intraflagelar (IFT), todas ellas detectadas en el interactoma de las microglías.

Las proteínas IFT son componentes necesarios para el correcto ensamblaje de cilios y flagelos, y se encargan del transporte bidireccional de moléculas a través de estas estructuras celulares. Estas proteínas constituyen el complejo IFT, cuya arquitectura consta de dos partes: IFT-A e IFT-B. La subunidad IFT-B se subdivide a su vez en una región periférica y otra nuclear. Las ocho proteínas IFT identificadas en el interactoma de las microglías pertenecen a la región nuclear de la subunidad IFT-B. Se sabe que varias proteínas pertenecientes a este subcomplejo nuclear, como IFT46, IFT52 e IFT56, son esenciales para el correcto desarrollo de muchos tejidos en vertebrados, incluido el tejido nervioso. Por tanto, cabe concluir que ANKRD55 podría desempeñar un papel relevante en este proceso, al unirse a proteínas que están directamente implicadas en el mismo.

Los cilios primarios actúan como transductores de señales externas y median la regulación de muchos procesos celulares y fisiológicos, incluyendo el correcto desarrollo tisular y la homeostasis. Se trata de estructuras que se encuentran en muchos tipos de células eucariotas, pero entre las cuales no se encuentran las microglías. Sin embargo, las proteínas IFT también se han hallado en células no ciliadas ejerciendo funciones importantes, como es el caso de la denominada sinapsis inmune, la cual se produce entre un linfocito T y una célula presentadora de antígenos. De forma análoga, existe también una sinapsis neuro-inmune, que se da entre neuronas y microglías en el sistema nervioso central, y que juega un papel clave en el inicio de enfermedades neurodegenerativas. Debido a su interacción con proteínas IFT en las microglías, ANKRD55 podría ejercer también funciones relacionadas con este tipo de comunicación celular.

Como conclusión, los resultados obtenidos en este trabajo ofrecen nuevas claves sobre la regulación y la función de ANKRD55. El aumento de expresión génica de *ANKRD55* en moDC inmaduras, en comparación con los monocitos, sugiere que este gen juega un papel relevante en el proceso de diferenciación de los monocitos precursores hasta las células dendríticas. Los cambios de expresión de *ANKRD55* en las moDC, producidos tanto por el ácido retinoico como por factores de maduración (aumentando y disminuyendo sus niveles, correspondientemente), refuerzan la idea de la influencia reguladora de *ANKRD55* sobre este tipo celular. Además, los resultados de los estudios de asociación en distintos tipos de células inmunes ponen de manifiesto la complejidad del efecto que ejercen sobre la expresión génica las variantes genómicas consideradas factores de riesgo para la esclerosis múltiple. Por ello es necesario seguir profundizando en este tipo de estudios, poniendo el foco sobre nuevas subpoblaciones de células inmunes para comprender mejor la patogénesis de la enfermedad. Por otro lado, los análisis del interactoma de proteínas en distintos tipos celulares revelan la diversidad de proteínas a las que se une ANKRD55 de manera directa o indirecta. La interacción de ANKRD55 con miembros de la familia de proteínas 14-3-3, observada anteriormente en otros tipos celulares, se confirma como una característica compartida también entre las líneas celulares de microglías, astrocitos, monocitos y neuroblastoma analizadas. En el caso de las microglías, cabe destacar además la unión de ANKRD55 a componentes de la maquinaria de transporte intraflagelar. Los datos obtenidos en estos interactomas sugieren una participación de ANKRD55 en procesos como el control del ciclo y la proliferación celular, el desarrollo del tejido nervioso y la sinapsis inmunológica, otorgándole así una especial importancia en el sistema inmune y nervioso.

1. Introduction

1. Introduction

1.1. Definition and epidemiology of multiple sclerosis

Definition and etiology of multiple sclerosis

Multiple sclerosis (MS) is an autoimmune[1], neuroinflammatory[2], neurodegenerative[3] and multifactorial[4] disease of the brain and spinal cord that causes chronic physical disability[5,6] and cognitive dysfunction[7]. In 1868, Jean Marie Charcot defined the concept of MS for the first time under the name *La Sclérose en Plaques*[8,9]. However, some earlier descriptions of symptoms indicative of MS can be found in the literature, such as the cases of Saint Ludwina of Schiedam in Holland (1380-1433)[10] and, later on, Augustus d'Esté (1794-1848), whose diary referred to a progressive disease causing physical disability[8,10,11]. Apart from Charcot, other authors, such as Marburg, Fromann and Babinski, provided a more detailed description of the pathology through the study of axonal damage, inflammation and clinical disability[12–16].

MS is caused by the destruction of myelin, a process named demyelination, in the central nervous system (CNS)[17]. Lesions caused by MS are present in both gray and white matter and result from the infiltration of immune cells through the blood-brain barrier (BBB), producing demyelination, inflammation, glial reaction and axon damage and disrupting neuronal transmission[18,19]. However, there is a high heterogeneity in lesions among MS patients[20].

Some proteins, such as myelin basic protein (MBP), myelin proteolipid protein (PLP) and myelin oligodendrocyte glycoprotein (MOG), among others, have been proposed as initial autoantigens of the disease[21,22]. These molecules have been proven to be recognized by $CD4^+$ T cells in patients with MS, but other investigations either were not able to reproduce these results or detected some of these autoantibodies in healthy people[23–25]. Recently, GDP-L-fucose synthase has been proposed as a $CD4^+$ T cell-specific autoantigen in some MS patients[26]. However, to date, none of these autoantigens have been well characterized as a single antigen in the initial autoimmunity of MS[23,27,28].

Epidemiology of MS

MS is considered a global disease, being the most common inflammatory disease of the CNS. Moreover, in many countries, it is the leading cause of non-traumatic neurological disability in young adults[29–31]. To date, the most complete information about MS epidemiology on a

global scale comes from the *Atlas of MS*, whose first edition was published in 2008. This project is a collaboration between the Multiple Sclerosis International Federation and the World Health Organization and contains data from 115 countries. The most recent version of this report is the 3rd edition, published in September 2020 (https://www.atlasofms.org)[31]. According to this source, the current estimation of people with MS globally is 2.8 million, which equates to 1 in 3,000 people and a prevalence (number of patients with the disease alive at a specific date) of 36 cases per 100,000 people. Specifically in Spain, there are 55,000 people currently living with MS, corresponding to a prevalence of 120 per 100,000 people (**Figure 1**). Apart from adults, there are at least 30,000 people under the age of 18 living with MS around the world[31]. Between 2013 and 2020, there has been an increase of the number of people with MS, from 2.3 million people to 2.8 million people, which is thought to be attributable to improvement of diagnosis, increase of registries and higher life expectancy of patients rather than to an increase of the risks of developing MS.

Sex ratio

Women are more often affected from MS than men, with a ratio between 2–3:1, reaching to 4:1 in some countries, such as Egypt, Iran and Sudan[31,32]. These differences in sex ratio are likely influenced by hormonal[33,34], genetic[35] and environmental response differences[36] between women and men. The sex ratio (female to male) has increased in recent decades[36,37], which can be explained by several factors, including an improvement in access to healthcare by women in some countries and therefore to correct diagnosis and/or changes in lifestyle, such as later pregnancy, use of oral contraception, obesity, smoking and increased stress[32,38]. On the other hand, a decrease in relapses of women with MS during pregnancy, especially during the third trimester, has been reported, probably due to pregnancy-related changes in the endocrine and immune systems [39,40].

Regarding age, there are no differences in the prevalence of MS between boys and girls during the first decade of life. However, during the adolescence, the prevalence in girls increases more than in boys, and this trend tends to continue in adulthood. From the seventh decade of life onwards, there is an attenuation of prevalence in men, but it still rises in women[41]. There is minimal evidence for the effect of sex on the clinical course of the disease, but, in general, women have an earlier onset of MS, slightly lower prevalence of primary progressive (PPMS), higher prevalence of relapsing-remitting (RRMS) and less progression of disability compared to men[36,42].

The geography of MS

MS is present in all regions of the world[31] but has an unequal distribution (**Figure 1**). This observation led researchers to perform studies that revealed a significant association between MS frequency and geographical latitude[43–45], showing differences in prevalence of up to nine times higher in high latitudes compared to equatorial regions[41]. However, there are some populations where this association does not align with these trends, such in Sardinia, located in Southern Europe, which has a high MS frequency; or the Inuit population, in the north of Canada, with a low frequency of the disease[30]. In general, in North America and western Europe, a positive correlation between prevalence and latitude has been observed, whereas in Australia, this gradient seems to be reversed[41]. Nevertheless, some authors question the idea of a latitudinal gradient in MS incidence, suggesting the existence of errors or bias in study methods or asserting that this gradient is disappearing[32,38,46,47]. It should also be noted that the relatively low frequency of MS in less-developed countries is likely related to underdiagnosis because of lack of access to healthcare systems[31,41].

One of the factors proposed to explain the irregular geographic distribution of MS is genetics. It has been demonstrated that specific ethnic groups, including Caucasian people from Scandinavia and Scotland, show a high susceptibility to the disease, while others, such as Mongolian, Japanese, Chinese, Native American, Eskimo, African Black, Aboriginal, Norwegian, Sami and Gypsy groups are less affected. However, studies conducted on migrant populations showed changes in MS frequency in a short period of time (one generation) which could not be explained by genetic factors alone[30]. The other main factor is environment. Studies indicate an association between lower disability of MS and better diet[48], which can in turn be related to cultural and geographic distribution and point as well to an association between lower MS risk and higher sun exposure[49], which is a likely explanation for the latitudinal distribution of MS.

Migration may also modulate the risk of developing MS[32]. The general consensus is that migrants who move from a high- to a low-MS risk region during childhood reduce their risk of MS, while migration from a low- to a high-MS risk area has not been associated with variation in MS risk[50]. This suggests that environmental risk factors may affect disease risk mainly in the first two decades of life[30,50,51], although some further studies suggest that risk factors may also operate over many more years[52].

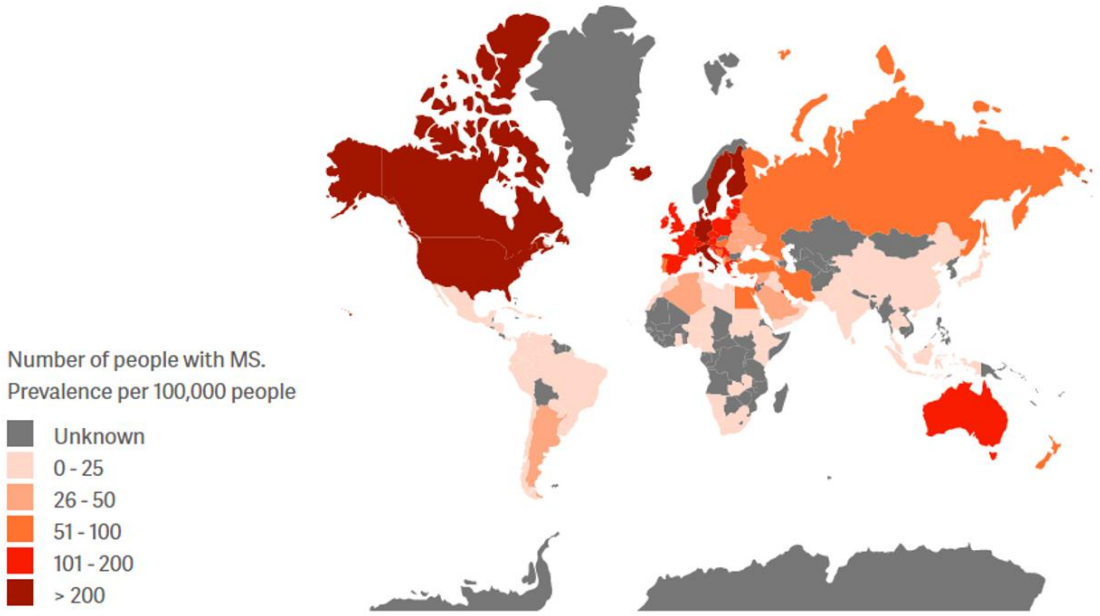


FIGURE 1. World map of MS prevalence (per 100,000 people).
Source: The Multiple Sclerosis International Federation (MSIF), 2020[31].

1.2. Clinical features of MS

1.2.1. Diagnosis of MS

The diagnosis of MS is based on neurological symptoms and laboratory tests and defined as focal demyelination affecting more than one part of the CNS on more than one occasion[53,54]. The most widely used diagnostic tool currently is the McDonald Criteria (latest revision in 2017)[55,56], which combines clinical, imaging and laboratory evidence and has improved with the advance of technology, making the diagnosis process simpler and easier. The symptoms are widely variable and heterogeneous but, in most patients, include alterations in the motor, sensory, visual and autonomic systems in addition to cognitive changes, mood perturbation, and fatigue. As the disease progresses, it leads to severe disability in the patient[19,53,57], which can be quantified with the Expanded Disability Status Scale (EDSS), developed by the neurologist John Kurtzke in 1983[58].

The main proof for the diagnosis of MS is the presence of focal plaques of demyelination in the brain (present in more than 95% of patients), located in both white and grey matter at all stages of the disease[59]. Magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) analysis can be used to detect these sites of inflammatory demyelination, named as active

lesions. In addition, most MS patients show oligoclonal bands in the CSF, indicating intrathecal immunoglobulin synthesis[53,54,60]. These oligoclonal bands are produced by clonally expanded infiltrating B cells and represent the clearest evidence of autoantibody production. More than 90% of patients show oligoclonal bands of immunoglobulin G (IgG), whereas those of immunoglobulin M (IgM) are only found in 30–40% of people with the disease. The oligoclonal bands are not absolutely specific for the disease but have been proposed as markers for prognosis in some groups of MS patients[23,25,61,62].

The average age of MS diagnosis is 32 years[31], and up to 5% of MS patients are children, mainly girls. Children with MS usually present encephalopathy and tend to develop RRMS[53,54]. On the other hand, there are cases of MS diagnosis after age 50, which usually occur in men and normally have a progressive onset[54]. The course of MS is difficult to predict, but several factors, such as older age, male sex, frequent relapses in the first two years, higher disability at baseline, progressive course from the onset and greater brain atrophy have been associated with a more severe clinical course[60,63]. Patients with advanced neurological disability have an increased risk of skin, chest and bladder infections. Life expectancy of MS patients is around 30 years from disease onset[53].

1.2.2. Subtypes of MS

The first classification of clinical subtypes of MS was established in 1996 by the United States National Multiple Sclerosis Society (NMSS) Advisory Committee on Clinical Trials in Multiple Sclerosis. It was based on subjective descriptions rather than clear pathophysiological evidence. Therefore, in 2013, this classification was reviewed and improved to include the following categories[63–65] (**Figure 2**):

Clinically Isolated Syndrome (CIS)

CIS was not initially recognized as an MS clinical category but now is considered the first clinical manifestation of inflammatory demyelinating diseases, including MS[64]. It has a duration of at least 24 hours, is usually monofocal and affects the optic nerve, brain stem or spinal cord, and it represents the onset of MS in 85% of patients[65,66]. The most common symptoms of CIS are unilateral optic neuritis, which can be detected by ocular coherence tomography[67], sensory disturbances, brainstem syndrome, limb weakness, fatigue, bowel and bladder

affectation[54,57,60]. In patients who have experienced a CIS event, the risk of developing MS within 10 years is 60–80% if they have brain lesions detectable by MRI, and it reduces to 20% if the CIS event is not accompanied by detected brain lesions[68]. According to the 2017 revision of McDonald Criteria, CIS together with MRI evidence of lesions in the CNS and the presence of specific oligoclonal bands in the CSF allows for a diagnosis of MS[56].

Relapsing-remitting MS (RRMS)

Relapsing-remitting MS is the most common form of MS, affecting around 85% of patients. It is characterized by repetitive cycles starting with an acute episode of neurological disfunction (relapse or exacerbation) lasting from days to weeks followed by a recovery (remission) period of weeks or months[19,53,54,69]. Relapses are related to foci of demyelination and inflammation in the CNS, specifically in the white matter, causing variable symptoms, such as fatigue or problems with vision and balance[19,65]. A patient is diagnosed with RRMS when they suffer from at least two relapses[63]. Although these relapses vary in intensity and frequency between patients, they usually decrease with age and disease progression, not exceeding one or two episodes per year[65,70]. RRMS can last for years or decades and is up to three times more common in women than men. The average age at onset is around 30 years[54,70].

Secondary progressive MS (SPMS)

Approximately 25–40% of patients diagnosed with RRMS progress to SPMS within 10 to 23 years[19,69–71]. Some factors, including higher age at RRMS onset and male sex, have been associated with earlier progression to SPMS, but no molecular markers have been identified as predictors of progression[65]. Therefore, this transition is mainly determined by retrospective observation of medical history rather than specific clinical evidence[64]. SPMS is characterized by a gradual and irreversible decline of neurological function with loss of brain volume, increase of axonal damage and absence of inflammatory lesions, leading to an increase of disability[19,54,72]. There are no relapses associated with degradation of the CNS in SPMS[57,73].

Primary progressive MS (PPMS)

10–15% of MS patients directly enter a slowly progressive decline phase, termed primary progressive MS[69]. This neurodegenerative stage causes disability, lasts more than six months and has few or no relapses[54,57,70]. Some of the typical symptoms of PPMS include upper motor neuron syndrome of the legs; gradually loss of vision; cognitive impairment; brainstem,

bladder and bowel affection; and sexual dysfunction[60]. Unlike RRMS, PPMS affects men and women equally, and it is usually diagnosed at older ages, with a mean age of 40 years[54,70]. Prior to the 2012 revision of MS clinical phenotypes, a category named progressive-relapsing multiple sclerosis (PRMS) was included in the clinical course descriptions. PRMS was defined as PPMS with acute attacks[57]. However, this category has been eliminated, and these patients are classified nowadays as PPMS with active disease[64].

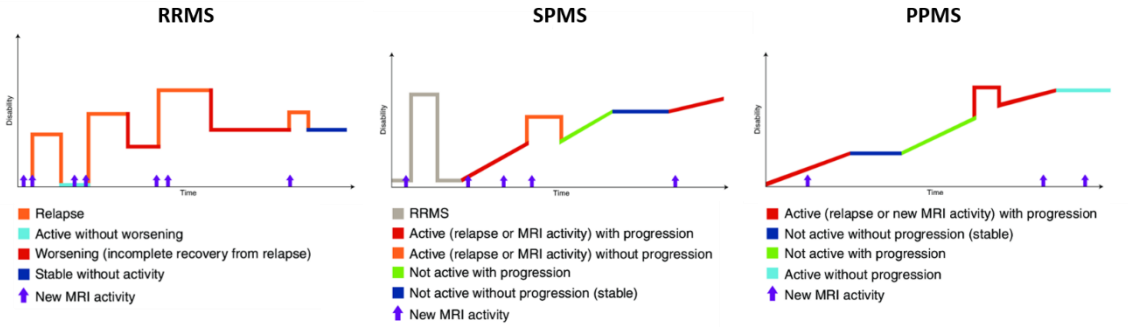


FIGURE 2. Clinical subtypes of MS.
Schematic representation of the main types of MS and their most characteristic features over time. RRMS: relapsing-remitting MS. SPMS: secondary progressive MS. PPMS: primary progressive MS. Adapted from Klineova *et al.*[65] and the National Multiple Sclerosis Society (www.nationalmssociety.org).

1.2.3 Treatments of MS

Although there is no cure for MS to date, many therapies have been developed in the last decades with the objective to reduce symptoms and slow the progression of the disease. Among these treatments are disease modifying therapies (DMTs), hematopoietic stem cell transplantation (HSCT), probiotic treatments and other immunotherapies.

Disease modifying therapies (DMTs)

DMTs include drugs developed for MS that target neuroinflammation by suppressing the immune system. These treatments act on different pathways such as protecting myelin of axons from degeneration or reducing the entrance of immune cells into the CNS[74]. All of these have the capacity to reduce the number of attacks and are effective in patients with RRMS and CIS[63,75]. On the other hand, there is only one drug (ocrelizumab) which can effectively reduce disease progression in patients with PPMS[63]. Approved medications for MS can be classified as following:

DMTs based on interferon beta (IFN-β)

Interferons were the first DMTs approved, in the early 1990s[76]. IFN-β regulates the immune system and has anti-inflammatory properties. It has been shown to slow disability progression and be more effective in RRMS than in SPMS[57]. DMTs based on IFN-β include the commercial medications Avonex®, Betaseron®, Extavia®, Rebif®, and Plegridy®[18,75,77].

DMTs based on glatiramer acetate

Glatiramer acetate induces proliferation of anti-inflammatory lymphocytes and regulatory cytokines[77]. Commercialized under the name of Copaxone®, or Glatopa® as generic form, this medication is recommended for RRMS patients[57] and can be also used in pregnant people and children[78].

DMTs based on monoclonal antibodies

Humanized monoclonal antibodies are commonly used in the treatment of MS due to their high specificity and few side effects. They target and neutralize key molecules involved in the immune system which play a role in the inflammatory stage of the disease. Among these antibodies are natalizumab (anti-α4β1 integrin, commercialized as Tysabri®) and alemtuzumab (anti-CD52, commercialized as Lemtrada®), both used for RRMS treatment, and ocrelizumab (anti-CD20, commercialized as Ocrevus®), recommended for RRMS and PPMS patients[63,77,79]. There are also other antibodies that have not been specifically approved for MS treatment but are used to treat the disease given their high level of efficacy, such as rituximab (anti-CD20, commercialized as Rituxan®) and ofatumumab (anti-CD20), both used to treat RRMS[77,79].

Other DMTs

Other approved DMTs with different mechanisms of action include dimethyl fumarate (commercialized as Tecfidera®), which reduces inflammatory cytokines; fingolimod (commercialized as Gilenya®), a sphingosine 1-phosphate inhibitor which was the first DMT licensed for children[18,57,75]; teriflunomide (commercialized as Aubagio®), a pyrimidine synthesis inhibitor which reduces the proliferation of T and B immune cells which are active in MS[57]; cladribine (commercialized as Mavenclad®), which temporarily reduces the number of T and B lymphocytes[63]; and dalfampridine (commercialized as Ampyra®), which blocks potassium channels on the surface of nerve fibers and has shown to improve mobility in MS patients by improving the transmission of nerve impulses[18,75]. Although all of these drugs are mainly recommended for RRMS, there are other medications specifically for SPMS, including the

sphingosine 1-phosphate receptor modulators siponimod (commercialized as Mayzent®) and ozanimod (commercialized as Zeposia®)[63,77] and the chemotherapeutic agent mitoxantrone (commercialized as Novantrone®)[18,57,75].

The majority of the medications previously mentioned target and may suppress the immune system, which increases the risk of cancer or infections[75], among other undesirable side effects[57,63]. Nevertheless, there are also new therapeutic agents, still under study, which are focused on neuroprotection, including phenytoin[80] and ibudilast[81]. Other agents, such as clemastine, are focused on remyelination and are intended for RRMS patients[82].

Hematopoietic stem cell transplantation (HSCT)

Autologous hematopoietic stem cell transplantation (HSCT), has proven increasingly successful in treating patients with active RRMS, who do not have significant disability and who do not respond to conventional drugs[63]. This procedure is based initially on high immunosuppression, causing the destruction of autoreactive cells, followed by stem cell transplantation from the person's own bone marrow, which leads to the renewal of immune cells and restoration of efficient immune tolerance that can last for years[83,84]. HSCT consists of multiple steps, starting with the mobilization of hematopoietic stem cells from the bone marrow of the patient, followed by collection of these cells, ablation of the immune system using chemotherapy, and finally reinfusion of the hematopoietic stem cells and post-transplant care.

Due to the powerful anti-inflammatory effect of HSCT, this therapy has resulted in better outcomes in early phases of the disease (RRMS) and aggressive MS cases than in progressive stages[83–85]. Despite the great advantage of using the patient's own bone marrow for immune system reconstitution, HSCT presents some side effects, including an increased long-term risk of developing infections and autoimmune conditions[86,87]. From 1997 to 2017, more than 1,000 patients have been treated with HSCT. The main problem associated with this therapy was the risk of mortality, which was 3.6% before 2005[83]. However, a more careful selection of patients has since reduced this percentage to 0.3%. Besides the type of MS, there are other important factors to consider a patient as a candidate for HSCT, including age, risk tolerance, duration of disease, pregnancy intention and comorbidities, among others. According to some studies, HSCT has been shown to be more efficient in patients with young age, short time from the onset of the disease (less than 5–10 years) and no severe disability[83,85,88].

Probiotic treatments

Many recent studies have demonstrated a link between the gut microbiota and the regulation of the immune system in several diseases, including MS[89,90]. Specifically, species belonging to *Clostridia* clusters and the phylum *Bacteroidetes* are less abundant in MS patients[91,92], while *Akkermansia* and *Acinetobacter* genera are more abundant[93] compared to healthy controls. In this respect, treatments based on probiotics to modulate gut microbiota have been proposed to complement the use of approved drugs. Preclinical studies have shown promising results with benefits in the motor and mental behaviors of patients, delay of progression and decrease of disease severity by modulating inflammatory pathways[94–96]. In addition, in a Theiler's murine encephalomyelitis virus (TMEV) mouse model of progressive MS, the administration of an oral antibiotic cocktail revealed a reduction in disability and protection from axon damage[97]. The most commonly used probiotics belong to the *Lactobacillus*, *Lactococcus*, *Bifidobacterium* and *Clostridium* genera, which promote an anti-inflammatory response[93,98]. A related approach is fecal microbiota transplantation, in which the entire network of microbial species is transplanted from a healthy donor to a patient through fecal matter to restore the microbiota[99].

Other immunotherapies

The majority of the immunotherapies for MS modulate the immune system in a non-antigen-specific manner. Nonetheless, new immunotherapeutic strategies such as DNA vaccines, nanoparticles and altered peptide ligands focus selectively on the autoimmune response[100–102]. For example, T-cell vaccination consists of the *in vitro* attenuation of autoimmune T lymphocytes derived from an MS patient, which are then inoculated in the donor's immune system to suppress the pathogenic potential of these cells[103]. Other strategies based on immunotherapy include administration of soluble antigens or antigen-coupled cells[28]. However, further investigation is needed to improve these techniques.

Apart from all these therapies, evidence suggests that physical exercise is also beneficial in RRMS and PPMS patients with functional disability[104], in addition to complementary care for MS patients, including the work of neuropsychologists, social workers and occupational therapists, among others[53].

1.3. Pathogenesis of MS

MS is widely considered an autoimmune disease of the CNS, although the initial antigen has not been identified yet. The pathological process is characterized by the breakdown of the BBB, migration and infiltration of immune cells, secretion of pro-inflammatory cytokines, formation of sclerotic plaques (lesions) in the brain and spinal cord, demyelination and axonal loss[30,75,105,106].

Two different hypotheses have been proposed to explain the onset of the disease. The "outside-in" hypothesis, most widely accepted, considers the first event to be the activation of autoreactive $CD4^+$ T lymphocytes in the periphery, which cross the BBB and enter the CNS, where they promote inflammation and ultimately neurodegeneration. On the other hand, the "inside-out" hypothesis suggests a primary pathogenesis inside the CNS, possibly initiated by inflammation or myelin injury, leading to an autoimmune response[19,63,107]. Several experimental murine models support each of these hypotheses. The classical and most used experimental autoimmune encephalomyelitis (EAE) mouse model is generated by injecting myelin antigens together with complete Freund's adjuvant into the mice, inducing an inflammatory demyelinating disease similar to RRMS and PPMS. Using a slightly different approach, the Theiler's Murine Encephalomyelitis Virus-Induced Demyelinating Disease (TMEV-IDD) is another relevant model for the study of MS. Both EAE and TMEV-IDD show robust evidence supporting the "outside-in" hypothesis. Others models such as the Diphtheria Toxin A Chain (DTA) model and the Cuprizone Autoimmune Encephalitis (CAE) model align with the "inside-out" hypothesis[107].

Whether by proteins from autoreactive antigens ("outside-in" or autoimmune hypothesis) or proteins from the CNS after degeneration ("inside-out" or degeneration hypothesis), it is clear that the beginning of the disease occurs in the lymphoid tissues, such as lymph nodes and the spleen. In these tissues, antigens are presented to $CD4^+$ and $CD8^+$ T cells by antigen-presenting cells (APCs), such as dendritic cells, through the major histocompatibility complex (MHC) I and II, respectively. In addition, B cells are capable of capturing antigens and acting as APCs. Consequently, there is an activation and clonal expansion of T cells and B cells, which, together with other immune cells, migrate across the BBB and into the CNS. Once in the CNS, these infiltrating immune cells promote tissue injury. B cells mature into plasma cells, which release a high number of IgG antibodies against myelin and glial cells. $CD8^+$ T cells encounter their specific peptide ligands in glial cells or neurons and cause direct damage. $CD4^+$ T cells interact with the resident microglia of the CNS, which secrete cytokines, thereby

attracting macrophages, promoting inflammation and leading to the attack of myelin and neural cells[19,108] (**Figure 3**).

The role of the immune system in the development of MS is therefore irrefutable. The immune system can be divided into the innate and adaptive (acquired) systems. The innate immune system represents the first line of defense, acts rapidly, and consists of macrophages, dendritic cells, complement factors and Toll-like receptors, among others. In contrast, the adaptive immune system is more specific due to its highly diverse antigen receptors, can establish immunologic memory and involves T and B cells[109]. In MS, demyelination and tissue damage is induced by macrophages and microglia of the innate immune system together with specific T and B cells of the adaptive immune system[75]. T cells play an important role in relapsing MS, whereas microglia and B cells are involved in the progressive phase, when the immune response is confined to the CNS[63]. Although autoreactive lymphocytes are detectable in both MS patients and healthy donors, activated $CD4^+$ T cells are still considered relevant at the start of the disease, as they can be found in the CNS in the early stages of MS[19,53]. More specifically, two $CD4^+$ T cells subsets, Th1 and Th17, are responsible for cytokine secretion and activation of the inflammatory response that ultimately leads to axonal damage and demyelination[75]. Other T cells, such as $CD8^+$ T lymphocytes and regulatory T cells (Tregs), also actively participate in the ongoing disease[110,111]. $CD8^+$ T lymphocytes are found in higher numbers than $CD4^+$ T cells in white matter and grey matter cortical demyelinating lesions. On the other hand, Tregs present functional deficits in MS patients. In addition to T cells, clonally expanded B cells are found in the CSF, meninges and brain parenchyma of MS patients, where they produce antibodies that can be used as a diagnostic tool[19,111]. Thus, the disease is the result of an interaction between the immune system, glia (oligodendrocytes, microglia and astrocytes) and neurons and includes the processes of inflammation, demyelination and neurodegeneration. These events are present in the three forms of MS (RRMS, SPMS and PPMS) but vary quantitatively and qualitatively over time and among patients[18,59,112].

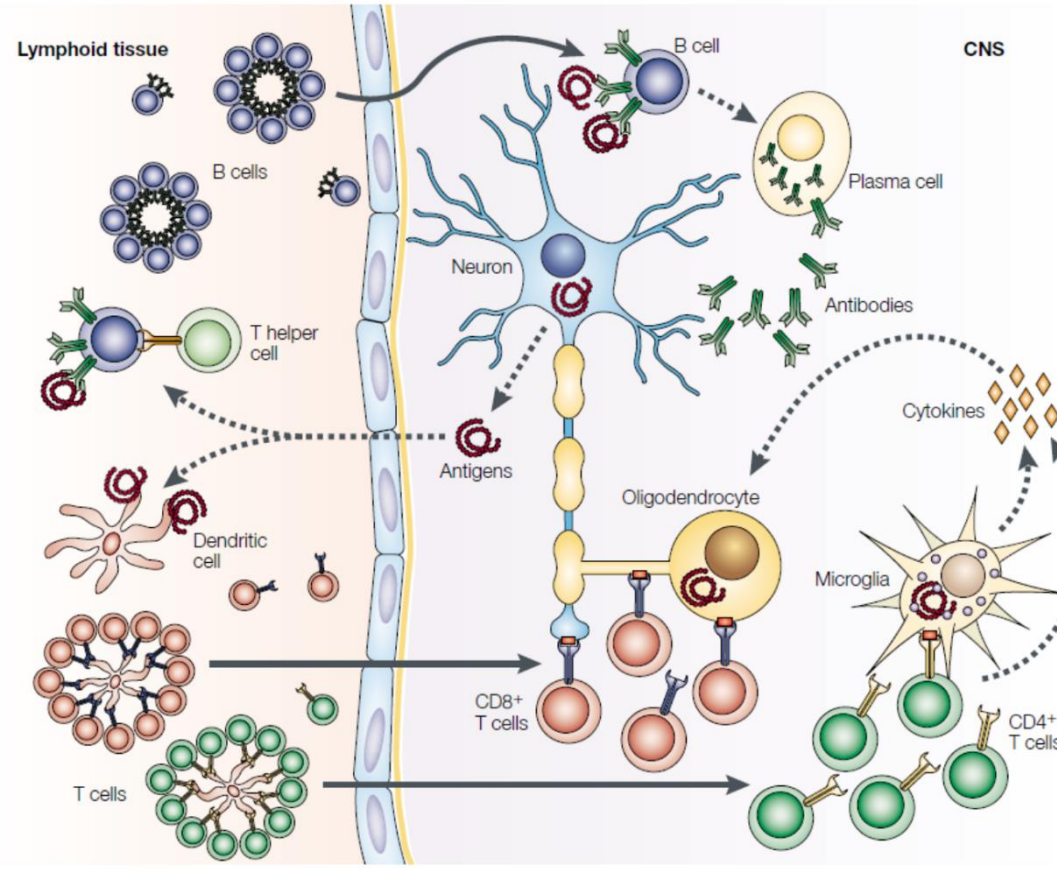


FIGURE 3. Immune response in MS.
Antigens released from neurons in the CNS are recognized by B cells and T cells in lymphoid tissue. Dendritic cells also stimulate T-cell responses. After clonal expansion, T cells and B cells infiltrate the CNS. B cells then mature to plasma cells and secrete large amounts of antibodies. Clonally expanded $CD8^+$ T cells induce direct damage to neurons and oligodendrocytes, whereas $CD4^+$ T cells encounter antigens presented by microglia, leading to production of inflammatory cytokines. These molecules, together with other immune mediators, cause the destruction of the myelin sheath. Source: Hemmer *et al*.[108]

Neuroinflammation

Inflammation is present at all stages of MS and is mediated by T and B lymphocytes, macrophages and activated microglia. Inflammatory lesions result from the infiltration of lymphocytes into the CNS, at which point the lymphocytes proliferate and secrete pro-inflammatory cytokines[59,63]. Two types of inflammation in MS patients have been proposed. The first occurs in acute and relapsing MS, when T and B lymphocytes cross and damage the BBB and invade the white matter; this leads to the classical active plaques. The other type of inflammation consists of a slow accumulation of T and B cells in the CNS without loss of BBB integrity. This form of inflammation is associated with the formation of demyelinated lesions in the cortex and expansion of pre-existing lesions and is found in progressive stages of the disease, when inflammatory activity diminishes in general[113]. Among T lymphocytes, $CD4^+$ T cells are implicated in the pathogenesis of the MS and contribute to the inflammatory process through

secretion of cytokines such as interferon gamma (IFN-γ) and interleukin-17 (IL-17), while $CD8^+$ T cells play a main role in active lesions. Reactive oxygen species (ROS) and reactive nitrogen species (RNS), produced mainly by macrophages and microglia, also contribute to the inflammatory process[60,69,111].

Demyelination

It is proposed that the inflammatory process leads to the development of demyelinating lesions, known as plaques, which are considered one of the hallmarks of MS[105]. The process of demyelination consists of loss of myelin. Under healthy conditions, myelin surrounds and protects axons, with the relative preservation of axons, and the formation of astrocytic scars[60] (**Figure 4**). Some of the mechanisms involved in demyelination include T cells, specific antibodies and death of primary oligodendrocytes[20]. Focal demyelination plaques are mostly observed in the white matter of the brain and spinal cord but can also be present in grey matter[19,59]. Active plaques from the same patient usually show homogeneous patterns of demyelination, but they can be heterogenous amongst patients[60]. RRMS patients tend to have more new brain and spinal cord lesions, while in PPMS and SPMS patients, a slow expansion of pre-existing lesions is observed[59].

The main and direct consequence of demyelinated axons is the conduction block, in which the nerve impulse cannot be properly propagated, thus causing the first neurological symptoms of the disease[105]. Due to demyelination, the energy demand of nerve conduction is increased, causing an ionic imbalance that ultimately produces more damage in the axon[70]. Despite the damage suffered, most demyelinated axons initially survive. Moreover, in around 20% of patients, plaques are eventually remyelinated, promoting restoration of nerve conduction. The process of remyelination is observed in a more active manner in acute plaques and early MS and also occurs in the primary and secondary progressive phases of the disease, although to a lower extent[53,60]. However, after remyelination, the conduction is more continuous than saltatory, and the new myelinated internodes are shorter and thinner than they were prior to disease onset[63,70]. Moreover, after repeated episodes of disease activity, the repair mechanisms of demyelination progressively fail in MS patients[59].

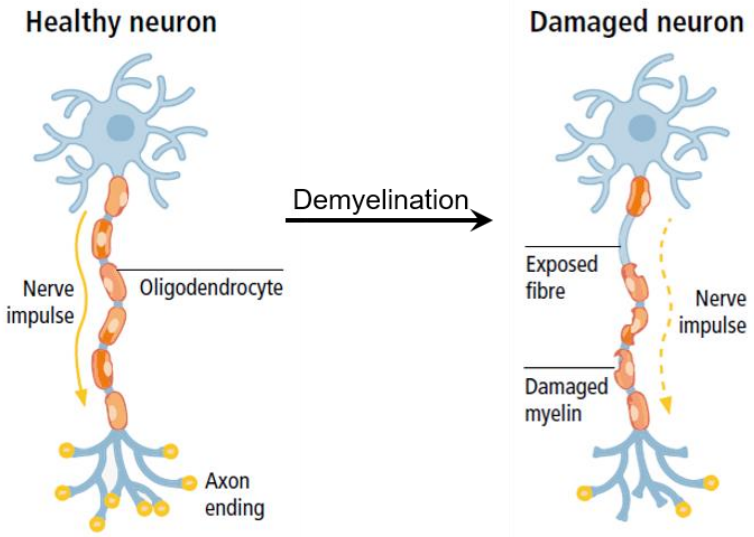


FIGURE 4. Demyelination process in neurons.
A typical feature of MS is the demyelination process, which consists of the loss of the myelin sheath that protects the axons of neurons, causing disturbances in the propagation of the nerve impulse.

Neurodegeneration

Inflammation and chronic demyelination slowly lead to neuronal and axonal degeneration. Inflammatory substances such as cytokines, proteolytic enzymes and oxidative products secreted by activated microglia and infiltrated macrophages injure vulnerable demyelinated axons[69,70]. In addition, direct attack by autoreactive antibodies or cytotoxic T cells, as well as loss of myelin trophic support from oligodendrocytes, may also contribute to axonal loss in MS[59,114].

Evidence suggests that mitochondrial dysfunction significantly contributes to tissue injury in MS. Damage to mitochondria produces an increase in ROS, and this oxidative damage contributes to the disturbance of ion channels for sodium, potassium and calcium, leading to a neuronal energy deficit. This energy deficit in neurons in turn produces more metabolic stress and further energy deficiency that, together with demyelination, results in a loss of neuronal connectivity, apoptosis of oligodendrocytes and degeneration of axons[59,69,105]. Glutamate, produced by macrophages and glial cells, is also important in the process of neuronal death: excess glutamate overstimulates glutaminergic receptors, resulting in excitotoxic effects[70,105].

Axonal or neuronal loss is considered to be the primary cause of irreversible neurological disability in MS patients and produces brain atrophy and cortical thinning as the diseases progresses. A reduction in brain volume at rates of 0.5–1.5% per year is observed in MS patients, and the pattern of atrophy is different from that observed in normal aging. Unlike the

inflammatory process, the brain atrophy continues in the later stages of the disease[63,70]. New techniques, including MRI and use of markers of axonal damage, have demonstrated that neurodegeneration in MS starts at disease onset [69,114]. Axonal injury is more acute in new inflammatory active lesions but also occurs in chronic lesions of slow expansion in the progressive stage of the disease[59,105]. The transition from RRMS to SPMS occurs when the CNS exhausts its capacity to compensate for further neuronal loss[70,106]. In SPMS, sites of demyelination coexist with axonal and neuronal degeneration[53].

1.3.1. Glial cells in MS

Microglia, astrocytes and oligodendrocytes are collectively the most abundant cells in the nervous system and compose the glial cells, neuroglia, or simply glia. The term 'neuroglia' was first introduced by Rudolph Virchow in the mid-19th century and described as a connective tissue that held nervous elements together. Other glial functions include myelination, modulation of homeostasis, neuronal protection, synapsis, nerve signal propagation and repair of nervous tissue[115].

Microglia

Microglia are the resident innate immune cells of the CNS, constituting around 10% of the total glial cell population in the brain and spinal cord. They have a myeloid origin, are capable of self-renewal and are mainly involved in maintenance and defense of the CNS[115]. They contribute to neurodegeneration but also promote remyelination. In MS, microglia are activated and present in all stages of lesion formation. Acute active lesions show more activated microglia in white matter compared to grey matter[18]. Activated microglia have been traditionally classified as 'classical activated microglia' (M1), which secrete proinflammatory cytokines, producing neural damage, and 'alternatively activated microglia' (M2), which release anti-inflammatory cytokines and regulate the immune response. However, this classification is currently under debate due to new evidence of further subgroups and plasticity of microglia[116,117].

The dual role of microglia involves, on the one hand, myelin injury and neuronal death through secretion of proteinases, nitric oxide and complement proteins, and on the other hand, protection of the CNS due to phagocytic function and tissue repair. In addition, activated microglia can also function as APCs, expressing antigens with MHC class I and II proteins and

initiating the immune response[118]. Microglia also play an important role in remyelination and phagocytosis of myelin debris, an essential step for later remyelination by oligodendrocytes. They also secrete neurotrophic molecules such as insulin-like growth factor 1 (IGF-1) and fibroblast growth factor 2 (FGF-2)[118].

Astrocytes

Astrocytes are considered to be the most abundant glial cell type in the CNS. Together with oligodendrocytes, they originate from a common lineage of neural progenitor cells. Astrocytes are defined by a star-shaped morphology and perform diverse functions in the CNS, including structural support of BBB, homeostasis maintenance, metabolic regulation, phagocytosis and synapse maintenance. One single astrocyte is capable of supporting up to hundreds of dendrites and protecting thousands of synapses[115,118]. Ramón y Cajal classified astrocytes over 100 years ago according to morphological differences and neuroanatomical locations into protoplasmic astrocytes, located in the grey matter, and fibrous astrocytes, located in the white matter. Although this classification is still considered valid, new research has shown that astrocytic heterogeneity is far more complex[119]. Several factors such as pro-inflammatory cytokines, oxidative stress, and molecules from pathogens promote astrocyte reactivity in diseases. In MS, the proliferation of astrocytes, named astrogliosis, produces neurotoxic molecules and enhances inflammation which causes tissue damage. After the demyelination process, astrocytes form glial scars at the sites of injury that isolate the inflamed area and provide structural support, but these cells also actively contribute to the development of MS lesions[120].

Oligodendrocytes

Derived from oligodendrocyte progenitor cells, oligodendrocytes are generated continuously in the healthy adult brain[121]. Oligodendrocytes were characterized by del Río-Hortega in 1928. They provide metabolic support to axons and contribute to neuroplasticity, but their main function is to myelinate neurons in the CNS[115]. Myelin is composed mainly of lipids as well as proteins such as PLP and MBP. Concentric layers of myelin are arranged to form the myelin sheath, which protects axons and ensures correct saltatory nerve conduction[118]. MS lesions are related to dysfunctional oligodendrocytes. Over the course of the disease, as a consequence of damage in the nervous system, oligodendrocyte progenitor cells gradually fail to respond to myelin injury, impairing the process of remyelination[18].

1.3.2. Immune cells in MS

Various immune cell types participate in the pathogenesis of MS, but this section will focus mainly in $CD4^+$ T lymphocytes, monocytes and dendritic cells, in concordance with the results shown in this thesis.

$CD4^+$ T cells

$CD4^+$ T cells, also named T helper (Th) cells, are part of the adaptive immune system. They recognize peptides presented on the surface of APCs by MHC class II. The interaction between T-cell receptors and MHC class II, together with costimulatory signals, induces proliferation and differentiation of $CD4^+$ T cells that promote effector functions in other immune cells, such as B cells and NK cells[122].

A body of evidence supports the assumption that autoreactive $CD4^+$ effector T cells are central drivers of the pathogenesis of MS. First, activated $CD4^+$ T cells are present in CNS lesions and in the CSF of patients. Furthermore, qualitative differences in the activation state and the cytokine profile of myelin-specific $CD4^+$ T cells have been found between MS patients and healthy individuals. In addition, the strongest genetic risk factor for MS is located in an MHC class II region. Upon activation, $CD4^+$ T cells invade the CNS, where they recognize antigens and initiate the inflammatory response, ultimately leading to tissue damage and neurodegeneration[25,28,106].

During the processes of activation, maturation and migration, $CD4^+$ T cells differentiate into various subtypes. These subtypes are classified depending on the cytokine secretion profile and effector functions and include Th1, Th2, Th9, Th17 and Th22, among others[122]. In general, Th1, Th9 and Th17 cells are involved in neuroinflammation and demyelination and therefore are considered key contributors to the pathogenesis and progression of MS[123–125]. On the other hand, Th2 and Treg cells play a primarily protective role in the disease due to their anti-inflammatory and immune response-modulating properties[126,127]. Other subtypes, such as Th22, may have both anti- and pro-inflammatory effects[128]. Thus, the whole picture is complex, and the specific dominant Th immune response in the MS is determined by the particular microenvironment of cytokines as well as the patient's genetic factors[127].

Monocytes

Monocytes are circulating blood non-dividing cells derived from a myeloid precursor in the bone marrow. They represent around 10% of peripheral leukocytes in humans and are considered part of the "mononuclear phagocyte system" together with macrophages and dendritic cells. Under inflammatory conditions, monocytes are rapidly mobilized to reach inflamed tissues, where they support the immune response[129,130]. These cells are precursors of macrophages and dendritic cells, but they also show specific functions: beside phagocytosis, monocytes can produce cytokines and present antigens, allowing them to participate in immune defense, homeostasis, inflammation and tissue repair[131,132].

Based on the expression of CD14 and CD16 surface markers, three different categories of human monocytes have been determined: classical ($CD14^{++}$ $CD16^{-}$), intermediate ($CD14^{++}$ $CD16^{+}$) and non-classical ($CD14^{+}$ $CD16^{++}$)[133] (**Figure 5**). As a clarification, here the $^{+}$ denotes an expression level of around 10-fold above the isotype control, while $^{++}$ is about 100-fold above the isotype control. It is widely accepted is that the three monocyte subpopulations are developmentally related, starting as classical monocytes and differentiating into intermediate and sequentially to non-classical monocytes in the peripheral blood[134]. Classical monocytes represent about 90% of blood monocytes and have a half-life of one to two days[135]. They show diverse differentiation potential, have the ability to migrate into tissues and are specialized in phagocytosis[132,136]. These monocytes secrete large amounts of cytokines and have a pro-inflammatory profile[131,137]. Intermediate monocytes, with a lifespan of around four days, are characterized by expressing the highest levels of antigen presentation-related molecules and are considered inflammatory cells[135,138–141]. Finally, non-classical monocytes are responsible for the antiviral response and have a half-life of seven days[135]. This monocyte subtype releases low levels of cytokines and mediates wound healing and tissue repair processes[131,136]. They are related to cytoskeletal dynamics and infiltration and are generally considered to have anti-inflammatory properties[137]. Although the functions mentioned above are the main functions, each monocyte subset may play different roles depending on tissue and microenvironmental signals[142]. In addition, as with other cell types, these monocyte categories are not homogenous populations, and further investigations are needed to characterize new specific subpopulations.

Regarding the contribution of monocytes to the pathogenesis of MS, they are known to promote inflammation[25]. The number of monocytes in the blood of MS patients has been associated with the clinical severity of the disease even in the early phase[143]. $CD16^{+}$ monocytes, mainly non-classical, are increased in the peripheral blood of MS patients, whereas classical

monocytes are reduced in comparison to healthy subjects[144]. Additionally, $CD16^+$ monocytes from MS patients show a highly inflammatory and neurotoxic phenotype and thus may be involved in T cell activation and contribute to nerve injury[145]. The percentage of monocytes producing IL-12, a cytokine secreted primarily by $CD16^+$ subsets, has been correlated with disease progression and activity[146]. Moreover, levels of monocytes secreting IL-6 and IL-12 are higher in MS patients compared to healthy controls[147]. Contrarily, $CD14^+$ monocytes from MS patients express lower levels of inflammatory cytokines and are the primary producers of IL-10, suggesting a regulatory role of this subset[145]. Currently, the use of novel techniques such as single-cell RNA sequencing (scRNA-seq) allow for the identification of tissue-specific monocyte profiles[148], description of new infiltrating monocyte subsets during neuroinflammation stages[149] and study of changes in the monocyte transcriptome of MS patients[150].

Dendritic cells (DCs)

DCs are professional APCs that act as a link between the innate and adaptive responses of the immune system and represent less than 1% of the total circulating peripheral blood mononuclear cells (PBMCs). They are located in lymphoid tissues (spleen, tonsils and lymph nodes) as resident DCs and in peripheral tissues (skin, lung, liver and intestine, among others) as migratory DCs[151].

In general, DCs can be subclassified based on anatomical location, origin, function and activation state. According to surface marker expression, two major DC subpopulations have been identified in peripheral blood: plasmacytoid DCs (pDCs) and conventional or classical DCs (cDCs) (**Figure 5**). pDCs are long-lived cells with a spherical shape, characterized by the expression of CD123 (IL-3R), CD303 (CLEC4C; BDCA-2) and CD304 (neuropilin; BDCA-4). They circulate in the blood as immature cells and are rapidly recruited to tissues and lymph nodes during inflammation. cDCs show a dendritic morphology and are migratory cells that move from tissues to lymphoid organs. They are short-lived and exhibit high phagocytic activity as immature cells and high cytokine production as mature cells. cDCs can be subclassified into two groups: cDC1, the minority group, which are $CD141^+$ (thrombomodulin; BDCA-3); and cDC2, which are characterized by the expression of CD1c (BDCA-1). Both subsets are found in blood, lymph nodes, tonsils, the spleen and bone marrow as well as in non-lymphoid tissues such as the skin, lungs, intestines and liver[152–154]. Despite the characterization of these markers, the full atlas of human DCs is not yet complete, since new RNA sequencing techniques are redefining the taxonomy of these cells in a more accurate way. In this line, Villani *et al.* proposed a new

classification of six DCs (DC1 to DC6), among which DC2 and DC3 were defined as two cDC2 subsets and where DC5 represented a new subtype with similar properties to pDCs that can additionally activate T cells[155].

Several immunity-related roles have been attributed to DCs. Their main function is to capture, process and present exogenous antigens to T cells[154]. In peripheral tissues, in the absence of inflammation or infection, DCs are in a resting immature state (iDCs), characterized by high capacity of uptake antigens through various mechanisms, including phagocytosis, endocytosis and macropinocytosis. After antigen interaction or inflammatory stimulation, DCs lose their ability to capture antigens and start to express MHC class II and costimulatory molecules on their surface. At this point, they become activated mature DCs (mDCs) and migrate to lymphoid organs, where they activate T cells, promoting the immune response. The shape of DCs, with elongated extensions, as well as their motility capacity, facilitate the antigen uptake and migratory functions of these cells[156,157]. Furthermore, due to their ability to present antigens, DCs are also involved in tolerance processes. Central and peripheral tolerance, both necessary to eliminate autoreactive T cell clones, are mediated by DCs[158]. DCs with immunosuppressive properties are called tolerogenic dendritic cells; they show an immature or semi-mature phenotype and produce large amounts of anti-inflammatory cytokines that induce Treg cells to suppress T cell reactivity[159,160].

Inflammatory DCs

In response to inflammatory stimuli, monocytes from the blood can reach tissues and differentiate into DCs[161,162]. These monocyte-derived DCs (moDCs) are generally considered inflammatory and can be found in the skin, lungs and intestines. They are able to present antigens and share some phenotypic traits with cDCs[162,163]. It has been reported that inflammatory moDCs mediate autoimmunity, as they are an important component of the onset and progression of autoimmune diseases[164]. Inflammatory moDCs express surface markers such as CD14, MHC class II, CD11c, CD1c (BDCA-1), CD206 (DC-SIGN), CD207 and CD209 [163,165]. However, the heterogeneity of these cells has recently begun to be revealed, with the discovery of up to seven moDC subtypes[166]. Due to the low levels of DCs in human blood, moDCs can also be generated *in vitro* using different methods for use in laboratory experiments [167]. One of the most commonly used protocols consists of the isolation of $CD14^+$ peripheral blood monocytes and their subsequent incubation with interleukin 4 (IL-4) and granulocyte macrophage colony-stimulating factor (GM-CSF)[168,169]. moDCs obtained *in vitro* show morphological, phenotypic,

physiological and functional similarities to immature DCs[170]. Gene expression analyses performed during moDC differentiation and in activated mature moDCs have revealed the upregulation of genes involved in cell adhesion, motility, antigen processing and presentation, regulation of the immune response, lipid metabolism, production of pro-inflammatory cytokines (such as TNF-α, IL-1β, and IL-6), signal transduction (such as the Wnt component WNT5A) and transcription factor production (such as NFAT)[171–174]. Other subtypes of DCs with pro-inflammatory properties include the DC3 subset, belonging to the cDC lineage and expanded in systemic lupus erythematosus patients[175], and the $SLAN^+$ DC, a cell type that expresses the Sulpho LacNAc (SLAN) carbohydrate modification of the P selectin glycoprotein ligand 1 and which has been linked to a $CD16^+$ monocyte subset and also to the DC4 subset[153,176].

Role of DCs in MS

DCs are considered critical for the induction and development of many autoimmune diseases, including MS, due to the loss of central and peripheral tolerance related to this kind of pathologies[158]. cDCs and pDCs have been detected in the CSF of MS patients[177]. Although it is unclear whether these DCs originate from CNS-resident microglia differentiation or are recruited from the periphery to the brain[157], evidence from EAE animal models suggests that DCs permit immune cell invasion to the CNS, leading to neurological damage[178]. DCs have also been identified within active demyelinating and inflammatory lesions in MS patients[179]. In comparison to healthy controls, the number of pDCs and cDCs in people suffering from MS is augmented in the CSF but not in the peripheral blood, and this number is related to the degree of inflammation[177,180,181]. Moreover, the $SLAN^+$ DC subpopulation is also present in the CSF of MS patients and accumulates in highly inflammatory brain lesions, suggesting an important role of these cells in the pathogenesis of the disease[182]. Besides alterations in the number of DCs, functional and phenotypic abnormalities of these cells have also been observed in MS patients[183], such as disturbed immunoregulatory function, impaired maturation and increased secretion of pro-inflammatory cytokines IFN-γ, TNF-α and IL-6[184–186]. All of these deficits and anomalies in DCs may contribute to MS pathology[187].

In addition to a pathological role, DCs may serve regulatory functions during autoimmunity[188]. Knowledge of DCs as protectors against or promoters of illnesses is being harnessed to design new therapies[189]. In the EAE mouse model, inhibition of disease development was achieved using lipopolysaccharide (LPS)-treated DCs[190]. Many current disease-modifying therapies used for MS have an effect on DCs, although none of them has been

specifically designed to target these cells[157]. New compounds affecting DC development and immunogenic function are under study[191]. The tolerogenic capacity of DCs is of special interest in strategies to suppress autoimmunity. Different protocols have been described to generate tolerogenic DCs *in vitro* with the use of anti-inflammatory cytokines (IL-10), immunosuppressive drugs (dexamethasone, rapamycin) and natural compounds (retinoic acid, vitamin D3, PGE2), among others. Each of these molecules induces a distinct transcriptional program, but they share functional properties[158,174]. These tolerogenic DCs obtained *in vitro* can inhibit pro-inflammatory T cells, induce tolerance and restore homeostasis, representing a powerful tool for clinical application in MS[192].

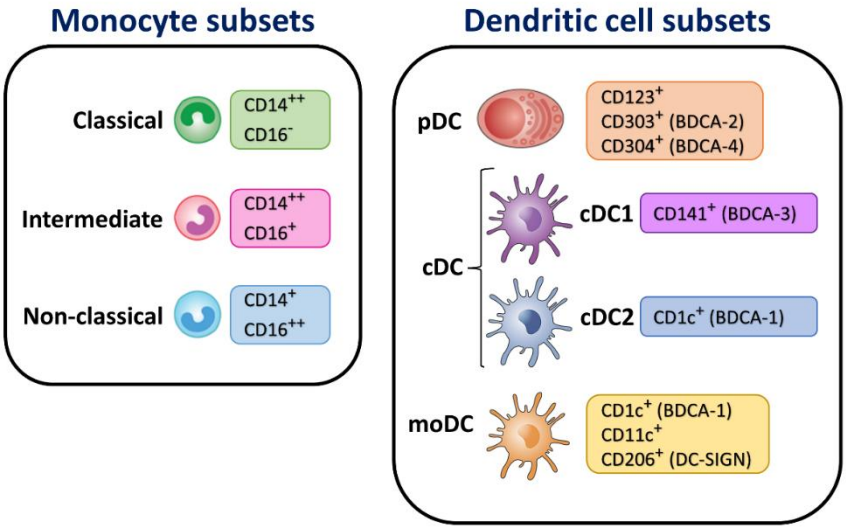


FIGURE 5. Monocyte and dendritic cell major subpopulations.
Monocytes can be classified according to expression of CD14 and CD16 surface markers into classical ($CD14^{++}$ $CD16^{-}$), intermediate ($CD14^{++}$ $CD16^{+}$) and non-classical ($CD14^{+}$ $CD16^{++}$) monocytes. Dendritic cells can originate from hematopoietic progenitors in the bone marrow, as occurs in plasmacytoid DCs (pDCs) and conventional or classical DCs (cDCs), or they can be derived from monocytes (moDCs).

