A deep transcriptome meta-analysis reveals sex-based molecular differences in Multiple Sclerosis

José Francisco Català-Senent[1†], Zoraida Andreu[2†], Francisco José Roig[1,3], Marta R. Hidalgo[1], 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

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

[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

Abstract

Background: Multiple Sclerosis (MS) is a chronic autoimmune, inflammatory, and degenerative disease of the central nervous system that affects both women and men but the risk of developing MS is higher in women (2-3:1 ratio compared to men). Current knowledge does not allow a precise definition of the sex-related factors intervening in MS. Here, we explore the role of sex in this neurodenegerative disease to identify molecular mechanisms underlying sex-based diferences that can guide novel therapeutics approaches aimed for men and women.

Methods: We performed a rigorous and systematic review of MS whole transcriptome studies that included sex patient information in Gene Expression Omnibus and ArrayExpress databases, following PRISMA statment guidelines. A differential gene expression (DGE) was performed for each selected study. Then, three meta-analyses based on genes and different contrasts were addressed to evaluate common features and sex bias in all scenarios: one meta-analysis in nervous tissue studies (4), another in blood studies (5), and a third that integrated the studies from both tissues, blood and nervous system (9). Finally, a gene set analysis (GSA) was performed on the meta-analyzed differential transcriptomic profile between women and men, to study their involvement in biological pathways and phenotypes (physiological and pathological states) sex related.

Results: After screening 122 publications, the systematic review provided a selection of nine studies (5 in blood and 4 in nervous tissue) with a total of 653 samples (269 MS women, 152 control women, 116 MS men, 116 control men). The blood sex-differences gene meta-analysis identified 1 gene overexpressed in women (KIR2DL3) while gene meta-analysis in the nervous tissue resulted in 5 genes differentially expressed (4 overexpressed in women: C7orf73, CECR7, TRAF3IP2-AS1, ZNF117; and 1 overexpressed in men: HIST1H2AE). In the global meta-analysis (blood and nervous tissue), 1 gene involved in sex differences were detected (more expressed in women: LOC102723701). 9 MS biomarkers were also identified in both sexes (4 overexpressed genes in MS patients: HIST1H2BC, SMEK2, STBD1, TMEM140) and 5 down-regulated in MS patients: C16orf59, C20orf177, C9orf23, FAM65B, PKI55). Additionally, the GSA related to gene meta-analysis revealed a set of functions differentially enriched in men and women, in the different analysed tissues.

Conclusion: These findings provide valuable opportunities for better understanding of MS related sex differences and develop effective and sex-specific interventions down further studies.

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

Background

Multiple Sclerosis (MS) is considered an special case of chronic neurodegenerative disease since it progresses with degeneration of the nervous system promoted by exacerbated inflammatory autoimmune reaction that causes demyelination of the central nervous system. Main MS symptoms are fatigue, lack of balance, pain, visual and cognitive alterations, speech difficulties, and tremor. MS is mainly diagnosed in young adults (around 25-30 years) and usually runs a relapsing-remitting course with non-traumatic neurological disability. This disease represents the main cause of disability among young adults not associated with accidents. Four types of MS have been identified: Clinical Isolated Syndrome (CIS), the earliest form diagnosed and characterized by 24 hour episodes; Relapsing-Remitting (RRMS), the most common form of MS, (85% of people affected), followed by a Secondary-Progressive course (SPMS), in which the remission period lasts less between relapses and the less frequently Progressive-Relapse (PRMS) that has a progressive course right from the onset [1].

MS etiology is unknown but a combination of genetic and environmental components are involved in the disease-onset with a high dimorfism sexual influence [2]. As other autoinmune and neurodegenerative diseases, MS presents a strongest sex bias in the incidence, prevalence, symptom severity, disease course, response to therapy and overall survival [3, 4]. Women have a higher susceptibility to MS than men. However, men with MS are, on average, older than women at diagnosis, have worse and faster progression. Men are also more likely to develop the progressive form of the disease than women [5, 6]. This rapid rise cannot result from genetic mutations, which require time. Sex-based differences in autoinmune and neurological diseases is well reported [3, 5] and highlight in the presence of sexual chromosomes (XX vs XY, gene dosed, genetic imprint and their transcriptomic profile) [6, 7], and the endogeneous sex hormones (estrogens, progesterone, androgens and prolactin) that influence several features of inmune system and promote changes in severity disease with estrogens as potent stimulators of autoimmunity and androgens playing a protective role [1]. Additionally, sex chromosomes transcriptomic profiles and epigenetic modifications can influence hormonal actions on inmune response in both sexes. Response to environmental factors as microbial exposure, diet, and sociological differences, among others, represent other sex-related factors in MS [8]. Therefore, differences in prevalence and severity of MS between males and females involve complex and vaguely understood interactions between genetic, hormonal and environmental closely sex-related factors. There is currently no cure for MS. Therefore, the therapies developed are focused on alleviating the symptoms, as well as slowing down the worsening of the disease.

Multi-omics studies in large and clinically well characterized patient cohorts with accessible sex information data, can help to identify critical pathways and networks related to sexual dimorphism. These bioinformatics strategies can be useful for the discovery of new biomarkers and targeted for sex-specific therapeutic intervention and/or prevention of MS. Highlight, that *in silico* approaches through a computational models with a large volume of biological information, accessible in public repositories of biomedical data, such as GEO (Genes Omnibus Expression) [9] represent a powerful tool to get a better knowledge in biomedical areas and contribute to solve biological or clinical problems. In particular, meta-analyses of transcriptomics studies allow a consensus pattern of expression to be achieved

among several studies, thanks to their capacity for integration and greater statistical power than individual studies.

In this context, we performed a systematic review and selection of whole transcriptome studies related to MS, from GEO and ArrayExpress databases [9, 10] that include sex patient information. A total of 653 samples were analyzed (269 MS women, 152 control women, 116 MS men, 116 control men). Given that MS presents both neurodegenerative and autoimmune features, the systematic review included studies focused on nervous tissue but also on blood studies where the autoimmune profile could be reflected and offer a possible source of non-invasive biomarkers. Different contrasts in men and women were evaluated to check common genes involved in MS and differential expression related to sex. After differential gene expression (DGE) analysis, nervous studies and blood studies were meta-analysed separately and a third meta-analysis including both tissues were also performed. Related to sex differences individual meta-analysis in blood detected one gene differentialy expressed (KIR2DL3), and 5 genes in nervous tissue. In the global meta-analysis (blood and nervous tissue), 1 gene involved in sex differences were detected (LOC102723701). Additionally, a functional enrichment analysis of nervous tissue meta-analysis provided key information about biologic processes and metabolic pathways in men and women, to better understand sex differences in MS. These identified molecular mechanisms may have important clinical implications for the prevention and treatment of MS, and could potentially favor the development of novel diagnostic and therapeutic applications with a personalized approach to this disease.

Methods

Literature review and selection of studies

The systematic review was performed by searching for the keyword “multiple sclerosis” on the GEO and ArrayExpress [9, 10] transcriptomic databases. The review 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", iii) sample count: at least 12 samples. To ensure the quality and transparency of the systematic review, the PRISMA Statement guidelines were followed [11].

The following exclusion criteria were applied: i) studies not based on MS, ii) experimental design other than patients vs. controls, iii) absence of information on the sex of the individuals. In addition, studies whose individuals had received some type of treatment or were immunosuppressed were also excluded.

