medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

A Deep Transcriptome Meta-Analysis Reveals Sex-based Molecular Differences in Multiple Sclerosis

José Francisco Català-Senent[1†], Zoraida Andreu[2†], Marta R. Hidalgo[1], Francisco José Roig[1,3], Natalia Yanguas-Casás[4], Almudena Neva-Alejo[1], Adolfo López-Cerdán[5], Irene Soler-Sáez[1], María de la Iglesia-Vayá[5], Francisco García-García[1]

Correspondence: Francisco García-García, fgarcia@cipf.es

[1]Bioinformatics and Biostatistics Unit, Principe Felipe Research Center (CIPF), 46012, Valencia, Spain

[2]Foundation Valencian Institute of Oncology (FIVO), 46009, Valencia, Spain

[3]Faculty of Health Sciences. San Jorge University, 50830, Zaragoza, Spain

[4]Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana (IDIPHISA), Grupo de Investigación en Linfomas, C/Joaquín Rodrigo 2, Majadahonda, 28222 Madrid, Spain

[5]Biomedical Imaging Unit FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, 46012, Valencia, Spain

[†]José Francisco Català-Senent and Zoraida Andreu contributed equally to this study

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Abstract

Background: Multiple sclerosis (MS), a chronic auto-immune, inflammatory, and degenerative disease of the central nervous system, affects both males and females; however, females suffer from a higher risk of developing MS (2-3:1 ratio compared to males). Current knowledge does not allow a precise definition of the sex-based factors influencing MS. Here, we explore the role of sex in MS to identify potential molecular mechanisms underlying sex-based differences that may guide novel therapeutic approaches tailored for males or females.

Methods: We performed a rigorous and systematic review of whole transcriptome studies of MS that included patient information regarding sex in Gene Expression Omnibus and ArrayExpress databases following PRISMA statement guidelines. We analyzed differential gene expression for each selected study and addressed 3 meta-analyses based on genes to evaluate common features and sex bias: the first meta-analysis of 4 nervous tissue studies, a second in 5 blood studies, and a third integrating 9 studies from both tissues. Finally, we performed a gene set analysis on the meta-analyzed differential transcriptomic profiles of the nervous system to study sex-based differences in biological pathways and phenotypes (physiological and pathological states).

Results: After screening 122 publications, the systematic review provided a selection of 9 studies (5 in blood and 4 in nervous tissue) with a total of 474 samples (189 MS females and 109 control females; 82 MS males and 94 control males). The tissue-specific meta-analysis identified the overexpression of KIR2DL3 in blood in females and 13 genes with a sex-based differential expression pattern in the nervous system (7 overexpressed in females: ARL17B, CECR7, CEP78, STMP1, TRAF3IP2-AS1, ZNF117 and ZNF488; and 6 overexpressed in males: IFFO2, LOC401127, NUDT18, RNF10, SLC17A5, and UBXN2B). The two-tissue meta-analysis detected a single gene overexpressed in females (LOC102723701). Functional analyses revealed different altered immune scenarios in females and males. A pro-inflammatory environment and innate immune responses related to myeloid linage predominate in females, while in males, adaptative responses associated with the lymphocyte linage. Additionally, MS females displayed alterations in mitochondrial respiratory chain complexes,

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

purine, and glutamate metabolism, while MS males displayed alterations in stress response to metal ion, amine, and amino acid transport.

Conclusion: We found transcriptomic and functional differences between MS males and females (especially in the immune system), which may support the development of sex-specific treatments. Our study highlights the importance of understanding sex as a variable in MS.

Keywords: Multiple sclerosis, biomarkers, functional profiling, meta-analysis, sex-based differences, neurodegeneration

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Background

Multiple sclerosis (MS), a chronic autoimmune disease of the central nervous system (CNS), is also considered a neurodegenerative disease as progression occurs alongside the degeneration of the CNS, with exacerbated inflammatory reactions causing demyelination. The main symptoms of MS include fatigue, lack of balance, pain, visual and cognitive alterations, speech difficulties, and tremors. MS is mainly diagnosed in young adults (around 20-40 years) and represents the leading cause of disability in this age group that does not derive from trauma [1, 2]. Clinical isolated syndrome (CIS), the earliest diagnosed form of MS, is characterized by 24-hour episodes; meanwhile, relapsing-remitting MS (RRMS) as the most common disease form (85% of people affected) has milder symptoms that tend to appear as brief flare-ups followed by remission for periods of a few weeks or even months. Between 30-50% of patients initially suffering from RRMS develop secondary-progressive MS (SPMS), which is characterized by continuous progression with or without occasional relapses, minor remissions, and phases of stability. Primary progressive MS (PPMS) only affects 10% of all MS patients and is characterized by the absence of definite flare-ups; however, PPMS patients suffer from slow disease onset followed by a steady worsening of symptoms without an intermediate period of remission. Progressive relapsing MS (PRMS), an atypical form, is characterized by clear acute flares, with or without complete recovery, with periods between outbreaks continuously progressing [3]. Finally, 20% of patients suffering from benign MS suffer from mild flare-ups and complete or almost complete remission [4].

While the etiology of MS remains unknown, genetic and environmental factors impact disease-onset, and sex significantly influences many levels [5]. As with other autoimmune and neurodegenerative diseases, MS presents a strong sex bias in incidence, prevalence, symptom severity, disease course, response to therapy, and overall survival [6, 7]. While females have a higher susceptibility to MS than males, males are (on average) older than females at MS diagnosis and suffer from worse and more rapid progression. Males are also more likely to develop progressive MS than females [8, 9]. Sex-based differences in autoimmune and neurological diseases have been previously described [6, 8] and related to the presence of sexual chromosomes (XX vs. XY, gene dosage, imprinting, and transcriptomic profile) [9, 10] and endogenous sex hormones (estrogens, progesterone, androgens, and prolactin) that

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

impact the immune system and promote changes in disease severity. While estrogens represent potent stimulators of autoimmunity, androgens play a protective role [3]. Additionally, the transcriptomic profiles and epigenetic modifications on sex chromosomes influence hormonal actions in immune responses in both sexes. Responses to environmental factors such as microbial exposure, diet, and sociological differences, also represent sex-based factors in MS [11]. Therefore, differences in the prevalence and severity of MS between males and females involve complex and poorly understood interactions between genetic, hormonal, and environmental factors. This knowledge gap, in turn, impairs the development of personalized therapeutics based on sex-specific information.

Multi-omics-based studies in large and clinically well-characterized patient cohorts with accessible sex information data may identify critical pathways and networks related to sexual dimorphism and thereby support the discovery of new biomarkers and targeted sex-specific therapeutic interventions. *In silico* approaches using the large volumes of biological information accessible in public repositories of biomedical data, such as GEO (Genes Omnibus Expression) [12], represent a powerful research tool. In particular, meta-analyses of transcriptomic studies can uncover consensus patterns of expression from several studies focused on the same question due to their capacity for integration and greater statistical power than individual studies.

In this context, we performed a systematic review and selection of transcriptomic studies related to MS from GEO and ArrayExpress databases [12, 13] that included information on patient sex. We analyzed a total of 474 samples from nine different studies (189 MS females, 109 control females, 82 MS males, 94 control males). Given that MS presents neurodegenerative and autoimmune features, the systematic review included studies focused on nervous tissue and blood that reflected the autoimmune profile and could offer a means of discovering non-invasive biomarkers. Comparisons between males and females highlighted common genes involved in MS and differentially-expressed genes related to sex. After differential gene expression (DGE) analysis, we meta-analyzed nervous tissue and blood studies separately and combined in a third meta-analysis. Specific blood and nervous tissue analyses highlighted 1 (KIR2DL3) and 13 (ARL17B, CECR7, CEP78, IFFO2, LOC401127, NUDT18, RNF10, SLC17A5, STMP1, TRAF3IP2-AS1, UBXN2B, ZNF117, ZNF488) differentially-expressed

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

genes that associated with sex in MS patients, respectively. The two-tissue meta-analysis highlighted an RNA gene (LOC102723701) related to sex in MS patients. Subsequent functional enrichment analyses of the nervous tissue data provided vital information about biologic processes in males and females (mainly immune features) that allowed a better understanding of sex-based differences in MS. The identified molecular mechanisms may have important clinical implications for the prevention and treatment of MS and could potentially favor the development of novel, personalized diagnostic and therapeutic applications.

Methods

Literature Review and Study Selection

The systematic review was performed by searching for the keyword "multiple sclerosis" on the GEO and ArrayExpress [12, 13] transcriptomic databases. The reviewed period was from January 2002 (generation of the public repositories) to October 2020. Results were filtered by i) organism: "*Homo sapiens*," ii) study type: "expression profiling by array" or "expression profiling by high throughput sequencing," and iii) sample count: at least 12 samples. PRISMA Statement guidelines were followed to ensure the quality and transparency of the systematic review [14].

The following exclusion criteria were applied: i) studies not based on MS, ii) experimental design other than patients vs. controls, and iii) the absence of information on patient sex. In addition, studies with patients receiving treatment or who were immunosuppressed were also excluded.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

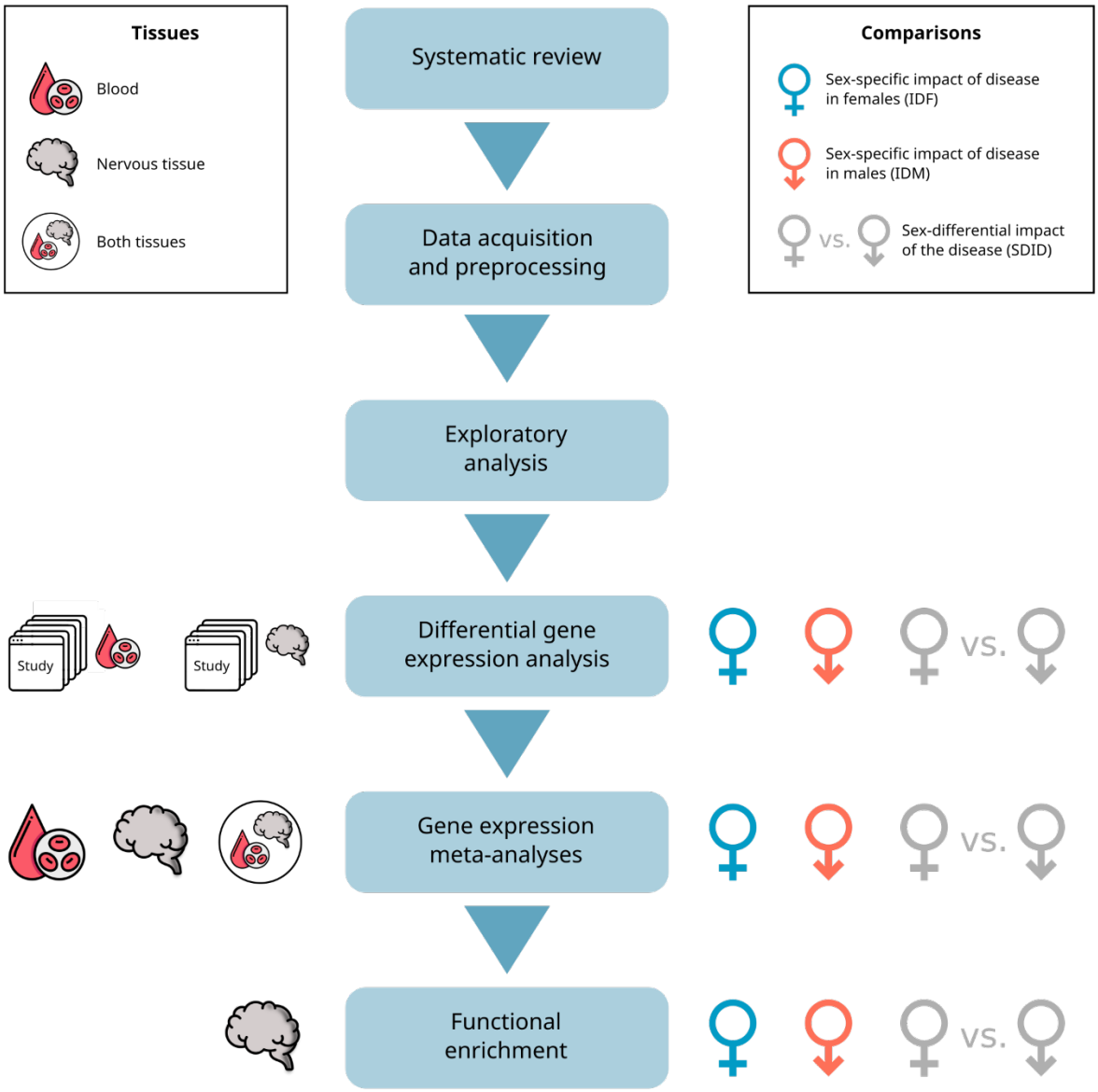

Figure 1. Flowchart describing the work performed, indicating tissues analyzed and comparisons performed in each step**.**

Bioinformatic Analyses

A defined analytical strategy was followed for each of the selected studies: i) data download and normalization, ii) exploratory analysis, iii) DGE analysis, iv) integration of DGE results using meta-analysis techniques (considering tissue of origin of samples to obtain robust results across all studies), and v) functional enrichment of meta-analyses data (Figure 1). By grouping the studies on a tissue-specific basis or all tissues together, this analysis allowed us to explore sex-based differences in the global disease and blood and nervous tissue. Bioinformatics analysis was carried out using R 4.1.2 [15]; the packages used and their version number are

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

described in Supplementary Table S1 and the following Zenodo repository (http://doi.org/10.5281/zenodo.6344450).