Other immune cells

In addition to those mentioned above, other immune cells are relevant in MS pathology, including $CD8^{+}$ T cells, NKT cells, B cells, NK cells and macrophages.

$CD8^{+}$ T cells are present in the early stages of the disease and exist in close proximity to demyelinated axons and oligodendrocytes, and their number positively correlates with the extent of axonal damage. The presence of self-reactive $CD8^{+}$ T lymphocytes in MS patients and the cytotoxic function of these cells support the hypothesis that they contribute to the pathogenesis of the disease[193–196].

NKT cells are a minority subset of T cells that share properties of both T and NK cells. They have been reported to participate in different autoimmune diseases, including MS, but their specific contribution to the disease has not yet been fully elucidated[197].

B cells mediate humoral and cellular immunity and can recognize components of pathogens. The two main markers of these cells are CD19 and CD20[198]. B cells can differentiate into plasma cells, which produce high affinity antibodies, and into memory B cells, which can act as APCs. In MS patients, abnormalities in the function of B cells have been found, leading to the production of autoantibodies which reach the CNS and mediate brain cell death. In addition, compared to healthy donors, B cells from MS patients secrete higher amounts of pro-inflammatory cytokines and lower levels of anti-inflammatory cytokines[24,25,62,199,200].

Natural killer (NK) cells are part of the innate immune system and have cytolytic and secretory functions. The two major subsets are $CD56^{dim}$ $CD16^{+}$ NK cells (around 90–95% of the total NK cell population), considered efficient killer cells with cytotoxic and cytolytic functions, and $CD56^{bright}$ $CD16^{-}$ NK cells (around 5–10% of the total NK cell population), which secrete large amounts of cytokines. In MS, NK cells play a dual role. On the one hand, activated NK cells induce direct neural tissue damage and contribute to disease progression though the release of pro-inflammatory cytokines. On the other hand, these cells also serve protective functions by secreting anti-inflammatory cytokines and killing myelin-specific T cells[201–203].

Macrophages are mononuclear phagocytes that can be derived from monocytes (in the skin and gut and during inflammatory conditions) or have an independent origin. These cells show a wide variety of phenotypes and activation states, but, in general, they can be divided into two subtypes: classically activated macrophages (M1), with pro-inflammatory and cytotoxic characteristics, and alternatively activated macrophages (M2), with anti-inflammatory capacities and expression of neurotrophic factors. Due to this heterogeneity, macrophages also assume a dual role in MS. Infiltrated activated macrophages in the CNS are found in high numbers in active demyelinating lesions, where they secrete soluble mediators that contribute to axonal damage and neurodegeneration. However, macrophages have also been proven to support repair mechanisms through the production of growth factors and the elimination of myelin debris by phagocytosis, a necessary process for efficient remyelination[204–206].

1.4. Risk factors for MS

MS is a multifactorial disease influenced by the combination of genetic and environmental factors, each of which is believed to have a modest effect on the overall pathogenesis (**Figure 6**). The way these risk factors operate, both individually and combined, to induce MS pathogenesis is complex and still poorly understood. Ongoing research is taking place to elucidate the causes of MS[207].

1.4.1. Environmental factors

Many environmental and lifestyle factors have been studied as risk factors for MS. Each can somehow influence innate and/or adaptive immunity and pathogenic pathways, but unlike a genetic predisposition to MS, these environmental risk factors can be modified through prevention strategies in most cases. Among the factors that have shown a stronger association with the disease are Epstein-Barr virus (EBV) infection, smoking and low vitamin D levels[208]. Systematic reviews of meta-analysis studies revealed null or small effects of other environmental factors, including vaccinations, exposure to toxic environmental agents, comorbid diseases, surgeries and allergies[46], while others, such as sun exposure, diet and obesity, possibly contribute to MS risk but need further investigation to confirm[209].

Epstein-Barr virus and other viral infections

EBV is a member of the herpes virus family and represents one of the most common human viruses. More than 95% of all adults carry a persistent latent infection, but in some cases this latent virus may reactivate[210]. One of the symptomatic manifestations of EBV is infectious mononucleosis, which has been associated with a two-fold increase in the risk of developing MS[211]. The strongest evidence for this link was found in a study conducted in a cohort of more than 10 million people over the course of 20 years, in which a 32-fold risk increase was observed after infection with EBV[212]. Previously, meta-analysis and systematic reviews examined the link between EBV and MS, concluding a causal relationship between the virus infection and the development of the disease[213,214]. Furthermore, other experimental investigations have observed an accumulation of EBV-infected immune cells and virus reactivation in the CNS of MS patients[215]. The interaction between the human leukocyte antigen (HLA) MS risk genetic

variants and antibodies against EBV have also reached statistical significance[216]. Together, these data support the hypothesis that EBV is involved in the pathogenesis of MS.

Herpes viruses other than EBV have also been suggested as potential causes of MS, but studies report conflicting results. Prior infection with herpes simplex virus type 1 (HSV-1) may be correlated with pediatric MS, and HSV-2 may correlate with adult MS[217]. However, other analyses have revealed no relationship between genetically predicted HSV infection and MS risk[218]. On the other hand, infection by cytomegalovirus (CMV), another member of the herpes virus family, has been found to reduce the risk of MS in several studies[212,219], but other authors attribute a more pathogenic role to CMV[220]. Among the proposed hypothesis of how viral infections could drive autoimmunity, one of the most accepted is molecular mimicry. This mechanism is based on structural similarities or shared sequences between a foreign antigen (in this case, from the virus) and self-antigens, leading to T or B cell activation[221].

Smoking tobacco

Another widely documented lifestyle element associated with MS susceptibility and faster development of disability is smoking[222]. The probability of developing a progressive form of the disease from a relapsing-remitting clinical course is higher among smokers than nonsmokers[223]. Not only cigarette smoking but also passive exposure to environmental tobacco smoke is related to an augmented risk of MS, and this risk increases with longer duration of exposure[224,225]. Much of the physiological damage caused by smoking acts directly on the innate and adaptive immune system. This includes lung inflammation that promotes an enhanced inflammatory response, attenuated immune recognition of pathogens and reduced functional capabilities of immune cells such as macrophages, dendritic cells, B cells and T cells[226]. In addition, the effect of smoking on MS risk depends on the genetic predisposition of an individual. Among carriers of genetic risk factors for MS, smokers have an increased risk of developing the disease[227]. Apart from the negative effects on the immune system, tobacco also induces the production of neutralizing antibodies against drugs, such as natalizumab, used in treatment of MS[228]. All this, positions smoking as an independent risk factor contributing to MS pathogenesis.

Vitamin D and sunlight exposure

Low circulating levels of 25-hydroxyvitamin D, a marker of vitamin D status, have been linked to a higher MS risk[229]. Vitamin D has broad effects in modulating both the innate and

adaptive immune systems, including the suppression of T cell and B cell proliferation, inhibition of inflammatory cytokines production, promotion of a tolerogenic phenotype in monocytes and dendritic cells and enhancement of the Treg response[230]. The vitamin D receptor (VDR) occupies an essential place in mediating the metabolism and biological effects of vitamin D and is associated with several autoimmune diseases. In fact, gene polymorphisms of the VDR have been related to MS susceptibility and disability[231]. Conversely, high levels of 25-hydroxyvitamin D have been associated with a lower risk of developing MS[232] and predict reduced disease activity, reduced brain atrophy and slower clinical progression[233]. Moreover, in MS patients, vitamin D shows an inverse association with CSF neurofilament light levels, used as marker for axonal injury[234]. These findings suggest that vitamin D confers a protective effect in the aetiology of MS. In this regard, vitamin D has emerged as a possible therapeutic approach to treat MS. Studies in EAE animal models have demonstrated that 25-hydroxyvitamin D slows disease severity and has regenerative properties. The application of this therapy in humans faces challenges such as the difficulty of finding the optimal dose, but ongoing clinical trials appear promising[235].

The major source of vitamin D is sunlight exposure. Ultraviolet radiation (UVR) from sunlight is absorbed in the skin and used to convert a precursor molecule into the active metabolite of vitamin D[236]. Thus, it is difficult to distinguish between the effects of UVR and vitamin D separately, although both factors have been found to independently decrease the risk of CNS demyelination[237]. Subjects with low exposure to sunlight have a significantly increased risk of MS compared to those who are exposed to higher intensities, demonstrating an important protective effect of UVR against the disease[238]. In relation to sunlight exposure, the association between geographical latitude and MS has been studied. A substantial body of research has established a latitude gradient in MS incidence and disease activity, showing an increasing MS risk from the equator, where the UVR intensity is higher, to the poles[48,239].

Obesity and diet

Large cohort studies have associated obesity during adolescence, in both boys and girls, with an increased risk of developing MS in the future[240,241]. Body mass index, used as a measure of body fat, is determined by many genetic variants, some of which are associated with higher MS risk, suggesting a link between MS pathogenesis and obesity[242,243]. The mechanistic pathways underlying this association are still unclear, but some hypotheses implicate a role of inflammation[208,243].

Results from dietary studies are difficult to interpret, and conclusions should be drawn with caution due to the complex interactions between different nutrients. For instance, alcohol and caffeine showed no correlation with MS risk in one study[244], while other research found a dose-dependent inverse association between alcohol consumption and the incidence of MS[245]. The same happens with high sodium intake (dietary salt), which has been associated in one case with increased disease activity in MS patients[246]. Another study, however, failed to find such an association[247].

Microbiota

Recent studies have highlighted the importance of the type and distribution of gut microbiota as an environmental risk factor for MS. Metagenomic analysis identified gut microbiome alterations in MS patients compared to healthy control individuals and revealed interactions between the gut microbiome and host genome. These studies also found correlations between some bacterial species and disease severity[248,249]. Furthermore, faecal samples from MS patients transplanted into germ-free mouse models promote EAE disease, providing evidence that the microbiota of MS patients contains factors that induce an autoimmune disease similar to MS[250].

The commensal microbiota interact with immune cells and participate in immunoregulatory pathways, conferring protection against diseases[251]. Changes in the microbiota composition associated with diet and the environment can lead to dysbiosis, defined as an imbalanced microbiota, which is associated with autoimmune pathologies. In MS, gut dysbiosis mediates inflammation and demyelination of the CNS[252]. Therefore, the regulation of the gut microbiota, using probiotics or other strategies, is a powerful tool for the treatment and prevention of MS.

Psychological stress

The link between stress and MS has generated controversial results, possibly due to heterogeneity in study design and different methods of stress measurement[253]. Some results have not supported psychological stress as a risk factor in the development of MS or have found only a weak association between stress and disease exacerbation[254,255]. However, many other investigations have shown an increased risk of exacerbation, defined as worsening of existing symptoms or appearance of new symptoms, in MS patients that have experienced at least one stressful event[256,257]. Among the stressful events included in these studies were home and

work stress, financial problems, sexual abuse during childhood or adolescence, job loss, problems in relationships and death of a close family member. Stressor duration, severity and frequency are more important than type and source of stress in influencing disease progression[258]. In conclusion, although the detailed biological mechanisms mediating the stress–MS relationship are still unknown, there is sufficient evidence to suggest that psychological stress may be a risk factor for MS exacerbation[259].

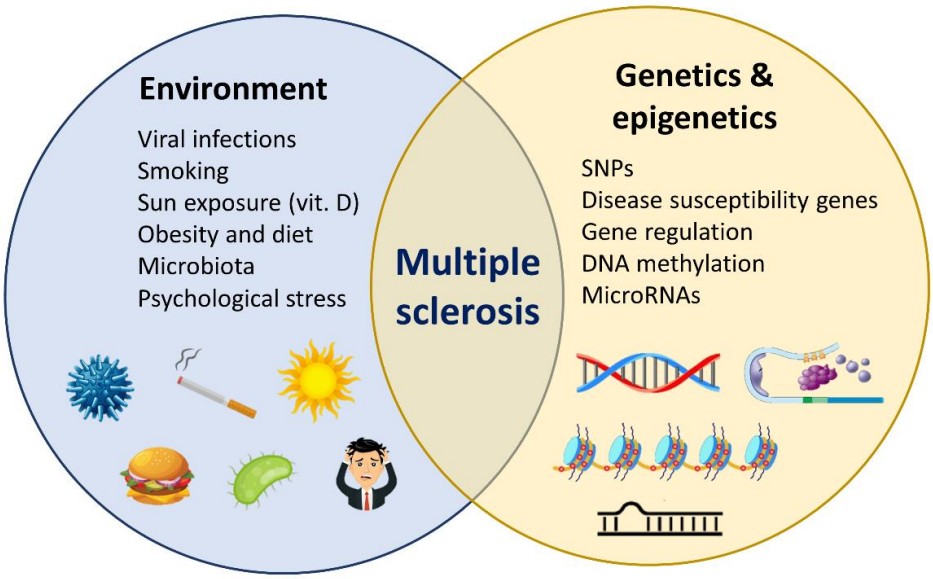


FIGURE 6. Environmental and genetic factors that contribute to MS pathogenesis.
Adapted from Oksenberg *et al.*[260]

1.4.2. Genetic factors

Genetic background of MS

The scientific evidence of recent decades has indicated a clear role of genetic factors in the onset and progression of MS. With the development of novel techniques and improved statistical methods, more than 230 genetic risk variants have been associated with MS susceptibility thus far. Most of these variants are single nucleotide polymorphisms (SNPs) located within or near genes expressed mainly in innate and acquired immune cell types, but they can also be expressed in some brain resident cells, such as microglia and astrocytes[261]. MS fits the common disease–common variant theory, which states that common diseases in a population are caused by many small, frequent genetic variations. In fact, the majority of MS susceptibility alleles are relatively common in the population, with a minor allele frequency of > 5%. None of these high frequency-associated alleles cause MS on their own, but it is known that

one locus (HLADRB1*15:01) contributes with moderate effect to the disease and many loci do with small effect, conferring an increased risk of less than 1.5 fold[262].

MS is considered a partially heritable disease. Studies performed in families have provided strong support for a complex genetic component in MS. Among the results found, a higher concordance rate is shown in monozygotic or identical twins (20–30%) compared to dizygotic or fraternal twins (2–5%)[263,264]. Twin studies have also helped distinguish the influence of genetic predisposition from the impact of environmental factors and identify some of the non-heritable immune alterations in MS[265,266]. Although MS is not considered a Mendelian disease, the risk of MS increases with a greater number of genetic traits shared with the affected family member. In fact, 15–20% of MS patients have a family history of the disease, a much higher prevalence than the general population, and first-, second- and third-degree relatives of MS patients are 10 to 30 times more likely to develop the pathology than the general population[267]. Rare variants, with allele frequencies < 5%, might contribute relatively large effects to the genetics of familial MS and explain in part the heritability of the disease[268]. In this regard, some investigations have revealed some genetic risk factors implicated in familial MS, including variants in genes of nuclear receptors *(NR1H3, NCOA3)*, plasminogen *(PLG)*, cation channels *(KCNG4, SLC24A6* and *SLC8B1)*, Wnt signaling *(UBR2, CTNNA3, NFATC2, RNF213)*, inflammasome assembly *(NLRP12)* and fibrinolysis and complement pathways *(PLAU, MASP1, C2)*[269–271].

Genome-wide associations studies (GWAS) in MS

A breakthrough in the study of the genetic background associated with the risk of MS has been determined by genome-wide associations studies (GWAS), which examine genetic variants, typically SNPs, across the whole genome of many different individuals to identify genetic factors that are associated with a trait, generally in human diseases[272]. Since the first GWAS performed on MS in 2007, many more studies have increased the number of identified risk variants associated with the disease, many of which were found outside the MHC[273]. The most recent GWAS by the International Multiple Sclerosis Genetics Consortium (IMSGC) was performed in 2018 using a sample size of more than 47,000 cases and over 68,000 healthy controls. This study revealed a detailed genetic map of common variants that explains almost half of the estimated heritability for MS and identified a total of 233 independent associations, of which 200 are located in the autosomal non-MHC genome, 32 are contained in the MHC region, and one is found on the X chromosome (**Figure 7**). MS susceptibility genes found in this

work are expressed in many different types of immune cells and are related to the development, maturation and differentiation of T, B, NK and dendritic cells. It is worth mentioning that this GWAS also showed an enrichment for MS genes in microglia, the resident immune cells of the brain. The *CLECL1* locus in particular suggested an important role of microglia in the mechanisms of neurodegeneration linked to the disease[274]. In summary, the majority of MS GWAS loci encode genes related to the immune system, particularly lymphocytes, corroborating the hypothesis that MS is an autoimmune disease. In addition, some of the genetic variants associated with MS have also been identified as genetic risk factors for other autoimmune diseases, such as psoriasis, Crohn's disease, lupus erythematosus, rheumatoid arthritis and type 1 diabetes, revealing a shared genetic background[275].

The human leukocyte antigen (HLA) locus and other genetic variants associated with MS

The first identified genetic risk factor related to MS was the human leukocyte antigen (HLA) locus, first reported in the 1970s[276,277]. HLA is the name for the human MHC. The MHC locus is composed of a genomic region at the chromosomal position 6p21 , occupies about 0.1% of the human genome and contains more than 150 genes[278]. This region is the most polymorphic region of the human genome, with over 25,000 alleles identified, and its genes encode surface proteins that play a key role in the immune response[279]. The MHC region includes the class I and class II subregions. The *HLA-DRB1* gene, belonging to MHC class II, constitutes the strongest MS risk genetic factor, showing a high statistically significant association in almost all populations studied[260]. This gene is expressed in APCs, and its corresponding protein displays antigenic peptides to $CD4^+$ T cells in order to induce tolerance or attack pathogenic agents[280]. The HLA-DRB1*15:01 allele has been shown to be the main MS susceptibility allele in European populations[281], and it has been also associated with disease severity[282]. In European people, homozygotes for the HLA-DRB1*15:01 allele carry an average odds ratio > 6, while heterozygotes have an odds ratio of around 3[283]. However, in African American people, the odds ratio for HLA-DRB1*15:01 declines to 1.7[284], and in other groups it is also lower than in most European populations[285]. Other alleles that increase the risk of MS are HLA-DRB1*03:01, HLA-DRB1*15:03 and HLA-DRB1*13:03, although all of these do so to a lesser extent than HLADRB1*15:01[286]. Moreover, there are other alleles that protect against the disease, such as HLA-DRB1*14 and HLA-DRB1*11. In the case of HLA-DRB1*14, it completely eliminates the risk effect of the HLA-DRB1*15 allele when they are inherited together[287]. In

addition to MHC class II alleles, several MHC class I alleles have also been implicated in MS, including HLA-A*02:01 and HLA*B44, both associated with a reduced risk of developing the disease[288,289].

Besides the HLA locus, variants in many other non-MHC genes have been related to MS susceptibility. Some of the proteins encoded by these genes include interleukin receptors (*IL7R*, *IL2RA*, *IL22RA2*) involved in immune tolerance and lymphocytes development; transcription factors (*IRF8*, *SOX8*) that act in the CNS and the immune system; transmembrane proteins (*CD58*, *CD6*, *CD226*) with roles in cell adhesion and T lymphocyte activation; zinc-finger proteins (*ZBTB46*) that function as transcriptional regulators; ribosomal proteins (*RPL5*, *RPS6KB1*); mannosidases (*MANBA*); proteins related to cytokine signaling (*TNFRSF1A*, *SOCS1*); and proteins involved in the regulation of the cell cycle (*EVI5*), autophagy (*CLEC16A*), cell growth (*TYK2*) and B cell migration (*CXCR5*)[290–294]. In addition, some other genetic variants have been linked not only to MS susceptibility but also to the progression and severity of the disease. These include variants associated with *CPXM2*, *IGSF9B* and *CYP2R1* genes[295,296]. In line with this, a recent GWAS identified a novel locus associated with MS severity. This work also suggested a key role of CNS resilience and neurocognitive reserve in determining MS outcomes, in contrast to the immune signals observed for disease susceptibility[297].

Epigenetics, gene–environment interactions and future perspectives in the genetics of MS

The still-missing heritability of MS could be explained by rare variants, epigenetic effects, gene–gene interactions, gene–environment interactions and genetic regulators, among other mechanisms[298]. Low-frequency variants (minor allele frequency between 1–5%) and rare variants (minor allele frequency < 1%) have been shown to contribute to MS risk and explain a fraction of disease heritability that cannot be attributed to common variants[299]. In addition, many epigenetic modifications, such as acetylation, methylation and citrullination, that affect genes related to inflammation and neurodegeneration are involved in the pathogenesis of MS[300]. As an example, increased histone acetylation and citrullination of MBP have been detected in the brains of MS patients, and these modifications have been suggested to exacerbate the disease course[301,302]. Epigenetic regulation can also be affected by microRNAs (miRNA), a family of nonprotein-coding small RNAs that downregulate the expression levels of specific mRNAs by mRNA degradation or inhibition of translation. Researchers from different groups have found differences in miRNA expression in various tissues of MS patients compared

to healthy controls, highlighting the importance of miRNAs in molecular processes implicated in the development of the disease[303,304]. Combined, these discoveries have led to a search for treatments targeting the epigenetic mechanisms involved in MS.

Some of the challenges that remain to fully complete the architecture of MS include understanding the exact mechanisms of gene–gene and gene–environment interactions as well as a detailed functional analysis of GWAS-associated genes. The majority of associations map to either intronic or intergenic regions, that is, parts of the genome likely involved in regulatory processes. Therefore, the study of regulatory regions, including the transient transcriptome, is also necessary to uncover these molecular mechanisms, their influence on the activity of other genes and the interaction between genetic and non-genetic factors[305,306]. RNA-seq and other novel high-throughput technologies are greatly increasing the knowledge in cell type-specific transcriptomics, showing heterogeneity in gene expression profiles of immune cells and glial cells and providing evidence for alterations in patients with MS[307–309].

In recent years, the information obtained from genetic studies has been used to develop diagnostic algorithms capable of predicting disease susceptibility. Based on the largest GWAS for MS performed to date, polygenic risk scores have been generated including all identified genome-wide significant susceptibility variants, with the goal of predicting the cumulative effects of genetic burden[310]. Some other MS risk scores have been created, integrating both genetic susceptibility risk variants and the effects of environmental risk factors[311,312]. Other models aimed to predict the transition from RRMS to SPMS, which is relevant to improve the prognoses for patients[313]. However, there are several theoretical and practical limitations that make it difficult to predict whether an individual will develop MS. Completing the missing heritability of MS, improving the models of interactions between genetic and environmental factors and use of GWAS data from non-European cohorts pose challenges to achieving more accurate predictive scoring systems[314]. These genetic and predictive methods, together with other “-omics” techniques, will provide a more complete picture of the disease, and their application to clinical practice will enable precision medicine, thus improving diagnosis, treatment choices and prognosis.

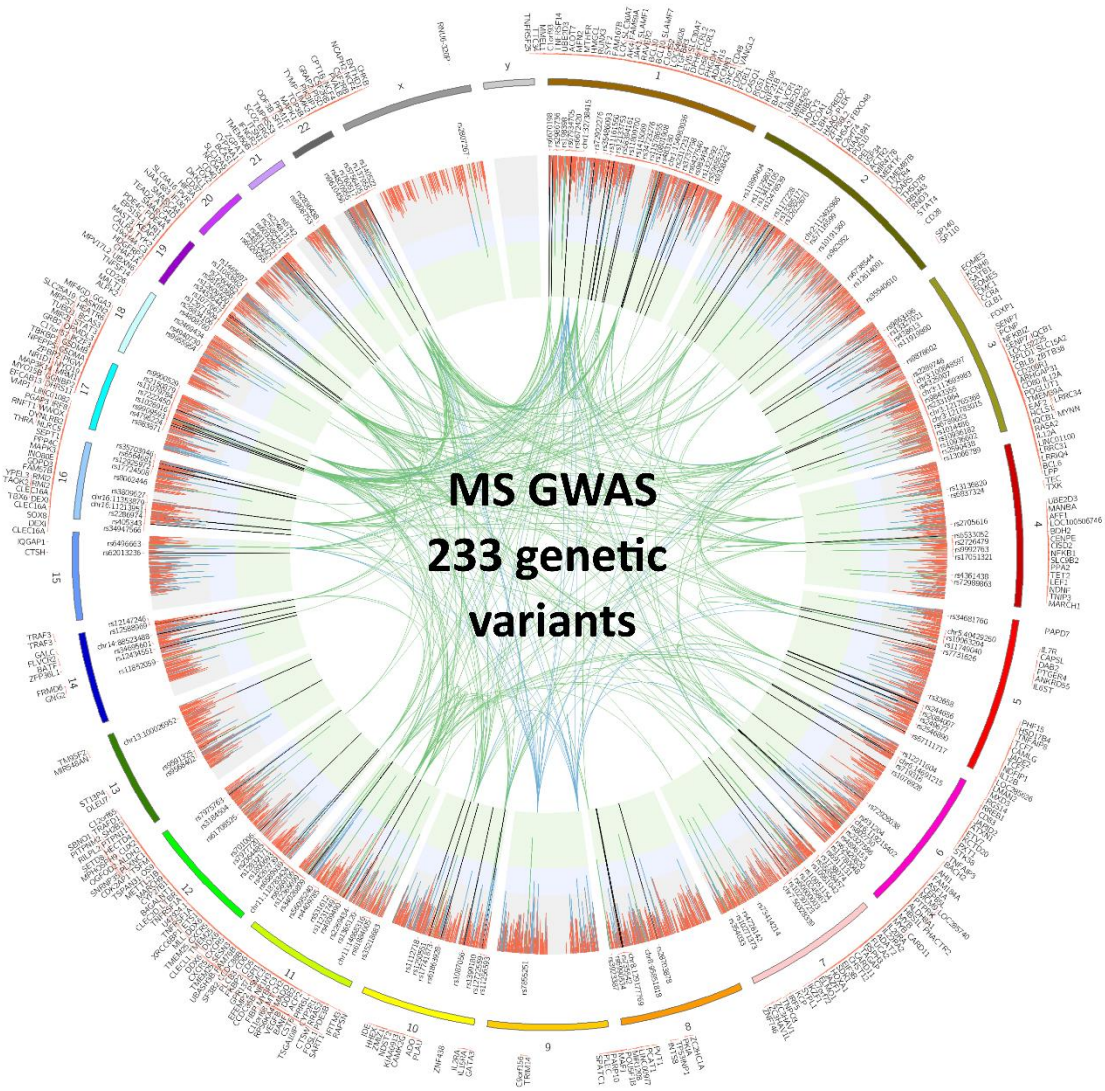


FIGURE 7. The genetic map of MS.
The outer surface shows 551 prioritized genes. Each vertical line in the inner layers represents one effect. A total of 200 autosomal non-MHC and one chromosome X genome-wide effects are listed. The lines in the inner circle represent protein–protein interactions. Source: International Multiple Sclerosis Genetics Consortium, 2019[274].

1.5. *ANKRD55* and neighboring genes

The Ankyrin repeat domain-containing protein 55 (*ANKRD55*) gene has been demonstrated to be of interest in MS and other diseases. It is located on chromosome 5q11.2, a region which also contains various protein-coding genes, including *IL6ST*, *IL31RA*, *DDX4* and *SLC38A9*, situated at 104.68 kbp, 176.82 kbp, 282.52 kbp and 362.48 kbp, respectively, from *ANKRD55* (**Figure 8**).

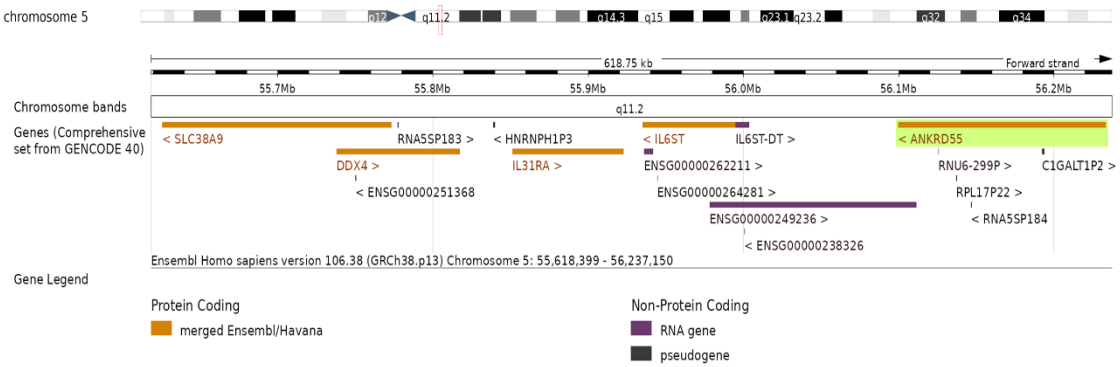


FIGURE 8. *ANKRD55* and neighboring genes.
Detail of chromosome 5q11.2 region, including the genes *ANKRD55*, *IL6ST*, *IL31RA*, *DDX4*, and *SLC38A9*. Image obtained from Ensembl human GRCh38.p13[315] (https://www.ensembl.org).

1.5.1. *ANKRD55*

ANKRD55 gene

Based on the dbSNP database[316] (https://www.ncbi.nlm.nih.gov/snp), the *ANKRD55* gene has a total length of 133,651 nucleotides. Six unique isoforms of this gene have been characterized to date, although not all with the same level of certainty (Ensembl database GRCh38.p13[315]). Four of these isoforms code for proteins (*ANKRD55-201*, *ANKRD55-203*, *ANKRD55-202* and *ANKRD55-205*), while two are non-coding processed transcripts (*ANKRD55-204* and *ANKRD55-206*). Based on transcript support levels (TSL) from the Ensembl database[315], which is a method to highlight the well-supported transcript models, *ANKRD55-201*, *ANKRD55-202* and *ANKRD55-204* are the most confirmed and best-supported isoforms. The full-length isoform, *ANKRD55-201*, contains 12 exons (2,471 base pairs) and codes for a protein of 614 amino acids. The *ANKRD55-203* isoform, with 2,139 base pairs and coding for a 571-amino acid protein, is formed by 10 exons and lacks exons 1 and 7 of *ANKRD55-201*. Isoform *ANKRD55-202* (1,518 base pairs and coding for a 326-amino acid protein) contains only the final four exons of

ANKRD55-201, while *ANKRD55-205* (626 base pairs and coding for a 178-amino acid protein) shares exons 3, 4, 5, 6 and 7 with the full-length isoform. The non-coding transcript *ANKRD55-204* (853 base pairs in 7 exons) contains exons 6, 7, 9 and 10 (partially) of the *ANKRD55-201* isoform and has three new exons. Finally, the non-coding *ANKRD55-206* isoform, with 349 base pairs contained in 4 exons, shares exons 1, 2, 3 and part of 4 with *ANKRD55-201*[317] (**Figure 9**).

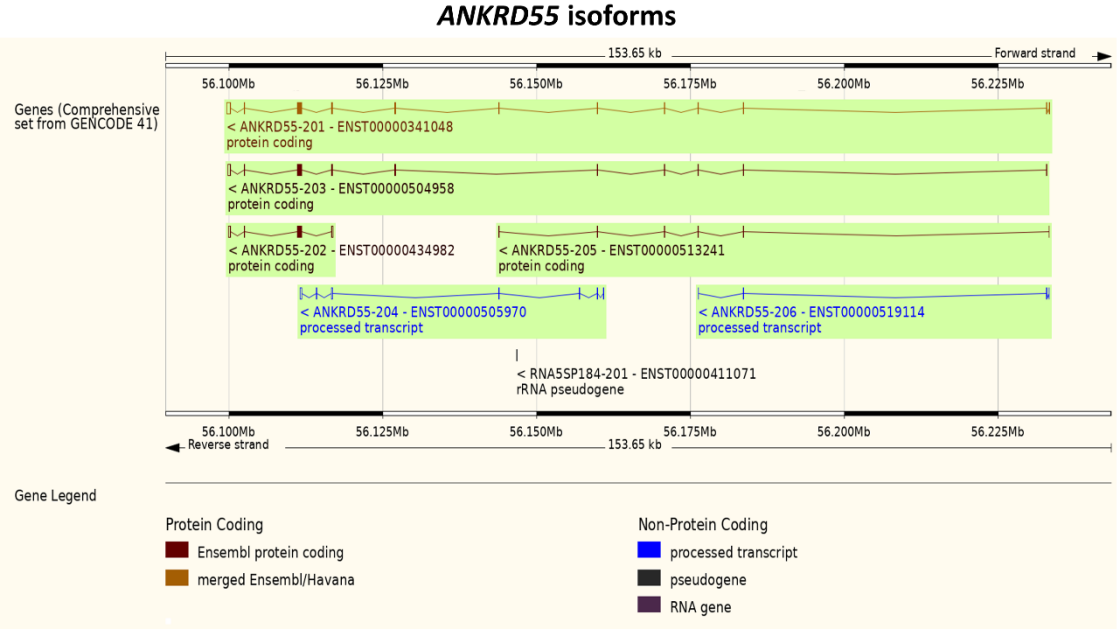


FIGURE 9. Schematic representation of the *ANKRD55* gene isoforms.
The *ANKRD55* gene has six isoforms: four protein-coding transcripts (*ANKRD55-201*, *ANKRD55-203*, *ANKRD55-202* and *ANKRD55-205*) and two non-coding processed transcripts (*ANKRD55-204* and *ANKRD55-206*). Vertical lines represent exons and horizontal lines represent introns. Image obtained from Ensembl human GRCh38.p13[315].

Several SNPs contained in the *ANKRD55* gene have emerged as risk variants for diseases with predominantly autoimmune or chronic inflammatory etiopathogenesis, such as MS[274,318–322], rheumatoid arthritis[323–331] and juvenile idiopathic arthritis[332–334]. Moreover, some of these SNPs, as well as others located in the intergenic region between *ANKRD55* and *MAP3K1*, have also been associated with other disorders, including pediatric autoimmune diseases[335], celiac disease[336,337], inflammatory bowel disease[338,339] (specifically Crohn's disease[340]), type 1 diabetes[341,342], type 2 diabetes[343–345], Hashimoto's thyroiditis[346], open-angle glaucoma[347], psoriasis[348], dermatomyositis and polymyositis[349], Alzheimer's disease[350], prostate cancer[351] and posttraumatic stress disorder[352] (**Figure 10** and **Table 1**). The intronic or intergenic location of these SNPs points to allelic variation in regulatory elements as a potential mechanism for transcriptional modulation of *ANKRD55* and/or adjacent genes.

In addition to the association with disease susceptibility, some SNPs in *ANKRD55* have been linked to different pathological, biological or genetic traits, including worsening of disability in MS[353], adiposity[354], reduced IgA levels[355], dendritic cell-specific enhancer[356], IgG glycosylation[357], enhancer (H3K4me1) with potential binding for STAT transcription factors[358] and expression of quantitative trait loci (eQTL) in $CD4^+$ T cells associated with MS, rheumatoid arthritis and Crohn's disease[359]. Differences in the expression of *ANKRD55* gene (upregulation or downregulation) have also been observed in thymoma-associated myasthenia gravis[360], male infertility[361], dermatomyositis and polymyositis[362], breast cancer[363,364] and central nervous system neuroblastoma[365]. Interestingly, one work suggests a protective role of ANKRD55 in pancreatic ductal adenocarcinoma, mediating survival benefits through the remodelling of the immune microenvironment[366].

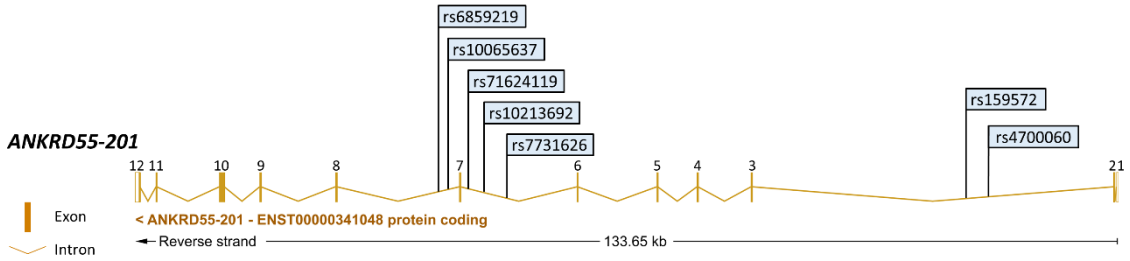


FIGURE 10. Locations of the main disease-associated SNPs in the full-length *ANKRD55-201* isoform.
Image obtained and modified from Ensembl human GRCh38.p13[315].

TABLE 1. SNPs in the *ANKRD55* region most strongly associated with diseases.
Each SNP position, indicated between brackets in the first column, is based on the human reference genome GRCh38.p13. MS: multiple sclerosis; RA: rheumatoid arthritis; CD: Crohn's disease; JIA: juvenile idiopathic arthritis; T1D: type 1 diabetes; pJIA: pediatric JIA; PTSD: posttraumatic stress disorder; AD: Alzheimer's disease; OAG: open-angle glaucoma; T2D: type 2 diabetes. The '*p*-value' column indicates the strongest association value of the GWAS studies indicated in the 'References' column for each SNP.

SNP	Gen	Disease	*p*-value	References
rs6859219 (Chr 5: 56142753)	*ANKRD55*	MS	$\leq 1.18 \times 10^{-8}$	[318–321]
		RA	9.60×10^{-12}	[323]
rs10065637 (Chr 5: 56143024)	*ANKRD55*	CD	3.86×10^{-12}	[340]
rs71624119 (Chr 5: 56144903)	*ANKRD55*	MS	3.40×10^{-13}	[322]
		RA	$\leq 4.90 \times 10^{-7}$	[324–328]
		JIA	4.40×10^{-11}	[332]
rs10213692 (Chr 5: 56146422)	*ANKRD55*	JIA	$\leq 3 \times 10^{-11}$	[332,333]
		T1D	2.85×10^{-9}	[341]
rs7731626 (Chr 5: 56148856)	*ANKRD55*	MS	3.89×10^{-15}	[274]
		RA	7.30×10^{-24}	[329]
		JIA	$\leq 1 \times 10^{-10}$	[333,334]
		pJIA and CD	1.4×10^{-10}	[335]
rs159572 (Chr 5: 56211219)	*ANKRD55*	PTSD	2.43×10^{-8}	[352]
rs4700060 (Chr 5: 56214829)	*ANKRD55*	AD	1.07×10^{-8}	[350]
rs61275591 (Chr 5: 56479729)	Intergenic (*ANKRD55- MAP3K1*)	OAG	5.78×10^{-11}	[347]
rs459193 (Chr 5: 56510924)	Intergenic (*ANKRD55- MAP3K1*)	T2D	$\leq 6.0 \times 10^{-9}$	[343,344]

ANKRD55 is expressed in many human tissues but generally at low levels (**Figure 11**). According to various databases (UniProt[367] https://www.uniprot.org; Human Protein Atlas[368] https://www.proteinatlas.org; and GeneCards[369] https://www.genecards.org), the highest expression occurs in the testes, brain, lymphoid tissues, white blood cells and liver. Other tissues expressing *ANKRD55* include the small intestine, salivary glands, placenta, adipose tissue, skin, kidneys, stomach, lungs and heart, among others. The RNA single cell type specificity shows enhanced expression in early spermatids, inhibitory neurons, spermatocytes, hepatic stellate cells and late spermatids (Human Protein Atlas[368]). Within immune cells, *ANKRD55* is mainly (but not exclusively) expressed in basophils, eosinophils and $CD4^+$ T cells (naïve[317], Th1, Th2,

Th17, Treg and effector memory cells[370], among other $CD4^+$ T cell subsets) (Human Protein Atlas[368]; DICE database https://dice-database.org).

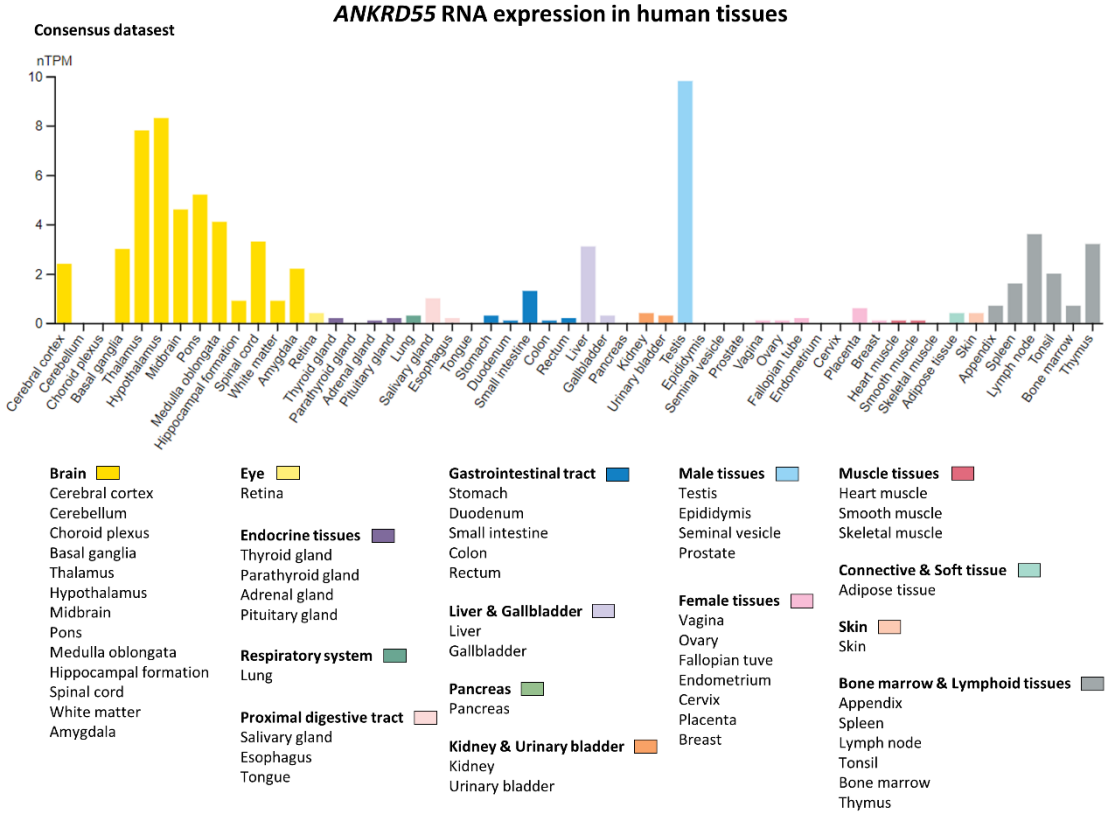


FIGURE 11. Expression of mRNA *ANKRD55* in human tissues.
The dataset shown, created by combining the HPA and the Genotype-Tissue Expression (GTEx) RNA-seq tissue datasets, is reported as nTPM (normalized protein-coding transcripts per million) for 54 tissue types. Color-coding is based on tissue groups, each consisting of tissues with common functional features. Imagen from Human Protein Atlas[368] version 21.1.

ANKRD55 protein

The ANKRD55-201 isoform (considered the canonical sequence) is a protein of 614 amino acids, has a molecular mass of 68.4 kDa and contains nine ankyrin repeat motifs (Ensembl[315]; UniProt[367]). Ankyrin repeat motifs are some of the most common protein motifs in nature. Each of the ankyrin repeat sequences consist of 33–36 amino acid residues, some of which are highly conserved in certain positions of the ankyrin repeat sequences. The number of ankyrin repeats can vary from 2 to 34 per protein, and most proteins contain six repeats. The ankyrin repeats have a well-defined structure, with an arrangement of two antiparallel α-helices followed by a β-harpin or a long loop. Consecutive repeats stack together, forming an L-conformational-shaped domain[371]. Proteins that contain ankyrin-like repeats in their structure

typically mediate protein–protein interactions that are involved in the regulation of transcription, membrane and cytoskeletal organization, initiation of immune response and modulation of the cell cycle, among many other biological processes[372]. Although ankyrin repeat proteins show great versatility in protein binding, they do not recognize specific sequences. Different ankyrin repeat proteins can bind to the same target but modulate it in different ways; the same ankyrin repeat protein may also bind to multiple proteins[373]. Ankyrin repeat proteins are present in bacteria, archaea and eukaryotes, as well as in some viral genomes. Genetic alterations in genes encoding these proteins have been linked to human diseases, especially some types of cancer[374]. The ANKRD55 protein has been poorly studied, and therefore its structural conformation is not well understood. The AlphaFold bioinformatic tool[375,376] (https://alphafold.ebi.ac.uk), based on an artificial intelligence system that predicts the three-dimensional structure of proteins from their amino acid sequences, currently offers the most accurate model (**Figure 12**). Still, in some regions of the protein, the structure is predicted with a very low confidence level, while the nine ankyrin repeat motifs show the highest confidence score. Additionally, recent data revealed six phosphorylation sites in the human ANKRD55 protein: three on threonine residues (T6, T11 and T189) and three on serine residues (S436, S475 and S597)[377].

Despite all the studies and information on ankyrin repeat motifs, little is still known about the biological function of the ANKRD55 protein. Recent interactome and knockdown studies have provided some initial, though heterogeneous, indications. On the basis of integrated mass spectrometry protein interaction datasets, ANKRD55 was first identified as a novel member of the intraflagellar transport machinery[378]. In HEK293 cells, recombinant ANKRD55 in the cytosol preferentially interacts with 14-3-3 isoforms and ATP- or nucleotide-binding proteins. It also appeared capable of entering the nucleus of HEK293 cells, where it is found to preferentially associate with RNA-binding proteins and proteins involved in sumoylation[377]. In preadipocytes, silencing of ANKRD55 by the lentiviral CRISPR/Cas9 system increased both their proliferation rate and lipolysis[379]. Other studies have shown a downregulation of *ANKRD55* by transforming growth factor beta (TGF-β) in eosinophils[380] and a reduction of *ANKRD55* gene expression level in mucosal-associated invariant T cells compared to conventional T cells[381]. However, a direct functional relevance of ANKRD55 to autoimmune processes has not been demonstrated.

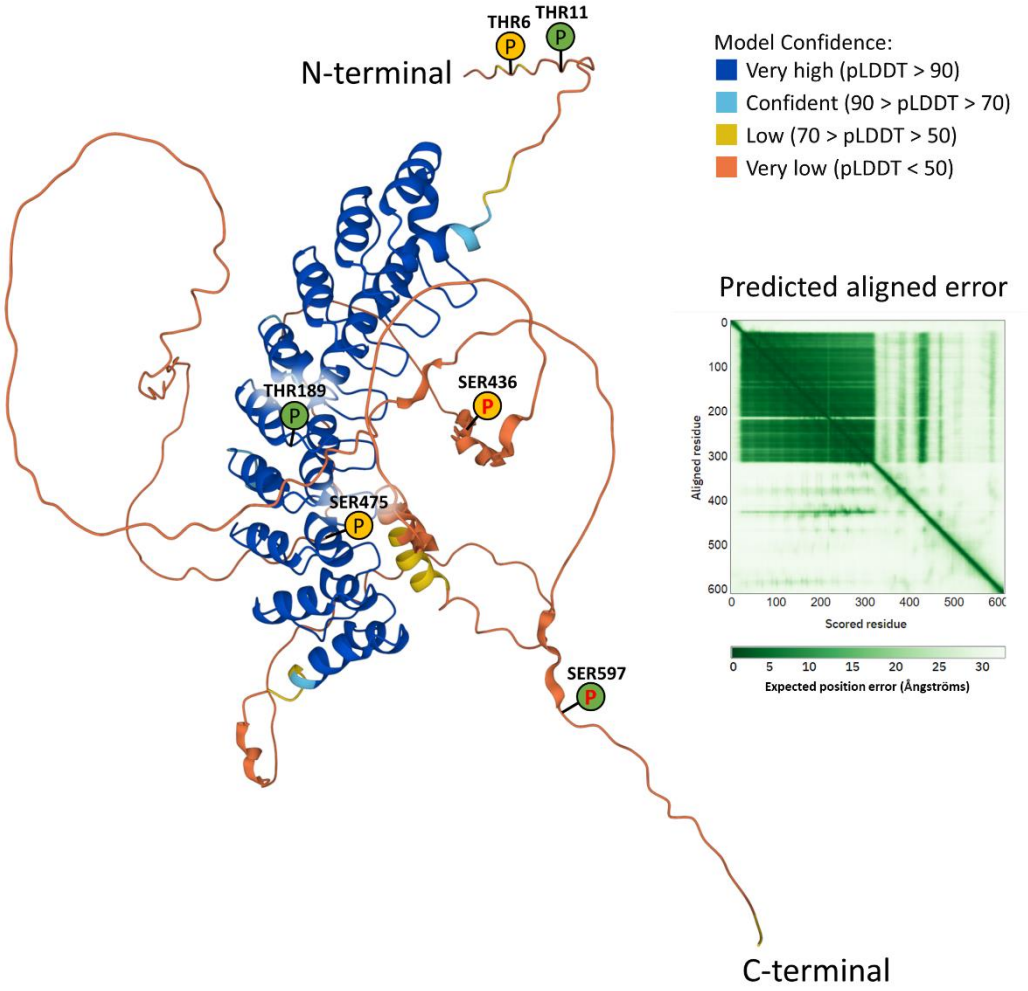


FIGURE 12. Three-dimensional structure of the ANKRD55 protein.

The structure of ANKRD55 was predicted using the AlphaFold artificial intelligence system[375,376] (https://alphafold.ebi.ac.uk). AlphaFold estimates errors using the predicted local distance difference test (pLDDT), which gives a per-residue confidence score from 0 to 100. The model is shown in colors, with high confidence residues colored in blue (including the nine ankyrin repeat domains) and low confidence residues in yellow and orange. Predicted aligned error gives a distance error for every pair of residues, indicating if domains are correctly positioned relative to one another. Based on data from Ugidos *et al.*[377], predicted phosphorylation sites (THR11, THR189 and SER597) are indicated with a green circled "P" and experimentally detected phosphorylation sites (THR6, SER436 and SER475) with a yellow circled "P". Of these, SER436 and SER597 (indicated by a red "P") exhibit the highest scores as likely 14-3-3 binding phosphosites.

1.5.2. Neighboring genes of *ANKRD55*

IL6ST

Interleukin 6 cytokine family signal transducer (*IL6ST*) codes for the interleukin-6 receptor subunit beta protein, also known as glycoprotein 130 (gp130) or CD130. This protein forms a part of membrane receptor complexes and functions as a signal transducer for members of the IL-6 family cytokines, such as interleukin 6 (IL-6), interleukin 11 (IL-11), interleukin 27 (IL-27), leukemia inhibitory factor (LIF), oncostatin M (OSM), ciliary neurotrophic factor (CNTF), cardiotrophin-1 (CT-1) and cardiotrophin-like cytokine (CLC), among others[382]. These cytokines activate the Janus kinase–signal transducer and activator of transcription (JAK-STAT) signaling pathway and are implicated in the regulation of the immune response, hematopoiesis, bone metabolism, and the development of the cardiovascular and neuronal systems[383].

Of the cytokines that bind to gp130, special interest has been focused on IL-6, which is considered the archetypal member of this cytokine family. IL-6 is a pleiotropic cytokine secreted mainly by APCs (dendritic cells, macrophages and B cells) but also by keratinocytes, fibroblasts and astrocytes. It participates in the inflammatory and immune responses by modulating hematopoiesis and inducing B cell maturation and antibody secretion, and it also promotes proliferation, survival and differentiation of $CD4^+$ T cells[384,385]. Given the pro-inflammatory properties of IL-6, this cytokine contributes to the pathogenesis of cancer and inflammatory diseases including MS, rheumatoid arthritis, systemic lupus erythematosus and inflammatory bowel disease[386,387].

IL-6 utilizes two different signaling mechanisms. In the classical signaling pathway, IL-6 binds to its membrane-bound receptor (IL-6R), while in the trans-signaling pathway, a soluble form of IL-6R (sIL-6R) binds to secreted IL-6 (**Figure 13A**). In both pathways, the complex formed by IL-6 and IL-6R (or sIL-6R) needs to associate with gp130 protein to initiate the intracellular signaling cascade[388]. All tissues express gp130, whereas IL-6R is only found in certain cell types, such as monocytes, macrophages, B cells, some T cell subtypes and hepatocytes. Thus, only cells expressing membrane-bound IL-6R respond to IL-6 through the classical signaling pathway. However, the soluble form of IL-6R (sIL-6R) increases the variety of cells responsive to IL-6, in this case by trans-signaling[389]. IL-6 classical signaling mediates regenerative and anti-inflammatory processes, whereas trans-signaling is involved in pro-inflammatory processes[390,391]. Alterations in gp130 expression and polymorphisms in this gene have been related to a variety of diseases, including many types of cancer (e.g., myeloma, prostate, ovarian,

breast, lung and liver)[392], juvenile idiopathic arthritis[393], lung diseases[394], atherosclerosis[395] and myocardial infarction[396]. These findings have increased interest in blocking the IL-6/gp130 signaling pathway as a therapeutic strategy[397]. Moreover, inborn errors of *IL6ST* gene, such as two homozygous mutations referred as p.N404Y and p.P498L, have been reported to produce elevated IgE levels, skeletal abnormalities, eosinophilia and recurrent infections[398,399]. These infections were caused by *Staphylococcus aureus*, *Haemophilus influenzae*, *Pseudomonas aeruginosa* and *Escherichia coli*, among other bacteria[400]. Another report described a pediatric patient with a novel syndrome of neonatal immunodeficiency, autoinflammation and immune dysregulation, associated with a mutation (Tyr186_Tyr190del) in the *IL6ST* gene present as a constitutive mosaic[401].

Soluble forms of gp130 (sgp130) interact with the IL-6/sIL-6R complex acting as natural specific inhibitors of IL-6 trans-signaling[402] (**Figure 13A**), although they can also interfere with classical signaling at high concentrations[403]. Distinct sgp130 isoforms can be generated by alternative gp130 mRNA splicing or alternative polyadenylation and, to a lesser extent, metalloprotease activity[404] (**Figure 13B**). These isoforms include full length sgp130 (110 kDa)[405], sgp130-RAPS (rheumatoid arthritis antigen peptide-bearing soluble form) (50 kDa)[406], and sgp130-E10 (90 kDa)[407]. Full-length sgp130 is produced through incorporation of a novel exon 16 after the regular exon 15 of gp130. The smallest variant, sgp130-RAPS, arises from a deletion of exon 9, and an alternative polyadenylation site after exon 10 generates the sgp130-E10 isoform, which represents 1–2% of total sgp130[404]. Increased levels of sgp130 have been linked to pathologies, among which are childhood asthma[408], heart disease[409], diabetic retinopathy[410], multiple myeloma[411], prostate cancer[412], chronic lung disease of prematurity[413] and chronic liver disease[414]. Furthermore, the IL-6-neutralizing sIL-6R-sgp130 buffer system has been found to be disturbed in patients with type 2 diabetes[415], and some SNPs have been associated with sgp130 levels and are related to subclinical atherosclerosis[416]. Finally, the use of sgp130 as a therapeutic tool has been shown to be effective in animal models for hepatocellular carcinoma and arthritis[417,418] and has been proposed to prevent preterm birth[419].

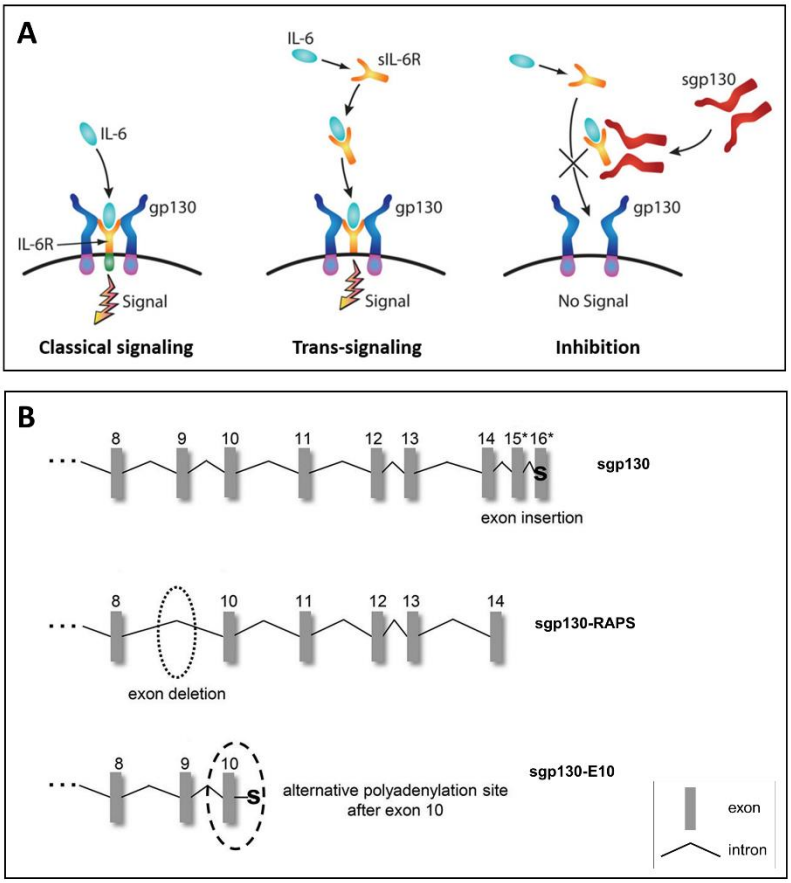


FIGURE 13. IL-6 signaling pathways and sgp130 isoforms.
(A) In IL-6 classical signaling, the cytokine IL-6 binds to a heterodimer receptor composed of IL-6R and gp130. Trans-signaling occurs when IL-6 binds to soluble IL-6 receptor (sIL-6R) and this complex associates with membrane gp130. Soluble gp130 (sgp130) can block the complex of IL-6 and sIL-6R, causing signal inhibition. **(B)** Differences in genetic structure between the three sgp130 isoforms. Adapted from (A) Lee, *et al*.[419] and (B) Wolf, *et al*.[404]

IL31RA

IL31RA is a protein-coding gene for the interleukin-31 receptor subunit alpha. This receptor, previously termed gp130-like receptor (GPL), forms a complex with the oncostatin M receptor beta (OSMRβ) and activates the signal transducer and activator of transcription factor 3 (STAT-3) and 5 (STAT-5), among others[420–422]. The cytokine IL-31 is produced by activated $CD4^+$ T cells and is involved in many physiological functions, including regulation of cell proliferation and hematopoiesis, induction of cytokines and regulation of the immune response. *IL31RA* gene is constitutively expressed in skin and reproductive tissues (testes and ovaries) and can be also detected in the brain, lungs, spleen, skeletal muscle, thymus, bone marrow and blood leukocytes[423,424]. Staphylococcal exotoxins act as potent inducers of *IL31RA* expression in monocytes and macrophages, whereas IFN-γ does so in dendritic cells. Subsequent IL-31 stimulation promotes the release of pro-inflammatory mediators by previously activated

monocytes, macrophages and dendritic cells[425,426]. Alterations in both *IL31RA* gene and protein, as well as in the IL-31 cytokine, have been related to several inflammatory skin diseases, such as atopic dermatitis[427], allergic contact dermatitis[428], primary cutaneous amyloidosis[429] and dermatomyositis[430]. More recently, IL31RA protein has been shown to promote breast cancer progression and metastasis[431].