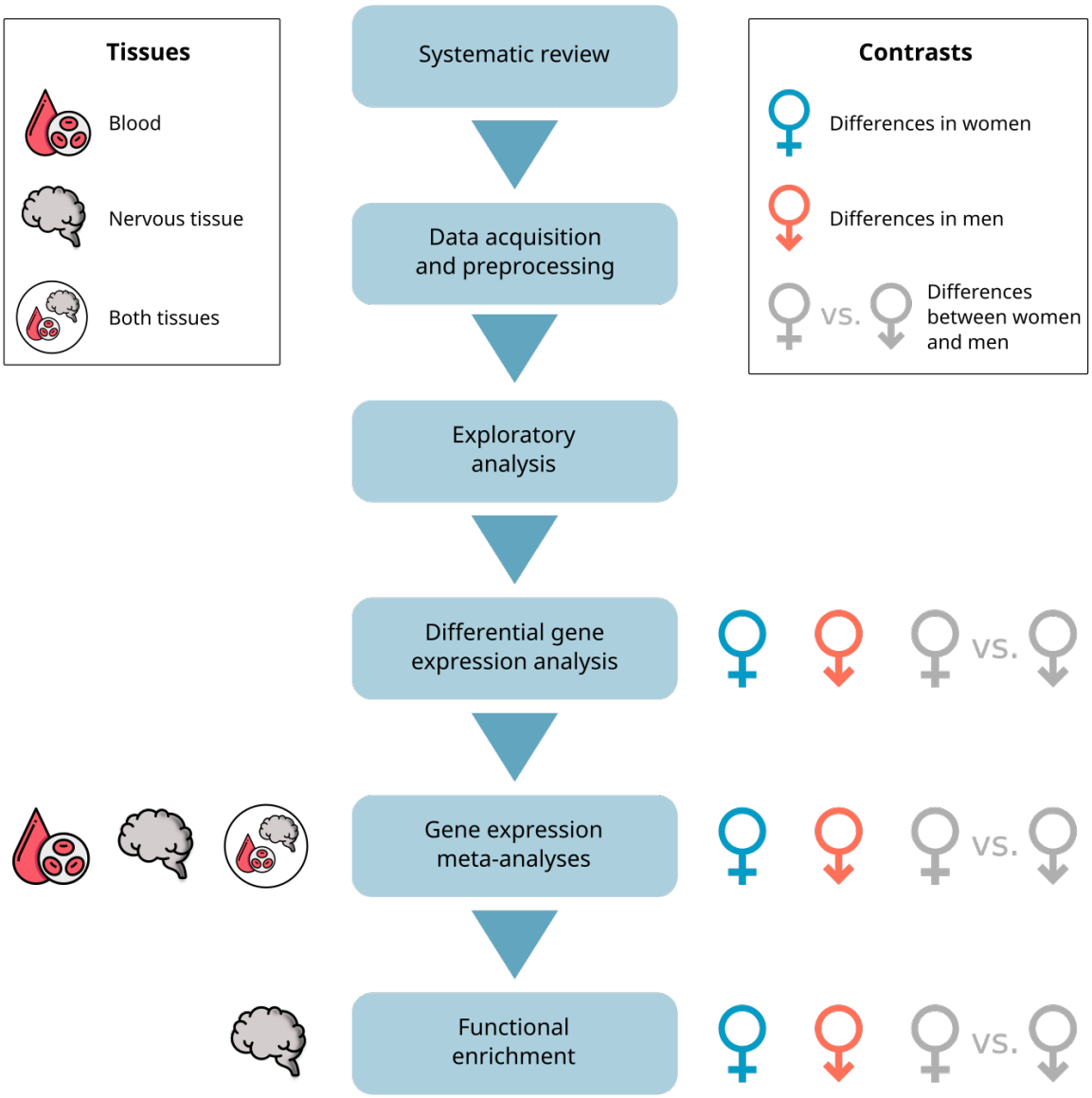

Figure 1. Flowchart of the work performed, indicating tissues analyzed and contrasts performed in each step**.**

Bioinformatics analysis

For each of the selected studies, the same analysis strategy was followed: i) data download and normalization, ii) exploratory analysis, iii) differential gene level expression analysis (DGE), iv) integration for DGE results using meta-analysis techniques, taking into account the tissue of origin of the samples (blood, nervous and both), to obtain robust results across all studies, v) functional enrichment of the meta-analysis results (Figure 1). This analysis of grouping the studies according to their tissue or together, allowed us to explore sex differences in the disease globally and specifically in blood and nerve tissue. Bioinformatics analysis was carried out using the programming language R 3.6.2 [12]; information regarding the packages used and their version is available in the zenodo repository (http://doi.org/10.5281/zenodo.5211962).

Data acquisition and preprocessing

The selected studies in the systematic review were downloaded and standardized. Gene names were converted to Gene Symbol nomenclature and the median expression was calculated in the case of multiple values for the same gene. The labels of the sex of the patients were homogenized and the individuals were grouped into healthy or affected by the disease. Paired samples were tagged to account for this factor in subsequent analyses. After data normalization, an exploratory analysis was performed, using clustering and principal component analysis (PCA), to detect possible batch effects and anomalous behavior in the data.

Differential gene expression

After downloading and normalizing the data, several DGE analyses were performed, for each individual study, to detect differences in gene level expression between groups. First, we designed a contrast to detect sex-related differencies in MS by the comparison: (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl). Additionally, differential expression analysis was also performed to discover which genes are affected in females (Women.MS - Women.Ctrl) and males (Men.MS - Men.Ctrl) with the disease. Comparative analysis of these results constituted a multiple approach for the detection of sex differences in MS. All contrasts were performed using the R package limma [13]. After calculation of the differential expression statistics, the p-values were adjusted by the Benjamini & Hochberg (BH) method.

Meta-analysis

To robustly integrate the results of the DGEs of each study, meta-analysis techniques were used. The results of each of the three contrasts were combined separately: (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl), (Women.MS - Women.Ctrl) and (Men.MS - Men.Ctrl). For each contrast, 3 meta-analyses were performed according to the tissue of origin of the samples: blood, nervous tissue or both tissues. In the latter case, only genes detected in at least one study from each of the two tissue types are integrated.

The meta-analyses were performed following the methodology described by García-García [14]. Briefly, the metafor R package [15] was used to combine the expression results of the corresponding individual studies, using for each gene a random-effects model, particularly suitable when there are different sources of variability in the studies. [16]. The measure of effect used in these meta-analyses of gene expression levels was the fold change (LogFC), which tells us whether a gene has a higher mean expression level in one group or another, when we compare two experimental classes. For the overall calculation of the LogFC, the less variable results were given greater weight by considering the variability of the individual studies in the global estimation of the measured effect. Better integration of the selected studies and statistically stronger results are achieved by incorporating the variability between experiments in the random effects model. The suitability of this method was tested by assessing the heterogeneity and influence of each study on the overall model and cross-validation techniques.

For each gene, the meta-analysis calculates a pooled measure of the expression effect across all studies, using the logFC and its 95% confidence interval. It also provides a contrast to assess the significance of the logFC, as well as the BH-adjusted p-value due to the multidisciplinary setting in which we work. Those genes with an adjusted p-value < 0.05 were considered significant. The Open Targets platform [17] (release 21.06) was used to test the association of genes to diseases.

Functional enrichment analysis

To detect functions or pathways associated to MS disease in the individual women and men contrasts but also sex related, a gene set analysis (GSA) of nervous tissue meta-analyses was performed. Firstly, the genes were ordered according to the p-value and sign of the statistic obtained in the DGE. Next, GSA was performed using the logistic regression model implemented in the mdgsa R package [18]. The logarithm of odds ratio (LOR) is the measure of effect in functional enrichment analysis, and is obtained directly from logistic regression. For example, in the Men.MS - Men.Ctrl contrast, a LOR > 0 informs us that there is a greater overrepresentation of this function in men with MS, while a LOR < 0 would indicate that this function or pathway is more inhibited in men with MS than in Men.Ctrl. A higher magnitude of the LOR will mean that there is a higher level of overrepresentation of that function or pathway. The functional annotation required for the GSA was obtained from the following databases: Gene Ontology (GO) [19], in the case of biological processes (BP), and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway [19] for pathways. Due to their hierarchical structure, the gene annotations with GO terms applied in the mdgsa package were propagated to inherit the annotations of the ancestor terms. Excessively specific or generic annotations (blocks smaller than ten or larger than 500 words) were subsequently filtered out. Finally, functions with a BH-adjusted p value under 0.05 were considered significant.

Metafun-MS platform

All data and results generated in the different steps of the meta-analysis are available in the Metafun-MS platform (https://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 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 [21] except for the exploratory analysis cluster plot, which was generated with ggplot2 [22].

This easy-to-use resource is divided into five sections: 1) Summary of analysis results in each phase. Then, for each of the studies, the detailed results of the 2) exploratory analysis, 3) differential expression, 4) meta-analysis and, 5) functional profiling. The user can interact with the web tool through graphics and tables and search information for specific genes or biological functions.

Results

Systematic review and exploratory analysis

Our systematic review screened 122 publications and resulted in a total of 9 studies eligible for our analyses: 5 studies with blood samples and 4 studies with nerve tissue (Figure 2, Table 1). It is noteworthy that 35% of the studies in which the presence of patients of both sexes was evaluated had to be discarded because they did not have this information or because they were performed with individuals of only one sex.

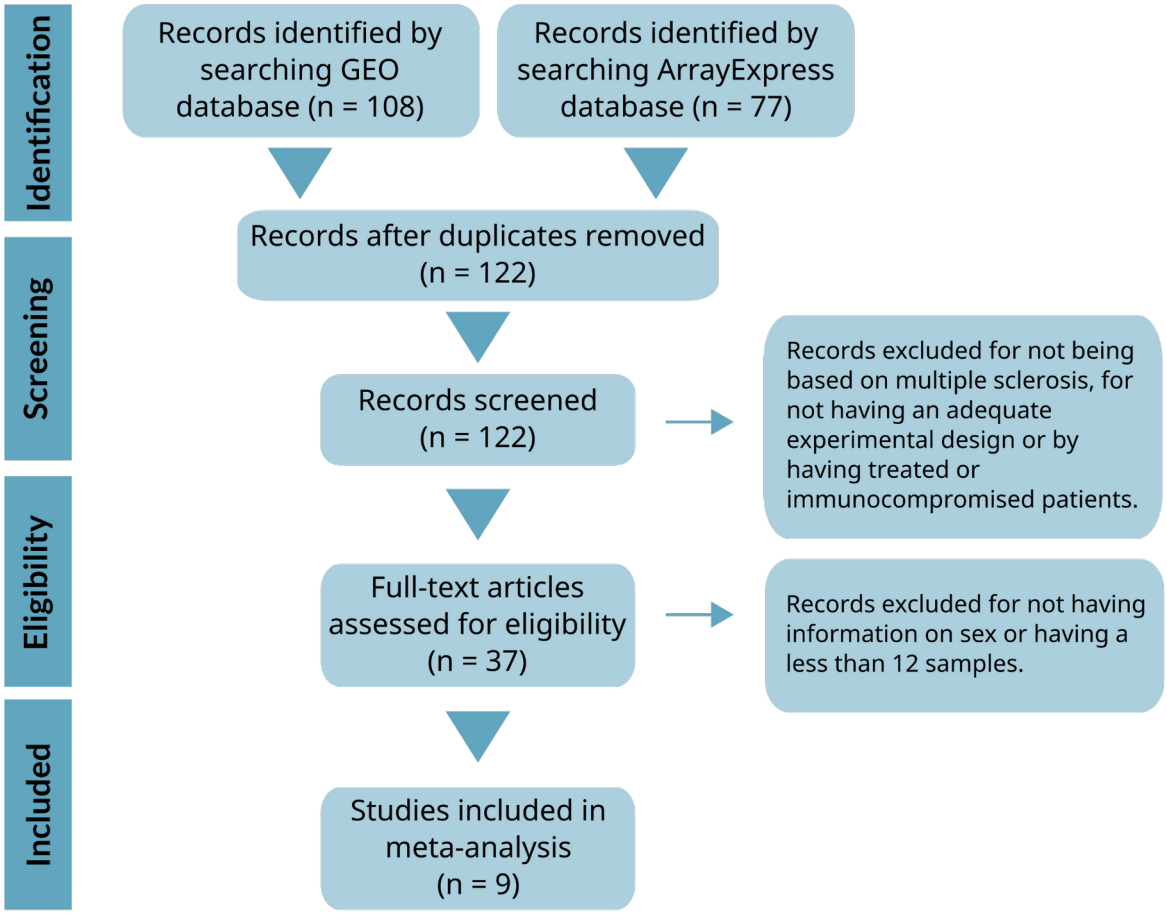

Figure 2. PRISMA diagram of the systematic review.