Data Acquisition and Preprocessing

First, study data were downloaded and standardized as follows: conversion and updating (if necessary) gene names to HGNC Gene Symbol nomenclature [16], followed by the calculation of the median expression in the case of multiple values for a gene.

The labels used for the sex of the patients were homogenized, and individuals were grouped into healthy or affected by MS. Paired samples were tagged to account for this factor in subsequent analyses. After normalization of the data, possible batch effects or anomalous behavior of the data were analyzed using clustering and principal component analysis (PCA).

Comparisons

Different comparisons were performed to detect sex-specific genes and find sex-differential functional patterns. The impact of the disease in females (IDF) was analyzed by comparing female MS patients to healthy female controls (Female.MS - Female.Ctrl), while the impact of the disease in males (IDM) was analyzed by comparing males MS patients to healthy male controls (Male.MS - Male.Ctrl). Significant features in both comparisons represent potential disease biomarkers, while those significant in just one comparison can be considered sex-specific biomarkers. Additionally, we analyzed the sex-differential impact of the disease (SDID) through an advanced comparison (Female.MS - Female.Ctrl) - (Male.MS - Male.Ctrl) to find differences in disease impact by sex (Figure 1).

As a measure of the statistical effect, the logarithm of the fold change (LFC) was used for DGE analysis, and the logarithm of the odds ratio (LOR) was used for functional enrichment analysis. Positive statistics indicate that the variable has a higher mean in the first element of the comparison, while negative statistics indicate higher mean values in the second element. Positive LFC or LOR values indicate higher values in MS patients (MS females and MS males in IDF and IDM comparisons) and, therefore, gene upregulation or function overrepresentation in MS. Negative LFC or LOR indicates higher values in control samples (healthy control females

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

and healthy control males), and therefore gene downregulation or function underrepresentation in MS. The SDID comparison, for which each element is a comparison itself, focuses on finding differences between the IDF and IDM comparisons. Thus, a positive statistic could indicate: i) an upregulation in IDF together with a downregulation in IDM, ii) an upregulation in IDF without significant changes in IDM, iii) an upregulation in IDF and IDM but higher in IDF, iv) a downregulation in IDM without a significant change in IDF, or v) a downregulation in IDF and IDM, but higher in IDF (Supplementary Figure 1A-E). A negative statistic associated with the opposite of these cases (Supplementary Figure 1F-J). In this comparison, the particular behavior of each feature across the 4 groups must be assessed *a posteriori*, examining IDF and IDM comparisons. For simplicity, features with positive LFC or LOR will be referred to as features with an incremental change in females or increased in females, and features with a negative LFC or LOR as features with an incremental change in males or increased in males.

Differential Gene Expression and Meta-analyses

3 DGE analyses were performed for each study to detect differences in gene-level expression between groups following comparisons between IDF, IDM, and SDID as described above. All comparisons were performed using the R package limma [17]. After calculating the differential expression statistics, p-values were adjusted using the Benjamini & Hochberg (BH) method [18].

To robustly integrate the DGE results of each study, the meta-analysis technique was employed for each comparison (IDF, IDM, and SDID) and available tissue of origin (blood, nervous tissue, and both together), which led to a total of 9 meta-analyses. The two-tissue meta-analysis included only genes detected in at least 1 study from each tissue. The meta-analyses were performed following the methodology previously described by García-García [19]. Briefly, the metafor R package [20] combined the expression results of the corresponding individual studies, using a random-effects model for each gene, which is particularly suitable when there exist diverse sources of variability [21]. For the overall calculation of LFC, the least variable results were given greater weight by considering the variability of the individual studies in the global estimation of the measured effect. Better

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

integration of the selected studies and statistically more robust results were achieved by incorporating the variability between experiments in the random-effects model. The suitability of this method was evaluated by assessing the heterogeneity and influence of each study on the overall model and cross-validation techniques.

The meta-analyses calculate a pooled measure of the expression effect across all studies for each gene (combined LFC and its 95% confidence interval) and provide a BH-adjusted p-value [18]. The significance cutoff was set to 0.05. The Open Targets platform [22] (release 21.09.5) was used to evaluate the association of genes with MS.

Functional Enrichment Analyses

To detect sex-specific and sex-differential functions or pathways associated with MS, a gene set analysis (GSA) [23] of nervous tissue meta-analysis was performed for each of the 3 comparisons (IDF, IDM, and SDID). First, genes were ordered according to the p-value and sign of the statistic obtained in the DGE, and GSA was performed using the logistic regression model implemented in the mdgsa R package [24]. The functional annotation required for GSA was obtained from the following databases: Gene Ontology (GO) [25] in the case of biological processes (BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway [26] for pathways. Due to their hierarchical structure, the gene annotations were propagated with GO terms applied in the mdgsa package to inherit the annotations of ancestor terms. Next, excessively specific or generic annotations were filtered out (blocks smaller than 10 or larger than 500 genes). Finally, functions with a Benjamini & Yekutieli (BY) adjusted p-value [27] under 0.05 were considered significant. To obtain a global view of sex-specific alterations in GO biological processes, the GSA results of each comparison were grouped into categories defined by hierarchically superior GO terms, counting the number of processes within each of them.

Metafun-MS Platform

All data and results generated in the different steps of the meta-analysis are available on the Metafun-MS platform (http://bioinfo.cipf.es/metafun-MS), which is freely accessible to any user and allows the confirmation of the results described in this manuscript and the

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

exploration of other results of interest. The front-end was developed using the Angular Framework. All graphics used in this web resource have been implemented with Plot.ly [28] except for the exploratory analysis cluster plot generated with ggplot2 [29].

This easy-to-use resource is divided into 5 sections: i) the summary of analysis results in each phase, followed by detailed results for the ii) exploratory analysis, iii) DGE, iv) meta-analysis, and v) functional profiling for each study. The user can interact with the web tool through graphics and tables and explore information associated with specific genes or biological functions. In addition, the detailed results of all analyses are available in the public repository Zenodo (http://doi.org/10.5281/zenodo.6344450).

Results

Systematic Review and Exploratory Analysis

Our systematic review screened 122 publications, which rendered a total of 9 studies eligible for our analyses: 5 studies with blood samples and 4 studies with nervous tissue samples (Figure 2, Table 1). Notably, we discarded 35% of the studies evaluated for not including sex information or only including individuals of a single sex.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

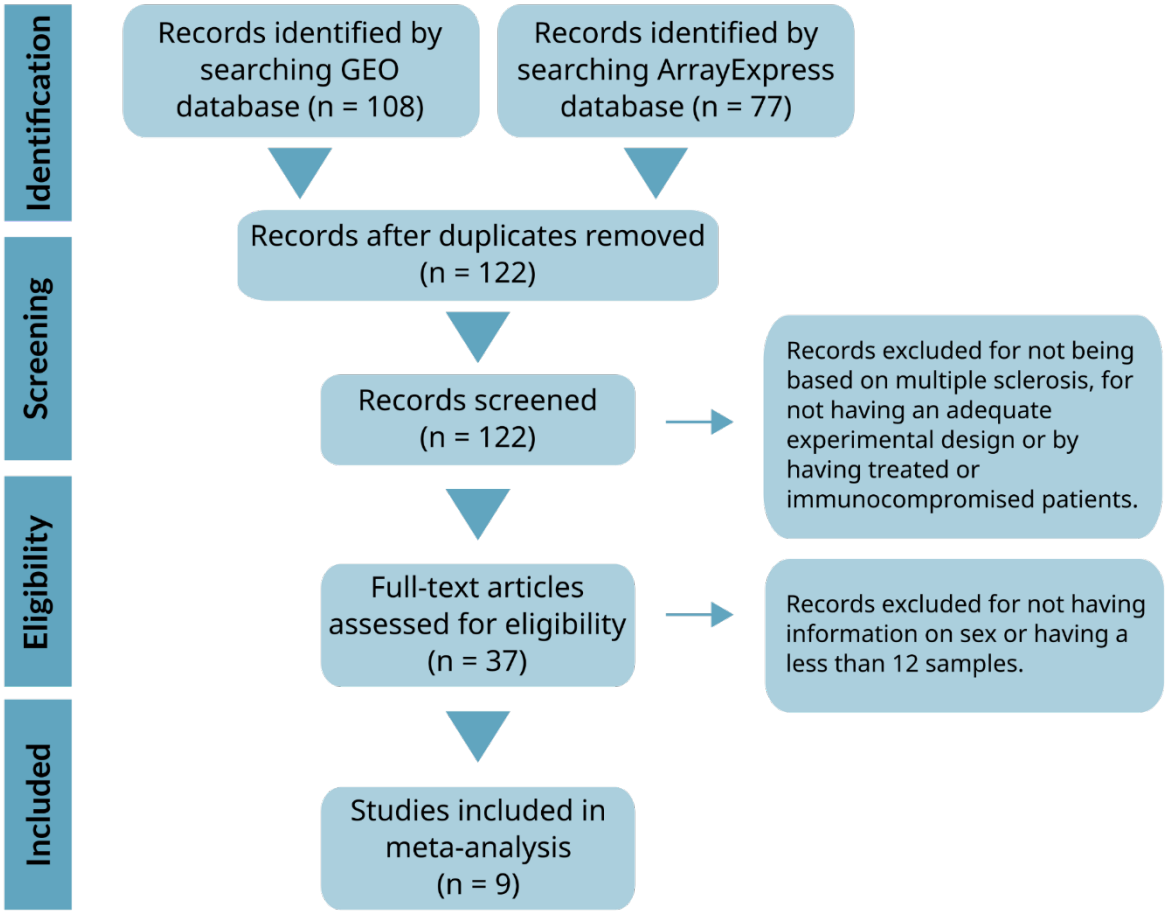

Figure 2. PRISMA diagram of the systematic review.

The 9 selected studies contained 474 samples - 271 MS patients and 203 healthy control patients (62.9% female samples and 37.1% male samples). Table 2 describes the distribution of the samples in detail.

Table 1. Description of selected studies. Platform, tissue, and MS subtype for the samples from 9 studies. The characteristics of the studies are shown as available in the GEO database. An asterisk (MS*) is used when the study does not provide MS subtype information (PPMS - progressive primary multiple sclerosis, SPMS - secondary-progressive multiple sclerosis).

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

GEO accession	Platform	Tissue	Subtypes of MS
GSE37750 [30]	Affymetrix Human Genome U133 Plus 2.0 Array	Plasmacytoid dendritic cells	MS*
GSE41848 [31]	Affymetrix Human Exon 1.0 ST Array	Whole blood	Relapse, remission
GSE41849 [31]	Affymetrix Human Exon 1.0 ST Array	Whole blood	MS*
GSE41890 [32]	Affymetrix Human Gene 1.0 ST Array	Peripheral blood leukocytes	MS*
GSE62584 [33]	Affymetrix Human Genome U133 Plus 2.0 Array	CD4+ blood cells, CD8+ blood cells, CD14+ blood cells, and CD19+ blood cells	ON patient in relapse
GSE108000 [34]	Agilent-026652 Whole Human Genome Microarray 4x44K v2	Brain	MS active, MS inactive
GSE111972 [35]	Illumina NextSeq 500	Corpus callosum and occipital cortex	MS*
GSE131281 [36]	Illumina HumanHT-12 V4.0 expression beadchip	Grey matter	SPMS, PPMS
GSE135511 [37]	Illumina HumanRef-8 v3.0 expression beadchip	Motor cortex	MS lesion, MS normal-appearing

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Table 2. Distribution of samples by sex and experimental group.