DDX4

DDX4 gene encodes the DEAD (Asp-Glu-Ala-Asp) box helicase 4 protein, also called VASA. DEAD box proteins are RNA helicases associated with many molecular processes such as splicing, transcription, ribosome biogenesis, translation, RNA transport and RNA degradation[432]. *DDX4* is specifically expressed in fetal and adult germ cell linages, both in males and females (testes and ovaries), and is undetectable in somatic tissues. It plays a role in germ cell development and maintenance, mainly spermatogenesis[433,434]. In ovarian tissue, two subpopulations of small DDX4-positive cells with aldehyde dehydrogenase (ALDH1) activity have been characterized[435]. Levels of DDX4 are decreased in patients with asthenozoospermia, showing the function of this protein in spermatic motility[436]. In addition, it has been shown that DDX4 protein is overexpressed in ovarian cancer[437] and downregulated by vitamin D treatment, producing a suppressive effect on the proliferation and invasion of cancer cells[438].

SLC38A9

The solute carrier family 38 member 9 (*SLC38A9*) gene codes for the sodium-coupled neutral amino acid transporter 9 protein. SLC38A9 is a lysosomal transmembrane protein that is part of the lysosomal amino acid-sensing machinery; it functions as an amino acid transporter according to the substrate availability and can activate the mechanistic target of rapamycin complex 1 (mTORC1) signaling[439–441]. mTORC1 acts as a nutrient sensor and regulates metabolism and cell growth. It is involved in processes such as protein synthesis, autophagy and maintenance of homeostasis[442]. SLC38A9 is an arginine sensor for the mTORC1 pathway and mediates the lysosomal transport and metabolism of the essential amino acid leucine[443]. SLC38A9 also enables mTORC1 activation via lysosomal cholesterol[444]. Due to the role of mTORC1, dysregulation of this signaling pathway has been associated with metabolic syndrome and cancer[442]. Specifically, SLC38A9 has been reported to be required for the growth of pancreatic tumors[443].

1.6. Proteomics studies in MS

For decades, most association studies for diseases, including MS, have focused on genomics and transcriptomics techniques. However, levels of mRNA do not always correlate with protein quantification. Therefore, proteomics studies are necessary to reveal more information about cellular functions. Integration of proteomics data with gene expression signatures allow for a better understanding of the mechanisms underlying the pathogenesis of diseases and discovery of therapeutic targets[445]. In the recent years, significant advances have been made in proteomics techniques, such as mass spectrometry and chromatography, which have led to greater sensitivity, resolution and accuracy. These approaches have the advantage of being able to identify post-translational modifications, such as phosphorylation, acetylation and glycosylation. These post-translational modifications also play an important functional role in the development of the diseases. Bioinformatic tools and databases have also been beneficial to the development of the field of proteomics[446].

In MS, analysis of protein–protein interaction networks, or interactomes, represents a powerful tool to identify causative factors of disease, such as new genetic regions that are definitively associated and genes that mediate the GWAS findings[447,448]. Cell type-specific interactome analyses and other proteomics-based approaches are also effective at identifying biomarkers for early diagnosis of MS[449,450]. Another application of these techniques is to determine the proteomic profile of immune cell subsets, which allows protein expression to be correlated with MS-associated SNPs for the identification of protein quantitative trait loci (pQTLs)[451]. More recently, single cell proteomic analyses, based on novel high-resolution mass cytometry techniques such as the CyTOF, are helping to characterize new cell populations involved in MS, including microglial subtypes[452] and brain myeloid cells in EAE models[453], providing more insight into cell heterogeneity. Interactome analyses are also used to identify common pathways between different autoimmune disorders, such as type 1 diabetes and MS, thus identifying shared biological signatures[454].

2. General aims and structure of the thesis

2. General aims and structure of the thesis

The main goal of this work is to deepen the knowledge of the MS risk factor ANKRD55 at the levels of gene expression and protein interaction. For this purpose, this thesis is divided into two different but clearly related parts, each of which includes its own materials and methods, results and discussion sections.

The first part is based on gene expression of *ANKRD55* and its neighboring genes in different immune cell subpopulations relevant to the development of the disease. The subpopulations of primary focus are $CD4^+$ T lymphocytes and monocyte-derived dendritic cells, the latter of which was studied in different stages of maturation and under the effect of several tolerogenic stimuli. Based on previous results showing the significant association between the genotype of MS risk SNP rs6859219 and *ANKRD55* transcript levels in $CD4^+$ T cells[317], the main aim of this first part is to elucidate the influence of three different SNPs of interest (rs6859219, rs7731626 and rs13186299) on *ANKRD55* and *IL6ST* expression in monocyte-derived dendritic cells from healthy donors and MS patients.

The second part presents the results of interactome analyses of ANKRD55 protein across different cell types. Expanding upon a previous work where the ANKRD55 interactome of HEK293 cells was studied[377], the study presented here uses cell lines of the microglia, astrocytes, monocytes and neuroblastoma types. The objectives of this second part are: 1) to identify the proteins that interact with ANKRD55 in these four different cell lines; 2) to perform bioinformatic analyses to identify molecular and cellular pathways related to said proteins; and 3) to experimentally demonstrate the physical interactions between ANKRD55 and some of the proteins detected in the interactomes.

3. Part I.
Gene expression study

3. Part I. Gene expression study

3.1. Context and objectives

Among the SNPs contained in the *ANKRD55* gene, rs7731626 displays the strongest association with both MS and rheumatoid arthritis (**Table 1**). Other SNPs in *ANKRD55*, located like rs7731626 in intronic areas next to exon 6, have also been associated with the risk for autoimmune disease with genome-wide significance levels. These include members of a trio of SNPs highly correlated among them, namely rs6859219, rs10065637 and rs71624119 ($D' > 0.98$; $R^2 > 0.84$ in European populations), but more modestly to rs7731626 ($R^2 < 0.53$) (**Figure 14** and **Table 1**). The intronic location of all these SNPs points to allelic variation in regulatory elements as a potential mechanism for transcriptional modulation of *ANKRD55* and/or adjacent genes.

ANKRD55 is highly expressed in peripheral blood $CD4^+$ T lymphocytes, but not in $CD8^+$, $CD14^+$, $CD19^+$ or $CD56^+$ subsets. The risk allele (C) of rs6859219 is associated with higher levels of *ANKRD55* mRNA in the $CD4^+$ subset[317]. This SNP is also associated with DNA methylation status at CpG sites mapping to *ANKRD55-IL6ST* intergenic and *ANKRD55* intronic areas. Increased methylation at these sites seen with the protective allele (A) correlates with decreased expression of both *ANKRD55* and *IL6ST* mRNA in $CD4^+$ T cells[455]. The highly correlated rs71624119 variant was confirmed to colocalize with a *cis*-eQTL in $CD4^+$ T cells that significantly affected the expression of *ANKRD55*, and to a lesser extent *IL6ST*[359]. In a study using peripheral blood samples, both rs10065637 and rs6589219 colocalized with eQTLs for the expression of *ANKRD55* but not of *IL6ST*[456]. Similarly, the Genotype-Tissue Expression (GTEx) Project lists the MS and RA lead risk variant rs7731626 as the SNP with the most significant *cis*-eQTL effect for the expression of *ANKRD55* in spleen and whole blood (GTEx Analysis Release V8 https://www.gtexportal.org/home/gene/ANKRD55). However, rs7731626 colocalized with $CD4^+$ T-cell specific eQTLs for expression of both *ANKRD55* and *IL6ST*. This suggests that this locus has a pleiotropic mechanistic effect on the transcription of the two genes in this specific cell subset[457]. Risk for autoimmune disease conferred by rs7731626 is correlated with increased mRNA expression of *ANKRD55* and *IL6ST* and with reduced *cis*-CpG methylation in $CD4^+$ T cells[457,458]. Using naïve $CD4^+$ T-cell capture Hi-C data, a significant chromatin interaction between the intronic area around rs7731626 and the promoter of *IL6ST* was observed, potentially clarifying this coregulation in 3D space[334,455,457]. On the whole, current evidence explains genetic risk for MS and RA conferred by *ANKRD55* SNPs in terms of a *cis*-CpG

methylation-dependent mechanism to increase the expression of two genes, *ANKRD55* and *IL6ST*, in $CD4^+$ T lymphocytic cells.

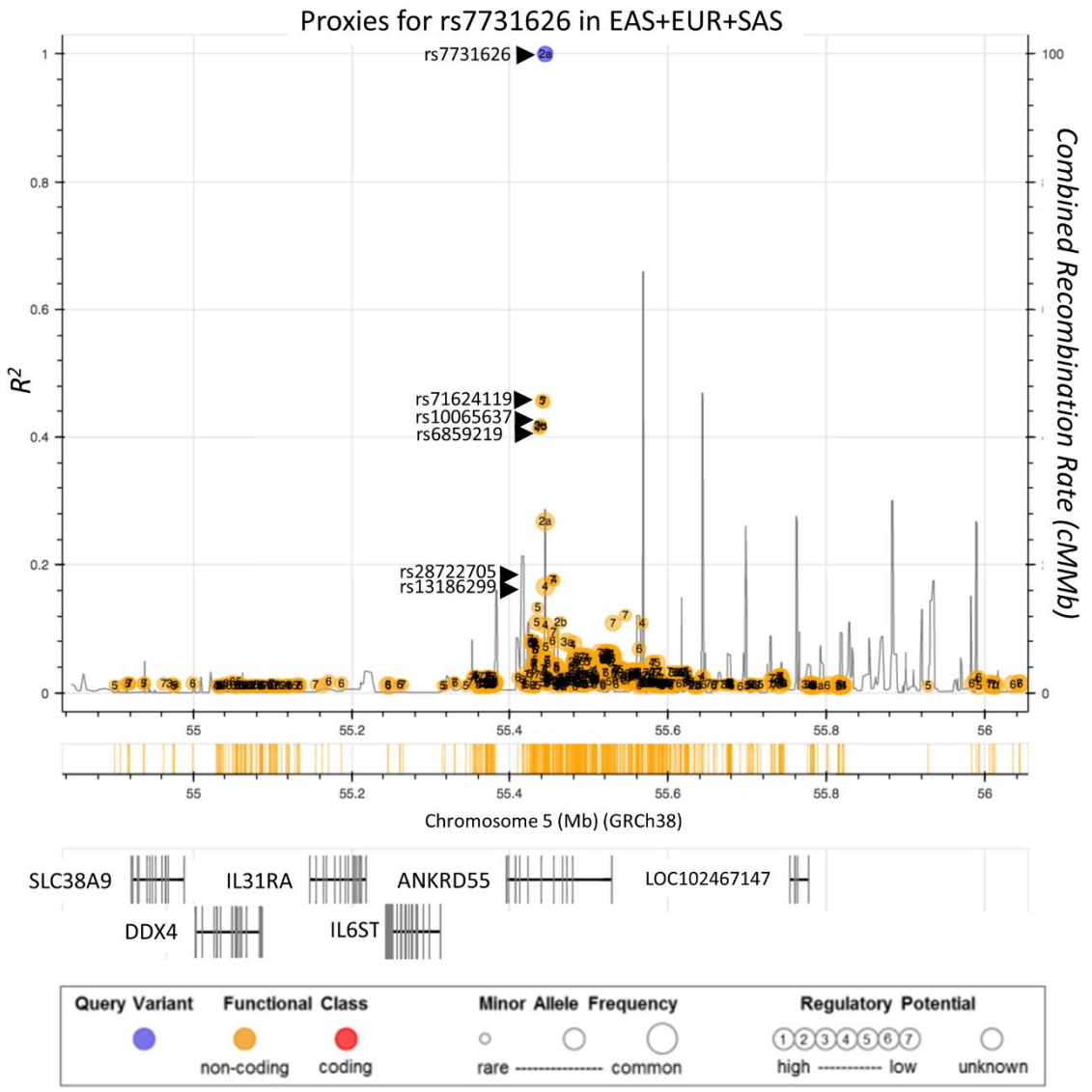


FIGURE 14. Autoimmune risk SNPs located in the *ANKRD55* gene area.
Plot representing individual linkage disequilibrium (LD) R^2 values of SNPs (ocher circles) located in the delineated genomic interval with the main autoimmune risk SNP rs7731626 (blue circle) in East Asian (EAS), European (EUR) and South Asian (SAS) populations. Indicated SNPs rs71624119, rs10065637 and rs6859219 are independently associated with autoimmune diseases with genome-wide significance, while SNPs rs28722705 and rs13186299 have been robustly associated with monocyte count[459,460] and monocyte percentage of white cells[461], respectively, in GWAS. Genes located in this interval are represented in the lower part of the figure and combined recombination rate is provided (black vertical lines). Image information was produced via LDlink (https://ldlink.nci.nih.gov).

IL6ST, an *ANKRD55*'s proximate neighbor (**Figure 9**), encodes the IL-6 receptor beta unit and is especially relevant due to its well-documented role in inflammation and autoimmune disease[382]. In MS, IL-6 blockage is thought to limit immune-mediated tissue damage. The humanized monoclonal antibody tocilizumab, a IL-6 receptor-blocking, has the ability to inhibit EAE[462]. On the other hand, as mentioned above, the precise biological function of ANKRD55 is currently unknown.

With the aim of deepening the knowledge of *ANKRD55* gene, in this study we determined its expression in moDCs treated with tolerogenic and maturation stimuli. Neighboring genes, including *IL6ST*, *IL31RA* and *SLC38A9*, were also analyzed. We additionally performed genetic analysis of the association between certain MS risk SNPs and gene expression levels in moDCs and $CD4^+$ T cells from healthy donors and MS patients.

3.2. Materials and methods

3.2.1. Patients and healthy controls

Buffy coats and serum samples from healthy donors and whole blood samples from MS patients were collected upon written informed consent in accordance with the local ethics committees in Bilbao (Comité Ético de Investigación Clínica de Euskadi [CEIC-E]), Barcelona (Comité Ético de Investigación Clínica Hospital Universitari Vall D'Hebron) and Madrid (Comité de Ética de la Investigación del Hospital Ramón y Cajal). All patients were receiving neither treatment nor placebo at the time of sample collection. A total of 128 biological samples of MS patients from three independent Spanish cohorts (IIS Biocruces Bizkaia MS Cohort, Barakaldo; Hospital Universitari Vall d'Hebron MS Cohort, Barcelona; and Hospital Universitario Ramón y Cajal MS Cohort, Madrid) were included in the study (54 primary progressive [PP], 70 relapsing remitting [RR], and 4 secondary progressive [SP]) in addition to 103 buffy coats from healthy control samples provided by the Biobanco Vasco. Patients' demographic and clinical data are summarized in **Table 2**.

TABLE 2. Clinical data of MS patients.

		N	%women/%men	Age (years)[a]	Duration of disease (years)[a]	EDSS[b]
Cohort 1[c]	PP	16	56.25/43.75	54.75 (± 10.98)	12.49 (± 6.80)	5.25 (4.00-6.50)
	RR	19	63.16/36.84	35.63 (± 8.85)	3.21 (± 4.62)	0.00 (0.00-2.00)
	SP	-	-	-	-	-
Cohort 2[d]	PP	28	39.29/60.71	51.50 (± 9.10)	12.59 (± 10.32)	5.50 (3.50-6.00)
	RR	11	81.82/18.18	39.00 (± 8.66)	11.18 (± 9.86)	2.50 (1.50-4.00)
	SP	4	75.00/25.00	44.50 (± 4.20)	10.67 (± 10.18)	3.50 (3.00-4.38)
Cohort 3[e]	PP	10	50.00/50.00	56.70 (± 9.55)	15.38 (± 12.58)	6.00 (2.38-6.63)
	RR	40	70.00/30.00	38.08 (± 11.48)	4.46 (± 5.67)	2.00 (1.50-2.00
	SP	-	-	-	-	-

[a]Data are expressed as mean (± standard deviation).
[b]Expanded Disability Status Scale (EDSS) data are expressed as median (Quartile1-Quartile3).
[c]Cohort 1 = IIS Biocruces Bizkaia MS Cohort, Barakaldo.
[d]Cohort 2 = Hospital Universitari Vall d'Hebron MS Cohort, Barcelona.
[e]Cohort 3 = Hospital Universitario Ramón y Cajal MS Cohort, Madrid.

3.2.2. Peripheral blood mononuclear cell subpopulations

PBMCs were isolated from buffy coats using a Ficoll®-Paque density gradient (Cytiva, Cat. No. 17-1440-03) as indicated by the manufacturer. Fresh PBMCs were used for all experiments on monocytes and moDCs. Comparative analysis of gene expression by qPCR was

performed in $CD4^+$, $CD8^+$, $CD14^+$, $CD19^+$ and $CD56^+$ cells purified from PBMCs that had been stored in liquid nitrogen. $CD4^+$ T-helper lymphocytes, $CD8^+$ T cytotoxic lymphocytes, $CD19^+$ B lymphocytes, $CD14^+$ monocytes and $CD56^+$ NK cells were separated by positive selection using CD4 (Cat. No. 130-045-101), CD8 (Cat. No. 130-045-201), CD19 (Cat. No. 130-050-301), CD14 (Cat. No. 130-050-201) and CD56 (Cat. No. 130-050-401) human MicroBeads (all from Miltenyi Biotec) correspondingly, following the manufacturer's instructions. For optimal separation, PBMCs from each donor were divided into two parts. One part was used for $CD8^+$ T cytotoxic lymphocyte purification, followed by $CD56^+$ NK selection from the previous negative fraction. The second part of the sample was used for isolation of $CD19^+$ B lymphocytes, and $CD14^+$ monocytes were recovered from the corresponding negative fraction. Finally, $CD4^+$ T-helper lymphocytes were obtained from the CD14 negative fraction.

3.2.3. Isolation of conventional dendritic cells and plasmacytoid dendritic cells

Human natural circulating conventional DCs (cDCs) and plasmacytoid DCs (pDCs) were isolated from healthy donors' fresh blood. PBMCs were obtained from 30 ml of blood by Ficoll®-Paque density gradient using SepMate™-50 tubes (STEMCELL Technologies, Cat. No. 85450). $CD14^+$ monocytes were depleted from PBMCs using CD14 human MicroBeads (Miltenyi Biotec, Cat. No. 130-050-201) and remaining PBMCs were separated into two equal parts and processed using an EasySep™ Human Myeloid DC Enrichment Kit (STEMCELL Technologies, Cat. No. 19061) for cDC isolation or EasySep™ Human Plasmacytoid DC Enrichment Kit (STEMCELL Technologies, Cat. No. 19062) for pDC isolation following the manufacturer's protocol. Both kits are based on immunomagnetic negative selection by targeting non-cDCs or non-pDCs with specific antibodies bound to magnetic particles. Unwanted cells were then separated using a magnet.

To verify the purity of each subpopulation, cells were analyzed by flow cytometry (MACSQuant® flow cytometer, Miltenyi Biotec) using antihuman CD1c (BCDA-1)-APC antibody (Miltenyi Biotec, Cat. No. 130-110-595) for cDCs and CD303 (BDCA-2)-APC antibody (Miltenyi Biotec, Cat. No. 130-114-177) for pDCs. Isotype control human IgG1-PE antibody (Miltenyi Biotec, Cat. No. 130-113-428) was used as negative control for CD14 and isotype control human IgG1-APC as negative control for CD1c and CD303 (Miltenyi Biotec, Cat. No. 130-113-434).

3.2.4. Sorting of classical, intermediate, and non-classical monocytes

Monocytes were obtained from human PBMCs with a Pan Monocyte Isolation kit (Miltenyi Biotec, Cat. No. 130-096-537) following the manufacturer's protocol, labeled with antihuman CD14-PE antibody (Miltenyi Biotec, Cat. No. 130-110-519) and CD16-FITC (Miltenyi Biotec, Cat. No. 130-113-392) and sorted in a FACSJazz™ cell sorter (BD Biosciences). The sorter strategy consisted of a size-complexity selection of monocytes followed by discrimination of three monocyte subpopulations: classical ($CD14^{++}/CD16^{-}$), intermediate ($CD14^{++}/CD16^{+}$) and non-classical ($CD14^{+}/CD16^{++}$), as described in a previous article[463].

3.2.5. Monocyte-derived dendritic cell differentiation and maturation

$CD14^{+}$ monocytes were isolated from fresh PBMCs by positive selection using CD14 MACS MicroBeads (Miltenyi Biotec, Cat. No. 130-050-201) and cultured at a density of 10^6 cells/ml in moDC differentiation medium (containing IL-4/GM-CSF; Miltenyi Biotec, Cat. No. 130-094-812) for 6 days, with fresh medium added on day 3. Maturation of moDCs was induced using combinations of different agents, including 0.2–2 µg/ml LPS (Sigma-Aldrich, Cat. No. L6143), 0.5 pg/ml IFN-γ (PeproTech, Cat. No. AF-300-02), 1 mg/ml polyinosinic-polycytidylic acid (PolyI:C; Sigma-Aldrich, Cat. No. P1530) and 1 mg/ml oligodeoxyribonucleotide CpG (Miltenyi Biotec, Cat. No. 130-100-243). To induce tolerogenicity, DCs were treated during the differentiation process with 100 nM retinoic acid receptor alpha agonist AM580 (Sigma-Aldrich, Cat. No. A8843), 10 nM vitamin D3 (STEMCELL Technologies, Cat. No. 72412), 1 mM prostaglandin E2 (PGE2; Sigma-Aldrich, Cat. No. P0409), 20 ng/ml IL-10 (PeproTech, Cat. No. AF-200-10), 1 mM dexamethasone (Sigma-Aldrich, Cat. No. D4902) and 10 ng/ml rapamycin (STEMCELL Technologies, Cat. No. 73362). All were added on day 0 and again on day 3 of the moDC differentiation period. To evaluate maturation, cells were incubated with Mo-DC Differentiation Inspector antibody cocktail (Miltenyi Biotec, Cat. No. 130-093-567; antihuman CD14-FITC antibody [clone Tük4, isotype: mouse IgG2a], anti-human CD83-APC [clone: HB15, isotype: mouse IgG1], and anti-human CD209-PE [clone DCN-47.5.4, isotype: mouse IgG1]) or isotype control cocktail following manufacturer's instructions and analyzed on a MACSQuant® flow cytometer (Miltenyi Biotec).

3.2.6. DNA extraction, RNA isolation and cDNA synthesis

Genomic DNA (gDNA) was obtained from PBMCs with a PureLink™ Genomic DNA purification kit (Invitrogen, Cat. No. K182001) and used for genotyping analyses. Total RNA was extracted from cells using TRI Reagent® (Sigma-Aldrich, Cat. No. T9424), NucleoSpin RNA Kit (Macherey-Nagel, Cat. No. 740955) or RNeasy Mini Kit (Qiagen, Cat. No. 74104) following each manufacturer's instructions. After extraction, samples were treated with DNase (Sigma-Aldrich Cat. No. AMPD1 or Qiagen Cat. No. 79254) to remove DNA and measured using a Nanodrop® 2000 spectrophotometer (Thermo Fisher Scientific) to determine the RNA concentration. cDNA was synthesized from 100 to 600 ng of total RNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Cat. No. 4368814) and the following cycling conditions: priming at 25 ºC for 10 minutes, retrotranscription at 37 ºC for 120 minutes and inactivation at 85 ºC for 5 minutes.

3.2.7. Primer design, qPCR, droplet digital PCR and genotyping

Quantitative polymerase chain reaction (qPCR) was performed using 5–30 ng cDNA, specific primer pairs (**Table 3**) and Fast SYBR® Green Master Mix (Applied Biosystems, Cat. No. 4385612) or SsoAdvanced™Universal SYBR® Green Supermix (Bio-Rad, Cat. No. 1725272), following manufacturer's instructions, in a 7300, 7500 Fast or 7900HT Fast Real-Time PCR System (Applied Biosystems). *IL6ST* primers are best-coverage primers that co-amplify the main splice variants of *IL6ST* by annealing to shared transcribed sequences (isoforms 206, 201, 205, 209, 204, 212, and 203). *ACTB* and/or *GAPDH* were used as reference genes to normalize mRNA expression levels, and data were analyzed using the $2^{-\Delta Ct}$ method. Samples were examined in triplicate for each condition, and no-template controls were included. The threshold was set within the linear phase of the logarithmic amplification plot.

Absolute gene transcript quantification was performed by droplet digital PCR (ddPCR) in a QX200 Droplet Digital PCR System (Bio-Rad). 350 pg or 30 ng of cDNA were mixed with ddPCR Supermix for Probes (no dUTP) (Bio-Rad, Cat. No. 1863023) and specific primer pairs (**Table 4**) according to manufacturer's instructions. 20 µl of this reaction mixture and 70 µl of Droplet Generation Oil for Probes (Bio-Rad, Cat. No. 186-3005) were used to generate droplets in a QX200 Droplet Generator (Bio-Rad). 40 µl of droplets were transferred to a 96-well plate which was heat-sealed with foil in a PX1 PCR Plate Sealer (Bio-Rad), and a PCR reaction was performed in a C1000 Touch Thermal Cycler (Bio-Rad). Cycling conditions consisted of a first enzyme

activation step at 95 ºC for 10 minutes followed by 40 cycles of denaturation at 94 ºC for 30 seconds and annealing/extension at 60 ºC for 1 minute and a final step of enzyme deactivation at 98 ºC for 10 minutes, all with 2 ºC/second ramp rate. The PCR plate was cooled at 4 ºC for 30 minutes before droplet fluorescence was measured in a QX200 Droplet Reader (Bio-Rad). Data were analyzed using QuantaSoft™ software (Bio-Rad).

gDNA Samples were genotyped for SNPs rs6859219, rs7731626 and rs13186299 using specific commercial TaqMan® probes (Thermo Fisher Scientific, Cat. No. 4351379) and TaqPath™ ProAmp™ master mix (Applied Biosystems, Cat. No. 30865) in a 7500 Fast Real-Time PCR System (Applied Biosystems). 20 ng of gDNA were used in a total volume reaction of 25 µl with the following cycling conditions according to manufacturer's instructions: a pre-read step at 60 ºC for 30 seconds, initial denaturation and enzyme activation at 95 ºC for 5 minutes, 40 cycles of denaturation at 95 ºC for 15 seconds and annealing/extension at 60 ºC for 1 minute and a final post-read step at 60 ºC for 30 seconds.

TABLE 3. Primers used in the study for qPCR (corresponding to Part I).

Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Company or Biblio. Refer.
ACTB_1	Cat. No. Hs_ACTB_1_SG		Qiagen
ACTB_2	Cat. No. Hs_ACTB_2_SG		Qiagen
GAPDH	GCAACAATATCCACTTTACCAGAG	CACATCGCTCAGACACCAT	Own design
ANKRD55 (best coverage)	Cat. No. Hs.PT.58.27501603		IDT
ANKRD55_201	CAGCCTCAACACACACAAATGC	TAGTTGATTATGGACGGCCCCTG	[317]
ANKRD55_202	CGGGCTCATTTAACACTTACTATTTC	CTTAGCCAGCAACAGCTCCTG	[317]
ANKRD55_204	TCCACTATGCTCGGCTGC	CTCGCTGATGTTCGACTGTTG	[317]
IL6ST	CCGCCACATAATTTATCAGT	AAGGTCTTGGACAGTGAATG	[464]
IL31RA	CTCTGCGATGTGCGGTCAA	GCCAAGTGTTTTCTCTAGGACTG	Own design
SLC38A9	TCCAGCTCCAGAAGAGTGCT	CTGTTTTATGCCCCAAGGAA	Own design
DDX4	TCATACTTGCAGGACGAGATTTG	AACGACTGGCAGTTATTCCATC	Own design
sgp130	GCAGCATACACAGATGAAGGTG	TAAGCTGTAAGGTCCTCGTTGG	[404]
sgp130-RAPS	TCCACCCGATCTTCATTCACTG	AGGAGGCAATGTTATCTTCATAGG*	[404]
sgp130-E10	TCCATCCCATACTCAAGGCTAC	TCACAGATACAAACCTTGAAAGTCAC	[404]
ALDH1A2	GAGGAGTTTGTGAGAAGAAGCGT	CTGTGGGCTCAATGAAAAACC	[465]
RARA	Cat. No. Hs.PT.58.442330		IDT
RARG	Cat. No. Hs.PT.58.38473235		IDT

*Reverse primer: first and last base from the original sequence were deleted to improve primer characteristics.

TABLE 4. Primers used in the study for ddPCR (corresponding to Part I).

Gene	Company
ANKRD55_201	Bio-Rad (Cat. No. dCNS104464519)
IL6ST	Bio-Rad (Cat. No. dHsaCPE5047372)
IL31RA	Bio-Rad (Cat. No. dHsaCPE5052967)
SLC38A9	Bio-Rad (Cat. No. dHsaCPE5057630)
IL3RA	Predesigned (Bio-Rad, Cat. No. dHsaCPE5058092)
CD1C	Predesigned (Bio-Rad, Cat. No. dHsaCPE5038928)
CLEC4C	Predesigned (Bio-Rad, Cat. No. dHsaCPE5048022)

3.2.8. Immunofluorescence of monocytes and moDCs

Human $CD14^+$ monocytes were plated on coverslips coated with 0.1 mg/ml poly-D-lysine (Sigma-Aldrich, Cat. No. P0899) and cultured in Roswell Park Memorial Institute (RPMI)-1640 medium (Sigma-Aldrich, Cat. No. R8758) or differentiated for 6 days into moDCs as described above. Cells were fixed and permeabilized with ice-cold methanol for 5 minutes, blocked with 1% bovine serum albumin (BSA; Sigma-Aldrich, Cat. No. A9418) in phosphate-buffered saline (PBS; Gibco, Cat. No. 18912014) for 30 minutes at room temperature, stained with defined primary antibodies for 1 hour at room temperature, incubated with corresponding fluorescent secondary antibodies for 1 hour at room temperature protected from light, stained with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich, Cat. No. D9542) for 5 minutes at room temperature and mounted on glass slides using Fluoromount-G (Invitrogen, Cat. No. 00-4958-02). Between each step, coverslips were washed 3 times with PBS. The primary and secondary antibodies used are specified in **Table 5** and **Table 6**.

For the visualization of fluorophores, the following laser beams and emission ranges were selected using a Zeiss LSM 880 Airyscan microscope: DAPI (Excitation: 405 nm; Emission: 410–486 nm), AlexaFluor 488 (Excitation: 488 nm; Emission: 495–584 nm) and AlexaFluor 594 (Excitation: 561 nm; Emission: 585–733 nm). Excitation and emission light passed through an oil-immersion Plan-Apochromat 63x objective (Numerical Aperture: 1.4), and square images of either 1024 × 1024, 1808 × 1808 or 2048 × 2048 pixels were acquired, achieving a final resolution of 0.131, 0.074 or 0.065 microns/pixel, respectively. Images were acquired at 16-bit grayscale resolution, exposure time was 2.3 µs/pixel and laser power and digital gain were adjusted to

avoid pixel saturation. The volume of nuclei was scanned by sequential imaging with a Z-interval of 0.7 microns, generating Z-stacks of 8–12 planes. Pictures were analyzed with Fiji software[466] (https://imagej.net/software/fiji).

For quantification of ANKRD55 expression in the nuclei of immature, mature and AM580-treated moDCs, a Fiji custom-made macro was created to ensure reproducibility. Briefly, this macro delineated the contours of DAPI stained nuclei in the maximal projection of the stack acquired, transferred them to the Alexa594 stained image and averaged the intensity of the pixels in each region. Nuclei with incomplete volume acquired into the stack were discarded. Individual ANKRD55 values were averaged across nuclei in a single stack and across stacks to provide a final value in each experimental condition analyzed. Nuclear ANKRD55 and DAPI signals were quantified in the immunofluorescent images for monocyte and moDC comparison following the method described by Shihan and colleagues[467]. Different regions of interest (ROIs) were selected in the corresponding DAPI channel, and the Otsu method was used for automatic thresholding. Then, the selection pattern of ROIs was applied to the corresponding ANKRD55 channel. Mean DAPI and ANKRD55 fluorescence intensities were calculated for each ROI.

TABLE 5. Primary antibodies used for immunofluorescence (corresponding to Part I).

Primary antibody	Host	Company	Dilution (concentration)
α-ANKRD55	Rabbit (polyclonal)	Human Protein Atlas (Cat. No. HPA 061649)	1:250
α-ALYREF	Mouse (monoclonal)	Thermo Fisher Scientific (Cat. No. MA1-26754)	1:500 (4 µg/ml)
α-HNRNPC	Mouse (monoclonal)	Sigma-Aldrich (Cat. No. AMAB91010)	1:100 (10 µg/ml)
α-DC-SIGN/CD209	Mouse (monoclonal)	R&D Systems (Cat. No. MAB161-SP)	1:36 (10 µg/ml)

TABLE 6. Secondary antibodies used for immunofluorescence (corresponding to Part I).

Secondary antibody	Host	Company	Dilution (concentration)
AlexaFluor 488-conjugated anti-mouse IgG H&L	Goat (polyclonal)	Abcam (Cat. No. ab150113)	1:1000 (2 µg/ml)
AlexaFluor 594-conjugated anti-rabbit F(ab')2	Goat (polyclonal)	Thermo Fisher Scientific (Cat. No. A-11072)	1:500 (4 µg/ml)

3.2.9. ELISA

Levels of gp130 were measured in serum samples of MS patients using commercial enzyme-linked immunosorbent assay (ELISA; R&D Systems, Cat. No. DY228). 96-well microplates were coated with 4 µg/ml mouse anti-human gp130 capture antibody (R&D System, Cat. No. 840928) at room temperature overnight. Wells were washed three times with wash buffer (0.05% Tween®-20 in PBS, pH 7.2–7.4; R&D System, Cat. No. WA126) and plates were blocked with reagent diluent (1% BSA in PBS, pH 7.2–7.4; R&D Systems, Cat. No. DY995) for 1 hour at room temperature. After three more washes, 100 µl of sample dilutions (1:100 and 1:200) or human gp130 standards (R&D Systems, Cat. No. 840930) diluted in reagent diluent (R&D Systems, Cat. No. DY995) were added in duplicate to wells for 2 hours at room temperature and washed again. Biotinylated goat anti-human gp130 detection antibody (R&D Systems, Cat. No. 840929) diluted in reagent diluent with normal goat serum (R&D Systems, Cat. No. DY005) was added to each well in a concentration of 400 ng/ml for 2 hours at room temperature, followed by three washes. Wells were incubated with streptavidin-HRP A (R&D Systems, Cat. No. 890803) for 20 minutes and with substrate solution (1:1 H_2O_2 and tetramethylbenzidine; R&D Systems, Cat. No. DY999) for another 20 minutes at room temperature in darkness. Finally, the reaction was stopped by adding stop solution (2N H_2SO_4; R&D Systems, Cat. No. DY994) for 20 minutes, and absorbance was read at 450 nm in Varioskan Flash (Thermo Fisher Scientific). Data were analyzed using GraphPad Prism v.7 software.

3.2.10. Statistical analysis

Data were analyzed with GraphPad Prism v.7 software and are presented as the mean ± standard error of the mean (SEM), unless otherwise indicated. Outliers identified via ROUT test were removed, and a Shapiro–Wilk test was performed to determine whether the variables were normally distributed. A Mann–Whitney test, Wilcoxon signed-rank test, unpaired t-test or paired t-test was performed to analyze the statistical difference between two groups. A Kruskal–Wallis test, Friedman test (followed by Dunn's multiple comparison test) or ANOVA test was used to compare differences between three or more groups. Correlation analyses were determined using the Pearson correlation coefficient. Values of $p \leq 0.05$ were considered significant (*), $p \leq 0.01$ very significant (**), $p \leq 0.001$ highly significant (***), $p \leq 0.0001$ extremely significant (****) and $p > 0.05$ not significant (n.s.). For analysis of group-by-time interactions between the $2^{-\Delta Ct}$ gene expression values of *IL6ST*, *IL31RA* and *SLC38A9* and those

of *ANKRD55*, a longitudinal analysis was performed in moDC samples for four treatment conditions (immature control, immature AM580-treated, IFN-γ/LPS-matured control and IFN-γ/LPS-matured AM580-treated) using generalized linear mixed univariable models. Random intercepts and unstructured variance-covariance matrices were used.

3.3. Results

3.3.1. Gene expression of *ANKRD55*, *IL6ST*, *IL31RA*, *DDX4*, *SLC38A9* and soluble gp130 isoforms in PBMCs

High expression of *ANKRD55* in $CD4^+$ T cells has been previously reported, showing co-expression of three distinct *ANKRD55* mRNA splice variants: *ANKRD55-201* (Ensembl nomenclature; full-length isoform; protein encoded Mr = 68.4 kDa), *ANKRD55-202* (shorter protein-coding isoform; Mr = 36.9 kDa) and the non-coding processed transcript *ANKRD55-204*. These three splice variants were originally designated as isoforms 001, 005 and 007, respectively[317].

To identify cellular sources of *ANKRD55* flanking transcripts in PBMC, five subpopulations were isolated from healthy donors: $CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, $CD14^+$ monocytes, $CD19^+$ B lymphocytes and $CD56^+$ NK cells. mRNA levels of *ANKRD55* and neighboring genes, including the soluble gp130 isoforms sgp130, sgp130-RAPS and sgp130-E10, were assessed by qPCR analysis using specific primer pairs. Similar to *ANKRD55*, *IL6ST* was most expressed in $CD4^+$ T-helper lymphocytes, however it was also detected in the other four subpopulations at lower levels of expression. *IL31RA* was only detected in $CD14^+$ monocytes, and *DDX4* was not detected in any of the PBMC subpopulations. *SLC38A9* was highly expressed in $CD19^+$ B lymphocytes but also detected in the remaining subpopulations. The three soluble gp130 isoforms (sgp130, sgp130-RAPS and sgp130-E10) were highly expressed in $CD4^+$ T helper lymphocytes but less so in the other subpopulations studied, similar to the patterns observed for *IL6ST.* sgp130-RAPS showed lower expression compared to the other two soluble gp130 isoforms (**Figure 15**).

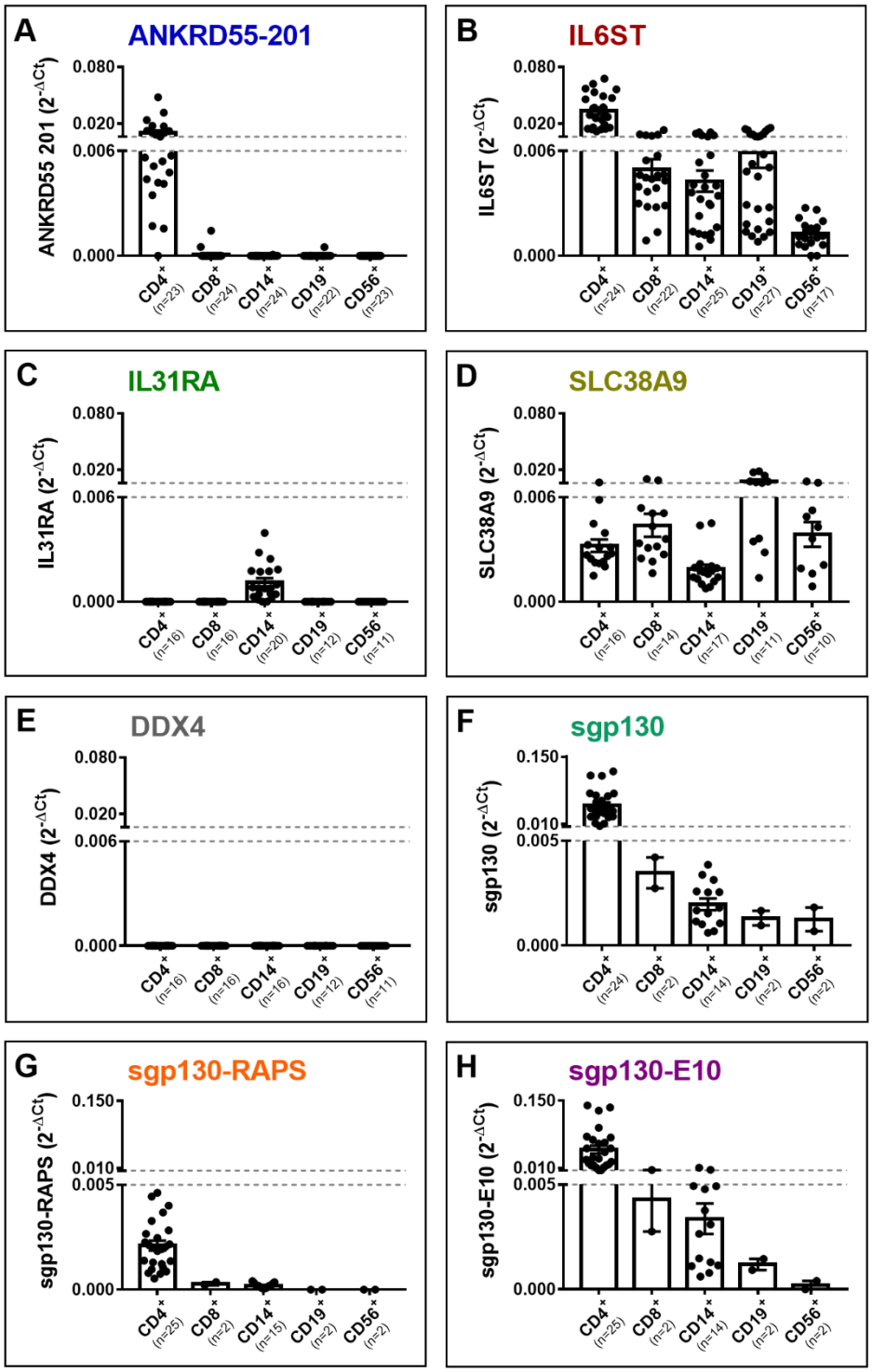


FIGURE 15. Gene expression in CD4⁺, CD8⁺, CD14⁺, CD19⁺ and CD56⁺ subpopulations of PBMCs from healthy controls.

(A) *ANKRD55* main isoform 201 (data from Lopez de Lapuente, *et al.*[317]), **(B)** *IL6ST*, **(C)** *IL31RA*, **(D)** *SLC38A9*, **(E)** *DDX4*, **(F)** sgp130, **(G)** sgp130-RAPS and **(H)** sgp130-E10 expression levels were measured using qPCR. *ACTB* and *GAPDH* were used as reference genes. Data are presented as mean ± SEM of $2^{-\Delta Ct}$ values, and the number (n) of healthy controls used in each subpopulation is indicated in brackets.

3.3.2. *ANKRD55* is expressed in immature moDCs and induced by retinoic acid receptor alpha agonist AM580

ANKRD55 gene expression was also studied in moDCs based on previous findings of the *IL22RA2* gene[468]. *IL22RA2* is located at an MS risk locus[274] and is expressed minimally or not at all in PBMC subsets. However, it has been confirmed to be highly expressed in immature moDCs but not in monocytes or mature moDCs[463,465]. This unique expression pattern may indicate a role for immature moDCs in MS pathogenesis and suggest that other MS risk genes are expressed during the monocyte-to-DC differentiation stage.

To investigate this hypothesis, purified $CD14^+$ monocytes from healthy donors were cultivated in moDC differentiation medium containing IL-4 and GM-CSF for 6 days to induce immature moDC formation. Proper differentiation was evaluated with the CD209 (DC-SIGN) marker by immunofluorescence microscopy, eliciting a positive staining in > 95% of cells. Cells were harvested for gene expression analysis by qPCR, and results showed expression of *ANKRD55*, *IL6ST*, *IL31RA* and *SLC38A9* in immature moDCs. Since expression of *DDX4* was not detected, this gene was not further studied. Moreover, the effects of a series of tolerogenic compounds known to modulate gene expression in moDCs[174] including AM580 (an analog of retinoic acid that acts as a selective retinoic acid receptor alpha [RAR-α] agonist), vitamin D3, IL-10, PGE2, dexamethasone and rapamycin were tested by being added to Mo-DC medium over the course of the differentiation period. Of these, AM580 appeared to enhance mRNA levels of both *ANKRD55* (3.5-fold) and *IL31RA* (379-fold), while the other compounds either inhibited the expression or had more modest effects on mRNA levels of any gene (**Figure 16**).

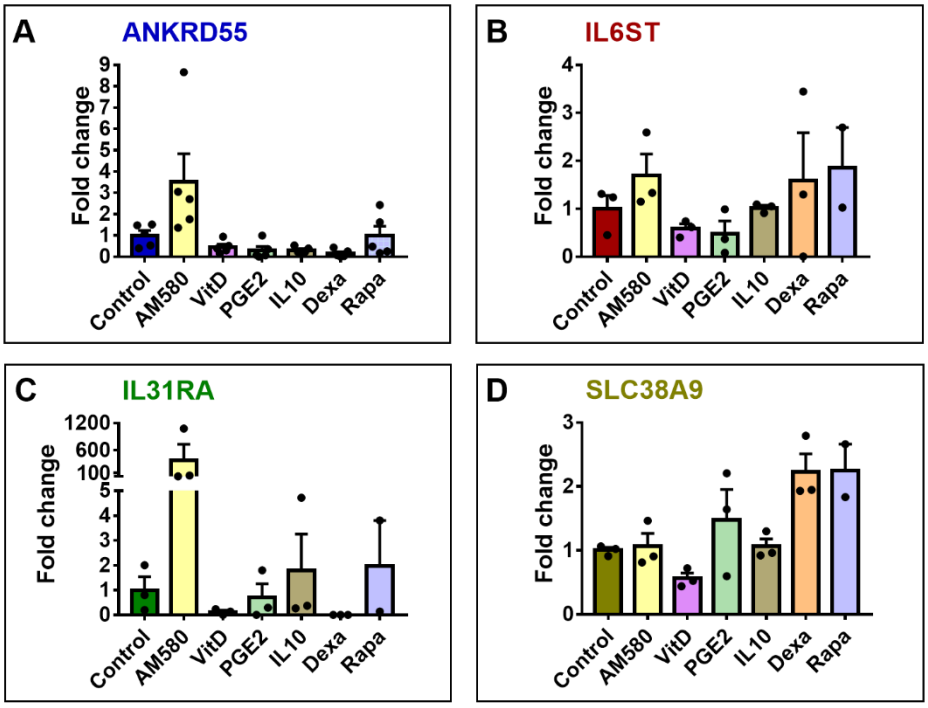


FIGURE 16. Expression of *ANKRD55* and neighboring genes in immature moDCs treated with tolerogenic compounds.
Monocytes were differentiated into moDCs with the addition of 100 nM AM580, 10 nM vitamin D3 (VitD), 1 mM PGE2, 20 ng/ml IL-10, 1 mM dexamethasone (Dexa), 10 ng/ml rapamycin (Rapa) or no treatment (control). Compounds were added on days 0 and 3. **(A)** *ANKRD55*, **(B)** *IL6ST*, **(C)** *IL31RA* and **(D)** *SLC38A9* gene expression was assessed by qPCR using *GAPDH* as a housekeeping gene. Mean ± SEM of fold change is shown. A maximum of 5 different donors were used for each condition. A Friedman test, followed by Dunn's multiple comparison test, was performed.

To further elucidate the changes in gene expression throughout the dendritic cell differentiation process, a qPCR time course analysis was performed. The *ANKRD55*, *IL6ST*, *IL31RA* and *SLC38A9* genes were assessed on days 0 (monocyte stage), 1, 2, 4 and 6 (**Figure 17**). *ANKRD55* gene expression started to rise from day 2 onward ($p < 0.01$ for comparison of levels on day 6 to those on day 0). In contrast, expression levels of *IL6ST* did not vary significantly, while those of *IL31RA* decreased from day 1 ($p < 0.01$ for comparison of levels at day 4 vs. 0). *SLC38A9* expression levels peaked at day 1 ($p < 0.01$).

At the same time, a time course of the effect of AM580 on each of these genes compared to untreated cells was also performed (gray curves in **Figure 17**). In the presence of AM580, moDCs produced slightly higher levels of *ANKRD55*, *IL6ST* and *SLC38A9* than untreated cells. However, none of the data points with AM580 showed significant differences compared to the corresponding untreated points in the time course for any of the genes studied. *ANKRD55* and *IL6ST* were significantly higher in the presence of AM580 on day 6 compared to day 0 ($p < 0.01$

and $p < 0.05$ correspondingly). *SLC38A9* was significantly higher on day 1 of AM580 treatment compared to day 0 ($p < 0.01$).

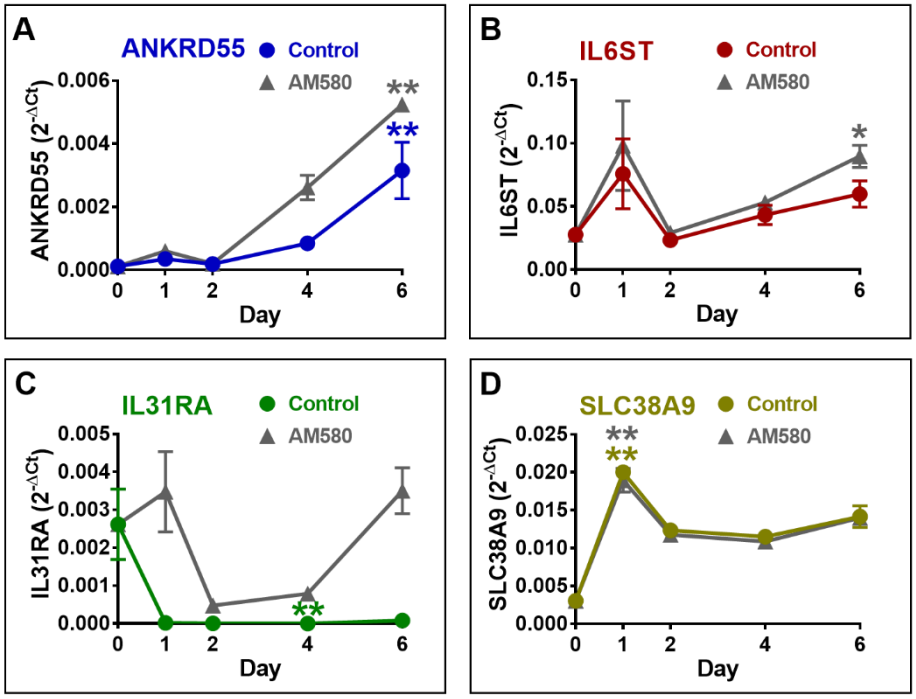


FIGURE 17. Time course analysis of *ANKRD55*, *IL6ST*, *IL31RA* and *SLC38A9* with and without AM580 treatment.

(A-D) Monocytes were cultivated for 6 days in IL-4/GM-CSF medium for differentiation into immature moDCs in the absence (colored curves) or presence (gray curves) of AM580, and expression levels of the indicated genes were measured by qPCR at the start of the cultivation (day 0) and on days 1, 2, 4, and 6. Mean ± SEM of $2^{-\Delta Ct}$ values from three independent measurements are shown and *GAPDH* was used as a housekeeping gene. Friedman test, followed by Dunn's multiple comparison test, was performed to compare data points in each curve with day 0, and Wilcoxon test was used to compare curves (control vs. AM580). * $p \leq 0.05$, ** $p \leq 0.01$.

Bioactivity of AM580 was ascertained through upregulation of the *ALDH1A2* gene in moDCs ($p < 0.01$ compared to untreated; **Figure 18A**), since retinoic acid is essential for GM-CSF-induced *ALDH1A2* expression in DCs[469]. Two retinoic acid receptors, RAR-α and RAR-γ[470], were also assessed through expression of their corresponding genes *RARA* and *RARG*. Both *RARA* and *RARG* were expressed in the moDCs, and the former was significantly affected ($p < 0.01$) only after treatment with IFN-γ/LPS maturation factors.

In addition, to further investigate the effect of AM580 on the increased *ANKRD55* expression in moDCs (**Figure 16A**), qPCR was performed using *ANKRD55* isoform-specific primer pairs, each amplifying a unique *ANKRD55* splice variant (primer sequences detailed in **Table 3**). This analysis showed that immature moDCs express the three splice variants and that AM580 significantly enhances the expression levels of all three. Fold upregulation of transcripts 201,

202 and 204 by AM580 amounted to 2.8 ($p < 0.01$ compared to untreated), 2.2 ($p < 0.05$) and 2.9 ($p < 0.01$), respectively (**Figure 18B**).

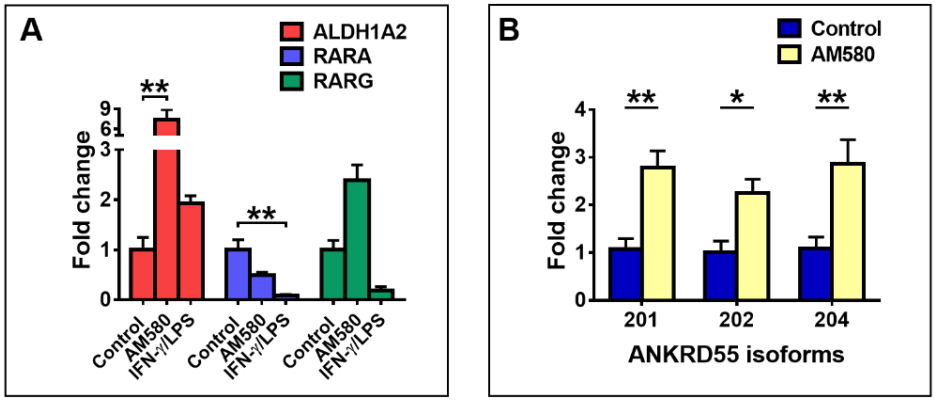


FIGURE 18. Effect of AM580 on *ANKRD55* isoforms in moDCs.
(A) Analysis of *ALDH1A2*, *RARA* and *RARG* expression via qPCR. Data are presented as mean ± SEM of fold change; $n = 6$. Expression was normalized to the housekeeping gene *ACTB*, and a Friedman test, followed by Dunn's multiple comparison test, was used. **(B)** Expression levels of three *ANKRD55* splice variants 201, 202 and 204 by qPCR in immature moDCs without (control) or with AM580. Mean ± SEM of fold change is shown; $n = 10$. Expression was normalized to *ACTB* and *GAPDH* genes, and a Wilcoxon test was performed. *$p \leq 0.05$, **$p \leq 0.01$.

We also assessed gene expression of *ANKRD55* isoforms *IL6ST* and *IL31RA* in different monocyte subtypes and in the corresponding immature DCs. Circulating monocytes can be distinguished based on relative expression levels of CD14 and CD16 surface proteins, namely classical ($CD14^{++}/CD16^{-}$), intermediate ($CD14^{+}/CD16^{+}$) and non-classical ($CD14^{+}/CD16^{++}$) subsets. These are thought to contribute differentially to immune surveillance of the CNS and inflammatory responses in MS[145,471]. Expression of indicated genes was measured in these three monocyte subsets using fluorescence-activated cell sorting (FACS) as well as in the corresponding immature DCs by cultivation of each subset in moDC differentiation medium for six days (**Figure 19**). Very low but detectable levels of *ANKRD55* isoforms 201 and 204, though not of 202, were detected in all three monocyte subsets. The highest expression level was found in the classical subset, in line with Schmiedel *et al*., who reported 0.4 transcripts per million (TPM) of *ANKRD55* in classical vs. 0.1 TPM in non-classical monocytes[472]. Differentiation of the monocyte subsets into immature DCs strongly enhanced expression levels of all three *ANKRD55* transcripts, with highest levels of *ANKRD55-201* (11.5-fold upregulation) and *ANKRD55-204* (15.2-fold upregulation) found in the classical monocyte-derived DC subset. *IL6ST* was expressed at much higher levels than *ANKRD55* in the three monocyte subsets but was less strongly increased in the moDCs from classical monocytes (1.4-fold upregulation). *IL31RA* was downregulated in each moDC subset.

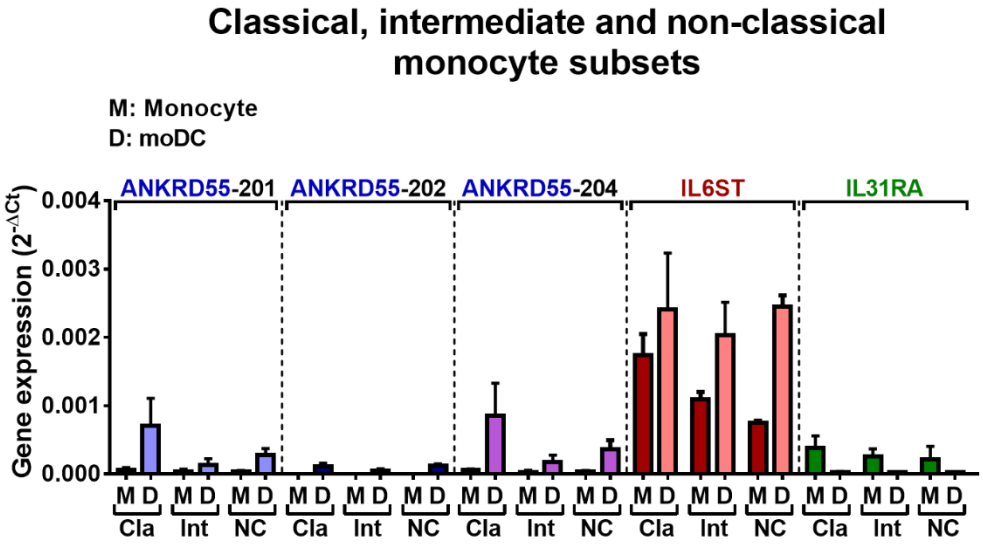


FIGURE 19. Gene expression in different monocyte subsets and the corresponding moDCs. Classical (Cla; $CD14^{++}/CD16^{-}$), intermediate (Int; $CD14^{+}/CD16^{+}$) and non-classical (NC; $CD14^{+}/CD16^{++}$) monocytes were isolated by FACS and separately cultivated for six days to differentiate into moDCs. Gene expression levels were measured in the original monocyte subsets (M) and derived moDCs (D) by qPCR normalized to *ACTB*. Mean ± SEM of $2^{-\Delta Ct}$ values is shown (n = 3), and a Mann–Whitney test was used to compare across monocyte and moDC genes and subsets.