The 9 selected studies contained 653 patients, of whom 64.5% are women and 35.5% men. Point that two-thirds of the samples were blood, probably, because of the ease of sample collection. The 59% of the analyzed samples belong to individuals with MS (358 MS samples) while the 41% are control individuals (268 control samples). The detailed distribution of the samples is shown in Table 2.

Table 1. Description of selected studies. Platform, tissue and MS subtype of the samples are detailed from each of the nine studies.

GEO accession	Platform	Tissue	Subtypes of MS
GSE37750 [23]	Affymetrix Human Genome U133 Plus 2.0 Array	Plasmacytoid dendritic cells	MS*
GSE41848 [24]	Affymetrix Human Exon 1.0 ST Array	Whole blood	Relapse, remission
GSE41849 [24]	Affymetrix Human Exon 1.0 ST Array	Whole blood	MS*
GSE41890 [25]	Affymetrix Human Gene 1.0 ST Array	Peripheral blood leukocytes	MS*
GSE62584 [26]	Affymetrix Human Genome U133 Plus 2.0 Array	CD4+ blood cells, CD8+ blood cells, CD14+ blood cells, CD19+ blood cells	ON patient in relapse
GSE108000 [27]	Agilent-026652 Whole Human Genome Microarray 4x44K v2	Brain	RIM chronic active, RIM inactive, non-lesional chronic active, non-lesional inactive
GSE111972 [28]	Illumina NextSeq 500	Corpus callosum, occipital cortex	MS*
GSE131281 [29]	Illumina HumanHT-12 V4.0 expression beadchip	Grey Matter	SPMS, PPMS
GSE135511 [30]	Illumina HumanRef-8 v3.0 expression beadchip	Motor cortex	MS lesion, MS normal appearing

MS*: when MS subtype is not indicated in the study.

Table 2. Distribution of the samples along each of the groups, in each study.

GEO accession	Tissue	Women		Men		Total
		MS	Control	MS	Control	
GSE37750	Blood	7	6	2	2	17
GSE41848	Blood	98	59	35	20	212
GSE41849	Blood	35	25	21	21	102
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		269	152	116	116	653

Exploratory analysis showed the absence of biases and anomalous behavior in the samples of most of the studies. Only in the GSE62584 study a batch effect was detected, which was considered in the following analyses, produced by the 4 cell types analyzed in that study.

Differential gene expression analysis

As described in material and methods section, for each selected study were evaluated 3 contrasts: (Women.MS - Women.Ctrl): women, (Men.MS - Men.Ctrl): men, (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl): sex differences. The individual contrasts, (only men or women) showed us which genes are common and significant for the disease in both sexes. Emphasize that these contrasts are not as robust to identify sex-specific genes involved in the disease, information that is provided by the contrast (Women.MS - Women.Ctrl): women, (Men.MS - Men.Ctrl), which jointly assesses the samples of both sexes. The number of significantly differentially expressed genes is shown in Table 3. Detailed results for each comparison are shown in the zenodo repository and on the Metafun-MS web platform (see below).

Table 3. Number of significantly differentially expressed genes by study. Description of significant genes by contrasts: (Women.MS - Women.Ctrl): women, (Men.MS - Men.Ctrl): men, (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl): sex differences. The last column reports the total number of genes assessed in each study, depending on the technology used.

Study (tissue)	Women	Men	Sex differences	Genes analyzed
GSE37750 (Blood)	0	0	0	22899
GSE41848 (Blood)	32	0	0	17479
GSE41849 (Blood)	0	0	0	17479
GSE41890 (Blood)	21	96	30	19914
GSE62584 (Blood)	0	80	501	13277
GSE108000 (Nervous)	5152	4020	384	21586
GSE111972 (Nervous)	0	0	0	20846
GSE131281 (Nervous)	35	4	0	20923
GSE135511 (Nervous)	2	409	4	15434

The disparity of significant results from each study shows the need for the use of powerful integration tools, such as meta-analysis, that combine expression results from several studies, and extract the common signal across them.

Meta-analysis of DGE profiles

Gene expression meta-analysis was performed, from the DGE results obtained in the previous section, for each of the three contrasts. In addition, 3 meta-analyses were performed, depending on whether blood studies, nerve tissue studies, or both were included in the meta-analysis. When analyzing the results of this 9 meta-analyses (3 contrasts x 3 tissues), we should not only take into account the adjusted p-value as an indicator of the significance level of the genes, but we should also valorate the logFC that indicates the magnitude of the observed change. The sign of the logFC shows in which component of the comparison the

gene is overexpressed. In the comparisons of women and men, a positive logFC indicates that the gene is overexpressed in patients, while a negative logFC indicates overexpression in control subjects. Whereas that the comparative with sex perspective: (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl), a positive logFC represents a higher gene expression in women (Women.MS - Women.Ctrl) and a negative logFC in men (Men.MS - Men.Ctrl).

Meta-analysis of DGE identified only one significant gene in blood, related to sex differences in MS, while 5 significant genes were detected in nervous tissue, and one gene in both tissues according to the amount of genes found in the individual contrasts, shown in Table 4.

All significant genes for each contrast and their detailed statistical results can be accessed in the zenodo repository and on the Metafun-MS web platform.

Table 4. Number of significant genes (adjusted p-value < 0.05) in the meta-analyses performed. Description of contrasts: (Women.MS - Women.Ctrl): women, (Men.MS - Men.Ctrl): men, (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl): sex differences.

	Women	Men	Sex differences
Blood	8	21	1
Nervous	59	79	5
Both tissues	3	19	1

Meta-analysis of DGE in blood

In blood samples, 8 genes are dysregulated in the women contrast (5 genes upregulated; 3 downregulated in Women.MS) and 21 genes in men contrast (16 upregulated; 5 downregulated in men.MS). No significant common genes between women and men were found in blood tissue and only the KIR2DL3 gene was found to be significant in the sex contrast (Women.MS - Men.MS), as shown in upset Figure 3. This gene is overexpressed in Women.MS compared to men.MS. KIR2DL3 encodes for the "Killer Cell Immunoglobulin-Like Receptor 2DL3, a transmembrane glycoprotein expressed by natural killer cells and subsets of T cells. The absence of the inhibitory KIR2DL3 gene is associated with the development of MS [31].

Table 5. MA results of blood samples. Described contrasts "sex differences", "women" and "men" refer (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl), (Women.MS - Women.Ctrl) and (Men.MS - Men.Ctrl), respectively. The red color in the genes means a positive logFC (overexpressed genes), while the blue color indicates a negative logFC (down expressed genes).

Contrast	**Significant genes**
Sex differences	KIR2DL3
Women	C2orf79, LAPTM4B, PCK2, RANGRF, SERPINF1, FNBP1, MYST2, SETD5
Men	COMMD9, COX6B1, COX8A, GLA, GRIK4, MIR6773, MRPL23, NDUFA2, PSMB2, SOCS3, SRP14, SSR4, TESC, TMX4, TXNL4A, UXT, FAM115A, KIR2DL3, LINC00965, LOC101930363, LOC102725213

In the individual women contrast, were identified as upregulated genes such as: **C2orf79** also known as PTRHD1 (Peptidyl-TRNA Hydrolase Domain Containing 1, intellectual disability and parkinsonism) [32]; **LAPTM4B** (lysosomal protein, dementia with lewy bodies) [33]; **PCK2** (Phosphoenolpyruvate Carboxykinase 2, Mitochondrial); **RANGRF** (MOG1 protein) and **SERPINF1 (Serpin Family F** Member 1, codes for pigment-epithelium derived factor (PEDF) and involve in osteogenesis imperfecta) [34]. The following genes appeared downregulated (Table 5): **FNBP1** (Formin Binding Protein 1; Eccrine Adenocarcinoma and Familial Hypocalciuric Hypercalcemia as associated diseases), **MYST2** (histone acetyltransferase complex, that plays crucial roles in gene regulation, DNA replication, development, as well as, maintaining pluripotency in embryonic stem cells (ESCs)) [35]; **SETD5** (SET Domain Containing 5, diseases associated with SETD5 include mental retardation, cognitive functions alteration) [36, 37].