GEO accession	Tissue	Females		Males		Total
		MS	Control	MS	Control	
GSE37750	Blood	7	6	2	2	17
GSE41848	Blood	41	28	13	10	92
GSE41849	Blood	12	13	9	9	43
GSE41890	Blood	6	12	8	12	38
GSE62584	Blood	16	20	8	16	60
GSE108000	Nervous tissue	22	7	8	3	40
GSE111972	Nervous tissue	5	9	10	7	31
GSE131281	Nervous tissue	58	11	6	31	106
GSE135511	Nervous tissue	22	3	18	4	47
Total		189	109	82	94	474

Exploratory analyses revealed the absence of biases and anomalous behavior in the samples from all studies except for GSE62584, in which we detected a batch effect produced by the 4

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

cell types analyzed in that study. We considered this effect during subsequent analyses.

Differential Gene Expression Analysis

As described in the material and methods section, we evaluated 3 comparisons for each study: IDF, IDM, and SDID. Table 3 summarizes the number of significantly differentially-expressed genes, while the associated Zenodo repository and Metafun-MS web platform contain the detailed results for each comparison.

Table 3. The number of significantly differentially-expressed genes by study (and tissue). Description of significant genes by comparisons: IDF (impact of the disease in females), IDM (impact of the disease in males), and SDID (sex-differential impact of the disease). Significant genes are separated according to the sign of their log fold-change (LFC). The final column reports the number of genes assessed in each study.

Study	Tissue	COMPARISON						Analyzed genes
		SDID		IDF		IDM		
		LFC > 0	LFC < 0	LFC > 0	LFC < 0	LFC > 0	LFC < 0	
GSE37750	Blood	0	0	0	0	0	0	22,899
GSE41848	Blood	0	0	23	9	0	0	17,479
GSE41849	Blood	0	0	0	0	0	0	17,479
GSE41890	Blood	15	15	17	4	93	3	19,914
GSE62584	Blood	145	356	0	0	55	25	13,277
GSE108000	Nervous	154	230	2038	3114	1811	2209	21,586
GSE111972	Nervous	0	0	0	0	0	0	20,846

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

GSE131281	Nervous	4	0	15	8	0	3	20,923
GSE135511	Nervous	0	0	12	0	100	114	15,434

Meta-analysis of DGE Profiles

The SDID meta-analysis in the different tissues reported 15 genes that displayed significant alterations between sexes (KIR2DL3 in blood; ARL17B, CECR7, CEP78, IFFO2, LOC401127, NUDT18, RNF10, SLC17A5, STEMP1, TRAF3IP2-AS1, UBXN2B, ZNF117, ZNF488 in nervous tissue; LOC 102723701 in both tissues). We also found that specific genes displayed significance for IDF and/or IDM comparisons, which allowed us to detect the change between sexes and the alterations between healthy and diseased individuals within each sex. We will discuss the alterations found in each tissue in the following sections.

We have made the complete results of the meta-analyses available in the Zenodo repository, the supplementary materials, and on the Metafun-MS web platform. In summary, Figure 3 shows the pattern of LFC change between sexes (SDID, dashed line), in females (IDF, orange), and males (IDM, green). In addition, Table 4 indicates the number of significantly altered genes for each tissue and comparison.

Table 4. Number of DGEs in the meta-analyses by comparison and tissue. Description of significant genes by comparisons: IDF (Impact of the disease in females), IDM (Impact of the disease in males), and SDID (Sex-differential impact of the disease). Significant genes are separated according to the sign of their log fold-change (LFC).

Tissue	**COMPARISON**					
	SDID		**IDF**		**IDM**	
	LFC > 0	**LFC < 0**	**LFC > 0**	**LFC < 0**	**LFC > 0**	**LFC < 0**

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Blood	1	0	5	4	27	5
Nervous	7	6	33	57	29	43
Both tissues	1	0	7	4	8	7

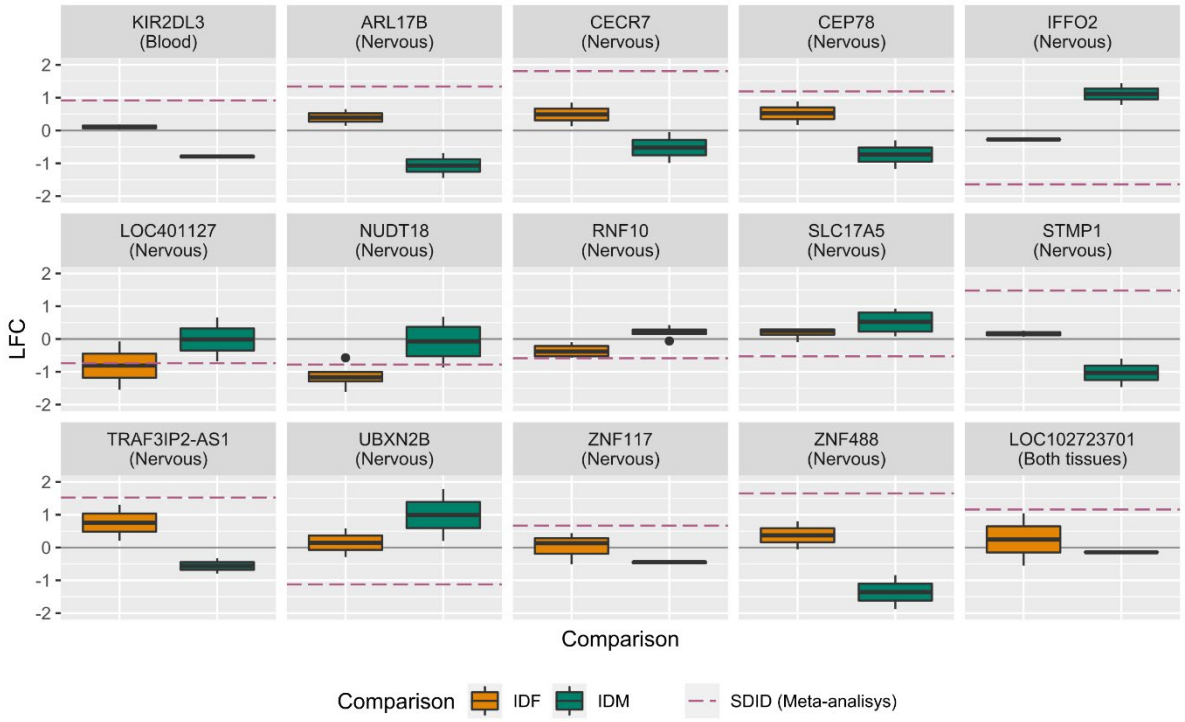

Figure 3. Boxplot of LFC values for the DGE in the SDID comparison. Each boxplot indicates the log fold-change (LFC) values of a gene in the individual DGEs. The dashed purple line indicates the LFC value of the meta-analysis for the SDID comparison. The gray line (at a value of 0) indicates the absence of change. The tissue in which the gene is significant is indicated in parentheses under the gene name.

Meta-analysis of DGE in Blood

In blood, the DGE meta-analysis only reported the KIR2DL3 gene, which encodes for a

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

transmembrane glycoprotein expressed by natural killer cells and subsets of T cells, as significantly altered in the SDID comparison. KIR2DL3 displays increased expression in MS females compared to MS males. In particular, the IDM comparison revealed lower KIR2DL3 expression in MS males than in healthy control males but did not show a significant change in the IDF comparison (Figure 3 and Figure 4, red box). This is a special case of Supplementary Figure 1D.

The IDF and IDM comparisons reported 40 significant sex-specific genes not implicated in the SDID comparison (Figure 4 and Table 5). Interestingly, some genes have known associations with MS in the Open Targets Platform (Table 5, in bold), while other genes, such as PSMB2, SRP14, TESC, and KAT7, have known associations with neurodegenerative diseases, the innate immune system, neuroprotective effects, and T cell development, respectively [38–41].

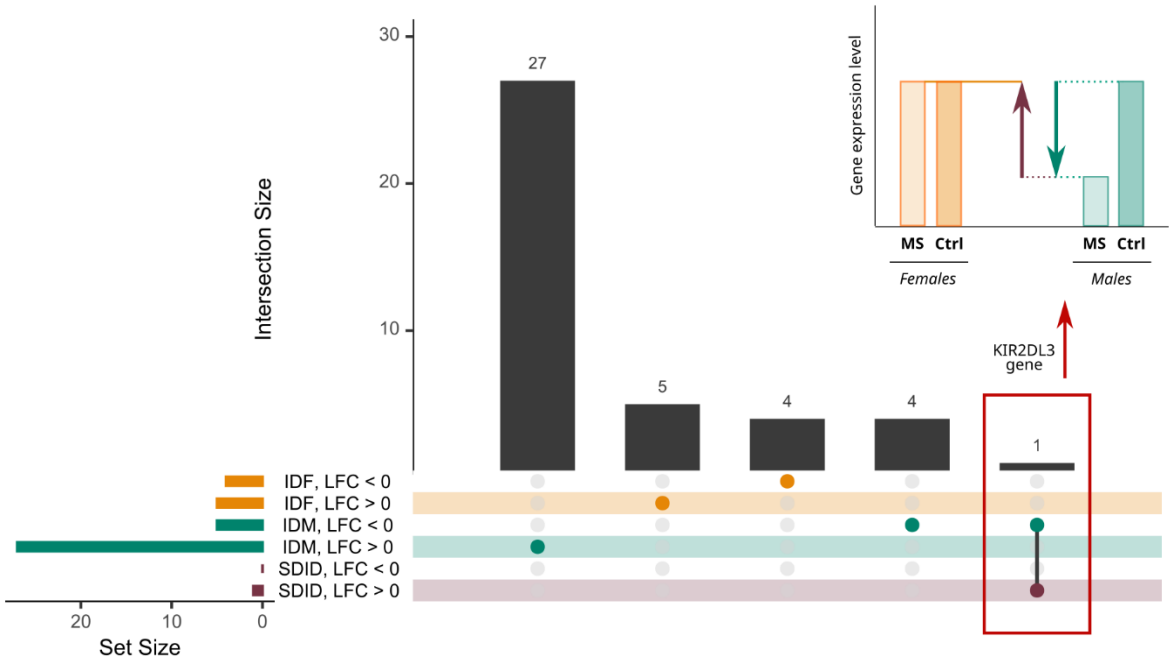

Figure 4. Upset plot of meta-analysis results in blood studies. The results of each comparison are separated according to the LFC sign. Horizontal bars indicate the number of significant genes in each comparison (1 specific color for each comparison). The vertical bars indicate the genes included in the intersection of the groups denoted with a colored dot underneath. A colored dot under a bar indicates that the genes are specific to this group. The red box indicates the case of the KIR2DL3 gene, which is significant SDID and IDM comparisons in the manner represented by the bar plot in the upper right corner.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Table 5. Meta-analysis results of blood samples. Significant genes for SDID (sex-differential impact of the disease), IDF (impact of the disease in females), and IDM (impact of the disease in males) comparisons. The results are separated into columns according to the LFC sign of the genes. Genes associated with MS in the Open Targets database are indicated in boldface. *KIR2DL3 was not associated with MS in Open Targets but was in a previous publication [42].