3.3.3. *ANKRD55* expression is downregulated in IFN-γ/LPS-matured moDCs

We next aimed to study the effect of moDC maturation on the expression of *ANKRD55* and neighboring genes. To do so, we treated immature cells on day 6 of the differentiation period for 6 hours with Toll-like receptor (TLR) 9 agonist CpG or combinations of TLR4 agonist LPS and TLR3 agonist poly(I:C) or IFN-γ and LPS (**Figure 20**). Maturation of moDCs was assessed via expression of CD83- and CD209- (DC-SIGN) specific markers as measured by flow cytometry. The percentage of mature cells over four independent experiments was (average ± STDEV) 79.0 ± 8.9% with LPS + IFN-γ, 62.8 ± 12.4% with LPS + poly(I:C) and 23.75 ± 6.4% with CpG. Spontaneous maturation under the control condition amounted to 9.25 ± 5.1%.

Expression levels of the *ANKRD55* isoforms 201, 202 and 204 were measured in immature (control) and differently matured moDCs. The three isoforms were significantly downregulated in moDCs matured with IFN-γ/LPS ($p < 0.05$ vs. control; **Figure 20B**), and similar, though weaker, trends were observed for the other two conditions. In conclusion, the extent of *ANKRD55* suppression appeared proportional to the degree of moDC maturation.

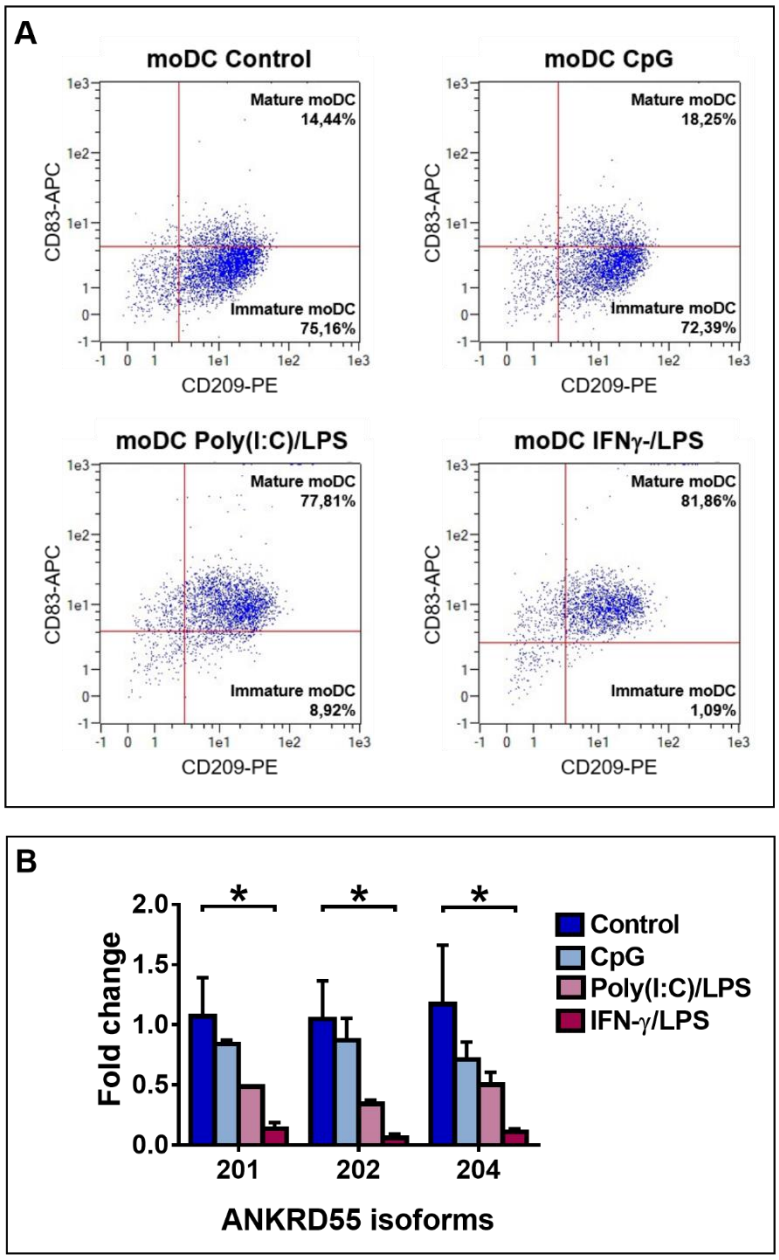


FIGURE 20. Evaluation of different moDC maturation conditions.
(A) Flow cytometry to assess the maturation percentage induced by the indicated stimuli based on the expression of CD209 (in moDCs) and CD83 (in mature moDCs only). One representative experiment out of the four performed is shown. **(B)** Comparison of the effect of maturation by CpG, poly(I:C)/LPS and IFN-γ/LPS on expression levels of the main *ANKRD55* splice variants, measured via qPCR. Gene expression was normalized to *ACTB* and *GAPDH*. Mean ± SEM of fold change is shown (n = 3), and data were analyzed using a Friedman test followed by Dunn's multiple comparison test. *$p \leq 0.05$.

A time course analysis was performed to determine the expression of *ANKRD55* and its neighboring genes following initiation of moDC maturation with IFN-γ/LPS in the presence or absence of AM580 (**Figure 21**). Maturation of moDCs induced rapid downregulation of *ANKRD55* and *SLC38A9* gene expression ($p < 0.05$ at 3 hours), with the lowest levels recorded from 3 to 12 hours following the addition of LPS. This effect was observed in both untreated and AM580-pretreated cells. IFN-γ/LPS had little effect on *IL6ST* gene expression but induced high *IL31RA* expression levels in the untreated moDCs ($p < 0.01$ at 12 hours and $p < 0.05$ at 24 hours) after LPS addition for comparison with levels at day 5. After 48 hours, the gene expression in untreated moDCs showed levels similar to pre-stimulation in the four genes analyzed, suggesting a disappearance of the maturation effect.

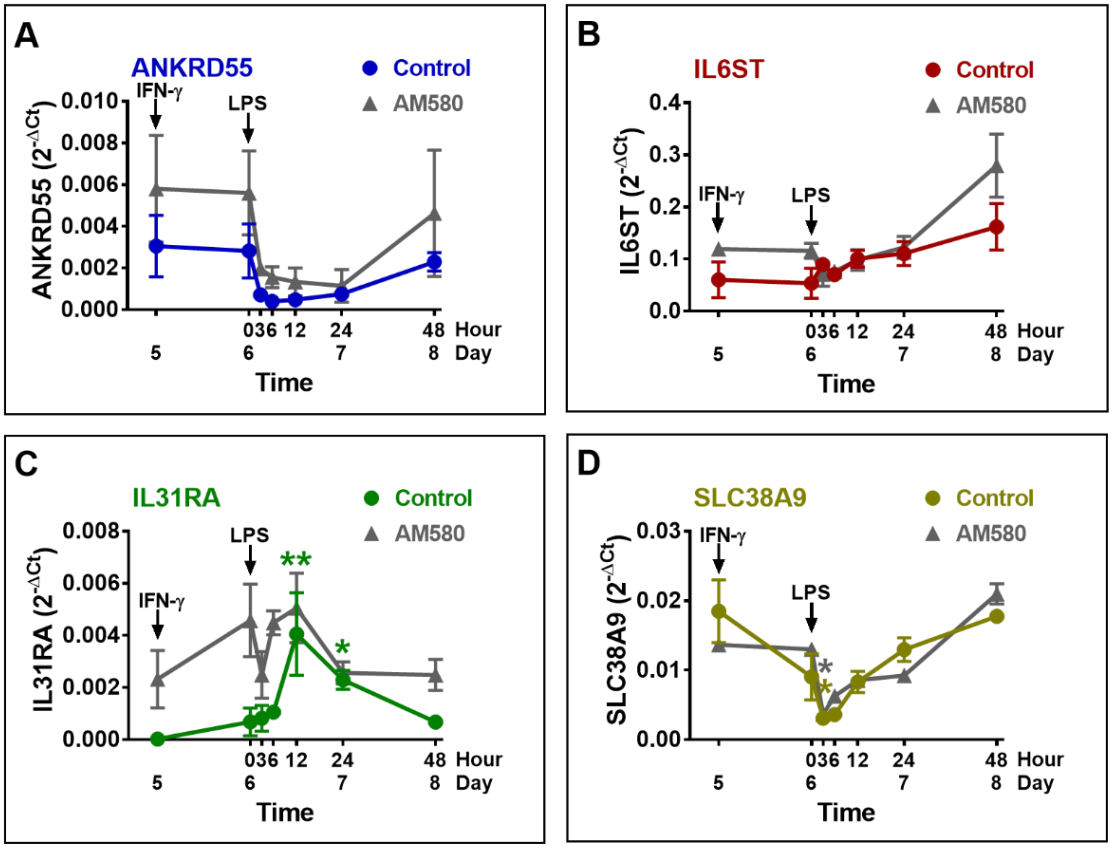


FIGURE 21. Effect of moDC maturation on the expression of *ANKRD55*, *IL6ST*, *IL31RA* and *SLC38A9*.

moDCs cultivated for 5 days in MoDC medium matured with the addition of IFN-γ and, 24 hours later, LPS. Gene expression of **(A)** *ANKRD55*, **(B)** *IL6ST*, **(C)** *IL31RA* and **(D)** *SLC38A9* was analyzed using qPCR coinciding with time of IFN-γ (day 5) and LPS (0 hours) addition and 3, 6, 12, 24 and 48 hours later. The experiment was also performed in cells differentiated in the presence of AM580 (gray curves). Mean ± SEM of the $2^{-\Delta Ct}$ values from three independent measurements is shown. *GAPDH* was used as a housekeeping gene. A Friedman test, followed by Dunn's multiple comparison test, was performed for comparison of data points in each curve with day 5, and a Wilcoxon test was performed to compare both curves (control vs. AM580). *$p \leq 0.05$, **$p \leq 0.01$.

In order to determine whether the curves representing *ANKRD55* and *IL6ST*, *IL31RA* and *SLC38A9* mRNA expression levels (**Figure 17** and **Figure 21**) are statistically dissimilar within each treatment condition, a longitudinal analysis was performed. The *p*-values corresponding to the group-by-time interaction term's significance in the adjusted generalized linear mixed univariable models are displayed in **Table 7**. *IL31RA* gene expression showed significant group-by-time interactions with *ANKRD55* in all tested conditions. However, group-by-time interactions were not statistically significant when comparing *ANKRD55* with *IL6ST* or *SLC38A9* in both control and AM580-administered groups of the immature sample and with *SLC39A9* in both control and AM580-administered groups of the mature sample. Therefore, in those cases, the similarity between the analyzed groups in terms of the trend among the evaluated times cannot be rejected. Subsequent genetic analysis was performed focusing on *ANKRD55* and *IL6ST* in the immature untreated moDCs.

TABLE 7. Corresponding *p*-values for the group-by-time interaction term's significance in the adjusted generalized linear mixed univariable models.

Compared evolutions	Immature		Mature	
	Control	AM580	Control	AM580
ANKRD55 vs. IL6ST	0.5772	0.2973	0.0016	0.0142
ANKRD55 vs. IL31RA	0.0001	0.0010	0.0304	0.0005
ANKRD55 vs. SLC38A9	0.7150	0.8674	0.4231	0.0972

3.3.4. Intracellular location of ANKRD55 in moDCs and colocalization with nuclear speckles

We used immunofluorescence (IF) microscopy to detect the intracellular location of ANKRD55. ANKRD55 was borderline visible in the nuclei of $CD14^+$ monocytes but stained clearly in the cytosol and nuclei of immature moDCs. Differentiation of moDCs was confirmed by co-staining with CD209, a membrane marker that is absent in monocytes but visible in moDCs (**Figures 22A, B**). The ratio of nuclear ANKRD55 to DAPI was significantly higher in moDCs than that in monocytes ($p < 0.001$), in accordance with results obtained by qPCR. We also compared immature, mature (IFN-γ/LPS-treated) and AM580-treated moDCs. In a quantitative analysis of IF images, IFN-γ/LPS treatment significantly reduced nuclear ANKRD55 levels ($p < 0.0001$), while AM580 had no effect (**Figure 22C**).

Given a nuclear pattern of irregular spots, we verified whether ANKRD55 could be associated with nuclear speckles. Nuclear speckles, also known as interchromatin granule clusters, are irregularly shaped, dynamic structures in the nucleus involved in RNA metabolism and gene expression[473,474]. As shown in **Figures 22D** and **E**, in moDCs, ANKRD55 consistently colocalized with ALYREF and HNRNPC, two confirmed constituents of nuclear speckles[474].

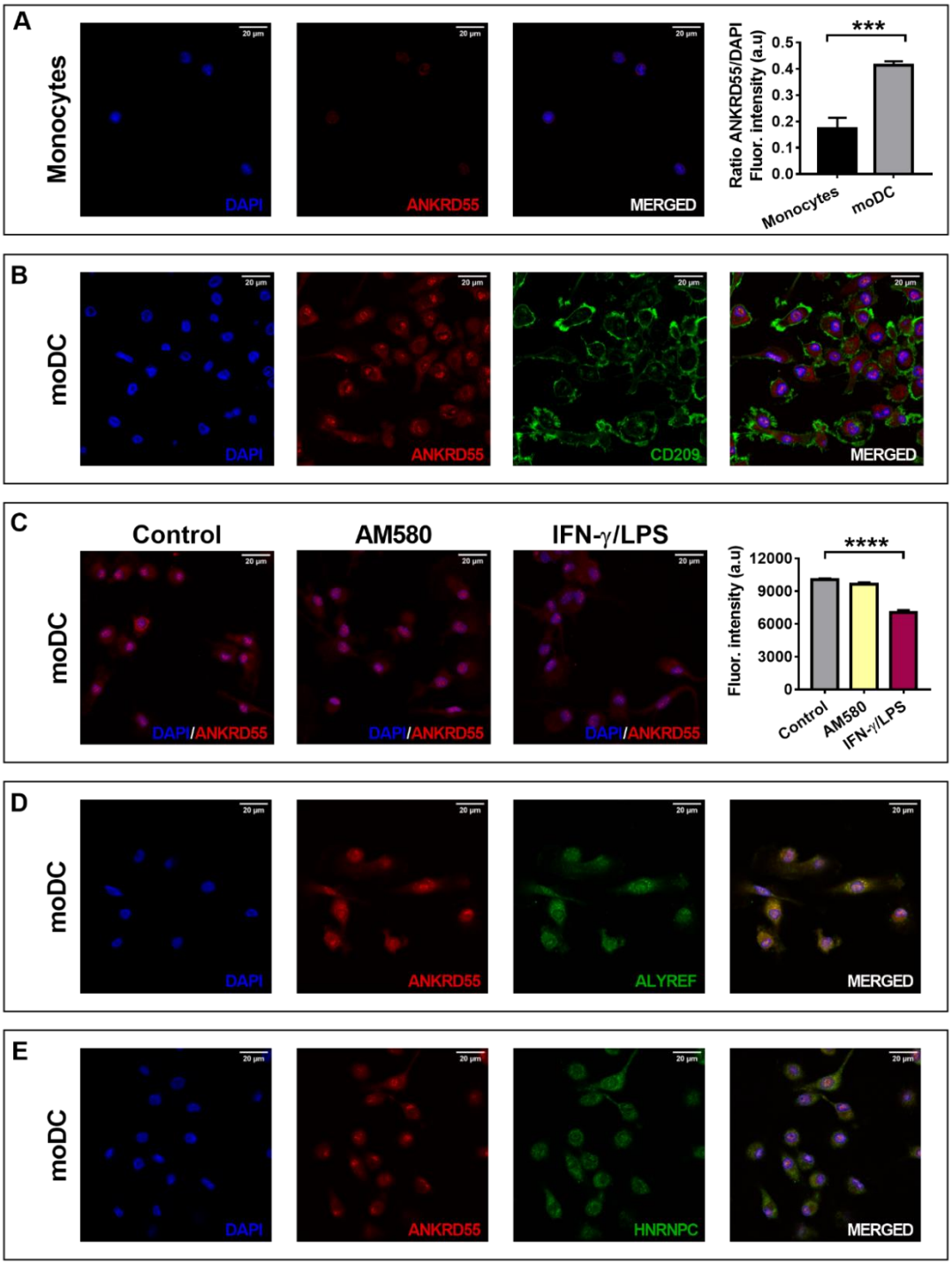


FIGURE 22. Localization of ANKRD55 in monocytes and moDCs and colocalization with nuclear speckles using immunofluorescence microscopy.

(A) Detection of ANKRD55 (red) and nuclei (DAPI; blue) in human $CD14^+$ monocytes via confocal microscopy (1024 × 1024 pixels = 0.131 microns/pixel). The graph on the right represents mean ± SEM of the ANKRD55/DAPI ratio fluorescence intensity in arbitrary units (a.u.). *n* = 5 cellular ROIs/group. An

unpaired t-test was used. ***$p \leq 0.001$. **(B)** Colocalization of ANKRD55 (red) with CD209 (green) in moDCs (2048 × 2048 pixels = 0.065 microns/pixel). **(C)** Effect of AM580 and IFN-γ/LPS treatment on ANKRD55 (red) in the nucleus (DAPI; blue) of moDCs (1808 × 1808 pixels = 0.074 microns/pixel). The diagram on the right provides a quantitative analysis of the nuclear ANKRD55 signal, presented as mean ± SEM of the fluorescence intensity (a.u., arbitrary units). $n \geq 59$ cellular ROIs/condition. A Kruskal–Wallis test, followed by Dunn's multiple comparison test, was used. ****$p \leq 0.0001$. **(D)** Colocalization of ANKRD55 (red) with ALYREF (green) in moDCs (1024 × 1024 pixels = 0.131 microns/pixel). **(E)** Colocalization of ANKRD55 (red) with HNRNPC (green) in moDCs (1024 × 1024 pixels = 0.131 microns/pixel).

3.3.5. Expression of *ANKRD55*, *IL6ST*, *IL31RA* and *SLC38A9* in pDCs and cDCs

In addition to moDCs, we aimed to detect *ANKRD55* and neighboring genes in natural circulating conventional dendritic cells (cDCs) and plasmacytoid dendritic cells (pDCs). To this aim, we used ddPCR technique due to the low number of these cells in the blood, and, therefore, the small amount of mRNA obtained from them. Absolute quantification of transcript levels was performed in cDCs, pDCs, $CD14^+$ monocytes and $CD4^+$ T lymphocytes as well as in untreated and treated moDCs (**Figure 23**). As measured in 350 pg cDNA, the total *ANKRD55* transcript copy numbers per µl (average ± STDEV) in pDCs (4.3 ± 2.8) and cDCs (1.4 ± 1.5) were slightly higher than those in $CD14^+$ monocytes (1 ± 1.15) (**Figure 23A**). In contrast, *IL6ST*, *IL31RA* and *SLC38A9* were expressed at much higher levels in $CD14^+$ cells than in pDCs or cDCs (**Figure 23E**). *IL31RA* was not detected in pDCs or cDCs.

We also used ddPCR to confirm and more accurately quantify the results obtained by qPCR of *ANKRD55* gene expression in AM580 and IFN-γ/LPS-treated moDCs. AM580 enhanced basal *ANKRD55* transcript levels in immature moDCs at day 6 from 88.8 ± 46.2 to 208.3 ± 63.8 copies/µl, and IFN-γ/LPS downregulated day 5 immature moDC *ANKRD55* levels from 89.3 ± 93.3 to 19.5 ± 14.9 copies/µl, both measured in 30 ng of cDNA (**Figures 23C, D**). In comparison, $CD4^+$ T cells expressed much higher levels of *ANKRD55* (1643 ± 1741 copies/µl), also measured in 30 ng cDNA (**Figure 23B**). The purity of the two subsets of natural circulating dendritic cells studied was assessed using specific membrane markers including CD303 (CLEC4C) and IL3RA for pDCs and CD1c (BCDA-1) for cDCs (**Figure 23F** and **Figure 24**).

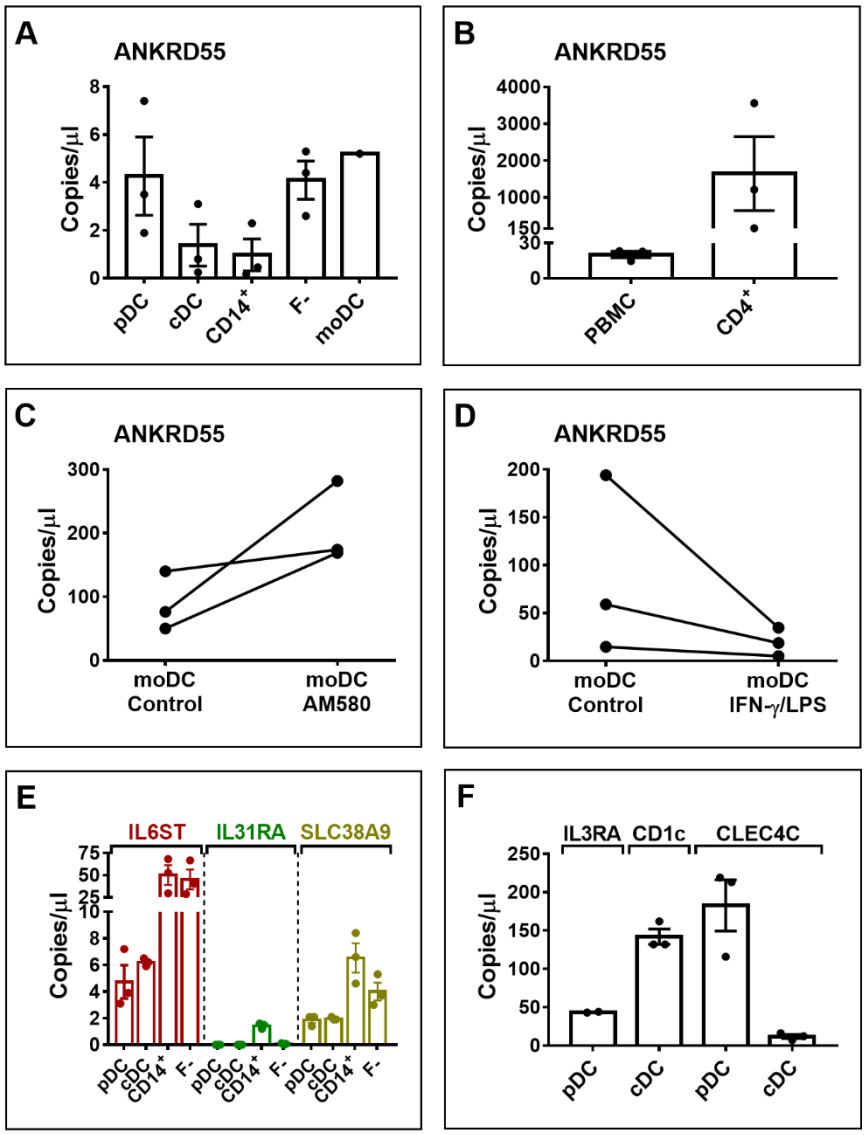


FIGURE 23. Quantification of gene transcript copy numbers using ddPCR.
ANKRD55 transcript copy number quantifications in **(A)** pDCs, cDCs, CD14$^+$ monocytes, PBMCs without monocytes (F-) and moDCs; **(B)** PBMCs and CD4$^+$ T lymphocytes; **(C)** moDCs treated with AM580; and **(D)** moDCs with or without IFN-γ/LPS maturation treatment. **(E)** ddPCR quantification of *IL6ST*, *IL31RA* and *SLC38A9* in pDCs, cDCs, CD14$^+$ and F-. **(F)** Transcript levels of markers specific for cDCs (CD1c) and pDCs (IL3RA, CLEC4C). **(A, B, E, F)** Data are shown as mean ± SEM of copies per µl, when $n > 2$. Analysis was performed in 350 pg cDNA **(A, E, F)** or 30 ng of cDNA **(B–D)**.

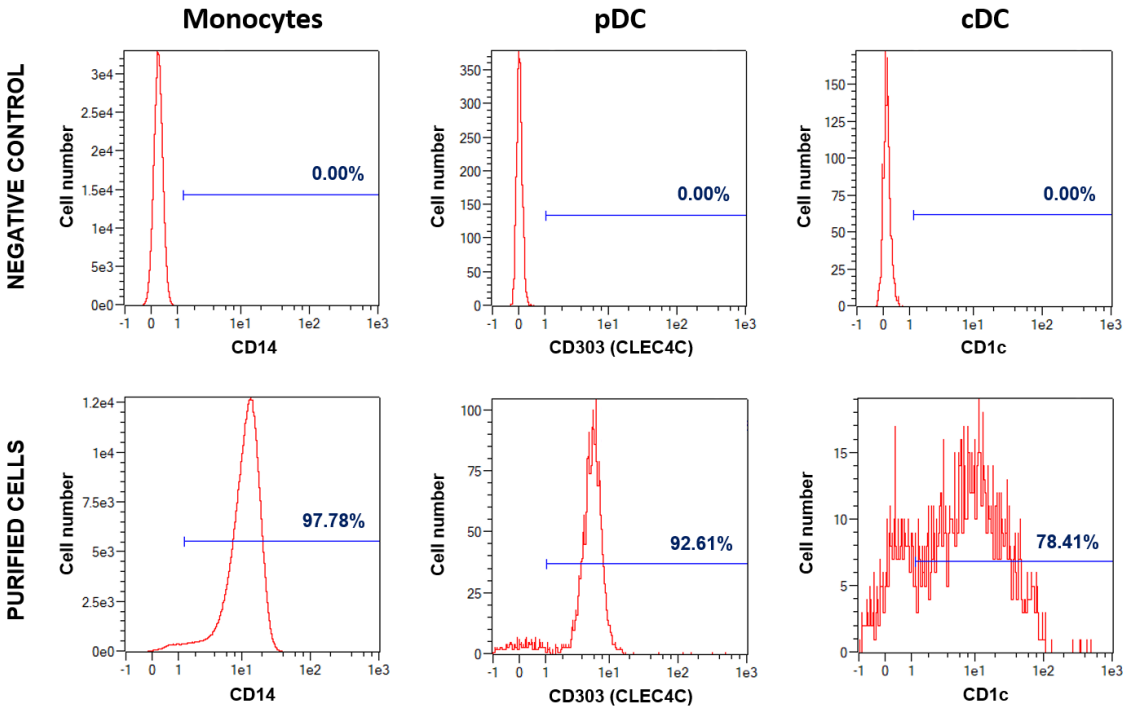


FIGURE 24. Purity of monocyte, pDC and cDC subpopulations.
Representative flow cytometry histograms with the percentage of positive cells obtained from purified (bottom row) $CD14^+$ monocytes, plasmacytoid dendritic cells (pDCs) and conventional dendritic cells (cDCs). Cells were stained with specific surface markers using CD14, CD303 (CLEC4C) and CD1c antibodies correspondingly or negative control (top row). The negative control antibodies used were the isotype control IgG1-PE antibody for CD14 and isotype control human IgG1-AP antibody for CD1c and CD303.

3.3.6. Influence of single-nucleotide polymorphisms on the expression of *ANKRD55* and neighboring genes in PBMC subpopulations

MS risk alleles are associated with enhanced expression of *ANKRD55* in $CD4^+$ T cells[317,359], but genotypic effects on the expression of the flanking genes in PBMC subsets have not been reported. We analyzed the influence of homozygosity for MS risk SNPs rs6859219 and rs7731626 and of the correlated SNP rs13186299 (explained below) on the expression of *ANKRD55*, *IL6ST*, the three sgp130 isoforms and *SLC38A9* in healthy control (HC) $CD4^+$ T cells (**Figure 25** and **Figure 26A**). Homozygotes for the risk alleles (CC for rs6859219, GG for rs7731626 and GG for rs13186299) expressed significantly higher levels of *ANKRD55* than those for the protective alleles, and directionally similar but weaker trends were observed for all other gene transcripts. Of note, homozygotes for the rs6859219 C risk allele also expressed significantly higher levels of sgp130-E10 than those for the protective allele (**Figure 25M**), and homozygotes for the rs13186299 G risk allele were associated with higher expression of *SLC38A9* compared to carriers of the protective allele (**Figure 26A**). The influence of these risk alleles on gene expression was also analyzed in four further PBMC subpopulations: $CD8^+$ T lymphocytes, $CD14^+$ monocytes, $CD19^+$ B lymphocytes and $CD56^+$ NK cells. Of all the analyses performed in these four subpopulations, only the rs13186299 G risk allele was significantly associated with lower *IL6ST* expression in $CD56^+$ cells ($p < 0.01$) (**Figure 26**).

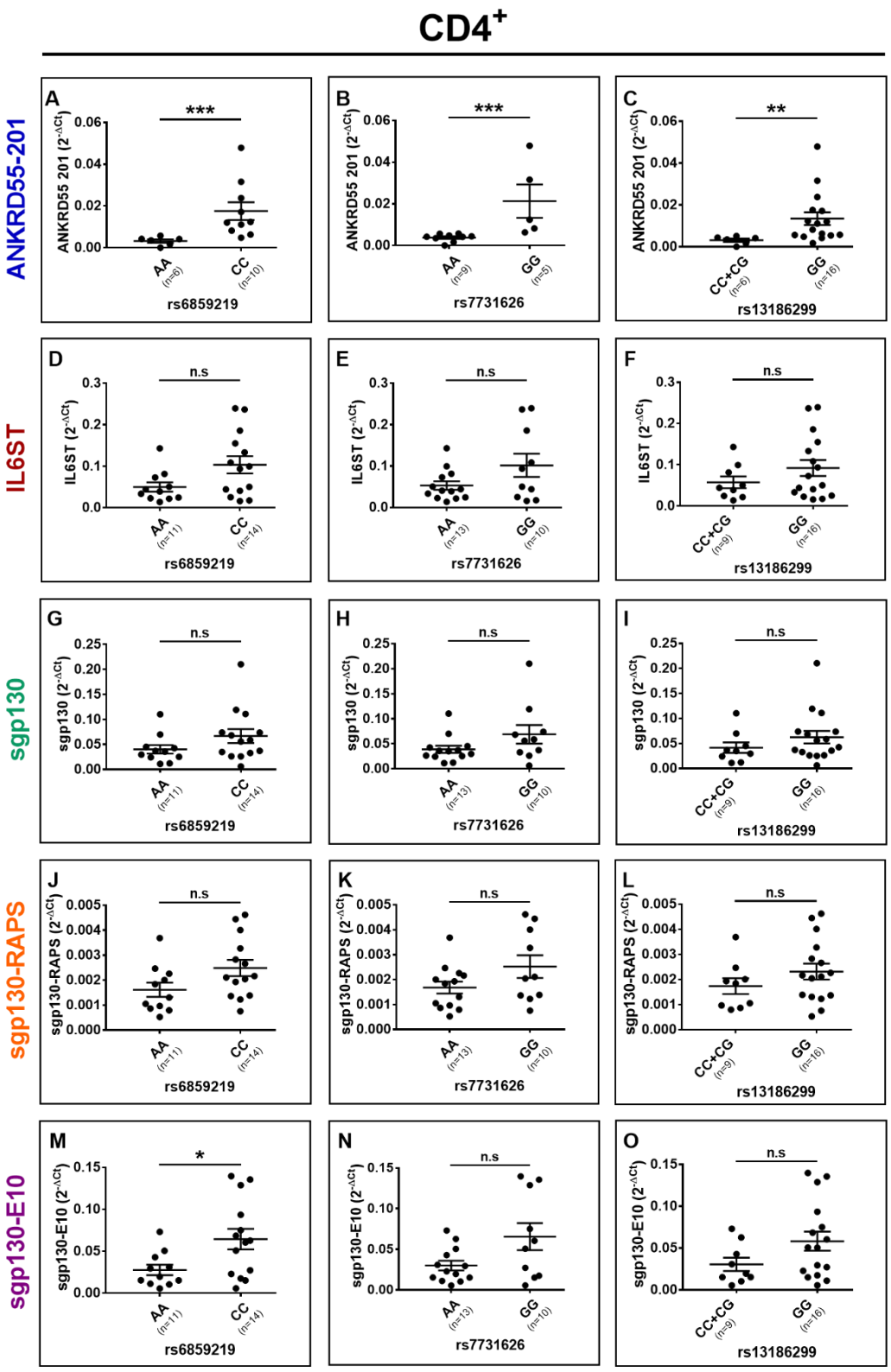


FIGURE 25. Effect of homozygosity for protective or risk alleles of the MS risk SNPs rs6859219 and rs7731626 and the correlated SNP rs13186299 on gene expression in CD4⁺ T cells.
(A, B, C) *ANKRD55* isoform 201, **(D, E, F)** *IL6ST*, **(G, H, I)** sgp130, **(J, K, L)** sgp130-RAPS and **(M, N, O)** sgp130-E10 were measured via qPCR in CD4⁺ T cells from healthy controls. **(A)** Data from Lopez de Lapuente *et al.*[317]). The risk genotypes are CC (rs6859219), GG (rs7731626) and GG (rs13186299). Data are mean ± SEM of $2^{-\Delta Ct}$ values. *ACTB* and *GAPDH* were used as housekeeping genes for normalization. The number of individual subjects analyzed per genotype is indicated in brackets. A Mann–Whitney test was performed for comparisons. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. n.s., not significant.

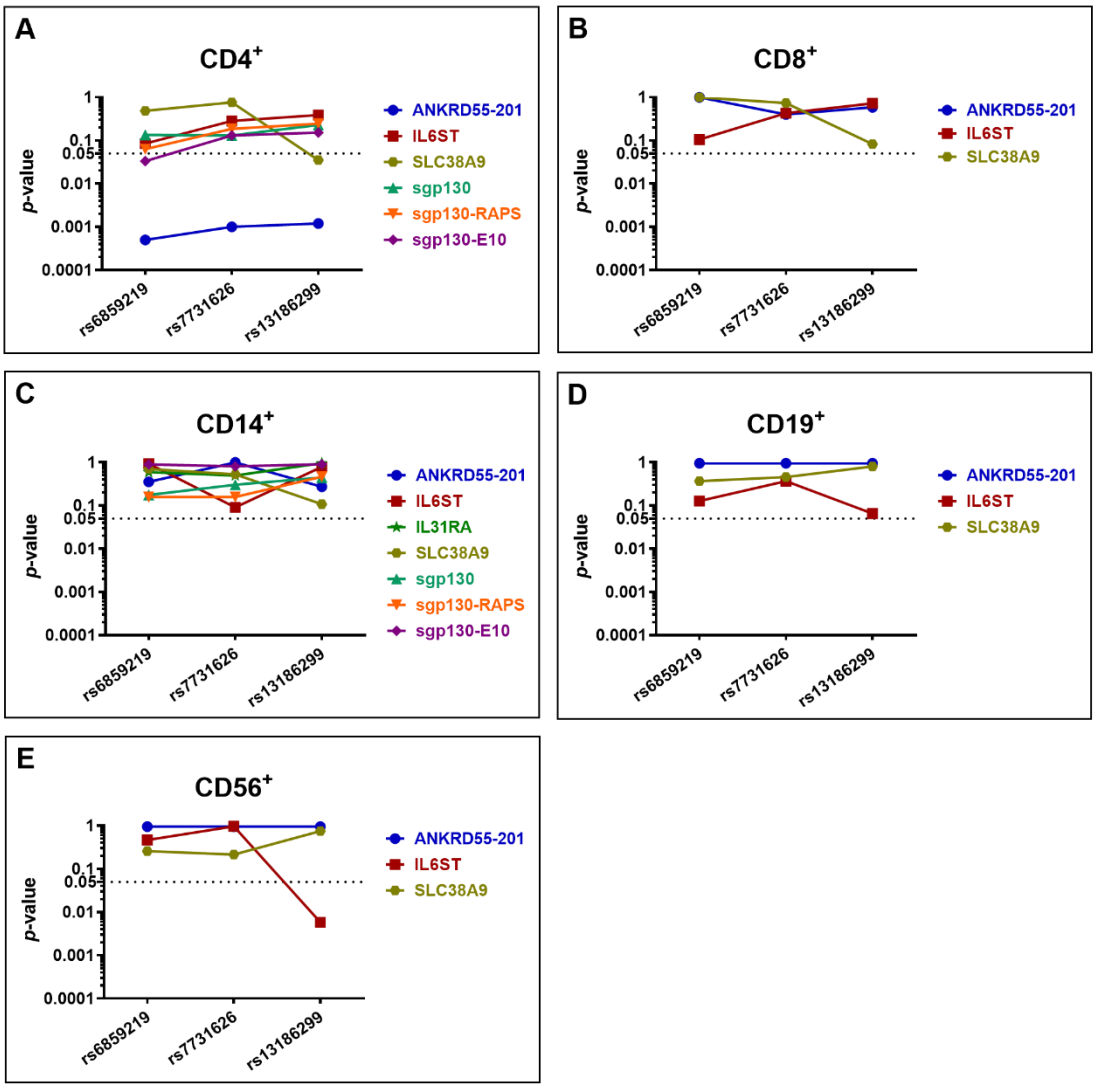


FIGURE 26. Influence of MS risk SNPs rs6859219 and rs7731626 and of the correlated SNP rs13186299 genotype on gene expression in PBMC subpopulations.

Plots of the *p*-values for association of the genotypes of indicated SNPs with *ANKRD55*, *IL6ST* and *SLC38A9* expression levels, measured via qPCR, in **(A)** $CD4^+$, **(B)** $CD8^+$, **(C)** $CD14^+$, **(D)** $CD19^+$ and **(E)** $CD56^+$ cells from healthy donors. In addition, sgp130, sgp130-RAPS and sgp130-E10 were analyzed in $CD4^+$ and $CD14^+$ cells, and *IL31RA* was analyzed in $CD14^+$ cells. Significance threshold of the *p*-value (0.005) is shown with the dotted line. **(A)** Data are from **Figure 25** adding *SLC38A9*. A Mann–Whitney test was performed for comparisons.

3.3.7. MS risk single-nucleotide polymorphisms modulate the expression of *ANKRD55* and *IL6ST* in immature moDCs

To assess whether *ANKRD55* MS risk genotypes regulate *ANKRD55* and *IL6ST* expression in immature moDCs, qPCR was performed in day 6 immature moDCs derived from healthy subjects who were genotyped for MS risk SNPs rs6859219 and rs7731626. Compared to the risk alleles, the protective A alleles of both SNPs were associated with higher mRNA expression of *ANKRD55* ($p < 0.01$ and $p < 0.05$, respectively) and *IL6ST* (not significant trend and $p < 0.01$, respectively) (**Figures 27A, B**). Notably, in $CD4^+$ T lymphocytes, the opposite directional effect is observed, as the A alleles were associated with lower levels of *ANKRD55* compared to the risk alleles[317,455,457] (**Figure 25**). The opposite effects in moDCs could be due to linkage disequilibrium (LD) with distinct genomic variants correlated to specific phenotypic traits of myeloid or monocytic cells. In this respect, we performed a search for GWAS variants in LD with these SNPs and identified two candidate SNPs meeting these criteria, rs28722705 (Chr 5: 56158115) and its proxy rs13186299 (Chr 5: 56159818) (D' = 1; R^2 = 1), which emerged with genome-wide significance from GWAS (GWAS Catalog[475] https://www.ebi.ac.uk/gwas/home) on monocyte count ($p \leq 4 \times 10^{-11}$)[459,460] and monocyte percentage of white cells ($p = 7 \times 10^{-22}$)[461], respectively. Both are intron variants located in *ANKRD55* (**Figure 14**) that occur in partial LD with both *ANKRD55* MS risk SNPs (D' = 0.62-0.98; R^2 = 0.22-0.26). Interestingly, in the PheWAS database from Open Targets Genetics[476, (https://genetics.opentargets.org), which integrates studies by the FinnGen, UK Biobank and GWAS Catalog consortia, MS risk SNPs rs7731626 and rs6859219 are reportedly associated with monocyte count with genome-wide significance ($p = 1.8 \times 10^{-14}$ and $p = 5.2 \times 10^{-19}$, respectively), suggesting colocalization of MS risk and monocyte count traits. The Variant-to-Gene pipeline from Open Targets Genetics[476], which integrates QTL data, Promoter Capture Hi-C data, *in silico* functional predictions and variant distance to gene, prioritizes *ANKRD55* and *IL6ST* as the genes functionally implicated by all four variants. In the GTEx portal (https://gtexportal.org/home), rs13186299 is recorded as the most significant single-tissue eQTL for *IL6ST* expression in whole blood ($p = 4.4 \times 10^{-11}$; normalized effect size 0.17) (**Figure 27D**).

Reanalysis of the expression data in the sample cohort now genotyped for rs13186299 (**Figure 27C**) revealed a highly significant association of *ANKRD55* and *IL6ST* mRNA levels with genotype in immature moDCs. The protective C allele of rs13186299 was related to a higher expression of both genes ($p < 0.001$ for *ANKRD55* and $p < 0.0001$ for *IL6ST*), identifying this SNP

as exerting a larger effect upon *ANKRD55* and *IL6ST* expression in moDCs than either MS risk SNP.

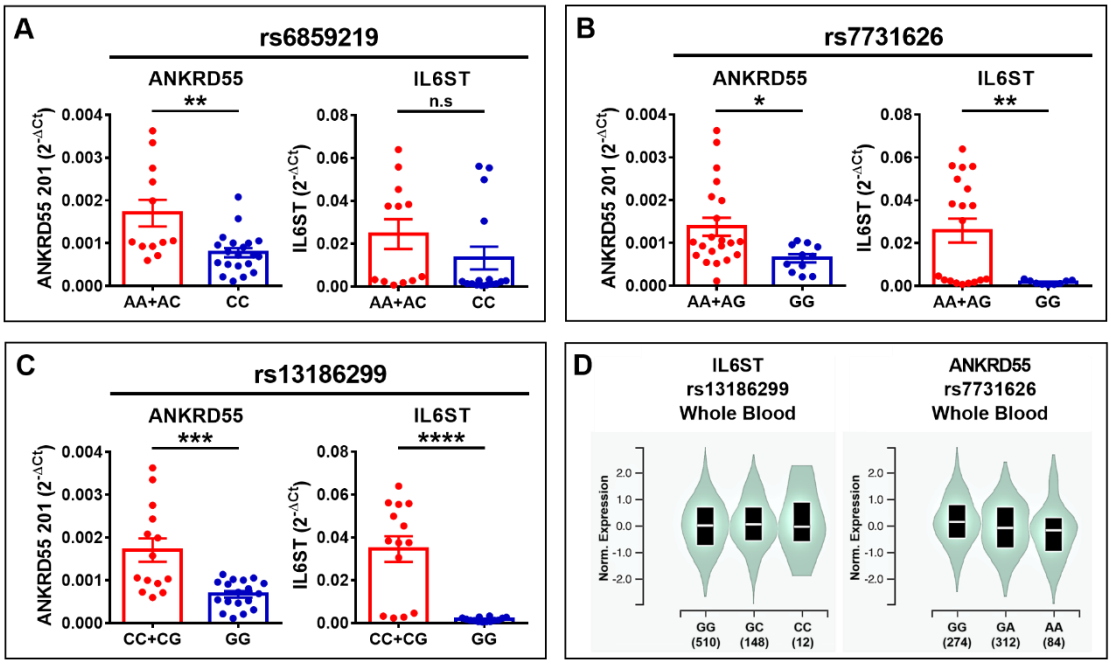
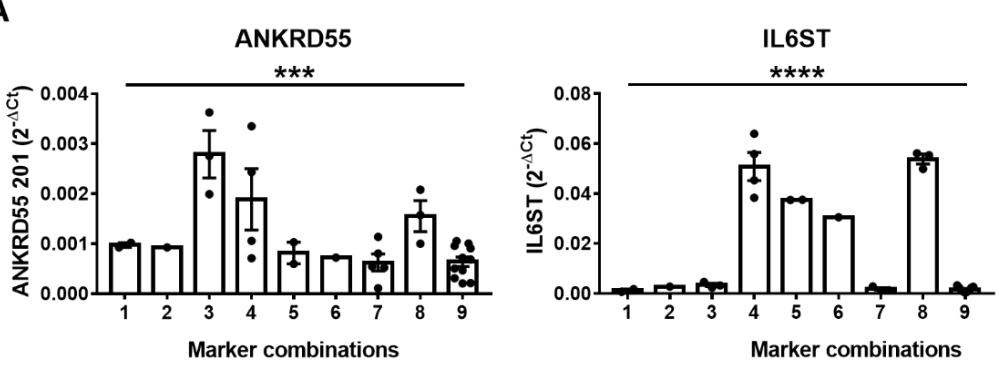


FIGURE 27. Modulation of *ANKRD55* and *IL6ST* gene expression in immature moDCs of healthy donors by SNPs rs6859219, rs7731626 and rs13186299.

Samples were genotyped for **(A)** rs6859219, **(B)** rs7731626 and **(C)** the correlated *ANKRD55*-intronic SNP rs13186299. Gene expression was measured via qPCR using *ACTB* and *GAPDH* as internal control genes for normalization. The risk genotypes are CC for rs6859219, GG for rs7731626 and GG for rs13186299, and the protective allele is A for rs6859219, A for rs7731626 and C for rs13186299. Mean ± SEM of $2^{-\Delta Ct}$ values is presented; $n \geq 28$. A Mann–Whitney test was used. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$. n.s., not significant. **(D)** Genomic variants rs13186299 (left) and rs7731626 (right) are the most significant single tissue eQTLs for *IL6ST* and *ANKRD55* in whole blood, respectively, based on information collected from the GTEx portal (https://gtexportal.org/home).

Expression levels showed roughly similar variations according to marker combination (exception made for combination 3: rs6859219 AA, rs7731626 AA and rs13186299 GC), with the rs13186299 GG genotype typically segregating with low levels of either gene product (**Figure 28A**). In the European population, the MS risk SNP rs7731626 G allele is uniquely found in a single haplotype together with the rs13186299 G allele (LDlink database; https://ldlink.nci.nih.gov) (**Figure 28B**). Therefore, homozygotes for the rs7731626 G allele, corresponding to 38% of European rs7731626 sample genotypes (Ensembl data base[315]; https://www.ensembl.org), are predicted to display genetically-determined low levels of both genes in moDCs.

A



Marker combination	rs6859219 genotype	rs7731626 genotype	rs13186299 genotype	Max. Number of samples
1	AA	AA	GG	2
2	AA	AA	CC	1
3	AA	AA	GC	3
4	AC	AG	GC	4
5	AC	AA	GC	2
6	CC	AA	CC	1
7	CC	AG	GG	5
8	CC	AG	GC	3
9	CC	GG	GG	11

B

RS Number	Position (GRCh37)	Allele Frequencies	Haplotypes				
rs6859219	chr5:55438580	C=0.78, A=0.22	C	C	A	A	C
rs7731626	chr5:55444683	G=0.622, A=0.378	G	A	A	A	A
rs13186299	chr5:55455645	G=0.859, C=0.141	G	G	G	C	C
		Haplotype Count	622	121	119	100	40
		Haplotype Frequency	0.6183	0.1203	0.1183	0.0994	0.0398

FIGURE 28. Marker combinations and haplotype distribution of MS risk SNPs rs6859219 and rs7731626 and of the correlated SNP rs13186299.

(A) Assessment of *ANKRD55* (left) and *IL6ST* (right) expression levels according to three SNP marker combinations. qPCR was performed using *ACTB* and *GAPDH* as housekeeping genes. Data are shown as mean ± SEM of $2^{-\Delta Ct}$ values ($n \geq 28$) and a one-way ANOVA test was used. ***$p \leq 0.001$, ****$p \leq 0.0001$. **(B)** Haplotype distribution in the European population of the three SNPs. Image information was produced by means of LDlink database (https://ldlink.nci.nih.gov).

We conducted further analyses to determine the effect of genotype on gene expression in immature, AM580-treated and mature (IFN-γ/LPS-treated) moDCs. Compared to untreated cells, levels of *ANKRD55* splice variants induced by AM580 in immature moDCs were significantly higher in homozygous carriers of risk alleles but were not significantly different in homozygotes of the protective alleles, though the scarcity of the latter genotype restricted the power of the test in this group (**Table 8**). In the case of sgp130 isoforms, none of the compared groups, according to genotype, reached significant differences (**Figure 29**). However, the mRNA levels of sgp130 and sgp130-RAPS in immature untreated moDCs were higher in carriers of protective alleles for the three SNPs and homozygotes of the risk alleles showed a slightly higher expression of sgp130 in the three SNPs.

TABLE 8. Effect of risk SNP homozygosity on AM580-induced expression of *ANKRD55* isoforms.
Healthy donor moDCs were cultured in the absence (control) or presence of AM580. Gene expression was assessed via PCR and normalized to internal control *ACTB*. Statistical comparisons were performed with a Mann–Whitney test or an unpaired *t*-test. **p* ≤ 0.05, ***p* ≤ 0.01, ****p* ≤ 0.001.

Isoform	SNP	Genotype	Average fold change (AM580/Control)	*p*-value	Number of paired samples
ANKRD55-201	rs6859219	AA	1.55	0.1875	5
		CC	3.32	0.0020 (**)	11
	rs7731626	AA	1.55	0.1875	5
		GG	3.12	0.0156 (*)	7
	rs13186299	CG	1.19	0.7500	3
		GG	3.21	0.0010 (**)	12
ANKRD55-202	rs6859219	AA	1.38	0.6250	5
		CC	2.56	0.0049 (**)	11
	rs7731626	AA	1.38	0.6250	5
		GG	2.64	0.0156 (*)	7
	rs13186299	CG	1.26	> 0.9999	3
		GG	2.51	0.0024 (**)	12
ANKRD55-204	rs6859219	AA	2.31	0.0625	5
		CC	2.89	0.0010 (***)	11
	rs7731626	AA	2.31	0.0625	5
		GG	2.75	0.0156 (*)	7
	rs13186299	CG	2.38	0.2500	3
		GG	2.79	0.0005 (***)	12

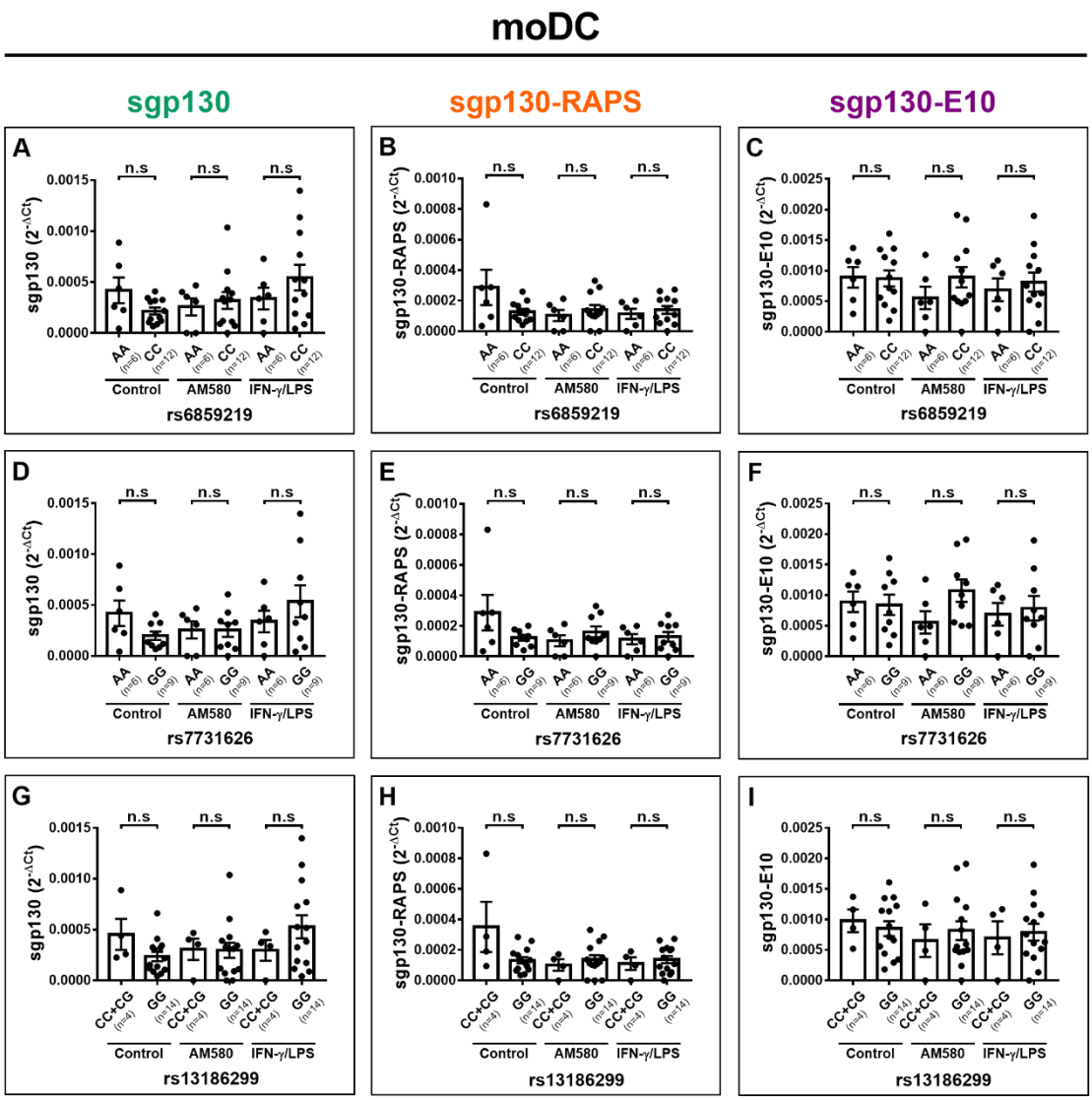


FIGURE 29. Effect of homozygosity for protective or risk alleles of the three SNPs on expression of sgp130 isoforms in immature, AM580-treated and IFN-γ/LPS-matured moDCs.
Monocytes from healthy donors were co-cultured in MoDC medium for six days in the presence or absence (control) of 100 nM AM580 from day 0, and new AM580 was added on day 3 of the moDC differentiation period. Alternatively, moDCs were matured with a combination of 0.5 pg/ml IFN-γ for 12 hours and 200 ng/ml LPS for 6 hours. Samples were grouped by genotype of the MS risk SNPs **(A-C)** rs6859219, **(D-F)** rs7731626 and **(G-I)** correlated SNP rs13186299. The expression of sgp130, sgp130-RAPS and sgp130-E10 was measured via qPCR and *ACTB* was used as housekeeping gene for normalization. Data are mean ± SEM of $2^{-\Delta Ct}$ values. The number of individual subjects analyzed per condition is indicated in brackets. A Mann–Whitney test or an unpaired *t*-test was used. The risk genotypes are CC (rs6859219), GG (rs7731626) and GG (rs13186299). n.s., not significant.

3.3.8. *ANKRD55* and *IL6ST* gene expression patterns in CD4$^+$ T cells and moDCs of MS patients

We next set out to verify whether these genetic influences on *ANKRD55* and *IL6ST* gene expression are measurable in CD4$^+$ T lymphocytes and immature moDCs derived from CD14$^+$ monocytes freshly isolated from venous blood samples from untreated MS patients (clinical and demographic data of MS patients in **Table 2**). In CD4$^+$ T cells, the rs7731626 genotype was significantly associated with mRNA expression levels of *ANKRD55* ($p < 0.01$) but not with that of *IL6ST*, while the other two SNPs were not associated with expression levels of any gene (**Figures 30A, C**). In immature moDCs, no association between genotype of any SNP and expression level of *ANKRD55* or *IL6ST* was observed (**Figures 30B, D**). Expression levels of the three sgp130 isoforms (sgp130, sgp130-RAPS, and sgp130-E10) were also measured by qPCR but did not vary with SNP genotype, though two weaker associations ($p < 0.05$; CD4$^+$ sgp130-RAPS for rs13186299 and moDC sgp130-E10 for rs7731626) were observed (**Figures 30C, D**). We also calculated Pearson correlation coefficients between individual *ANKRD55* and *IL6ST* mRNA expression levels in both cell subsets. In immature moDCs, expression levels of *IL6ST* and *ANKRD55* were moderately positively correlated, with Pearson's $r = 0.45$ ($p < 0.0001$). In CD4$^+$ T cells, however, their expression levels were more strongly correlated, with Pearson's $r = 0.76$ ($p < 0.0001$) (**Figures 30E, F**). It is worth noting that the expression levels of both genes in moDCs of rs13186299 GG homozygotes were higher in MS than in healthy controls (trending at $p = 0.12$ for *ANKRD55* and significant for *IL6ST* at $p < 0.0001$ upon comparison of both groups), while those of C carriers were not significantly different ($p > 0.05$ for comparison of both groups) (**Figure 31**). This observation is suggestive of allele-restricted preconditioning of CD14$^+$ monocytes in MS patients.