In the individual men contrast, the meta-analysis found overexpression in men.MS of the genes: **COMMD9** (COMM Domain Containing 9, related to mental retardation syndrome and related to innate immune system); **COX genes** (COX6B1 and COX8A, two cytochrome oxidases involved in mitochondrial function); **GLA** (gene related to galactosidase metabolism); **GRIK4** (a glutamate receptor ionotropic kainate 4 with a role in cognitive dysfunction and neurodegenerative diseases) [38]; **PSMB2** (proteasome B-type family, whose dysregulation has been described in Parkinson's disease) [39]; **SOCS3 protein**, (a suppressor of cytokine signaling); **SRP14** (Signal Recognition Particle 14,whose related pathways are metabolism of proteins and innate immune system); **TESC** (Tescalcin, that include calcium ion binding and magnesium ion binding)**.** This gene has been described as a potential target of class I histone deacetylase inhibitors in neurons with a neuroprotective effect [40]; **TMX** (transmembrane Trx-like protein 4, with reductase activity) [41]; TXNL4A, (subunit of a protein complex called the major spliceosome, related to craniofacial developmental disorder) [42]; UXT, (Ubiquitously Expressed Prefoldin Like Chaperone, that functions as a cofactor that modulates androgen receptor-dependent transcription), essential for cell viability and cell transformation. Among these genes, COX8A, GLA, GRIK4, NDUFA2, SOCS3 genes have been described in Open Targets related to other neural disorders. Other genes as **FAM115A** also know as **TCAF1** (TRPM8 channel associated

factor 1, regulates the plasma membrane cation channel TRPM8 activity); **KIR2DL3** also involved in sex differences as described before; **LINC00965** (long intergenic non-protein coding RNA 965); **LOC101930363** and **LOC102725213** were downregulated in men.MS (Table 5).

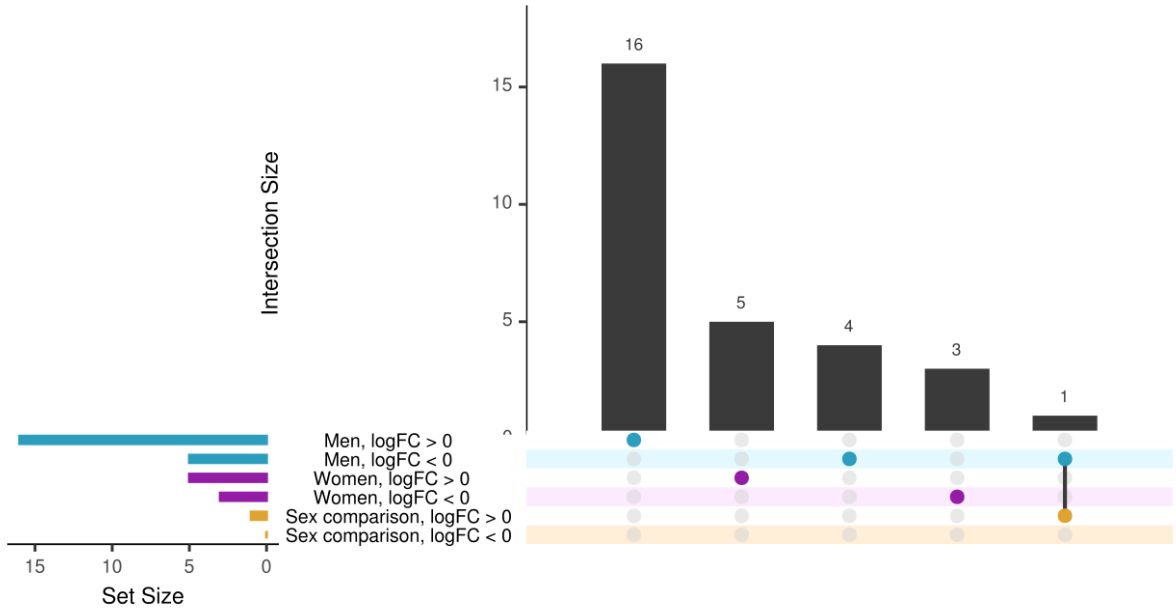

Figure 3. Upset plot of the meta-analysis results in blood studies. The results of each comparison are separated according to the sign of the logFC. Horizontal bars indicate the number of significant genes in each contrast (one specific color for each contrast). Results are exhibited separately by up-regulated (logFC > 0) and down-regulated (logFC < 0) significant genes. The vertical bars display the common genes in the sets indicated with dots under each bar. The single points represent the number of unique genes in each group.

Meta-analysis of DGE in Nervous tissue

A higher number of significant DGEs were detected in nervous tissue than in blood or both tissues together, consistent with the fact that this tissue is the main affected in MS. According to sex differences, 5 genes (Table 6) were found differentially expressed in nervous tissue (4 up, 1 down). The up genes in Women.MS corresponded to **C7orf73** also known as STMP1 (a Short Transmembrane Mitochondrial Protein 1 related to osteitis deformans in Open Targets); **CECR7** (Cat Eye Syndrome Chromosome Region, Candidate 7); **TRAF31IP2-AS1** (Antisense RNA 1 involved in osteosarcoma reported by Open Targets); **ZNF117** (Zinc Finger Protein related, associated with Metopic craniosynostosis in Open Targets). In women **HIST1H2AE** (Histone H2A type 1-B/E related to Alzheimer and osteosarcoma) [43, 44] appeared downregulated.

Meta-analysis of individual contrasts in women and men separately detected a higher

number of significant genes in nerve tissue than in blood or both tissues together: 50 (24 up and 26 down) and 79 (33 up; 46 down) dysregulated significant genes, respectively (Table 6). Results for all these genes are included in the zenodo repository and on the Metafun-MS website. Given the high number of the genes identified in DGE analysis in nervous tissue, a GSA was also performed to provide an integrative functional and biological perspective of these genes in women and men with MS following GO biological processes and KEGG pathways that is shown later.

Sex-stratified meta-analyses identified 9 MS biomarkers in both men and women (Figure 4, columns 7 and 8). 5 downregulated genes in MS patients: **C16orf59**, **C20orf177**, **C9orf23**, **FAM65B**, **PKI55**; and 4 upregulated genes in MS patients: **HIST1H2BC**, **SMEK2**, **STBD1**, **TMEM140**.

Table 6. MA results of nervous tissue samples. Described contrasts: "sex differences", "women" and "men" refer (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl), (Women.MS - Women.Ctrl) and (Men.MS - Men.Ctrl), respectively. The red color in the genes means a positive logFC, while the blue color indicates a negative logFC.

Contrast	**Significant genes**
Sex differences	C7orf73, CECR7, TRAF3IP2-AS1, ZNF117, HIST1H2AE
Women	ASCL2, BMP6, CCNA2, CECR7, HIST1H2AJ, HIST1H2BC, HIST2H3D, HLA-DRB1, HLA-DRB5, IGLL5, LTBP2, MGAM, MICA, OLFML2B, PHF16, RGMB, SLC30A7, SMEK2, STBD1, STON1-GTF2A1L, TMEM140, TTC17, TXLNB, ZSCAN29, ANKRD26P3, C16orf59, C18orf19, C19orf47, C19orf48, C1orf204, C20orf177, C9orf23, FAM65B, FAM84A, FDX1L, FGFBP3, GTF2H4, HPDL, HRASLS, KIAA1644, KIAA2022, LAMTOR5, LOC100130093, LOC100130370, LOC285696, LOC339529, MAGED4B, MROH8, NLRP2, NUDT18, P2RX5, PKI55, PRDM12, SARS, SREBF1, TMC4, UBL7, WISP2, ZNF839
Men	ANKLE2, ATP13A3, B3GNT2, B4GALT1, C12orf23, C1orf63, CASP7, CTAGE4, ENAH, FAM150B, FAM46B, FAM46C, FITM2, HIST1H2BC, HIST1H2BL, HIST1H2BO, HIST1H3F, HNRPLL, HYAL1, IPO5, KIAA1432, LOC100129617, LOC100506136, MB21D1, RBM14-RBM4, SMEK2, SNORA59B, SNORA63, STBD1, STON1, TMEM140, TMEM163, TRPM8, ACTL6B, ANKRD39, AP3S2, APOA1BP, ATP5G1, C11orf51, C12orf24, C15orf23, C16orf59, C20orf177, C21orf119, C21orf33, C2orf72, C8orf38, C8orf44-SGK3, C9orf23, CECR7, COMMD4, CSF3, DEFA3, EPHA7, FAM19A5, FAM58A, FAM65B, FLJ32063, GTF2IRD2, LOC100652758, LOC730183, LPP-AS2, LRRC16A, MDP1, NIP7, NSUN4, OR7D2, PIPOX, PKI55, PKM2, PPAPDC2, PYCRL, RABL5, RWDD3, SERF1A, SSSCA1, TTC30A, ZNF117, ZNF771