Comparison	**LFC > 0**	**LFC < 0**
SDID	KIR2DL3*	
IDF	DDX39A, LAPTM4B, PCK2, **RANGRF**, **SERPINF1**	FNBP1, KAT7, KMT2C, SETD5
IDM	**B2M**, COMMD9, COX6B1, COX8A, GLA, **GRIK4**, MIR6773, MRPL23, MRPS16, NAA38, **NDUFA2**, **NDUFS7**, OSER1, PSMB2, RP11-79P5.2, RPS19, **SOCS3**, SRP14, SSNA1, SSR4, TESC, TMEM222, TMX4, TXNL4A, UXT, WDR83OS, ZNF816-ZNF321P	KIR2DL3*, LINC00965, LOC101930363, LOC102725213, TCAF1

Meta-analysis of DGE in Nervous Tissue

In nervous tissue, our meta-analysis detected 13 significant genes in the SDID comparison (Table 6). Of these 13, 6 genes (RNF10, IFFO2, SLC17A5, UBXN2B, ZNF117, and ZNF488) have previous links to neurodegenerative diseases, while IFFO2 participates in muscle control [43–48] and 5 other genes (CECR7, STMP1, TRAF31P2-AS1, and LOC401127) [49–52] possess links to inflammation, bone, and autoimmune diseases. The remaining two genes - CEP78 and NUDT18 - encode for a centrosomal protein and a hydrolase of the Nudix superfamily, respectively. Interestingly, the CECR7 gene displayed significance in all 3 comparisons (Figure 5, red box and upper right corner), being overexpressed in MS females and healthy males.

We found 90 and 72 significant genes in the IDF and IDM comparisons, respectively (Figure 5). In the IDF and IDM comparisons, 9 genes displayed the same pattern of change, suggesting

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

that they may represent possible sex-independent disease markers (Figure 5). Table 6 describes all significant genes in nervous tissue, indicating those associated with MS according to the Open Targets Platform (in bold) and those common to the IDF and IDM comparisons (underlined). We display a functional enrichment analysis in the following sections, given the high number of genes involved in IDF and IDM comparisons.

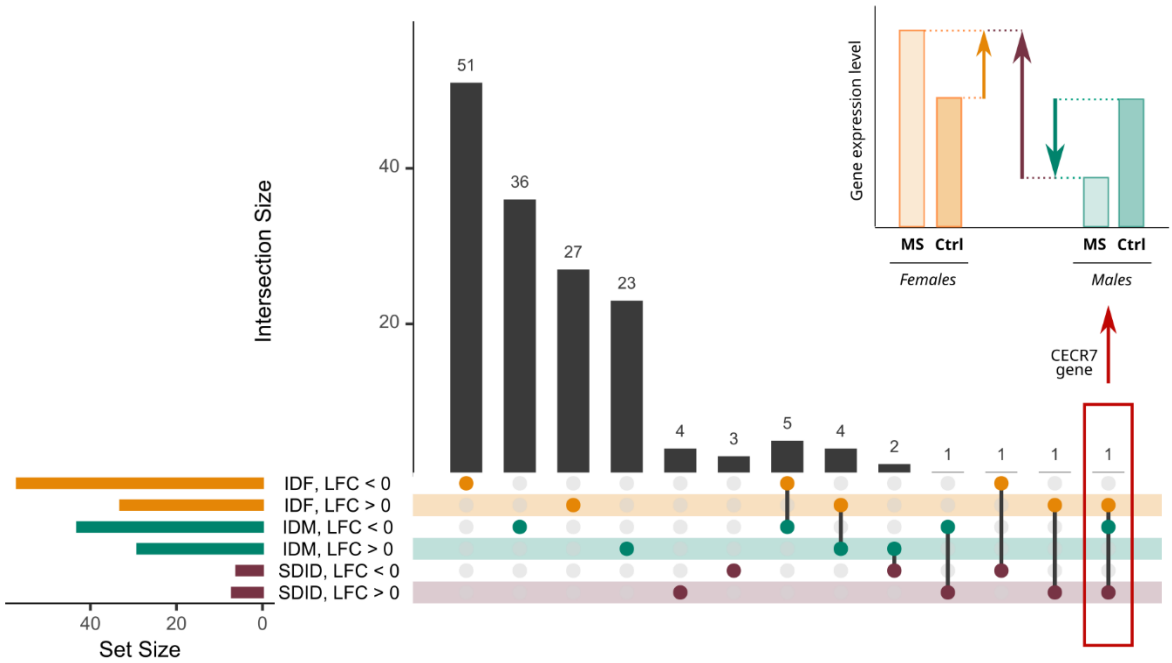

Figure 5. Upset plot of the meta-analysis results in nervous tissue samples. The results of each comparison are separated according to the LFC sign. Horizontal bars indicate the number of significant genes in each comparison (a specific color for each comparison). The vertical bars indicate the genes included in the intersection of the groups denoted with a colored dot underneath. A colored dot under a bar indicates that the genes are specific to this group. The red box indicates the case of the CECR7 gene, which is significant in all comparisons in the manner represented by the bar plot in the upper right corner.

Table 6. Meta-analysis results from nervous tissue samples. Significant genes for SDID (sex-differential impact of the disease), IDF (impact of the disease in females), and IDM (impact of the disease in males) comparisons. The results are separated into columns according to the LFC sign of the genes. Genes associated with MS in the Open Targets database are indicated in boldface. Genes altered in the same direction in the IDF and IDM comparisons are underlined.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Comparison	LFC > 0	LFC < 0
SDID	ARL17B, CECR7, CEP78, STMP1, TRAF3IP2-AS1, ZNF117, ZNF488	IFFO2, LOC401127, NUDT18, RNF10, **SLC17A5**, UBXN2B
IDF	BMP6, **CCNA2**, CECR7, CEP78, DNAJB5, ERI1, G2E3, GAS2L3, HCG4, **HLA-DRB1**, **HLA-DRB5**, IGLL5, LDAH, LOC729970, MGAM, OLFML2B, **RTP4**, SFT2D2, STBD1, STON1-GTF2A1L, STYXL1, SUPT3H, TCP10L, TDRD3, **TLR8**, TMEM140, TTC17, TXLNB, **UQCC2**, XRCC4, ZFAS1, ZNF615, ZSCAN29	ANKRD26P3, ATP1A2, **BMPER**, C19orf47, C2CD4C, CATSPERZ, CCDC57, CES2, CMYA5, DGCR9, EBF4, ERCC2, **FAM156B**, FBXL6, FGFBP3, GCAT, **HBB**, HOXA4, HPDL, KRTAP10-9, **KRTAP5-7**, LINGO3, LOC100130093, LOC100130370, LOC100288911, LOC285696, LOC339529, LOC728743, LRFN1, MAGED4B, **MLXIPL**, NLRP2, NOXA1, NPIPB15, NUDT18, NUTM2F, ODAD4, OVGP1, PET100, PKI55, PMM2, PRDM12, SAXO1, SBF1P1, SNHG20, SNHG28, SP9, **SREBF1**, STKLD1, TADA2B, TESPA1, TMC4, TPBG, TRABD2A, VPREB3, XXYLT1, ZNF839
IDM	ATP13A3, B4GALT1, BAG2, **CASP7**, CR1L, CTAGE4, EFNA1, **HBEGF**, HS3ST3B1, HYAL1, IFFO2, LIMS4, LOC100129617, LOC100506136, LOC401127, LOC729970, MIR100HG, **MPZL2**, PRO0628, RBM14-RBM4, RIC1, SNORA59B, ST7, STBD1, STON1, TMEM140, TMEM163, **TRPM8**, ZFAS1	ACTL6B, ARIH2OS, C2orf72, C8orf44-SGK3, CA5BP1, CECR7, COQ4, DEFA3, DGCR10, DGCR6L, DGCR9, **DHRS9**, DOLK, EPHA7, FAM66D, FLJ32063, GPR89A, GTF2IRD2, JMJD7, LINC00685, LOC100652758, LOC338963, LOC644189, LOC730183, LOH12CR2, LPP-AS2, MDP1, NIP7, OR7D2, PHF24, PIPOX, PKI55, RWDD3, SBF1P1, SERF1A, SNHG20, SYNE4, TADA2B, TRIM66, TTC30A, URB1, ZNF117, ZNF771

Blood and Nervous Tissue Meta-analysis of DGE

The two-tissue meta-analysis revealed an altered gene in the SDID comparison (LOC102723701), whose expression increases in females (Table 7). This lncRNA class is

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

described as an antisense transcript of ERLIN2, a gene associated with juvenile primary lateral sclerosis and immune cells infiltration [53–55]. We found 11 and 15 genes significantly altered in the IDF and IDM comparisons, respectively; however, none displayed commonality (Figure 6, Table 7). We highlight KAT7 (also significant in the blood meta-analysis), SERF2 (a positive regulator of myeloid protein aggregation) [56], and NTN3 (a netrin, which controls guidance of CNS commissural axons and peripheral motor axons) [57] as well as others previously associated with disease (highlighted in bold in Table 7).

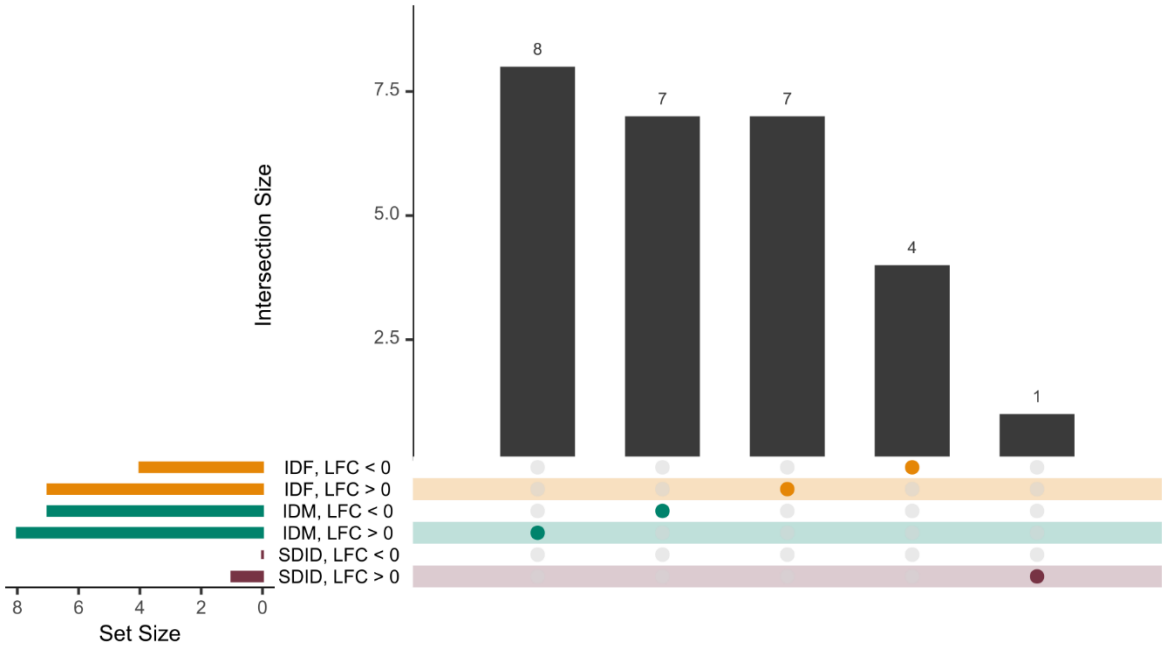

Figure 6. Upset plot of the two-tissue meta-analysis results. The results of each comparison are separated according to the LFC sign. Horizontal bars indicate the number of significant elements in each comparison (one specific color for each comparison). The vertical bars indicate the genes included in the intersection of the groups denoted with a colored dot underneath. A colored dot under a bar indicates that the genes are specific to this group.

Table 7. Results from the two-tissue meta-analysis. Significant genes for SDID (sex-differential impact of the disease), IDF (impact of the disease in females), and IDM (impact of the disease in males) comparisons. The results are separated into columns according to the LFC sign of the genes. Genes associated with MS in the Open Targets database are indicated in boldface.

Comparison	**LFC > 0**	**LFC < 0**

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

SDID	LOC102723701	
IDF	BANF1, **CCL19**, **GABRE**, LOC101927972, NTN3, RPL35, SERF2	**EP400**, FNBP1, KAT7, ZZEF1
IDM	B4GALT1, **DUSP1**, DUSP18, FAU, **RPL19**, TXNL1, VDAC2, ZNF844	A2M-AS1, FLJ30403, LOC283861, LOC286272, LOC400968, LOC728903, OR7D2

Functional Enrichment of DGE Meta-analysis in Nervous Tissue

We performed a GSA on the genes altered in all comparisons (IDF, IDM, and SDID) of the nervous tissue meta-analyses to provide a functional and integrative perspective of these genes in BP GO terms and KEGG pathways. In the SDID comparison, we detected the alteration of the "Staphylococcus aureus infection" KEGG pathway (hsa05150), which includes genes such as MHC class II antigen and Interleukin (IL-)10. In the IDF and IDM comparisons, we found significant results for both BP GO terms and KEGG pathways (Figure 7). We detected sex-specific elements and elements common to both sexes in both ontologies. Detailed results of the functional enrichment can be found in the supplementary material, the Zenodo repository, and the metafun-MS web platform.