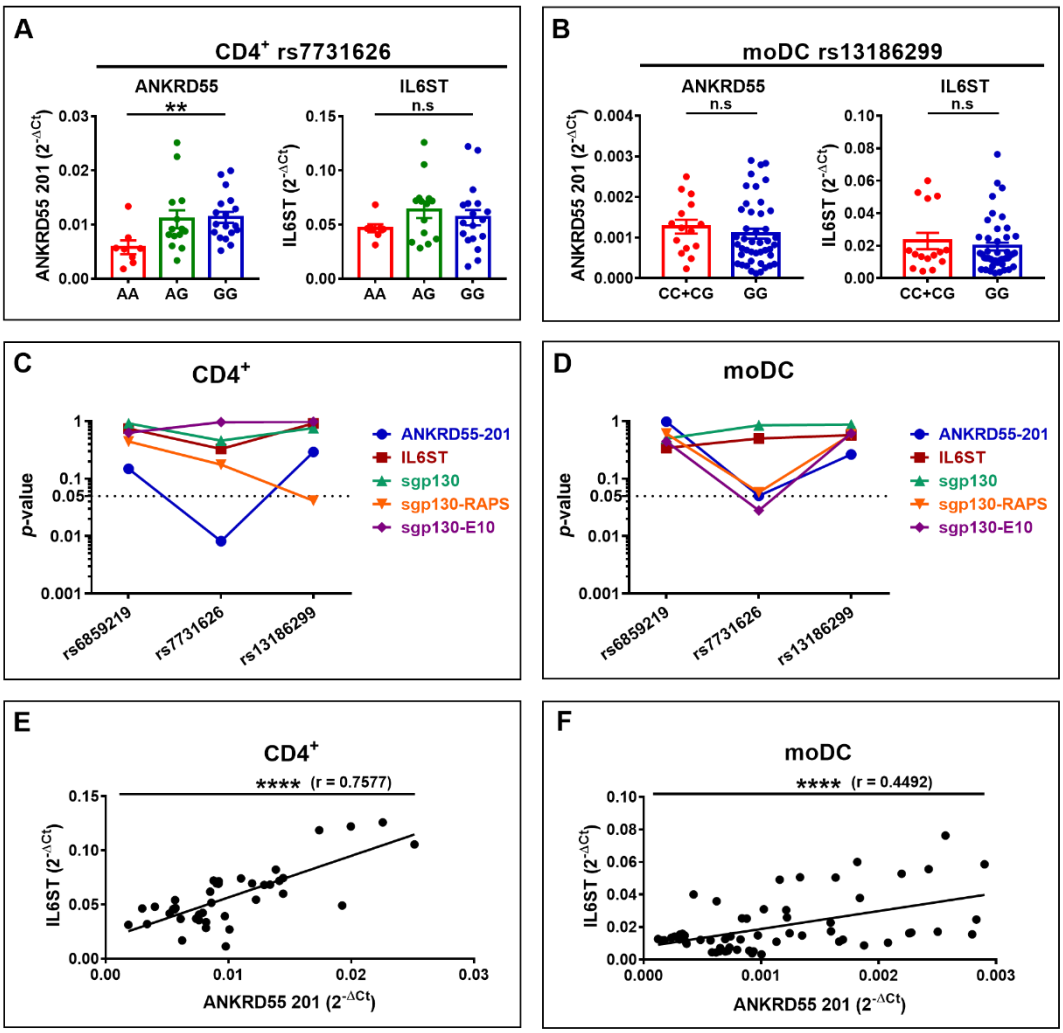


FIGURE 30. Expression of *ANKRD55*, *IL6ST* and sgp130 isoforms in CD4⁺ T cells and immature moDCs of MS patients.

(A, B) Effects of the rs7731626 genotype in CD4⁺ T cells (mean ± SEM of $2^{-\Delta Ct}$ values; *n* = 40, Cohort 2 [see **Table 2**]; Kruskal–Wallis test) and the rs13186299 genotype in moDCs (mean ± SEM of $2^{-\Delta Ct}$ values; *n* = 59, Cohorts 1 and 2 [see **Table 2**]; Mann–Whitney test), respectively. Expression levels of *ANKRD55* and *IL6ST* were measured via qPCR, using *ACTB* and *GAPDH* as reference genes for normalization. **(C, D)** Plots display *p*-values for the effects of three SNPs (rs6859219, rs7731626 and rs13186299) on the expression of *ANKRD55*, *IL6ST*, sgp130, sgp130-RAPS and sgp130-E10 in CD4⁺ T cells and moDCs. Mann–Whitney or Kruskal–Wallis tests were used for statistical comparisons. Significance threshold of the *p*-value (0.005) is shown with the dotted line. **(E, F)** Pearson correlation of *ANKRD55* and *IL6ST* $2^{-\Delta Ct}$ values in CD4⁺ T cells and in moDCs of MS patients from panels (A, B), respectively. $^{**}p \leq 0.01$, $^{****}p \leq 0.0001$. n.s., not significant.

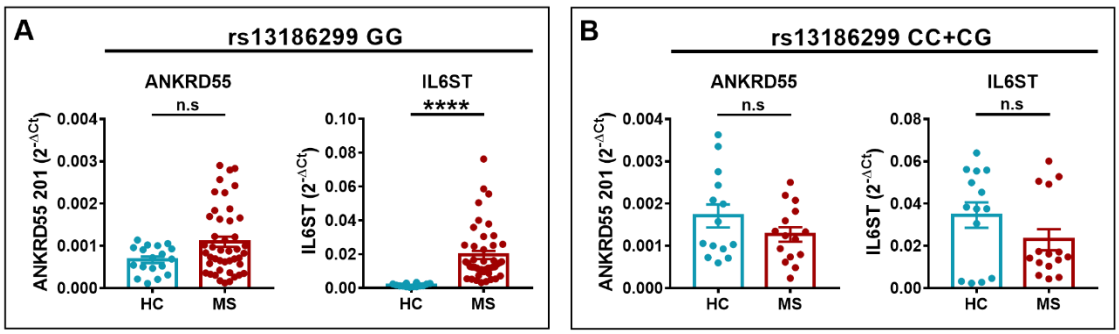


FIGURE 31. Comparison of levels of *ANKRD55* and *IL6ST* between healthy controls and MS patients in moDCs according to rs13186299 genotype.

Expression levels of *ANKRD55* and *IL6ST* assessed via qPCR in moDCs of **(A)** GG homozygotes and **(B)** carriers of the C allele (CC + CG) of rs13186299 were compared between healthy controls (HC; $n \geq 14$; data from **Figure 27C**) and MS patients (MS; $n \geq 14$; data from **Figure 30B**). Mean ± SEM of $2^{-\Delta Ct}$ values is shown and a Mann–Whitney test was performed for comparisons between groups. ****$p \leq 0.0001$. n.s., not significant.

Clinical course of MS patients was also considered as a variable of study. Due to the frequent transition from RRMS to SPMS, these two patient cohorts were grouped for the analysis. As result, in moDCs, increased *ANKRD55* gene expression was observed in PPMS patients compared to healthy controls ($p < 0.05$) as well as decreased *IL6ST* expression in both PPMS and RR + SPMS groups ($p < 0.01$; **Figure 32**).

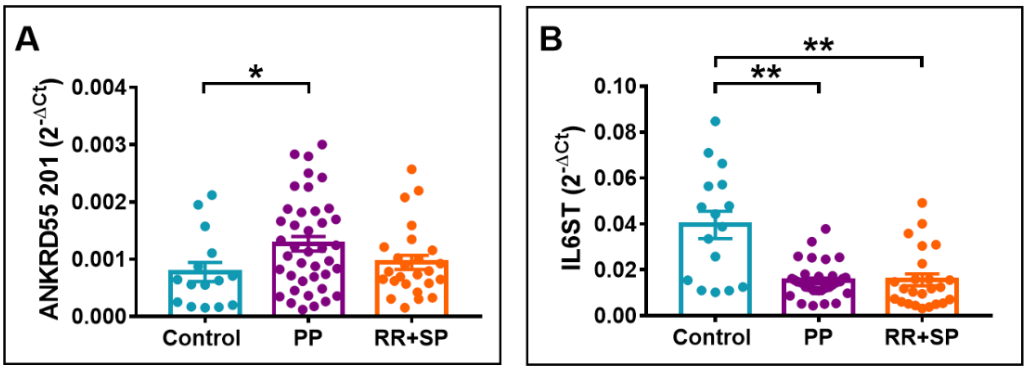


FIGURE 32. Differential *ANKRD55* and *IL6ST* gene expression levels in moDCs according to clinical course of MS.

(A) *ANKRD55* and **(B)** *IL6ST* gene expression was measured via qPCR in moDCs from healthy controls, primary progressive (PP) and relapsing remitting (RR) + secondary progressive (SP) MS patients. *ACTB* and *GAPDH* were used as internal controls for normalization. Data are presented as mean ± SEM of $2^{-\Delta Ct}$ values; $n = 80$ from Cohorts 1 and 2 (see **Table 2**). A Kruskal–Wallis test, followed by Dunn's multiple comparison test, was performed. *$p \leq 0.05$, **$p \leq 0.01$.

Finally, we speculate on the role of sgp130 in serum of MS patients. Sgp130 has been linked to many diseases, including MS[477]. A recent publication showed significantly increased serum sgp130 levels in MS carriers of the protective allele of the correlated MS risk SNP rs71624119[478]. To verify these results in our cohorts and SNP of interest, we quantified serum sgp130 levels in MS patients using ELISA. Sgp130 levels did not differ in PPMS compared to RR + SPMS patients (**Figure 33A**) but showed significant variation with respect to the rs7731626 genotype. Higher levels of sgp130 were observed in homozygotes of the protective A allele ($p < 0.05$) for this SNP, and similar trends were observed with the SNPs rs6859219 and rs13186299 (**Figure 33B**).

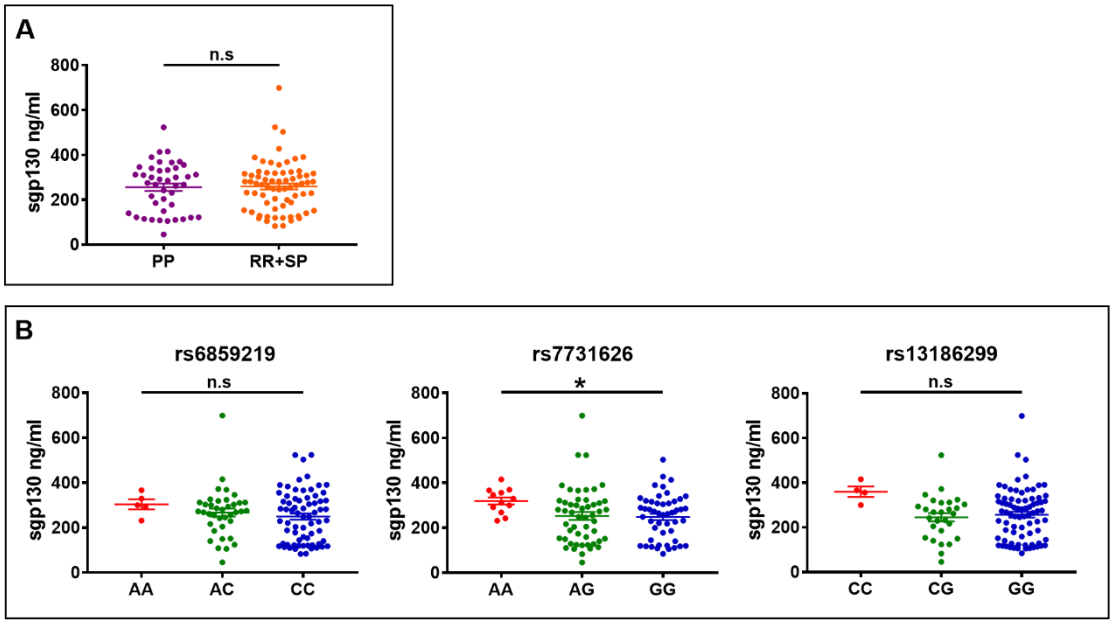


FIGURE 33. Levels of serum sgp130 protein in MS patients and association with SNP genotypes. **(A)** The sgp130 protein was measured via ELISA in the serum of PPMS and RR + SPMS. Mean ± SEM of concentration (ng/ml) is shown; $n = 112$ from Cohorts 1, 2, and 3 (see **Table 2**). A Mann–Whitney test was performed. **(B)** Serum sgp130 levels in MS patients from panel (A) stratified for rs6859219, rs7731626 or rs13186299 genotype. A Kruskal–Wallis test, followed by Dunn's multiple comparison test, was used for statistical comparisons between groups. *$p \leq 0.05$, **$p \leq 0.01$. n.s., not significant.

3.4. Discussion

The results of this study demonstrate that the *ANKRD55* gene is induced in monocytes during differentiation into immature moDCs in the presence of IL-4/GM-CSF. Moreover, the expression of *ANKRD55* in these cells is enhanced by retinoic acid agonist AM580 and suppressed by maturation with IFN-γ/LPS treatment, while expression of *IL6ST*, the adjacent gene, appears less affected by either of these treatments. Expression levels of both genes in immature moDCs are similarly modulated by the MS risk SNPs rs6859219 and rs7731626, with opposite directional effect to that seen in $CD4^+$ T lymphocytes: MS risk alleles are associated with higher levels of *ANKRD55* and *IL6ST* in $CD4^+$ T lymphocytes of healthy controls[317,455] but with lower levels in immature moDCs compared to the protective alleles. Though we did not perform a formal eQTL study, further analysis revealed that these allelic effects in moDCs are arguably due to LD of the MS risk SNPs with another nearby intronic SNP in *ANKRD55*, rs13186299. This SNP has not arisen from GWAS in MS but from a GWAS on the monocyte percentage of white blood cells[461] and is the main eQTL SNP for *IL6ST* expression in whole blood, according to GTEx database (https://gtexportal.org/home). Tissue-specific opposite eQTL effects have been observed for GWAS SNPs associated with complex traits including MS, with a fraction of these showing discordant effects in untreated $CD4^+$ T cells and monocytes[472,479,480]. Around 30% of the 200 GWAS autosomal risk variants (SNPs) contained within the current MS genomic map can be tagged to *cis*-QTL effects on gene expression in either naïve $CD4^+$ T cells or monocytes[274]. Of these SNPs, only five passed the false discovery rate (FDR) threshold for expression of *ANKRD55* in $CD4^+$ T cells and PBMCs but not in monocytes or brain[274]. Thus, MS risk SNPs in *ANKRD55* have been previously associated with a constitutive regulatory *cis*-effect on *ANKRD55* and *IL6ST* gene expression in $CD4^+$ T cells. Our work now links these MS risk SNPs in *ANKRD55* to a distinct regulatory element mediating an opposite IL-4/GM-CSF-induced *cis*-effect in monocytes.

It has been previously reported that the expression levels of *ANKRD55* and *IL6ST* genes are similarly affected by SNP rs7731626 in healthy controls[457]. However, in $CD4^+$ T cells of MS patients, we show that rs7731626 is significantly associated with *ANKRD55* expression but not with *IL6ST* (**Figure 30A**). One study showed a *cis*-eQTL effect for this SNP on *ANKRD55* expression in both naïve and memory $CD4^+$ T cells of MS patients[458]. In immature moDCs of MS patients, we found that *ANKRD55* and *IL6ST* expression was independent of the tested SNPs, in contrast to healthy controls, whose expression levels of both genes were significantly associated with rs7731626 and rs13186299 (**Figure 30D** and **Figures 27B, C**). This discrepancy between healthy

subjects and patients may be the result of preconditioning of blood monocytes by factors enriched in the serum of MS patients prior to their isolation and differentiation into moDCs and was only seen in rs13186299 GG homozygotes (**Figure 31A**). Although we did not perform a functional analysis, we assessed whether clinical course could be a priming factor affecting expression of both genes in moDCs. In PPMS, but not in RRMS or SPMS, blood monocytes are characterized by an IL-1β signature, a consequence of NLRP3 inflammasome overactivation[481]. Our analysis of immature moDCs generated from circulating monocytes in PPMS, RRMS and SPMS patients revealed enhanced expression of *ANKRD55* uniquely in the PPMS group, while *IL6ST* was decreased in all MS patients compared to healthy controls regardless of clinical course (**Figure 32**). These effects were independent of genotype, which therefore in healthy controls, but less so in MS, indicates moDC coregulation of *ANKRD55* and *IL6ST*.

moDCs are used widely as an *in vitro* DC model and are considered a potential target for therapy in MS[482]. They are, in fact, a highly plastic cell type capable of many discrete functions. They are closely related to and share transcriptomic gene signatures with naturally occurring inflammatory DCs that emerge from monocytes under inflammatory conditions[153,155,176] and that do not develop via a common DC precursor, unlike pDCs and cDCs. Inflammatory DCs show considerable phenotypic heterogeneity, and they function mainly at a site of inflammation rather than migrating to lymph nodes. One such subset, the $SLAN^+$ DCs, represent a subfraction of $CD14^+/CD16^+$ monocytes with pronounced pro-inflammatory functions that has been shown to accumulate in highly inflammatory brain lesions in MS patients[182]. More recently, single-cell RNA sequencing studies identified the DC subset DC4, which is marked by high CD16, shares signatures with monocytes and may be related to $SLAN^+$ DCs[153,155]. Though their role in MS is still poorly understood, inflammatory DCs are known to stimulate $CD4^+$ T cells production of IL-17 and induce Th17 differentiation from naïve $CD4^+$ T cells[483]. These effects may be mediated through IL-6 trans-signaling into T cells that abrogates the *de novo* induction of autoimmune inhibitory Tregs[484] and promotes differentiation of pathogenic Th17 cells[485], in turn contributing to CNS inflammation in MS[486]. These effects can be modulated by sgp130, which inhibits IL-6 trans-signaling by binding to the IL-6/sIL-6R complex, thus blocking its interaction with gp130 cell surface receptors[487].

Given this context, as part of this study, we analyzed the expression of three secreted gp130 isoforms (sgp130, sgp130-RAPS, and sgp130-E10) produced by the *IL6ST* gene using qPCR, as well as the levels of the circulating sgp130 using ELISA. In $CD4^+$ T cells of healthy controls, MS risk alleles were associated with higher mRNA levels of sgp130 isoforms compared to protective

alleles, while in moDCs, the lowest levels of sgp130 isoforms were observed in the MS risk alleles. Serum levels of sgp130 were higher in MS patients homozygous for the protective allele of MS risk SNPs, reaching significance for rs7731626 (**Figure 33B**). This is concordant with a recent publication showing significantly increased serum sgp130 levels in MS carriers of the protective allele of the correlated MS risk SNP rs71624119[478]. In this recent study, serum levels of sgp130 increased in all patients treated with fingolimod but to a higher degree in patients who are homozygous for the risk allele[478]. Findings from our work and the aforementioned study are compatible with higher protection against MS risk through blockade of the trans-signaling pathway by sgp130 and the consequent downregulation of IL-6/sIL6R-dependent pro-inflammatory activities, highlighting these MS risk SNPs as potential determinants of circulating sgp130. However, this pattern is incongruent with the rs7731626 genotype effect on sgp130 mRNA levels in $CD4^+$ T lymphocytes (**Figure 25**), which shows a pattern in homozygotes that opposes that of serum sgp130 levels (**Figure 33B**). One possible explanation is that rs7731626, probably via LD with other eQTL-linked genomic variants such as rs13186299, may affect serum sgp130 production by other PBMC subsets in opposite allelic direction to that of $CD4^+$ T lymphocytes. In line with this hypothesis, the DICE database (https://dice-database.org) indicates opposite genotype effects for rs7731626 on *IL6ST* transcription in B cells, classical and non-classical monocytes and NK cells when compared to $CD4^+$ T cells. Still, among PBMCs, $CD4^+$ T lymphocytes, the most abundant PBMC subset, are the highest producers of *IL6ST* and sgp130 mRNA (Human Protein Atlas[368]; **Figure 15**). The contrasting genotype-serum sgp130 ELISA data in this study could therefore refer to a relatively limited contribution of $CD4^+$ T lymphocytes to circulating levels of sgp130 or to additional sgp130-producer contributors from vessel walls or non-PBMC subsets. It may also point to still unknown effects of genomic *cis*-variants affecting cell-specific sgp130 isoform splicing or stability and/or posttranslational processing and trafficking of sgp130, which are correlated to the main MS risk SNP. A missense variant (Gly148Arg) in exon 5 of the *IL6ST* gene, rs2228044, that has been associated with higher serum sgp130 levels[488] is not in LD with rs7731626 (R^2 = 0.0008 in the European population), ruling out a mechanistic correlation. In light of these current observations, and considering serum sgp130 an apparently polygenic trait influenced by genomic variants acting in *trans*[416], the decreased levels of circulating sgp130 may constitute a secondary effect of the major causative link to MS risk conferred by rs7731626, which thus far is understood as an upregulation of both *ANKRD55* and *IL6ST* expression in $CD4^+$ T lymphocytes[334,461].

Among the tolerogenic compounds with the ability to modulate gene expression in moDCs, we found retinoic acid to be a common inducer of most of genes analyzed in this cell

subset. Retinoic acid is the active metabolite of vitamin A, also called retinol, which must be obtained from the diet. Retinol is metabolized into retinal, which is subsequently converted into retinoic acid by retinaldehyde dehydrogenase (RALDH)[489]. Once in the nucleus, retinoic acid can activate its receptors, including retinoic acid receptors (RARs) and retinoid X receptors (RXRs), which bind to DNA to regulate transcription and induce gene expression[470]. We observed that a RAR-α agonist, AM580, significantly increased *ANKRD55* expression in moDCs. This same agonist has been shown to upregulate many genes in moDCs[490]. In leukaemia cells, retinoic acid is known to induce *IL6ST* (gp130) expression, which is in line with our results in moDCs[491]. Several studies in DCs have shown that IL-4, GM-CSF and retinoic acid coordinately induce the expression of *ALDH1A2* (also known *RALDH2*), one of the retinaldehyde dehydrogenase isoforms necessary for the metabolism of the retinoic acid[492–494]. We found the same result in moDCs, whose *in vitro* differentiation is driven by IL-4 and GM-CSF cytokines, suggesting that moDCs may be capable of retinoic acid synthesis and signalling, as proposed by other studies[495]. Based on previous findings of RAR-α and RAR-γ expression in moDCs[496], we also analyzed these retinoic acid receptors in mature and AM580-treated moDCs and observed a significantly reduced expression of RAR-α with IFN-γ/LPS maturation factor treatment. Retinoic acid is crucial for a wide range of physiological and biological processes, such as development, cell growth and differentiation, vision and metabolism[497]. In immunity, it is involved in tolerance and the inflammatory response by regulating T, B and dendritic cell populations. Vitamin A deficiency leads to dysregulation of immune tolerance and pathogenic immune cell production and has therefore been linked to immune diseases such as MS[498]. Retinoic acid can suppress EAE[499] and reduce proliferation of myelin oligodendrocyte glycoprotein-reactive lymphocytes in MS patients[500] and may be useful for therapeutic application. All of these findings highlight the importance of studying the functions of this molecule in immune cells, such as dendritic cells, and its relationship with diseases such as MS.

Finally, this work revealed the presence of ANKRD55 protein in both cytosolic and nuclear compartments of moDCs. We observed that nuclear ANKRD55 colocalized with proteins typically contained within nuclear speckles. Nuclear speckles were originally thought to be sites for storage and modification of RNA splicing factors but are now recognized as hubs for integration of all the nuclear gene expression regulation steps[474]. The nuclear speckles-based location of ANKRD55 observed here is compatible with the enrichment of RNA-binding proteins previously found in the HEK293 cell line nuclear interactome[377]. Through the use of microscopy techniques, some of the results obtained in the mRNA expression studies were validated at the

protein level, including a higher expression of ANKRD55 in moDCs compared to $CD14^+$ monocytes and a decreased level of ANKRD55 protein in mature compared to immature moDCs.

Although little is known about the biological relevance of ANKRD55 in autoimmunity, its expression in immature moDCs, upregulation by retinoic acid and downregulation by maturation factors suggest a specific role in the differentiation stage of monocyte precursors into DCs. In addition, the coregulation of *ANKRD55* and *IL6ST* in immature moDCs by genomic variants associated with MS risk reported in this work highlights the importance of performing more extensive genome-wide quantitative trait loci studies in intermediate or induced states of immune cells in order to gather an integrated understanding of the multifarious expression of MS risk conferred by associated genomic variants.

4. Part II. Protein interactome study

4. Part II. Protein interactome study

4.1. Context and objectives

This second part of the thesis focuses on the ANKRD55 protein. As mentioned in the Introduction (section 1.5.1. *ANKRD55*), the structure of the full-length isoform of the ANKRD55 protein is composed of nine ankyrin repeat domains (**Figure 12**), which are known to mediate protein–protein interactions. The biological function of ANKRD55, however, requires clarification, and its functional relevance to autoimmunity remains to be elucidated. In this line, the main objective of this part of the project is to determine, study and analyze the protein interactome of ANKRD55 in different cell lines to understand the role and cellular function of ANKRD55. Protein–protein interactomes, defined as the complete set of all protein interactions that occur in a cell, including functional relationships and biophysical interactions, are essential to better understand the individual and collective functions of proteins as well as the biological processes underlying human diseases[501].

Our research group previously identified the ANKRD55 protein interactomes from the HEK293 cell line[377], and in the current project we aimed to extend and validate these prior data by expanding the number of cell types studied and enhancing detection of interacting proteins with the help of novel techniques. Another remarkable difference of this work was the use of synthetic mRNA for transient protein expression, which eliminates the risk of genomic insertion events presented by viral systems[502]. For a further increase in the purity and stability of the mRNA molecule, a cap-1 structure was added to the 5' end of the mRNA, based on previous methods[502].

Compared to primary cells, the use of established cell lines has the advantage of minimizing the possible effect of genetic variation, thus improving reproducibility. Due to the role of the *ANKRD55* gene as risk factor for MS, cell lines derived from astrocytes, microglia, neuroblastoma cells and monocytes were used to assess cell-specific effects on the ANKRD55 interactome composition in the context of a relevant proteome for neuroinflammation. Immortalized human fetal astrocytes-SV40 were utilized as the astrocytic cell line. This cell line has been employed previously in other proteomic studies[503,504]. The limited availability of human microglial cell lines makes poses an obstacle to using this cell type in research, but the new immortalized human microglia-SV40 cell line, derived from primary human microglial cells, has been demonstrated as a suitable model to investigate this cell type, as it maintains specific

microglial markers[505–507]. Neuroblastoma cells are represented in this work by the SH-SY5Y cell line, originally derived from a metastatic bone marrow biopsy of a neuroblastoma patient. This cell line is commonly used for *in vitro* experiments requiring neuron-like cells[508–510]. Finally, for the study of human monocytic cells, the THP-1 cell line was used. This immortalized monocyte-like cell line, originally isolated from the peripheral blood of a childhood patient suffering from acute monocytic leukemia, is considered a reliable model for *in vitro* studies of monocyte functions and signaling pathways[511,512].

4.2. Materials and methods

4.2.1. Cell lines and primary cell culture

Immortalized human astrocytes-SV40 (IMhu-A) (Applied Biological Materials, Cat. No. T0280) cells were grown in RPMI-HEPES medium (Sigma-Aldrich, Cat. No. R5886) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, Cat. No. F9665), 2 mM L-glutamine (Sigma-Aldrich, Cat. No. G7513) and 1% penicillin-streptomycin solution (Sigma-Aldrich, Cat. No. P4458) and were cultured in 0.1% collagen I (Sigma-Aldrich, Cat. No. C3867) pre-coated flasks. Immortalized human microglia-SV40 (IMhu-M) (Applied Biological Materials, Cat. No. T0251) cells were grown in high glucose Dulbecco's Modified Eagle's medium (DMEM) (Sigma-Aldrich, Cat. No. D5796) supplemented with 10% FBS, 2 mM L-glutamine and 1% penicillin-streptomycin solution or in IL-4/GM-CSF moDC differentiation medium (Miltenyi Biotec, Cat. No. 130-094-812) when indicated., and cultured in 0.1% collagen I pre-coated flasks. The SH-SY5Y (Sigma-Aldrich, Cat. No. 94030304) cell line was grown in DMEM/F-12 GlutaMAX™ supplement medium (Thermo Fisher Scientific, Cat. No. 10565-018) complemented with 10% FBS, 2 mM L-glutamine and 1% penicillin-streptomycin solution. THP-1 cells were maintained in RPMI-1640 medium (Sigma-Aldrich, Cat. No. R8758), complemented with 10% FBS, 2 mM L-glutamine and 1% penicillin-streptomycin solution, or cultured in moDC differentiation medium when indicated. HEK293 cells were maintained in DMEM medium (Thermo Fisher Scientific, Cat. No. 31885-023) complemented with 10% FBS, 2 mM L-glutamine and 1% penicillin-streptomycin solution. Jurkat cells were grown in RPMI-1640 medium with 10% FBS, 2 mM L-glutamine and 1% penicillin-streptomycin solution. All cell lines were maintained under standard cell culture conditions at 37 °C and 5% CO_2 in a humidified atmosphere. $CD14^+$ human monocytes were isolated from fresh PBMCs using CD14 MACS MicroBeads (Miltenyi Biotec, Cat. No. 130-050-201) and cultured in moDC differentiation medium at a density of 10^6 cells/ml for 6 days. Fresh medium was added on day 3.

4.2.2. Synthesis and transfection of *ANKRD55* mRNA

mRNA was synthesized encompassing the open reading frame of a fusion protein coding for the full-length *ANKRD55* isoform 201 coupled to C-terminal MYC-FLAG tags as provided by a commercial vector (Origene, Cat. No. RC221211). Unmodified synthesized *ANKRD55* mRNA transcript was capped at the 5' end using wild-type bases CleanCap® AG (TriLink

BioTechnologies, Cat. No. N-7113) to generate a natural cap-1 structure, which is known to reduce activation of pattern recognition receptors and yield more active mRNA. The RNA was subsequently polyadenylated (120 A). The mRNA was treated with DNase and phosphatase, purified via silica membrane adsorption and reconstituted in 1 mM sodium citrate (pH 6.4). Cells were transfected at 60–80% confluency with *ANKRD55* MYC-FLAG isoform 201 mRNA using Viromer® mRNA (Lipocalyx, Cat. No. VmR-01LB-00) following manufacturer's recommendations and incubated for 24 hours unless otherwise indicated.

4.2.3. RNA extraction and cDNA synthesis

A NucleoSpin RNA Kit (Macherey-Nagel, Cat. No. 740955) was used for total RNA isolation from cells following manufacturer's protocol. After measuring its concentration with a Nanodrop® 2000 spectrophotometer (Thermo Fisher Scientific), cDNA synthesis was performed with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Cat. No. 4368814) with the following temperature cycles: priming at 25 ºC for 10 minutes, reverse transcription at 37 ºC for 120 minutes and enzyme inactivation at 85 ºC for 5 minutes.

4.2.4. ddPCR and qPCR

Droplet digital PCR (ddPCR) was used to quantify to absolute gene expression. 30 ng of cDNA, ddPCR supermix for probes (no dUTP) (Bio-Rad, Cat. No. 1863023), droplet generation oil (Bio-Rad, Cat. No. 1863005) and specific primer pair for *ANKRD55* isoform 201 (Bio-Rad) were used to generate droplets in a QX200 Droplet Generator (Bio-Rad). Droplets were then amplified by PCR in a thermal cycler C1000 (Bio-Rad) following manufacturer's instructions, and fluorescence intensity was measured in a QX200 Droplet Reader (Bio-Rad). QuantaSoft™ software (Bio-Rad) was used for data analysis. Quantitative PCR (qPCR) was performed with 5–10 ng cDNA, Fast SYBR® Green Master Mix (Applied Biosystems, Cat. No. 4385612) and specific primer pairs (**Table 9**) in a 7500 Fast Real-Time PCR system (Applied Biosystems) following manufacturer's instructions. Analysis was performed using the $2^{-\Delta Ct}$ method. Samples were assessed in triplicate. No-template controls were included, and *ACTB* and *GAPDH* (**Table 9**) primers were used for normalization.

TABLE 9. Primers used in the study for qPCR (corresponding to Part II).

Gene	Company
ACTB_1	Qiagen (Cat. No. Hs_ACTB_1_SG)
GAPDH	IDT (Cat. No. Hs.PT.39a.22214836)
ANKRD55 (best coverage)	IDT (Cat. No. Hs.PT.58.27501603)

4.2.5. Western blot

Protein cell lysates, together with pre-stained protein ladder (Fisher Scientific, Cat. No. BP3603-500), were resolved on 10% SDS-PAGE, transferred to PVDF membranes (Merck Millipore, Cat. No. IPVH00010), blocked with 2% casein (Sigma-Aldrich; Cat. No. C5890) in tris-buffered saline (TBS) for 1 hour at room temperature and incubated overnight at 4 ºC with anti-ANKRD55 rabbit polyclonal (1:500; Sigma-Aldrich, Cat. No. HPA051049) or anti-FLAG (DDDDK tag) mouse monoclonal (1:800; Proteintech, Cat. No. 66008-2-Ig) primary antibody (**Table 10**), followed by incubation with HRP-conjugated donkey anti-rabbit (1:10000; Jackson ImmunoResearch, Cat. No. 711-035-152) or donkey anti-mouse (1:10000; Jackson ImmunoResearch, Cat. No. 715-035-150) secondary antibody (**Table 11**) for 1 hour at room temperature. Chemiluminescent signal was detected using enhanced chemiluminescence (ECL) substrate (Bio-Rad, Cat. No. 1705061).

4.2.6. Flow cytometry

Untransfected or transfected Jurkat cells were harvested, centrifuged and washed three times with PBS. Cells were fixed with 1% paraformaldehyde for 10 minutes at 37 ºC, permeabilized using ice-cold methanol for 30 minutes on ice, blocked with 3% BSA (Sigma-Aldrich, Cat. No. A9418) for 30 minutes at 37 ºC and incubated with FITC conjugated anti-FLAG mouse monoclonal antibody (GenScript, Cat. No. A01632) (**Table 10**) at a concentration of 2 µg per million cells for 1 hour at 37 ºC in the dark. The percentage of transfected cells was analyzed in a MACSQuant® flow cytometer (Miltenyi Biotec).

4.2.7. Immunofluorescence microscopy

Cell lines were cultured on coverslips pre-coated with poly-D-lysine (Sigma-Aldrich, Cat. No. P0899) and transfected as described above or left untransfected. After 24 hours, culture media was removed, and fixation and permeabilization were performed using ice-cold methanol for 5 minutes, followed by a blocking step with 1% BSA in for 30 minutes at room temperature. Cells were incubated with anti-ANKRD55 rabbit polyclonal (1:250; Human Protein Atlas, Cat. No. HPA 061649) primary antibody (**Table 10**) for 1 hour at room temperature and stained with AlexaFluor 594-conjugated goat anti-rabbit F(ab')2 (1:500; Thermo Fisher Scientific, Cat. No. A-11072) fluorescent secondary antibody (**Table 11**) for 1 hour at room temperature protected from light. 1 µg/ml DAPI (Sigma-Aldrich, Cat. No. D9542) was added for 5 minutes to stain nuclei and Fluoromount-G (Invitrogen, Cat. No. 00-4958-02) was used to mount coverslips on glass slides. Images were obtained using a Zeiss LSM 880 Airyscan microscope with a 63x immersion objective. The fluorescent signal was quantified with Fiji software[466] (https://imagej.net/software/fiji) and the ratio between ANKRD55 and region of interest (ROI) signals was calculated. Statistical analysis was performed with GraphPad Prism v.7 software and data are shown as the mean ± standard error of the mean (SEM). After applying a Shapiro–Wilk test to determine if the variables followed a normal distribution, a Mann–Whitney test or an unpaired *t*-test was used, considering values of $p \leq 0.05$ as significant (*), $p \leq 0.01$ as very significant (**), $p \leq 0.001$ as highly significant (***), $p \leq 0.0001$ as extremely significant (****) and $p > 0.05$ as not significant (n.s.).

TABLE 10. Primary antibodies used for western blotting, flow cytometry and immunofluorescence microscopy (corresponding to Part II).

Application	Primary antibody	Host	Company	Dilution or concentration
Western blot	α-ANKRD55	Rabbit (polyclonal)	Human Protein Atlas (Cat. No. HPA 061649)	1:500
	α-FLAG	Mouse (monoclonal)	Proteintech (Cat. No. 66008-2-Ig)	1:800
Flow cytometry	α-FLAG (FITC conjugated)	Mouse (monoclonal)	GenScript (Cat. No. A01632)	2 µg/10^6 cells
Immunofluorescence microscopy	α-ANKRD55	Rabbit (polyclonal)	Human Protein Atlas (Cat. No. HPA 061649)	1:250

TABLE 11. Secondary antibodies used for western blotting and immunofluorescence microscopy (corresponding to Part II).

Application	Secondary antibody	Host	Company	Dilution (concentration)
Western blot	HRP-conjugated anti-rabbit IgG H&L	Donkey (polyclonal)	Jackson ImmunoResearch (Cat. No. 711-035-152)	1:10000
	HRP-conjugated anti-mouse IgG H&L	Donkey (polyclonal)	Jackson ImmunoResearch (Cat. No. 715-035-150)	1:10000
Immunofluorescence microscopy	AlexaFluor 594-conjugated anti-rabbit F(ab')2	Goat (polyclonal)	Thermo Fisher Scientific (Cat. No. A-11072)	1:500 (4 µg/ml)

4.2.8. Affinity purification

Untransfected or transfected cells were harvested at 24 hours after transfection, either directly from flasks (THP-1 suspension cells) or using a rubber scraper (IMhu-A, IMhu-M, and SH-SY5Y adherent cells). Cells were then centrifuged for 5 minutes at 300 × g and washed three times with PBS. Protein extracts were obtained by incubation of cells with lysis buffer (50 mM NaH_2PO_4 [Sigma-Aldrich, Cat. No. S8282], 200 mM NaCl, 1% Triton™ X-100 [Sigma-Aldrich, Cat. No. T8787] and 1% cOmplete™ EDTA-free protease inhibitor cocktail [Roche, Cat. No. 11873580001], pH 8.0) for 45 minutes on ice, followed by centrifugation at 22.000 × g for 15 minutes at 4 ºC and recovery of supernatant. Protein concentration was measured with a Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, Cat. No. 23225) and, for each experiment, identical amounts of protein lysate from either untransfected or transfected cells was added to 20 µl of anti-DYKDDDDK G1 Affinity Resin (GenScript, Cat. No. L00432), previously washed with TBS 0.1% Tween™ 20 (TBS-T) and incubated overnight on a rotating wheel at 4 ºC. Resin was then washed three times with TBS-T, followed by three washes with TBS on ice, and elution of ANKRD55 complexes was performed with 100 µl of CLB buffer (2 M thiourea [Sigma-Aldrich, Cat. No. T7875], 7 M urea [PanReac AppliChem, Cat. No. A1049] and 4% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate [CHAPS; Sigma-Aldrich, Cat. No. 226947]). All experiments were independently replicated at least three times. Transfection and purification conditions were identical for all cell lines and replicates.

4.2.9. In-solution digestion and mass spectrometry analysis

Protein samples were digested in CLB buffer supplemented with 5 mM dithiothreitol (DTT; Fisher Scientific, Cat. No. 11896744) following the filter-aided sample preparation (FASP) protocol described by Wisniewski *et al.*[513] with minor modifications. Trypsin (Sigma-Aldrich, Cat. No. T1143) was added to generate a trypsin:protein ratio of 1:20, and the mixture was incubated overnight at 37 ºC, dried in an RVC 2-25 rotational vacuum concentrator (Martin Christ) and resuspended in 0.1% formic acid (Fisher Scientific, Cat. No. A117). Samples were analyzed in a timsTOF Pro with parallel accumulation serial fragmentation (PASEF®) (Bruker) coupled online to a nanoElute® liquid chromatograph (Bruker). 200 ng samples were directly loaded onto the Evosep One platform (Evosep) and resolved using 30-minute gradient runs. Database searching was performed using Mascot software 2.2.07 (Matrix Science) through Proteome Discoverer software 1.4 (Thermo Fisher Scientific) against a Uniprot/Swiss-Prot database[367] (https://www.uniprot.org) filled exclusively with entries corresponding to *Homo sapiens*. For protein identification, the following parameters were adopted: carbamidomethylation of cysteines (C) as fixed modification and oxidation of methionines (M) as variable modification, 20 ppm of peptide mass tolerance, 0.5 Da fragment mass tolerance and up to two missed cleavage points and peptide charges of +2 and +3. Relative quantification was carried out using a modified spectral counting method, Normalized Spectral Abundance Factor (NSAF)[514]. Briefly, protein spectral counts (the sum of all peptide identifications obtained for a certain protein) are corrected by protein length, yielding the Spectral Abundance Factor (SAF) for each protein. These SAF values are further normalized against the sum of all SAF values in a certain sample and expressed as a percentage of the total.

4.2.10. Protein ranking

Proteins identified in the interactomes were assigned to one out of four categories, defined as follows: Category A, proteins enriched with an NSAF ratio in transfected (T) versus untransfected control (C) cells (NSAF (T)/NSAF (C)) > 2 in at least two of three replicates (and absent in the third replicate, if not enriched) and identified with ≥ 2 unique peptides in at least two of three replicates; Category B, proteins enriched with NSAF (T)/NSAF (C) > 2 in at least two of three replicates (and absent in the third replicate, if not enriched) and identified with no unique peptide count requirements; Category C, proteins enriched with NSAF (T)/NSAF (C) > 2 in two replicates but failed to provide enrichment in the third replicate (present but not

enriched) and with no unique peptide count requirements; and Category D, proteins that failed to provide an enrichment in the third replicate when analyzing NSAF but did when spectral counting was analyzed. Within each category (A, B, C, D), proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged over three replicates. Proteins absent in untransfected control cells and exclusive in transfected cells in a single replicate were given an NSAF (T)/NSAF (C) value of 100, while proteins absent in both untransfected control and transfected cells in a single replicate were given an NSAF (T)/NSAF (C) value of 1.

4.2.11. RNA-seq transcriptomic analysis of IMhu-M cells

RNA samples were isolated using a NucleoSpin RNA extraction kit (Macherey-Nagel, Cat. No. 740955). The concentration and the purity of the final RNA solution were measured using a Nanodrop® One spectrophotometer (Thermo Fisher Scientific). The isolated RNA was quantified with a Qubit™ 4 fluorometer (Thermo Fisher Scientific) and RNA integrity was assessed with an Agilent RNA 6000 Nano kit (Aligent, Cat. No. 5067-1511) using an Agilent 2100 bioanalyzer (Agilent). The average A260/280 ratio was 1.9 (range 1.8-2.1), and the average RNA integrity ratio was between 8 and 10. Libraries were constructed with 750 pg of RNA using the NEBNext® Single Cell/Low input library preparation kit from Illumina® (New England Biolabs, Cat. No. E6420S). Library concentrations and fragment sizes were evaluated with an Agilent High Sensitivity DNA kit (Aligent, Cat. No. 5067-4626). Libraries were sequenced on an Illumina® NovaSeq™ 6000 sequencer (2 × 150 pb, paired-end) (Illumina) and generated an average of 300 million reads per sample. The FASTQ files were processed for quality control with the FastQC program (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Raw paired-end reads were mapped against the *Homo sapiens* reference genome GRCh38 using TopHat 2.1.0. Low quality reads were removed using the packages Samtools 1.2 and Picard tools 2.9.0, followed by transcript assembly and gene identification using Bayesian inference methods with Cufflinks v2.2.2. Pearson correlation analysis was used to determined correlation between IMhu-M RNA read counts (from RNA-seq data) and IMhu-M NSAF values (from interactome data).

4.2.12. Bioinformatics analysis

Proteins resulting from ANKRD55 interactomes were analyzed with the use of Ingenuity Pathway Analysis (IPA) software from Qiagen (https://digitalinsights.qiagen.com/IPA) to identify protein pathways. Fisher's exact test-adjusted *p*-values obtained from IPA analysis were used for data representation as heatmaps. The STRING database[515] (version 11.5) of known and predicted protein–protein interactions (https://string-db.org) was used to assign molecular and cellular functions to the combined interactome proteins as well as the non-14-3-3 and IFT interactants. The text mining option was omitted from the interaction source settings for STRING results, and false discovery rate (FDR)-corrected *p*-values were used to generate the corresponding heatmaps. A Venn diagram was created using the InteractiVenn[516] online tool (http://www.interactivenn.net), and heatmaps were designed with Microsoft Excel software.

4.3. Results

4.3.1. Expression of synthetic *ANKRD55* mRNA

In order to verify the expression of *ANKRD55*, absolute quantification of the naturally expressed *ANKRD55* full-length isoform 201 mRNA was performed in the untransfected cell lines IMhu-A, IMhu-M, SH-SY5Y and THP-1 using ddPCR. Low transcript copy numbers were observed in THP-1 (0.45 copies/ml) and IMhu-A (1 copy/ml) cells, but *ANKRD55* transcript copy numbers were higher in the other two cell lines, reaching 6.6 copies/ml in IMhu-M cells and 8.7 copies/ml in SH-SY5Y cells (**Figure 34**).

Using Viromer polymer nanoparticles, transfection scale and expression of synthetic *ANKRD55* mRNA were tested first in HEK293 cells, a cell line known for being easily transfectable. 74-kD protein bands, corresponding to the molecular weight of the full-length protein isoform ANKRD55-201 (69 kDa) with added C-terminal MYC-FLAG tag that reacted with antibodies to ANKRD55 and FLAG, were observed in all conditions tested (**Figure 35A**). All further experiments were routinely performed using 1 µg of *ANKRD55* mRNA per ml of culture media for adherent cells (IMhu-A, IMhu-M and SH-SY5Y) or 2 µg per ml of culture media for suspension cells (THP-1) according to the manufacturer's protocol; both concentrations corresponded to a transfection scale of 1x. The transfection method was efficient in all four cell lines, at both the mRNA and protein level. Measurements by ddPCR resulted in highly increased *ANKRD55* transcript copy number counts per ml post-transfection (**Figure 34**). Additionally, in western blots of the cell lysates of the IMhu-A, IMhu-M, SH-SY5Y and THP-1 cell lines, 74-kD ANKRD55 bands appeared in the transfected cells, and the strongest signal was produced in the IMhu-M (microglial) cell line. Both transfected and control cells also expressed the native 69-kD ANKRD55 protein. In THP-1 suspension cells, cell density affected transfection or expression levels of ANKRD55, with the highest signal observed for intermediate density cultures of 1×10^6 THP-1 cells per well (**Figure 35B**). Moreover, using IF microscopy, native ANKRD55 was detected in the four untransfected cell lines (**Figure 36**), and an increased signal was observed 24 hours post-transfection. Quantitative analysis of IF images showed that the ANKRD55/DAPI fluorescence signal was significantly higher in the four transfected cell lines compared to the control cells, with the strongest difference found in the IMhu-M cell line ($p < 0.0001$).

The T-lymphocytic cell line Jurkat was also tested for ANKRD55 transfection with this approach, but recombinant ANKRD55 protein was undetectable in transfected cells via sensitive flow cytometry (**Figure 37A**), discarding this cell line for subsequent interactome analyses.

Moreover, based on results from Part I of this thesis showing that *ANKRD55* is induced in primary $CD14^+$ monocytes during differentiation into immature dendritic cells in the presence of IL-4/GM-CSF, we aimed to test whether this cytokine treatment affected native *ANKRD55* expression levels in THP-1 and IMhu-M cell lines. However, *ANKRD55* expression was not affected by this condition (**Figure 37B**), so further interactome analysis from these cell lines was performed in untreated cells.

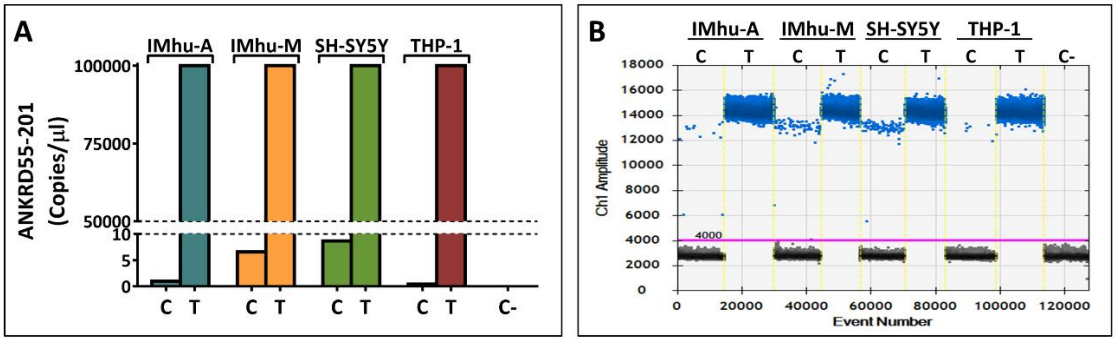


FIGURE 34. Absolute quantification of native and synthetic *ANKRD55* mRNA using ddPCR.
(A) *ANKRD55-201* mRNA copies per µl were quantified in untransfected control (C) and transfected (T) immortalized human fetal astrocytes (IMhu-A), immortalized human microglia (IMhu-M), neuroblastoma cells (SH-SY5Y) and monocytic cells (THP-1). Data from one representative experiment are shown. **(B)** Number of events (occurrences of each amplitude) are presented. Positive events are data points in blue above the pink horizontal threshold line, and negative events are those in black below the threshold line.

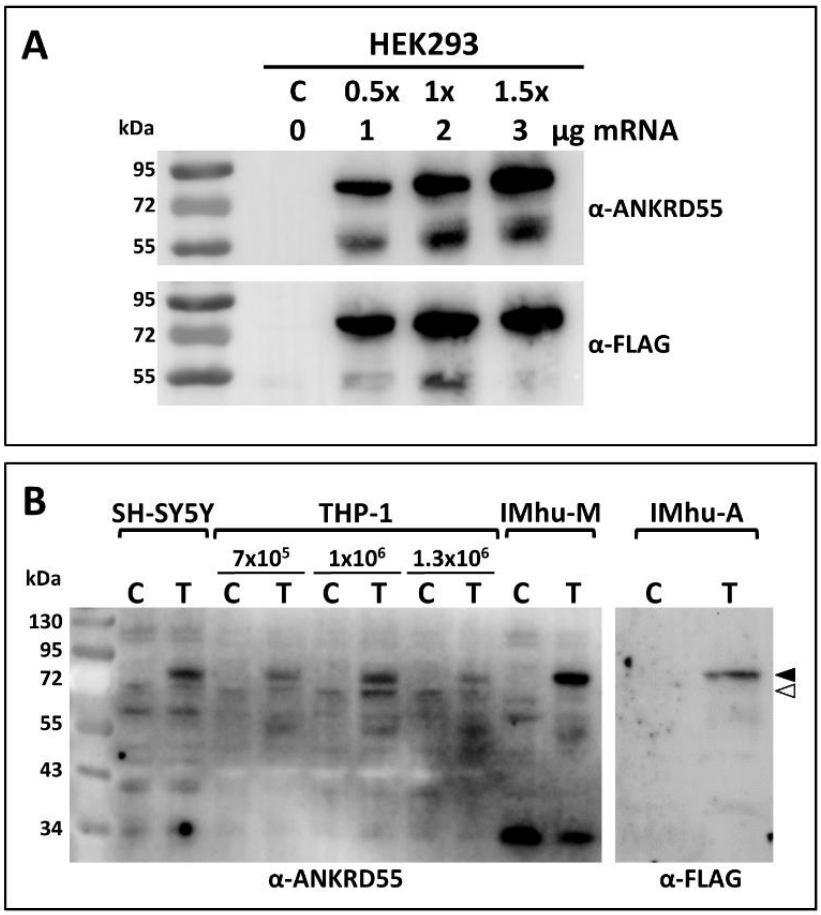
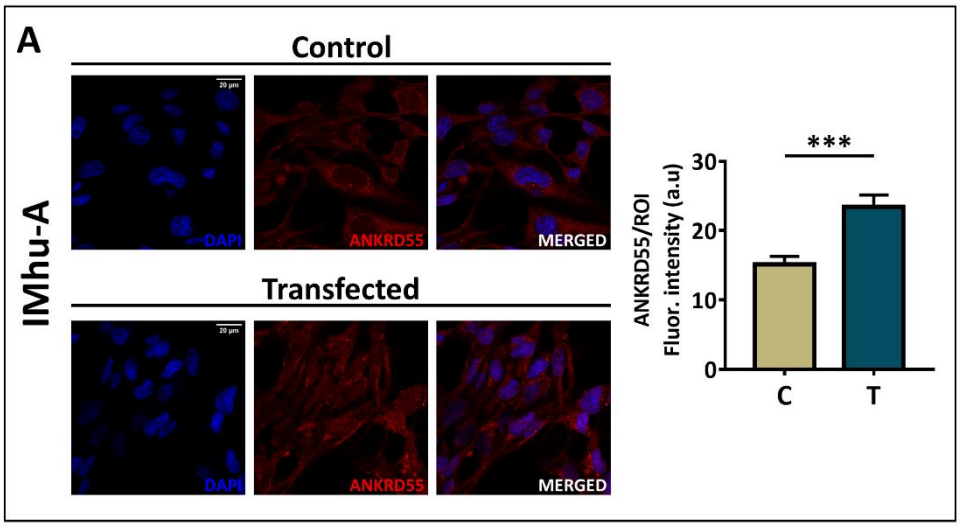
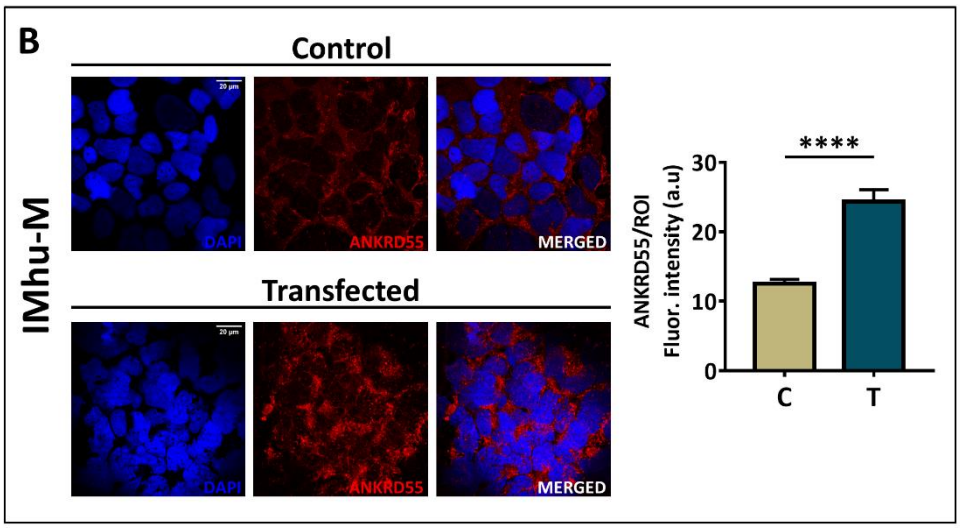
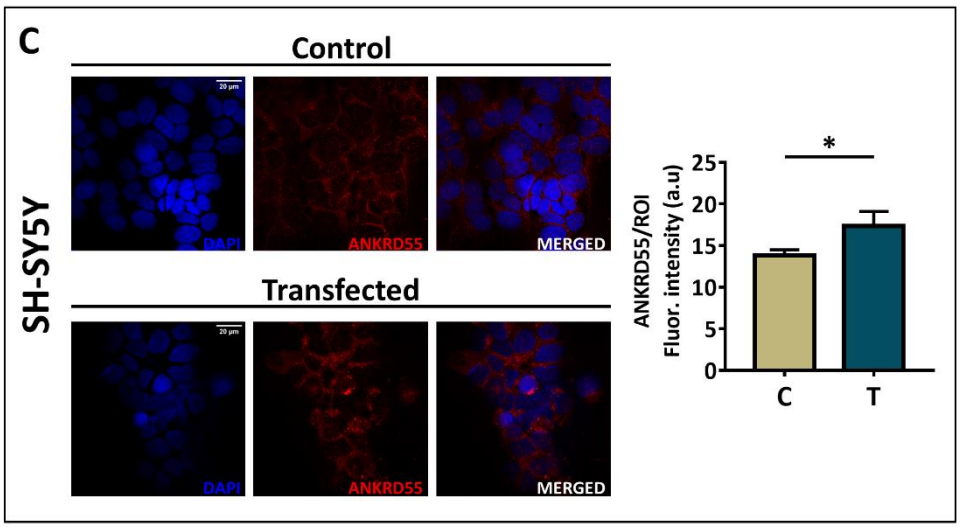


FIGURE 35. Western blot detection of endogenous and synthetic mRNA-expressed ANKRD55 protein.

(A) Reactivity in a western blot of recombinant ANKRD55 protein with anti-ANKRD55 and anti-FLAG antibodies in lysates from untransfected (C) and transfected HEK293 cells. Cells were cultivated in 6-well plates containing 2 ml of culture media, and 1 µg, 2 µg and 3 µg of ANKRD55 mRNA (corresponding to transfection scales of 0.5x, 1x and 1.5x, respectively) were tested. **(B)** Western blot analysis of native (white arrowhead) and recombinant (black arrowhead) ANKRD55 protein in untransfected (C) and transfected (T) cells of the four target cell lines. Equal amounts of cell lysate (10 µg) were loaded in each lane. Numbers above THP-1 lanes indicate the density of cells tested per well.


A
IMhu-A
Control
Transfected
DAPI
ANKRD55
MERGED
ANKRD55/ROI
Fluor. intensity (a.u)
***
30
20
10
0
C
T


B
IMhu-M
Control
Transfected
DAPI
ANKRD55
MERGED
ANKRD55/ROI
Fluor. intensity (a.u)
****
30
20
10
0
C
T


C
SH-SY5Y
Control
Transfected
DAPI
ANKRD55
MERGED
ANKRD55/ROI
Fluor. intensity (a.u)
*
25
20
15
10
5
0
C
T

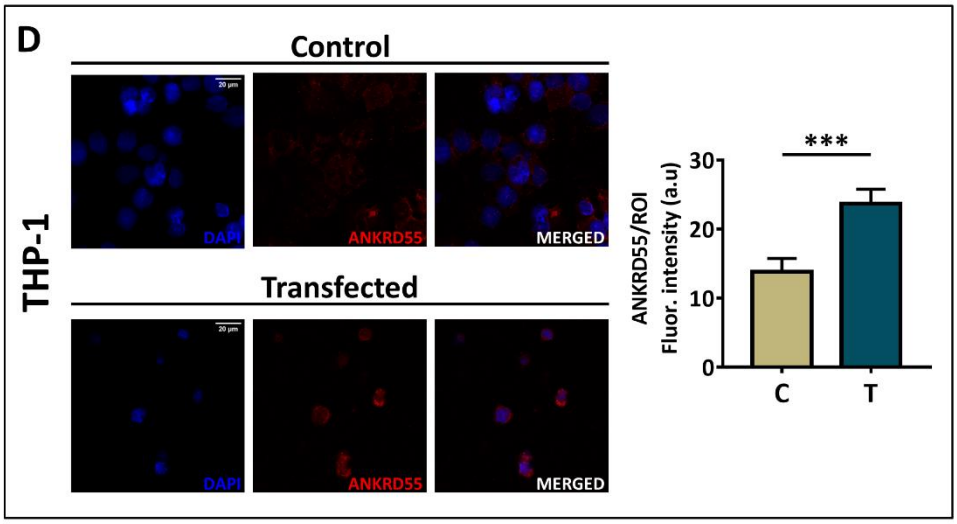


FIGURE 36. Immunofluorescence detection of native (control) and recombinant (transfected) ANKRD55 protein.