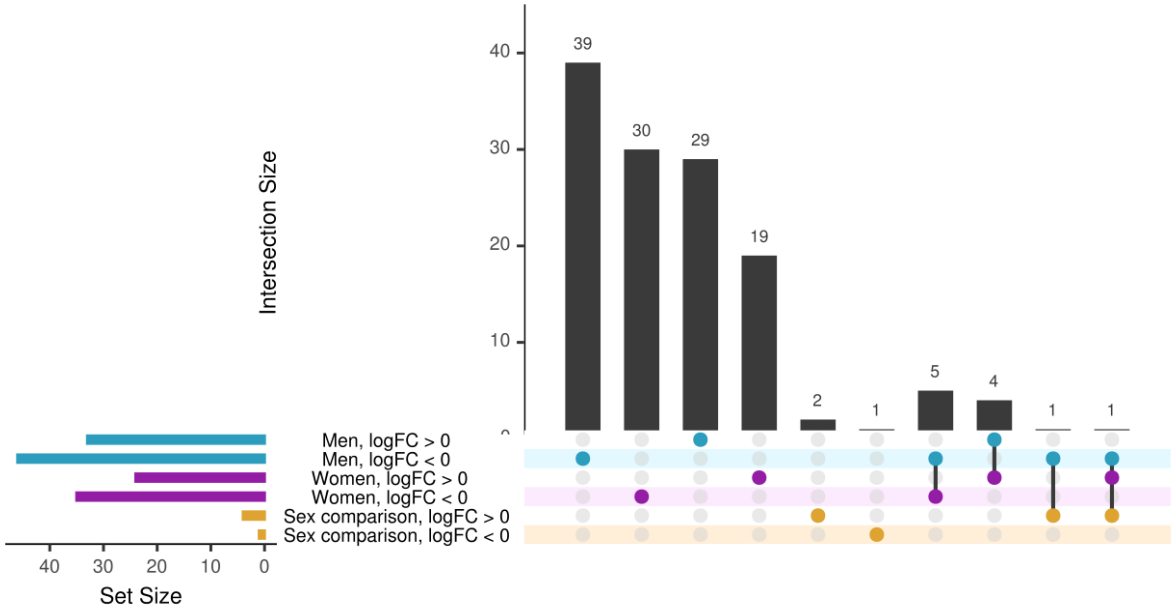

Figure 4. Upset plot of the meta-analysis results in nervous tissue samples. The results of each comparison are separated according to the sign of the logFC. Horizontal bars indicate the number of significant genes in each contrast (one specific color for each contrast). Results are exhibited separately by up-regulated (logFC > 0) and down-regulated (logFC < 0) significant genes. The vertical bars display the common genes in the sets indicated with dots under each bar. The single points represent the number of unique genes in each group.

Meta-analysis of DGE in blood together with nervous tissues

When both tissues were meta-analyzed together, 3 genes appeared dysregulated in the women contrast (2 genes upregulated; 1 downregulated in Women.MS) and 19 genes in men contrast (11 upregulated; 8 downregulated in Men. MS). Gene intersection did not show common genes in the analysis of both tissues between men and women. In the sex contrast (Women.MS - Wen.MS), the LOC102723701 gene was found to be significant, as shown in Figure 5 and Table 7. This gene is described by GeneCards and NCBI PubMed as uncharacterized. The meta-analysis in the individual women contrast indicated two genes upregulated: **CCL19** (Chemokine (C-C motif) ligand 19, involved in inflammatory and immunological responses, and skeletal myopathy in Open Targets) and **RPL35** (Ribosomal Protein L35) and one downregulated: **MYST2** also known as KAT7 (Catalytic subunit of histone acetyltransferase HBO1 complexes, related to fallopian tube carcinoma) (Table 7).

In men, the meta-analysis showed upregulated genes such as: **B4GALT1** (beta-1,4-galactosyltransferase 1, that catalyze the production of lactose); **DUSP1** (dual specificity phosphatase 1, this kind of phosphatases play a regulation role in neurons and glial cells) [45]; **FAU** (ubiquitin like and ribosomal protein S30 fusion, involved in Alzheimer diseases according to Open Targets); **LOC100129617** (RNA gene uncharacterized according to GeneCards); **MTPN** (myotrophin, related to duchenne muscular dystrophy in Open Targets); **OSER1** (oxidative stress responsive serine rich 1); **PPP4R1L** (protein phosphatase 4

regulatory subunit 1 like (pseudogene)); PSMB8 (proteasome 20S subunit beta 8); **PTPLAD2** (also known as HACD4, a 3-hydroxyacyl-CoA dehydratase 4); **TXNL1** (thioredoxin like 1, related to Alzheimer diseases in Open Targets); and **VDAC2** (voltage dependent anion channel 2, also related to Alzheimer diseases in Open Targets). Other genes resulted downregulated in Men.MS: **A2M-AS1** (A2M antisense RNA 1, related to other autoimmune diseases as lupus erythematosus, cutaneus in Open Targets) **FAM115A** (that also appeared downregulated in blood in Men.MS); **FLJ30403** (known as LINC01530, Long Intergenic Non-Protein Coding RNA), several uncharacterized genes **(LOC283861, LOC286272, LOC400968, LOC728903)** and the gene **OR7D2** (olfactory receptor family 7 subfamily D member 2, related to juvenile idiopathic arthritis in Open Targets) (Table 7).

Finally, and as a summary of the results in the different meta-analyses that detect a differential transcriptomic profile by sex, we represented this change in mean expression by means of a line plot (Figure 6). For each of the 7 significant genes identified in the sex difference contrasts, in the different tissues (1 gene in blood, 5 genes in nervous tissue and 1 gene when combining both tissues), we described the pattern of logFC change between men and women.

Table 7. MA results of both tissues samples. Described contrasts "sex differences", "women" and "men" refer (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl), (Women.MS - Women.Ctrl) and (Men.MS - Men.Ctrl), respectively. The red color in the genes means a positive logFC, while the blue color indicates a negative logFC.

Contrast	**Significant genes**
Sex differences	LOC102723701
Women	CCL19, RPL35, MYST2
Men	B4GALT1, DUSP1, FAU, LOC100129617, MTPN, OSER1, PPP4R1L, PSMB8, PTPLAD2, TXNL1, VDAC2, A2M-AS1, FAM115A, FLJ30403, LOC283861, LOC286272, LOC400968, LOC728903, OR7D2

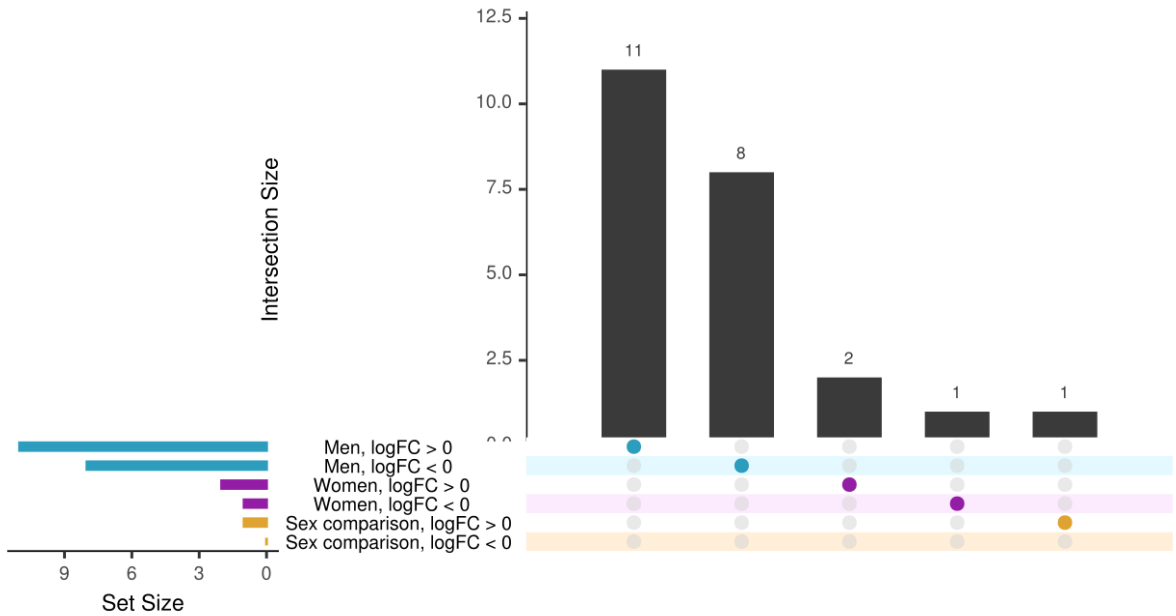

Figure 5. Upset plot of the meta-analysis results in both tissues together. The results of each comparison are separated according to the sign of the logFC. Horizontal bars indicate the number of significant elements in each contrast (one specific color for each contrast). Results are exhibited separately by up-regulated (logFC > 0) and down-regulated (logFC < 0) significant genes. The vertical bars display the common genes in the sets indicated with dots under each bar. The single points represent the number of unique genes in each group.

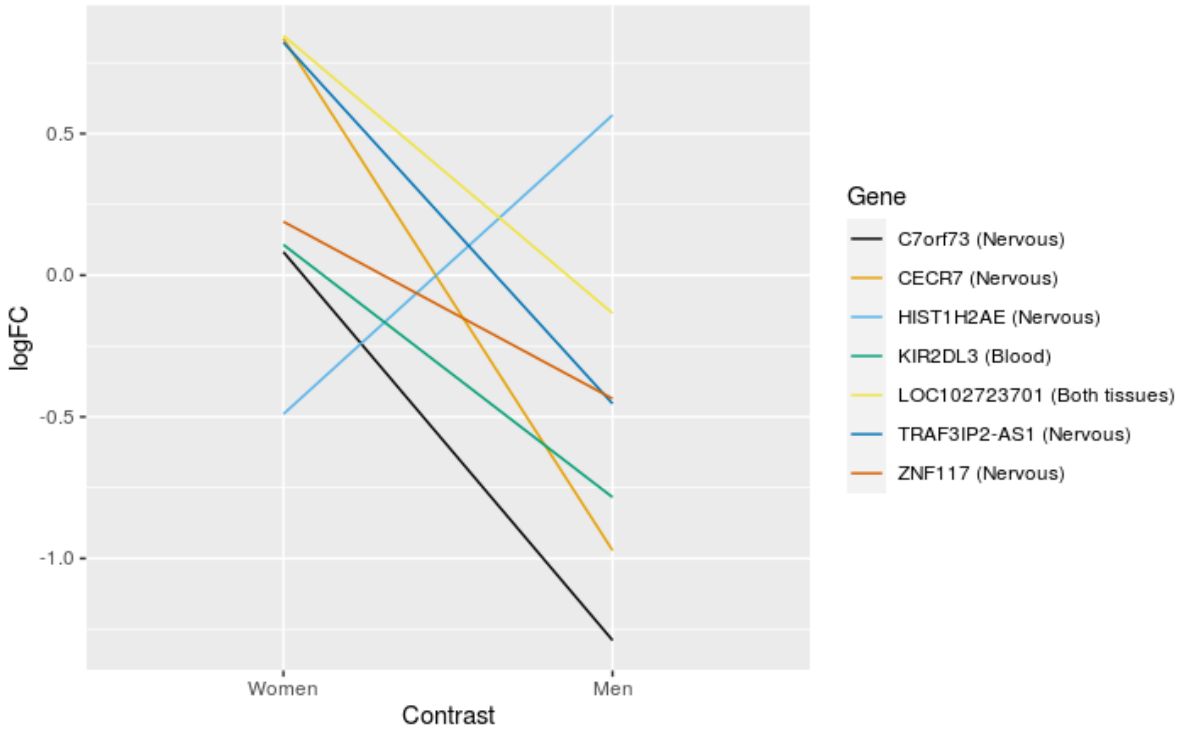

Figure 6. Line plot of the seven significant genes, in the different tissues, in the contrast of sex differences. The lines show the logFC value of each gene for the women and men contrasts.