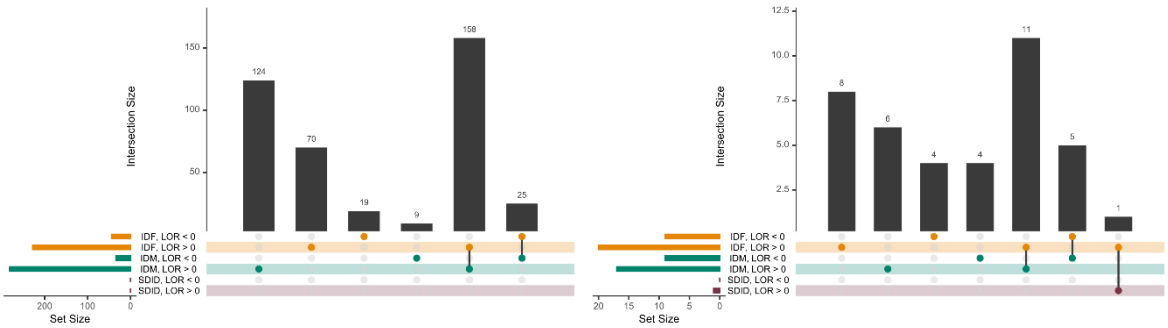

Figure 7. Upset plot of the functional enrichment results in nervous tissue for BP GO terms. Horizontal bars indicate the number of significant BP GO terms in each comparison (a specific color for each comparison). Results are exhibited separately for significantly upregulated (LOR > 0) and downregulated (LOR < 0) BP GO terms. The vertical bars indicate the genes included in the intersection of the groups denoted with a colored dot

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

underneath. A colored dot under a bar indicates that the genes are specific to this group.

To summarize the high number of significant BP GO terms in the IDF and IDM comparisons, we classified them into 15 manually selected parental GO classes (hierarchically superior GO terms) (Figure 8). We classified BP GO terms with no ancestor among the selected parental GO classes as "Other." Interestingly, the number of BP GO terms overrepresented in patients with MS (LOR > 0, Figure 8) is notably superior to those overrepresented in healthy patients (Figure 8, LOR < 0). The parental GO classes with the highest number of alterations were "Immune system process", "Response to stimulus" and, to a lesser extent, "Cellular response to stimulus" and "Anatomical structure development". Regarding the immune system, MS females specifically presented altered BP GO terms related to the myeloid lineage, regulation, and production of cytokines (IL-1, 2, 6, and 10), while MS males presented alterations in the lymphocyte lineage (Figure 9). MS females also presented altered mitochondrial respiratory chain complexes and purine and glutamate metabolism, while MS males presented alterations in amine and amino acid transport and stress response to metal ions (Figure 9).

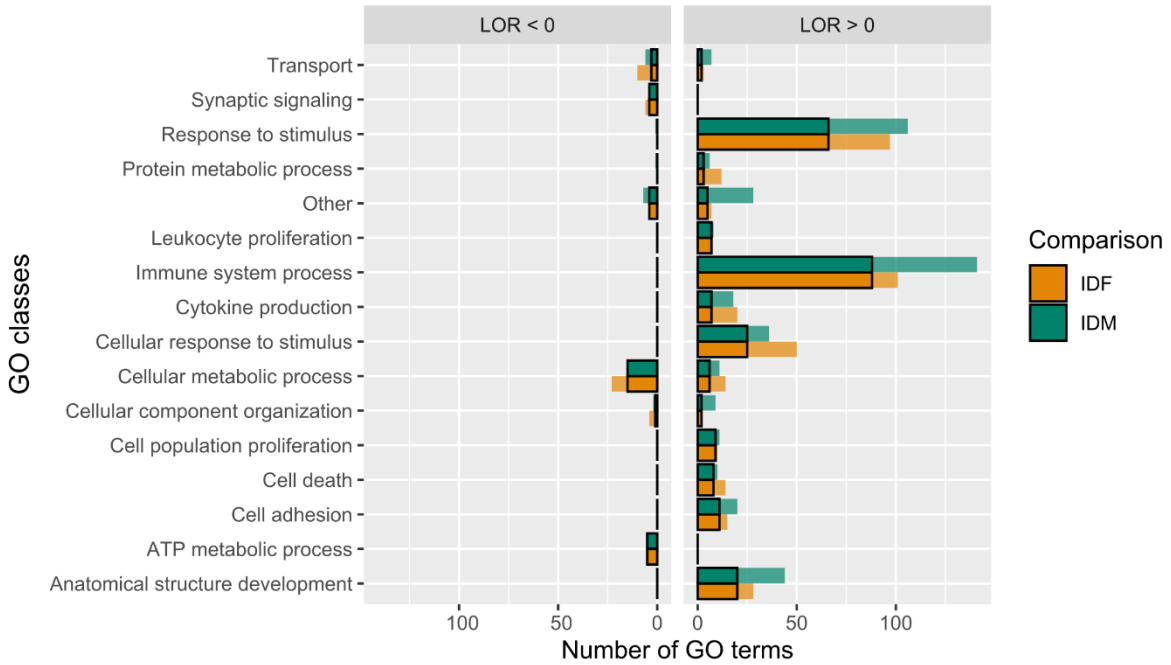

Figure 8. Bar chart of the number of significant BP GO terms in the IDF and IDM comparisons. The horizontal

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

bars show the number of BP GO terms, separated between overexpressed in MS patients (LOR > 0, right) and healthy patients (LOR < 0, right) in each parental GO class. The orange and green bars (IDF and IDM, respectively) show the number of GO terms within each parental GO class. Opaque bars surrounded by a black border indicate items common to the comparisons, while translucent bars indicate items unique to IDF or IDM.

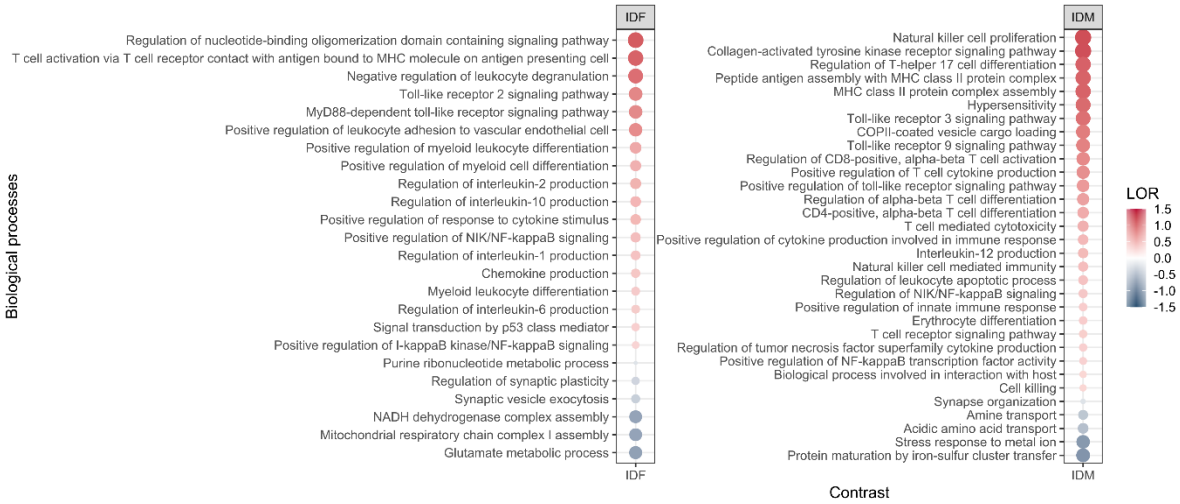

Figure 9. Dotplot of a representative selection of the specific significant BP GO terms in IDF and IDM. GO terms affected exclusively in IDF (left) or IDM (right). The color and size of the dots represent the sign and magnitude of the change (LOR).

In the case of IDF, the KEGG analysis demonstrated altered signaling pathways via Toll-like receptor and NOD-like receptor (constitute two major forms of innate immune sensors), as well as osteoclast differentiation, p53 pathway, and metabolism pathways related to biotin, alanine, aspartate and glutamate, xenobiotics, and drugs. KEGG pathway analysis of the IDF comparison also demonstrated alterations in terms related to diseases other than MS, including Chagas disease, prostate cancer, and Staphylococcus aureus infection. The IDM comparison showed alterations in KEGG pathways related to cytokine-cytokine receptor interactions, the Jak-STAT signaling pathway, protein processing in the endoplasmic reticulum, the peroxisome, olfactory transduction, calcium signaling, metabolism related to valine, leucine and isoleucine degradation, butanoate, and diseases such as small cell lung cancer and auto-immune thyroid disease (Supplementary Table 9).

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Metafun-MS platform

We have developed an open-source web platform, Metafun-MS (http://bioinfo.cipf.es/metafun-MS), containing a detailed description of the 9 datasets and 474 samples involved in this study. This resource includes detailed information on all genes, BP GO terms, and KEGG pathways that users can explore to identify transcriptomic profiles of interest for each study.

For each significant gene in a meta-analysis, Metafun-MS depicts the global activation level for all studies and each study's specific contribution using statistical indicators (LFC/LOR, confidence interval, p-value, and adjusted p-value) and graphical representations by function (forest and funnel plots). This open resource aims to contribute to data sharing between researchers, the elaboration of innovative studies, and the discovery of new findings.

Discussion

A range of previous studies have integrated data from numerous sources to characterize sex-based differences [58–60]; however, to our knowledge, this study represents the first meta-analysis of transcriptomic studies in MS and the first identification of sex-specific biomarkers in several tissues, which provides a better understanding of the molecular mechanisms underlying MS. This pathology has an important impact on society as one of the significant causes of disability in young adults; however, data regarding the causes of MS remain limited. Furthermore, the significant variability in clinical manifestations makes diagnosis and prognosis challenging tasks [2]. MS remains a clinical challenge, as current treatments focus on symptom amelioration and the slowing of the clinical course of the disease rather than patient recovery.

We carried out an *in-silico* approach using MS data related to sex from public repositories as a rapid and cost-effective means to predict and validate the effects of relevant biological variables on human health and guide further clinical research. A lack of standardization in the

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

description, nomenclature, and format of public data represented significant challenges to our approach. Therefore, in addition to encouraging free access to research data (Open Science), we also wish to promote the FAIR principles (Findable, Accessible, Interoperable, Reusable) [61], which aim to standardize the deposition of data in repositories.

In this study, we performed a meta-analysis using a robust random-effects model, which allows for the integration of information from individual studies to obtain greater statistical power and precision and reveals findings that cannot be easily obtained through the intersection or addition of results in individual studies. Table 3 depicts the diverse nature of each study - study GSE108000 in nervous tissue provided many differentially expressed genes, while others failed to provide any. This variability highlights the need to employ strategies such as meta-analysis to integrate information from individual studies and obtain a transcriptional consensus profile as a final result. We applied the same bioinformatics strategy from normalized expression matrices to DGE results with updated gene names to ensure comparable results and reduce biases. The design and application of this strategy identified sex-specific MS biomarkers. The further development of studies with sex information within each MS subgroup may extend the identification of biomarkers into specific subtypes and affected tissues.

Meta-analysis of transcriptomics profiles. Since MS is a chronic autoinflammatory disorder of the CNS, we performed separate meta-analyses in blood and nervous tissue and a third meta-analysis that integrated both tissues. Identifying sex-based differences represented our primary goal by studying genes differentially altered in the SDID comparison. Meanwhile, the IDF and IDM comparison allowed us to detect specific effects in each sex, understand the basis of the disease for each sex, and find common patterns in both sexes.