(A) IMhu-A, **(B)** IMhu-M, **(C)** SH-SY5Y (all in previous page) and **(D)** THP-1 cells were transfected with synthetic *ANKRD55* mRNA or left untransfected (control), fixed 24 hours later and analyzed by immunofluorescence confocal microscopy (1024 × 1024 pixels = 0.131 microns/pixel). ANKRD55 (red) and nuclei (DAPI; blue) were detected. Graphs on the right show the mean ± SEM of the fluorescence intensity, expressed in arbitrary units (a.u.), of the ratio cellular ANKRD55/regions of interest (ROI) in control (C) and transfected (T) cell lines. $n \geq 15$ ROIs per group (C or T). A Mann–Whitney test was used. * $p \leq 0.05$, *** $p \leq 0.005$, **** $p \leq 0.0001$.

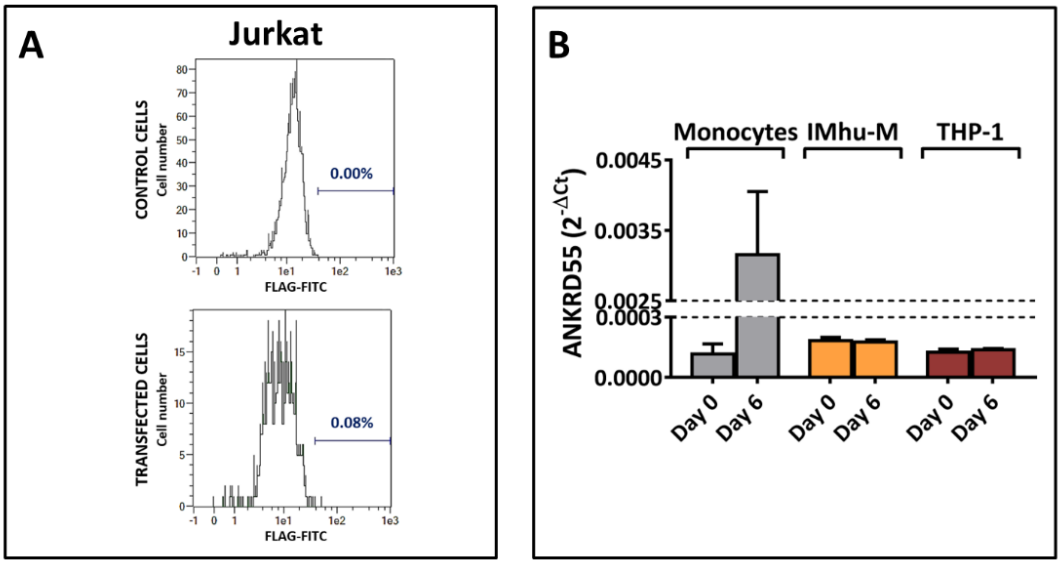


FIGURE 37. Absence of ANKRD55 expression in Jurkat cells and lack of *ANKRD55* induction in IMhu-M and THP-1.

(A) Lack of detection of ANKRD55 protein expression in transfected cells (lower panel) of the T-lymphocytic cell line Jurkat compared to untransfected cells (upper panel). Measurements were performed using flow cytometry. **(B)** Induction of *ANKRD55* mRNA in primary $CD14^+$ monocytes, but not in IMhu-M and THP-1, following a 6-day incubation in moDC differentiation medium containing IL-4/GM-CSF. Mean ± SEM of $2^{-\Delta Ct}$ values from two qPCR measurements is shown. *ACTB* and *GAPDH* were used as housekeeping genes.

4.3.2. Proteins shared by ANKRD55 interactomes

Proteins present in the ANKRD55 and control interactomes were identified as described in the "Materials and methods" from Part II (point 4.2.9). To rank identified proteins, three parameters were considered: enrichment, specificity and abundance of protein detection. The degree of enrichment was represented by the ratio of normalized spectral abundance factor (NSAF) of a protein in the interactomes from transfected to untransfected cells. A higher number of unique peptides was considered to enhance the reliability of the identified proteins, and overall NSAF values represented peptide abundance in the sample. Four categories (A, B, C, D; see "Materials and methods" from Part II [point 4.2.10]) were devised, and, within each category, proteins were ranked according to the average NSAF transfected/control ratio of three independent replicates from highest to lowest.

The total number of identified proteins in ANKRD55 interactomes varied according to cell line, ranging from 128 unique proteins detected in SH-SY5Y cells to 25 in IMhu-A (astrocytes) (**Figure 38**). In the four cell lines, proteins belonging to categories A and B accounted for more than half of the total proteins in the interactome. The 20 highest-ranking proteins identified in the ANKRD55 interactomes are represented in **Tables 12-15**, and additional identified proteins are listed in **Supplementary Tables 1-4**. Subsequent bioinformatic analysis was performed separately on proteins in the combined AB category (proteins identified according to the most stringent criteria) or ABCD category (all identified proteins). Considering proteins categorized as A or B (AB), two proteins, 14-3-3η (encoded by the gen *YWHAH*) and 14-3-3β/α (*YWHAB*), were found to be shared among the ANKRD55 interactomes of all four cell lines (**Figure 39**). By relaxing the ranking criteria to additionally include category C and D proteins (ABCD), an additional member of the 14-3-3 isoform family was added to this group, 14-3-3θ (*YWHAQ*). Three further proteins were shared by the IMhu-A, IMhu-M and SH-SY5Y cell lines, but absent in THP-1. These were 14-3-3γ (*YWHAG*) and BAG cochaperone 2 (*BAG2*), both considered as category AB proteins, as well as 14-3-3ε (*YWHAE*), a category ABCD protein. No further proteins were shared by any other three-cell line combinations, though several more proteins were reproduced between pairs of cell lines. The IMhu-M and SH-SY5Y ANKRD55 interactomes had a further four unique AB proteins in common: CAD, a trifunctional multi-domain enzyme involved in the first three steps of pyrimidine biosynthesis; the intraflagellar transport proteins IFT74 and IFT22; and the ribosomal 60S subunit protein L32 (*RPL32*). IMhu-A and IMhu-M shared the sixth 14-3-3 isoform identified as an ANKRD55 interactant in this study, named 14-3-3ζ/δ (*YWHAZ*). Two category AB mitochondrial membrane proteins, the ATPase subunit d (*ATP5PD*) and the

glutamate carrier 1 (*SLC25A22)*, were shared by SH-SY5Y and THP-1 cells. Moreover, IMhu-M and THP-1 shared the category AB 40S ribosomal protein S28 (*RPS28*), and the IMhu-A cell line shared the cullin-associated and neddylation-dissociated protein 1 (*CAND1*) with the THP-1 line. Despite these shared proteins, the majority of the ANKRD55 interactome proteins were cell line-specific, ranging from 50% in IMhu-A to 78% in SH-SY5Y for category ABCD proteins.

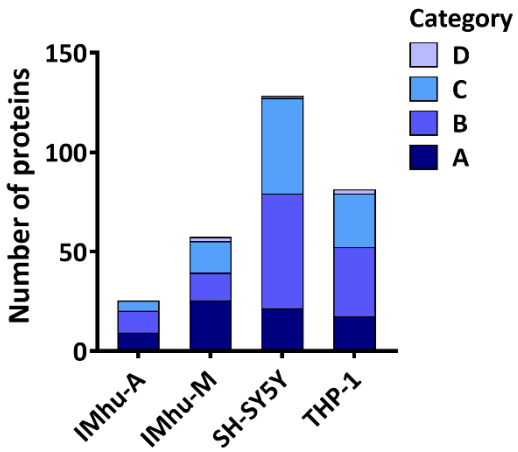


FIGURE 38. Number of proteins corresponding to prioritization categories A–D identified by FASP-based mass spectrometry of affinity-purified ANKRD55 interactomes.

Three independently replicated experiments were performed for each cell line. Category requirements are detailed as followed:

Category A: Proteins enriched with NSAF (T)/NSAF (C) > 2 in at least two of three replicates (and absent in the third replicate, if not enriched) and identified with ≥ 2 unique peptides in at least two of three replicates.

Category B: Proteins enriched with NSAF (T)/NSAF (C) > 2 in at least two of three replicates (and absent in the third replicate, if not enriched) and identified with no unique peptide count requirements.

Category C: Proteins enriched with NSAF (T)/NSAF (C) > 2 in two replicates but failed to provide such enrichment in the third replicate (present but not enriched) and with no unique peptide count requirements.

Category D: Proteins that failed to provide an enrichment in the third replicate when analyzing NSAF but did when spectral counting was analyzed.

TABLE 12. Top 20 proteins identified in the IMhu-A interactome.
Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

IMhu-A			
Gene symbol	**Average ratio NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Category**
ANKRD55	100	3	A
YWHAG	100	2	A
YWHAZ	70.31	2	A
VIM	68.36	3	A
YWHAH	67	2	A
YWHAQ	40.34	3	A
CAND1	34.55	2	A
ARF4	34.48	2	A
RPL3	34.35	2	A
YWHAB	67	1	B
PRDX2	67	1	B
BAG2	67	1	B
ARL8A	67	0	B
ELAVL1	67	1	B
TNFAIP6	67	0	B
NDUFA10	67	0	B
FASN	67	0	B
SLC16A3	35.19	1	B
RPL28	34.85	1	B
COPA	34.58	1	B

TABLE 13. Top 20 proteins identified in the IMhu-M interactome.
Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

IMhu-M			
Gene symbol	**Average ratio NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Category**
ANKRD55	100	3	A
YWHAB	100	3	A
YWHAH	100	3	A
YWHAG	100	3	A
RPS27L	100	3	A
IFT74	100	3	A
IFT22	100	2	A
IFT81	100	3	A
IFT52	100	2	A
PRKDC	100	2	A
CAD	100	2	A
YWHAQ	73.64	3	A
YWHAZ	70.68	3	A
BAG2	68.17	2	A
ATAD3B	67.35	3	A
HSPA2	67	2	A
IFT70A	67	2	A
IFT56	67	2	A
XRCC6	67	2	A
EIF2S2	67	2	A

TABLE 14. Top 20 proteins identified in the SH-SY5Y interactome.
Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

SH-SY5Y			
Gene symbol	**Average ratio NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Category**
YWHAB	100	3	A
YWHAH	100	3	A
SLC25A10	100	2	A
CAD	100	3	A
ANKRD55	84.95	3	A
ATP5F1C	71.53	3	A
ATP5PD	68.77	3	A
CCT6A	67.42	2	A
RAB1B	67	2	A
SLC25A22	67	2	A
RTL1	44.72	3	A
SLC25A11	38.06	3	A
VDAC1	35.32	3	A
ATAD3A	34.79	2	A
ACADM	34.79	2	A
PDHB	34.66	2	A
PSMC6	34.42	2	A
YWHAG	5.76	3	A
YWHAE	5.24	3	A
YWHAQ	4.56	3	A

TABLE 15. Top 20 proteins identified in the THP-1 interactome.
Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

THP-1			
Gene symbol	**Average ratio NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Category**
H2BC11	100	3	A
ILF3	100	3	A
ANKRD55	71.11	2	A
RPS19	67.42	2	A
ATP5PD	67	2	A
DPM1	36.18	2	A
HADHB	34.83	2	A
MT-CO2	34.80	2	A
COLGALT1	34.80	2	A
RPN2	34.50	2	A
SDHA	34.42	2	A
UQCRC2	4.89	2	A
ATP5MK	3.47	2	A
RPL13A	2.35	2	A
DARS1	2.30	2	A
SRP14	2.23	2	A
RPS26	2.23	2	A
NACA	100	1	B
YWHAH	67	1	B
NUDT21	67	1	B

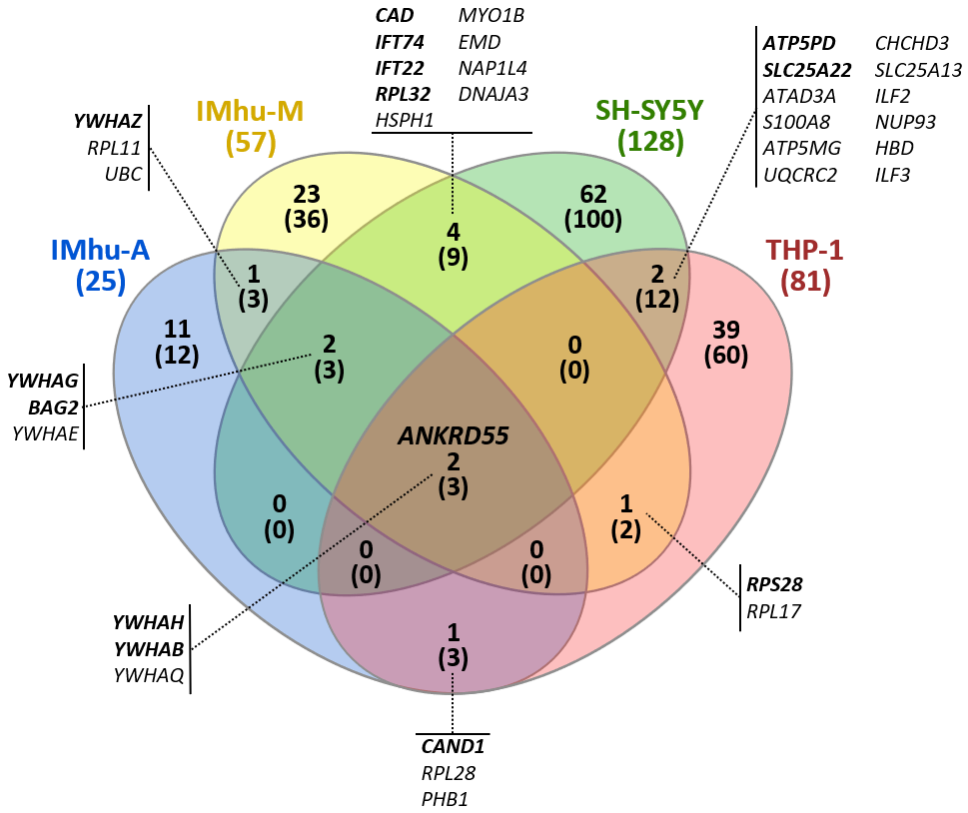


FIGURE 39. Venn diagram representing proteins shared among ANKRD55 interactomes.
ANKRD55 interactome proteins shared between two, three or all four cell lines over three independent replicates. Number of shared proteins belonging to category AB are shown in bold, with total number of shared proteins (ABCD) indicated below in brackets. Gene symbols of the shared proteins are provided (category AB proteins, italic and bold; category CD proteins, italic).

4.3.3. Pathways and functions of ANKRD55 interactome proteins

Qiagen IPA software was used to identify canonical pathways and biofunctions associated with ANKRD55 interactome proteins. Heatmaps of -log *p*-values of canonical pathways, as well as diseases and biofunctions, are provided for each cell line in **Figure 40** (red panels). Top-ranked pathways were shared by the IMhu-A, IMhu-M and SH-SY5Y cell lines. The canonical pathway term "Cell Cycle: G2/M DNA Damage Checkpoint Regulation" emerged with high significance from these three cell lines but with lower significance from THP-1 (*p*-values of 1.14×10^{-9}, 2.19×10^{-12}, 7.89×10^{-7} and 6×10^{-3}, respectively). Similar observations were made for the terms "HIPPO signaling", "ERK5 signaling", "IGF-1 signaling" and "Protein kinase A signaling". Different 14-3-3 isoforms were highly enriched in the top-ranked pathways of the IMhu-A, IMhu-M and SH-SY5Y cell lines. Blue heatmaps in **Figure 40 A–D** represent diversity and relative enrichment of 14-3-3 isoforms in the protein networks associated with each pathway. For instance, the "Cell Cycle: G2/M DNA Damage Checkpoint Regulation" pathway emerging

from category AB proteins includes five distinct 14-3-3 isoforms and no other proteins in IMhu-A and SH-SY5Y, six 14-3-3 isoforms in addition to PRKDC in IMhu-M and two 14-3-3 isoforms as only associated network proteins in THP-1 interactomes.

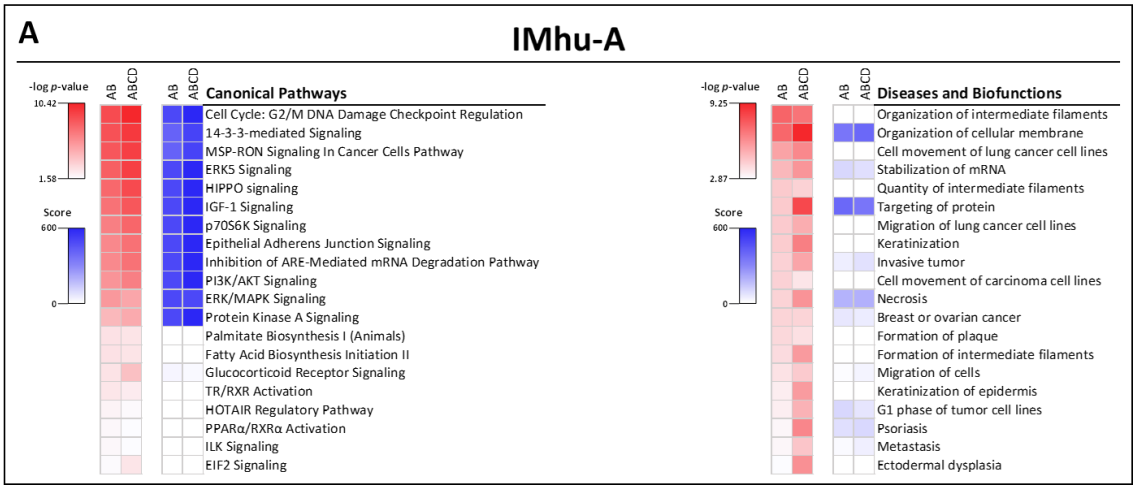
With the aim of relating the interactome proteins to their corresponding gene expressions, an RNA-seq transcriptomic analysis from the transfected and untransfected IMhu-M cells was performed. The plot in **Figure 44A** shows that the RNA-seq mRNA read counts of ANKRD55 interactome proteins were positively correlated with their corrected interactome NSAF values (Pearson's $r = 0.397$, p-value = 0.002). However, the correlation was entirely driven by the six 14-3-3 isoforms (**Figure 44C**; Pearson's $r = 0.942$; p-value = 0.005), the most abundant proteins in the interactome, and it trended towards negative in their absence (**Figure 44B** Pearson's $r = -0.253$, p-value = 0.076). 14-3-3 proteins may bind to ANKRD55 via its predicted SER436 and SER597 14-3-3-binding phophosites (**Figure 12**), though phospho-independent 14-3-3 protein–protein interactions are known to occur[517].

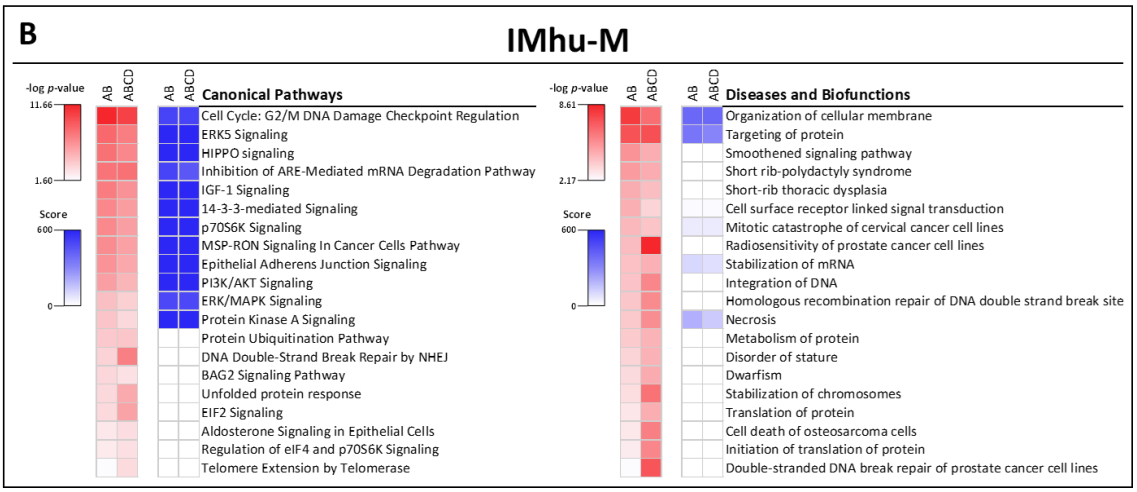
Canonical pathways from which 14-3-3 proteins were absent had lower significance levels and were less frequently shared among cell lines (**Figure 40**). One of these, "Mitochondrial Dysfunction", the top THP-1 canonical pathway (**Figure 40D**) (p-value for AB proteins: 4.33×10^{-9}; genes coding for protein network: *ATP5MG*, *ATP5PD*, *CPT1A*, *NDUFA4*, *NDUFB4*, *SDHA*, *UQCRC2* and *VDAC3*) was also identified in SH-SY5Y cells (**Figure 40C**) (p-value for AB proteins: 2.75×10^{-5}; genes coding for protein network: *ATP5F1C*, *ATP5PB*, *ATP5PD*, *CAT*, *NDUFA13* and *VDAC1*). Only ATP5PD, corresponding to mitochondrial ATP synthase subunit d, was shared between these networks. However, additional discrete subunits of the mitochondrial ATP synthase complex[518] were identified, and diverse mitochondrial NADH-Ubiquinone Oxidoreductase respiratory chain complex proteins (NDUF) and voltage-dependent anion channel (VDAC) isoforms were also present. Functional enrichment analysis using the STRING database was performed following the removal of all 14-3-3 isoform proteins from the interactomes to search for patterns in the remaining protein sets. "Acetylation" (UniProt Keyword) was the most enriched term in the combined interactomes from the four cell lines (**Figure 41**; FDR $p = 2.71 \times 10^{-32}$ in AB and 1.2×10^{-53} in ABCD protein categories), as well as in the individual SH-SY5Y and THP-1 interactomes (**Figure 42**). The terms "Extracellular exosome" and "RNA binding" were found in the individual interactomes of three cell lines when considering only category AB proteins, but they were present in the four cell lines with the inclusion of category C and D proteins (**Figure 42**; AB protein FDR $p < 4.7 \times 10^{-2}$; ABCD protein FDR $p < 3.3 \times 10^{-2}$). Compartment terms referring to organelle membrane/envelope and mitochondrial membrane/envelope were significantly enriched only in the SH-SY5Y and THP-1

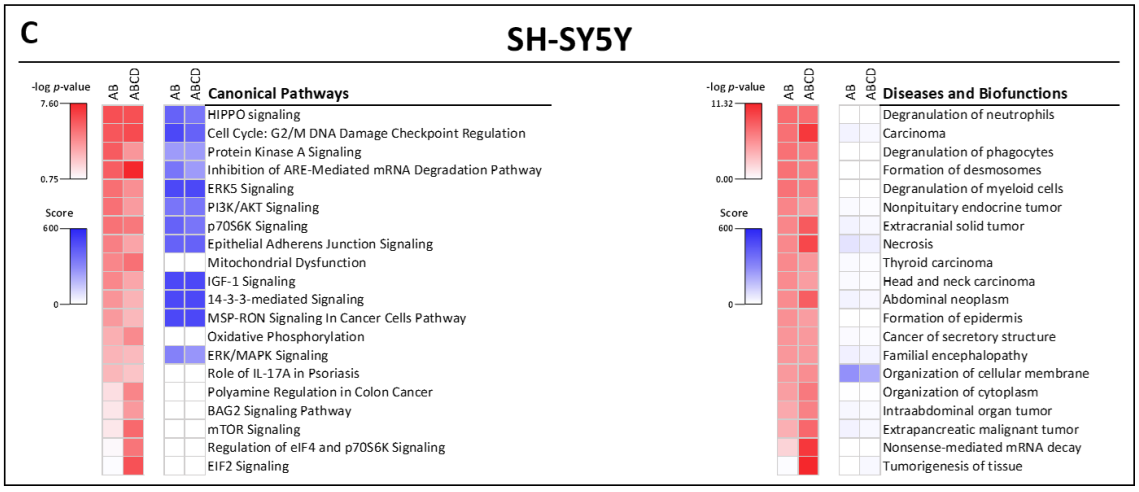
interactomes (**Figure 42**; AB protein FDR $p < 7.5 \times 10^{-3}$ and $< 1.36 \times 10^{-5}$, respectively), in line with the "Mitochondrial Dysfunction" pathway term identified in the 14-3-3-free Qiagen analysis (**Figure 40C, D**).

Interestingly, GO Component "Intraciliary transport particle b" (GO:0030992) emerged as the most enriched term in the non-14-3-3 ANKRD55 interactome of IMhu-M (category AB FDR $p = 6.7 \times 10^{-14}$) but was not identified in any other cell line (**Figure 42**). The IMhu-M interactome protein network of this pathway contained eight category AB proteins: IFT74, IFT22, IFT81, IFT52, TTC30A (also named IFT70A), TTC26 (also named IFT56), IFT46 and IFT27, corresponding to all the IFT proteins identified in the IMhu-M interactome (**Table 13** and **Supplementary Table 2**). The inclusion of category CD proteins to this pathway did not increase the strength of association (**Figure 42**; category ABCD FDR $p = 3.0 \times 10^{-12}$). In the Qiagen IPA IMhu-M ANKRD55 interactome analysis (**Figure 40B**), diseases or biofunctions (for category AB proteins) assigned to members of this group of IFT proteins included the following: the smoothened branch of the Hedgehog pathway for signaling across the membrane (IFT27, IFT46, IFT52, IFT81 and TTC26), cell surface receptor-linked signal transduction (IFT27, IFT46, IFT52, IFT81 and TTC26, in addition to UBC and YWHAZ), the skeletal ciliopathies short-rib polydactyly syndrome and thoracic dysplasia (IFT52, IFT74 and IFT81), disorder of stature (IFT52, IFT74 and IFT81, in addition to DSP and XRCC6), and dwarfism (IFT52, IFT74 and IFT81, in addition to XRCC6). Although two of these proteins, IFT22 and IFT74, were also identified in the SH-SY5Y interactome, further analysis with the STRING database did not uncover shared informative categories beyond GO terms reflecting moderately enriched basic processes (**Figure 43**).

No correlation was found between IMhu-M RNA-seq read counts of the identified IFTs and their individual corrected NSAF values in the ANKRD55 interactome (**Figure 44D**; Pearson's $r = -0.372$; p-value = 0.364). In contrast to the 14-3-3 proteins (NSAF [T] – NSAF [C] average range = [1.28 – 4.39]), IFT proteins interacted with ANKRD55 over a lower and narrower normalized spectral abundancy interval (NSAF [T] – NSAF [C] average range = [0.06 – 0.28]). Correlation analysis was also performed on a group of four proteins present in the IMhu-M ANKRD55 interactome that, according to the BioGRID repository of interactions (https://thebiogrid.org), constitute a protein complex. These are DNA-PK, PARP-1, Ku80 and Ku70, encoded by the genes *PRKDC*, *PARP1*, *XRCC5* and *XRCC6*, respectively (**Table 13, Supplementary Table 2,** and **Figures 44A, B**); and have been demonstrated to participate in DNA repair. The RNA-seq read counts of these genes and corrected NSAF values of the corresponding proteins were not significantly correlated (Pearson's $r = 0.639$; p-value = 0.361).


A
IMhu-A
-log p-value
10.42
1.58
Score
600
0
AB
ABCD
Canonical Pathways
Cell Cycle: G2/M DNA Damage Checkpoint Regulation
14-3-3-mediated Signaling
MSP-RON Signaling In Cancer Cells Pathway
ERK5 Signaling
HIPPO signaling
IGF-1 Signaling
p70S6K Signaling
Epithelial Adherens Junction Signaling
Inhibition of ARE-Mediated mRNA Degradation Pathway
PI3K/AKT Signaling
ERK/MAPK Signaling
Protein Kinase A Signaling
Palmitate Biosynthesis I (Animals)
Fatty Acid Biosynthesis Initiation II
Glucocorticoid Receptor Signaling
TR/RXR Activation
HOTAIR Regulatory Pathway
PPARα/RXRα Activation
ILK Signaling
EIF2 Signaling
-log p-value
9.25
2.87
Score
600
0
Diseases and Biofunctions
Organization of intermediate filaments
Organization of cellular membrane
Cell movement of lung cancer cell lines
Stabilization of mRNA
Quantity of intermediate filaments
Targeting of protein
Migration of lung cancer cell lines
Keratinization
Invasive tumor
Cell movement of carcinoma cell lines
Necrosis
Breast or ovarian cancer
Formation of plaque
Formation of intermediate filaments
Migration of cells
Keratinization of epidermis
G1 phase of tumor cell lines
Psoriasis
Metastasis
Ectodermal dysplasia


B
IMhu-M
-log p-value
11.66
1.60
Score
600
0
AB
ABCD
Canonical Pathways
Cell Cycle: G2/M DNA Damage Checkpoint Regulation
ERK5 Signaling
HIPPO signaling
Inhibition of ARE-Mediated mRNA Degradation Pathway
IGF-1 Signaling
14-3-3-mediated Signaling
p70S6K Signaling
MSP-RON Signaling In Cancer Cells Pathway
Epithelial Adherens Junction Signaling
PI3K/AKT Signaling
ERK/MAPK Signaling
Protein Kinase A Signaling
Protein Ubiquitination Pathway
DNA Double-Strand Break Repair by NHEJ
BAG2 Signaling Pathway
Unfolded protein response
EIF2 Signaling
Aldosterone Signaling in Epithelial Cells
Regulation of eIF4 and p70S6K Signaling
Telomere Extension by Telomerase
-log p-value
8.61
2.17
Score
600
0
Diseases and Biofunctions
Organization of cellular membrane
Targeting of protein
Smoothened signaling pathway
Short rib-polydactyly syndrome
Short-rib thoracic dysplasia
Cell surface receptor linked signal transduction
Mitotic catastrophe of cervical cancer cell lines
Radiosensitivity of prostate cancer cell lines
Stabilization of mRNA
Integration of DNA
Homologous recombination repair of DNA double strand break site
Necrosis
Metabolism of protein
Disorder of stature
Dwarfism
Stabilization of chromosomes
Translation of protein
Cell death of osteosarcoma cells
Initiation of translation of protein
Double-stranded DNA break repair of prostate cancer cell lines


C
SH-SY5Y
-log p-value
7.60
0.75
Score
600
0
AB
ABCD
Canonical Pathways
HIPPO signaling
Cell Cycle: G2/M DNA Damage Checkpoint Regulation
Protein Kinase A Signaling
Inhibition of ARE-Mediated mRNA Degradation Pathway
ERK5 Signaling
PI3K/AKT Signaling
p70S6K Signaling
Epithelial Adherens Junction Signaling
Mitochondrial Dysfunction
IGF-1 Signaling
14-3-3-mediated Signaling
MSP-RON Signaling In Cancer Cells Pathway
Oxidative Phosphorylation
ERK/MAPK Signaling
Role of IL-17A in Psoriasis
Polyamine Regulation in Colon Cancer
BAG2 Signaling Pathway
mTOR Signaling
Regulation of eIF4 and p70S6K Signaling
EIF2 Signaling
-log p-value
11.32
0.00
Score
600
0
Diseases and Biofunctions
Degranulation of neutrophils
Carcinoma
Degranulation of phagocytes
Formation of desmosomes
Degranulation of myeloid cells
Nonpituitary endocrine tumor
Extracranial solid tumor
Necrosis
Thyroid carcinoma
Head and neck carcinoma
Abdominal neoplasm
Formation of epidermis
Cancer of secretory structure
Familial encephalopathy
Organization of cellular membrane
Organization of cytoplasm
Intraabdominal organ tumor
Extrapancreatic malignant tumor
Nonsense-mediated mRNA decay
Tumorigenesis of tissue

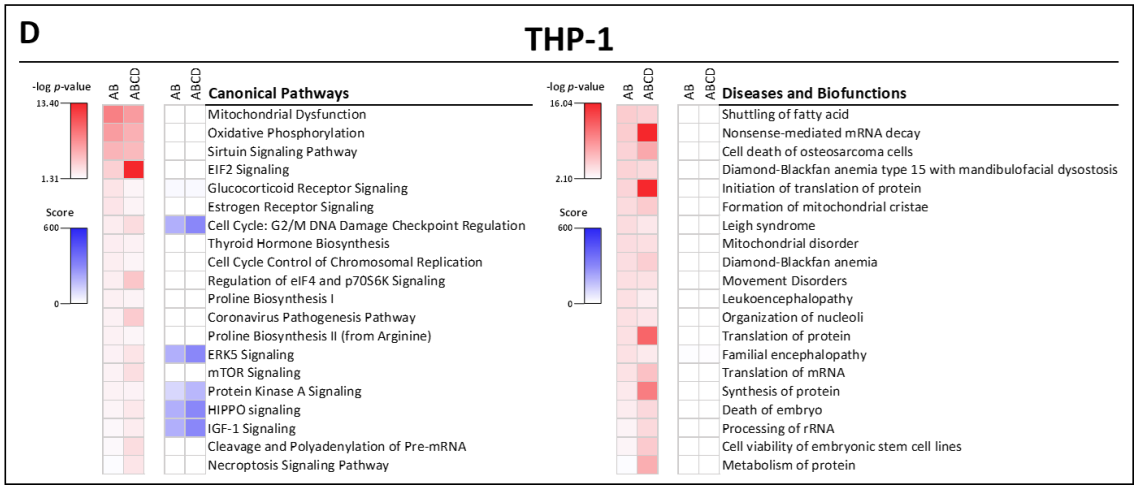


FIGURE 40. Ingenuity pathway analysis (IPA) of ANKRD55 interactomes.

Protein data from **(A)** IMhu-A, **(B)** IMhu-M, **(C)** SH-SY5Y (all in previous page) and **(D)** THP-1 cell interactomes were analyzed using Qiagen IPA software. Red heatmaps provide the significance (-log *p*-value) of canonical pathways and of diseases and biofunctions detected in category AB and ABCD interactomes, ranked following the significance levels (high to low) found in the category AB proteins. Blue heatmaps represent the diversity and relative enrichment of 14-3-3 isoforms in the identified protein network assigned to canonical pathways or diseases and biofunctions in each cell line. The score of the blue heatmaps was calculated by multiplying the number of distinct 14-3-3 isoform proteins present in the pathway or biofunction with the % of 14-3-3 proteins in the associated protein network. The maximum score of 600 corresponds to the presence of six diverse 14-3-3 isoforms and no other proteins and is seen in nine of the 12 top ranked pathways in Imhu-M category AB and ABCD proteins.

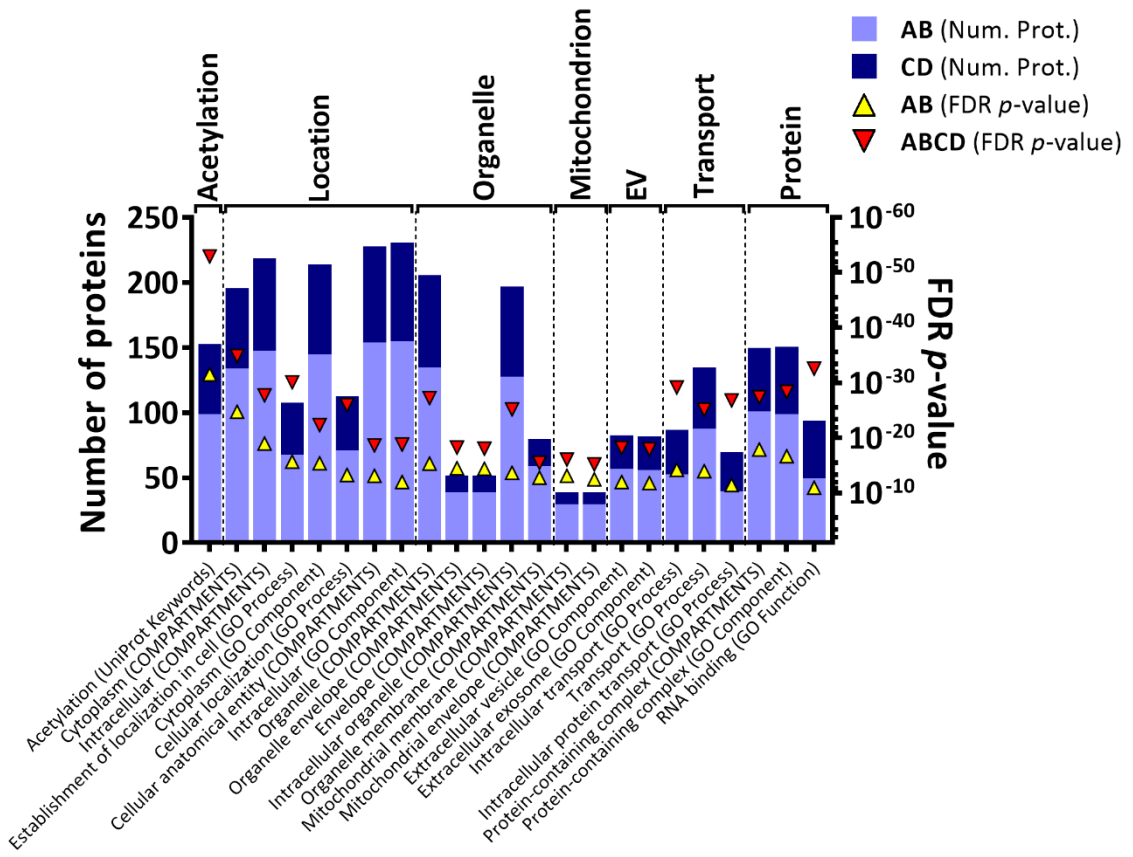
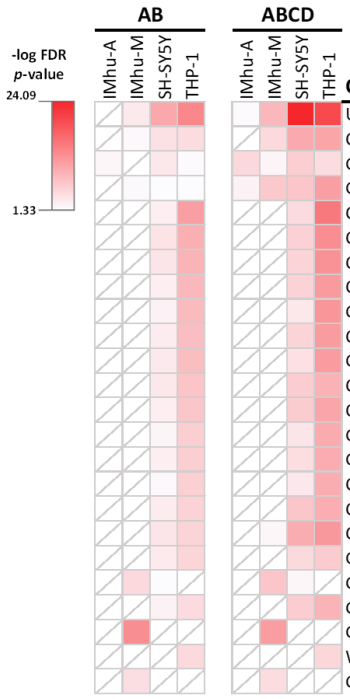


FIGURE 41. Top categories from the STRING database associated with the combined 14-3-3-free interactomes of IMhu-A, IMhu-M, SH-SY5Y and THP-1 cell lines.

The total number of proteins present in each STRING term from analysis of the combined four-cell line interactomes, omitting all 14-3-3 isoforms, is indicated with bars, showing differences between category AB proteins (light blue) and category CD proteins (dark blue). Significance associated to each STRING term, expressed as an FDR *p*-value, is represented with yellow triangles for category AB proteins and red triangles for category ABCD proteins. Terms obtained from the STRING analysis were grouped according to similarity in cell function and location. (EV: extracellular vesicles).

Non 14-3-3 proteins analysis

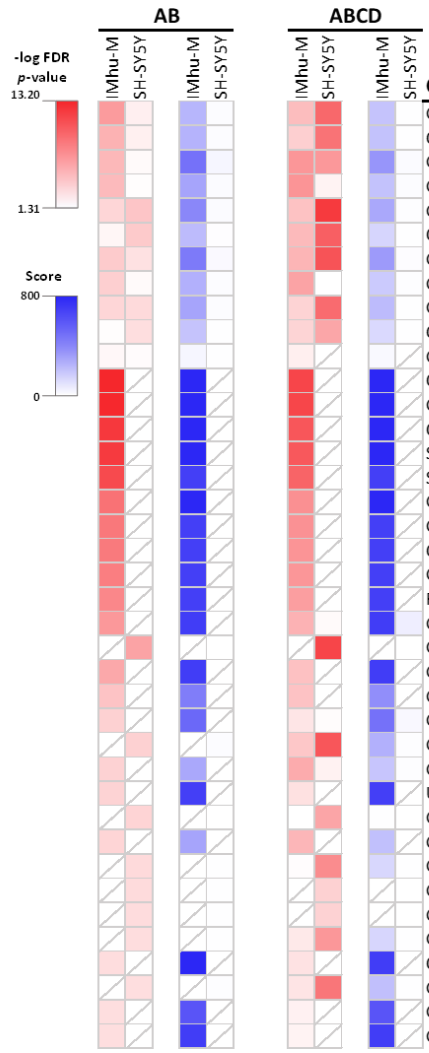


Category	Term ID	Term description
UniProt Keywords	KW-0007	Acetylation
COMPARTMENTS	GOCC:0043226	Organelle
GO Component	GO:0070062	Extracellular exosome
GO Function	GO:0003723	RNA binding
COMPARTMENTS	GOCC:0031967	Organelle envelope
COMPARTMENTS	GOCC:0031966	Mitochondrial membrane
COMPARTMENTS	GOCC:0005740	Mitochondrial envelope
COMPARTMENTS	GOCC:0019866	Organelle inner membrane
GO Component	GO:0031967	Organelle envelope
COMPARTMENTS	GOCC:0005743	Mitochondrial inner membrane
GO Component	GO:0031966	Mitochondrial membrane
COMPARTMENTS	GOCC:0031090	Organelle membrane
COMPARTMENTS	GOCC:0043231	Intracellular membrane-bounded organelle
GO Component	GO:0098800	Inner mitochondrial membrane protein complex
GO Component	GO:0005743	Mitochondrial inner membrane
COMPARTMENTS	GOCC:0098800	Inner mitochondrial membrane protein complex
COMPARTMENTS	GOCC:0043227	Membrane-bounded organelle
COMPARTMENTS	GOCC:0043229	Intracellular organelle
GO Component	GO:0031090	Organelle membrane
GO Process	GO:0006996	Organelle organization
GO Component	GO:0043231	Intracellular membrane-bounded organelle
GO Component	GO:0030992	Intraciliary transport particle b
WikiPathways	WP111	Electron transport chain: OXPHOS system in mitochondria
GO Process	GO:0070925	Organelle assembly

FIGURE 42. Functional analysis of ANKRD55 14-3-3-free interactomes for each individual cell line by the STRING database.

Heatmap from STRING analysis performed in category AB (left) and ABCD (right) ANKRD55 interactomes of astrocytic (IMhu-A), microglial (IMhu-M), neuroblastoma (SH-SY5Y) and monocytic (THP-1) cell lines, excluding all 14-3-3 protein isoforms. Significance values, provided as -log FDR *p*-value, are represented on a color scale from dark red (with maximum value of 24.09) to white (with minimum value of 1.33). Cells with a diagonal line indicate an absence of a STRING term.

Processes, components and compartments containing IFT terms



Category	Term ID	Term description
GO Component	GO:0032991	Protein-containing complex
COMPARTMENTS	GOCC:0032991	Protein-containing complex
GO Process	GO:0006886	Intracellular protein transport
GO Process	GO:0006996	Organelle organization
GO Process	GO:0051649	Establishment of localization in cell
COMPARTMENTS	GOCC:0043226	Organelle
GO Process	GO:0046907	Intracellular transport
GO Process	GO:0016043	Cellular component organization
GO Process	GO:0006810	Transport
COMPARTMENTS	GOCC:0110165	Cellular anatomical entity
GO Function	GO:0017076	Purine nucleotide binding
COMPARTMENTS	GOCC:0030992	Intraciliary transport particle b
GO Component	GO:0030992	Intraciliary transport particle b
COMPARTMENTS	GOCC:0030990	Intraciliary transport particle
STRING clusters	CL:10480	Intraciliary transport particle
STRING clusters	CL:10482	Intraciliary transport particle b
GO Process	GO:0042073	Intraciliary transport
GO Component	GO:0097542	Ciliary tip
COMPARTMENTS	GOCC:0097542	Ciliary tip
GO Process	GO:0035735	Intraciliary transport involved in cilium assembly
Reactome	HSA-5620924	Intraflagellar transport
GO Process	GO:0030705	Cytoskeleton-dependent intracellular transport
GO Component	GO:0005737	Cytoplasm
GO Process	GO:0031503	Protein-containing complex localization
GO Process	GO:0070925	Organelle assembly
GO Component	GO:0015630	Microtubule cytoskeleton
GO Process	GO:0051641	Cellular localization
GO Component	GO:0043232	Intracellular non-membrane-bounded organelle
UniProt Keywords	KW-0969	Cilium
GO Component	GO:0031982	Vesicle
GO Process	GO:0022607	Cellular component assembly
GO Component	GO:0005622	Intracellular
GO Component	GO:0031410	Cytoplasmic vesicle
GO Function	GO:0036094	Small molecule binding
GO Component	GO:0043229	Intracellular organelle
COMPARTMENTS	GOCC:0005929	Cilium
GO Process	GO:0051179	Localization
GO Component	GO:0005813	Centrosome
GO Component	GO:0005815	Microtubule organizing center

FIGURE 43. Processes, components and compartments from the STRING database containing IFT terms in IMhu-M and SH-SY5Y cell lines.

STRING analysis was applied to ANKRD55 interactomes that contain IFT proteins (IMhu-M and SH-SY5Y), following exclusion of all 14-3-3 isoforms. Statistical significance (expressed as -log FDR *p*-value) of STRING terms containing any IFT protein associated with category AB and ABCD protein interactomes is shown as red heatmaps (highest value: 13.20; lowest value: 1.31). Blue heatmaps show the diversity and relative abundance of IFT proteins in each STRING term. A score, calculated by multiplying the number of distinct IFT proteins detected with the % of IFTs in the STRING terms, was assigned to each case. The maximum score of 800 corresponds to presence of eight different IFTs and no other proteins. Absence of a STRING term in a cell line is represented with a diagonal line in the corresponding cell.

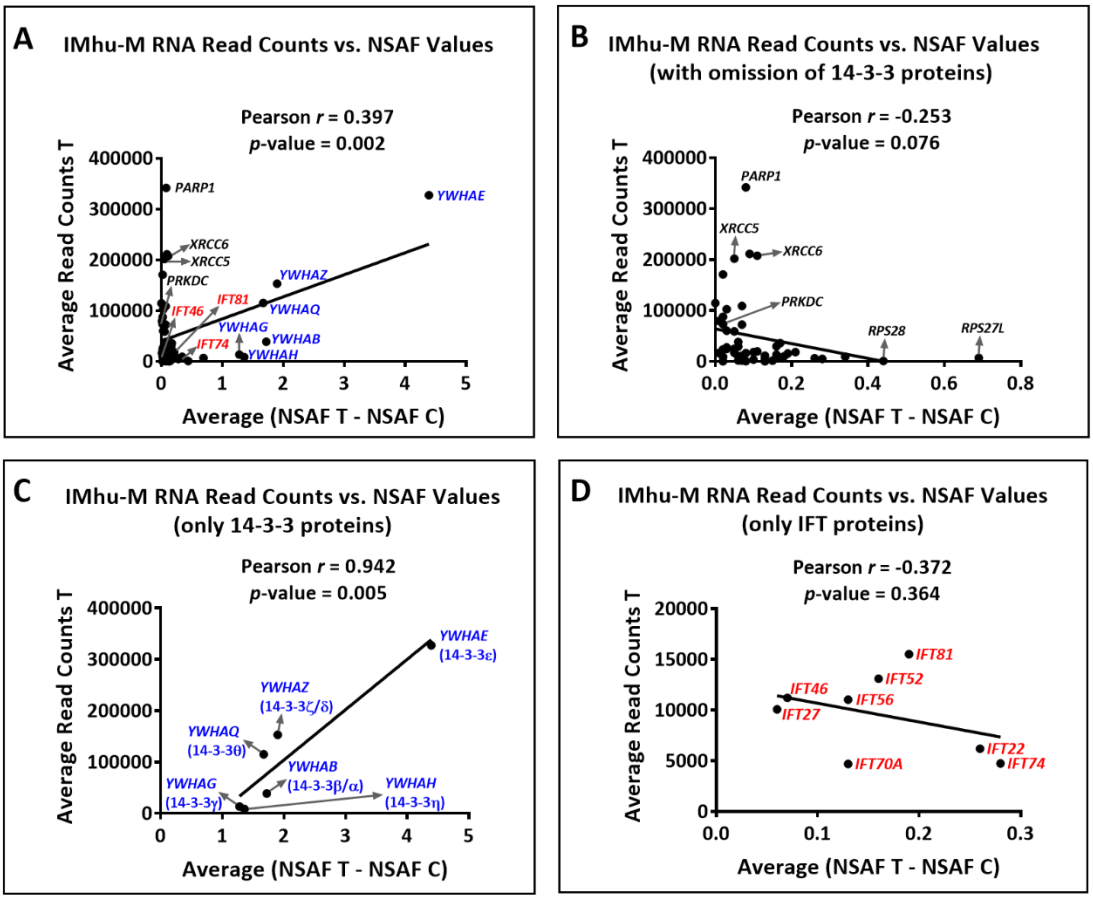


FIGURE 44. Correlation analysis of RNA-seq and interactome data from IMhu-M cells.
A Pearson correlation was applied between RNA-seq read counts of *ANKRD55* mRNA transfected (T) IMhu-M cells (average of two independent experiments) and normalized spectral abundance factor (NSAF) values, taken as the difference between transfected (NSAF T) and untransfected control (NSAF C) IMhu-M (average of three independent replicates). The analysis was performed including **(A)** all proteins identified in the interactome, **(B)** excluding 14-3-3 isoforms from total proteins, **(C)** considering only 14-3-3 isoforms or **(D)** considering only IFT proteins. Pearson correlation coefficients (*r*) and *p*-values are shown. Relevant proteins are indicated, including PRKDC-PARP1-XRCC5-XRCC6 complex (black), 14-3-3 isoforms (blue) and IFT proteins (red).

The IFT machinery is organized into two distinct multi-subunit complexes: IFT-A and IFT-B. The latter, in turn, consists of a core and a peripheral subcomplex[519]. The eight IFT subunits identified in the IMhu-M ANKRD55 interactome are scattered over the known IFT-B core-1 and IFT-B core-2 subcomplexes[519] and are not present in the peripheral subcomplex of this structure (**Figure 45**).

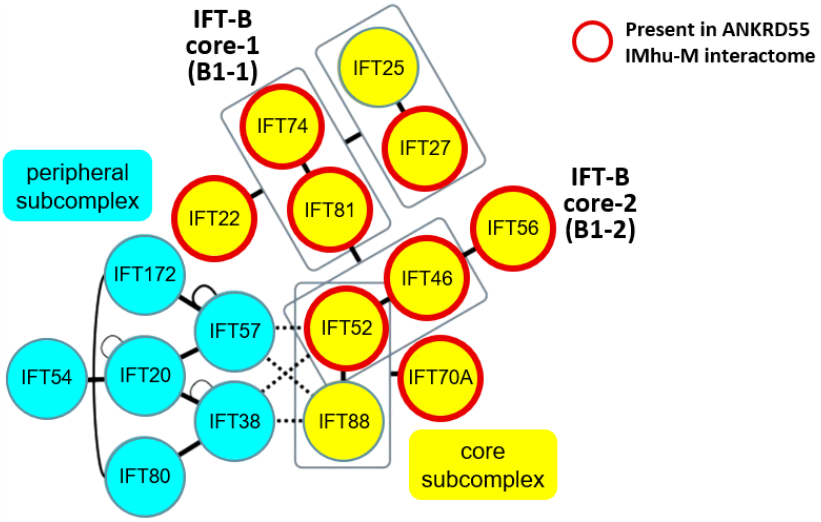


FIGURE 45. Schematic representation of the IFT-B complex.
The IFT-B complex is formed by the peripheral subcomplex (IFTs in blue) and core subcomplex (IFTs in yellow). The core subcomplex can be further divided into IFT-B core-1 (B1-1), formed by IFT22, IFT25, IFT27, IFT74 and IFT81; and IFT-B core-2 (B1-2), composed of IFT46, IFT52, IFT56, IFT70A and IFT88. IFT proteins detected in the IMhu-M interactome are circled in red, and the ANKRD55 interaction site is indicated. Figure adapted from Katoh *et al*.[520]

We verified that transfection with the synthetic ANKRD55 RNA/nanoparticle combination used in this study did not artefactually induce IFT gene expression in the IMhu-M cell line. In this regard, RNA-seq data in **Figure 46** show that IFT-B core-1 and IFT-B core-2 genes were expressed at similar or lower levels in transfected cells compared to untransfected control cells. In fact, the same observation was made in all genes involved in the complete IFT apparatus, including those forming the IFT-A complex, the heterotrimeric kinesin-II (involved in motor functions) and the BBSome complex (acts as a scaffold to maintain the integrity of IFT particles[519]) (**Figure 46**).

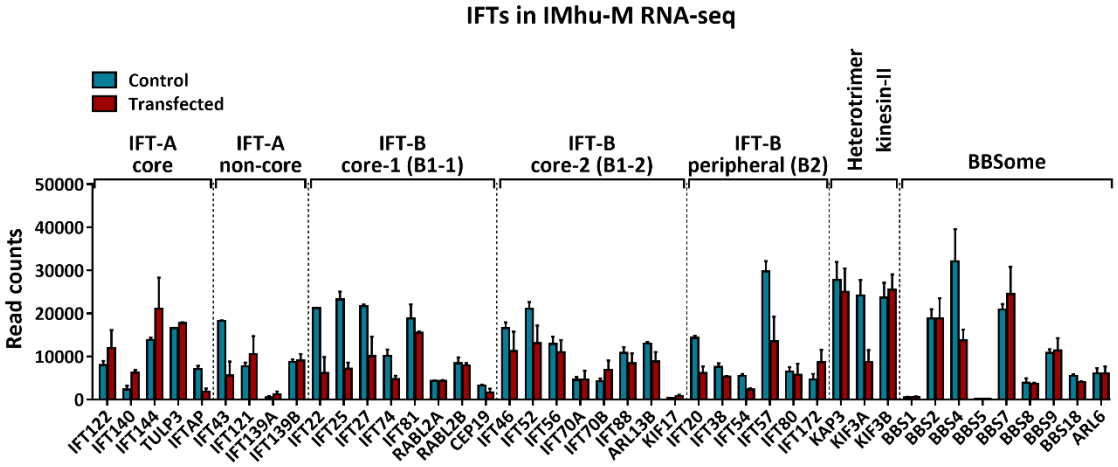


FIGURE 46. Transcript levels of IFT genes in IMhu-M.
Read counts from RNA-seq data of untransfected (blue) and ANKRD55 mRNA transfected (red) IMhu-M cells. Mean ± SEM from two independent experiments is represented. IFT elements were grouped into different subcomplexes based on work by Nakayama *et al*.[519]

4.4. Discussion

In this work, the molecular partners of the ANKRD55 protein were determined through interactome analysis in different types of cells, including astrocytic, microglial, neuroblastoma and monocytic cell lines. It has been shown that the use of synthetic mRNA is highly efficient in the protein expression of ANKRD55. The highest number of total ANKRD55-interacting proteins were detected in the SH-SY5Y cell line, often used as a neuronal model. Regarding the ANKRD55 interactome profile, although the majority of the proteins detected in the interactomes are unique to each cell line, the results show a common "core" of ANKRD55-interacting proteins. This "core" of proteins is shared by the four cell lines analyzed and formed by the 14-3-3 protein family, specifically isoforms 14-3-3η, 14-3-3β/α, and 14-3-3θ (encoded by *YWHAH*, *YWHAB*, and *YWHAQ* genes, respectively). Interestingly, these three 14-3-3 isoforms were also found in a previous ANKRD55 interactome study performed by our group in HEK293 cells[377], suggesting a universal binding between ANKRD55 and 14-3-3 in all cell types. Some other 14-3-3 isoforms are ANKRD55 interactants shared between two or three of the cell lines; an exception to this was the THP-1 cell line, whose interactome showed only the three common 14-3-3 isoforms with the other cellular interactomes studied.

The 14-3-3 protein family encompasses a group of proteins, expressed in all eukaryotic cells, which interact with hundreds of diverse proteins and function as regulators of many biological processes[521]. The name "14-3-3" derives from the 14th fraction on the diethylaminoethyl-cellulose chromatography and the 3.3 migration position in the subsequent starch-gel electrophoresis during its initial biochemical purification process[521]. In mammals, seven different isoforms of 14-3-3 have been identified: β/α, γ, ε, ζ/δ, η, σ and τ/θ[522]. 14-3-3 proteins are very abundant in the brain, accounting for about 1% of the total amount of soluble brain protein, but they are also present in almost all tissues[523]. These proteins lack enzymatic activity but are crucial regulators of many kinases, which in turn are implicated in cellular processes such as transcription, cell cycle progression and apoptosis, among others[524]. Furthermore, this family of proteins also has the ability to modulate the immune system (particularly T cells and B cells) and the inflammatory response[525]. In fact, alterations of 14-3-3 protein levels have been linked to chronic inflammatory disorders, including autoimmune diseases[525]. In cancer, 14-3-3 proteins have also been shown to participate in the regulation of the development, metastasis and invasion of tumors[526]. Moreover, 14-3-3 proteins act as molecular chaperones, preventing the aggregation of unfolded proteins, and thus, their dysregulation may contribute to the pathogenesis of neurodegenerative diseases characterized

by excessive protein aggregation, such as Alzheimer's disease, Parkinson's disease and dementia with Lewy bodies[526,527]. In fact, the 14-3-3 protein has been proposed as an early prognostic biomarker of MS due to the correlation between the levels of this protein in the CSF and an increased rate of relapse and severity of disability[528,529]. In this context, isoform 14-3-3γ (encoded by the *YWHAG* gene) is especially relevant, since it is mainly expressed in neurons and is highly produced in the brain during development[530]. 14-3-3γ deficiency was found to produce oligodendrocyte apoptosis, leading to increased demyelination in a murine model that mimics the pathological features of MS. This suggests that this 14-3-3 isoform may confer protection in autoimmune demyelination[525,530,531]. In this regard, it should be noted that 14-3-3γ is detected in our results as a common ANKRD55 interactant in IMhu-A, IMhu-M and SH-SY5Y. The three cell types represented by these lines (astrocytes, microglia and neuron-like cells) are all crucial for neurological system function and are implicated in MS disease.