Functional enrichment of DGE meta-analysis in nervous tissue

Since the meta-analysis resulted in an elevated number of DGEs in nervous tissue, a gene set analysis (GSA) was performed to get a functional and integrative perspective of the genes in GO biological processes and KEGG pathways in individual contrasts and in the sex related. As is shown in Table 8 only one biological process (GO:0023061, signal release) and one KEGG pathway (hsa00472, D-Arginine and D-ornithine metabolism) were significantly related to sex differences, but an elevated number of biological functions was identified in individual contrasts (255 up in women, 236 up in men; 71 down in women and 21 down in men) (Figure 7A, Table 8). Regarding KEGG pathways, a smaller number of significant dysregulated pathways was identified: 28 up in women, 21 in men; and a minor number appeared downregulated 8 and 5 respectively (Figure 7B, Table 8). The top 5 biological processes and KEGG pathways with the most extreme LOR values for the female and male contrasts are shown in Figure 8. Focus on KEGG pathways database, the intersection analysis of functional enrichment revealed only one pathway in sex bias (D-Arginine and D-ornithine metabolism). In women and men contrasts 15 pathways resulted up and 4 down common to both sexes. While 13 pathways were detected exclusively up in women (Biotin metabolism, Spliceosome, p53 signaling pathway, TGF-beta signaling pathway, Toll-like receptor signaling pathway, NOD-like receptor signaling pathway, B cell receptor signaling pathway, Intestinal immune network for IgA production, Chagas disease (American trypanosomiasis), Prostate cancer, Chronic myeloid leukemia, Systemic lupus erythematosus, Rheumatoid arthritis) and 4 pathways were down (Metabolism of xenobiotics by cytochrome P450, Drug metabolism - cytochrome, Cardiac muscle, Alzheimer’s disease). In men were detected 6 pathways up (RNA transport, Cytokine-cytokine receptor interaction, Protein processing in endoplasmic reticulum, Phagosome, Jak-STAT signaling pathway, Olfactory transduction) and one down (Valine, leucine and isoleucine degradation). In addition, common elements throughout the different contrasts (Figure 7) are detailed in the supplementary files deposited in the zenodo repository.

Table 8. Number of significant biological processes or KEGG pathways.

Contrast	**Biological processes**		**KEGG**	
	LOR > 0	LOR < 0	LOR > 0	LOR < 0
Sex differences	0	1	1	0
Women	255	71	28	8
Men	236	33	21	5

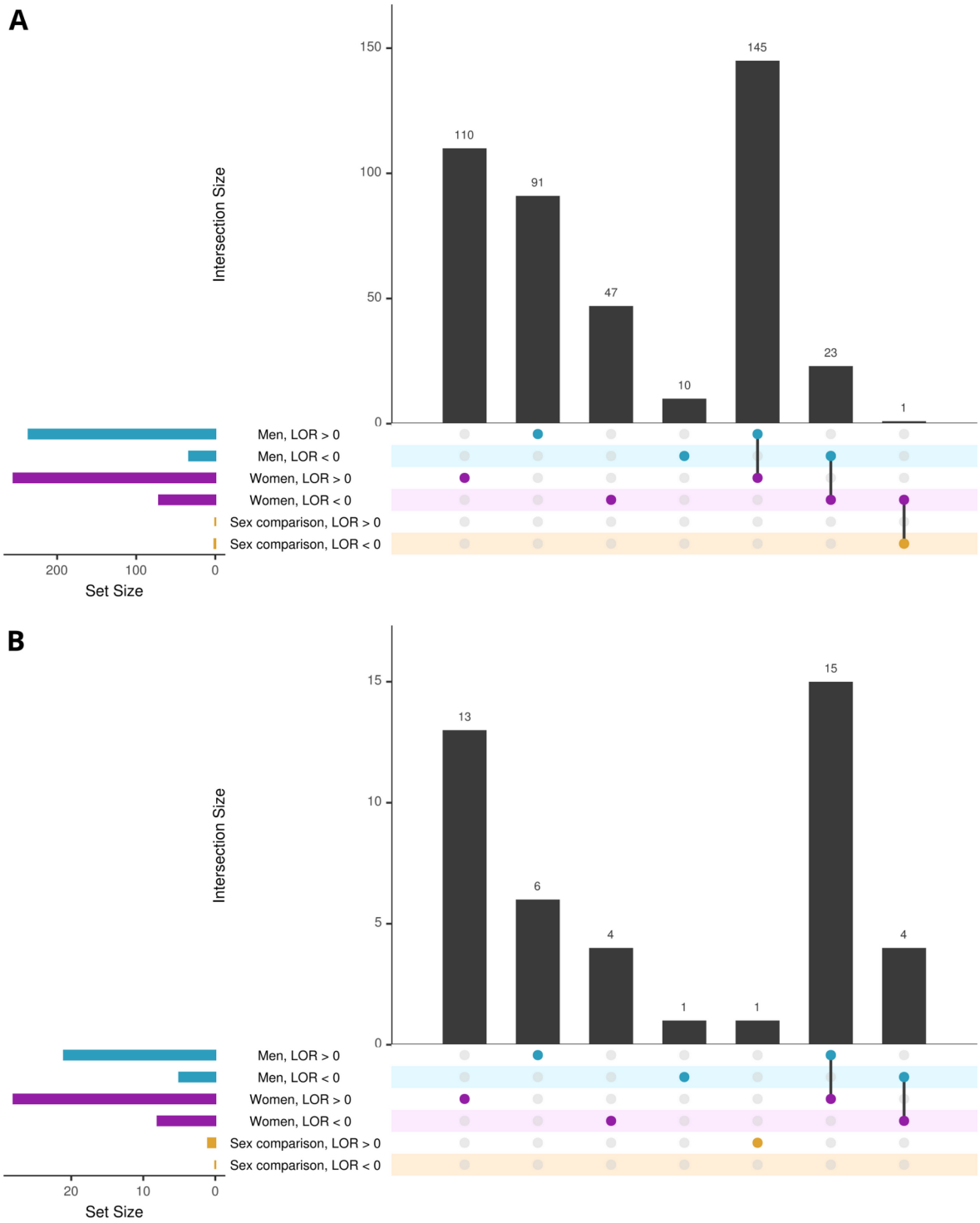

Figure 7. Upset plot of the functional enrichment results, for biological processes (A) and KEGG pathways (B). Horizontal bars indicate the number of significant biological functions in each contrast (one specific color for each contrast). Results are exhibited separately by up-regulated (LOR > 0) and down-regulated (LOR < 0) significant biological processes (A) or KEGG pathways (B). The vertical bars display the common functional elements in the sets indicated with dots under each bar. The single points represent the number of unique processes (A) or pathways (B) in each group.