Meta-analysis in blood. Our meta-analysis in blood identified a single gene, KIR2DL3, in the SDID comparison. Interestingly, we also observed alterations to KIR2DL3 in the IDM comparison, with this gene downregulated in male MS patients. KIR2DL3 is related to MHC I mediated antigen processing and presentation and innate immune system pathways and has a protective effect; however, the absence of KIR2DL3 has been linked to MS development [42, 62]. Moreover, the protective role of KIR2DL3 has been documented in other autoimmune

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

diseases such as rheumatoid vasculitis [63]. Identifying biomarkers in blood that facilitate the non-invasive diagnosis, prognosis, and evaluation of treatment response remains crucial for MS patients who occasionally undergo cerebrospinal fluid analysis. Unfortunately, identifying blood-based biomarkers remains challenging due to factors such as molecule availability in the blood, the dilution factor, sensitivity, specificity, and disease stage. Together, these factors may explain why we only identified a single gene, KIR2DL3, in the SDID meta-analysis in blood.

The blood meta-analysis in IDF comparison revealed a gene dysregulation profile related to the immune system and symptoms associated with MS. For example, DDX39A, LAPTM4B, PCK2, RANGRF, SERPINF1, FNBP1, KAT7, KMT2C, and SETD5 displayed dysregulation in MS females. DDX39A and LAPTM4B may play roles in controlling immune responses [64, 65], while RANGRF has been related to cardiac arrhythmia in Brugada Syndrome [66]. The number of cardiac ventricular arrhythmias worsens the prognosis of MS patients with a higher risk of sudden death [67]. SERPINF1 may influence immune and vascular injury [68]. Studies have generally described roles in inflammation, immunity, and dementia for serpins, while serpin family members have been identified in MS plaques and neurodegenerative conditions such as Alzheimer's disease [69]. FNBP1 may represent a significant biomarker, as the expression of this gene links to prognosis and level of immune infiltration in different cancers [70] and could play a similar role in MS. Interestingly, KAT7, which gene encodes a histone acetyltransferase crucial for the development, maintenance, and survival of T cells and the prevention of auto-immune diseases [71], appeared downregulated in MS females. This study represents the first description of a sex-based association of KAT7 with MS to the best of our understanding.

The IDM blood meta-analysis reflected the dysregulation of genes related to inflammation, immune system, and autoimmune diseases (B2M, COMMD9, GLA, PSMB2, SOCS3, SRP14, SSNA1, SSR4, TESC, UXT, and RPS19 upregulated in MS males and KIR2DL3 downregulated in MS males), genes related to redox activity/stress oxidative and mitochondrial function (COX6B1, COX8A, MRPL23, MRPS16, NDUFA2, NDUFS7, OSER1, TMX4, and TXNL4A upregulated in MS males) and genes related to neurodegenerative diseases and intellectual disabilities (WDR83OS, GRIK4, NAA38, TMEM222 upregulated in MS males). Other short and

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

long non-coding (lnc)RNA genes (MIR6773, LINC00965, LOC101930363, and LOC102725213) could be critical for post-transcriptional regulation of other genes, while TCAF1, a transient receptor potential channel-associated factor, can regulate cell migration [72]. These results suggest that MS hallmarks such as inflammation and immunity occur in both MS females and males through different target genes in each case, which may be critical to guide medical decisions. Unfortunately, the blood meta-analyses failed to identify common MS biomarkers for males and females, perhaps due to difficulties finding stable markers in the blood, as discussed above.

Meta-analysis in nervous tissue. Analyses in nervous tissue yielded a higher number of differentially expressed genes than in blood, reflecting the notable impact of MS in this tissue. The SDID nervous tissue meta-analysis found 13 differentially affected genes by sex, with specific genes related to the loss of mental functions and neural disorders, blood cell abnormalities, inflammation, and autoimmune disorders.

ARL17B, CECR7, CEP78, STMP1, TRAF31P2-AS1, ZNF117, and ZNF488 displayed increased expression in MS females, as did the CECR7 and CEP78 genes (LFC > 0 in IDF). CECR7 has been previously related to immune infiltration in hepatocellular carcinoma [73] and regulates the expression of CTLA4 by targeting miR-429 [74]; meanwhile, CEP78 has been recognized as an auto-antigen in prostate cancer [75]. Thus, both genes represent potential candidates for immunotherapies. STMP1 has a reported involvement in the NLRP3 inflammasome and Paget's disease bone [50], while TRAF31P2-AS1 participates in IL-17 regulation. Interestingly, this interleukin possesses pro-inflammatory properties and contributes to immune dysregulation in other autoimmune diseases such as colitis and rheumatoid arthritis [76].

IFFO2, LOC401127, NUDT18, RNF10, SLC17A5, or UBXN2B, presented a negative LFC in this SDID comparison. In particular, IFFO2 and LOC401127 displayed significant upregulation in MS males (LFC > 0 in IDM) while NUTD18 displayed a significant downregulation in MS females (LFC < 0 in IDF). The presence of intermediate filament-associated proteins (IFAPs) in MS plaques tissue has been previously documented [77]; therefore, IFFO2 may represent a novel IFAP for MS male plaque tissues. The LOC401127 lncRNA could play an essential role in MS regulation or prognosis, as for other described lncRNAs in chronic musculoskeletal disorders

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[78]. We also highlight the protective role against oxidative stress and DNA damage of NUTD18 (also known as MTH3), which displays downregulation in MS females [79].

In the IDF and IDM meta-analyses in nervous tissue, we identified genes that appeared dysregulated in both sexes; therefore, these genes may represent potential MS biomarkers - LOC729970, STBD1, TMEM140, and ZAFS1 displayed upregulation and DGCR9, PKI55, SBF1P1, SNHG20, and TADA2B displayed downregulation in MS in both sexes. Specific genes identified in the IDF and IDM comparisons have not been previously linked to MS and may represent interesting future research directions. In general, we observed an increase in genes related to inflammatory and immune responses in female MS patients. Autoimmune diseases are generally exacerbated in female patients, and a more pro-inflammatory phenotype of XX mice than XY mice has been demonstrated in autoimmune diseases such as experimental and spontaneous lupus models [80, 81]. Sex-based differences in immune responses are well-documented and influenced by multiple factors (chromosomes organization, reproductive organs, and sex hormone levels), which, in turn, influence the incidence and other aspects of autoimmune diseases [82].

The two-tissue meta-analysis identified a single gene - LOC102723701 - as displaying SDID. This lncRNA gene, not previously related to MS, is described as a novel transcript antisense to the ERLIN2 gene. The misfunction of ERLIN2, which mediates intracellular calcium signaling and participates in neurodevelopment and neurotransmission, leads to motor disorders and spastic paraplegia [83, 84]. Importantly, ERLIN2 mutations lead to abnormal splicing and nonsense-mediated mRNA decay, promoting motor neurodegeneration and juvenile amyotrophic lateral sclerosis [55]. Additionally, ERLIN2 has been related to the infiltration of immune cells in the tumor microenvironment [54]. Therefore, LOC102723701 may represent an exciting and novel target in MS for future research. Regarding the IDF and IDM comparisons in both tissues, some of the identified genes (e.g., CCL19, GABRE, and EP400 in IDF; DUSP1 and RPL19 in IDM) have been previously associated with MS in the Open Targets platform, thereby supporting our results. In general, the two-tissue meta-analysis mirrored the results from the meta-analysis in individual tissues. Genes related to oxidative stress, mitochondrial functions, metabolism, immune system, and other neural or autoimmune disorders such as Alzheimer's disease or lupus erythematosus displayed dysregulation in MS

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

patients.

Focusing on the **functional enrichment in nervous tissue,** we found a single significantly dysregulated KEGG pathway in the SDID comparison: "Staphylococcus aureus infection" that involved important immune genes important immune genes (MHC class II antigen and IL-10). We also discovered the upregulation of this pathway in MS females, further supporting the differential role of the immune system in MS between males and females. When we analyzed GO terms in the IDF and IDM comparisons, many biological processes related to inflammation and immunity displayed sex-dependent dysregulation in MS patients. Specifically, MS females presented alterations related to myeloid linage and pro-inflammatory environment, influenced by alterations in interleukin regulation and response (e.g.., IL-1, IL-2, IL-6, IL-10, and interferon-gamma). These alterations are closely related to innate immune responses and have been previously linked to MS and autoimmune diseases [85–89]. In contrast, MS males presented immune alterations related to immunomodulation of lymphoid linage (e.g., T cell subsets, natural killer cells, dendritic cells), affecting activation, proliferation, or differentiation, which have a closer relation to adaptative immune responses [90–92]. Identifying T cell clones associated with human autoimmunity remains challenging and could have critical connotations in MS risk and outcomes. Taken together, the different scenarios observed in the immune system could partly explain sex-based differences in susceptibility to MS. Pro-inflammatory cytokines, which are significantly altered in MS females, have been related to autoimmune disorders such as arthritis rheumatoid [93]. Meanwhile, MS males suffer from a more rapid and worsened progression, which could be explained by more intensive participation of T cells in demyelination. Further studies must be carried out to clarify these hypotheses and help guide decisions in the design of personalized immunotherapies.

Other sex-related differences in GO terms included alterations in mitochondrial respiratory chain complexes, purine, and glutamate metabolism in MS females and alterations in amine and amino acid transport and stress response to metal ions in MS males. These biologic processes have important connotations in MS - the glutamatergic system plays a vital role in neurotoxicity and MS pathogenesis [94], while amino acid metabolism affects immune homeostasis, as activated immune cells present a high demand for amino acids [96].

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Analysis of the KEGG pathway data revealed that MS females displayed an upregulation in several pathways related to immune signaling receptors (Toll-like receptor signaling pathway, NOD-like receptor signaling pathway, B cell receptor signaling pathway), the upregulated pro-inflammatory microenvironment observed in GO terms, and the innate responses explained above. Toll-like and Nod-like receptors are two major forms of innate immune sensors and provide immediate responses against pathogenic invasion or tissue injury [99, 100]. Activation of these sensors induces the recruitment of innate immune cells such as macrophages and neutrophils (myeloid linage) (which displayed dysregulation in MS females), initiates tissue repair processes, and results in adaptive immune activation. Abnormalities in these innate sensor-mediated processes may cause excessive inflammation due to hyper-responsive innate immune signaling or sustained compensatory adaptive immune activation [101]. Other KEGG pathways enriched in MS females have associations with chronic or autoimmune diseases (e.g., chronic myeloid leukemia, systemic lupus erythematosus, and rheumatoid arthritis), which may indicate that autoimmune diseases share altered signaling pathways. We also discovered the upregulation of intestinal immune networks for IgA production and Chagas disease in MS females. Diet, microbiota gut, parasite, and viral infections are associated with chronic inflammatory diseases and could play a significant role in MS etiology [102]; meanwhile, intestinal disorders also represent common symptoms of MS [103, 104]. The KEGG pathways downregulated in MS females include the metabolism of xenobiotics by cytochrome P450, drug metabolism - cytochrome, cardiac muscle, and Alzheimer's disease, which could have significant implications in treatment responses and therapy development. In MS males, upregulated KEGG pathways related to RNA transport, cytokine-cytokine receptor interaction, protein processing in the endoplasmic reticulum, the phagosome, JAK-STAT signaling pathway, and olfactory transduction, while valine, leucine, and isoleucine degradation represent the unique downregulated pathway. RNA localization controls gene expression, while neurological diseases such as amyotrophic lateral sclerosis and fragile X syndrome have been associated with dysregulated RNA localization in neurons [105]. Endoplasmic reticulum-associated protein modifications, folding, and degradation could mediate the pathogenesis of autoimmune disorders [106], while antigen processing and cross-presentation by dendritic cells require phagosome maturation [107]. Finally, metabolism and diet have relevant implications in MS, as in other autoimmune diseases [108], so amino acid degradation may impact MS at this level.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Strengths and limitations. Patient sex influences the observed pathophysiological differences in neurodegenerative and autoimmune diseases [3, 109], with sex-based differences described in MS risk and disease progression [110]. Unfortunately, sex is not frequently considered a biological variable during biomedical research; this became evident during our systematic review, where the absence of sex information prompted us to discard a substantial number of studies. The small number of studies with sex information has not allowed us to stratify MS subtypes; however, our global approach could identify a set of biomarkers common to MS. Encouragingly, the sex perspective has gained recent relevance as a crucial aspect of personalized medicine and has been recognized as a necessary biological variable for inclusion in research by the National Institutes of Health (US Department of Health and Human Services) [111] and the European Union [112, 113].