Beyond 14-3-3, there is only one protein present in three of the interactomes of the cell types analyzed. This protein is BAG-2, a cochaperone belonging to the BAG family which interacts with other molecules, regulates multiple cellular processes and is involved in the pathogenesis of neurodegenerative diseases[532]. In our study, BAG-2 was found in the ANKRD55 interactome of astrocytic, microglial and neuroblastoma cell lines. In addition, some other proteins were shared between two different cell type interactomes. These included the CAD protein, which participates in pyrimidine synthesis[533]; the 60S ribosomal protein L32 (*RPL32*), involved in the protein synthesis[534]; and two intraflagellar transport proteins (*IFT74* and *IFT22*), all of them identified in the IMhu-M and SH-SY5Y interactomes. The SH-SY5Y and THP-1 interactomes shared two mitochondrial proteins: the mitochondrial glutamate carrier 1 (encoded by the *SLC25A22* gene) and the ATPase subunit d (encoded by the *ATP5PD* gene). The 40S ribosomal protein S28 (*RPS28*) was shared between the IMhu-M and THP-1 interactomes; and the cullin-associated and neddylation-dissociated protein 1 (*CAND1*), involved in the regulation of ubiquitination process for protein degradation[535], was found in the IMhu-A and THP-1 interactomes.

In the context of the connections and relationships identified between ANKRD55 interactant proteins, several cell signaling pathways emerged with significant relevance. Again, THP-1 cells were less related to the other three cell lines, as several of the pathways present in all the types of cells studied showed less significance in the monocytic cell line. Among these pathways are the G2/M DNA damage checkpoint regulation pathway, where cyclin-dependent kinases play a critical role[536]; the Hippo signaling pathway, which is involved in the regulation

of cell proliferation, tissue regeneration and organ growth[537]; the ERK5 signaling pathway, which is activated under stress and whose functions are to maintain the integrity of tissues exposed to mechanical stress and promote cell proliferation and control cell cycle progression[538]; the IGF-1 signaling pathway, relevant to the processes of aging and longevity[539], as well as CNS development, maturation and neuroplasticity[540]; and the protein kinase A signaling pathway, implicated in metabolic regulation and the maintenance of cellular homeostasis[541], and essential to neuronal development[542]. Other molecular networks with high significance in the IMhu-A, IMhu-M and SH-SY5Y interactomes include p70S6k signaling, related to protein synthesis and control of cell cycle, growth and survival[543]; PI3K/AKT signaling, critical for cellular proliferation and involved in lymphocyte development[544]; and ERK/MAPK signaling, necessary for proper skeletal development[545], responsible for the induction of cellular proliferation, differentiation and neuronal plasticity and whose dysregulation is related to neurodegenerative diseases[546]. All of these signaling pathways have common links, such as the control of cell proliferation or the cell cycle and involvement in neurological development, which provide new clues about the possible function of ANKRD55. In THP-1 cells, the most significant terms associated with the interactome results are the so-called "mitochondrial dysfunction" when only category A and B proteins are considered, or "EIF2 signaling" if proteins from categories B and C are also included. EIF2 signaling is also present in the analyses results of the other three cell lines and constitutes a key component in initiation of protein biosynthesis[547].

In concordance with the results of the ANKRD55 interacting proteins, "14-3-3-mediated signaling" also emerged with high significance in all cell lines, with the exception of THP-1. Analyses of the presence of 14-3-3 isoforms in signaling pathways revealed a high abundance of this protein in the most significant canonical pathways of the IMhu-A, IMhu-M and SH-SY5Y interactomes. For instance, ERK5 signaling and the IGF-1 signaling pathways are characterized by the presence of only 14-3-3 isoforms in the four cell lines. In addition, correlation analysis between the microglia interactome data and RNA-seq performed in this cell line reinforced this observation. Thus, it is concluded that 14-3-3 isoforms are the main contributors to the ANKRD55 interactomes. Discarding 14-3-3 isoform proteins from the analyses, "acetylation" was the most significant term associated to combined interactomes from the four cell lines. Acetylation, one of the main post-translational protein modifications in the cell, is involved in protein-level and transcriptional activation, and malfunctions in this process can lead to neurogenerative diseases[548].

Another group of proteins to highlight is the formed by the intraflagellar transport (IFT) proteins. Up to eight unique members of this family were detected in the microglial cell line interactome: IFT22, IFT27, IFT46, IFT52, IFT56 (also named TTC26), IFT70A (also named TTC30A), IFT74 and IFT81, of which only IFT22 and IFT74 were also identified in the interactome of the neuroblastoma cell line. IFT proteins were first discovered as components necessary for the assembly of flagella and cilia, as they are responsible for the bidirectional transport of proteins within these dynamic structures[549]. The IFT trafficking machinery can be dissociated into two complexes, named IFT-A and IFT-B, which are connected to ciliary membranes through the eight-protein BBSome complex [519]. The IFT-A complex consists of six subunits (IFT122, IFT140, IFT144, IFT143, IFT121 and IFT139) and the adaptor protein TULP3, and it mediates retrograde trafficking of ciliary proteins, driven by a dynein motor[519]. On the other hand, the IFT-B complex is involved in anterograde ciliary protein trafficking powered by kinesin and is composed of sixteen subunits. The overall IFT-B complex is divided into two subcomplexes: the core subcomplex (or IFT-B1 subcomplex), composed of IFT22, IFT25, IFT27, IFT46, IFT52, IFT56, IFT70, IFT74, IFT81 and IFT88; and the peripheral subcomplex (also known as IFT-B2 subcomplex), which is formed by IFT20, IFT38, IFT54, IFT57, IFT80 and IFT172[520]. The IFT-B1 core subcomplex can be further divided into two components: the core-1 (B1-1) subgroup, including IFT22, IFT25, IFT27, IFT74 and IFT81; and the core-2 (B1-2) subgroup, including IFT46, IFT52, IFT56, IFT70 and IFT88[519].

All of the IFT proteins found in our microglial interactome belonged to the core subcomplex. Some of the components of this core subcomplex, such as IFT46 and IFT56, are known to be essential for the stability of the IFT-B complex and consequently for proper development of ciliated tissues in vertebrates[550,551]. In addition, previously to our work, the integrated hu.MAP human protein complex map (http://proteincomplexes.org) created by Drew *et al.* revealed ANKRD55 as a component of the IFT-B complex that mainly interacts with IFT52 protein[378]. Experiments performed *in vivo* showed that the ANKRD55-GFP fusion protein localized to the cilia of *Xenopus laevis* multi-ciliated epithelial cells, in which it was seen to traffic up and down and to move coordinately with the IFT38 (*CLUAP1*) protein[378]. Moreover, morpholino antisense oligonucleotide knockdown of *ANKRD55* led to a reduction in both number and length of cilia, similar to the effect of knockdown of *IFT52*, and resulted in defective vertebrate neural tube closure in *Xenopus* embryos, which was consistent with previous observations following IFT disruption[378,552,553]. Thus, this may imply a role of ANKRD55 in development.

Primary cilia are non-motile cilia present in a wide variety of eukaryotic cells and whose structure and length are regulated by IFT proteins[554]. These organelles act as an "antenna" to sense the extracellular environment and transduce signals from different pathways, regulating cell proliferation, differentiation and communication[554]. One of these pathways is the Hedgehog signaling pathway, which is vital to animal development and tissue homeostasis and which requires IFTs for proper functioning[552,555]. Several of these functions or diseases related to IFT proteins, such as cell surface receptors, signal transduction, skeletal ciliopathies, disorder of stature and dwarfism, were found in the bioinformatic analysis from the microglial interactome data performed in this work. Interestingly, microglial cells do not display primary cilia[556,557]. However, IFT presence and functions are not only restricted to cilia and flagella: these proteins are also found outside these organelles and in non-ciliated cells, where they are relevant for the control of cell division geometry[558], and in vesicle trafficking and exocytosis associated with the Golgi apparatus[559]. In addition, IFTs have been demonstrated to be essential for the formation of the immune synapse in T lymphocytes, which do not possess cilia[560,561]. The term "immune synapse" refers to the physical interaction between a T cell and an antigen-presenting cell and is orchestrated by the reorganization of the cytoskeleton and surface receptors[561]. In the CNS, a similar interaction is produced between neurons and microglia (or T cells), forming the so-called "neuro-immune synapse", which is thought to have regulatory effects on activated microglia at the onset of neurodegenerative diseases[364]. Furthermore, during the progression of neurodegeneration, when neuronal damage occurs, microglia are the main MHC-II APCs in the brain, and crosstalk occurs between activated microglia and infiltrating T cells[562]. It has been hypothesized that the immune synapse could represent the functional homolog of the primary cilium in non-ciliated hematopoietic cells[561]. Therefore, it is reasonable to think that the same happens in microglia, which also lack primary cilia, meaning ANKRD55 may play a role, along with IFT proteins, in this form of cell–cell communication. Apart from the IFT machinery, data from the microglial interactome suggest that ANKRD55 may also recognize another well-studied complex, formed by DNA-PK, PARP-1, Ku80 and Ku70 proteins (encoded by the genes *PRKDC*, *PARP1*, *XRCC5* and *XRCC6*, respectively). This complex is involved in DNA repair processes, specifically in the non-homologous DNA end-joining (NHEJ) mechanism[563–565], so ANKRD55 might also indirectly participate in this molecular process.

The different ANKRD55 interactome studies described in the literature, both experimentally and bioinformatically, exhibit similarities and differences. Various isoforms of

the 14-3-3 family of proteins are shared among the ANKRD55 interactomes identified by Ugidos *et al*. in HEK293 cells[377], the BioPlex consortium (https://bioplex.hms.harvard.edu) in HEK293 and HCT116 cells[566], and the current work in IMhu-M, IMhu-A, SHSY-5 and THP-1 cell lines. IFT proteins were absent from the ANKRD55 interactome integrated in the human reference interactome (HuRI) map of binary protein interactions (http://www.interactome-atlas.org)[567] and from that of the colon cancer cell line HCT116 (BioPlexHCT_1_0 network), determined by the BioPlex consortium[566]. However, some elements of the IFT-B complex are highly enriched in our interactome study. In this line, the study by Drew *et al*.[378] and in the two ANKRD55 interactomes in HEK293 cells, performed by the BioPlex consortium (BioPlex 3.0)[566] and by Ugidos *et al*.[377], also showed the presence of IFT-B proteins as ANKRD55 interactants. On the other hand, the HEK293 ANKRD55 interactome by Ugidos *et al*. was also enriched in ATP- or nucleotide-binding proteins of diverse functionality with no apparent relationship to IFT-mediated ciliogenesis[377]. Nuclear transport proteins and terms were also identified in the analysis of this interactome performed in HEK293 cells, suggesting a classical importin/exportin-mediated nuclear transport of ANKRD55, even though a different nuclear import system, independent of importins and predetermined by the amino acid protein sequence, has been shown to be frequently used by many ankyrin repeat-containing proteins[377,568].

Taken together, the results from the comparative ANKRD55 interactomes of this work demonstrate the direct or indirect molecular interactions of ANKRD55 with specific proteins in each cell type, in addition to the existence of a common core of ANKRD55-interacting proteins, which is shared among different cell types. This implies a possible role of ANKRD55 in neural tissue development, immune synapse maintenance, cell signaling and control of cell cycle and proliferation, among others, and reveals the importance of this protein in the immune and nervous systems. The *ANKRD55* gene is a risk modulator involved in many pathologies, including MS, thus the study of the corresponding protein is essential for better understanding this and other diseases which currently lack effective cures. Although further investigations will be needed to better describe the specific function of ANKRD55 in certain pathways, this research adds new insight into the partners and cell processes associated with this poorly studied but increasingly interesting protein.

5. Conclusions

5. Conclusions

The following conclusions were drawn from this work:

- *ANKRD55* gene expression is induced by retinoic acid receptor alpha agonist AM580 in immature monocyte-derived dendritic cells and is downregulated in IFN-γ/LPS-matured monocyte-derived dendritic cells.

- The protein ANKRD55 is visible in the cytosol and nuclei of monocyte-derived dendritic cells and colocalizes with components of nuclear speckles in this cell type.

- The protective A alleles of MS risk SNPs rs6859219 and rs7731626 are associated with higher mRNA expression of *ANKRD55* and *IL6ST* in immature monocyte-derived dendritic cells of healthy donors. These protective alleles are also associated with lower levels of both genes in $CD4^+$ T lymphocytes. The same effects, but with greater significance, are observed for the protective C allele of the correlated SNP rs13186299, which is related to the monocyte percentage of white blood cells.

- In $CD4^+$ T cells from MS patients, the rs7731626 SNP genotype is associated with *ANKRD55* gene expression, with the homozygotes for the protective A allele showing the lowest expression levels. The genotype of SNPs rs6859219, rs7731626 and rs13186299 does not significantly influence the expression levels of *ANKRD55* or *IL6ST* in immature monocyte-derived dendritic cells from patients with MS.

- 14-3-3 family proteins, specifically isoforms 14-3-3η, 14-3-3β/α and 14-3-3θ, are the main ANKRD55 interactants shared by astrocytic, microglial, neuroblastoma and monocytic cell lines.

- In the microglia cell line, ANKRD55 interacts with intraflagellar transport proteins belonging to the IFT-B complex.

- Bioinformatic analyses performed on proteins from ANKRD55 interactomes resulted in the identification of cell signaling pathways related to cell cycle control, cell proliferation and neurological development.

6. Supplementary material

6. Supplementary material

SUPPLEMENTARY TABLE 1. List of proteins, excluding top 20, identified in the IMhu-A interactome.

Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged (Aver. Rat. NSAF T/C) over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category (Cat.) requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

IMhu-A			
Gene symbol	**Aver. Rat. NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Cat.**
TUFM	34.98	2	C
RPL11	34.78	3	C
UBC	34.74	3	C
PHB1	34.38	1	C
YWHAE	34.37	1	C

SUPPLEMENTARY TABLE 2. List of proteins, excluding top 20, identified in the IMhu-M interactome.

Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged (Aver. Rat. NSAF T/C) over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category (Cat.) requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

IMhu-M			
Gene symbol	**Aver. Rat. NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Cat.**
HNRNPA3	67	2	A
YWHAE	42.15	3	A
UBC	36.67	3	A
VPS4A	35.17	2	A
PSMA3	34.42	2	A
PRSS1	67	0	B
IFT46	67	1	B
AKAP8L	67	1	B
IFT27	67	1	B
SEC13	67	1	B
TPM4	67	0	B
GCDH	67	1	B
HSPH1	67	1	B
DNAJB12	67	1	B
SUPT16H	67	1	B
RPS28	34.80	0	B
MARS1	34.52	1	B
RPL32	34.42	0	B
MYO1B	34.42	0	B
PCOLCE	67.23	2	C
DNAJB6	67.18	2	C
EMD	67.04	2	C

SUPPLEMENTARY TABLE 2. (Continued from right side of table on previous page)

RPL17	67.04	1	C
NAP1L4	67.04	1	C
RPL11	67.01	3	C
RAB10	66.99	1	C
SLC25A3	66.92	2	C
SAFB	66.89	0	C
RPS17	66.88	2	C
PUF60	66.86	1	C

PARP1	66.79	3	C
TRIP6	35.35	3	C
DNAJA3	34.61	2	C
AIFM1	34.24	2	C
DCD	34.19	1	C
XRCC5	34.31	2	D
EXOSC8	34.31	1	D

SUPPLEMENTARY TABLE 3. List of proteins, excluding top 20, identified in the SH-SY5Y interactome.

Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged (Aver. Rat. NSAF T/C) over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category (Cat.) requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

SH-SY5Y			
Gene symbol	**Aver. Rat. NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Cat.**
RAB11A	3.06	2	A
RPS20	100	0	B
BAG2	100	0	B
ATP5PB	100	0	B
NAP1L4	100	0	B
RBFOX2	100	0	B
CAP1	100	0	B
ATP1A3	68.16	0	B
RPL32	67.54	0	B
LYZ	67	0	B
TMED3	67	0	B
S100A8	67	0	B
SLC25A15	67	0	B
S100A7	67	0	B
KDELR1	67	0	B
TXNDC5	67	0	B
TMED7	67	0	B
DAD1	67	0	B
CASP14	67	0	B

LCN1P1	67	0	B
IFT22	67	0	B
GALK1	67	0	B
RUFY2	67	0	B
TSN	67	0	B
PLOD1	67	0	B
IFT74	67	0	B
TECR	67	0	B
TIMM50	67	0	B
PTPA	67	0	B
GOT1L1	67	0	B
CCDC47	67	0	B
SARS1	67	0	B
PPP3CA	67	0	B
DYNC1LI1	67	0	B
CAT	67	0	B
EIF4B	67	0	B
PGM2	67	0	B
TMCC1	67	0	B
DGKA	67	0	B
HNRNPUL2	67	0	B
H6PD	67	0	B
OCRL	67	0	B

SUPPLEMENTARY TABLE 3. (Continued from right side of table on previous page)

APC	67	0	B
CCDC80	67	0	B
FLNB	67	0	B
MAP2	67	0	B
DYNC1H1	67	0	B
RYR2	67	0	B
DBN1	35.54	0	B
MMP19	34.92	1	B
DNAJA3	35.16	0	B
JUP	34.82	0	B
MYO1C	34.59	0	B
NDUFA13	34.42	0	B
HAX1	34.42	0	B
AKR1B10	34.42	0	B
PLOD3	34.42	0	B
ETF1	2.76	0	B
APRT	2.01	0	B
HSPH1	67.23	3	C
SSR1	67.09	1	C
SEC61A1	67.09	2	C
GMPPB	67.09	1	C
RPL10	67.04	1	C
PSMA4	67.04	2	C
SEC61B	67.04	1	C
RPS16	66.98	3	C
DBH	66.95	2	C
SCYL2	66.95	2	C
ATP5MG	66.88	1	C
YTHDF2	66.88	1	C
HYOU1	66.87	3	C
RPL31	66.85	1	C
ERLIN2	66.84	1	C
UQCRC2	66.81	2	C
ISOC1	66.81	1	C
RPS23	66.79	1	C
CHCHD3	66.77	1	C
PSMB1	66.73	1	C
PSMD12	66.72	1	C
ACTL6A	66.72	1	C
RPS12	36.16	3	C
CALU	35.89	3	C
RPS25	35.53	2	C
AHSA1	35.31	3	C
SLC25A13	35.09	3	C
RBM39	34.97	1	C
EMD	34.67	1	C
HADHA	34.43	3	C
PFDN1	34.40	1	C
SEMA3C	34.40	3	C
STT3A	34.37	2	C
MAP1B	34.35	2	C
C1QBP	34.30	2	C
ILF2	34.28	2	C
RBMXL1	34.28	1	C
RPS13	34.22	1	C
NUP93	34.22	2	C
HBD	34.20	0	C
ILF3	3.77	3	C
RPS6	3.47	3	C
HSPB1	3.29	3	C
RCN2	3.28	3	C
RPN1	2.53	3	C
SKP1	2.30	3	C
SNRPD3	2.17	0	C
MYO1B	1.84	2	C
HDGF	2.78	1	D

SUPPLEMENTARY TABLE 4. List of proteins, excluding top 20, identified in the THP-1 interactome.

Gene symbol corresponding to each protein is shown. Proteins were ranked from highest to lowest NSAF (T)/NSAF (C) ratio averaged (Aver. Rat. NSAF T/C) over three independent replicates. Proteins absent in the control and exclusively found in transfected cells in a single replicate are given a value of 100 in the average ratio NSAF T/C, while proteins absent in both control and transfected cells in a single replicate are given a value of 1. Number of replicates in which the protein was identified with more than one unique peptide is indicated (Num. Rep. > 1 Un. Pep.). Category (Cat.) requirements are detailed in the "Materials and methods" from part II (point 4.2.10. Protein ranking).

THP-1			
Gene symbol	**Aver. Rat. NSAF T/C**	**Num. Rep. > 1 Un. Pep.**	**Cat.**
NDUFA4	67	0	B
CHCHD3	67	1	B
NDUFB4	67	0	B
SDCBP	67	0	B
PYCR3	67	1	B
CTSD	67	1	B
PRPF19	67	1	B
H1-10	67	0	B
AGPAT5	67	0	B
EHD4	67	1	B
MCM5	67	0	B
CAND1	67	1	B
MX2	67	0	B
MCM4	67	1	B
PPP2R2D	67	0	B
ARCN1	67	0	B
NOP56	67	0	B
TOP1	67	0	B
SLC25A13	67	0	B
IPO7	67	0	B
CPT1A	38	1	B
YWHAB	35.56	1	B
VDAC3	35.18	1	B
GNB2	35	0	B
RPS28	34.80	0	B
HBD	34.80	0	B
SCAMP3	34.80	1	B
LRPPRC	34.80	1	B
S100A6	34.67	0	B
ATP5MG	34.42	0	B
SLC25A22	3.47	1	B
RPL19	2.10	1	B
CPSF6	67.11	2	C
SLC25A1	67.05	2	C
IMMT	66.92	2	C
EIF2S1	66.89	0	C
LUC7L2	66.83	0	C
YWHAQ	35.75	3	C
ATAD3A	35.16	3	C
NUP93	35.13	2	C
ILF2	35.01	2	C
NEU1	34.98	2	C
RPL17	34.93	2	C
CALR	34.72	3	C
PHB2	34.55	1	C
PHB1	34.51	2	C
SFPQ	34.50	3	C
RPL28	34.47	3	C
FLG2	34.34	1	C
HNRNPA1L2	34.34	2	C
PGAM5	4.32	3	C
RPS18	3.45	3	C
RPS3A	2.82	3	C
RPL9	2.59	2	C
NPM1	2.55	2	C
RPL14	2.32	3	C
RPL18	2.25	3	C
RPL7A	1.93	3	C
SLC25A6	1.72	3	C
S100A8	34.63	2	D
SRSF7	34.25	1	D

7. Bibliography

7. Bibliography

1. McFarland, H. F. & Martin, R. Multiple sclerosis: a complicated picture of autoimmunity. Nat Immunol 8, 913–919 (2007).

2. Frischer, J. M. et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. Brain 132, 1175–1189 (2009).

3. Hauser, S. L. & Oksenberg, J. R. The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. Neuron 52, 61–76 (2006).

4. Ramagopalan, S. V., Dobson, R., Meier, U. C. & Giovannoni, G. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. Lancet Neurol 9, 727–739 (2010).

5. Benedict, R. H. B. et al. Predicting quality of life in multiple sclerosis: accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change. J Neurol Sci 231, 29–34 (2005).

6. Confavreux, C., Vukusic, S., Moreau, T. & Adeleine, P. Relapses and progression of disability in multiple sclerosis. N Engl J Med 343, 1430–1438 (2000).

7. Oreja-Guevara, C. et al. Cognitive Dysfunctions and Assessments in Multiple Sclerosis. Front Neurol 10, 581 (2019).

8. Landtblom, A.-M., Fazio, P., Fredrikson, S. & Granieri, E. The first case history of multiple sclerosis: Augustus d'Esté (1794-1848). Neurol Sci 31, 29–33 (2010).

9. Charcot, J. Histologie de la sclérose en plaques. Gaz Hopital 41, 554–566 (1868).

10. Murray, T. J. The history of multiple sclerosis: the changing frame of the disease over the centuries. J Neurol Sci 277 Suppl 1, S3-8 (2009).

11. Firth, D. The Case of Augustus d'Este (1794–1848): The First Account of Disseminated Sclerosis. Proceedings of the Royal Society of Medicine 34, 381–384 (1941).

12. Kornek, B. & Lassmann, H. Axonal pathology in multiple sclerosis. A historical note. Brain Pathol 9, 651–656 (1999).

13. Charcot, J. Leçons sur les maladies du système nerveux: Faites à la Salpêtrière. Tome 1 (1880).

14. Marburg, O. Die sogenannte "akute multiple Sklerose" (Encephalomyelitis peraxialis scleroticans). Jahrb Neurol Psych 27, 211–312 (1906).

15. Frommann, C. Untersuchungen über die Gewebsveränderungen bei der multiplen Sklerose des Gehirns und Rückenmarks. Jena (1878).

16. Babinski, J. Recherche sur l'anatomie pathologique de la sclérose en plaque et étude comparative des diverses variétés de sclérose de la moelle. Arch Physiol 5–6, 186–207 (1885).

17. Kutzelnigg, A. et al. Cortical demyelination and diffuse white matter injury in multiple sclerosis. Brain 128, 2705–2712 (2005).

18. Reich, D. S., Lucchinetti, C. F. & Calabresi, P. A. Multiple Sclerosis. N Engl J Med 378, 169–180 (2018).

19. Dendrou, C. A., Fugger, L. & Friese, M. A. Immunopathology of multiple sclerosis. Nat Rev Immunol 15, 545–558 (2015).

20. Lucchinetti, C. et al. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. Ann Neurol 47, 707–717 (2000).

21. Schmidt, S. Candidate autoantigens in multiple sclerosis. Mult Scler 5, 147–160 (1999).

22. Mirshafiey, A. & Kianiaslani, M. Autoantigens and autoantibodies in multiple sclerosis. Iran J Allergy Asthma Immunol 12, 292–303 (2013).

23. Hohlfeld, R., Dornmair, K., Meinl, E. & Wekerle, H. The search for the target antigens of multiple sclerosis, part 2: CD8+ T cells, B cells, and antibodies in the focus of reverse-translational research. Lancet Neurol 15, 317–331 (2016).

24. Cencioni, M. T., Mattoscio, M., Magliozzi, R., Bar-Or, A. & Muraro, P. A. B cells in multiple sclerosis - from targeted depletion to immune reconstitution therapies. Nat Rev Neurol 17, 399–414 (2021).

25. Haase, S. & Linker, R. A. Inflammation in multiple sclerosis. Ther Adv Neurol Disord 14, 17562864211007688 (2021).

26. Planas, R. et al. GDP-l-fucose synthase is a CD4+ T cell-specific autoantigen in DRB3*02:02 patients with multiple sclerosis. Sci Transl Med 10, eaat4301 (2018).

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

28. Hohlfeld, R., Dornmair, K., Meinl, E. & Wekerle, H. The search for the target antigens of multiple sclerosis, part 1: autoreactive CD4+ T lymphocytes as pathogenic effectors and therapeutic targets. Lancet Neurol 15, 198–209 (2016).

29. Browne, P. et al. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. Neurology 83, 1022–1024 (2014).

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

31. The Multiple Sclerosis International Federation (MSIF). Atlas of MS, 3rd Edition. Part 1: Mapping multiple sclerosis around the world key epidemiology findings. (September 2020).

32. Koch-Henriksen, N. & Sørensen, P. S. The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol 9, 520–532 (2010).

33. Russi, A. E., Ebel, M. E., Yang, Y. & Brown, M. A. Male-specific IL-33 expression regulates sex-dimorphic EAE susceptibility. Proc Natl Acad Sci U S A 115, E1520–E1529 (2018).

34. Ysrraelit, M. C. & Correale, J. Impact of sex hormones on immune function and multiple sclerosis development. Immunology 156, 9–22 (2019).

35. Voskuhl, R. R., Sawalha, A. H. & Itoh, Y. Sex chromosome contributions to sex differences in multiple sclerosis susceptibility and progression. Mult Scler 24, 22–31 (2018).

36. Harbo, H. F., Gold, R. & Tintoré, M. Sex and gender issues in multiple sclerosis. Ther Adv Neurol Disord 6, 237–248 (2013).

37. Trojano, M. et al. Geographical variations in sex ratio trends over time in multiple sclerosis. PLoS One 7, e48078 (2012).

38. Leray, E., Moreau, T., Fromont, A. & Edan, G. Epidemiology of multiple sclerosis. Rev Neurol (Paris) 172, 3–13 (2016).

39. Krysko, K. M. et al. Sex effects across the lifespan in women with multiple sclerosis. Ther Adv Neurol Disord 13, 1756286420936166 (2020).

40. Alhomoud, M. A., Khan, A. S. & Alhomoud, I. The Potential Preventive Effect of Pregnancy and Breastfeeding on Multiple Sclerosis. Eur Neurol 84, 71–84 (2021).

41. GBD 2016 Multiple Sclerosis Collaborators. Global, regional, and national burden of multiple sclerosis 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 18, 269–285 (2019).

42. Simpson, S., Taylor, B. V. & van der Mei, I. The role of epidemiology in MS research: Past successes, current challenges and future potential. Mult Scler 21, 969–977 (2015).

43. Tao, C. et al. Higher latitude is significantly associated with an earlier age of disease onset in multiple sclerosis. J Neurol Neurosurg Psychiatry 87, 1343–1349 (2016).

44. Simpson, S. et al. Latitude continues to be significantly associated with the prevalence of multiple sclerosis: an updated meta-analysis. J Neurol Neurosurg Psychiatry 90, 1193–1200 (2019).

45. Lam, T., VoPham, T., Munger, K. L., Laden, F. & Hart, J. E. Long-term effects of latitude, ambient temperature, and ultraviolet radiation on the incidence of multiple sclerosis in two cohorts of US women. Environ Epidemiol 4, e0105 (2020).

46. Belbasis, L., Bellou, V., Evangelou, E., Ioannidis, J. P. A. & Tzoulaki, I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol 14, 263–273 (2015).

47. Alonso, A. & Hernán, M. A. Temporal trends in the incidence of multiple sclerosis: a systematic review. Neurology 71, 129–135 (2008).

48. Jelinek, G. A. et al. Associations of Lifestyle, Medication, and Socio-Demographic Factors with Disability in People with Multiple Sclerosis: An International Cross-Sectional Study. PLoS One 11, e0161701 (2016).

49. Tremlett, H., Zhu, F., Ascherio, A. & Munger, K. L. Sun exposure over the life course and associations with multiple sclerosis. Neurology 90, e1191–e1199 (2018).

50. Gale, C. R. & Martyn, C. N. Migrant studies in multiple sclerosis. Prog Neurobiol 47, 425–448 (1995).

51. Ebers, G. C. Environmental factors and multiple sclerosis. Lancet Neurol 7, 268–277 (2008).

52. Hammond, S. R. The age-range of risk of developing multiple sclerosis: Evidence from a migrant population in Australia. Brain 123, 968–974 (2000).

53. Compston, A. & Coles, A. Multiple sclerosis. Lancet 372, 1502–1517 (2008).

54. Brownlee, W. J., Hardy, T. A., Fazekas, F. & Miller, D. H. Diagnosis of multiple sclerosis: progress and challenges. Lancet 389, 1336–1346 (2017).

55. McNicholas, N., Hutchinson, M., McGuigan, C. & Chataway, J. 2017 McDonald diagnostic criteria: A review of the evidence. Mult Scler Relat Disord 24, 48–54 (2018).

56. Thompson, A. J. et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 17, 162–173 (2018).

57. Loma, I. & Heyman, R. Multiple sclerosis: pathogenesis and treatment. Curr Neuropharmacol 9, 409–416 (2011).

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

59. Lassmann, H., van Horssen, J. & Mahad, D. Progressive multiple sclerosis: pathology and pathogenesis. Nat Rev Neurol 8, 647–656 (2012).

60. Noseworthy, J. H., Lucchinetti, C., Rodriguez, M. & Weinshenker, B. G. Multiple sclerosis. N Engl J Med 343, 938–952 (2000).

61. Disanto, G., Morahan, J. M., Barnett, M. H., Giovannoni, G. & Ramagopalan, S. V. The evidence for a role of B cells in multiple sclerosis. Neurology 78, 823–832 (2012).

62. Sospedra, M. B cells in multiple sclerosis. Curr Opin Neurol 31, 256–262 (2018).

63. Thompson, A. J., Baranzini, S. E., Geurts, J., Hemmer, B. & Ciccarelli, O. Multiple sclerosis. Lancet 391, 1622–1636 (2018).

64. Lublin, F. D. et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 83, 278–286 (2014).

65. Klineova, S. & Lublin, F. D. Clinical Course of Multiple Sclerosis. Cold Spring Harb Perspect Med 8, a028928 (2018).

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

67. Pérez-Rico, C. et al. Evaluation of visual structural and functional factors that predict the development of multiple sclerosis in clinically isolated syndrome patients. Invest Ophthalmol Vis Sci 55, 6127–6131 (2014).

68. Marcus, J. F. & Waubant, E. L. Updates on clinically isolated syndrome and diagnostic criteria for multiple sclerosis. Neurohospitalist 3, 65–80 (2013).

69. Friese, M. A., Schattling, B. & Fugger, L. Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. Nat Rev Neurol 10, 225–238 (2014).

70. Trapp, B. D. & Nave, K.-A. Multiple sclerosis: an immune or neurodegenerative disorder? Annu Rev Neurosci 31, 247–269 (2008).

71. Ziemssen, T. et al. A mixed methods approach towards understanding key disease characteristics associated with the progression from RRMS to SPMS: Physicians' and patients' views. Mult Scler Relat Disord 38, 101861 (2020).

72. Katz Sand, I. Classification, diagnosis, and differential diagnosis of multiple sclerosis. Curr Opin Neurol 28, 193–205 (2015).

73. Lorscheider, J. et al. Defining secondary progressive multiple sclerosis. Brain 139, 2395–2405 (2016).

74. Weiner, H. L. The challenge of multiple sclerosis: how do we cure a chronic heterogeneous disease? Ann Neurol 65, 239–248 (2009).

75. Dargahi, N. et al. Multiple Sclerosis: Immunopathology and Treatment Update. Brain Sci 7, 78 (2017).

76. The Multiple Sclerosis International Federation (MSIF). Atlas of MS, 3rd Edition. Part 2: Clinical management of multiple sclerosis around the world. (April 2021).

77. Hauser, S. L. & Cree, B. A. C. Treatment of Multiple Sclerosis: A Review. Am J Med 133, 1380-1390.e2 (2020).

78. McDonell, J. et al. World Health Organization Essential Medicines List: Multiple sclerosis disease-modifying therapies application. Mult Scler 26, 153–158 (2020).

79. Voge, N. V. & Alvarez, E. Monoclonal Antibodies in Multiple Sclerosis: Present and Future. Biomedicines 7, 20 (2019).

80. Raftopoulos, R. et al. Phenytoin for neuroprotection in patients with acute optic neuritis: a randomised, placebo-controlled, phase 2 trial. Lancet Neurol 15, 259–269 (2016).

81. Fox, R. J. et al. Phase 2 Trial of Ibudilast in Progressive Multiple Sclerosis. N Engl J Med 379, 846–855 (2018).

82. Green, A. J. et al. Clemastine fumarate as a remyelinating therapy for multiple sclerosis (ReBUILD): a randomised, controlled, double-blind, crossover trial. Lancet 390, 2481–2489 (2017).

83. Mancardi, G., Sormani, M. P., Muraro, P. A., Boffa, G. & Saccardi, R. Intense immunosuppression followed by autologous haematopoietic stem cell transplantation as a therapeutic strategy in aggressive forms of multiple sclerosis. Mult Scler 24, 245–255 (2018).

84. Massey, J. C., Sutton, I. J., Ma, D. D. F. & Moore, J. J. Regenerating Immunotolerance in Multiple Sclerosis with Autologous Hematopoietic Stem Cell Transplant. Front Immunol 9, 410 (2018).

85. Sormani, M. P. et al. Autologous hematopoietic stem cell transplantation in multiple sclerosis: A meta-analysis. Neurology 88, 2115–2122 (2017).

86. Karussis, D., Petrou, P., Vourka-Karussis, U. & Kassis, I. Hematopoietic stem cell transplantation in multiple sclerosis. Expert Rev Neurother 13, 567–578 (2013).

87. Radaelli, M. et al. Autologous bone marrow transplantation for the treatment of multiple sclerosis. Curr Neurol Neurosci Rep 14, 478 (2014).

88. Rush, C. A., Atkins, H. L. & Freedman, M. S. Autologous Hematopoietic Stem Cell Transplantation in the Treatment of Multiple Sclerosis. Cold Spring Harb Perspect Med 9, a029082 (2019).

89. Jangi, S. et al. Alterations of the human gut microbiome in multiple sclerosis. Nat Commun 7, 12015 (2016).

90. Chen, J. et al. Multiple sclerosis patients have a distinct gut microbiota compared to healthy controls. Sci Rep 6, 28484 (2016).

91. Miyake, S. et al. Dysbiosis in the Gut Microbiota of Patients with Multiple Sclerosis, with a Striking Depletion of Species Belonging to Clostridia XIVa and IV Clusters. PLoS One 10, e0137429 (2015).

92. Takewaki, D. et al. Alterations of the gut ecological and functional microenvironment in different stages of multiple sclerosis. Proc Natl Acad Sci U S A 117, 22402–22412 (2020).

93. Kohl, H. M., Castillo, A. R. & Ochoa-Repáraz, J. The Microbiome as a Therapeutic Target for Multiple Sclerosis: Can Genetically Engineered Probiotics Treat the Disease? Diseases 8, 33 (2020).

94. Salami, M., Kouchaki, E., Asemi, Z. & Tamtaji, O. R. How probiotic bacteria influence the motor and mental behaviors as well as immunological and oxidative biomarkers in multiple sclerosis? A double blind clinical trial. Journal of Functional Foods 52, 8–13 (2019).

95. Tankou, S. K. et al. Investigation of probiotics in multiple sclerosis. Mult Scler 24, 58–63 (2018).

96. Jiang, J. et al. Efficacy of probiotics in multiple sclerosis: a systematic review of preclinical trials and meta-analysis of randomized controlled trials. Food Funct 12, 2354–2377 (2021).

97. Mestre, L. et al. Manipulation of Gut Microbiota Influences Immune Responses, Axon Preservation, and Motor Disability in a Model of Progressive Multiple Sclerosis. Front Immunol 10, 1374 (2019).

98. 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 17, 165–174 (2018).

99. Makkawi, S., Camara-Lemarroy, C. & Metz, L. Fecal microbiota transplantation associated with 10 years of stability in a patient with SPMS. Neurol Neuroimmunol Neuroinflamm 5, e459 (2018).

100. Bar-Or, A. et al. Induction of antigen-specific tolerance in multiple sclerosis after immunization with DNA encoding myelin basic protein in a randomized, placebo-controlled phase 1/2 trial. Arch Neurol 64, 1407–1415 (2007).

101. Bielekova, B. et al. Encephalitogenic potential of the myelin basic protein peptide (amino acids 83-99) in multiple sclerosis: results of a phase II clinical trial with an altered peptide ligand. Nat Med 6, 1167–1175 (2000).

102. Al-Ghobashy, M. A. et al. Development and Pre-Clinical Evaluation of Recombinant Human Myelin Basic Protein Nano Therapeutic Vaccine in Experimental Autoimmune Encephalomyelitis Mice Animal Model. Sci Rep 7, 46468 (2017).

103. Karussis, D. et al. T cell vaccination benefits relapsing progressive multiple sclerosis patients: a randomized, double-blind clinical trial. PLoS One 7, e50478 (2012).

104. Motl, R. W. et al. Exercise in patients with multiple sclerosis. Lancet Neurol 16, 848–856 (2017).

105. Ciccarelli, O. et al. Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. Lancet Neurol 13, 807–822 (2014).

106. Baecher-Allan, C., Kaskow, B. J. & Weiner, H. L. Multiple Sclerosis: Mechanisms and Immunotherapy. Neuron 97, 742–768 (2018).

107. Titus, H. E. et al. Pre-clinical and Clinical Implications of 'Inside-Out' vs. 'Outside-In' Paradigms in Multiple Sclerosis Etiopathogenesis. Front Cell Neurosci 14, 599717 (2020).

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

109. Waisman, A., Liblau, R. S. & Becher, B. Innate and adaptive immune responses in the CNS. Lancet Neurol 14, 945–955 (2015).

110. Salou, M., Nicol, B., Garcia, A. & Laplaud, D.-A. Involvement of CD8(+) T Cells in Multiple Sclerosis. Front Immunol 6, 604 (2015).

111. Pegoretti, V. et al. Inflammation and Oxidative Stress in Multiple Sclerosis: Consequences for Therapy Development. Oxid Med Cell Longev 2020, 7191080 (2020).

112. Macchi, B. et al. Role of inflammation and apoptosis in multiple sclerosis: Comparative analysis between the periphery and the central nervous system. J Neuroimmunol 287, 80–87 (2015).

113. Lassmann, H. Pathogenic Mechanisms Associated With Different Clinical Courses of Multiple Sclerosis. Front Immunol 9, 3116 (2018).

114. Peterson, L. K. & Fujinami, R. S. Inflammation, demyelination, neurodegeneration and neuroprotection in the pathogenesis of multiple sclerosis. J Neuroimmunol 184, 37–44 (2007).

115. Domingues, H. S., Portugal, C. C., Socodato, R. & Relvas, J. B. Oligodendrocyte, Astrocyte, and Microglia Crosstalk in Myelin Development, Damage, and Repair. Front Cell Dev Biol 4, 71 (2016).

116. Guerrero, B. L. & Sicotte, N. L. Microglia in Multiple Sclerosis: Friend or Foe? Front Immunol 11, 374 (2020).

117. Luo, C. et al. The role of microglia in multiple sclerosis. Neuropsychiatr Dis Treat 13, 1661–1667 (2017).

118. Wilbanks, B., Maher, L. J. & Rodriguez, M. Glial cells as therapeutic targets in progressive multiple sclerosis. Expert Rev Neurother 19, 481–494 (2019).

119. Miller, S. J. Astrocyte Heterogeneity in the Adult Central Nervous System. Front Cell Neurosci 12, 401 (2018).

120. Ponath, G., Park, C. & Pitt, D. The Role of Astrocytes in Multiple Sclerosis. Front Immunol 9, 217 (2018).

121. Yeung, M. S. Y. et al. Dynamics of oligodendrocyte generation and myelination in the human brain. Cell 159, 766–774 (2014).

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

123. Raphael, I., Nalawade, S., Eagar, T. N. & Forsthuber, T. G. T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. Cytokine 74, 5–17 (2015).

124. Høglund, R. A. & Maghazachi, A. A. Multiple sclerosis and the role of immune cells. World J Exp Med 4, 27–37 (2014).

125. González, H. & Pacheco, R. T-cell-mediated regulation of neuroinflammation involved in neurodegenerative diseases. J Neuroinflammation 11, 201 (2014).

126. Kleinewietfeld, M. & Hafler, D. A. Regulatory T cells in autoimmune neuroinflammation. Immunol Rev 259, 231–244 (2014).

127. Kostic, M., Stojanovic, I., Marjanovic, G., Zivkovic, N. & Cvetanovic, A. Deleterious versus protective autoimmunity in multiple sclerosis. Cell Immunol 296, 122–132 (2015).

128. Jiang, Q. et al. Role of Th22 Cells in the Pathogenesis of Autoimmune Diseases. Front Immunol 12, 688066 (2021).

129. van Furth, R., Raeburn, J. A. & van Zwet, T. L. Characteristics of human mononuclear phagocytes. Blood 54, 485–500 (1979).

130. Ziegler-Heitbrock, L. Blood Monocytes and Their Subsets: Established Features and Open Questions. Front Immunol 6, 423 (2015).

131. Boyette, L. B. et al. Phenotype, function, and differentiation potential of human monocyte subsets. PLoS One 12, e0176460 (2017).

132. Guilliams, M., Mildner, A. & Yona, S. Developmental and Functional Heterogeneity of Monocytes. Immunity 49, 595–613 (2018).

133. Ziegler-Heitbrock, L. et al. Nomenclature of monocytes and dendritic cells in blood. Blood 116, e74-80 (2010).

134. Zawada, A. M. et al. Monocyte heterogeneity in human cardiovascular disease. Immunobiology 217, 1273–1284 (2012).

135. Patel, A. A. et al. The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. J Exp Med 214, 1913–1923 (2017).

136. Kapellos, T. S. et al. Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. Front Immunol 10, 2035 (2019).

137. Anbazhagan, K., Duroux-Richard, I., Jorgensen, C. & Apparailly, F. Transcriptomic network support distinct roles of classical and non-classical monocytes in human. Int Rev Immunol 33, 470–489 (2014).

138. Grage-Griebenow, E. et al. Identification of a novel dendritic cell-like subset of CD64(+) / CD16(+) blood monocytes. Eur J Immunol 31, 48–56 (2001).

139. Belge, K.-U. et al. The proinflammatory CD14+CD16+DR++ monocytes are a major source of TNF. J Immunol 168, 3536–3542 (2002).

140. Zawada, A. M. et al. SuperSAGE evidence for CD14++CD16+ monocytes as a third monocyte subset. Blood 118, e50-61 (2011).

141. Lee, J. et al. The MHC class II antigen presentation pathway in human monocytes differs by subset and is regulated by cytokines. PLoS One 12, e0183594 (2017).

142. Wong, K. L. et al. The three human monocyte subsets: implications for health and disease. Immunol Res 53, 41–57 (2012).

143. Akaishi, T., Takahashi, T. & Nakashima, I. Peripheral blood monocyte count at onset may affect the prognosis in multiple sclerosis. J Neuroimmunol 319, 37–40 (2018).

144. Gjelstrup, M. C. et al. Subsets of activated monocytes and markers of inflammation in incipient and progressed multiple sclerosis. Immunol Cell Biol 96, 160–174 (2018).

145. Chuluundorj, D., Harding, S. A., Abernethy, D. & La Flamme, A. C. Expansion and preferential activation of the CD14(+)CD16(+) monocyte subset during multiple sclerosis. Immunol Cell Biol 92, 509–517 (2014).

146. Makhlouf, K., Weiner, H. L. & Khoury, S. J. Increased percentage of IL-12+ monocytes in the blood correlates with the presence of active MRI lesions in MS. J Neuroimmunol 119, 145–149 (2001).

147. Kouwenhoven, M., Teleshova, N., Ozenci, V., Press, R. & Link, H. Monocytes in multiple sclerosis: phenotype and cytokine profile. J Neuroimmunol 112, 197–205 (2001).

148. Mulder, K. et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. Immunity 54, 1883-1900.e5 (2021).

149. Jordão, M. J. C. et al. Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. Science 363, eaat7554 (2019).

150. Almsned, F., Lipsky, R. H. & Jafri, M. S. Transcriptomic analysis of Multiple Sclerosis patient-derived monocytes by RNA-Sequencing for candidate gene discovery. Informatics in Medicine Unlocked 23, 100563 (2021).

151. Segura, E. et al. Characterization of resident and migratory dendritic cells in human lymph nodes. J Exp Med 209, 653–660 (2012).

152. Haniffa, M., Bigley, V. & Collin, M. Human mononuclear phagocyte system reunited. Semin Cell Dev Biol 41, 59–69 (2015).

153. Collin, M. & Bigley, V. Human dendritic cell subsets: an update. Immunology 154, 3–20 (2018).

154. Boltjes, A. & van Wijk, F. Human dendritic cell functional specialization in steady-state and inflammation. Front Immunol 5, 131 (2014).

155. Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356, eaah4573 (2017).

156. Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. Nature 392, 245–252 (1998).

157. Luessi, F., Zipp, F. & Witsch, E. Dendritic cells as therapeutic targets in neuroinflammation. Cell Mol Life Sci 73, 2425–2450 (2016).

158. Manicassamy, S. & Pulendran, B. Dendritic cell control of tolerogenic responses. Immunol Rev 241, 206–227 (2011).

159. Domogalla, M. P., Rostan, P. V., Raker, V. K. & Steinbrink, K. Tolerance through Education: How Tolerogenic Dendritic Cells Shape Immunity. Front Immunol 8, 1764 (2017).

160. Boks, M. A. et al. IL-10-generated tolerogenic dendritic cells are optimal for functional regulatory T cell induction--a comparative study of human clinical-applicable DC. Clin Immunol 142, 332–342 (2012).

161. Schraml, B. U. & Reis e Sousa, C. Defining dendritic cells. Curr Opin Immunol 32, 13–20 (2015).

162. Qian, C. & Cao, X. Dendritic cells in the regulation of immunity and inflammation. Semin Immunol 35, 3–11 (2018).

163. Tang-Huau, T.-L. & Segura, E. Human in vivo-differentiated monocyte-derived dendritic cells. Semin Cell Dev Biol 86, 44–49 (2019).

164. Miyagawa, F., Tagaya, Y., Ozato, K., Horie, K. & Asada, H. Inflammatory monocyte-derived dendritic cells mediate autoimmunity in murine model of systemic lupus erythematosus. J Transl Autoimmun 3, 100060 (2020).

165. Coutant, F., Pin, J.-J. & Miossec, P. Extensive Phenotype of Human Inflammatory Monocyte-Derived Dendritic Cells. Cells 10, 1663 (2021).

166. Gao, Y. et al. Single-Cell Analysis Reveals the Heterogeneity of Monocyte-Derived and Peripheral Type-2 Conventional Dendritic Cells. J Immunol 207, 837–848 (2021).

167. Nair, S., Archer, G. E. & Tedder, T. F. Isolation and generation of human dendritic cells. Curr Protoc Immunol Chapter 7, 7.32.1-7.32.23 (2012).

168. Posch, W., Lass-Flörl, C. & Wilflingseder, D. Generation of Human Monocyte-derived Dendritic Cells from Whole Blood. J Vis Exp 54968 (2016).

169. Chometon, T. Q. et al. A protocol for rapid monocyte isolation and generation of singular human monocyte-derived dendritic cells. PLoS One 15, e0231132 (2020).

170. Chapuis, F. et al. Differentiation of human dendritic cells from monocytes in vitro. Eur J Immunol 27, 431–441 (1997).

171. Hashimoto, S. I. et al. Identification of genes specifically expressed in human activated and mature dendritic cells through serial analysis of gene expression. Blood 96, 2206–2214 (2000).

172. Le Naour, F. et al. Profiling changes in gene expression during differentiation and maturation of monocyte-derived dendritic cells using both oligonucleotide microarrays and proteomics. J Biol Chem 276, 17920–17931 (2001).

173. Lehtonen, A. et al. Gene expression profiling during differentiation of human monocytes to macrophages or dendritic cells. J Leukoc Biol 82, 710–720 (2007).

174. Schinnerling, K., García-González, P. & Aguillón, J. C. Gene Expression Profiling of Human Monocyte-derived Dendritic Cells - Searching for Molecular Regulators of Tolerogenicity. Front Immunol 6, 528 (2015).

175. Dutertre, C.-A. et al. Single-Cell Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies Circulating Inflammatory Dendritic Cells. Immunity 51, 573-589.e8 (2019).

176. Schlitzer, A., McGovern, N. & Ginhoux, F. Dendritic cells and monocyte-derived cells: Two complementary and integrated functional systems. Semin Cell Dev Biol 41, 9–22 (2015).

177. Pashenkov, M. et al. Two subsets of dendritic cells are present in human cerebrospinal fluid. Brain 124, 480–492 (2001).

178. Greter, M. et al. Dendritic cells permit immune invasion of the CNS in an animal model of multiple sclerosis. Nat Med 11, 328–334 (2005).

179. Wu, G. F. & Laufer, T. M. The role of dendritic cells in multiple sclerosis. Curr Neurol Neurosci Rep 7, 245–252 (2007).

180. Lande, R. et al. Plasmacytoid dendritic cells in multiple sclerosis: intracerebral recruitment and impaired maturation in response to interferon-beta. J Neuropathol Exp Neurol 67, 388–401 (2008).

181. Longhini, A. L. F. et al. Plasmacytoid dendritic cells are increased in cerebrospinal fluid of untreated patients during multiple sclerosis relapse. J Neuroinflammation 8, 2 (2011).

182. Thomas, K. et al. Accumulation and therapeutic modulation of 6-sulfo LacNAc(+) dendritic cells in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 1, e33 (2014).

183. Nuyts, A. H., Lee, W. P., Bashir-Dar, R., Berneman, Z. N. & Cools, N. Dendritic cells in multiple sclerosis: key players in the immunopathogenesis, key players for new cellular immunotherapies? Mult Scler 19, 995–1002 (2013).

184. Stasiolek, M. et al. Impaired maturation and altered regulatory function of plasmacytoid dendritic cells in multiple sclerosis. Brain 129, 1293–1305 (2006).

185. López, C., Comabella, M., Al-zayat, H., Tintoré, M. & Montalban, X. Altered maturation of circulating dendritic cells in primary progressive MS patients. J Neuroimmunol 175, 183–191 (2006).

186. Huang, Y. M. et al. Multiple sclerosis is associated with high levels of circulating dendritic cells secreting pro-inflammatory cytokines. J Neuroimmunol 99, 82–90 (1999).

187. Quintana, F. J., Yeste, A. & Mascanfroni, I. D. Role and therapeutic value of dendritic cells in central nervous system autoimmunity. Cell Death Differ 22, 215–224 (2015).

188. Deshpande, P., King, I. L. & Segal, B. M. Cutting edge: CNS CD11c+ cells from mice with encephalomyelitis polarize Th17 cells and support CD25+CD4+ T cell-mediated immunosuppression, suggesting dual roles in the disease process. J Immunol 178, 6695–6699 (2007).

189. Steinman, R. M. & Banchereau, J. Taking dendritic cells into medicine. Nature 449, 419–426 (2007).

190. Zhou, F., Ciric, B., Zhang, G.-X. & Rostami, A. Immunotherapy using lipopolysaccharide-stimulated bone marrow-derived dendritic cells to treat experimental autoimmune encephalomyelitis. Clin Exp Immunol 178, 447–458 (2014).

191. Comabella, M., Montalban, X., Münz, C. & Lünemann, J. D. Targeting dendritic cells to treat multiple sclerosis. Nat Rev Neurol 6, 499–507 (2010).

192. Flórez-Grau, G., Zubizarreta, I., Cabezón, R., Villoslada, P. & Benitez-Ribas, D. Tolerogenic Dendritic Cells as a Promising Antigen-Specific Therapy in the Treatment of Multiple Sclerosis and Neuromyelitis Optica From Preclinical to Clinical Trials. Front Immunol 9, 1169 (2018).

193. Saxena, A., Martin-Blondel, G., Mars, L. T. & Liblau, R. S. Role of CD8 T cell subsets in the pathogenesis of multiple sclerosis. FEBS Lett 585, 3758–3763 (2011).

194. Huseby, E. S., Huseby, P. G., Shah, S., Smith, R. & Stadinski, B. D. Pathogenic CD8 T cells in multiple sclerosis and its experimental models. Front Immunol 3, 64 (2012).

195. Mars, L. T., Saikali, P., Liblau, R. S. & Arbour, N. Contribution of CD8 T lymphocytes to the immuno-pathogenesis of multiple sclerosis and its animal models. Biochim Biophys Acta 1812, 151–161 (2011).

196. Legroux, L. & Arbour, N. Multiple Sclerosis and T Lymphocytes: An Entangled Story. J Neuroimmune Pharmacol 10, 528–546 (2015).

197. Cui, Y. & Wan, Q. NKT Cells in Neurological Diseases. Front Cell Neurosci 13, 245 (2019).

198. Forsthuber, T. G., Cimbora, D. M., Ratchford, J. N., Katz, E. & Stüve, O. B cell-based therapies in CNS autoimmunity: differentiating CD19 and CD20 as therapeutic targets. Ther Adv Neurol Disord 11, 1756286418761697 (2018).

199. Thomas, M. D., Srivastava, B. & Allman, D. Regulation of peripheral B cell maturation. Cell Immunol 239, 92–102 (2006).

200. Wekerle, H. B cells in multiple sclerosis. Autoimmunity 50, 57–60 (2017).

201. Morandi, B. et al. Role of natural killer cells in the pathogenesis and progression of multiple sclerosis. Pharmacol Res 57, 1–5 (2008).

202. Chanvillard, C., Jacolik, R. F., Infante-Duarte, C. & Nayak, R. C. The role of natural killer cells in multiple sclerosis and their therapeutic implications. Front Immunol 4, 63 (2013).

203. Gross, C. C. et al. Regulatory Functions of Natural Killer Cells in Multiple Sclerosis. Front Immunol 7, 606 (2016).

204. Ginhoux, F. & Jung, S. Monocytes and macrophages: developmental pathways and tissue homeostasis. Nat Rev Immunol 14, 392–404 (2014).

205. Hendriks, J. J. A., Teunissen, C. E., de Vries, H. E. & Dijkstra, C. D. Macrophages and neurodegeneration. Brain Res Brain Res Rev 48, 185–195 (2005).

206. Vogel, D. Y. S. et al. Macrophages in inflammatory multiple sclerosis lesions have an intermediate activation status. J Neuroinflammation 10, 35 (2013).

207. Waubant, E. et al. Environmental and genetic risk factors for MS: an integrated review. Ann Clin Transl Neurol 6, 1905–1922 (2019).

208. Olsson, T., Barcellos, L. F. & Alfredsson, L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. Nat Rev Neurol 13, 25–36 (2017).

209. Hempel, S. et al. A systematic review of modifiable risk factors in the progression of multiple sclerosis. Mult Scler 23, 525–533 (2017).

210. Thorley-Lawson, D. A. Epstein-Barr virus: exploiting the immune system. Nat Rev Immunol 1, 75–82 (2001).

211. Handel, A. E. et al. An updated meta-analysis of risk of multiple sclerosis following infectious mononucleosis. PLoS One 5, e12496 (2010).

212. Bjornevik, K. et al. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. Science 375, 296–301 (2022).

213. Almohmeed, Y. H., Avenell, A., Aucott, L. & Vickers, M. A. Systematic review and meta-analysis of the sero-epidemiological association between Epstein Barr virus and multiple sclerosis. PLoS One 8, e61110 (2013).

214. Jacobs, B. M., Giovannoni, G., Cuzick, J. & Dobson, R. Systematic review and meta-analysis of the association between Epstein-Barr virus, multiple sclerosis and other risk factors. Mult Scler 26, 1281–1297 (2020).

215. Serafini, B. et al. Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. J Exp Med 204, 2899–2912 (2007).

216. Sundqvist, E. et al. Epstein-Barr virus and multiple sclerosis: interaction with HLA. Genes Immun 13, 14–20 (2012).

217. Xu, L. et al. Positive association of herpes simplex virus-IgG with multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord 47, 102633 (2021).

218. Zhang, W. et al. Mendelian Randomization Analysis Suggests No Associations of Herpes Simplex Virus Infections With Multiple Sclerosis. Front Neurosci 16, 817067 (2022).