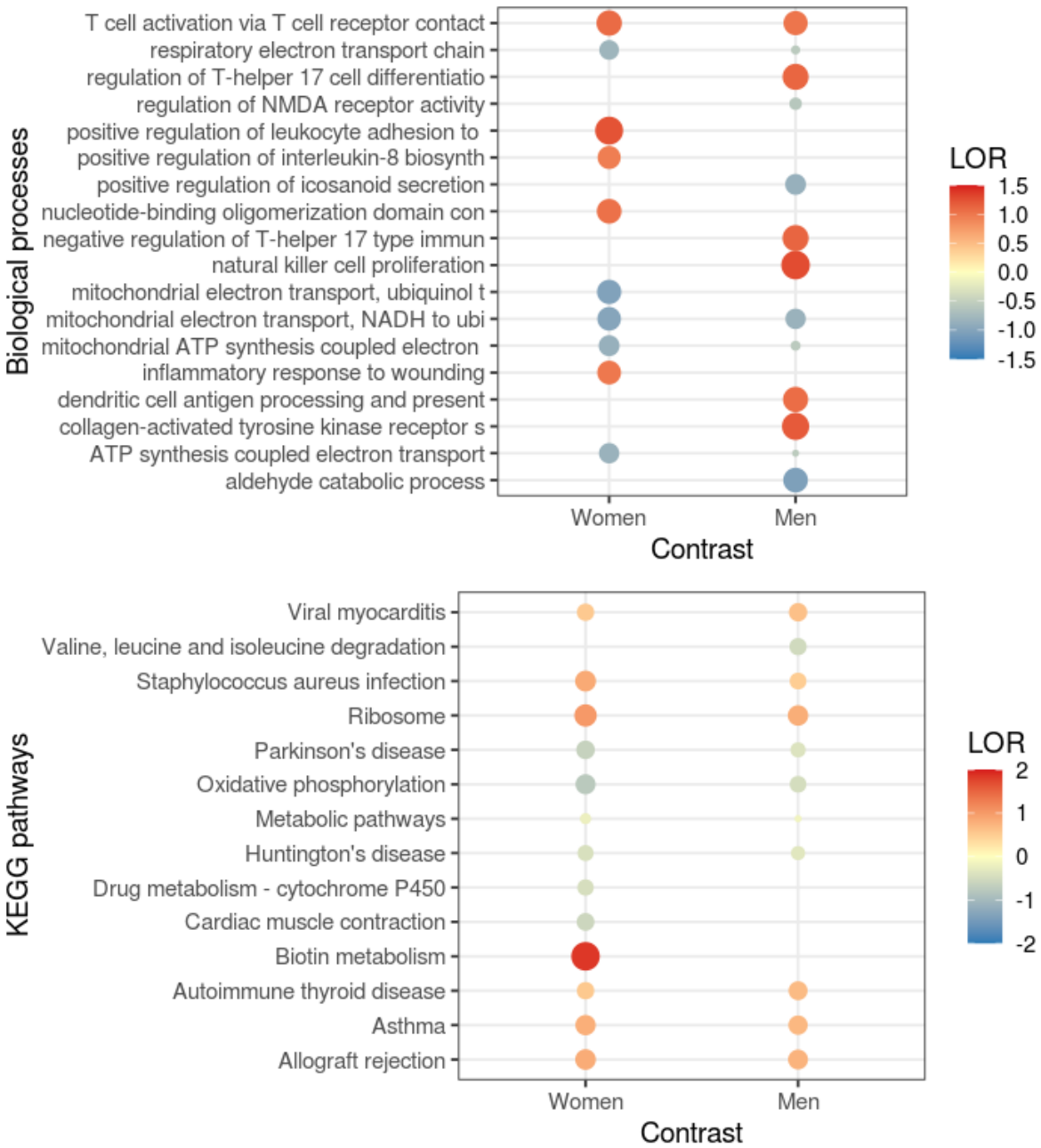

Figure 8. Significant biological processes and KEGG pathways with highest absolute LOR values in female and male contrasts. The size of the dot indicates the absolute value of the LOR.

Metafun-MS platform

The Metafun-MS platform (https://bioinfo.cipf.es/metafun-MS) contains an exhaustive description related to the nine studies and 653 samples involved in this study. For each study, this resource includes detailed information on all genes and biological processes assessed that users can explore to identify transcriptomics profiles of interest.

We carried out a total of 9 meta-analyses, one for each combination of tissue and contrast.

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

Here, we also highlight the impact of including and reporting sex-related data in the results of clinical studies, given their general importance in the progression, the severity and other outcomes in multiple sclerosis, as well as in other neurodegenerative disorders. The integration of sex-based differences in this manner has the potential to significantly impact the neurodegenerative biology field.

Discussion

Strengths and limitations. Pathophysiological differences of a multitude of diseases as neurodegenerative and auntoimmune diseases are influenced by patient sex [1, 46]. In MS sex differences has been described in risk and disease progression [47]. Unfortunately, the sex as biological variable has been infrequently considered by the scientific community during biomedical research that was evident in our systematic review in which an important number of studies were discarded by the absence of sex information. In recent years, sex perspective has begun to gain importance, as it represents one of the keys to personalised medicine. In fact, relatively recently sex variable has been recognized by the National Institute of Health (NIH) [48] and the European Union [49, 50] as a biological variable to include in research.

Previous works have applied this strategy of integrating data from numerous studies to better characterise sex differences [51–53]. However, to our knowledge, this is the first time that a meta-analysis of transcriptomic studies in MS has been carried out, where several tissues have been evaluated with the aim of identifying sex-specific biomarkers that provide a better understanding of the molecular mechanisms in MS. As described in the introduction section, MS has a great impact on society, being the majority cause of disability not associated with accidents in the young adult population but MS causes knowledge is very limited. Their etiology is still unknown, and due to the great variability of clinical manifestations, its diagnosis and prognosis are highly complex [54]. Likewise, MS continues to be an incurable chronic disease and a clinical challenge, since current treatments are palliatives. With this objective, an *in silico* approach has been carried out using MS data deposited in public repositories with sex information for each sample. *In silico* studies are fast and low cost-effective tools that can help predict and validate the effects of a variable on human health and guide further clinical research. But the lack of standardization in the description, nomenclature, and format of public data has been evident. Therefore, in addition to encouraging the data generated in the research process to be publicly accessible (Open science), it is highly relevant to promote the FAIR principles (Findable, Accessible, Interoperable, Reusable) [55]. These bases, established by scientists from different areas, aim to standardize the process so that the data is deposited in the repositories in a homogeneous way.

In this work, we have performed a meta-analysis using a random-effects model. This robust method allows integrating information from different individual studies to obtain a greater statistical power and precision, and reveals findings that cannot be obtained through the intersection or addition of results in individual studies. A representative example can be seen in Table 3, the contribution of each study is diverse, specially, the study GSE108000 (nervous tissue) that showed an elevated number of significant genes compared to other studies in which no genes were detected. This highlights the need for the use of strategies such as meta-analysis, capable of integrating information from individual studies and obtaining a transcriptional consensus profile as a final result. To make results comparable and reduce biases (different sample sizes, platforms, protocols, etc.) we applied the same bioinformatics strategy from normalized expression matrices to DGE results. The design and application of this thorough strategy has provided the identification of sex-specific biomarkers in the set of the various subtypes of multiple sclerosis. Further generation of studies with sex information within each of these subgroups will make it possible in the future to extend this methodology by specific subtypes and tissues.

Meta-analysis of transcriptomics profiles. Since MS is a chronic autoinflammatory disorder of the central nervous system, we performed a meta-analysis in each of the tissues involved, blood and nervous system, and a last meta-analysis at the integrative level of both tissues. Our major goal was identified sex differences in MS through the contrast (Women.MS - Women.Ctrl) - (Men.MS - Men.Ctrl). This contrast is best suited to identify differences between men and women with multiple sclerosis, having considered their respective controls. For example, it allows us to detect genes that have the same expression profile in both sexes (more active in patients), but have a much higher level of expression in men or women. By means of this contrast, it is also possible to identify genes that present an opposite expression pattern by sex (in one group it is upregulated and in the other it is downregulated). On some occasions it may be that some of these genes are significant in the specific contrasts for men and women, however, on other occasions, they may not be significant in the individual contrasts, but we can detect this difference in expression through this joint evaluation of the data for men and women.. In addition, we also performed individual contrasts (Women.MS - Women.Ctrl) and (Men.MS - Men.Ctrl) to find out the common and specific events in women and men with MS. Therefore, we offer all possible perspectives for a better knowledge about MS. The findings presented in this study provide information that could be exploited for understanding etiology of disease, design novel therapeutics approaches and for biomarker development. Identifying biomarkers in blood is an alternative less invasive than other methods.

Meta-analysis in blood with sex perspective identified the disregulation of KIR2DL3. The overexpression of this gene in Women.MS could suggest an advantage over Men.MS according to their described protective effect [31, 56]. In general, the identification of biomarkers in blood is a challenge made difficult by several factors such as: the availability of the molecules in blood, dilution factor, sensitivity, specificity and disease progression condition. These reasons could explain because only one gene (KIR2DL3) has been identified in blood and also taking account the sex perspective. The circulating levels in blood of molecules related to several processes involved in MS, such as inflammation and oxidative stress, represent a valuable tool for early diagnosis and evaluation of MS progression. In women, the meta-analyses in blood revealed a gene dysregulation related to

dementia and intellectual tasks and also was more evident that in men skeletal and osteo disturbances, that could indicate a relation with menopause and sexual hormone levels in women [57]. While blood men reflected a dysregulation in mitochondrial and metabolism genes.

Regarding nervous tissue, the number of DGEs identified was higher than in blood, that is according to nervous tissue, the main tissue affected by MS. Again, women observed the same trend as in blood. The genes affected were related to osteo and muscle problems, mainly, and moreover sex related. When both tissues were analyzed only LOC102723701 was related to sex differences. This gene is not described in the literature. In women, genes related to inflammatory and immune response and skeletal disorders were upregulated. Autoimmune diseases are exacerbated in women and a more proinflammatory phenotype of XX compared to XY- mice has been also demonstrated in experimental and spontaneous lupus models [58, 59]. Sex bias in immune responses are well documented [60]. Although the focus of this work is on sex differences in MS, also in the meta-analysis of nerve tissue studies, a group of 9 candidate biomarkers of MS, common in both men and women, could be identified by stratified analysis, which so far had not been described in previous works.

In men, the meta-analysis in both tissues also recapitulated, in general, the meta-analysis performed in individual tissues where genes related to oxidative stress, mitochondrial functions, metabolism and other neural disorders as Alzheimer were dysregulated, mainly. Aiming at the concordance of the data obtained among three meta-analyses.