Conclusions

We have performed a gene and functional meta-analysis of 9 MS transcriptomic studies from a sex perspective, identifying sex-based differences and similarities at gene, function, and pathway levels. The obtained results allow for a better understanding of those mechanisms crucially involved in MS and the identification of potential biomarkers, whose detailed study may improve personalized therapies. In addition, we highlight the importance of providing information on sex in studies and the need for open platforms for data sharing to allow for efficient advancements in research.

Additional files

Detailed results of all analyses and the versions of the R packages used can be found in the Zenodo repository: http://doi.org/10.5281/zenodo.6344450

Abbreviations

BH: Benjamini & Hochberg p-value adjust method; BP: Biological processes of Gene Ontology; CIS: Clinical Isolated Syndrome; CNS: central nervous system; DGE: Differential gene expression; GEO: Genes Omnibus Expression database; GO: Gene Ontology; GSA: Gene set

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

analysis; IDF: Impact of the disease in females; IDM: Impact of the disease in males; IL: Interleukin; KEGG: Kyoto Encyclopedia of Genes and Genomes; LFC: Logarithm of the fold change; LOR: Logarithm of odds ratio; MS: Multiple sclerosis; PCA: Principal component analysis; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PPMS: Progressive Primary Multiple Sclerosis; PRMS: Progressive-Relapse Multiple Sclerosis; RRMS: Relapsing-Remitting Multiple Sclerosis; SDID: Sex-differential impact of the disease; SPMS: Secondary-Progressive Multiple Sclerosis.

Funding

Research supported by the Principe Felipe Research Center and a GV/2020/186 grant.

Availability of data and materials

The large volumes of data and results generated in this study are freely available through the Metafun-MS web tool: http://bioinfo.cipf.es/metafun-MS and the Zenodo repository: http://doi.org/10.5281/zenodo.6344450. This study analyzed transcriptomic data available in the Gene Expression Omnibus database with accession numbers GSE37750, GSE41848, GSE41849, GSE41890, GSE62584, GSE108000, GSE111972, GSE131281, and GSE135511.

Computer Code and Software

The code developed for the analyses described in this work and the software and its versions are publicly available at Zenodo: http://doi.org/10.5281/zenodo.6344450.

Authors' contributions

Conceptualization, F.G.-G.; Data curation, J.F.C.-S., F.J.R., A.N.-A., and I.S.-S.; Formal analysis, J.F.C.-S., and F.J.R.; Funding acquisition, F.G.-G; Investigation, J.F.C.-S., Z.A., M.R.H., N.Y.-C., and F.G.-G.; Methodology, J.F.C.-S., and F.G.-G.; Project administration, F.G.-G.; Software, J.F.C.-S., and M.R.H.; Supervision, F.G.-G.; Validation, Z.A., and N.Y.-C.; Visualization, J.F.C.-S., I.S.-S., M.R.H., and F.G.-G.; Writing – original draft, J.F.C.-S., Z.A., and F.G.-G.; Writing – review & editing, J.F.C.-S., Z.A., F.J.R., M.R.H., A.N.-A., I.S.-S., A.L.-C., N.Y.-C., M.I.-V., and F.G.-G. All

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgments

The authors thank the Principe Felipe Research Center (CIPF) for providing access to the cluster, co-funded by European Regional Development Funds (FEDER) in Valencian Community 2014-2020. The authors also thank Stuart P. Atkinson for reviewing the manuscript.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Ford H. Clinical presentation and diagnosis of multiple sclerosis. Clin Med. 2020;20:380–3.

2. Dobson R, Giovannoni G. Multiple sclerosis - a review. Eur J Neurol. 2019;26:27–40.

3. Ortona E, Pierdominici M, Maselli A, Veroni C, Aloisi F, Shoenfeld Y. Sex-based differences in autoimmune diseases. Ann Ist Super Sanita. 2016;52:205–12.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

4. Correale J, Ysrraelit MC, Fiol MP. Benign Multiple Sclerosis: Does it exist? Curr Neurol Neurosci Rep. 2012;12:601–9.

5. Bishop M, Rumrill PD. Multiple sclerosis: Etiology, symptoms, incidence and prevalence, and implications for community living and employment. Work. 2015;52:725–34.

6. Voskuhl RR. The effect of sex on multiple sclerosis risk and disease progression. Mult Scler. 2020;26:554–60.

7. Muñoz-Culla M, Irizar H, Sáenz-Cuesta M, Castillo-Triviño T, Osorio-Querejeta I, Sepúlveda L, et al. SncRNA (microRNA & snoRNA) opposite expression pattern found in multiple sclerosis relapse and remission is sex dependent. Sci Rep. 2016;6:20126.

8. Lasrado N, Jia T, Massilamany C, Franco R, Illes Z, Reddy J. Mechanisms of sex hormones in autoimmunity: focus on EAE. Biol Sex Differ. 2020;11:50.

9. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. The Lancet. 2018;391:1622–36.

10. Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. The Lancet. 2017;389:1336–46.

11. Pierdominici M, Maselli A, Colasanti T, Giammarioli AM, Delunardo F, Vacirca D, et al. Estrogen receptor profiles in human peripheral blood lymphocytes. Immunology Letters. 2010;132:79–85.

12. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Research. 2012;41:D991–5.

13. Athar A, Füllgrabe A, George N, Iqbal H, Huerta L, Ali A, et al. ArrayExpress update – from bulk to single-cell expression data. Nucleic Acids Research. 2019;47:D711–5.

14. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6:e1000097.

15. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.

16. European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI). HGNC Database, HUGO Gene Nomenclature Committee (HGNC). www.genenames.org. Accessed 28 Feb 2022.

17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43:e47.

18. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57:289–300.

19. García-García F. Métodos de análisis de enriquecimiento funcional en estudios genómicos. University of Valencia; 2016.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

20. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software. 2010;36:1–48.

21. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7:177–88.

22. Ochoa D, Hercules A, Carmona M, Suveges D, Gonzalez-Uriarte A, Malangone C, et al. Open Targets Platform: supporting systematic drug–target identification and prioritisation. Nucleic Acids Research. 2021;49:D1302–10.

23. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.

24. Montaner D, Dopazo J. Multidimensional Gene Set Analysis of Genomic Data. PLoS One. 2010;5:e10348.

25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25:25–9.

26. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research. 2000;28:27–30.

27. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Statist. 2001;29.

28. Sievert C. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC; 2020.

29. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016.

30. Aung LL, Brooks A, Greenberg SA, Rosenberg ML, Dhib-Jalbut S, Balashov KE. Multiple sclerosis-linked and interferon-beta-regulated gene expression in plasmacytoid dendritic cells. J Neuroimmunol. 2012;250:99–105.

31. Nickles D, Chen HP, Li MM, Khankhanian P, Madireddy L, Caillier SJ, et al. Blood RNA profiling in a large cohort of multiple sclerosis patients and healthy controls. Hum Mol Genet. 2013;22:4194–205.

32. Irizar H, Muñoz-Culla M, Sepúlveda L, Sáenz-Cuesta M, Prada Á, Castillo-Triviño T, et al. Transcriptomic profile reveals gender-specific molecular mechanisms driving multiple sclerosis progression. PLoS One. 2014;9:e90482.

33. Feldman A, Gurevich M, Huna-Baron R, Achiron A. The role of B cells in the early onset of the first demyelinating event of acute optic neuritis. Invest Ophthalmol Vis Sci. 2015;56:1349–56.

34. Hendrickx DAE, van Scheppingen J, van der Poel M, Bossers K, Schuurman KG, van Eden CG, et al. Gene Expression Profiling of Multiple Sclerosis Pathology Identifies Early Patterns of Demyelination Surrounding Chronic Active Lesions. Front Immunol. 2017;8:1810.

35. van der Poel M, Ulas T, Mizee MR, Hsiao C-C, Miedema SSM, Adelia, et al. Transcriptional profiling of human microglia reveals grey-white matter heterogeneity and multiple sclerosis-associated changes. Nat Commun. 2019;10:1139.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

36. Enz LS, Zeis T, Schmid D, Geier F, van der Meer F, Steiner G, et al. Increased HLA-DR expression and cortical demyelination in MS links with HLA-DR15. Neurol Neuroimmunol Neuroinflamm. 2020;7:e656.

37. Magliozzi R, Howell OW, Durrenberger P, Aricò E, James R, Cruciani C, et al. Meningeal inflammation changes the balance of TNF signalling in cortical grey matter in multiple sclerosis. J Neuroinflammation. 2019;16:259.

38. Martinez B, Peplow PV. MicroRNAs in Parkinson's disease and emerging therapeutic targets. Neural Regen Res. 2017;12:1945–59.

39. SRP14 Gene - GeneCards. https://www.genecards.org/cgi-bin/carddisp.pl?gene=SPR14. Accessed 28 Feb 2022.

40. Takamatsu G, Katagiri C, Tomoyuki T, Shimizu-Okabe C, Nakamura W, Nakamura-Higa M, et al. Tescalcin is a potential target of class I histone deacetylase inhibitors in neurons. Biochem Biophys Res Commun. 2017;482:1327–33.

41. Heinlein M, Gandolfo LC, Zhao K, Teh CE, Nguyen N, Baell JB, et al. The acetyltransferase KAT7 is required for thymic epithelial cell expansion, expression of AIRE target genes, and thymic tolerance. Science Immunology. 2022;7:eabb6032.

42. Jelčić I, Hsu KC, Kakalacheva K, Breiden P, Dupont B, Uhrberg M, et al. Killer immunoglobulin-like receptor locus polymorphisms in multiple sclerosis. Mult Scler. 2012;18:951–8.

43. Hoshikawa S, Ogata T, Fujiwara S, Nakamura K, Tanaka S. A novel function of RING finger protein 10 in transcriptional regulation of the myelin-associated glycoprotein gene and myelin formation in Schwann cells. PLoS One. 2008;3:e3464.

44. Santana MHA, Gomes RC, Utsunomiya YT, Neves HHR, Novais FJ, Bonin MN, et al. Short Communication Genome-wide association with residual body weight gain in Bos indicus cattle. Genet Mol Res. 2015;14:5229–33.

45. Robak LA, Jansen IE, van Rooij J, Uitterlinden AG, Kraaij R, Jankovic J, et al. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain. 2017;140:3191–203.

46. UBXN2B Gene - GeneCards. https://www.genecards.org/cgi-bin/carddisp.pl?gene=UBXN2B. Accessed 28 Feb 2022.

47. Kewitz S, Staege MS. Expression and Regulation of the Endogenous Retrovirus 3 in Hodgkin's Lymphoma Cells. Front Oncol. 2013;3.

48. Soundarapandian MM, Selvaraj V, Lo U-G, Golub MS, Feldman DH, Pleasure DE, et al. Zfp488 promotes oligodendrocyte differentiation of neural progenitor cells in adult mice after demyelination. Sci Rep. 2011;1:2.

49. Varier KM, Dhandapani H, Liu W, Song J, Wang C, Hu A, et al. An immunotherapeutic approach to decipher the role of long non-coding RNAs in cancer progression, resistance and epigenetic regulation of immune cells. J Exp Clin Cancer Res. 2021;40:242.

50. Mullin BH, Zhu K, Brown SJ, Mullin S, Tickner J, Pavlos NJ, et al. Genetic regulatory mechanisms in human osteoclasts suggest a role for the STMP1 and DCSTAMP genes in Paget's disease of bone. Sci Rep. 2019;9:1052.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

51. He R, Wu S, Gao R, Chen J, Peng Q, Hu H, et al. Identification of a Long Noncoding RNA TRAF3IP2-AS1 as Key Regulator of IL-17 Signaling through the SRSF10–IRF1–Act1 Axis in Autoimmune Diseases. JI. 2021;206:2353–65.

52. Subhi Y, Nielsen MK, Molbech CR, Liisborg C, Søndergaard HB, Sellebjerg F, et al. The transcriptome of peripheral blood mononuclear cells in patients with clinical subtypes of late age-related macular degeneration. Immun Ageing. 2019;16:20.

53. LOC102723701 Gene - GeneCards. https://www.genecards.org/cgi-bin/carddisp.pl?gene=LOC102723701. Accessed 28 Feb 2022.

54. Liu Y, Xie P, Jiang D, Liu J, Zhang J, Bian T, et al. Molecular and Immune Characteristics for Lung Adenocarcinoma Patients With ERLIN2 Overexpression. Front Immunol. 2020;11:568440.