219. Grut, V. et al. Cytomegalovirus seropositivity is associated with reduced risk of multiple sclerosis-a presymptomatic case-control study. Eur J Neurol 28, 3072–3079 (2021).

220. Vanheusden, M., Stinissen, P., ’t Hart, B. A. & Hellings, N. Cytomegalovirus: a culprit or protector in multiple sclerosis? Trends Mol Med 21, 16–23 (2015).

221. Cusick, M. F., Libbey, J. E. & Fujinami, R. S. Molecular mimicry as a mechanism of autoimmune disease. Clin Rev Allergy Immunol 42, 102–111 (2012).

222. Wingerchuk, D. M. Smoking: effects on multiple sclerosis susceptibility and disease progression. Ther Adv Neurol Disord 5, 13–22 (2012).

223. Hernán, M. A. et al. Cigarette smoking and the progression of multiple sclerosis. Brain 128, 1461–1465 (2005).

224. Sundström, P., Nyström, L. & Hallmans, G. Smoke exposure increases the risk for multiple sclerosis. Eur J Neurol 15, 579–583 (2008).

225. Hedström, A. K., Bäärnhielm, M., Olsson, T. & Alfredsson, L. Exposure to environmental tobacco smoke is associated with increased risk for multiple sclerosis. Mult Scler 17, 788–793 (2011).

226. Alrouji, M., Manouchehrinia, A., Gran, B. & Constantinescu, C. S. Effects of cigarette smoke on immunity, neuroinflammation and multiple sclerosis. J Neuroimmunol 329, 24–34 (2019).

227. Hedström, A. K., Olsson, T. & Alfredsson, L. Smoking is a major preventable risk factor for multiple sclerosis. Mult Scler 22, 1021–1026 (2016).

228. Hedström, A. K. et al. Smokers run increased risk of developing anti-natalizumab antibodies. Mult Scler 20, 1081–1085 (2014).

229. Rhead, B. et al. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. Neurol Genet 2, e97 (2016).

230. Aranow, C. Vitamin D and the immune system. J Investig Med 59, 881–886 (2011).

231. Smolders, J. et al. The relevance of vitamin D receptor gene polymorphisms for vitamin D research in multiple sclerosis. Autoimmun Rev 8, 621–626 (2009).

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

233. Ascherio, A. et al. Vitamin D as an early predictor of multiple sclerosis activity and progression. JAMA Neurol 71, 306–314 (2014).

234. Sandberg, L. et al. Vitamin D and axonal injury in multiple sclerosis. Mult Scler 22, 1027–1031 (2016).

235. Wasnik, S., Sharma, I., Baylink, D. J. & Tang, X. Vitamin D as a Potential Therapy for Multiple Sclerosis: Where Are We? Int J Mol Sci 21, 3102 (2020).

236. Nair, R. & Maseeh, A. Vitamin D: The 'sunshine' vitamin. J Pharmacol Pharmacother 3, 118–126 (2012).

237. Lucas, R. M. et al. Sun exposure and vitamin D are independent risk factors for CNS demyelination. Neurology 76, 540–548 (2011).

238. Bäärnhielm, M. et al. Sunlight is associated with decreased multiple sclerosis risk: no interaction with human leukocyte antigen-DRB1*15. Eur J Neurol 19, 955–962 (2012).

239. Lucas, R. M., Byrne, S. N., Correale, J., Ilschner, S. & Hart, P. H. Ultraviolet radiation, vitamin D and multiple sclerosis. Neurodegener Dis Manag 5, 413–424 (2015).

240. Munger, K. L., Chitnis, T. & Ascherio, A. Body size and risk of MS in two cohorts of US women. Neurology 73, 1543–1550 (2009).

241. Munger, K. L. et al. Childhood body mass index and multiple sclerosis risk: a long-term cohort study. Mult Scler 19, 1323–1329 (2013).

242. Mokry, L. E. et al. Obesity and Multiple Sclerosis: A Mendelian Randomization Study. PLoS Med 13, e1002053 (2016).

243. Gianfrancesco, M. A. et al. Causal Effect of Genetic Variants Associated With Body Mass Index on Multiple Sclerosis Susceptibility. Am J Epidemiol 185, 162–171 (2017).

244. Massa, J., O'Reilly, E. J., Munger, K. L. & Ascherio, A. Caffeine and alcohol intakes have no association with risk of multiple sclerosis. Mult Scler 19, 53–58 (2013).

245. Hedström, A. K., Hillert, J., Olsson, T. & Alfredsson, L. Alcohol as a modifiable lifestyle factor affecting multiple sclerosis risk. JAMA Neurol 71, 300–305 (2014).

246. Farez, M. F., Fiol, M. P., Gaitán, M. I., Quintana, F. J. & Correale, J. Sodium intake is associated with increased disease activity in multiple sclerosis. J Neurol Neurosurg Psychiatry 86, 26–31 (2015).

247. Fitzgerald, K. C. et al. Sodium intake and multiple sclerosis activity and progression in BENEFIT. Ann Neurol 82, 20–29 (2017).

248. Kishikawa, T. et al. A Metagenome-Wide Association Study of Gut Microbiome in Patients With Multiple Sclerosis Revealed Novel Disease Pathology. Front Cell Infect Microbiol 10, 585973 (2020).

249. Cantoni, C. et al. Alterations of host-gut microbiome interactions in multiple sclerosis. EBioMedicine 76, 103798 (2022).

250. Berer, K. et al. Gut microbiota from multiple sclerosis patients enables spontaneous autoimmune encephalomyelitis in mice. Proc Natl Acad Sci U S A 114, 10719–10724 (2017).

251. Zheng, D., Liwinski, T. & Elinav, E. Interaction between microbiota and immunity in health and disease. Cell Res 30, 492–506 (2020).

252. Joscelyn, J. & Kasper, L. H. Digesting the emerging role for the gut microbiome in central nervous system demyelination. Mult Scler 20, 1553–1559 (2014).

253. Artemiadis, A. K., Anagnostouli, M. C. & Alexopoulos, E. C. Stress as a risk factor for multiple sclerosis onset or relapse: a systematic review. Neuroepidemiology 36, 109–120 (2011).

254. Riise, T. et al. Stress and the risk of multiple sclerosis. Neurology 76, 1866–1871 (2011).

255. Gasperini, C. et al. A controlled study of potential risk factors preceding exacerbation in multiple sclerosis. J Neurol Neurosurg Psychiatry 59, 303–305 (1995).

256. Buljevac, D. et al. Self reported stressful life events and exacerbations in multiple sclerosis: prospective study. BMJ 327, 646 (2003).

257. Mohr, D. C., Hart, S. L., Julian, L., Cox, D. & Pelletier, D. Association between stressful life events and exacerbation in multiple sclerosis: a meta-analysis. BMJ 328, 731 (2004).

258. Briones-Buixassa, L. et al. Stress and multiple sclerosis: A systematic review considering potential moderating and mediating factors and methods of assessing stress. Health Psychol Open 2, 2055102915612271 (2015).

259. Mohr, D. C. Stress and multiple sclerosis. J Neurol 254 Suppl 2, II65-68 (2007).

260. Oksenberg, J. R. Decoding multiple sclerosis: an update on genomics and future directions. Expert Rev Neurother 13, 11–19 (2013).

261. Canto, E. & Oksenberg, J. R. Multiple sclerosis genetics. Mult Scler 24, 75–79 (2018).

262. Oksenberg, J. R. & Baranzini, S. E. Multiple sclerosis genetics--is the glass half full, or half empty? Nat Rev Neurol 6, 429–437 (2010).

263. Willer, C. J. et al. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci U S A 100, 12877–12882 (2003).

264. Hansen, T. et al. Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. Mult Scler 11, 504–510 (2005).

265. Ingelfinger, F. et al. Twin study reveals non-heritable immune perturbations in multiple sclerosis. Nature 603, 152–158 (2022).

266. Fagnani, C. et al. Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. Mult Scler 21, 1404–1413 (2015).

267. Patsopoulos, N. A. & De Jager, P. L. Genetic and gene expression signatures in multiple sclerosis. Mult Scler 26, 576–581 (2020).

268. Hoppenbrouwers, I. A. & Hintzen, R. Q. Genetics of multiple sclerosis. Biochim Biophys Acta 1812, 194–201 (2011).

269. Wang, Z. et al. Nuclear Receptor NR1H3 in Familial Multiple Sclerosis. Neuron 90, 948–954 (2016).

270. Sadovnick, A. D. et al. Analysis of Plasminogen Genetic Variants in Multiple Sclerosis Patients. G3 (Bethesda) 6, 2073–2079 (2016).

271. Vilariño-Güell, C. et al. Exome sequencing in multiple sclerosis families identifies 12 candidate genes and nominates biological pathways for the genesis of disease. PLoS Genet 15, e1008180 (2019).

272. Muñoz-Culla, M., Irizar, H. & Otaegui, D. The genetics of multiple sclerosis: review of current and emerging candidates. Appl Clin Genet 6, 63–73 (2013).

273. Patsopoulos, N. A. Genetics of Multiple Sclerosis: An Overview and New Directions. Cold Spring Harb Perspect Med 8, a028951 (2018).

274. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science 365, eaav7188 (2019).

275. Cotsapas, C. & Mitrovic, M. Genome-wide association studies of multiple sclerosis. Clin Transl Immunology 7, e1018 (2018).

276. Jersild, C., Svejgaard, A. & Fog, T. HL-A antigens and multiple sclerosis. Lancet 1, 1240–1241 (1972).

277. Naito, S., Namerow, N., Mickey, M. R. & Terasaki, P. I. Multiple sclerosis: association with HL-A3. Tissue Antigens 2, 1–4 (1972).

278. Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J. K. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet 54, 15–39 (2009).

279. Robinson, J. et al. IPD-IMGT/HLA Database. Nucleic Acids Res 48, D948–D955 (2020).

280. Parnell, G. P. & Booth, D. R. The Multiple Sclerosis (MS) Genetic Risk Factors Indicate both Acquired and Innate Immune Cell Subsets Contribute to MS Pathogenesis and Identify Novel Therapeutic Opportunities. Front Immunol 8, 425 (2017).

281. Gourraud, P.-A., Harbo, H. F., Hauser, S. L. & Baranzini, S. E. The genetics of multiple sclerosis: an up-to-date review. Immunol Rev 248, 87–103 (2012).

282. Okuda, D. T. et al. Genotype-Phenotype correlations in multiple sclerosis: HLA genes influence disease severity inferred by 1HMR spectroscopy and MRI measures. Brain 132, 250–259 (2009).

283. Barcellos, L. F. et al. HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. Am J Hum Genet 72, 710–716 (2003).

284. Oksenberg, J. R. et al. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. Am J Hum Genet 74, 160–167 (2004).

285. Schmidt, H., Williamson, D. & Ashley-Koch, A. HLA-DR15 haplotype and multiple sclerosis: a HuGE review. Am J Epidemiol 165, 1097–1109 (2007).

286. Hollenbach, J. A. & Oksenberg, J. R. The immunogenetics of multiple sclerosis: A comprehensive review. J Autoimmun 64, 13–25 (2015).

287. Sadovnick, A. D. Genetic background of multiple sclerosis. Autoimmun Rev 11, 163–166 (2012).

288. Brynedal, B. et al. HLA-A confers an HLA-DRB1 independent influence on the risk of multiple sclerosis. PLoS One 2, e664 (2007).

289. Healy, B. C. et al. HLA B*44: protective effects in MS susceptibility and MRI outcome measures. Neurology 75, 634–640 (2010).

290. McElroy, J. P. & Oksenberg, J. R. Multiple sclerosis genetics 2010. Neurol Clin 29, 219–231 (2011).

291. Vandenbroeck, K. et al. A cytokine gene screen uncovers SOCS1 as genetic risk factor for multiple sclerosis. Genes Immun 13, 21–28 (2012).

292. Beyeen, A. D. et al. IL-22RA2 associates with multiple sclerosis and macrophage effector mechanisms in experimental neuroinflammation. J Immunol 185, 6883–6890 (2010).

293. Gómez-Fernández, P. et al. The Rare IL22RA2 Signal Peptide Coding Variant rs28385692 Decreases Secretion of IL-22BP Isoform-1, -2 and -3 and Is Associated with Risk for Multiple Sclerosis. Cells 9, 175 (2020).

294. International Multiple Sclerosis Genetics Consortium et al. MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. Brain 136, 1778–1782 (2013).

295. Gil-Varea, E. et al. Exome sequencing study in patients with multiple sclerosis reveals variants associated with disease course. J Neuroinflammation 15, 265 (2018).

296. Scazzone, C. et al. Association of CYP2R1 rs10766197 with MS risk and disease progression. J Neurosci Res 96, 297–304 (2018).

297. Baranzini, S., Sawcer, S., International Multiple Sclerosis Genetics Consortium & MultipleMS Consortium. Genetic analysis of multiple sclerosis severity identifies a novel locus and implicates CNS resilience as a major determinant of outcome. Research Square (2022).

298. Lill, C. M. Recent advances and future challenges in the genetics of multiple sclerosis. Front Neurol 5, 130 (2014).

299. International Multiple Sclerosis Genetics Consortium. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. Cell 175, 1679-1687.e7 (2018).

300. Mo, X.-B. et al. Integrative analysis revealed potential causal genetic and epigenetic factors for multiple sclerosis. J Neurol 266, 2699–2709 (2019).

301. Koch, M. W., Metz, L. M. & Kovalchuk, O. Epigenetic changes in patients with multiple sclerosis. Nat Rev Neurol 9, 35–43 (2013).

302. Küçükali, C. İ., Kürtüncü, M., Çoban, A., Çebi, M. & Tüzün, E. Epigenetics of multiple sclerosis: an updated review. Neuromolecular Med 17, 83–96 (2015).

303. Otaegui, D. et al. Differential micro RNA expression in PBMC from multiple sclerosis patients. PLoS One 4, e6309 (2009).

304. Gandhi, R. miRNA in multiple sclerosis: search for novel biomarkers. Mult Scler 21, 1095–1103 (2015).

305. Baranzini, S. E. & Oksenberg, J. R. The Genetics of Multiple Sclerosis: From 0 to 200 in 50 Years. Trends Genet 33, 960–970 (2017).

306. Umeton, R. et al. Multiple sclerosis genetic and non-genetic factors interact through the transient transcriptome. Sci Rep 12, 7536 (2022).

307. Kim, K. et al. Cell type-specific transcriptomics identifies neddylation as a novel therapeutic target in multiple sclerosis. Brain 144, 450–461 (2021).

308. International Multiple Sclerosis Genetics Consortium. A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. Nat Commun 10, 2236 (2019).

309. Jäkel, S. et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. Nature 566, 543–547 (2019).

310. Shams, H. et al. Polygenic risk score association with multiple sclerosis susceptibility and phenotype in Europeans. Brain awac092 (2022).

311. De Jager, P. L. et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol 8, 1111–1119 (2009).

312. Dobson, R. et al. A Risk Score for Predicting Multiple Sclerosis. PLoS One 11, e0164992 (2016).

313. Misicka, E., Sept, C. & Briggs, F. B. S. Predicting onset of secondary-progressive multiple sclerosis using genetic and non-genetic factors. J Neurol 267, 2328–2339 (2020).

314. Hone, L., Giovannoni, G., Dobson, R. & Jacobs, B. M. Predicting Multiple Sclerosis: Challenges and Opportunities. Front. Neurol. 12, 761973 (2022).

315. Cunningham, F. et al. Ensembl 2022. Nucleic Acids Res 50, D988–D995 (2022).

316. Sherry, S. T., Ward, M. & Sirotkin, K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res 9, 677–679 (1999).

317. Lopez de Lapuente, A. et al. Novel Insights into the Multiple Sclerosis Risk Gene ANKRD55. J Immunol 196, 4553–4565 (2016).

318. Alloza, I. et al. ANKRD55 and DHCR7 are novel multiple sclerosis risk loci. Genes Immun 13, 253–257 (2012).

319. Lill, C. M. et al. Genome-wide significant association of ANKRD55 rs6859219 and multiple sclerosis risk. J Med Genet 50, 140–143 (2013).

320. Dankowski, T. et al. Successful Replication of GWAS Hits for Multiple Sclerosis in 10,000 Germans Using the Exome Array. Genet Epidemiol 39, 601–608 (2015).

321. Andlauer, T. F. M. et al. Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. Sci Adv 2, e1501678 (2016).

322. International Multiple Sclerosis Genetics Consortium (IMSGC) et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 45, 1353–1360 (2013).

323. Stahl, E. A. et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42, 508–514 (2010).

324. Eyre, S. et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 44, 1336–1340 (2012).

325. Kim, K. et al. High-density genotyping of immune loci in Koreans and Europeans identifies eight new rheumatoid arthritis risk loci. Ann Rheum Dis 74, e13 (2015).

326. Wei, W.-H. et al. Major histocompatibility complex harbors widespread genotypic variability of non-additive risk of rheumatoid arthritis including epistasis. Sci Rep 6, 25014 (2016).

327. Márquez, A. et al. Meta-analysis of Immunochip data of four autoimmune diseases reveals novel single-disease and cross-phenotype associations. Genome Med 10, 97 (2018).

328. Laufer, V. A. et al. Genetic influences on susceptibility to rheumatoid arthritis in African-Americans. Hum Mol Genet 28, 858–874 (2019).

329. Okada, Y. et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506, 376–381 (2014).

330. Salehi, R., Motaghi, M., Salehi, A., Karimzadeh, H. & Pakzad, B. Impact of Single Nucleotide Polymorphism in the ANKRD55 Gene on Occurrence and Clinical Characteristics of Rheumatoid Arthritis. Avicenna J Med Biotechnol 14, 259–263 (2022).

331. Mathebula, E. M. et al. A genome-wide association study for rheumatoid arthritis replicates previous HLA and non-HLA associations in a cohort from South Africa. Hum Mol Genet 31, 4286–4294 (2022).

332. Hinks, A. et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet 45, 664–669 (2013).

333. Jia, J. et al. Genetic architecture study of rheumatoid arthritis and juvenile idiopathic arthritis. PeerJ 8, e8234 (2020).

334. López-Isac, E. et al. Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms. Ann Rheum Dis 80, 321–328 (2021).

335. Li, Y. R. et al. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. Nat Med 21, 1018–1027 (2015).

336. Dubois, P. C. A. et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42, 295–302 (2010).

337. Zhernakova, A. et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. PLoS Genet 7, e1002004 (2011).

338. Liu, J. Z. et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet 47, 979–986 (2015).

339. Huang, H. et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. Nature 547, 173–178 (2017).

340. Jostins, L. et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491, 119–124 (2012).

341. Robertson, C. C. et al. Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. Nat Genet 53, 962–971 (2021).

342. Westra, H.-J. et al. Fine-mapping and functional studies highlight potential causal variants for rheumatoid arthritis and type 1 diabetes. Nat Genet 50, 1366–1374 (2018).

343. Morris, A. P. et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 44, 981–990 (2012).

344. Imamura, M. et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. Nat Commun 7, 10531 (2016).

345. Harder, M. N. et al. Type 2 diabetes risk alleles near BCAR1 and in ANK1 associate with decreased β-cell function whereas risk alleles near ANKRD55 and GRB14 associate with

decreased insulin sensitivity in the Danish Inter99 cohort. J Clin Endocrinol Metab 98, E801-806 (2013).

346. Fang, Y.-D. et al. Association of ANKRD55 Gene Polymorphism with HT: A Protective Factor for Disease Susceptibility. Int J Endocrinol 2022, 7300796 (2022).

347. Shiga, Y. et al. Genome-wide association study identifies seven novel susceptibility loci for primary open-angle glaucoma. Hum Mol Genet 27, 1486–1496 (2018).

348. Wei, W.-H., Massey, J., Worthington, J., Barton, A. & Warren, R. B. Genotypic variability-based genome-wide association study identifies non-additive loci HLA-C and IL12B for psoriasis. J Hum Genet 63, 289–296 (2018).

349. Li, L. et al. Positive Association between ANKRD55 Polymorphism 7731626 and Dermatomyositis/Polymyositis with Interstitial Lung Disease in Chinese Han Population. Biomed Res Int 2017, 2905987 (2017).

350. Sherva, R. et al. Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. Alzheimers Dement 10, 45–52 (2014).

351. Xu, J. et al. Inherited genetic variant predisposes to aggressive but not indolent prostate cancer. Proc Natl Acad Sci U S A 107, 2136–2140 (2010).

352. Stein, M. B. et al. Genome-wide Association Studies of Posttraumatic Stress Disorder in 2 Cohorts of US Army Soldiers. JAMA Psychiatry 73, 695–704 (2016).

353. Fuh-Ngwa, V. et al. Ensemble machine learning identifies genetic loci associated with future worsening of disability in people with multiple sclerosis. Sci Rep 12, 19291 (2022).

354. Ji, Y. et al. Genome-Wide and Abdominal MRI Data Provide Evidence That a Genetically Determined Favorable Adiposity Phenotype Is Characterized by Lower Ectopic Liver Fat and Lower Risk of Type 2 Diabetes, Heart Disease, and Hypertension. Diabetes 68, 207–219 (2019).

355. Jonsson, S. et al. Identification of sequence variants influencing immunoglobulin levels. Nat Genet 49, 1182–1191 (2017).

356. Freudenberg, J., Gregersen, P. & Li, W. Enrichment of Genetic Variants for Rheumatoid Arthritis within T-Cell and NK-Cell Enhancer Regions. Mol Med 21, 180–184 (2015).

357. Lauc, G. et al. Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. PLoS Genet 9, e1003225 (2013).

358. James, T. et al. Impact of genetic risk loci for multiple sclerosis on expression of proximal genes in patients. Hum Mol Genet 27, 912–928 (2018).

359. Chun, S. et al. Limited statistical evidence for shared genetic effects of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. Nat Genet 49, 600–605 (2017).

360. Xi, J. et al. The Cancer Genome Atlas dataset-based analysis of aberrantly expressed genes by GeneAnalytics in thymoma associated myasthenia gravis: focusing on T cells. J Thorac Dis 11, 2315–2323 (2019).

361. Razavi, S. M. et al. Comprehensive functional enrichment analysis of male infertility. Sci Rep 7, 15778 (2017).

362. Houtman, M. et al. T-cell transcriptomics from peripheral blood highlights differences between polymyositis and dermatomyositis patients. Arthritis Res Ther 20, 188 (2018).

363. Luo, J. et al. Identifying Functional Modules in Co-Regulatory Networks Through Overlapping Spectral Clustering. IEEE Trans Nanobioscience 17, 134–144 (2018).

364. Tian, Y. et al. The Significance of Tumor Microenvironment Score for Breast Cancer Patients. Biomed Res Int 2022, 5673810 (2022).

365. Korshunov, A. et al. Molecular analysis of pediatric CNS-PNET revealed nosologic heterogeneity and potent diagnostic markers for CNS neuroblastoma with FOXR2-activation. Acta Neuropathol Commun 9, 20 (2021).

366. Tang, R. et al. Role of tumor mutation burden-related signatures in the prognosis and immune microenvironment of pancreatic ductal adenocarcinoma. Cancer Cell Int 21, 196 (2021).

367. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res 51, D523–D531 (2023).

368. Uhlén, M. et al. Proteomics. Tissue-based map of the human proteome. Science 347, 1260419 (2015).

369. Stelzer, G. et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics 54, 1.30.1-1.30.33 (2016).

370. Hu, X. et al. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. Am J Hum Genet 89, 496–506 (2011).

371. Mosavi, L. K., Cammett, T. J., Desrosiers, D. C. & Peng, Z.-Y. The ankyrin repeat as molecular architecture for protein recognition. Protein Sci 13, 1435–1448 (2004).

372. Voronin, D. A. & Kiseleva, E. V. Functional role of proteins containing ankyrin repeats. Cell Tiss. Biol. 2, 1–12 (2008).

373. Li, J., Mahajan, A. & Tsai, M.-D. Ankyrin repeat: a unique motif mediating protein-protein interactions. Biochemistry 45, 15168–15178 (2006).

374. Kumar, A. & Balbach, J. Folding and Stability of Ankyrin Repeats Control Biological Protein Function. Biomolecules 11, 840 (2021).

375. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).

376. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 50, D439–D444 (2022).

377. Ugidos, N. et al. Interactome of the Autoimmune Risk Protein ANKRD55. Front Immunol 10, 2067 (2019).

378. Drew, K. et al. Integration of over 9,000 mass spectrometry experiments builds a global map of human protein complexes. Mol Syst Biol 13, 932 (2017).

379. Chen, Z. et al. Functional Screening of Candidate Causal Genes for Insulin Resistance in Human Preadipocytes and Adipocytes. Circ Res 126, 330–346 (2020).

380. Shen, Z.-J., Hu, J., Esnault, S., Dozmorov, I. & Malter, J. S. RNA Seq profiling reveals a novel expression pattern of TGF-β target genes in human blood eosinophils. Immunol Lett 167, 1–10 (2015).

381. Park, D. et al. Differences in the molecular signatures of mucosal-associated invariant T cells and conventional T cells. Sci Rep 9, 7094 (2019).

382. Rose-John, S. Interleukin-6 Family Cytokines. Cold Spring Harb Perspect Biol 10, a028415 (2018).

383. Murakami, M., Kamimura, D. & Hirano, T. Pleiotropy and Specificity: Insights from the Interleukin 6 Family of Cytokines. Immunity 50, 812–831 (2019).

384. Dienz, O. & Rincon, M. The effects of IL-6 on CD4 T cell responses. Clin Immunol 130, 27–33 (2009).

385. Hunter, C. A. & Jones, S. A. IL-6 as a keystone cytokine in health and disease. Nat Immunol 16, 448–457 (2015).

386. Yao, X. et al. Targeting interleukin-6 in inflammatory autoimmune diseases and cancers. Pharmacol Ther 141, 125–139 (2014).

387. Janssens, K., Slaets, H. & Hellings, N. Immunomodulatory properties of the IL-6 cytokine family in multiple sclerosis. Ann N Y Acad Sci 1351, 52–60 (2015).

388. Calabrese, L. H. & Rose-John, S. IL-6 biology: implications for clinical targeting in rheumatic disease. Nat Rev Rheumatol 10, 720–727 (2014).

389. Wolf, J., Rose-John, S. & Garbers, C. Interleukin-6 and its receptors: a highly regulated and dynamic system. Cytokine 70, 11–20 (2014).

390. Rose-John, S. IL-6 trans-signaling via the soluble IL-6 receptor: importance for the pro-inflammatory activities of IL-6. Int J Biol Sci 8, 1237–1247 (2012).

391. Scheller, J., Chalaris, A., Schmidt-Arras, D. & Rose-John, S. The pro- and anti-inflammatory properties of the cytokine interleukin-6. Biochim Biophys Acta 1813, 878–888 (2011).

392. Xu, S. & Neamati, N. gp130: a promising drug target for cancer therapy. Expert Opin Ther Targets 17, 1303–1328 (2013).

393. Honke, N. et al. The p38-mediated rapid down-regulation of cell surface gp130 expression impairs interleukin-6 signaling in the synovial fluid of juvenile idiopathic arthritis patients. Arthritis Rheumatol 66, 470–478 (2014).

394. Silver, J. S. & Hunter, C. A. gp130 at the nexus of inflammation, autoimmunity, and cancer. J Leukoc Biol 88, 1145–1156 (2010).

395. Luchtefeld, M. et al. Signal transducer of inflammation gp130 modulates atherosclerosis in mice and man. J Exp Med 204, 1935–1944 (2007).

396. Benrick, A. et al. A non-conservative polymorphism in the IL-6 signal transducer (IL6ST)/gp130 is associated with myocardial infarction in a hypertensive population. Regul Pept 146, 189–196 (2008).

397. Jones, S. A., Scheller, J. & Rose-John, S. Therapeutic strategies for the clinical blockade of IL-6/gp130 signaling. J Clin Invest 121, 3375–3383 (2011).

398. Schwerd, T. et al. A biallelic mutation in IL6ST encoding the GP130 co-receptor causes immunodeficiency and craniosynostosis. J Exp Med 214, 2547–2562 (2017).

399. Shahin, T. et al. Selective loss of function variants in IL6ST cause Hyper-IgE syndrome with distinct impairments of T-cell phenotype and function. Haematologica 104, 609–621 (2019).

400. Chen, Y.-H., Spencer, S., Laurence, A., Thaventhiran, J. E. & Uhlig, H. H. Inborn errors of IL-6 family cytokine responses. Curr Opin Immunol 72, 135–145 (2021).

401. Materna-Kiryluk, A. et al. Mosaic IL6ST variant inducing constitutive GP130 cytokine receptor signaling as a cause of neonatal onset immunodeficiency with autoinflammation and dysmorphy. Hum Mol Genet 30, 226–233 (2021).

402. Jostock, T. et al. Soluble gp130 is the natural inhibitor of soluble interleukin-6 receptor transsignaling responses. Eur J Biochem 268, 160–167 (2001).

403. Garbers, C. et al. Inhibition of classic signaling is a novel function of soluble glycoprotein 130 (sgp130), which is controlled by the ratio of interleukin 6 and soluble interleukin 6 receptor. J Biol Chem 286, 42959–42970 (2011).

404. Wolf, J. et al. Different Soluble Forms of the Interleukin-6 Family Signal Transducer gp130 Fine-tune the Blockade of Interleukin-6 Trans-signaling. J Biol Chem 291, 16186–16196 (2016).

405. Diamant, M. et al. Cloning and expression of an alternatively spliced mRNA encoding a soluble form of the human interleukin-6 signal transducer gp130. FEBS Lett 412, 379–384 (1997).

406. Tanaka, M. et al. Cloning of novel soluble gp130 and detection of its neutralizing autoantibodies in rheumatoid arthritis. J. Clin. Invest. 106, 137–144 (2000).

407. Sommer, J. et al. Alternative intronic polyadenylation generates the interleukin-6 trans-signaling inhibitor sgp130-E10. J Biol Chem 289, 22140–22150 (2014).

408. Xu, H. et al. Exploration of early-life candidate biomarkers for childhood asthma using antibody arrays. Pediatr Allergy Immunol 27, 696–701 (2016).

409. Askevold, E. T. et al. Soluble glycoprotein 130 predicts fatal outcomes in chronic heart failure: analysis from the Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA). Circ Heart Fail 6, 91–98 (2013).

410. Chen, H., Zhang, X., Liao, N. & Wen, F. Increased levels of IL-6, sIL-6R, and sgp130 in the aqueous humor and serum of patients with diabetic retinopathy. Mol Vis 22, 1005–1014 (2016).

411. Van Zaanen, H. C. et al. Blocking interleukin-6 activity with chimeric anti-IL6 monoclonal antibodies in multiple myeloma: effects on soluble IL6 receptor and soluble gp130. Leuk Lymphoma 31, 551–558 (1998).

412. Shariat, S. F. et al. Soluble gp130 regulates prostate cancer invasion and progression in an interleukin-6 dependent and independent manner. J Urol 186, 2107–2114 (2011).

413. Chakraborty, M. et al. Role of interleukin-6, its receptor and soluble gp130 in chronic lung disease of prematurity. Neonatology 104, 161–167 (2013).

414. Lemmers, A. et al. An inhibitor of interleukin-6 trans-signalling, sgp130, contributes to impaired acute phase response in human chronic liver disease. Clin Exp Immunol 156, 518–527 (2009).

415. Aparicio-Siegmund, S. et al. The IL-6-neutralizing sIL-6R-sgp130 buffer system is disturbed in patients with type 2 diabetes. Am J Physiol Endocrinol Metab 317, E411–E420 (2019).

416. Bonomi, A. et al. Analysis of the genetic variants associated with circulating levels of sgp130. Results from the IMPROVE study. Genes Immun 21, 100–108 (2020).

417. Hong, J. et al. Recombinant soluble gp130 protein reduces DEN-induced primary hepatocellular carcinoma in mice. Sci Rep 6, 24397 (2016).

418. Richards, P. J. et al. Functional characterization of a soluble gp130 isoform and its therapeutic capacity in an experimental model of inflammatory arthritis. Arthritis Rheum 54, 1662–1672 (2006).

419. Lee, S. Y. et al. IL-6 trans-signaling system in intra-amniotic inflammation, preterm birth, and preterm premature rupture of the membranes. J Immunol 186, 3226–3236 (2011).

420. Diveu, C. et al. GPL, a novel cytokine receptor related to GP130 and leukemia inhibitory factor receptor. J Biol Chem 278, 49850–49859 (2003).

421. Diveu, C. et al. Predominant expression of the long isoform of GP130-like (GPL) receptor is required for interleukin-31 signaling. Eur Cytokine Netw 15, 291–302 (2004).

422. Dreuw, A. et al. Characterization of the signaling capacities of the novel gp130-like cytokine receptor. J Biol Chem 279, 36112–36120 (2004).

423. Dillon, S. R. et al. Interleukin 31, a cytokine produced by activated T cells, induces dermatitis in mice. Nat Immunol 5, 752–760 (2004).

424. Zhang, Q., Putheti, P., Zhou, Q., Liu, Q. & Gao, W. Structures and biological functions of IL-31 and IL-31 receptors. Cytokine Growth Factor Rev 19, 347–356 (2008).

425. Horejs-Hoeck, J. et al. Dendritic cells activated by IFN-γ/STAT1 express IL-31 receptor and release proinflammatory mediators upon IL-31 treatment. J Immunol 188, 5319–5326 (2012).

426. Kasraie, S., Niebuhr, M. & Werfel, T. Interleukin (IL)-31 induces pro-inflammatory cytokines in human monocytes and macrophages following stimulation with staphylococcal exotoxins. Allergy 65, 712–721 (2010).

427. Sonkoly, E. et al. IL-31: a new link between T cells and pruritus in atopic skin inflammation. J Allergy Clin Immunol 117, 411–417 (2006).

428. Neis, M. M. et al. Enhanced expression levels of IL-31 correlate with IL-4 and IL-13 in atopic and allergic contact dermatitis. J Allergy Clin Immunol 118, 930–937 (2006).

429. Lin, M.-W. et al. Novel IL31RA gene mutation and ancestral OSMR mutant allele in familial primary cutaneous amyloidosis. Eur J Hum Genet 18, 26–32 (2010).

430. Kim, H. J. et al. Itch in dermatomyositis: the role of increased skin interleukin-31. Br J Dermatol 179, 669–678 (2018).

431. He, Y. et al. Interleukin-31 Receptor α Is Required for Basal-Like Breast Cancer Progression. Front Oncol 10, 816 (2020).

432. Rocak, S. & Linder, P. DEAD-box proteins: the driving forces behind RNA metabolism. Nat Rev Mol Cell Biol 5, 232–241 (2004).

433. Castrillon, D. H., Quade, B. J., Wang, T. Y., Quigley, C. & Crum, C. P. The human VASA gene is specifically expressed in the germ cell lineage. Proc Natl Acad Sci U S A 97, 9585–9590 (2000).

434. Boellaard, W. P. A., Stoop, H., Gillis, A. J. M., Oosterhuis, J. W. & Looijenga, L. H. J. VASA mRNA (DDX4) detection is more specific than immunohistochemistry using poly- or monoclonal antibodies for germ cells in the male urogenital tract. Medicine (Baltimore) 96, e7489 (2017).

435. Clarkson, Y. L. et al. Initial characterisation of adult human ovarian cell populations isolated by DDX4 expression and aldehyde dehydrogenase activity. Sci Rep 8, 6953 (2018).

436. Li, H., Yu, N., Zhang, X., Jin, W. & Li, H. Spermatozoal protein profiles in male infertility with asthenozoospermia. Chin Med J (Engl) 123, 2879–2882 (2010).

437. Hashimoto, H. et al. Germ cell specific protein VASA is over-expressed in epithelial ovarian cancer and disrupts DNA damage-induced G2 checkpoint. Gynecol Oncol 111, 312–319 (2008).

438. Chen, Y. et al. Vitamin D and DDX4 regulate the proliferation and invasion of ovarian cancer cells. Oncol Lett 16, 905–909 (2018).

439. Rebsamen, M. et al. SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. Nature 519, 477–481 (2015).

440. Wang, S. et al. Metabolism. Lysosomal amino acid transporter SLC38A9 signals arginine sufficiency to mTORC1. Science 347, 188–194 (2015).

441. Scalise, M. et al. Insights into the transport side of the human SLC38A9 transceptor. Biochim Biophys Acta Biomembr 1861, 1558–1567 (2019).

442. Liu, G. Y. & Sabatini, D. M. mTOR at the nexus of nutrition, growth, ageing and disease. Nat Rev Mol Cell Biol 21, 183–203 (2020).

443. Wyant, G. A. et al. mTORC1 Activator SLC38A9 Is Required to Efflux Essential Amino Acids from Lysosomes and Use Protein as a Nutrient. Cell 171, 642-654.e12 (2017).

444. Castellano, B. M. et al. Lysosomal cholesterol activates mTORC1 via an SLC38A9-Niemann-Pick C1 signaling complex. Science 355, 1306–1311 (2017).

445. Uzozie, A. C. & Aebersold, R. Advancing translational research and precision medicine with targeted proteomics. J Proteomics 189, 1–10 (2018).

446. Singh, V., Tripathi, A. & Dutta, R. Proteomic Approaches to Decipher Mechanisms Underlying Pathogenesis in Multiple Sclerosis Patients. Proteomics 19, e1800335 (2019).

447. Ragnedda, G. et al. Protein-protein interaction analysis highlights additional loci of interest for multiple sclerosis. PLoS One 7, e46730 (2012).

448. Safari-Alighiarloo, N., Rezaei-Tavirani, M., Taghizadeh, M., Tabatabaei, S. M. & Namaki, S. Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. PeerJ 4, e2775 (2016).

449. Yurduseven, K., Babal, Y. K., Celik, E., Kerman, B. E. & Kurnaz, I. A. Multiple Sclerosis Biomarker Candidates Revealed by Cell-Type-Specific Interactome Analysis. OMICS 26, 305–317 (2022).

450. Sen, M. K., Almuslehi, M. S. M., Shortland, P. J., Mahns, D. A. & Coorssen, J. R. Proteomics of Multiple Sclerosis: Inherent Issues in Defining the Pathoetiology and Identifying (Early) Biomarkers. Int J Mol Sci 22, 7377 (2021).

451. Berge, T. et al. Quantitative proteomic analyses of CD4+ and CD8+ T cells reveal differentially expressed proteins in multiple sclerosis patients and healthy controls. Clin Proteomics 16, 19 (2019).

452. Miedema, A., Wijering, M. H. C., Eggen, B. J. L. & Kooistra, S. M. High-Resolution Transcriptomic and Proteomic Profiling of Heterogeneity of Brain-Derived Microglia in Multiple Sclerosis. Front Mol Neurosci 13, 583811 (2020).

453. Ajami, B. et al. Single-cell mass cytometry reveals distinct populations of brain myeloid cells in mouse neuroinflammation and neurodegeneration models. Nat Neurosci 21, 541–551 (2018).

454. Safari-Alighiarloo, N., Taghizadeh, M., Mohammad Tabatabaei, S., Namaki, S. & Rezaei-Tavirani, M. Identification of common key genes and pathways between type 1 diabetes and multiple sclerosis using transcriptome and interactome analysis. Endocrine 68, 81–92 (2020).

455. Clark, A. D. et al. Lymphocyte DNA methylation mediates genetic risk at shared immune-mediated disease loci. J Allergy Clin Immunol 145, 1438–1451 (2020).

456. Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45, 1238–1243 (2013).

457. Kundu, K. et al. Genetic associations at regulatory phenotypes improve fine-mapping of causal variants for 12 immune-mediated diseases. Nat Genet 54, 251–262 (2022).

458. Roostaei, T. et al. Proximal and distal effects of genetic susceptibility to multiple sclerosis on the T cell epigenome. Nat Commun 12, 7078 (2021).

459. Astle, W. J. et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 167, 1415-1429.e19 (2016).

460. Chen, M.-H. et al. Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. Cell 182, 1198-1213.e14 (2020).

461. Vuckovic, D. et al. The Polygenic and Monogenic Basis of Blood Traits and Diseases. Cell 182, 1214-1231.e11 (2020).

462. Brod, S. A. & Bauer, V. L. Ingested (oral) tocilizumab inhibits EAE. Cytokine 68, 86–93 (2014).

463. Gómez-Fernández, P. et al. Long Interleukin-22 Binding Protein Isoform-1 Is an Intracellular Activator of the Unfolded Protein Response. Front Immunol 9, 2934 (2018).

464. Boniface, K. et al. Oncostatin M secreted by skin infiltrating T lymphocytes is a potent keratinocyte activator involved in skin inflammation. J Immunol 178, 4615–4622 (2007).

465. Martin, J. C. J. et al. Interleukin-22 binding protein (IL-22BP) is constitutively expressed by a subset of conventional dendritic cells and is strongly induced by retinoic acid. Mucosal Immunol 7, 101–113 (2014).

466. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682 (2012).

467. Shihan, M. H., Novo, S. G., Le Marchand, S. J., Wang, Y. & Duncan, M. K. A simple method for quantitating confocal fluorescent images. Biochem Biophys Rep 25, 100916 (2021).

468. Lindahl, H. et al. IL-22 Binding Protein Promotes the Disease Process in Multiple Sclerosis. J Immunol 203, 888–898 (2019).

469. Iwata, M. & Yokota, A. Retinoic acid production by intestinal dendritic cells. Vitam Horm 86, 127–152 (2011).

470. Erkelens, M. N. & Mebius, R. E. Retinoic Acid and Immune Homeostasis: A Balancing Act. Trends Immunol 38, 168–180 (2017).

471. Waschbisch, A. et al. Pivotal Role for CD16+ Monocytes in Immune Surveillance of the Central Nervous System. J Immunol 196, 1558–1567 (2016).

472. Schmiedel, B. J. et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell 175, 1701-1715.e16 (2018).

473. Spector, D. L. & Lamond, A. I. Nuclear speckles. Cold Spring Harb Perspect Biol 3, a000646 (2011).

474. Galganski, L., Urbanek, M. O. & Krzyzosiak, W. J. Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res 45, 10350–10368 (2017).

475. Sollis, E. et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res 51, D977–D985 (2023).

476. Ghoussaini, M. et al. Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res 49, D1311–D1320 (2021).

477. Padberg, F. et al. CSF and serum levels of soluble interleukin-6 receptors (sIL-6R and sgp130), but not of interleukin-6 are altered in multiple sclerosis. J Neuroimmunol 99, 218–223 (1999).

478. Bedri, S. K. et al. Multiple sclerosis treatment effects on plasma cytokine receptor levels. Clin Immunol 187, 15–25 (2018).

479. Mizuno, A. & Okada, Y. Biological characterization of expression quantitative trait loci (eQTLs) showing tissue-specific opposite directional effects. Eur J Hum Genet 27, 1745–1756 (2019).

480. Raj, T. et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science 344, 519–523 (2014).

481. Malhotra, S. et al. NLRP3 inflammasome as prognostic factor and therapeutic target in primary progressive multiple sclerosis patients. Brain 143, 1414–1430 (2020).

482. Duddy, M. E., Dickson, G., Hawkins, S. A. & Armstrong, M. A. Monocyte-derived dendritic cells: a potential target for therapy in multiple sclerosis (MS). Clin Exp Immunol 123, 280–287 (2001).

483. Segura, E. et al. Human inflammatory dendritic cells induce Th17 cell differentiation. Immunity 38, 336–348 (2013).

484. Dominitzki, S. et al. Cutting edge: trans-signaling via the soluble IL-6R abrogates the induction of FoxP3 in naive CD4+CD25 T cells. J Immunol 179, 2041–2045 (2007).

485. Bettelli, E. et al. Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. Nature 441, 235–238 (2006).

486. Kebir, H. et al. Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. Nat Med 13, 1173–1175 (2007).

487. Narazaki, M. et al. Soluble forms of the interleukin-6 signal-transducing receptor component gp130 in human serum possessing a potential to inhibit signals through membrane-anchored gp130. Blood 82, 1120–1126 (1993).

488. Wonnerth, A. et al. Glycoprotein 130 polymorphism predicts soluble glycoprotein 130 levels. Metabolism 63, 647–653 (2014).

489. Das, B. C. et al. Retinoic acid signaling pathways in development and diseases. Bioorg Med Chem 22, 673–683 (2014).

490. Széles, L. et al. Research resource: transcriptome profiling of genes regulated by RXR and its permissive and nonpermissive partners in differentiating monocyte-derived dendritic cells. Mol Endocrinol 24, 2218–2231 (2010).

491. Xie, P., Chan, F. S., Ip, N. Y. & Leung, M. F. IL-6 enhanced the retinoic acid-induced differentiation of human acute promyelocytic leukemia cells. Cancer Lett 148, 207–213 (2000).

492. Zhu, B. et al. IL-4 and retinoic acid synergistically induce regulatory dendritic cells expressing Aldh1a2. J Immunol 191, 3139–3151 (2013).

493. Ohoka, Y., Yokota-Nakatsuma, A., Maeda, N., Takeuchi, H. & Iwata, M. Retinoic acid and GM-CSF coordinately induce retinal dehydrogenase 2 (RALDH2) expression through cooperation between the RAR/RXR complex and Sp1 in dendritic cells. PLoS One 9, e96512 (2014).

494. Nagy, L., Szanto, A., Szatmari, I. & Széles, L. Nuclear hormone receptors enable macrophages and dendritic cells to sense their lipid environment and shape their immune response. Physiol Rev 92, 739–789 (2012).

495. Gyöngyösi, A. et al. RDH10, RALDH2, and CRABP2 are required components of PPARγ-directed ATRA synthesis and signaling in human dendritic cells. J Lipid Res 54, 2458–2474 (2013).

496. Fritsche, J., Stonehouse, T. J., Katz, D. R., Andreesen, R. & Kreutz, M. Expression of retinoid receptors during human monocyte differentiation in vitro. Biochem Biophys Res Commun 270, 17–22 (2000).

497. Zhang, R., Wang, Y., Li, R. & Chen, G. Transcriptional Factors Mediating Retinoic Acid Signals in the Control of Energy Metabolism. Int J Mol Sci 16, 14210–14244 (2015).

498. Reza Dorosty-Motlagh, A., Mohammadzadeh Honarvar, N., Sedighiyan, M. & Abdolahi, M. The Molecular Mechanisms of Vitamin A Deficiency in Multiple Sclerosis. J Mol Neurosci 60, 82–90 (2016).

499. Massacesi, L. et al. Suppression of experimental allergic encephalomyelitis by retinoic acid. J Neurol Sci 80, 55–64 (1987).

500. Honarvar, N. M. et al. In vitro effect of human serum and fetal calf serum on CD4+ T cells proliferation in response to myelin oligodendrocyte glycoprotein (MOG) in correlation with RBP/TTR ratio in multiple sclerotic patients. J Mol Neurosci 50, 571–576 (2013).

501. Vidal, M., Cusick, M. E. & Barabási, A.-L. Interactome networks and human disease. Cell 144, 986–998 (2011).

502. Henderson, J. M. et al. Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap® Analog by In Vitro Transcription. Curr Protoc 1, e39 (2021).

503. Prabhu, A., Sarcar, B., Kahali, S., Shan, Y. & Chinnaiyan, P. Targeting the unfolded protein response in glioblastoma cells with the fusion protein EGF-SubA. PLoS One 7, e52265 (2012).

504. Dozio, V. & Sanchez, J.-C. Profiling the proteomic inflammatory state of human astrocytes using DIA mass spectrometry. J Neuroinflammation 15, 331 (2018).

505. Chiavari, M., Ciotti, G. M. P., Navarra, P. & Lisi, L. Pro-Inflammatory Activation of A New Immortalized Human Microglia Cell Line. Brain Sci 9, 111 (2019).

506. Zhu, W. et al. Glioma-mediated microglial activation promotes glioma proliferation and migration: roles of Na+/H+ exchanger isoform 1. Carcinogenesis 37, 839–851 (2016).

507. Patel, A. B., Tsilioni, I., Leeman, S. E. & Theoharides, T. C. Neurotensin stimulates sortilin and mTOR in human microglia inhibitable by methoxyluteolin, a potential therapeutic target for autism. Proc Natl Acad Sci U S A 113, E7049–E7058 (2016).

508. Kovalevich, J. & Langford, D. Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods Mol Biol 1078, 9–21 (2013).

509. Lopez-Suarez, L., Awabdh, S. A., Coumoul, X. & Chauvet, C. The SH-SY5Y human neuroblastoma cell line, a relevant in vitro cell model for investigating neurotoxicology in human: Focus on organic pollutants. Neurotoxicology 92, 131–155 (2022).

510. Gilany, K. et al. The proteome of the human neuroblastoma cell line SH-SY5Y: an enlarged proteome. Biochim Biophys Acta 1784, 983–985 (2008).

511. Bosshart, H. & Heinzelmann, M. THP-1 cells as a model for human monocytes. Ann Transl Med 4, 438 (2016).

512. Chanput, W., Mes, J. J. & Wichers, H. J. THP-1 cell line: an in vitro cell model for immune modulation approach. Int Immunopharmacol 23, 37–45 (2014).

513. Wiśniewski, J. R. Filter Aided Sample Preparation - A tutorial. Anal Chim Acta 1090, 23–30 (2019).

514. Zybailov, B. L., Florens, L. & Washburn, M. P. Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. Mol Biosyst 3, 354–360 (2007).

515. Szklarczyk, D. et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 49, D605–D612 (2021).

516. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16, 169 (2015).

517. Ottmann, C. et al. Phosphorylation-independent interaction between 14-3-3 and exoenzyme S: from structure to pathogenesis. EMBO J 26, 902–913 (2007).

518. Jonckheere, A. I., Smeitink, J. A. M. & Rodenburg, R. J. T. Mitochondrial ATP synthase: architecture, function and pathology. J Inherit Metab Dis 35, 211–225 (2012).

519. Nakayama, K. & Katoh, Y. Architecture of the IFT ciliary trafficking machinery and interplay between its components. Crit Rev Biochem Mol Biol 55, 179–196 (2020).

520. Katoh, Y. et al. Overall Architecture of the Intraflagellar Transport (IFT)-B Complex Containing Cluap1/IFT38 as an Essential Component of the IFT-B Peripheral Subcomplex. J Biol Chem 291, 10962–10975 (2016).

521. Fu, H., Subramanian, R. R. & Masters, S. C. 14-3-3 proteins: structure, function, and regulation. Annu Rev Pharmacol Toxicol 40, 617–647 (2000).

522. Liu, J. et al. The role of 14-3-3 proteins in cell signalling pathways and virus infection. J Cell Mol Med 25, 4173–4182 (2021).

523. Cornell, B. & Toyo-Oka, K. 14-3-3 Proteins in Brain Development: Neurogenesis, Neuronal Migration and Neuromorphogenesis. Front Mol Neurosci 10, 318 (2017).

524. Obsilova, V. & Obsil, T. The 14-3-3 Proteins as Important Allosteric Regulators of Protein Kinases. Int J Mol Sci 21, 8824 (2020).

525. Munier, C. C., Ottmann, C. & Perry, M. W. D. 14-3-3 modulation of the inflammatory response. Pharmacol Res 163, 105236 (2021).

526. Fan, X. et al. 14-3-3 Proteins Are on the Crossroads of Cancer, Aging, and Age-Related Neurodegenerative Disease. Int J Mol Sci 20, 3518 (2019).

527. Pair, F. S. & Yacoubian, T. A. 14-3-3 Proteins: Novel Pharmacological Targets in Neurodegenerative Diseases. Trends Pharmacol Sci 42, 226–238 (2021).

528. Martínez-Yélamos, A. et al. 14-3-3 protein in the CSF as prognostic marker in early multiple sclerosis. Neurology 57, 722–724 (2001).

529. LoPresti, P. Serum-Based Biomarkers in Neurodegeneration and Multiple Sclerosis. Biomedicines 10, 1077 (2022).

530. Cho, E. & Park, J.-Y. Emerging roles of 14-3-3γ in the brain disorder. BMB Rep 53, 500–511 (2020).

531. Lee, D.-H. et al. Role of glial 14-3-3 gamma protein in autoimmune demyelination. J Neuroinflammation 12, 187 (2015).

532. Qin, L., Guo, J., Zheng, Q. & Zhang, H. BAG2 structure, function and involvement in disease. Cell Mol Biol Lett 21, 18 (2016).

533. Li, G., Li, D., Wang, T. & He, S. Pyrimidine Biosynthetic Enzyme CAD: Its Function, Regulation, and Diagnostic Potential. Int J Mol Sci 22, 10253 (2021).

534. Xu, L. et al. Biological effect of ribosomal protein L32 on human breast cancer cell behavior. Mol Med Rep 22, 2478–2486 (2020).

535. Chua, Y. S., Boh, B. K., Ponyeam, W. & Hagen, T. Regulation of cullin RING E3 ubiquitin ligases by CAND1 in vivo. PLoS One 6, e16071 (2011).

536. Kciuk, M., Gielecińska, A., Mujwar, S., Mojzych, M. & Kontek, R. Cyclin-dependent kinases in DNA damage response. Biochim Biophys Acta Rev Cancer 1877, 188716 (2022).

537. Misra, J. R. & Irvine, K. D. The Hippo Signaling Network and Its Biological Functions. Annu Rev Genet 52, 65–87 (2018).

538. Paudel, R., Fusi, L. & Schmidt, M. The MEK5/ERK5 Pathway in Health and Disease. Int J Mol Sci 22, 7594 (2021).

539. Katic, M. & Kahn, C. R. The role of insulin and IGF-1 signaling in longevity. Cell Mol Life Sci 62, 320–343 (2005).

540. Dyer, A. H., Vahdatpour, C., Sanfeliu, A. & Tropea, D. The role of Insulin-Like Growth Factor 1 (IGF-1) in brain development, maturation and neuroplasticity. Neuroscience 325, 89–99 (2016).

541. London, E., Bloyd, M. & Stratakis, C. A. PKA functions in metabolism and resistance to obesity: lessons from mouse and human studies. J Endocrinol 246, R51–R64 (2020).

542. Dagda, R. K. & Das Banerjee, T. Role of protein kinase A in regulating mitochondrial function and neuronal development: implications to neurodegenerative diseases. Rev Neurosci 26, 359–370 (2015).

543. Bahrami-B, F., Ataie-Kachoie, P., Pourgholami, M. H. & Morris, D. L. p70 Ribosomal protein S6 kinase (Rps6kb1): an update. J Clin Pathol 67, 1019–1025 (2014).

544. Patel, R. K. & Mohan, C. PI3K/AKT signaling and systemic autoimmunity. Immunol Res 31, 47–55 (2005).

545. Kim, J.-M. et al. The ERK MAPK Pathway Is Essential for Skeletal Development and Homeostasis. Int J Mol Sci 20, 1803 (2019).

546. Plotnikov, A., Zehorai, E., Procaccia, S. & Seger, R. The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. Biochim Biophys Acta 1813, 1619–1633 (2011).

547. Stolboushkina, E. A. & Garber, M. B. Eukaryotic type translation initiation factor 2: structure-functional aspects. Biochemistry (Mosc) 76, 283–294 (2011).

548. Drazic, A., Myklebust, L. M., Ree, R. & Arnesen, T. The world of protein acetylation. Biochim Biophys Acta 1864, 1372–1401 (2016).

549. Ishikawa, H. & Marshall, W. F. Intraflagellar Transport and Ciliary Dynamics. Cold Spring Harb Perspect Biol 9, a021998 (2017).

550. Lee, M.-S. et al. IFT46 plays an essential role in cilia development. Dev Biol 400, 248–257 (2015).

551. Xin, D., Christopher, K. J., Zeng, L., Kong, Y. & Weatherbee, S. D. IFT56 regulates vertebrate developmental patterning by maintaining IFTB complex integrity and ciliary microtubule architecture. Development 144, 1544–1553 (2017).

552. Huangfu, D. et al. Hedgehog signalling in the mouse requires intraflagellar transport proteins. Nature 426, 83–87 (2003).

553. Toriyama, M. et al. The ciliopathy-associated CPLANE proteins direct basal body recruitment of intraflagellar transport machinery. Nat Genet 48, 648–656 (2016).

554. Wang, W. et al. Intraflagellar Transport Proteins as Regulators of Primary Cilia Length. Front Cell Dev Biol 9, 661350 (2021).

555. Briscoe, J. & Thérond, P. P. The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol 14, 416–429 (2013).

556. Sipos, É., Komoly, S. & Ács, P. Quantitative Comparison of Primary Cilia Marker Expression and Length in the Mouse Brain. J Mol Neurosci 64, 397–409 (2018).

557. Sterpka, A. & Chen, X. Neuronal and astrocytic primary cilia in the mature brain. Pharmacol Res 137, 114–121 (2018).

558. Vitre, B., Guesdon, A. & Delaval, B. Non-ciliary Roles of IFT Proteins in Cell Division and Polycystic Kidney Diseases. Front Cell Dev Biol 8, 578239 (2020).

559. Baldari, C. T. & Rosenbaum, J. Intraflagellar transport: it's not just for cilia anymore. Curr Opin Cell Biol 22, 75–80 (2010).

560. Finetti, F. et al. Intraflagellar transport is required for polarized recycling of the TCR/CD3 complex to the immune synapse. Nat Cell Biol 11, 1332–1339 (2009).

561. Finetti, F., Paccani, S. R., Rosenbaum, J. & Baldari, C. T. Intraflagellar transport: a new player at the immune synapse. Trends Immunol 32, 139–145 (2011).

562. Schetters, S. T. T., Gomez-Nicola, D., Garcia-Vallejo, J. J. & Van Kooyk, Y. Neuroinflammation: Microglia and T Cells Get Ready to Tango. Front Immunol 8, 1905 (2017).

563. Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol 4, 712–720 (2003).

564. Ray Chaudhuri, A. & Nussenzweig, A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. Nat Rev Mol Cell Biol 18, 610–621 (2017).

565. Ray, U. & Raghavan, S. C. Understanding the DNA double-strand break repair and its therapeutic implications. DNA Repair (Amst) 106, 103177 (2021).

566. Huttlin, E. L. et al. Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. Cell 184, 3022-3040.e28 (2021).

567. Luck, K. et al. A reference map of the human binary protein interactome. Nature 580, 402–408 (2020).

568. Lu, M. et al. A code for RanGDP binding in ankyrin repeats defines a nuclear import pathway. Cell 157, 1130–1145 (2014).