Functional enrichment. Regarding sex meta-analysis in nervous system only one biological process (signal release) and one KEGG pathways (D-Arginine and D-ornithine metabolism) were evidenced, the former overrepresented in men (LOR < 0) and the latter in women (LOR > 0). The biological process we have found is rather generic, as it is related to the secretion of signals into the extracellular medium from a cellular source, but without further details about the signal or the origin of the secretion. For the KEGG pathway, previous studies have shown the role of pathways associated with these amino acids in the modulation of the autoimmune response, as well as mechanisms of cytotoxicity and neuroprotection [61, 62]. However, these studies did not take into account sex as factor and did not detect its influence on these alterations.

We performed a GSA analysis since the high number of genes identified in the individual contrasts men and women. GSA showed that women presented an enrichment in immune responses as previously evidenced by the DGEs meta-analyses and functions related to other autoimmune diseases as Systemic lupus erythematosus and Rheumatoid arthritis and in concordance with stronger immune responses and higher incidence of autoimmunity diseases in women [63] and down functions related to Alzheimer diseases and metabolism (drugs and xenobiotics) that in contrast seem upregulated in men.

Our study has characterized transcriptomics differences between the sexes in MS, shedding light on the molecular basis behind this pathology in male and female patients. While our meta-analysis confirmed the conclusions of other studies, we also reported previously novel sex-specific genes and undescribed functional alterations that may broaden this field of study.

Further knowledge regarding how those factors related to the molecular mechanisms, described above, differentially impact male and female MS patients and may improve our understanding of the disease and improve treatment and diagnosis through biomarker identification.

Conclusions

This study provide novel and suitable information about MS with sex perspective but also in women and men as individual groups. The list of genes and pathways identified in the different contrasts and tissues involved in MS allow a better understanding of the mechanisms involved in MS, and the identification of potential biomarkers by sex and no-related sex, whose detailed study will improve personalized medicine applications for this disease. The researchers must be aware of providing sex information in the studies. Finally, the use of open platforms to share data is key for research progression.

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.5211962

Abbreviations

The following abbreviations are used in this manuscript:

BH: Benjamini & Hochberg p-value adjust method; BP: Biological processes of Gene Ontology; CIS: Clinical Isolated Syndrome; DGE: Differential gene expression; GEO: Genes Omnibus Expression database; GO: Gene Ontology; GSA: Gene set analysis; KEGG: Kyoto Encyclopedia of Genes and Genomes; LogFC: Logarithm of the fold change; LOR: Logarithm of odds ratio; MS: Multiple Sclerosis; PCA: Principal component analysis; PRMS: Progressive-Relapse Multiple Sclerosis; RRMS: Relapsing-Remitting Multiple Sclerosis; SPMS: Secondary-Progressive Multiple Sclerosis.

Funding

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

Availability of data and materials

The large volume 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.5211962. 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 is publicly available at GitLab (https://gitlab.com/ubb-cipf/metafunr). All software and versions used are detailed in the zenodo repository.

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

Competing interests

The authors declare no competing interests.

Acknowledgements

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.

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. 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.
2. Bishop M, Rumrill PD. Multiple sclerosis: Etiology, symptoms, incidence and prevalence, and implications for community living and employment. Work Read Mass. 2015;52:725–34.
3. Voskuhl RR. The effect of sex on multiple sclerosis risk and disease progression. Mult Scler J. 2020;26:554–60.
4. 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.
5. 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.
6. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. The Lancet. 2018;391:1622–36.
7. Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. The Lancet. 2017;389:1336–46.
8. Pierdominici M, Maselli A, Colasanti T, Giammarioli AM, Delunardo F, Vacirca D, et al. Estrogen receptor profiles in human peripheral blood lymphocytes. Immunol Lett. 2010;132:79–85.
9. 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 Res. 2012;41:D991–5.
10. 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 Res. 2019;47:D711–5.
11. 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.
12. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019. https://www.R-project.org/.
13. 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 Res. 2015;43:e47.
14. García-García F. Métodos de análisis de enriquecimiento funcional en estudios genómicos. University of Valencia; 2016. https://www.educacion.gob.es/teseo/mostrarRef.do?ref=1307283.
15. Viechtbauer W. Conducting Meta-Analyses in *R* with the **metafor** Package. J Stat Softw. 2010;36. doi:10.18637/jss.v036.i03.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177–88.
17. 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 Res. 2021;49:D1302–10.
18. Montaner D, Dopazo J. Multidimensional Gene Set Analysis of Genomic Data. PLoS One. 2010;5:e10348.
19. 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.
20. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28:27–30.
21. Sievert C. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC; 2020. https://plotly-r.com.
22. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016. https://ggplot2.tidyverse.org.
23. 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.
24. 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.
25. 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.
26. 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.
27. 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.
28. van der Poel M, Ulas T, Mizee MR, Hsiao C-C, Miedema SSM, Adelia null, et al. Transcriptional profiling of human microglia reveals grey-white matter heterogeneity and multiple sclerosis-associated changes. Nat Commun. 2019;10:1139.
29. 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 Neuroinflammation. 2020;7:e656.
30. 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.
31. 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 J. 2012;18:951–8.
32. Khodadadi H, Azcona LJ, Aghamollaii V, Omrani MD, Garshasbi M, Taghavi S, et al. PTRHD1 (C2orf79) mutations lead to autosomal-recessive intellectual disability and parkinsonism. Mov Disord Off J Mov Disord Soc. 2017;32:287–91.
33. Kun-Rodrigues C, Orme T, Carmona S, Hernandez DG, Ross OA, Eicher JD, et al. A comprehensive screening of copy number variability in dementia with Lewy bodies. Neurobiol Aging. 2019;75:223.e1-223.e10.
34. Al-Jallad H, Palomo T, Roughley P, Glorieux FH, McKee MD, Moffatt P, et al. The effect of SERPINF1 in-frame mutations in osteogenesis imperfecta type VI. Bone. 2015;76:115–20.
35. Cho HI, Kim MS, Lee J, Yoo BC, Kim KH, Choe K-M, et al. BRPF3-HUWE1-mediated regulation of MYST2 is required for differentiation and cell-cycle progression in embryonic stem cells. Cell Death Differ. 2020;27:3273–88.
36. Fernandes IR, Cruz ACP, Ferrasa A, Phan D, Herai RH, Muotri AR. Genetic variations on SETD5 underlying autistic conditions. Dev Neurobiol. 2018;78:500–18.
37. Deliu E, Arecco N, Morandell J, Dotter CP, Contreras X, Girardot C, et al. Haploinsufficiency of the intellectual disability gene SETD5 disturbs developmental gene expression and cognition. Nat Neurosci. 2018;21:1717–27.
38. Pu M, Wang X, Zhang J. Impact of GRIK4 gene polymorphisms on cognitive dysfunction in patients with major depression. Rev Neurol (Paris). 2020;176:823–8.
39. Martinez B, Peplow PV. MicroRNAs in Parkinson’s disease and emerging therapeutic targets. Neural Regen Res. 2017;12:1945–59.
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. Sugiura Y, Araki K, Iemura S, Natsume T, Hoseki J, Nagata K. Novel thioredoxin-related transmembrane protein TMX4 has reductase activity. J Biol Chem. 2010;285:7135–42.
42. Wood KA, Rowlands CF, Thomas HB, Woods S, O’Flaherty J, Douzgou S, et al. Modelling the developmental spliceosomal craniofacial disorder Burn-McKeown syndrome using induced pluripotent stem cells. PloS One. 2020;15:e0233582.
43. Fu D, Liu S, Liu J, Chen W, Long X, Chen X, et al. iTRAQ-based proteomic analysis of the molecular mechanisms and downstream effects of fatty acid synthase in osteosarcoma cells. J Clin Lab Anal. 2021;35:e23653.
44. Madar IH, Sultan G, Tayubi IA, Hasan AN, Pahi B, Rai A, et al. Identification of marker genes in Alzheimer’s disease using a machine-learning model. Bioinformation. 2021;17:348–55.
45. Pérez-Sen R, Queipo MJ, Gil-Redondo JC, Ortega F, Gómez-Villafuertes R, Miras-Portugal MT, et al. Dual-Specificity Phosphatase Regulation in Neurons and Glial Cells. Int J Mol Sci.

2019;20:E1999.
46. Yanguas-Casás N. Sex Differences in Neurodegenerative Diseases. SM J Neurol Disord Stroke. 2017;3:1014.
47. Voskuhl RR, Sawalha AH, Itoh Y. Sex chromosome contributions to sex differences in multiple sclerosis susceptibility and progression. Mult Scler J. 2018;24:22–31.
48. Clayton JA, Collins FS. Policy: NIH to balance sex in cell and animal studies. Nature. 2014;509:282–3.
49. European Commission. Directorate General for Justice and Consumers., Fondazione Giacomo Brodolini. Gender equality and health in the EU. LU: Publications Office; 2021. https://data.europa.eu/doi/10.2838/991480. Accessed 3 Aug 2021.
50. Accounting for sex and gender makes for better science. Nature. 2020;588:196–196.
51. 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.
52. 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.
53. 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.
54. Dobson R, Giovannoni G. Multiple sclerosis - a review. Eur J Neurol. 2019;26:27–40.
55. 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.
56. 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.
57. Miclea A, Salmen A, Zoehner G, Diem L, Kamm CP, Chaloulos-Iakovidis P, et al. Age-dependent variation of female preponderance across different phenotypes of multiple sclerosis: A retrospective cross-sectional study. CNS Neurosci Ther. 2019;25:527–31.
58. 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.
59. 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.
60. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016;16:626–38.
61. Correale J. Immunosuppressive Amino-Acid Catabolizing Enzymes in Multiple Sclerosis. Front Immunol. 2021;11:600428.
62. 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.
63. Moulton VR. Sex Hormones in Acquired Immunity and Autoimmune Disease. Front Immunol. 2018;9:2279.