55. Al-Saif A, Bohlega S, Al-Mohanna F. Loss of ERLIN2 function leads to juvenile primary lateral sclerosis. Ann Neurol. 2012;72:510–6.

56. van Ham TJ, Holmberg MA, van der Goot AT, Teuling E, Garcia-Arencibia M, Kim H, et al. Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. Cell. 2010;142:601–12.

57. UniProtKB - O00634 (NET3_HUMAN). http://www.uniprot.org/uniprot/O00634. Accessed 28 Feb 2022.

58. Casanova Ferrer F, Pascual M, Hidalgo MR, Malmierca-Merlo P, Guerri C, García-García F. Unveiling Sex-Based Differences in the Effects of Alcohol Abuse: A Comprehensive Functional Meta-Analysis of Transcriptomic Studies. Genes. 2020;11:1106.

59. Pérez-Díez I, Hidalgo MR, Malmierca-Merlo P, Andreu Z, Romera-Giner S, Farràs R, et al. Functional Signatures in Non-Small-Cell Lung Cancer: A Systematic Review and Meta-Analysis of Sex-Based Differences in Transcriptomic Studies. Cancers. 2021;13:143.

60. Català-Senent JF, Hidalgo MR, Berenguer M, Parthasarathy G, Malhi H, Malmierca-Merlo P, et al. Hepatic steatosis and steatohepatitis: a functional meta-analysis of sex-based differences in transcriptomic studies. Biol Sex Differ. 2021;12:29.

61. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.

62. Lorentzen AR, Karlsen TH, Olsson M, Smestad C, Mero I-L, Woldseth B, et al. Killer immunoglobulin-like receptor ligand HLA-Bw4 protects against multiple sclerosis. Ann Neurol. 2009;65:658–66.

63. Nishimura WE, Sachetto Z, Costallat LTL, Yazbek MA, Londe ACS, Guariento EG, et al. The role of KIR2DL3/HLA-C*0802 in Brazilian patients with rheumatoid vasculitis. Clinics (Sao Paulo). 2015;70:408–12.

64. Al-Mossawi H, Yager N, Taylor CA, Lau E, Danielli S, de Wit J, et al. Context-specific regulation of surface and soluble IL7R expression by an autoimmune risk allele. Nat Commun. 2019;10:4575.

65. Huygens C, Liénart S, Dedobbeleer O, Stockis J, Gauthy E, Coulie PG, et al. Lysosomal-associated Transmembrane Protein 4B (LAPTM4B) Decreases Transforming Growth Factor

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

β1 (TGF-β1) Production in Human Regulatory T Cells. Journal of Biological Chemistry. 2015;290:20105–16.

66. Kattygnarath D, Maugenre S, Neyroud N, Balse E, Ichai C, Denjoy I, et al. *MOG1*: A New Susceptibility Gene for Brugada Syndrome. Circ Cardiovasc Genet. 2011;4:261–8.

67. Espinosa Collado A, Ramos Ortiz C, Burillo Gómez F, Moreno Terribas G. Esclerosis múltiple y cardiopatía isquémica: nuevo factor de riesgo cardiovascular. Med Gen Fam. 2021;10:156–8.

68. Dai E, Guan H, Liu L, Little S, McFadden G, Vaziri S, et al. Serp-1, a Viral Anti-inflammatory Serpin, Regulates Cellular Serine Proteinase and Serpin Responses to Vascular Injury. Journal of Biological Chemistry. 2003;278:18563–72.

69. Heit C, Jackson BC, McAndrews M, Wright MW, Thompson DC, Silverman GA, et al. Update of the human and mouse SERPINgene superfamily. Hum Genomics. 2013;7:22.

70. Wang Z, Tian Z, Song X, Zhang J. Membrane tension sensing molecule-FNBP1 is a prognostic biomarker related to immune infiltration in BRCA, LUAD and STAD. BMC Immunol. 2022;23:1.

71. Newman DM, Voss AK, Thomas T, Allan RS. Essential role for the histone acetyltransferase KAT7 in T cell development, fitness, and survival. J Leukoc Biol. 2017;101:887–92.

72. Gkika D, Lemonnier L, Shapovalov G, Gordienko D, Poux C, Bernardini M, et al. TRP channel–associated factors are a novel protein family that regulates TRPM8 trafficking and activity. Journal of Cell Biology. 2015;208:89–107.

73. Yin L, Zhou L, Xu R. Identification of Tumor Mutation Burden and Immune Infiltrates in Hepatocellular Carcinoma Based on Multi-Omics Analysis. Front Mol Biosci. 2020;7:599142.

74. Yu W, Wang H, He Q, Xu Y, Wang X. Long noncoding RNAs in cancer-immunity cycle. J Cell Physiol. 2018;233:6518–23.

75. Nesslinger NJ, Sahota RA, Stone B, Johnson K, Chima N, King C, et al. Standard treatments induce antigen-specific immune responses in prostate cancer. Clin Cancer Res. 2007;13:1493–502.

76. Swaidani S, Liu C, Zhao J, Bulek K, Li X. TRAF Regulation of IL-17 Cytokine Signaling. Front Immunol. 2019;10:1293.

77. Jeffcoat S, Givens L, Bolanos S, Malhotra S, Koke J. A novel intermediate filament-associated protein: further characterization of the G.3.5 antigen. Cytobios. 1995;82:81–99.

78. Huang H, Xing D, Zhang Q, Li H, Lin J, He Z, et al. LncRNAs as a new regulator of chronic musculoskeletal disorder. Cell Proliferation. 2021;54:e13113.

79. Hashiguchi K, Hayashi M, Sekiguchi M, Umezu K. The roles of human MTH1, MTH2 and MTH3 proteins in maintaining genome stability under oxidative stress. Mutat Res. 2018;808:10–9.

80. Smith-Bouvier DL, Divekar AA, Sasidhar M, Du S, Tiwari-Woodruff SK, King JK, et al. A role for sex chromosome complement in the female bias in autoimmune disease. J Exp Med. 2008;205:1099–108.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

81. Sasidhar MV, Itoh N, Gold SM, Lawson GW, Voskuhl RR. The XX sex chromosome complement in mice is associated with increased spontaneous lupus compared with XY. Ann Rheum Dis. 2012;71:1418–22.

82. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016;16:626–38.

83. Srivastava S, D'Amore A, Cohen JS, Swanson LC, Ricca I, Pini A, et al. Expansion of the genetic landscape of ERLIN2-related disorders. Ann Clin Transl Neurol. 2020;7:573–8.

84. Amador M-D-M, Muratet F, Teyssou E, Banneau G, Danel-Brunaud V, Allart E, et al. Spastic paraplegia due to recessive or dominant mutations in ERLIN2 can convert to ALS. Neurol Genet. 2019;5:e374.

85. Campbell G, Mahad DJ. Mitochondrial dysfunction and axon degeneration in progressive multiple sclerosis. FEBS Lett. 2018;592:1113–21.

86. Bargiela D, Chinnery PF. Mitochondria in neuroinflammation - Multiple sclerosis (MS), leber hereditary optic neuropathy (LHON) and LHON-MS. Neurosci Lett. 2019;710:132932.

87. Bartosik-Psujek H, Belniak E, Mitosek-Szewczyk K, Dobosz B, Stelmasiak Z. Interleukin-8 and RANTES levels in patients with relapsing-remitting multiple sclerosis (RR-MS) treated with cladribine. Acta Neurol Scand. 2004;109:390–2.

88. Matejčíková Z, Mareš J, Přikrylová Vranová H, Klosová J, Sládková V, Doláková J, et al. Cerebrospinal fluid inflammatory markers in patients with multiple sclerosis: a pilot study. J Neural Transm (Vienna). 2015;122:273–7.

89. Lu CZ, Jensen MA, Arnason BG. Interferon gamma- and interleukin-4-secreting cells in multiple sclerosis. J Neuroimmunol. 1993;46:123–8.

90. Saligrama N, Zhao F, Sikora MJ, Serratelli WS, Fernandes RA, Louis DM, et al. Opposing T cell responses in experimental autoimmune encephalomyelitis. Nature. 2019;572:481–7.

91. Kunkl M, Frascolla S, Amormino C, Volpe E, Tuosto L. T Helper Cells: The Modulators of Inflammation in Multiple Sclerosis. Cells. 2020;9:E482.

92. Cuevas-García CF, Segura-Méndez NH, Herrera-Sánchez DA. Actualidades en la inmunopatología de la esclerosis múltiple. Gac Med Mex. 2018;154:588–97.

93. Ridgley LA, Anderson AE, Pratt AG. What are the dominant cytokines in early rheumatoid arthritis? Curr Opin Rheumatol. 2018;30:207–14.

94. Kuzmina US, Zainullina LF, Vakhitov VA, Bakhtiyarova KZ, Vakhitova YV. The role of glutamate in the pathogenesis of multiple sclerosis. Zh Nevrol Psikhiatr Im S S Korsakova. 2019;119:160–7.

95. Zahoor I, Rui B, Khan J, Datta I, Giri S. An emerging potential of metabolomics in multiple sclerosis: a comprehensive overview. Cell Mol Life Sci. 2021;78:3181–203.

96. Correale J. Immunosuppressive Amino-Acid Catabolizing Enzymes in Multiple Sclerosis. Front Immunol. 2021;11:600428.

97. Palumbo S, Bosetti F. Alterations of brain eicosanoid synthetic pathway in multiple

medRxiv preprint doi: https://doi.org/10.1101/2021.08.31.21262175; this version posted May 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

sclerosis and in animal models of demyelination: role of cyclooxygenase-2. Prostaglandins Leukot Essent Fatty Acids. 2013;89:273–8.

98. Rassart E, Desmarais F, Najyb O, Bergeron K-F, Mounier C. Apolipoprotein D. Gene. 2020;756:144874.

99. Fitzgerald KA, Kagan JC. Toll-like Receptors and the Control of Immunity. Cell. 2020;180:1044–66.

100. Vijay K. Toll-like receptors in immunity and inflammatory diseases: Past, present, and future. International Immunopharmacology. 2018;59:391–412.

101. Kim YK, Shin JS, Nahm MH. NOD-Like Receptors in Infection, Immunity, and Diseases. Yonsei Med J. 2016;57:5–14.

102. Haase S, Haghikia A, Wilck N, Müller DN, Linker RA. Impacts of microbiome metabolites on immune regulation and autoimmunity. Immunology. 2018;154:230–8.

103. Martinez-Altarriba MC, Ramos-Campoy O, Luna-Calcaño IM, Arrieta-Antón E. A review of multiple sclerosis (2). Diagnosis and treatment. Semergen. 2015;41:324–8.

104. Torre-Fuentes L, Moreno-Jiménez L, Pytel V, Matías-Guiu JA, Gómez-Pinedo U, Matías-Guiu J. Experimental models of demyelination and remyelination. Neurologia (Engl Ed). 2020;35:32–9.

105. Engel KL, Arora A, Goering R, Lo H-YG, Taliaferro JM. Mechanisms and consequences of subcellular RNA localization across diverse cell types. Traffic. 2020;21:404–18.

106. Junjappa RP, Patil P, Bhattarai KR, Kim H-R, Chae H-J. IRE1α Implications in Endoplasmic Reticulum Stress-Mediated Development and Pathogenesis of Autoimmune Diseases. Front Immunol. 2018;9:1289.

107. Cruz FM, Colbert JD, Rock KL. The GTPase Rab39a promotes phagosome maturation into MHC-I antigen-presenting compartments. EMBO J. 2020;39:e102020.

108. Riccio P, Rossano R. Diet, Gut Microbiota, and Vitamins D + A in Multiple Sclerosis. Neurotherapeutics. 2018;15:75–91.

109. Yanguas-Casás N. Sex Differences in Neurodegenerative Diseases. SM Journal of Neurological Disorders and Stroke. 2017;3:1014.

110. Voskuhl RR, Sawalha AH, Itoh Y. Sex chromosome contributions to sex differences in multiple sclerosis susceptibility and progression. Mult Scler. 2018;24:22–31.

111. Clayton JA, Collins FS. Policy: NIH to balance sex in cell and animal studies. Nature. 2014;509:282–3.

112. European Commission. Directorate General for Justice and Consumers., Fondazione Giacomo Brodolini. Gender equality and health in the EU. LU: Publications Office; 2021.

113. Accounting for sex and gender makes for better science. Nature. 2020;588:196–196.